

IBM Spectrum Scale  
Version 4 Release 2.1

*Problem Determination Guide*





IBM Spectrum Scale  
Version 4 Release 2.1

*Problem Determination Guide*



**Note**

Before using this information and the product it supports, read the information in “Notices” on page 593.

This edition applies to version 4 release 2 modification 1 of the following products, and to all subsequent releases and modifications until otherwise indicated in new editions:

- IBM Spectrum Scale ordered through Passport Advantage® (product number 5725-Q01)
- IBM Spectrum Scale ordered through AAS/eConfig (product number 5641-GPF)
- IBM Spectrum Scale for Linux on z Systems (product number 5725-S28)

Significant changes or additions to the text and illustrations are indicated by a vertical line (|) to the left of the change.

IBM welcomes your comments; see the topic “How to send your comments” on page xiv. When you send information to IBM, you grant IBM a nonexclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

© Copyright IBM Corporation 2014, 2016.

US Government Users Restricted Rights – Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

# Contents

<b>Tables</b> . . . . .	<b>ix</b>	<b>Chapter 3. Monitoring system health by using the mmhealth command</b> . . . . .	<b>85</b>
<b>About this information</b> . . . . .	<b>xi</b>	<b>Chapter 4. Monitoring events through callbacks.</b> . . . . .	<b>91</b>
Prerequisite and related information . . . . .	xiii	<b>Chapter 5. Monitoring AFM</b> . . . . .	<b>93</b>
Conventions used in this information . . . . .	xiii	Fileset states for AFM . . . . .	93
How to send your comments . . . . .	xiv	Fileset states for AFM DR . . . . .	96
<b>Summary of changes</b> . . . . .	<b>xv</b>	Callback events for AFM and AFM DR . . . . .	100
<b>Chapter 1. Performance monitoring</b> . . . . .	<b>1</b>	Monitoring with <b>mmpmon</b> . . . . .	100
Network performance monitoring . . . . .	1	Monitor prefetch . . . . .	101
Monitoring GPFS I/O performance with the mmpmon command . . . . .	3	Policies used for monitoring AFM and AFM DR . . . . .	102
Overview of mmpmon . . . . .	3	<b>Chapter 6. GPFS SNMP support</b> . . . . .	<b>105</b>
Specifying input to the mmpmon command . . . . .	3	Installing Net-SNMP . . . . .	105
Display I/O statistics per mounted file system . . . . .	5	Configuring Net-SNMP . . . . .	106
Display I/O statistics for the entire node . . . . .	6	Configuring management applications . . . . .	106
Understanding the node list facility . . . . .	8	Installing MIB files on the collector node and management node . . . . .	107
Reset statistics to zero . . . . .	15	Collector node administration . . . . .	107
Understanding the request histogram facility . . . . .	16	Starting and stopping the SNMP subagent . . . . .	108
Understanding the Remote Procedure Call (RPC) facility . . . . .	28	The management and monitoring subagent . . . . .	108
Displaying mmpmon version . . . . .	31	SNMP object IDs . . . . .	109
Example mmpmon scenarios and how to analyze and interpret their results . . . . .	31	MIB objects . . . . .	109
Other information about mmpmon output . . . . .	40	Cluster status information . . . . .	109
Performance monitoring tool overview . . . . .	41	Cluster configuration information . . . . .	109
Configuring the performance monitoring tool . . . . .	43	Node status information . . . . .	110
Restarting the performance monitoring tool . . . . .	63	Node configuration information . . . . .	110
Configuring the metrics to collect performance data . . . . .	63	File system status information . . . . .	111
Viewing performance data with <b>mmpmon</b> . . . . .	64	File system performance information . . . . .	112
Performance monitoring through IBM Spectrum Scale GUI . . . . .	69	Storage pool information . . . . .	112
Configuring performance monitoring options in GUI . . . . .	70	Disk status information . . . . .	113
Configuring performance metrics and display options in the Statistics page of the GUI . . . . .	71	Disk configuration information . . . . .	113
Configuring the dashboard to view performance charts . . . . .	74	Disk performance information . . . . .	114
Querying performance data shown in the GUI through CLI . . . . .	76	Net-SNMP traps . . . . .	115
Monitoring performance of nodes . . . . .	76	<b>Chapter 7. Monitoring the IBM Spectrum Scale system by using call home</b> . . . . .	<b>117</b>
Performance monitoring limitations . . . . .	77	Collecting details that need to be shared with IBM Support . . . . .	117
<b>Chapter 2. Monitoring system health through the IBM Spectrum Scale GUI</b> . . . . .	<b>79</b>	Configuring call home . . . . .	119
Set up event notifications . . . . .	79	Sharing collected details with IBM Support . . . . .	124
Configuring email notifications . . . . .	80	Prerequisites for configuring call home . . . . .	124
Configuring SNMP manager . . . . .	80	<b>Chapter 8. Monitoring the health of Transparent Cloud Tiering service</b> . . . . .	<b>127</b>
<b>Chapter 3. Monitoring system health by using the mmhealth command</b> . . . . .	<b>85</b>	<b>E! Chapter 9. Best practices for E! troubleshooting</b> . . . . .	<b>129</b>
<b>Chapter 4. Monitoring events through callbacks.</b> . . . . .	<b>91</b>	E! How to get started with troubleshooting . . . . .	129
<b>Chapter 5. Monitoring AFM</b> . . . . .	<b>93</b>	E! Set up event notifications . . . . .	129
Fileset states for AFM . . . . .	93		
Fileset states for AFM DR . . . . .	96		
Callback events for AFM and AFM DR . . . . .	100		
Monitoring with <b>mmpmon</b> . . . . .	100		
Monitor prefetch . . . . .	101		
Policies used for monitoring AFM and AFM DR . . . . .	102		
<b>Chapter 6. GPFS SNMP support</b> . . . . .	<b>105</b>		
Installing Net-SNMP . . . . .	105		
Configuring Net-SNMP . . . . .	106		
Configuring management applications . . . . .	106		
Installing MIB files on the collector node and management node . . . . .	107		
Collector node administration . . . . .	107		
Starting and stopping the SNMP subagent . . . . .	108		
The management and monitoring subagent . . . . .	108		
SNMP object IDs . . . . .	109		
MIB objects . . . . .	109		
Cluster status information . . . . .	109		
Cluster configuration information . . . . .	109		
Node status information . . . . .	110		
Node configuration information . . . . .	110		
File system status information . . . . .	111		
File system performance information . . . . .	112		
Storage pool information . . . . .	112		
Disk status information . . . . .	113		
Disk configuration information . . . . .	113		
Disk performance information . . . . .	114		
Net-SNMP traps . . . . .	115		
<b>Chapter 7. Monitoring the IBM Spectrum Scale system by using call home</b> . . . . .	<b>117</b>		
Collecting details that need to be shared with IBM Support . . . . .	117		
Configuring call home . . . . .	119		
Sharing collected details with IBM Support . . . . .	124		
Prerequisites for configuring call home . . . . .	124		
<b>Chapter 8. Monitoring the health of Transparent Cloud Tiering service</b> . . . . .	<b>127</b>		
<b>E! Chapter 9. Best practices for E! troubleshooting</b> . . . . .	<b>129</b>		
E! How to get started with troubleshooting . . . . .	129		
E! Set up event notifications . . . . .	129		

E! Back up your data . . . . .	130
E! Resolve events in a timely manner . . . . .	131
E! Keep your software up to date . . . . .	131
E! Subscribe to the support notification. . . . .	131
E! Know your IBM warranty and maintenance agreement details . . . . .	131
E! Know how to report a problem . . . . .	131
E! Other problem determination hints and tips . . . . .	132
E! Which physical disk is associated with a logical volume in AIX systems?. . . . .	133
E! Which nodes in my cluster are quorum nodes? . . . . .	133
E! What is stored in the /tmp/mmfs directory and why does it sometimes disappear? . . . . .	134
E! Why does my system load increase significantly during the night? . . . . .	134
E! What do I do if I receive message 6027-648? . . . . .	134
E! Why can't I see my newly mounted Windows file system? . . . . .	134
E! Why is the file system mounted on the wrong drive letter? . . . . .	134
E! Why does the offline mmfscck command fail with "Error creating internal storage"? . . . . .	135
E! Why do I get timeout executing function error message? . . . . .	135
E! Questions related to active file management . . . . .	135

**Chapter 10. Understanding the system limitations. . . . . 137**

**Chapter 11. Collecting details of the issues . . . . . 139**

E! Collecting details of issues by using logs, dumps, and traces . . . . .	139
Time stamp in GPFS log entries . . . . .	139
E! Logs . . . . .	140
E! Setting up core dumps on a client system . . . . .	161
Trace facility. . . . .	162
E! Collecting diagnostic data through GUI . . . . .	174
E! CLI commands for collecting issue details . . . . .	176
E! Using the gpfs.snap command. . . . .	176
E! mmdumpperfdata command . . . . .	184
E! mmfsadm command . . . . .	186
E! Commands for GPFS cluster state information . . . . .	187
E! GPFS file system and disk information commands . . . . .	191
E! Collecting details of the issues from performance monitoring tools . . . . .	205
Other problem determination tools . . . . .	206

**Chapter 12. Managing deadlocks . . . . . 207**

Debug data for deadlocks . . . . .	207
Automated deadlock detection . . . . .	208
Automated deadlock data collection. . . . .	209
Automated deadlock breakup . . . . .	210
Deadlock breakup on demand. . . . .	211

**Chapter 13. Installation and configuration issues . . . . . 213**

Resolving most frequent problems related to installation, deployment, and upgrade . . . . .	213
Finding deployment related error messages more easily and using them for failure analysis . . . . .	213
Problems due to missing prerequisites . . . . .	218
Problems due to mixed operating system levels in the cluster . . . . .	221
Problems due to lack of clarity on using the spectrumscale installation toolkit for functions or configurations not fully supported . . . . .	222
Problems due to lack of clarity on supported upgrade functions with spectrumscale installation toolkit . . . . .	225
Post installation and configuration problems . . . . .	226
E! Cluster is crashed after reinstallation . . . . .	226
Node cannot be added to the GPFS cluster . . . . .	227
Problems with the /etc/hosts file. . . . .	227
Linux configuration considerations . . . . .	227
Problems with running commands on other nodes . . . . .	228
Authorization problems . . . . .	228
Connectivity problems . . . . .	229
GPFS error messages for rsh problems . . . . .	229
E! Cluster configuration data file issues . . . . .	229
E! GPFS cluster configuration data file issues. . . . .	229
GPFS error messages for cluster configuration data file problems . . . . .	230
Recovery from loss of GPFS cluster configuration data file . . . . .	230
Automatic backup of the GPFS cluster data . . . . .	231
E! GPFS application calls . . . . .	231
Error numbers specific to GPFS applications calls . . . . .	231
GPFS modules cannot be loaded on Linux. . . . .	232
E! GPFS daemon issues . . . . .	232
GPFS daemon will not come up . . . . .	232
GPFS daemon went down . . . . .	236
GPFS commands are unsuccessful . . . . .	237
GPFS error messages for unsuccessful GPFS commands . . . . .	238
Quorum loss . . . . .	239
E! CES configuration issues . . . . .	239
Application program errors. . . . .	239
GPFS error messages for application program errors . . . . .	240
E! Windows issues . . . . .	241
Home and .ssh directory ownership and permissions . . . . .	241
Problems running as Administrator . . . . .	241
GPFS Windows and SMB2 protocol (CIFS serving) . . . . .	241

**E! Chapter 14. Network issues . . . . . 243**

E! IBM Spectrum Scale failures due to a network failure . . . . .	243
E! OpenSSH connection delays . . . . .	243

**Chapter 15. File system issues . . . . . 245**

E! File system fails to mount . . . . .	245
---	-----

GPFS error messages for file system mount problems . . . . .	247	Errors encountered with storage pools . . . . .	267
Error numbers specific to GPFS application calls when a file system mount is not successful . . . . .	248	Snapshot problems . . . . .	267
Mount failure due to client nodes joining before NSD servers are online . . . . .	248	Problems with locating a snapshot . . . . .	268
E! File system fails to unmount . . . . .	249	Problems not directly related to snapshots. . . . .	268
Remote node expelled after remote file system successfully mounted. . . . .	250	Snapshot usage errors . . . . .	268
File system forced unmount . . . . .	250	Snapshot status errors . . . . .	269
Additional failure group considerations . . . . .	251	Snapshot directory name conflicts . . . . .	269
GPFS error messages for file system forced unmount problems . . . . .	252	Errors encountered when restoring a snapshot . . . . .	270
Error numbers specific to GPFS application calls when a file system has been forced to unmount . . . . .	252	Errors encountered when restoring a snapshot . . . . .	271
Automount file system will not mount . . . . .	252	Failures using the mmpmon command . . . . .	271
Steps to follow if automount fails to mount on Linux . . . . .	253	Failures using the mmbackup command . . . . .	273
Steps to follow if automount fails to mount on AIX . . . . .	254	GPFS error messages for mmbackup errors . . . . .	273
Remote file system will not mount . . . . .	255	IBM Spectrum Protect error messages . . . . .	273
Remote file system I/O fails with the "Function not implemented" error message when UID mapping is enabled . . . . .	255	Data integrity . . . . .	273
Remote file system will not mount due to differing GPFS cluster security configurations . . . . .	256	Error numbers specific to GPFS application calls when data integrity may be corrupted . . . . .	274
Cannot resolve contact node address . . . . .	256	Messages requeuing in AFM . . . . .	274
The remote cluster name does not match the cluster name supplied by the mmremotecluster command . . . . .	256		
Contact nodes down or GPFS down on contact nodes . . . . .	257	<b>Chapter 16. Disk issues . . . . . 275</b>	
GPFS is not running on the local node . . . . .	257	NSD and underlying disk subsystem failures. . . . .	275
The NSD disk does not have an NSD server specified and the mounting cluster does not have direct access to the disks . . . . .	258	Error encountered while creating and using NSD disks . . . . .	275
The cipherList option has not been set properly . . . . .	258	Displaying NSD information . . . . .	276
Remote mounts fail with the "permission denied" error message . . . . .	259	NSD creation fails with a message referring to an existing NSD . . . . .	278
Unable to determine whether a file system is mounted . . . . .	259	GPFS has declared NSDs as down . . . . .	278
GPFS error messages for file system mount status . . . . .	259	Unable to access disks . . . . .	279
Multiple file system manager failures . . . . .	259	Guarding against disk failures. . . . .	280
GPFS error messages for multiple file system manager failures . . . . .	260	Disk connectivity failure and recovery . . . . .	281
Error numbers specific to GPFS application calls when file system manager appointment fails . . . . .	260	Partial disk failure. . . . .	281
Discrepancy between GPFS configuration data and the on-disk data for a file system. . . . .	260	GPFS has declared NSDs built on top of AIX logical volumes as down . . . . .	282
Errors associated with storage pools, filesets and policies . . . . .	261	Verify logical volumes are properly defined for GPFS use. . . . .	282
A NO_SPACE error occurs when a file system is known to have adequate free space . . . . .	261	Check the volume group on each node. . . . .	282
Negative values occur in the 'predicted pool utilizations', when some files are 'ill-placed' . . . . .	263	Volume group varyon problems . . . . .	283
Policies - usage errors . . . . .	263	Disk accessing commands fail to complete due to problems with some non-IBM disks . . . . .	283
Errors encountered with policies . . . . .	264	Disk media failure. . . . .	283
Filesets - usage errors . . . . .	265	Replicated metadata and data . . . . .	284
Errors encountered with filesets . . . . .	265	Replicated metadata only . . . . .	285
Storage pools - usage errors . . . . .	266	Strict replication . . . . .	285
		No replication . . . . .	285
		GPFS error messages for disk media failures . . . . .	286
		Error numbers specific to GPFS application calls when disk failure occurs . . . . .	286
		Persistent Reserve errors . . . . .	287
		Understanding Persistent Reserve . . . . .	287
		Checking Persistent Reserve . . . . .	288
		Clearing a leftover Persistent Reserve reservation . . . . .	288
		Manually enabling or disabling Persistent Reserve . . . . .	289
		GPFS is not using the underlying multipath device . . . . .	290
		Kernel panics with a 'GPFS dead man switch timer has expired, and there's still outstanding I/O requests' message . . . . .	291
		<b>Chapter 17. Security issues . . . . . 293</b>	
		Encryption issues . . . . .	293

Unable to add encryption policies . . . . .	293		Swift PUT returns the 202 error and S3 PUT	
Receiving "Permission denied" message . . . . .	293		returns the 500 error due to the missing time	
"Value too large" failure when creating a file	293		synchronization . . . . .	321
Mount failure for a file system with encryption			Unable to generate the accurate container listing	
rules . . . . .	293		by performing the GET operation for unified file	
"Permission denied" failure of key rewrap . . . . .	294		and object access container . . . . .	322
E! Authentication issues . . . . .	294		Fatal error of object configuration during	
E! File protocol authentication setup issues . . . . .	294		deployment . . . . .	322
E! Protocol authentication issues . . . . .	294		Object authentication configuration fatal error	
Authentication error events . . . . .	295		during deployment . . . . .	323
			Fatal error of object authentication during	
			deployment . . . . .	323
<b>E! Chapter 18. Protocol issues . . . . .</b>	<b>297</b>			
E! NFS issues . . . . .	297		<b>Chapter 19. Disaster recovery issues 325</b>	
CES NFS failure due to network failure . . . . .	297		Disaster recovery setup problems . . . . .	325
E! NFS client with stale inode data . . . . .	297		Protocols cluster disaster recovery issues . . . . .	326
E! NFSV4 problems . . . . .	297		Other problems with disaster recovery . . . . .	326
NFS mount issues . . . . .	298			
NFS error events . . . . .	301		<b>Chapter 20. Performance issues . . . 327</b>	
NFS error scenarios . . . . .	302		Issues caused by the low-level system components	327
E! SMB issues . . . . .	304		Suboptimal performance due to high utilization	
E! Determining the health of integrated SMB			of the system level components . . . . .	327
E! server . . . . .	304		Suboptimal performance due to long IBM	
E! File access failure from an SMB client with			Spectrum Scale waiters . . . . .	327
E! sharing conflict . . . . .	305		Suboptimal performance due to networking	
SMB client on Linux fails with an "NT status			issues caused by faulty system components . . . . .	328
logon failure" . . . . .	306		Issues caused by the suboptimal setup or	
SMB client on Linux fails with the NT status			configuration of the IBM Spectrum Scale cluster . . . . .	329
password must change error message . . . . .	307		Suboptimal performance due to unbalanced	
SMB mount issues . . . . .	308		architecture and improper system level settings . . . . .	329
Net use on Windows fails with "System error			Suboptimal performance due to low values	
86" . . . . .	308		assigned to IBM Spectrum Scale configuration	
SMB error events . . . . .	309		parameters . . . . .	330
SMB access issues . . . . .	310		Suboptimal performance due to new nodes with	
E! Object issues . . . . .	311		default parameter values added to the cluster . . . . .	330
E! Getting started with troubleshooting object			Suboptimal performance due to low value	
E! issues . . . . .	311		assigned to QoSIO operation classes . . . . .	331
Authenticating the object service . . . . .	312		Suboptimal performance due to improper	
Authenticating or using the object service . . . . .	312		mapping of the file system NSDs to the NSD	
Accessing resources . . . . .	313		servers . . . . .	332
Connecting to the object services . . . . .	313		Suboptimal performance due to incompatible	
Creating a path . . . . .	313		file system block allocation type . . . . .	334
Constraints for creating objects and containers	314		Issues caused by the unhealthy state of the	
The Bind password is used when the object			components used . . . . .	335
authentication configuration has expired . . . . .	314		Suboptimal performance due to failover of	
The password used for running the keystone			NSDs to secondary server - NSD server failure . . . . .	335
command has expired or is incorrect . . . . .	315		Suboptimal performance due to failover of	
The LDAP server is not reachable . . . . .	315		NSDs to secondary server - Disk connectivity	
The TLS certificate has expired . . . . .	316		failure . . . . .	335
The TLS CACERT certificate has expired . . . . .	316		Suboptimal performance due to file system	
The TLS certificate on the LDAP server has			being fully utilized . . . . .	337
expired . . . . .	317		Suboptimal performance due to VERBS RDMA	
The SSL certificate has expired . . . . .	317		being inactive . . . . .	338
The signing certificate has expired . . . . .	318		Issues caused by the use of configurations or	
Users are not listed in the OpenStack user list	318		commands related to maintenance and operation . . . . .	340
IBM Security Lifecycle Manager cannot be			Suboptimal performance due to maintenance	
installed . . . . .	319		commands in progress . . . . .	340
The IBM Security Lifecycle Manager			Suboptimal performance due to frequent	
prerequisites cannot be installed . . . . .	320		invocation or execution of maintenance	
The error code signature does not match . . . . .	320		commands . . . . .	341



Suboptimal performance when a tracing is active on a cluster . . . . .	342	E! Start GPFS daemon . . . . .	363
Suboptimal performance due to replication settings being set to 2 or 3 . . . . .	342	E! Increase fileset space . . . . .	363
Suboptimal performance due to updates made on a file system or fileset with snapshot . . . . .	343	E! Synchronize node clocks. . . . .	363
Delays and deadlocks . . . . .	344	E! Start performance monitoring collector service	364
<b>E! Chapter 21. GUI issues . . . . .</b>	<b>347</b>	E! Start performance monitoring sensor service . . . . .	364
E! Understanding GUI support matrix and limitations	347	<b>E! Chapter 26. Support for troubleshooting . . . . .</b>	<b>367</b>
E! Examples for GUI issues and their resolutions . . . . .	347	E! Contacting IBM support center . . . . .	367
E! GUI fails to start . . . . .	347	E! Information to be collected before contacting the IBM Support Center . . . . .	367
E! GUI login page does not open. . . . .	348	E! How to contact the IBM Support Center . . . . .	369
E! GUI performance monitoring issues . . . . .	348	E! Call home . . . . .	370
E! GUI is showing "Server was unable to process the request" error . . . . .	350	<b>Chapter 27. References . . . . .</b>	<b>371</b>
E! GUI is displaying outdated information . . . . .	350	Events. . . . .	371
E! Capacity information is not available in GUI pages . . . . .	352	Messages . . . . .	443
<b>  Chapter 22. AFM issues . . . . .</b>	<b>353</b>	Message severity tags . . . . .	443
<b>  Chapter 23. AFM DR issues . . . . .</b>	<b>357</b>	<b>Accessibility features for IBM Spectrum Scale . . . . .</b>	<b>591</b>
<b>Chapter 24. Transparent Cloud Tiering issues . . . . .</b>	<b>359</b>	Accessibility features . . . . .	591
<b>E! Chapter 25. Recovery procedures . . . . .</b>	<b>361</b>	Keyboard navigation . . . . .	591
E! Restoring data and system configuration . . . . .	361	IBM and accessibility . . . . .	591
E! Automatic recovery . . . . .	361	<b>Notices . . . . .</b>	<b>593</b>
E! Upgrade recovery . . . . .	361	Trademarks . . . . .	594
E! Directed maintenance procedures. . . . .	362	Terms and conditions for product documentation	595
E! Start NSD . . . . .	362	IBM Online Privacy Statement. . . . .	595
E! Start cluster . . . . .	362	<b>Glossary . . . . .</b>	<b>597</b>
		<b>Index . . . . .</b>	<b>603</b>



## Tables

1.	IBM Spectrum Scale library information units	xi		35.	Attributes with their description	. . . . .	102
2.	Conventions	. . . . .		36.	gpfsClusterStatusTable: Cluster status information	. . . . .	109
3.	Documentation guides in this release.	. . . . .		37.	gpfsClusterConfigTable: Cluster configuration information	. . . . .	110
4.	Documentation structure in this release	xvi		38.	gpfsNodeStatusTable: Node status information	. . . . .	110
5.	Input requests to the <b>mmpmon</b> command	4		39.	gpfsNodeConfigTable: Node configuration information	. . . . .	110
6.	Keywords and values for the <b>mmpmon fs_io_s</b> response	. . . . .		40.	gpfsFileSystemStatusTable: File system status information	. . . . .	111
7.	Keywords and values for the <b>mmpmon io_s</b> response	. . . . .		41.	gpfsFileSystemPerfTable: File system performance information.	. . . . .	112
8.	nlist requests for the <b>mmpmon</b> command	8		42.	gpfsStgPoolTable: Storage pool information		113
9.	Keywords and values for the <b>mmpmon nlist add</b> response	. . . . .		43.	gpfsDiskStatusTable: Disk status information		113
10.	Keywords and values for the <b>mmpmon nlist del</b> response	. . . . .		44.	gpfsDiskConfigTable: Disk configuration information	. . . . .	114
11.	Keywords and values for the <b>mmpmon nlist new</b> response	. . . . .		45.	gpfsDiskPerfTable: Disk performance information	. . . . .	114
12.	Keywords and values for the <b>mmpmon nlist s</b> response	. . . . .		46.	Net-SNMP traps	. . . . .	115
13.	Keywords and values for the <b>mmpmon nlist failures</b>	. . . . .		E! 47.	IBM websites for help, services, and information	. . . . .	132
14.	Keywords and values for the <b>mmpmon reset</b> response	. . . . .		E! 48.	Core object log files in /var/log/swift		151
15.	rhist requests for the <b>mmpmon</b> command	16		E! 49.	Additional object log files in /var/log/swift		152
16.	Keywords and values for the <b>mmpmon rhist nr</b> response	. . . . .		E! 50.	General system log files in /var/adm/ras		152
17.	Keywords and values for the <b>mmpmon rhist off</b> response	. . . . .		E! 51.	Authentication log files	. . . . .	153
18.	Keywords and values for the <b>mmpmon rhist on</b> response	. . . . .		52.	. . . . .	. . . . .	319
19.	Keywords and values for the <b>mmpmon rhist p</b> response	. . . . .		E! 53.	Sensors available for each resource type		349
20.	Keywords and values for the <b>mmpmon rhist reset</b> response.	. . . . .		E! 54.	GUI refresh tasks	. . . . .	350
21.	Keywords and values for the <b>mmpmon rhist s</b> response	. . . . .		E! 55.	Troubleshooting details for capacity data display issues in GUI	. . . . .	352
22.	rpc_s requests for the <b>mmpmon</b> command	28		56.	Common questions in AFM with their resolution.	. . . . .	353
23.	Keywords and values for the <b>mmpmon rpc_s</b> response	. . . . .		57.	Common questions in AFM DR with their resolution.	. . . . .	357
24.	Keywords and values for the <b>mmpmon rpc_s size</b> response	. . . . .		58.	DMPs	. . . . .	362
25.	Keywords and values for the <b>mmpmon ver</b> response	. . . . .		59.	Events for the AUTH component	. . . . .	371
26.	Sensors available for each resource type	72		60.	Events for the CESNetwork component		373
27.	Sensors available to capture capacity details	73		61.	Events for the CloudGateway component		375
28.	Notification levels	. . . . .		62.	Events for the DISK component	. . . . .	375
29.	SNMP objects included in event notifications	80		63.	Events for the file system component		375
30.	SNMP OID ranges	. . . . .		64.	Events for the GPFS component	. . . . .	390
31.	AFM states and their description	. . . . .		65.	Events for the GUI component.	. . . . .	396
32.	AFM DR states and their description	. . . . .		66.	Events for the HadoopConnector component		397
33.	List of events that can be added using <b>mmaddcallback</b>	. . . . .		67.	Events for the KEYSTONE component		398
34.	Field description of the example	. . . . .		68.	Events for the NFS component.	. . . . .	399
				69.	Events for the Network component	. . . . .	402
				70.	Events for the object component	. . . . .	404
				71.	Events for the Performance component		409
				72.	Events for the SMB component	. . . . .	410
				73.	GUI events	. . . . .	411
				74.	Message severity tags ordered by priority		444



---

## About this information

This edition applies to IBM Spectrum Scale™ version 4.2.1 for AIX®, Linux, and Windows.

IBM Spectrum Scale is a file management infrastructure, based on IBM® General Parallel File System (GPFS™) technology, that provides unmatched performance and reliability with scalable access to critical file data.

To find out which version of IBM Spectrum Scale is running on a particular AIX node, enter:

```
lslpp -l gpfs\*
```

To find out which version of IBM Spectrum Scale is running on a particular Linux node, enter:

```
rpm -qa | grep gpfs
```

To find out which version of IBM Spectrum Scale is running on a particular Windows node, open the **Programs and Features** control panel. The IBM Spectrum Scale installed program name includes the version number.

## Which IBM Spectrum Scale information unit provides the information you need?

The IBM Spectrum Scale library consists of the information units listed in Table 1.

To use these information units effectively, you must be familiar with IBM Spectrum Scale and the AIX, Linux, or Windows operating system, or all of them, depending on which operating systems are in use at your installation. Where necessary, these information units provide some background information relating to AIX, Linux, or Windows; however, more commonly they refer to the appropriate operating system documentation.

**Note:** Throughout this documentation, the term “Linux” refers to all supported distributions of Linux, unless otherwise specified.

*Table 1. IBM Spectrum Scale library information units*

Information unit	Type of information	Intended users
<i>IBM Spectrum Scale: Administration Guide</i>	This information unit explains how to do the following: <ul style="list-style-type: none"><li>• Use the commands, programming interfaces, and user exits unique to GPFS</li><li>• Manage clusters, file systems, disks, and quotas</li><li>• Export a GPFS file system using the Network File System (NFS) protocol</li></ul>	System administrators or programmers of GPFS systems

Table 1. IBM Spectrum Scale library information units (continued)

Information unit	Type of information	Intended users
<i>IBM Spectrum Scale: Administration Guide</i>	<p>This information unit explains how to use the following advanced features of GPFS:</p> <ul style="list-style-type: none"> <li>• Accessing GPFS file systems from other GPFS clusters</li> <li>• Policy-based data management for GPFS</li> <li>• Creating and maintaining snapshots of GPFS file systems</li> <li>• Establishing disaster recovery for your GPFS cluster</li> <li>• Monitoring GPFS I/O performance with the <b>mmpmon</b> command</li> <li>• Miscellaneous advanced administration topics</li> </ul>	System administrators or programmers seeking to understand and use the advanced features of GPFS
<i>IBM Spectrum Scale: Concepts, Planning, and Installation Guide</i>	<p>This information unit provides information about the following topics:</p> <ul style="list-style-type: none"> <li>• Introducing GPFS</li> <li>• GPFS architecture</li> <li>• Planning concepts for GPFS</li> <li>• Installing GPFS</li> <li>• Migration, coexistence and compatibility</li> <li>• Applying maintenance</li> <li>• Configuration and tuning</li> <li>• Uninstalling GPFS</li> </ul>	System administrators, analysts, installers, planners, and programmers of GPFS clusters who are very experienced with the operating systems on which each GPFS cluster is based

Table 1. IBM Spectrum Scale library information units (continued)

Information unit	Type of information	Intended users
<i>IBM Spectrum Scale: Command and Programming Reference</i>	<p>This information unit describes the Data Management Application Programming Interface (DMAPI) for GPFS.</p> <p>This implementation is based on The Open Group's System Management: Data Storage Management (XD SM) API Common Applications Environment (CAE) Specification C429, The Open Group, ISBN 1-85912-190-X specification. The implementation is compliant with the standard. Some optional features are not implemented.</p> <p>The XD SM DMAPI model is intended mainly for a single-node environment. Some of the key concepts, such as sessions, event delivery, and recovery, required enhancements for a multiple-node environment such as GPFS.</p> <p>Use this information if you intend to write application programs to do the following:</p> <ul style="list-style-type: none"> <li>• Monitor events associated with a GPFS file system or with an individual file</li> <li>• Manage and maintain GPFS file system data</li> </ul>	Application programmers who are experienced with GPFS systems and familiar with the terminology and concepts in the XD SM standard
<i>IBM Spectrum Scale: Problem Determination Guide</i>	This information unit contains explanations of GPFS error messages and explains how to handle problems you may encounter with GPFS.	System administrators of GPFS systems who are experienced with the subsystems used to manage disks and who are familiar with the concepts presented in the <i>IBM Spectrum Scale: Concepts, Planning, and Installation Guide</i>

## Prerequisite and related information

For updates to this information, see IBM Spectrum Scale in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/STXKQY/ibmspectrumscale\\_welcome.html](http://www.ibm.com/support/knowledgecenter/STXKQY/ibmspectrumscale_welcome.html)).

For the latest support information, see the IBM Spectrum Scale FAQ in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html](http://www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html)).

## Conventions used in this information

Table 2 on page xiv describes the typographic conventions used in this information. UNIX file name conventions are used throughout this information.

**Note: Users of IBM Spectrum Scale for Windows** must be aware that on Windows, UNIX-style file names need to be converted appropriately. For example, the GPFS cluster configuration data is stored in the `/var/mmfs/gen/mmsdrfs` file. On Windows, the UNIX namespace starts under the `%SystemDrive%\cygwin64` directory, so the GPFS cluster configuration data is stored in the `C:\cygwin64\var\mmfs\gen\mmsdrfs` file.

Table 2. Conventions

Convention	Usage
<b>bold</b>	<p><b>Bold</b> words or characters represent system elements that you must use literally, such as commands, flags, values, and selected menu options.</p> <p>Depending on the context, <b>bold</b> typeface sometimes represents path names, directories, or file names.</p>
<b><u>bold underlined</u></b>	<p><b><u>bold underlined</u></b> keywords are defaults. These take effect if you do not specify a different keyword.</p>
constant width	<p>Examples and information that the system displays appear in constant-width typeface.</p> <p>Depending on the context, constant-width typeface sometimes represents path names, directories, or file names.</p>
<i>italic</i>	<p><i>Italic</i> words or characters represent variable values that you must supply.</p> <p><i>Italics</i> are also used for information unit titles, for the first use of a glossary term, and for general emphasis in text.</p>
<key>	<p>Angle brackets (less-than and greater-than) enclose the name of a key on the keyboard. For example, &lt;Enter&gt; refers to the key on your terminal or workstation that is labeled with the word <i>Enter</i>.</p>
\	<p>In command examples, a backslash indicates that the command or coding example continues on the next line. For example:</p> <pre>mkcondition -r IBM.FileSystem -e "PercentTotUsed &gt; 90" \ -E "PercentTotUsed &lt; 85" -m p "FileSystem space used"</pre>
{item}	<p>Braces enclose a list from which you must choose an item in format and syntax descriptions.</p>
[item]	<p>Brackets enclose optional items in format and syntax descriptions.</p>
<Ctrl-x>	<p>The notation &lt;Ctrl-x&gt; indicates a control character sequence. For example, &lt;Ctrl-c&gt; means that you hold down the control key while pressing &lt;c&gt;.</p>
item...	<p>Ellipses indicate that you can repeat the preceding item one or more times.</p>
	<p>In <i>synopsis</i> statements, vertical lines separate a list of choices. In other words, a vertical line means <i>Or</i>.</p> <p>In the left margin of the document, vertical lines indicate technical changes to the information.</p>

## How to send your comments

Your feedback is important in helping us to produce accurate, high-quality information. If you have any comments about this information or any other IBM Spectrum Scale documentation, send your comments to the following e-mail address:

**[mhvrcfs@us.ibm.com](mailto:mhvrcfs@us.ibm.com)**

Include the publication title and order number, and, if applicable, the specific location of the information about which you have comments (for example, a page number or a table number).

To contact the IBM Spectrum Scale development organization, send your comments to the following e-mail address:

**[gpfs@us.ibm.com](mailto:gpfs@us.ibm.com)**



---

## Summary of changes

This topic summarizes changes to the IBM Spectrum Scale licensed program and the IBM Spectrum Scale library. Within each information unit in the library, a vertical line (|) to the left of text and illustrations indicates technical changes or additions made to the previous edition of the information.

### Summary of changes for IBM Spectrum Scale version 4 release 2.1 as updated, July 2016

This release of the IBM Spectrum Scale licensed program and the IBM Spectrum Scale library include the following improvements:

#### Auditing configuration changes

A syslog entry is automatically written whenever a GPFS command makes a configuration change. Adding the information to the syslog gives flexibility in mining, processing, and redirecting these events. Entries can also be written to the standard GPFS log. The **commandAudit** parameter of the **mmchconfig** command controls this option. For more information, see the topic *Audit messages for cluster configuration changes* in the *IBM Spectrum Scale: Problem Determination Guide*.

#### Automated configuration of sensors for performance monitoring

IBM Spectrum Scale now supports automated configuration of sensors for its performance monitoring tool. For more information on automated configuration of sensors, see the *Automated configuration* section in the *IBM Spectrum Scale: Administration Guide*.

#### Callback event for file system structure errors

A new user callback event **fsstruct** (file system structure error) is triggered when the file system detects an error in the metadata. Immediate notification enables the callback program to act to mitigate further errors. For more information, see the topic *mmaddcallback command* in the *IBM Spectrum Scale: Command and Programming Reference* guide.

#### CES, NFS, and SMB protocols: Support for SLES V12 on x86 systems

Cluster Export Services (CES) partially supports SUSE Linux Enterprise Server (SLES) V12 on x86 systems. The SMB and NFS protocols are now supported via a manual installation process. For more information, see the topic *Manually installing IBM Spectrum Scale on SLES 12 systems* in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.

#### Compression support for FPO environments

File compression is expanded to support the File Placement Optimizer (FPO) environment. For the FPO environment, you must set the block group factor to a multiple of 10 to avoid degrading file system performance. For more information, see the topic *File compression* in the *IBM Spectrum Scale: Administration Guide*.

#### Deadlock management and debug data control

Deadlock management is extended with the following features:

- The detection thresholds for deadlocks are automatically adjusted according to waiter length and cluster overloaded status.
- New defaults more suitable for customer environments are established for the configuration variables **deadlockDataCollectionDailyLimit**, **deadlockDataCollectionMinInterval**, and others.

A new configuration variable **debugDataControl** controls the amount of debug data that is collected. The default setting is a minimal amount of debug information that is the most important for debugging issues. For more information, see the topic *Managing deadlocks* in the *IBM Spectrum Scale: Problem Determination Guide*.

## **/dev/<fs\_name> device for a file system on Linux**

On Linux, GPFS no longer creates the /dev/<fs\_name> device for a file system. Applications that relied on the file system device under /dev to be present, or that relied on "/dev" to be displayed in the output of the mount command, must find other ways to obtain the information. As a substitute, consider the information provided by the /etc/fstab file and /proc/mounts entries.

## **Documentation changes**

The IBM Spectrum Scale documentation guides are changed as follows:

*Table 3. Documentation guides in this release*

<b>Guides in release 4.2 and earlier</b>	<b>Guides in release 4.2.1</b>
<i>IBM Spectrum Scale: Concepts, Planning, and Installation Guide</i>	<i>IBM Spectrum Scale: Concepts, Planning, and Installation Guide</i>
IBM Spectrum Scale: Administration and Programming Reference Except commands, programming interfaces, and user exits	<i>IBM Spectrum Scale: Administration Guide</i>
IBM Spectrum Scale: Advanced Administration Guide	
IBM Spectrum Scale: Problem Determination Guide	
IBM Spectrum Scale: Data Management API Guide	<i>IBM Spectrum Scale: Command and Programming Reference</i>
IBM Spectrum Scale: Administration and Programming Reference (Commands, programming interfaces, and user exits)	

The top-level structure of the IBM Spectrum Scale documentation in the IBM Knowledge Center is changed as follows:

*Table 4. Documentation structure in this release*

<b>Documentation structure in release 4.2 and earlier</b>	<b>Documentation structure in release 4.2.1</b>
1. IBM Spectrum Scale: Concepts, Planning, and Installation Guide 2. IBM Spectrum Scale: Administration and Programming Reference 3. IBM Spectrum Scale: Advanced Administration Guide 4. IBM Spectrum Scale: Data Management API Guide 5. IBM Spectrum Scale: Problem Determination Guide	1. Summary of changes 2. Quick reference 3. Product overview 4. Planning 5. Installing and upgrading 6. Configuring 7. Administering 8. Monitoring 9. Troubleshooting 10. Command reference 11. Programming reference 12. Library and related publications 13. Glossary

## **Encryption: Simplified setup and Vormetric DSM support**

- A new console command **mmkeyserv** greatly simplifies the setup of encryption both on the key server and the client node. IBM Security Key Lifecycle Manager (SKLM) V2.5.0.4 or later (including V2.6) is required.
- Encryption support is added for key servers that run Vormetric Data Security Manager (DSM) V5.2.3 or later.

For more information, see the topic *Establishing an encryption-enabled environment* in the *IBM Spectrum Scale: Administration Guide*.

### **Federation in the performance monitoring tool**

A performance monitoring tool installation with multiple collectors is called a federation. Federation is introduced in performance monitoring to increase the scalability or the fault-tolerance of the performance monitoring system. For more information on federation, see the *Configuring multiple collectors* section in the *IBM Spectrum Scale: Administration Guide*.

### **Guided installation**

The **spectrumscale** installation toolkit now provides next step hints that are designed to help customers new to IBM Spectrum Scale with an easy workflow that helps customers to install and configure an IBM Spectrum Scale cluster.

### **Hadoop Support for IBM Spectrum Scale**

HDFS transparency now supports running the Hadoop Map/Reduce workload inside the virtual machine container, Docker.

Federation is introduced in HDFS to solve the HDFS NameNode scaling problem.

Hadoop distcp is used for data migration from HDFS to the IBM Spectrum Scale file system and between two IBM Spectrum Scale file systems.

For more information, see the following sections in the *IBM Spectrum Scale: Administration Guide*:

- *Docker support*
- *The HDFS transparency federation*
- *Hadoop distcp support*

HDFS transparency security has been introduced for the simple security mode and the Kerberos mode.

User authentication and authorization is weak in the simple mode. The data transfers and RPCs from the clients to the NameNode and DataNode are not encrypted. The Kerberos mode introduced in the Hadoop ecosystem provides a secure Hadoop environment.

For more information, see the *HDFS transparency security* section in the *IBM Spectrum Scale: Administration Guide*.

### **InfiniBand and RDMA performance**

Performance is improved for clusters that use InfiniBand and RDMA for their intranode communications network.

### **Linux on z Systems™: Expanded features**

The following IBM Spectrum Scale features are now available on Linux for z Systems:

- Quality of Service (QoS) support.
- Improved extended count key data (ECKD™) device handling: On different nodes, different bus IDs can refer to the same device.
- IBM Spectrum Scale GUI now supported on Linux for z Systems.

### **--metadata-only parameter for mmrestripefs**

A **--metadata-only** option for the **mmrestripefs** command allows the restripe to complete in less time than a full restripe of metadata and data. The savings in time is useful in situations where there is a concern about file system operations and you want to restripe. This operation is supported for migrating data off disks, rebalancing, restoring replication, and comparing replicas. For more information, see the topic *mmrestripefs command* in the *IBM Spectrum Scale: Command and Programming Reference* guide.

### **mmhealth: Monitoring services hosted on cluster nodes**

A new command, **mmhealth** is added to monitor the health status of nodes and different services hosted on nodes. The **mmhealth** command also displays the event logs responsible for the

unhealthy status of nodes and services, to analyze and determine the problem responsible for the service failure. For more information, see *mmhealth* command in the *IBM Spectrum Scale: Command and Programming Reference*.

### **Object storage improvements**

- Added support for starting and stopping the *ibmobjectizer* service. For information about *ibmobjectizer* service, see *Starting and Stopping the ibmobjectizer service* in *IBM Spectrum Scale: Administration Guide*.
- For problem determination, added potential problem scenarios with proposed solutions. The problem determination scenarios are listed here: *Object issues* in *IBM Spectrum Scale: Problem Determination Guide*.
- Added support for object encryption. For information about object encryption, see *Creating storage policy for encryption* in *IBM Spectrum Scale: Administration Guide*.
- Added new constraints for unified file and object access. The constraints are listed here: *Constraints applicable to unified file and object access* in *IBM Spectrum Scale: Administration Guide*.
- Added support for simplified enablement of S3. For information on S3, see *Changing the object base configuration to enable S3 API* in *IBM Spectrum Scale: Administration Guide*.
- Added support for multi-region object deployment with a highly available keystone service. For information about the multi-region object deployment, see *Authentication considerations for multi-region object deployment* in *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.
- Added support for OpenStack Liberty packages. For more information on Liberty packages, see *Protocol support overview: Integration of protocol access methods with GPFS* in *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.
- Added support to execute *mmobj* commands from any IBM Spectrum Scale client node. For more information on *mmobj*, see *mmobj* command in *IBM Spectrum Scale: Command and Programming Reference*.
- Added support for monitoring support for external AD and LDAP server for object authentication and main object services. For more information on external AD and LDAP server, see *Configuring an AD-based authentication for object access* in *IBM Spectrum Scale: Administration Guide*.

### **Quality of Service for I/O operations (QoS) improvements**

Quality of Service for I/O operations is expanded to support the File Placement Optimizer (FPO) environment. For more information, see the topic *Setting the Quality of Service for I/O operations (QoS)* in the *IBM Spectrum Scale: Administration Guide*.

### **Re-create and restore options for protocols cluster failover**

The failover procedure can choose between re-create and restore options. For more information on the failover options, see the *Performing failover for protocols cluster when primary cluster fails* section in the *IBM Spectrum Scale: Administration Guide*.

### **Re-create and restore options for failing back to an old primary for protocols cluster**

When failing back to an old primary, the file protocol configuration can either be re-created or restored. For more information on failing back to an old primary, see the *Performing failback to old primary for protocols cluster* section in the *IBM Spectrum Scale: Administration Guide*.

### **Re-create and restore options for failing back to a new primary for protocols cluster**

When failing back to a new primary, the file protocol configuration can either be re-created or restored. For more information on failing back to an old primary, see the *Performing failback to new primary for protocols cluster* section in the *IBM Spectrum Scale: Administration Guide*.

### **Support for Transparent Cloud Tiering**

The Transparent Cloud Tiering feature leverages the existing ILM policy available in IBM Spectrum Scale, and administrators can define policies to migrate cold data to a cloud storage tier or recall data from the cloud storage tier on reaching certain threshold levels.

A new command, *mmcloudgateway*, is added to manage and configure the cloud storage tier.

**Note:** To enable Transparent Cloud Tiering nodes, you must first enable the Transparent Cloud Tiering feature. This feature provides a new level of storage tiering capability to IBM Spectrum Scale customers. Please contact your IBM Client Technical Specialist (or send an email to <mailto:scale@us.ibm.com>) to review your use case of the Transparent Cloud Tiering feature and to obtain the instructions to enable the feature in your environment.

### **workerThreads tunes file system performance**

The **workerThreads** parameter of the **mmchconfig** command controls an integrated group of variables that tune file system performance. Use this variable to tune file systems in environments that are capable of high sequential or random read/write workloads or small-file activity. This variable can be used in any installation and is preferred over **worker1Threads** and **prefetchThreads** in new installations. For more information, see the topic *mmchconfig command* in the *IBM Spectrum Scale: Command and Programming Reference* guide.

### **IBM Spectrum Scale GUI changes**

The following main changes are added in the IBM Spectrum Scale management GUI:

- Renamed Monitoring > Topology page to NSDs. The NSDs page facilitates monitoring the status of Network Shared Disks (NSD) and nodes to NSD mapping in the system.
- Added new Monitoring > Nodes page in the GUI. The Nodes page provides an easy way to monitor the performance, health status, and configuration aspects of all available nodes in the IBM Spectrum Scale cluster. The properties of a node display the status of various CES services such as Object, NFS, and SMB as well as the authentication status of these services if they are enabled. It also displays other details such as network status, information on attached NSDs and file systems, and so on.
- Monitoring performance of transparent cloud tiering services through Performance and Dashboard pages.
- Renamed Monitoring > Performance page to Statistics.
- Added capacity monitoring options in the Statistics page.
- Added monitoring options for GPFS waiters in Monitoring > Statistics panel.
- The following improvements are made in the Dashboards page:
  - You can assign a name to the dashboards and the user can switch between dashboards.
  - The dashboards are now stored on the server instead of the browser. Therefore, it can be shared among users and browsers.
  - Default dashboards are shipped with the GUI. When you open the IBM Spectrum Scale™ GUI after the installation or upgrade, you can see the default dashboards. You can further modify or delete the default dashboards to suit your requirements.
- Renamed Download Logs page as Diagnostic Data. Now, the GUI can be used instead of the **gpfs.snap** command to collect the details of the issue. For more information on collecting diagnostic data through GUI, see *Collecting diagnostic data through GUI* topic in the *IBM Spectrum Scale: Problem Determination Guide*.
- The Files > Information Lifecycle page facilitates defining compression and deletion rules.
- The new Settings > Object Service page facilitates start and stop feature for object services.
- Up to 1000 nodes are supported.
- The GUI can now be used on an IBM Spectrum Scale cluster where sudo wrappers are used. For more information on how to configure IBM Spectrum Scale GUI to use sudo wrapper, see *Configuring IBM Spectrum Scale GUI to use sudo wrapper* in *IBM Spectrum Scale: Administration Guide*.
- IBM Spectrum Scale GUI support for System z® platform is available on RHEL7.2 and SLES12.
- By default, GUI commands that change the configuration of the cluster cause an audit message to be sent to syslog. Optionally, an audit message can also be sent to the GPFS log. For more information, see the topic *Audit messages for cluster configuration changes* in the *IBM Spectrum Scale: Problem Determination Guide*.

## **NFS and SMB protocol troubleshooting information added**

New AD Discovery tool to query and validate several AD settings.

New troubleshooting information for NFS issues:

- NFS mount issues
- NFS error events
- NFS error scenarios

New troubleshooting information for SMB issues:

- SMB client on Linux failures
- SMB mount errors
- SMB error events
- SMB access issues

## **Documented commands, structures, and subroutines**

The following lists the modifications to the documented commands, structures, and subroutines:

### **New commands**

The following commands are new:

- **mmadquery**
- **mmcloudgateway**
- **mmhealth**
- **mmkeyserv**

### **New structures**

There are no new structures.

### **New subroutines**

There are no new subroutines.

### **Changed commands**

The following commands were changed:

- **gpfs.snap**
- **mmafmlocal**
- **mmcallhome**
- **mmchconfig**
- **mmchnode**
- **mmcesdr**:
- **mmcrsnapshot**: You can create multiple snapshots in the same command.
- **mmdelsnapshot**: You can delete multiple snapshots in the same command.
- **mmlscluster**
- **mmnfs**
- **mmobj**
- **mmrestripefs**
- **mmsmb**
- **mmprotocoltrace**
- **mmsmb**
- **mmuserauth**

### **Changed structures**

There are no changed structures.

### **Changed subroutines**

**gpfs\_iopen()** subroutine

**gpfs\_iopen64()** subroutine

**Deleted commands**

There are no deleted commands.

**Deleted structures**

There are no deleted structures.

**Deleted subroutines**

There are no deleted subroutines.

**Messages**

The following lists the new, changed, and deleted messages:

**New messages**

6027-1826, 6027-2363, 6027-2364, 6027-2365, 6027-2366, 6027-2367, 6027-2368, 6027-2369,  
6027-2370, 6027-2371, 6027-2372, 6027-2373, 6027-2374, 6027-2375, 6027-2376, 6027-2377,  
6027-2378, 6027-3108, 6027-3720, 6027-3721, 6027-3722, 6027-3723, 6027-3724, 6027-3725,  
6027-3726, 6027-3727, 6027-3728, 6027-3915, 6027-3916, 6027-3594, 6027-3595, 6027-3596

**Changed messages**

6027-1368, 6027-1235, 6027-1545, 6027-2271, 6027-2272, 6027-2273, 6027-2274, 6027-2951

**Deleted messages**

6027-1997





---

# Chapter 1. Performance monitoring

With IBM Spectrum Scale, system administrators can monitor the performance of GPFS and the communications protocols that it uses.

Performance topics include:

- “Network performance monitoring”
- “Performance monitoring tool overview” on page 41

---

## Network performance monitoring

Network performance can be monitored with Remote Procedure Call (RPC) statistics.

The GPFS daemon caches statistics relating to RPCs. Most statistics are related to RPCs sent to other nodes. This includes a set of up to seven statistics cached per node and one statistic that is cached per size of the RPC message. For RPCs received from other nodes, one statistic is cached for each type of RPC message. The counters are measured in seconds and milliseconds

The statistics cached per node are the following:

### Channel wait time

The amount of time the RPC must wait for access to a communication channel to the target node.

### Send time TCP

The amount of time to transfer an RPC message to an Ethernet interface.

### Send time verbs

The amount of time to transfer an RPC message to an InfiniBand interface.

### Receive time TCP

The amount of time to transfer an RPC message from an Ethernet interface into the daemon.

### Latency TCP

The latency of the RPC when sent and received over an Ethernet interface.

### Latency verbs

The latency of the RPC when sent and received over an InfiniBand interface.

### Latency mixed

The latency of the RPC when sent over one type of interface (Ethernet or InfiniBand) and received over the other (InfiniBand or Ethernet).

If an InfiniBand network is not configured, no statistics are cached for send time verbs, latency verbs, and latency mixed.

The latency of an RPC is defined as the round-trip time minus the execution time on the target node. The round-trip time is measured from the start of writing the RPC message to the interface until the RPC reply is completely received. The execution time is measured on the target node from the time the message is completely received until the time the reply is sent. The latency, therefore, is the amount of time the RPC is being transmitted and received over the network and is a relative measure of the network performance as seen by the GPFS daemon.

There is a statistic associated with each of a set of size ranges, each with an upper bound that is a power of 2. The first range is 0 through 64, then 65 through 128, then 129 through 256, and then continuing until the last range has an upper bound of twice the **maxBlockSize**. For example, if the **maxBlockSize** is 1 MB, the upper bound of the last range is 2,097,152 (2 MB). For each of these ranges, the associated statistic is

the latency of the RPC whose size falls within that range. The size of an RPC is the amount of data sent plus the amount of data received. However, if one amount is more than 16 times greater than the other, only the larger amount is used as the size of the RPC.

The final statistic associated with each type of RPC message, on the node where the RPC is received, is the execution time of the RPC.

Each of the statistics described so far is actually an aggregation of values. By default, an aggregation consists of 60 one-second intervals, 60 one-minute intervals, 24 one-hour intervals, and 30 one-day intervals. Each interval consists of a sum of values accumulated during the interval, a count of values added into the sum, the minimum value added into the sum, and the maximum value added into the sum. Sixty seconds after the daemon starts, each of the one-second intervals contains data and every second thereafter the oldest interval is discarded and a new one entered. An analogous pattern holds for the minute, hour, and day periods.

As each RPC reply is received, the following information is saved in a *raw statistics buffer*:

- channel wait time
- send time
- receive time
- latency
- length of data sent
- length of data received
- flags indicating if the RPC was sent or received over InfiniBand
- target node identifier

As each RPC completes execution, the execution time for the RPC and the message type of the RPC is saved in a *raw execution buffer*. Once per second these raw buffers are processed and the values are added to the appropriate aggregated statistic. For each value, the value is added to the statistic's sum, the count is incremented, and the value is compared to the minimum and maximum, which are adjusted as appropriate. Upon completion of this processing, for each statistic the sum, count, minimum, and maximum values are entered into the next one-second interval.

Every 60 seconds, the sums and counts in the 60 one-second intervals are added into a one-minute sum and count. The smallest of the 60 minimum values is determined, and the largest of the 60 maximum values is determined. This one-minute sum, count, minimum, and maximum are then entered into the next one-minute interval.

An analogous pattern holds for the minute, hour, and day periods. For any one particular interval, the sum is the sum of all raw values processed during that interval, the count is the count of all values during that interval, the minimum is the minimum of all values during that interval, and the maximum is the maximum of all values during that interval.

When statistics are displayed for any particular interval, an average is calculated from the sum and count, then the average, minimum, maximum, and count are displayed. The average, minimum and maximum are displayed in units of milliseconds, to three decimal places (one microsecond granularity).

The following **mmchconfig** attributes are available to control the RPC buffers and intervals:

- **rpcPerfRawStatBufferSize**
- **rpcPerfRawExecBufferSize**
- **rpcPerfNumberSecondIntervals**
- **rpcPerfNumberMinuteIntervals**
- **rpcPerfNumberHourIntervals**
- **rpcPerfNumberDayIntervals**

The **mmdiag** command with the **--rpc** parameter can be used to query RPC statistics.

For more information, see the topics *mmchconfig command* and *mmdiag command* in the *IBM Spectrum Scale: Administration Guide*.

---

## Monitoring GPFS I/O performance with the mmpmon command

Use the **mmpmon** command to monitor GPFS performance on the node in which it is run, and other specified nodes.

Before attempting to use the **mmpmon** command, review the command documentation in the *IBM Spectrum Scale: Administration Guide*.

Next, read all of the following relevant **mmpmon** topics.

- “Overview of mmpmon”
- “Specifying input to the mmpmon command”
- “Example mmpmon scenarios and how to analyze and interpret their results” on page 31
- “Other information about mmpmon output” on page 40

### Overview of mmpmon

The **mmpmon** facility allows the system administrator to collect I/O statistics from the point of view of GPFS servicing application I/O requests.

The collected data can be used for many purposes, including:

- Tracking I/O demand over longer periods of time - weeks or months.
- Recording I/O patterns over time (when peak usage occurs, and so forth).
- Determining if some nodes service more application demand than others.
- Monitoring the I/O patterns of a single application which is spread across multiple nodes.
- Recording application I/O request service times.

Figure 1 shows the software layers in a typical system with GPFS. **mmpmon** is built into GPFS.

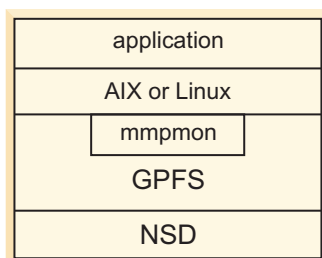


Figure 1. Node running mmpmon

### Specifying input to the mmpmon command

The input requests to the **mmpmon** command allow the system administrator to collect I/O statistics per mounted file system (**fs\_io\_s**) or for the entire node (**io\_s**).

The **mmpmon** command must be run using root authority. For command syntax, see **mmpmon** in the *IBM Spectrum Scale: Administration Guide*.

The **mmpmon** command is controlled by an input file that contains a series of requests, one per line. This input can be specified with the **-i** flag, or read from standard input (stdin). Providing input using stdin allows **mmpmon** to take keyboard input or output piped from a user script or application.

Leading blanks in the input file are ignored. A line beginning with a pound sign (#) is treated as a comment. Leading blanks in a line whose first non-blank character is a pound sign (#) are ignored.

Table 5 describes the **mmpmon** input requests.

*Table 5. Input requests to the mmpmon command*

Request	Description
fs_io_s	"Display I/O statistics per mounted file system" on page 5
io_s	"Display I/O statistics for the entire node" on page 6
nlist add <i>name</i> [ <i>name</i> ...]	"Add node names to a list of nodes for mmpmon processing" on page 8
nlist del	"Delete a node list" on page 9
nlist new <i>name</i> [ <i>name</i> ...]	"Create a new node list" on page 10
nlist s	"Show the contents of the current node list" on page 10
nlist sub <i>name</i> [ <i>name</i> ...]	"Delete node names from a list of nodes for mmpmon processing" on page 12
once <i>request</i>	Indicates that the request is to be performed only once.
reset	"Reset statistics to zero" on page 15
rhist nr	"Changing the request histogram facility request size and latency ranges" on page 18
rhist off	"Disabling the request histogram facility" on page 20. This is the default.
rhist on	"Enabling the request histogram facility" on page 21
rhist p	"Displaying the request histogram facility pattern" on page 22
rhist reset	"Resetting the request histogram facility data to zero" on page 24
rhist s	"Displaying the request histogram facility statistics values" on page 25
rpc_s	"Displaying the aggregation of execution time for Remote Procedure Calls (RPCs)" on page 28
rpc_s size	"Displaying the Remote Procedure Call (RPC) execution time according to the size of messages" on page 29
source <i>filename</i>	"Using request <i>source</i> and prefix directive <i>once</i> " on page 34
ver	"Displaying mmpmon version" on page 31
vio_s	"Displaying vdisk I/O statistics". See <i>IBM Spectrum Scale RAID: Administration</i> for more information.
vio_s_reset	"Resetting vdisk I/O statistics". See <i>IBM Spectrum Scale RAID: Administration</i> for more information.

## Running mmpmon on multiple nodes

Invoke **mmpmon** list requests on a single node for mmpmon request processing on multiple nodes in a local cluster.

The **mmpmon** command may be invoked on one node to submit requests to multiple nodes in a local GPFS cluster by using the **nlist** requests. See "Understanding the node list facility" on page 8.

## Running mmpmon concurrently from multiple users on the same node

Multiple instances of **mmpmon** can run on the same node so that different performance analysis applications and scripts can use the same performance data.

Five instances of **mmpmon** may be run on a given node concurrently. This is intended primarily to allow different user-written performance analysis applications or scripts to work with the performance data. For example, one analysis application might deal with **fs\_io\_s** and **io\_s** data, while another one deals with **rhist** data, and another gathers data from other nodes in the cluster. The applications might be separately written or separately maintained, or have different sleep and wake-up schedules.

Be aware that there is only one set of counters for **fs\_io\_s** and **io\_s** data, and another, separate set for **rhist** data. Multiple analysis applications dealing with the same set of data must coordinate any activities that could reset the counters, or in the case of **rhist** requests, disable the feature or modify the ranges.

## Display I/O statistics per mounted file system

The **fs\_io\_s** input request to the **mmpmon** command allows the system administrator to collect I/O statistics per mounted file system.

The **fs\_io\_s** (file system I/O statistics) request returns strings containing I/O statistics taken over all mounted file systems as seen by that node, and are presented as total values for each file system. The values are cumulative since the file systems were mounted or since the last **reset** request, whichever is most recent. When a file system is unmounted, its statistics are lost.

Read and write statistics are recorded separately. The statistics for a given file system are for the file system activity on the node running **mmpmon**, not the file system in total (across the cluster).

Table 6 describes the keywords for the **fs\_io\_s** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag.

Table 6. Keywords and values for the **mmpmon fs\_io\_s** response

Keyword	Description
<b>_n_</b>	IP address of the node responding. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.
<b>_cl_</b>	Name of the cluster that owns the file system.
<b>_fs_</b>	The name of the file system for which data are being presented.
<b>_d_</b>	The number of disks in the file system.
<b>_br_</b>	Total number of bytes read, from both disk and cache.
<b>_bw_</b>	Total number of bytes written, to both disk and cache.
<b>_oc_</b>	Count of <b>open()</b> call requests serviced by GPFS. This also includes <b>creat()</b> call counts.
<b>_cc_</b>	Number of <b>close()</b> call requests serviced by GPFS.
<b>_rdc_</b>	Number of application read requests serviced by GPFS.
<b>_wc_</b>	Number of application write requests serviced by GPFS.
<b>_dir_</b>	Number of <b>readdir()</b> call requests serviced by GPFS.
<b>_iu_</b>	Number of inode updates to disk.

### Example of mmpmon fs\_io\_s request

This is an example of the **fs\_io\_s** input request to the **mmpmon** command and the resulting output that displays the I/O statistics per mounted file system.

Assume that **commandFile** contains this line:

fs\_io\_s

and this command is issued:

```
mmpmon -p -i commandFile
```

The output is two lines in total, and similar to this:

```
_fs_io_s_n_199.18.1.8_nn_node1_rc_0_t_1066660148_tu_407431_cl_myCluster.xxx.com  
_fs_gpfs2_d_2_br_6291456_bw_314572800_oc_10_cc_16_rdc_101_wc_300_dir_7_iu_2  
_fs_io_s_n_199.18.1.8_nn_node1_rc_0_t_1066660148_tu_407455_cl_myCluster.xxx.com  
_fs_gpfs1_d_3_br_5431636_bw_173342800_oc_6_cc_8_rdc_54_wc_156_dir_3_iu_6
```

The output consists of one string per mounted file system. In this example, there are two mounted file systems, **gpfs1** and **gpfs2**.

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node1 fs_io_s OK  
cluster: myCluster.xxx.com  
filesystem: gpfs2  
disks: 2  
timestamp: 1066660148/407431  
bytes read: 6291456  
bytes written: 314572800  
opens: 10  
closes: 16  
reads: 101  
writes: 300  
readdir: 7  
inode updates: 2
```

```
mmpmon node 199.18.1.8 name node1 fs_io_s OK  
cluster: myCluster.xxx.com  
filesystem: gpfs1  
disks: 3  
timestamp: 1066660148/407455  
bytes read: 5431636  
bytes written: 173342800  
opens: 6  
closes: 8  
reads: 54  
writes: 156  
readdir: 3  
inode updates: 6
```

When no file systems are mounted, the responses are similar to:

```
_fs_io_s_n_199.18.1.8_nn_node1_rc_1_t_1066660148_tu_407431_cl_- _fs_-
```

The **\_rc\_** field is nonzero and the both the **\_fs\_** and **\_cl\_** fields contains a minus sign. If the **-p** flag is not specified, the results are similar to:

```
mmpmon node 199.18.1.8 name node1 fs_io_s status 1  
no file systems mounted
```

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

## Display I/O statistics for the entire node

The **io\_s** input request to the **mmpmon** command allows the system administrator to collect I/O statistics for the entire node.

The **io\_s** (I/O statistics) request returns strings containing I/O statistics taken over all mounted file systems as seen by that node, and are presented as total values for the entire node. The values are

cumulative since the file systems were mounted or since the last **reset**, whichever is most recent. When a file system is unmounted, its statistics are lost and its contribution to the total node statistics vanishes. Read and write statistics are recorded separately.

Table 7 describes the keywords for the **io\_s** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag.

Table 7. Keywords and values for the **mmpmon io\_s** response

Keyword	Description
<b>_n_</b>	IP address of the node responding. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.
<b>_br_</b>	Total number of bytes read, from both disk and cache.
<b>_bw_</b>	Total number of bytes written, to both disk and cache.
<b>_oc_</b>	Count of <b>open()</b> call requests serviced by GPFS. The open count also includes <b>creat()</b> call counts.
<b>_cc_</b>	Number of <b>close()</b> call requests serviced by GPFS.
<b>_rdc_</b>	Number of application read requests serviced by GPFS.
<b>_wc_</b>	Number of application write requests serviced by GPFS.
<b>_dir_</b>	Number of <b>readdir()</b> call requests serviced by GPFS.
<b>_iu_</b>	Number of inode updates to disk. This includes inodes flushed to disk because of access time updates.

### Example of mmpmon io\_s request

This is an example of the **io\_s** input request to the **mmpmon** command and the resulting output that displays the I/O statistics for the entire node.

Assume that **commandFile** contains this line:

```
io_s
```

and this command is issued:

```
mmpmon -p -i commandFile
```

The output is one line in total, and similar to this:

```
_io_s_ _n_ 199.18.1.8 _nn_ node1 _rc_ 0 _t_ 1066660148 _tu_ 407431 _br_ 6291456
_bw_ 314572800 _oc_ 10 _cc_ 16 _rdc_ 101 _wc_ 300 _dir_ 7 _iu_ 2
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node1 io_s OK
timestamp: 1066660148/407431
bytes read: 6291456
bytes written: 314572800
opens: 10
closes: 16
reads: 101
writes: 300
readdir: 7
inode updates: 2
```

## Understanding the node list facility

The node list facility can be used to invoke **mmpmon** on multiple nodes and gather data from other nodes in the cluster. The following table describes the **nlist** requests for the **mmpmon** command.

Table 8. *nlist* requests for the **mmpmon** command

Request	Description
<b>nlist add</b> <i>name</i> [ <i>name...</i> ]	“Add node names to a list of nodes for mmpmon processing”
<b>nlist del</b>	“Delete a node list” on page 9
<b>nlist new</b> <i>name</i> [ <i>name...</i> ]	“Create a new node list” on page 10
<b>nlist s</b>	“Show the contents of the current node list” on page 10
<b>nlist sub</b> <i>name</i> [ <i>name...</i> ]	“Delete node names from a list of nodes for mmpmon processing” on page 12

When specifying node names, keep these points in mind:

1. A node name of '.' (dot) indicates the current node.
2. A node name of '\*' (asterisk) indicates all currently connected local cluster nodes.
3. The nodes named in the node list must belong to the local cluster. Nodes in remote clusters are not supported.
4. A node list can contain nodes that are currently down. When an inactive node comes up, **mmpmon** will attempt to gather data from it.
5. If a node list contains an incorrect or unrecognized node name, all other entries in the list are processed. Suitable messages are issued for an incorrect node name.
6. When **mmpmon** gathers responses from the nodes in a node list, the full response from one node is presented before the next node. Data is not interleaved. There is no guarantee of the order of node responses.
7. The node that issues the **mmpmon** command need not appear in the node list. The case of this node serving only as a collection point for data from other nodes is a valid configuration.

### Add node names to a list of nodes for mmpmon processing

The **nlist add** (node list add) request is used to add node names to a list of nodes for **mmpmon** to collect their data. The node names are separated by blanks.

Table 9 describes the keywords for the **nlist add** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag.

Table 9. *Keywords and values for the mmpmon nlist add response*

Keyword	Description
<b>_n_</b>	IP address of the node processing the node list. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_req_</b>	The action requested. In this case, the value is add.
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.
<b>_c_</b>	The number of nodes in the user-supplied list.
<b>_ni_</b>	Node name input. A user-supplied node name from the offered list of names.



Table 9. Keywords and values for the **mmpmon nlist add** response (continued)

Keyword	Description
<b>_nx_</b>	Node name translation. The preferred GPFS name for the node.
<b>_nxip_</b>	Node name translated IP address. The preferred GPFS IP address for the node.
<b>_did_</b>	The number of nodes names considered valid and processed by the requests.
<b>_nlc_</b>	The number of nodes in the node list now (after all processing).

If the **nlist add** request is issued when no node list exists, it is handled as if it were an **nlist new** request.

### Example of mmpmon nlist add request:

This topic is an example of the **nlist add** request to add node names to a list of nodes for **mmpmon** processing and the output that displays.

A two- node cluster has nodes **node1** (199.18.1.2), a non-quorum node, and **node2** (199.18.1.5), a quorum node. A remote cluster has node **node3** (199.18.1.8). The **mmpmon** command is run on **node1**.

Assume that **commandFile** contains this line:

```
nlist add n2 199.18.1.2
```

and this command is issued:

```
mmpmon -p -i commandFile
```

Note in this example that an alias name **n2** was used for **node2**, and an IP address was used for **node1**. Notice how the values for **\_ni\_** and **\_nx\_** differ in these cases.

The output is similar to this:

```
_nlist_n_ 199.18.1.2 _nn_ node1 _req_ add _rc_ 0 _t_ 1121955894 _tu_ 261881 _c_ 2
_nlist_n_ 199.18.1.2 _nn_ node1 _req_ add _rc_ 0 _t_ 1121955894 _tu_ 261881 _ni_ n2 _nx_
node2 _nxip_ 199.18.1.5
_nlist_n_ 199.18.1.2 _nn_ node1 _req_ add _rc_ 0 _t_ 1121955894 _tu_ 261881 _ni_
199.18.1.2 _nx_ node1 _nxip_ 199.18.1.2
_nlist_n_ 199.18.1.2 _nn_ node1 _req_ add _rc_ 0 _t_ 1121955894 _tu_ 261881 _did_ 2 _nlc_
2
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.2 name node1 nlist add
initial status 0
name count 2
timestamp 1121955879/468858
node name n2, OK (name used: node2, IP address 199.18.1.5)
node name 199.18.1.2, OK (name used: node1, IP address 199.18.1.2)
final status 0
node names processed 2
current node list count 2
```

The requests **nlist add** and **nlist sub** behave in a similar way and use the same keyword and response format.

These requests are rejected if issued while quorum has been lost.

### Delete a node list

The **nlist del** (node list delete) request deletes a node list if one exists. If no node list exists, the request succeeds and no error code is produced.

Table 10 describes the keywords for the **nlist del** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag.

Table 10. Keywords and values for the **mmpmon nlist del** response

Keyword	Description
<b>_n_</b>	IP address of the node responding. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_req_</b>	The action requested. In this case, the value is <b>del</b> .
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.

### Example of mmpmon nlist del request:

This topic is an example of the **nlist del** request to delete a node list and the output that displays.

Assume that **commandFile** contains this line:

```
nlist del
```

and this command is issued:

```
mmpmon -p -i commandFile
```

The output is similar to this:

```
_nlist_ _n_ 199.18.1.2 _nn_ node1 _req_ del _rc_ 0 _t_ 1121956817 _tu_ 46050
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.2 name node1 nlist del status OK timestamp 1121956908/396381
```

### Create a new node list

The **nlist new** (node list new) request deletes the current node list if one exists, creates a new, empty node list, and then attempts to add the specified node names to the node list. The node names are separated by blanks.

Table 11 describes the keywords for the **nlist new** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag.

Table 11. Keywords and values for the **mmpmon nlist new** response

Keyword	Description
<b>_n_</b>	IP address of the node responding. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_req_</b>	The action requested. In this case, the value is <b>new</b> .
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.

### Show the contents of the current node list

The **nlist s** (node list show) request displays the current contents of the node list. If no node list exists, a count of zero is returned and no error is produced.

Table 12 describes the keywords for the **nlist s** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag.

Table 12. Keywords and values for the **mmpmon nlist s** response

Keyword	Description
<b>_n_</b>	IP address of the node processing the request. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_req_</b>	The action requested. In this case, the value is <b>s</b> .
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.
<b>_c_</b>	Number of nodes in the node list.
<b>_mbr_</b>	GPFS preferred node name for the list member.
<b>_ip_</b>	GPFS preferred IP address for the list member.

### Example of mmpmon nlist s request:

This topic is an example of the **nlist s** request to show the contents of the current node list and the output that displays.

Assume that **commandFile** contains this line:

```
nlist s
```

and this command is issued:

```
mmpmon -p -i commandFile
```

The output is similar to this:

```
_nlist_n_ 199.18.1.2 _nn_ node1 _req_ s _rc_ 0 _t_ 1121956950 _tu_ 863292 _c_ 2
_nlist_n_ 199.18.1.2 _nn_ node1 _req_ s _rc_ 0 _t_ 1121956950 _tu_ 863292 _mbr_ node1
_ip_ 199.18.1.2
_nlist_n_ 199.18.1.2 _nn_ node1 _req_ s _rc_ 0 _t_ 1121956950 _tu_ 863292 _mbr_
node2 _ip_ 199.18.1.5
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.2 name node1 nlist s
status 0
name count 2
timestamp 1121957505/165931
node name node1, IP address 199.18.1.2
node name node2, IP address 199.18.1.5
```

If there is no node list, the response looks like:

```
_nlist_n_ 199.18.1.2 _nn_ node1 _req_ s _rc_ 0 _t_ 1121957395 _tu_ 910440 _c_ 0
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.2 name node1 nlist s
status 0
name count 0
timestamp 1121957436/353352
the node list is empty
```

The **nlist s** request is rejected if issued while quorum has been lost. Only one response line is presented.

```
_failed_ _n_ 199.18.1.8 _nn_ node2 _rc_ 668 _t_ 1121957395 _tu_ 910440
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node2: failure status 668 timestamp 1121957395/910440  
lost quorum
```

## Delete node names from a list of nodes for mmpmon processing

The **nlist sub** (subtract a node from the node list) request removes a node from a list of node names.

These keywords and responses are similar to the **nlist add** request. The **\_req\_** keyword (action requested) for **nlist sub** is **sub**.

For more information, see the topic “Add node names to a list of nodes for mmpmon processing” on page 8.

## Node list examples and error handling

The **nlist** facility can be used to obtain GPFS performance data from nodes other than the one on which the **mmpmon** command is invoked. This information is useful to see the flow of GPFS I/O from one node to another, and spot potential problems.

### A successful **fs\_io\_s** request propagated to two nodes:

This topic is an example of a successful **fs\_io\_s** request to two nodes to display the I/O statistics per mounted file system and the resulting system output.

This command is issued:

```
mmpmon -p -i command_file
```

where **command\_file** has this:

```
nlist new node1 node2  
fs_io_s
```

The output is similar to this:

```
_fs_io_s_ _n_ 199.18.1.2 _nn_ node1 _rc_ 0 _t_ 1121974197 _tu_ 278619 _cl_  
xxx.localdomain _fs_ gpfs2 _d_ 2 _br_ 0 _bw_ 0 _oc_ 0 _cc_ 0 _rdc_ 0 _wc_ 0  
_dir_ 0 _iu_ 0  
_fs_io_s_ _n_ 199.18.1.2 _nn_ node1 _rc_ 0 _t_ 1121974197 _tu_ 278619 _cl_  
xxx.localdomain _fs_ gpfs1 _d_ 1 _br_ 0 _bw_ 0 _oc_ 0 _cc_ 0 _rdc_ 0 _wc_ 0  
_dir_ 0 _iu_ 0  
_fs_io_s_ _n_ 199.18.1.5 _nn_ node2 _rc_ 0 _t_ 1121974167 _tu_ 116443 _cl_  
cl1.xxx.com _fs_ fs3 _d_ 3 _br_ 0 _bw_ 0 _oc_ 0 _cc_ 0 _rdc_ 0 _wc_ 0 _dir_ 0  
_iu_ 3  
_fs_io_s_ _n_ 199.18.1.5 _nn_ node2 _rc_ 0 _t_ 1121974167 _tu_ 116443 _cl_  
cl1.xxx.comm _fs_ fs2 _d_ 2 _br_ 0 _bw_ 0 _oc_ 0 _cc_ 0 _rdc_ 0 _wc_ 0 _dir_ 0  
_iu_ 0  
_fs_io_s_ _n_ 199.18.1.5 _nn_ node2 _rc_ 0 _t_ 1121974167 _tu_ 116443 _cl_  
xxx.localdomain _fs_ gpfs2 _d_ 2 _br_ 0 _bw_ 0 _oc_ 0 _cc_ 0 _rdc_ 0 _wc_ 0  
_dir_ 0 _iu_ 0
```

The responses from a propagated request are the same as they would have been if issued on each node separately.

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.2 name node1 fs_io_s OK  
cluster: xxx.localdomain  
filesystem: gpfs2  
disks: 2  
timestamp: 1121974088/463102  
bytes read: 0
```

```
bytes written: 0
opens: 0
closes: 0
reads: 0
writes: 0
readdir: 0
inode updates: 0

mmpmon node 199.18.1.2 name node1 fs_io_s OK
cluster: xxx.localdomain
filesystem: gpfs1
disks: 1
timestamp: 1121974088/463102
bytes read: 0
bytes written: 0
opens: 0
closes: 0
reads: 0
writes: 0
readdir: 0
inode updates: 0

mmpmon node 199.18.1.5 name node2 fs_io_s OK
cluster: c11.xxx.com
filesystem: fs3
disks: 3
timestamp: 1121974058/321741
bytes read: 0
bytes written: 0
opens: 0
closes: 0
reads: 0
writes: 0
readdir: 0
inode updates: 2

mmpmon node 199.18.1.5 name node2 fs_io_s OK
cluster: c11.xxx.com
filesystem: fs2
disks: 2
timestamp: 1121974058/321741
bytes read: 0
bytes written: 0
opens: 0
closes: 0
reads: 0
writes: 0
readdir: 0
inode updates: 0

mmpmon node 199.18.1.5 name node2 fs_io_s OK
cluster: xxx.localdomain
filesystem: gpfs2
disks: 2
timestamp: 1121974058/321741
bytes read: 0
bytes written: 0
opens: 0
closes: 0
reads: 0
writes: 0
readdir: 0
inode updates: 0
```

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

## Failure on a node accessed by mmpmon:

This is an example of the system output for a failed request to two nodes to display the I/O statistics per mounted file system.

In this example, the same scenario described in “A successful fs\_io\_s request propagated to two nodes” on page 12 is run on **node2**, but with a failure on **node1** (a non-quorum node) because **node1** was shutdown:

```
_failed_n_199.18.1.5_nn_node2_fn_199.18.1.2_fnn_node1_rc_233
_t_1121974459_tu_602231
_fs_io_s_n_199.18.1.5_nn_node2_rc_0_t_1121974459_tu_616867_cl_
c11.xxx.com_fs_fs2_d_2_br_0_bw_0_oc_0_cc_0_rdc_0_wc_0_dir_0
_iu_0
_fs_io_s_n_199.18.1.5_nn_node2_rc_0_t_1121974459_tu_616867_cl_
c11.xxx.com_fs_fs3_d_3_br_0_bw_0_oc_0_cc_0_rdc_0_wc_0_dir_0
_iu_0
_fs_io_s_n_199.18.1.5_nn_node2_rc_0_t_1121974459_tu_616867_cl_
node1.localdomain_fs_gpfs2_d_2_br_0_bw_0_oc_0_cc_0_rdc_0_wc_0
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.5 name node2:
from node 199.18.1.2 from name node1: failure status 233 timestamp 1121974459/602231
node failed (or never started)
mmpmon node 199.18.1.5 name node2 fs_io_s OK
cluster: c11.xxx.com
filesystem: fs2
disks: 2
timestamp: 1121974544/222514
bytes read: 0
bytes written: 0
opens: 0
closes: 0
reads: 0
writes: 0
readdir: 0
inode updates: 0

mmpmon node 199.18.1.5 name node2 fs_io_s OK
cluster: c11.xxx.com
filesystem: fs3
disks: 3
timestamp: 1121974544/222514
bytes read: 0
bytes written: 0
opens: 0
closes: 0
reads: 0
writes: 0
readdir: 0
inode updates: 0

mmpmon node 199.18.1.5 name node2 fs_io_s OK
cluster: xxx.localdomain
filesystem: gpfs2
disks: 2
timestamp: 1121974544/222514
bytes read: 0
bytes written: 0
opens: 0
closes: 0
reads: 0
writes: 0
readdir: 0
inode updates: 0
```

**Node shutdown and quorum loss:** In this example, the quorum node (**node2**) is shutdown, causing quorum loss on **node1**. Running the same example on **node2**, the output is similar to:

```
_failed_ _n_ 199.18.1.2 _nn_ node1 _rc_ 668 _t_ 1121974459 _tu_ 616867
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.2 name node1: failure status 668 timestamp 1121974459/616867
lost quorum
```

In this scenario there can be a window where **node2** is down and **node1** has not yet lost quorum. When quorum loss occurs, the **mmpmon** command does not attempt to communicate with any nodes in the node list. The goal with failure handling is to accurately maintain the node list across node failures, so that when nodes come back up they again contribute to the aggregated responses.

### Node list failure values:

Table 13 describes the keywords and values produced by the **mmpmon** command on a node list failure:

Table 13. Keywords and values for the **mmpmon nlist** failures

Keyword	Description
<b>_n_</b>	IP address of the node processing the node list. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_fn_</b>	IP address of the node that is no longer responding to <b>mmpmon</b> requests.
<b>_fnn_</b>	The name by which GPFS knows the node that is no longer responding to <b>mmpmon</b> requests
<b>_rc_</b>	Indicates the status of the operation. See “Return codes from mmpmon” on page 41.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.

## Reset statistics to zero

The **reset** request resets the statistics that are displayed with **fs\_io\_s** and **io\_s** requests. The **reset** request *does not* reset the histogram data, which is controlled and displayed with **rhist** requests.

Table 14 describes the keywords for the **reset** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag. The response is a single string.

Table 14. Keywords and values for the **mmpmon reset** response

Keyword	Description
<b>_n_</b>	IP address of the node responding. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.

### Example of mmpmon reset request

This topic is an example of how to reset file system I/O and I/O statistics to zero.

Assume that **commandFile** contains this line:

```
reset
```

and this command is issued:

```
mmpmon -p -i commandFile
```

The output is similar to this:

```
_reset_ _n_ 199.18.1.8 _nn_ node1 _rc_ 0 _t_ 1066660148 _tu_ 407431
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node1 reset OK
```

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

## Understanding the request histogram facility

Use the **mmpmon rhist** requests to control the request histogram facility.

The request histogram facility tallies I/O operations using a set of counters. Counters for reads and writes are kept separately. They are categorized according to a pattern that may be customized by the user. A default pattern is also provided. The **size range** and **latency range** input parameters to the **rhist nr** request are used to define the pattern.

The first time that you run the **rhist** requests, assess if there is a noticeable performance degradation. Collecting histogram data may cause performance degradation. This is possible once the histogram facility is enabled, but will probably not be noticed while the commands themselves are running. It is more of a long term issue as the GPFS daemon runs with histograms enabled.

The histogram lock is used to prevent two **rhist** requests from being processed simultaneously. If an **rhist** request fails with an **\_rc\_** of 16, the lock is in use. Reissue the request.

The histogram data survives file system mounts and unmounts. In order to reset this data, use the **rhist reset** request.

Table 15 describes the **rhist** requests:

*Table 15. rhist requests for the mmpmon command*

Request	Description
rhist nr	“Changing the request histogram facility request size and latency ranges” on page 18
rhist off	“Disabling the request histogram facility” on page 20. This is the default.
rhist on	“Enabling the request histogram facility” on page 21
rhist p	“Displaying the request histogram facility pattern” on page 22
rhist reset	“Resetting the request histogram facility data to zero” on page 24
rhist s	“Displaying the request histogram facility statistics values” on page 25

## Specifying the size ranges for I/O histograms

The I/O histogram size ranges are used to categorize the I/O according to the size, in bytes, of the I/O operation.

The size ranges are specified using a string of positive integers separated by semicolons (;). No white space is allowed within the size range operand. Each number represents the upper bound, in bytes, of the I/O request size for that range. The numbers must be monotonically increasing. Each number may be optionally followed by the letters K or k to denote multiplication by 1024, or by the letters M or m to denote multiplication by 1048576 (1024\*1024).

For example, the size range operand:

```
512;1m;4m
```



represents these four size ranges

```
0          to      512 bytes
513        to 1048576 bytes
1048577    to 4194304 bytes
4194305    and greater bytes
```

In this example, a read of size 3 MB would fall in the third size range, a write of size 20 MB would fall in the fourth size range.

A size range operand of = (equal sign) indicates that the current size range is not to be changed. A size range operand of \* (asterisk) indicates that the current size range is to be changed to the default size range. A maximum of 15 numbers may be specified, which produces 16 total size ranges.

The default request size ranges are:

```
0          to      255 bytes
256        to      511 bytes
512        to     1023 bytes
1024       to     2047 bytes
2048       to     4095 bytes
4096       to     8191 bytes
8192       to    16383 bytes
16384      to    32767 bytes
32768      to    65535 bytes
65536      to   131071 bytes
131072     to   262143 bytes
262144     to   524287 bytes
524288     to  1048575 bytes
1048576    to  2097151 bytes
2097152    to  4194303 bytes
4194304    and greater bytes
```

The last size range collects all request sizes greater than or equal to 4 MB. The request size ranges can be changed by using the **rhist nr** request.

For more information, see “Processing of rhist nr” on page 18.

## Specifying the latency ranges for I/O

The I/O histogram latency ranges are used to categorize the I/O according to the latency time, in milliseconds, of the I/O operation.

A full set of latency ranges are produced for each size range. The latency ranges are the same for each size range.

The latency ranges are changed using a string of positive decimal numbers separated by semicolons (;). No white space is allowed within the latency range operand. Each number represents the upper bound of the I/O latency time (in milliseconds) for that range. The numbers must be monotonically increasing. If decimal places are present, they are truncated to tenths.

For example, the latency range operand:

```
1.3;4.59;10
```

represents these four latency ranges:

```
0.0 to 1.3 milliseconds
1.4 to 4.5 milliseconds
4.6 to 10.0 milliseconds
10.1 and greater milliseconds
```

In this example, a read that completes in 0.85 milliseconds falls into the first latency range. A write that completes in 4.56 milliseconds falls into the second latency range, due to the truncation.

A latency range operand of = (equal sign) indicates that the current latency range is not to be changed. A latency range operand of \* (asterisk) indicates that the current latency range is to be changed to the default latency range. If the latency range operand is missing, \* (asterisk) is assumed. A maximum of 15 numbers may be specified, which produces 16 total latency ranges.

The latency times are in milliseconds. The default latency ranges are:

```
0.0    to 1.0  milliseconds
1.1    to 10.0 milliseconds
10.1   to 30.0 milliseconds
30.1   to 100.0 milliseconds
100.1  to 200.0 milliseconds
200.1  to 400.0 milliseconds
400.1  to 800.0 milliseconds
800.1  to 1000.0 milliseconds
1000.1 and greater milliseconds
```

The last latency range collects all latencies greater than or equal to 1000.1 milliseconds. The latency ranges can be changed by using the **rhist nr** request.

For more information, see “Processing of rhist nr.”

### Changing the request histogram facility request size and latency ranges

The **rhist nr** (new range) request allows the user to change the size and latency ranges used in the request histogram facility.

The use of **rhist nr** implies an **rhist reset**. Counters for read and write operations are recorded separately. If there are no mounted file systems at the time **rhist nr** is issued, the request still runs. The size range operand appears first, followed by a blank, and then the latency range operand.

Table 16 describes the keywords for the **rhist nr** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag.

Table 16. Keywords and values for the **mmpmon rhist nr** response

Keyword	Description
<b>_n_</b>	IP address of the node responding. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_req_</b>	The action requested. In this case, the value is nr.
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.

An **\_rc\_** value of 16 indicates that the histogram operations lock is busy. Retry the request.

### Processing of rhist nr:

The **rhist nr** request changes the request histogram facility request size and latency ranges.

Processing of **rhist nr** is as follows:

1. The size range and latency range operands are parsed and checked for validity. If they are not valid, an error is returned and processing terminates.
2. The histogram facility is disabled.
3. The new ranges are created, by defining the following histogram counters:
  - a. Two sets, one for read and one for write.

- b. Within each set, one category for each size range.
- c. Within each size range category, one counter for each latency range.

For example, if the user specifies 11 numbers for the size range operand and 2 numbers for the latency range operand, this produces 12 size ranges, each having 3 latency ranges, because there is one additional range for the top endpoint. The total number of counters is 72: 36 read counters and 36 write counters.

4. The new ranges are made current.
5. The old ranges are discarded. Any accumulated histogram data is lost.

The histogram facility must be explicitly enabled again using **rhist on** to begin collecting histogram data using the new ranges.

The **mmpmon** command does not have the ability to collect data only for read operations, or only for write operations. The **mmpmon** command does not have the ability to specify size or latency ranges that have different values for read and write operations. The **mmpmon** command does not have the ability to specify latency ranges that are unique to a given size range.

For more information, see “Specifying the size ranges for I/O histograms” on page 16 and “Specifying the latency ranges for I/O” on page 17.

#### Example of **mmpmon rhist nr** request:

This topic is an example of using **rhist nr** to change the request histogram facility request size and latency changes.

Assume that **commandFile** contains this line:

```
rhist nr 512;1m;4m 1.3;4.5;10
```

and this command is issued:

```
mmpmon -p -i commandFile
```

The output is similar to this:

```
_rhist_ _n_ 199.18.2.5 _nn_ node1 _req_ nr 512;1m;4m 1.3;4.5;10 _rc_ 0 _t_ 1078929833 _tu_ 765083
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node1 rhist nr 512;1m;4m 1.3;4.5;10 OK
```

In this case, **mmpmon** has been instructed to keep a total of 32 counters. There are 16 for read and 16 for write. For the reads, there are four size ranges, each of which has four latency ranges. The same is true for the writes. They are as follows:

```
size range      0    to    512    bytes
  latency range  0.0  to    1.3    milliseconds
  latency range  1.4  to    4.5    milliseconds
  latency range  4.6  to   10.0    milliseconds
  latency range  10.1 and greater milliseconds
size range     513  to 1048576 bytes
  latency range  0.0  to    1.3    milliseconds
  latency range  1.4  to    4.5    milliseconds
  latency range  4.6  to   10.0    milliseconds
  latency range  10.1 and greater milliseconds
size range    1048577 to 4194304 bytes
  latency range  0.0  to    1.3    milliseconds
  latency range  1.4  to    4.5    milliseconds
  latency range  4.6  to   10.0    milliseconds
  latency range  10.1 and greater milliseconds
size range    4194305 and greater bytes
```

```

latency range    0.0 to 1.3 milliseconds
latency range    1.4 to 4.5 milliseconds
latency range    4.6 to 10.0 milliseconds
latency range    10.1 and greater milliseconds

```

In this example, a read of size 15 MB that completes in 17.8 milliseconds would fall in the last latency range listed here. When this read completes, the counter for the last latency range will be increased by one.

An `_rc_` value of 16 indicates that the histogram operations lock is busy. Retry the request.

An example of an unsuccessful response is:

```
_rhist_ _n_ 199.18.2.5 _nn_ node1 _req_ nr 512;1m;4m 1;4;8;2 _rc_ 22 _t_ 1078929596 _tu_ 161683
```

If the `-p` flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node1 rhist nr 512;1m;4m 1;4;8;2 status 22 range error
```

In this case, the last value in the latency range, 2, is out of numerical order.

Note that the request `rhist nr = =` does not make any changes. It is ignored.

For information on interpreting `mmpmon` output results, see “Other information about `mmpmon` output” on page 40.

## Disabling the request histogram facility

The `rhist off` request disables the request histogram facility. This is the default value.

The data objects remain persistent, and the data they contain is not disturbed. This data is not updated again until `rhist on` is issued. `rhist off` may be combined with `rhist on` as often as desired. If there are no mounted file systems at the time `rhist off` is issued, the facility is still disabled. The response is a single string.

Table 17 describes the keywords for the `rhist off` response, in the order that they appear in the output. These keywords are used only when `mmpmon` is invoked with the `-p` flag.

Table 17. Keywords and values for the `mmpmon rhist off` response

Keyword	Description
<code>_n_</code>	IP address of the node responding. This is the address by which GPFS knows the node.
<code>_nn_</code>	The hostname that corresponds to the IP address (the <code>_n_</code> value).
<code>_req_</code>	The action requested. In this case, the value is <code>off</code> .
<code>_rc_</code>	Indicates the status of the operation.
<code>_t_</code>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<code>_tu_</code>	Microseconds part of the current time of day.

An `_rc_` value of 16 indicates that the histogram operations lock is busy. Retry the request.

### Example of `mmpmon rhist off` request:

This topic is an example of the `rhist off` request to disable the histogram facility and the output that displays.

Assume that `commandFile` contains this line:

```
rhist off
```

and this command is issued:

```
mmpmon -p -i commandFile
```

The output is similar to this:

```
_rhist_ _n_ 199.18.1.8 _nn_ node1 _req_ off _rc_ 0 _t_ 1066938820 _tu_ 5755
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node1 rhist off OK
```

An **\_rc\_** value of 16 indicates that the histogram operations lock is busy. Retry the request.

```
mmpmon node 199.18.1.8 name node1 rhist off status 16  
lock is busy
```

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

## Enabling the request histogram facility

The **rhist on** request enables the request histogram facility.

When **rhist on** is invoked the first time, this request creates the necessary data objects to support histogram data gathering. This request may be combined with **rhist off** (or another **rhist on**) as often as desired. If there are no mounted file systems at the time **rhist on** is issued, the facility is still enabled. The response is a single string.

Table 18 describes the keywords for the **rhist on** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag.

Table 18. Keywords and values for the **mmpmon rhist on** response

Keyword	Description
<b>_n_</b>	IP address of the node responding. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_req_</b>	The action requested. In this case, the value is on.
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.

An **\_rc\_** value of 16 indicates that the histogram operations lock is busy. Retry the request.

### Example of mmpmon rhist on request:

This topic is an example of the **rhist on** request to enable the request histogram facility and the output that displays.

Assume that **commandFile** contains this line:

```
rhist on
```

and this command is issued:

```
mmpmon -p -i commandFile
```

The output is similar to this:

```
_rhist_ _n_ 199.18.1.8 _nn_ node1 _req_ on _rc_ 0 _t_ 1066936484 _tu_ 179346
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node1 rhist on OK
```

An **\_rc\_** value of 16 indicates that the histogram operations lock is busy. Retry the request.

```
mmpmon node 199.18.1.8 name node1 rhist on status 16  
lock is busy
```

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

## Displaying the request histogram facility pattern

The **rhist p** request displays the request histogram facility pattern.

The **rhist p** request returns the entire enumeration of the request size and latency ranges. The facility must be enabled for a pattern to be returned. If there are no mounted file systems at the time this request is issued, the request still runs and returns data. The pattern is displayed for both read and write.

Table 19 describes the keywords for the **rhist p** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag.

Table 19. Keywords and values for the **mmpmon rhist p** response

Keyword	Description
<b>_n_</b>	IP address of the node responding. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_req_</b>	The action requested. In this case, the value is <b>p</b> .
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.
<b>_k_</b>	The kind, <b>r</b> or <b>w</b> , (read or write) depending on what the statistics are for.
<b>_R_</b>	Request size range, minimum and maximum number of bytes.
<b>_L_</b>	Latency range, minimum and maximum, in milliseconds.

The request size ranges are in bytes. The zero value used for the upper limit of the last size range means 'and above'. The request size ranges can be changed by using the **rhist nr** request.

The latency times are in milliseconds. The zero value used for the upper limit of the last latency range means 'and above'. The latency ranges can be changed by using the **rhist nr** request.

The **rhist p** request allows an application to query for the entire latency pattern. The application can then configure itself accordingly. Since latency statistics are reported only for ranges with nonzero counts, the statistics responses may be sparse. By querying for the pattern, an application can be certain to learn the complete histogram set. The user may have changed the pattern using the **rhist nr** request. For this reason, an application should query for the pattern and analyze it before requesting statistics.

If the facility has never been enabled, the **\_rc\_** field will be nonzero. An **\_rc\_** value of 16 indicates that the histogram operations lock is busy. Retry the request.

If the facility has been previously enabled, the **rhist p** request will still display the pattern even if **rhist off** is currently in effect.

If there are no mounted file systems at the time **rhist p** is issued, the pattern is still displayed.

### Example of mmpmon rhist p request:

This topic is an example of the **rhist p** request to display the request histogram facility pattern and the output that displays.

Assume that **commandFile** contains this line:

```
rhist p
```

and this command is issued:

```
mmpmon -p -i commandFile
```

The response contains all the latency ranges inside each of the request ranges. The data are separate for read and write:

```
_rhist_ _n_ 199.18.1.8 _nn_ node1 _req_ p _rc_ 0 _t_ 1066939007 _tu_ 386241 _k_ r
... data for reads ...
_rhist_ _n_ 199.18.1.8 _nn_ node1 _req_ p _rc_ 0 _t_ 1066939007 _tu_ 386241 _k_ w
... data for writes ...
_end_
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node1 rhist p OK read
... data for reads ...
mmpmon node 199.18.1.8 name node1 rhist p OK write
... data for writes ...
```

Here is an example of data for reads:

```
_rhist_ _n_ 199.18.1.8 _nn_ node1 _req_ p _rc_ 0 _t_ 1066939007 _tu_ 386241 _k_ r
_R_      0          255
_L_      0.0        1.0
_L_      1.1        10.0
_L_      10.1       30.0
_L_      30.1       100.0
_L_      100.1      200.0
_L_      200.1     400.0
_L_      400.1     800.0
_L_      800.1    1000.0
_L_     1000.1      0
_R_      256        511
_L_      0.0        1.0
_L_      1.1        10.0
_L_      10.1       30.0
_L_      30.1       100.0
_L_      100.1      200.0
_L_      200.1     400.0
_L_      400.1     800.0
_L_      800.1    1000.0
_L_     1000.1      0
_R_      512        1023
_L_      0.0        1.0
_L_      1.1        10.0
_L_      10.1       30.0
_L_      30.1       100.0
_L_      100.1      200.0
_L_      200.1     400.0
_L_      400.1     800.0
_L_      800.1    1000.0
_L_     1000.1      0
...
_R_    4194304      0
_L_      0.0        1.0
_L_      1.1        10.0
_L_      10.1       30.0
```

```

_L_      30.1      100.0
_L_      100.1     200.0
_L_      200.1     400.0
_L_      400.1     800.0
_L_      800.1    1000.0
_L_     1000.1      0

```

If the **-p** flag is not specified, the output is similar to:

```

mmpmon node 199.18.1.8 name node1 rhist p OK read
size range      0 to      255
  latency range  0.0 to    1.0
  latency range  1.1 to   10.0
  latency range 10.1 to   30.0
  latency range 30.1 to  100.0
  latency range 100.1 to 200.0
  latency range 200.1 to 400.0
  latency range 400.1 to 800.0
  latency range 800.1 to 1000.0
  latency range 1000.1 to 0
size range     256 to     511
  latency range  0.0 to    1.0
  latency range  1.1 to   10.0
  latency range 10.1 to   30.0
  latency range 30.1 to  100.0
  latency range 100.1 to 200.0
  latency range 200.1 to 400.0
  latency range 400.1 to 800.0
  latency range 800.1 to 1000.0
  latency range 1000.1 to 0
size range     512 to    1023
  latency range  0.0 to    1.0
  latency range  1.1 to   10.0
  latency range 10.1 to   30.0
  latency range 30.1 to  100.0
  latency range 100.1 to 200.0
  latency range 200.1 to 400.0
  latency range 400.1 to 800.0
  latency range 800.1 to 1000.0
  latency range 1000.1 to 0
...
size range  4194304 to      0
  latency range  0.0 to    1.0
  latency range  1.1 to   10.0
  latency range 10.1 to   30.0
  latency range 30.1 to  100.0
  latency range 100.1 to 200.0
  latency range 200.1 to 400.0
  latency range 400.1 to 800.0
  latency range 800.1 to 1000.0
  latency range 1000.1 to 0

```

If the facility has never been enabled, the **\_rc\_** field will be nonzero.

```

_rhist_ _n_ 199.18.1.8 _nn_ node1 _req_ p _rc_ 1 _t_ 1066939007 _tu_ 386241

```

If the **-p** flag is not specified, the output is similar to this:

```

mmpmon node 199.18.1.8 name node1 rhist p status 1
not yet enabled

```

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

## Resetting the request histogram facility data to zero

The **rhist reset** request resets the histogram statistics.



Table 20 describes the keywords for the **rhist reset** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag. The response is a single string.

Table 20. Keywords and values for the **mmpmon rhist reset** response

Keyword	Description
<b>_n_</b>	IP address of the node responding. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_req_</b>	The action requested. In this case, the value is reset.
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.

If the facility has been previously enabled, the reset request will still reset the statistics even if **rhist off** is currently in effect. If there are no mounted file systems at the time **rhist reset** is issued, the statistics are still reset.

An **\_rc\_** value of 16 indicates that the histogram operations lock is busy. Retry the request.

#### Example of mmpmon rhist reset request:

This topic is an example of the **rhist reset** request to reset the histogram facility data to zero and the output that displays.

Assume that **commandFile** contains this line:

```
rhist reset
```

and this command is issued:

```
mmpmon -p -i commandFile
```

The output is similar to this:

```
_rhist_ _n_ 199.18.1.8 _nn_ node1 _req_ reset _rc_ 0 _t_ 1066939007 _tu_ 386241
```

If the **-p** flag is not specified, the output is similar to:

```
_rhist_ _n_ 199.18.1.8 _nn_ node1 _req_ reset _rc_ 0 _t_ 1066939007 _tu_ 386241
```

If the facility has never been enabled, the **\_rc\_** value will be nonzero:

```
_rhist_ _n_ 199.18.1.8 _nn_ node1 _req_ reset _rc_ 1 _t_ 1066939143 _tu_ 148443
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node1 rhist reset status 1
not yet enabled
```

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

## Displaying the request histogram facility statistics values

The **rhist s** request returns the current values for all latency ranges which have a nonzero count.

Table 21 on page 26 describes the keywords for the **rhist s** response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag.

Table 21. Keywords and values for the **mmpmon rhist s** response

Keyword	Description
<b>_n_</b>	IP address of the node responding. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_req_</b>	The action requested. In this case, the value is <b>s</b> .
<b>_rc_</b>	Indicates the status of the operation.
<b>_t_</b>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<b>_tu_</b>	Microseconds part of the current time of day.
<b>_k_</b>	The kind, <b>r</b> or <b>w</b> , (read or write) depending on what the statistics are for.
<b>_R_</b>	Request size range, minimum and maximum number of bytes.
<b>_NR_</b>	Number of requests that fell in this size range.
<b>_L_</b>	Latency range, minimum and maximum, in milliseconds.
<b>_NL_</b>	Number of requests that fell in this latency range. The sum of all <b>_NL_</b> values for a request size range equals the <b>_NR_</b> value for that size range.

If the facility has been previously enabled, the **rhist s** request will still display the statistics even if **rhist off** is currently in effect. This allows turning the histogram statistics on and off between known points and reading them later. If there are no mounted file systems at the time **rhist s** is issued, the statistics are still displayed.

An **\_rc\_** value of 16 indicates that the histogram operations lock is busy. Retry the request.

#### Example of mmpmon rhist s request:

This topic is an example of the **rhist s** request to display the request histogram facility statistics values and the output that displays.

Assume that **commandFile** contains this line:

```
rhist s
```

and this command is issued:

```
mmpmon -p -i commandFile
```

The output is similar to this:

```
_rhist_ _n_ 199.18.2.5 _nn_ node1 _req_ s _rc_ 0 _t_ 1066939007 _tu_ 386241 _k_ r
_R_      65536      131071 _NR_      32640
_L_       0.0        1.0 _NL_      25684
_L_       1.1       10.0 _NL_      4826
_L_      10.1       30.0 _NL_     1666
_L_      30.1      100.0 _NL_      464
_R_     262144     524287 _NR_      8160
_L_       0.0        1.0 _NL_     5218
_L_       1.1       10.0 _NL_      871
_L_      10.1       30.0 _NL_     1863
_L_      30.1      100.0 _NL_      208
_R_     1048576     2097151 _NR_     2040
_L_       1.1       10.0 _NL_      558
_L_      10.1       30.0 _NL_      809
_L_      30.1      100.0 _NL_      673
_rhist_ _n_ 199.18.2.5 _nn_ node1 _req_ s _rc_ 0 _t_ 1066939007 _tu_ 386241 _k_ w
_R_     131072     262143 _NR_     12240
_L_       0.0        1.0 _NL_     10022
_L_       1.1       10.0 _NL_     1227
_L_      10.1       30.0 _NL_      783
```

```

_L_      30.1      100.0 _NL_      208
_R_     262144    524287 _NR_     6120
_L_       0.0       1.0 _NL_     4419
_L_       1.1      10.0 _NL_     791
_L_      10.1     30.0 _NL_     733
_L_      30.1     100.0 _NL_     177
_R_     524288    1048575 _NR_    3060
_L_       0.0       1.0 _NL_    1589
_L_       1.1      10.0 _NL_     581
_L_      10.1     30.0 _NL_     664
_L_      30.1     100.0 _NL_     226
_R_    2097152    4194303 _NR_     762
_L_       1.1       2.0 _NL_     203
_L_      10.1     30.0 _NL_     393
_L_      30.1     100.0 _NL_     166
_end_

```

This small example shows that the reports for read and write may not present the same number of ranges or even the same ranges. Only those ranges with nonzero counters are represented in the response. This is true for both the request size ranges and the latency ranges within each request size range.

If the **-p** flag is not specified, the output is similar to:

```

mmpmon node 199.18.2.5 name node1 rhist s OK timestamp 1066933849/93804 read
size range      65536 to 131071 count 32640
  latency range 0.0 to 1.0 count 25684
  latency range 1.1 to 10.0 count 4826
  latency range 10.1 to 30.0 count 1666
  latency range 30.1 to 100.0 count 464
size range     262144 to 524287 count 8160
  latency range 0.0 to 1.0 count 5218
  latency range 1.1 to 10.0 count 871
  latency range 10.1 to 30.0 count 1863
  latency range 30.1 to 100.0 count 208
size range    1048576 to 2097151 count 2040
  latency range 1.1 to 10.0 count 558
  latency range 10.1 to 30.0 count 809
  latency range 30.1 to 100.0 count 673
mmpmon node 199.18.2.5 name node1 rhist s OK timestamp 1066933849/93968 write
size range      131072 to 262143 count 12240
  latency range 0.0 to 1.0 count 10022
  latency range 1.1 to 10.0 count 1227
  latency range 10.1 to 30.0 count 783
  latency range 30.1 to 100.0 count 208
size range     262144 to 524287 count 6120
  latency range 0.0 to 1.0 count 4419
  latency range 1.1 to 10.0 count 791
  latency range 10.1 to 30.0 count 733
  latency range 30.1 to 100.0 count 177
size range     524288 to 1048575 count 3060
  latency range 0.0 to 1.0 count 1589
  latency range 1.1 to 10.0 count 581
  latency range 10.1 to 30.0 count 664
  latency range 30.1 to 100.0 count 226
size range    2097152 to 4194303 count 762
  latency range 1.1 to 2.0 count 203
  latency range 10.1 to 30.0 count 393
  latency range 30.1 to 100.0 count 166

```

If the facility has never been enabled, the **\_rc\_** value will be nonzero:

```
_rhist_n_ 199.18.1.8 _nn_ node1 _req_ reset _rc_ 1 _t_ 1066939143 _tu_ 148443
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node1 rhist reset status 1
not yet enabled
```

An `_rc_` value of 16 indicates that the histogram operations lock is busy. Retry the request.

For information on interpreting `mmpmon` output results, see “Other information about `mmpmon` output” on page 40.

## Understanding the Remote Procedure Call (RPC) facility

The `mmpmon` requests that start with `rpc_s` display an aggregation of execution time taken by RPCs for a time unit, for example the last 10 seconds. The statistics displayed are the average, minimum, and maximum of RPC execution time over the last 60 seconds, 60 minutes, 24 hours, and 30 days.

Table 22 describes the `rpc_s` requests:

Table 22. `rpc_s` requests for the `mmpmon` command

Request	Description
<code>rpc_s</code>	“Displaying the aggregation of execution time for Remote Procedure Calls (RPCs)”
<code>rpc_s size</code>	“Displaying the Remote Procedure Call (RPC) execution time according to the size of messages” on page 29

The information displayed with `rpc_s` is similar to what is displayed with the `mmdiag --rpc` command.

## Displaying the aggregation of execution time for Remote Procedure Calls (RPCs)

The `rpc_s` request returns the aggregation of execution time for RPCs.

Table 23 describes the keywords for the `rpc_s` response, in the order that they appear in the output.

Table 23. Keywords and values for the `mmpmon rpc_s` response

Keyword	Description
<code>_req_</code>	Indicates the action requested. The action can be either <code>size</code> , <code>node</code> , or <code>message</code> . If no action is requested, the default is the <code>rpc_s</code> action.
<code>_n_</code>	Indicates the IP address of the node responding. This is the address by which GPFS knows the node.
<code>_nn_</code>	Indicates the hostname that corresponds to the IP address (the <code>_n_</code> value).
<code>_rn_</code>	Indicates the IP address of the remote node responding. This is the address by which GPFS knows the node. The statistics displayed are the averages from <code>_nn_</code> to this <code>_rnn_</code> .
<code>_rnn_</code>	Indicates the hostname that corresponds to the remote node IP address (the <code>_rn_</code> value). The statistics displayed are the averages from <code>_nn_</code> to this <code>_rnn_</code> .
<code>_rc_</code>	Indicates the status of the operation.
<code>_t_</code>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<code>_tu_</code>	Indicates the microseconds part of the current time of day.
<code>_rpcObj_</code>	Indicates the beginning of the statistics for <code>_obj_</code> .
<code>_obj_</code>	Indicates the RPC object being displayed.
<code>_nsecs_</code>	Indicates the number of one-second intervals maintained.
<code>_nmins_</code>	Indicates the number of one-minute intervals maintained.
<code>_nhours_</code>	Indicates the number of one-hour intervals maintained.
<code>_ndays_</code>	Indicates the number of one-day intervals maintained.
<code>_stats_</code>	Indicates the beginning of the RPC statistics.

Table 23. Keywords and values for the **mmpmon rpc\_s** response (continued)

Keyword	Description
<b>_tmu_</b>	Indicates the time unit (seconds, minutes, hours, or days).
<b>_av_</b>	Indicates the average value of execution time for <b>_cnt_</b> RPCs during this time unit.
<b>_min_</b>	Indicates the minimum value of execution time for <b>_cnt_</b> RPCs during this time unit.
<b>_max_</b>	Indicates the maximum value of execution time for <b>_cnt_</b> RPCs during this time unit.
<b>_cnt_</b>	Indicates the count of RPCs that occurred during this time unit.

The values allowed for **\_rpcObj\_** are the following:

- **AG\_STAT\_CHANNEL\_WAIT**
- **AG\_STAT\_SEND\_TIME\_TCP**
- **AG\_STAT\_SEND\_TIME\_VERBS**
- **AG\_STAT\_RECEIVE\_TIME\_TCP**
- **AG\_STAT\_RPC\_LATENCY\_TCP**
- **AG\_STAT\_RPC\_LATENCY\_VERBS**
- **AG\_STAT\_RPC\_LATENCY\_MIXED**
- **AG\_STAT\_LAST**

**Example of mmpmon rpc\_s request:**

This topic is an example of the **rpc\_s** request to display the aggregation of execution time for remote procedure calls (RPCs).

Assume that the file **commandFile** contains the following line:

```
rpc_s
```

The following command is issued:

```
mmpmon -p -i commandFile
```

The output is similar to the following example:

```
_response_begin mmpmon rpc_s
_mmpmon::rpc_s_req_node_n_192.168.56.168_nn_node3_rn_192.168.56.167_rnn_node2_rc_0_t_1388417709_tu_641530
_rpcObj_obj_AG_STAT_CHANNEL_WAIT_nsecs_60_nmins_60_nhours_24_ndays_30
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
.....
.....
.....
_rpcObj_obj_AG_STAT_SEND_TIME_TCP_nsecs_60_nmins_60_nhours_24_ndays_30
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
.....
.....
.....
_response_end
```

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

**Displaying the Remote Procedure Call (RPC) execution time according to the size of messages**

The **rpc\_s size** request returns the cached RPC-related size statistics.

Table 24 describes the keywords for the `rpc_s size` response, in the order that they appear in the output.

Table 24. Keywords and values for the `mmpmon rpc_s size` response

Keyword	Description
<code>_req_</code>	Indicates the action requested. In this case, the value is <code>rpc_s size</code> .
<code>_n_</code>	Indicates the IP address of the node responding. This is the address by which GPFS knows the node.
<code>_nn_</code>	Indicates the hostname that corresponds to the IP address (the <code>_n_</code> value).
<code>_rc_</code>	Indicates the status of the operation.
<code>_t_</code>	Indicates the current time of day in seconds (absolute seconds since Epoch (1970)).
<code>_tu_</code>	Indicates the microseconds part of the current time of day.
<code>_rpcSize_</code>	Indicates the beginning of the statistics for this <code>_size_</code> group.
<code>_size_</code>	Indicates the size of the messages for which statistics are collected.
<code>_nsecs_</code>	Indicates the number of one-second intervals maintained.
<code>_nmins_</code>	Indicates the number of one-minute intervals maintained.
<code>_nhours_</code>	Indicates the number of one-hour intervals maintained.
<code>_ndays_</code>	Indicates the number of one-day intervals maintained.
<code>_stats_</code>	Indicates the beginning of the RPC-size statistics.
<code>_tmu_</code>	Indicates the time unit.
<code>_av_</code>	Indicates the average value of execution time for <code>_cnt_</code> RPCs during this time unit.
<code>_min_</code>	Indicates the minimum value of execution time for <code>_cnt_</code> RPCs during this time unit.
<code>_max_</code>	Indicates the maximum value of execution time for <code>_cnt_</code> RPCs during this time unit.
<code>_cnt_</code>	Indicates the count of RPCs that occurred during this time unit.

### Example of `mmpmon rpc_s size` request:

This topic is an example of the `rpc_s size` request to display the RPC execution time according to the size of messages.

Assume that the file `commandFile` contains the following line:

```
rpc_s size
```

The following command is issued:

```
mmpmon -p -i commandFile
```

The output is similar to the following example:

```
_mmpmon::rpc_s_req_size_n_192.168.56.167_nn_node2_rc_0_t_1388417852_tu_572950
_rpcSize_size_64_nsecs_60_nmins_60_nhours_24_ndays_30
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
.....
.....
.....
_rpcSize_size_256_nsecs_60_nmins_60_nhours_24_ndays_30
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
_stats_tmu_sec_av_0.000_min_0.000_max_0.000_cnt_0
```

```

_stats_ _tmu_ sec_ av_ 0.000, _min_ 0.000, _max_ 0.000, _cnt_ 0
_stats_ _tmu_ sec_ av_ 0.000, _min_ 0.000, _max_ 0.000, _cnt_ 0
.....
.....
_stats_ _tmu_ min_ av_ 0.692, _min_ 0.692, _max_ 0.692, _cnt_ 1
_stats_ _tmu_ min_ av_ 0.000, _min_ 0.000, _max_ 0.000, _cnt_ 0
_stats_ _tmu_ min_ av_ 0.000, _min_ 0.000, _max_ 0.000, _cnt_ 0
_stats_ _tmu_ min_ av_ 0.000, _min_ 0.000, _max_ 0.000, _cnt_ 0
_response_ end

```

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

## Displaying mmpmon version

The **ver** request returns a string containing version information.

Table 25 Describes the keywords for the **ver** (version) response, in the order that they appear in the output. These keywords are used only when **mmpmon** is invoked with the **-p** flag.

Table 25. Keywords and values for the **mmpmon ver** response

Keyword	Description
<b>_n_</b>	IP address of the node responding. This is the address by which GPFS knows the node.
<b>_nn_</b>	The hostname that corresponds to the IP address (the <b>_n_</b> value).
<b>_v_</b>	The version of <b>mmpmon</b> .
<b>_lv_</b>	The level of <b>mmpmon</b> .
<b>_vt_</b>	The fix level variant of <b>mmpmon</b> .

## Example of mmpmon ver request

This topic is an example of the **ver** request to display the **mmpmom** version and the output that displays.

Assume that **commandFile** contains this line:

```
ver
```

and this command is issued:

```
mmpmon -p -i commandFile
```

The output is similar to this:

```
_ver_ _n_ 199.18.1.8 _nn_ node1 _v_ 3 _lv_ 3 _vt_ 0
```

If the **-p** flag is not specified, the output is similar to:

```
mmpmon node 199.18.1.8 name node1 version 3.3.0
```

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

## Example mmpmon scenarios and how to analyze and interpret their results

This topic is an illustration of how **mmpmon** is used to analyze I/O data and draw conclusions based on it.

The `fs_io_s` and `io_s` requests are used to determine a number of GPFS I/O parameters and their implication for overall performance. The `rhist` requests are used to produce histogram data about I/O sizes and latency times for I/O requests. The request `source` and prefix directive `once` allow the user of `mmpmon` to more finely tune its operation.

## fs\_io\_s and io\_s output - how to aggregate and analyze the results

The `fs_io_s` and `io_s` requests can be used to determine a number of GPFS I/O parameters and their implication for overall performance.

The output from the `fs_io_s` and `io_s` requests can be used to determine:

1. The I/O service rate of a node, from the application point of view. The `io_s` request presents this as a sum for the entire node, while `fs_io_s` presents the data per file system. A rate can be approximated by taking the `_br_` (bytes read) or `_bw_` (bytes written) values from two successive invocations of `fs_io_s` (or `io_s_`) and dividing by the difference of the sums of the individual `_t_` and `_tu_` values (seconds and microseconds).

This must be done for a number of samples, with a reasonably small time between samples, in order to get a rate which is reasonably accurate. Since we are sampling the information at a given interval, inaccuracy can exist if the I/O load is not smooth over the sampling time.

For example, here is a set of samples taken approximately one second apart, when it was known that continuous I/O activity was occurring:

```

_fs_io_s_n_199.18.1.3_nn_node1_rc_0_t_1095862476_tu_634939_cl_cluster1.xxx.com
_fs_gpfs1m_d_3_br_0_bw_3737124864_oc_4_cc_3_rdc_0_wc_3570_dir_0_iu_5
_fs_io_s_n_199.18.1.3_nn_node1_rc_0_t_1095862477_tu_645988_cl_cluster1.xxx.com
_fs_gpfs1m_d_3_br_0_bw_3869245440_oc_4_cc_3_rdc_0_wc_3696_dir_0_iu_5
_fs_io_s_n_199.18.1.3_nn_node1_rc_0_t_1095862478_tu_647477_cl_cluster1.xxx.com
_fs_gpfs1m_d_3_br_0_bw_4120903680_oc_4_cc_3_rdc_0_wc_3936_dir_0_iu_5
_fs_io_s_n_199.18.1.3_nn_node1_rc_0_t_1095862479_tu_649363_cl_cluster1.xxx.com
_fs_gpfs1m_d_3_br_0_bw_4309647360_oc_4_cc_3_rdc_0_wc_4116_dir_0_iu_5
_fs_io_s_n_199.18.1.3_nn_node1_rc_0_t_1095862480_tu_650795_cl_cluster1.xxx.com
_fs_gpfs1m_d_3_br_0_bw_4542431232_oc_4_cc_3_rdc_0_wc_4338_dir_0_iu_5
_fs_io_s_n_199.18.1.3_nn_node1_rc_0_t_1095862481_tu_652515_cl_cluster1.ibm.com
_fs_gpfs1m_d_3_br_0_bw_4743757824_oc_4_cc_3_rdc_0_wc_4530_dir_0_iu_5
_fs_io_s_n_199.18.1.3_nn_node1_rc_0_t_1095862482_tu_654025_cl_cluster1.xxx.com
_fs_gpfs1m_d_3_br_0_bw_4963958784_oc_4_cc_3_rdc_0_wc_4740_dir_0_iu_5
_fs_io_s_n_199.18.1.3_nn_node1_rc_0_t_1095862483_tu_655782_cl_cluster1.xxx.com
_fs_gpfs1m_d_3_br_0_bw_5177868288_oc_4_cc_3_rdc_0_wc_4944_dir_0_iu_5
_fs_io_s_n_199.18.1.3_nn_node1_rc_0_t_1095862484_tu_657523_cl_cluster1.xxx.com
_fs_gpfs1m_d_3_br_0_bw_5391777792_oc_4_cc_3_rdc_0_wc_5148_dir_0_iu_5
_fs_io_s_n_199.18.1.3_nn_node1_rc_0_t_1095862485_tu_665909_cl_cluster1.xxx.com
_fs_gpfs1m_d_3_br_0_bw_5599395840_oc_4_cc_3_rdc_0_wc_5346_dir_0_iu_5

```

This simple `awk` script performs a basic rate calculation:

```

BEGIN {
    count=0;
    prior_t=0;
    prior_tu=0;
    prior_br=0;
    prior_bw=0;
}

{
    count++;

    t = $9;
    tu = $11;
    br = $19;
    bw = $21;

    if(count > 1)

```



```

    {
    delta_t = t-prior_t;
    delta_tu = tu-prior_tu;
    delta_br = br-prior_br;
    delta_bw = bw-prior_bw;
    dt = delta_t + (delta_tu / 1000000.0);
    if(dt > 0) {
        rrate = (delta_br / dt) / 1000000.0;
        wrate = (delta_bw / dt) / 1000000.0;

printf("%5.1f MB/sec read %5.1f MB/sec write\n",rrate,wrate);
    }
    }

    prior_t=t;
    prior_tu=tu;
    prior_br=br;
    prior_bw=bw;
}

```

The calculated service rates for each adjacent pair of samples is:

```

0.0 MB/sec read    130.7 MB/sec write
0.0 MB/sec read    251.3 MB/sec write
0.0 MB/sec read    188.4 MB/sec write
0.0 MB/sec read    232.5 MB/sec write
0.0 MB/sec read    201.0 MB/sec write
0.0 MB/sec read    219.9 MB/sec write
0.0 MB/sec read    213.5 MB/sec write
0.0 MB/sec read    213.5 MB/sec write
0.0 MB/sec read    205.9 MB/sec write

```

Since these are discrete samples, there can be variations in the individual results. For example, there may be other activity on the node or interconnection fabric. I/O size, file system block size, and buffering also affect results. There can be many reasons why adjacent values differ. This must be taken into account when building analysis tools that read **mmpmon** output and interpreting results.

For example, suppose a file is read for the first time and gives results like this.

```

0.0 MB/sec read    0.0 MB/sec write
0.0 MB/sec read    0.0 MB/sec write
92.1 MB/sec read   0.0 MB/sec write
89.0 MB/sec read   0.0 MB/sec write
92.1 MB/sec read   0.0 MB/sec write
90.0 MB/sec read   0.0 MB/sec write
96.3 MB/sec read   0.0 MB/sec write
0.0 MB/sec read    0.0 MB/sec write
0.0 MB/sec read    0.0 MB/sec write

```

If most or all of the file remains in the GPFS cache, the second read may give quite different rates:

```

0.0 MB/sec read    0.0 MB/sec write
0.0 MB/sec read    0.0 MB/sec write
235.5 MB/sec read  0.0 MB/sec write
287.8 MB/sec read  0.0 MB/sec write
0.0 MB/sec read    0.0 MB/sec write
0.0 MB/sec read    0.0 MB/sec write

```

Considerations such as these need to be taken into account when looking at application I/O service rates calculated from sampling **mmpmon** data.

2. Usage patterns, by sampling at set times of the day (perhaps every half hour) and noticing when the largest changes in I/O volume occur. This does not necessarily give a rate (since there are too few samples) but it can be used to detect peak usage periods.
3. If some nodes service significantly more I/O volume than others over a given time span.
4. When a parallel application is split across several nodes, and is the only significant activity in the nodes, how well the I/O activity of the application is distributed.

- The total I/O demand that applications are placing on the cluster. This is done by obtaining results from `fs_io_s` and `io_s` in aggregate for all nodes in a cluster.
- The rate data may appear to be erratic. Consider this example:

```

0.0 MB/sec read    0.0 MB/sec write
6.1 MB/sec read    0.0 MB/sec write
92.1 MB/sec read   0.0 MB/sec write
89.0 MB/sec read   0.0 MB/sec write
12.6 MB/sec read   0.0 MB/sec write
0.0 MB/sec read    0.0 MB/sec write
0.0 MB/sec read    0.0 MB/sec write
8.9 MB/sec read    0.0 MB/sec write
92.1 MB/sec read   0.0 MB/sec write
90.0 MB/sec read   0.0 MB/sec write
96.3 MB/sec read   0.0 MB/sec write
4.8 MB/sec read    0.0 MB/sec write
0.0 MB/sec read    0.0 MB/sec write

```

The low rates which appear before and after each group of higher rates can be due to the I/O requests occurring late (in the leading sampling period) and ending early (in the trailing sampling period.) This gives an apparently low rate for those sampling periods.

The zero rates in the middle of the example could be caused by reasons such as no I/O requests reaching GPFS during that time period (the application issued none, or requests were satisfied by buffered data at a layer above GPFS), the node becoming busy with other work (causing the application to be undispached), or other reasons.

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

### Request histogram (**rhist**) output - how to aggregate and analyze the results

The **rhist** requests are used to produce histogram data about I/O sizes and latency times for I/O requests.

The output from the **rhist** requests can be used to determine:

- The number of I/O requests in a given size range. The sizes may vary based on operating system, explicit application buffering, and other considerations. This information can be used to help determine how well an application or set of applications is buffering its I/O. For example, if there are many very small or many very large I/O transactions. A large number of overly small or overly large I/O requests may not perform as well as an equivalent number of requests whose size is tuned to the file system or operating system parameters.
- The number of I/O requests in a size range that have a given latency time. Many factors can affect the latency time, including but not limited to: system load, interconnection fabric load, file system block size, disk block size, disk hardware characteristics, and the operating system on which the I/O request is issued.

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output” on page 40.

### Using request *source* and prefix directive *once*

The request *source* and prefix directive *once* allow **mmpmon** users to more finely tune their operations.

The **source** request causes **mmpmon** to read requests from a file, and when finished return to reading requests from the input stream.

The prefix directive **once** can be placed in front of any **mmpmon** request. The **once** prefix indicates that the request be run only once, irrespective of the setting of the **-r** flag on the **mmpmon** command. It is useful for requests that do not need to be issued more than once, such as to set up the node list or turn on the request histogram facility.

These rules apply when using the **once** prefix directive and **source** request:

1. **once** with nothing after it is an error that terminates **mmpmon** processing.
2. A file invoked with the **source** request may contain **source** requests, causing file nesting of arbitrary depth. No check is done for loops in this situation.
3. The request **once source filename** causes the **once** prefix to be applied to all the **mmpmon** requests in *filename*, including any **source** requests in the file.
4. If a *filename* specified with the **source** request cannot be opened for read, an error is returned and **mmpmon** terminates.
5. If the **-r** flag on the **mmpmon** command has any value other than one, and all requests are prefixed with **once**, **mmpmon** runs all the requests once, issues a message, and then terminates.

**An example of *once* and *source* usage:**

This topic provides an example of the **once** and **source** requests and the output that displays.

This command is issued:

```
mmpmon -p -i command.file -r 0 -d 5000 | tee output.file
```

File **command.file** consists of this:

```
once source mmpmon.header
once rhist nr 512;1024;2048;4096 =
once rhist on
source mmpmon.commands
```

File **mmpmon.header** consists of this:

```
ver
reset
```

File **mmpmon.commands** consists of this:

```
fs_io_s
rhist s
```

The **output.file** is similar to this:

```
_ver_n 199.18.1.8_nn_node1_v_2_lv_4_vt_0
_reset_n 199.18.1.8_nn_node1_rc_0_t_1129770129_tu_511981
_rhist_n 199.18.1.8_nn_node1_req_nr 512;1024;2048;4096 =_rc_0_t_1129770131_tu_524674
_rhist_n 199.18.1.8_nn_node1_req_on_rc_0_t_1129770131_tu_524921
_fs_io_s_n 199.18.1.8_nn_node1_rc_0_t_1129770131_tu_525062_cl_node1.localdomain
_fs_gpfs1_d_1_br_0_bw_0_oc_0_cc_0_rdc_0_wc_0_dir_0_iu_0
_fs_io_s_n 199.18.1.8_nn_node1_rc_0_t_1129770131_tu_525062_cl_node1.localdomain
_fs_gpfs2_d_2_br_0_bw_0_oc_0_cc_0_rdc_0_wc_0_dir_0_iu_0
_rhist_n 199.18.1.8_nn_node1_req_s_rc_0_t_1129770131_tu_525220_k_r
_rhist_n 199.18.1.8_nn_node1_req_s_rc_0_t_1129770131_tu_525228_k_w
_end
_fs_io_s_n 199.18.1.8_nn_node1_rc_0_t_1129770136_tu_526685_cl_node1.localdomain
_fs_gpfs1_d_1_br_0_bw_0_oc_0_cc_0_rdc_0_wc_0_dir_0_iu_0
_fs_io_s_n 199.18.1.8_nn_node1_rc_0_t_1129770136_tu_526685_cl_node1.localdomain
_fs_gpfs2_d_2_br_0_bw_395018_oc_504_cc_252_rdc_0_wc_251_dir_0_iu_147
_rhist_n 199.18.1.8_nn_node1_req_s_rc_0_t_1129770136_tu_526888_k_r
_rhist_n 199.18.1.8_nn_node1_req_s_rc_0_t_1129770136_tu_526896_k_w
_R_0_512_NR_169
_L_0.0_1.0_NL_155
_L_1.1_10.0_NL_7
_L_10.1_30.0_NL_1
_L_30.1_100.0_NL_4
_L_100.1_200.0_NL_2
_R_513_1024_NR_16
_L_0.0_1.0_NL_15
_L_1.1_10.0_NL_1
```

```

_R 1025 2048 _NR 3
_L 0.0 1.0 _NL 32
_R 2049 4096 _NR 18
_L 0.0 1.0 _NL 18
_R 4097 0 _NR 16
_L 0.0 1.0 _NL 16
_end
_fs_io_s_n 199.18.1.8 _nn_node1_rc_0_t_1129770141_tu_528613_cl_node1.localdomain
_fs_gpfs1_d_1_br_0_bw_0_oc_0_cc_0_rdc_0_wc_0_dir_0_iu_0
_fs_io_s_n 199.18.1.8 _nn_node1_rc_0_t_1129770141_tu_528613_cl_node1.localdomain
_fs_gpfs2_d_2_br_0_bw_823282_oc_952_cc_476_rdc_0_wc_474_dir_0_iu_459
_rhist_n 199.18.1.8 _nn_node1_req_s_rc_0_t_1129770141_tu_528812_k_r
_rhist_n 199.18.1.8 _nn_node1_req_s_rc_0_t_1129770141_tu_528820_k_w
_R 0 512 _NR 255
_L 0.0 1.0 _NL 241
_L 1.1 10.0 _NL 7
_L 10.1 30.0 _NL 1
_L 30.1 100.0 _NL 4
_L 100.1 200.0 _NL 2
_R 513 1024 _NR 36
_L 0.0 1.0 _NL 35
_L 1.1 10.0 _NL 1
_R 1025 2048 _NR 90
_L 0.0 1.0 _NL 90
_R 2049 4096 _NR 55
_L 0.0 1.0 _NL 55
_R 4097 0 _NR 38
_L 0.0 1.0 _NL 37
_L 1.1 10.0 _NL 1
_end
_fs_io_s_n 199.18.1.8 _nn_node1_rc_0_t_1129770146_tu_530570_cl_node1.localdomain
_fs_gpfs1_d_1_br_0_bw_0_oc_0_cc_0_rdc_0_wc_0_dir_0_iu_1
_fs_io_s_n 199.18.1.8 _nn_node1_rc_0_t_1129770146_tu_530570_cl_node1.localdomain
_fs_gpfs2_d_2_br_0_bw_3069915_oc_1830_cc_914_rdc_0_wc_901_dir_0_iu_1070
_rhist_n 199.18.1.8 _nn_node1_req_s_rc_0_t_1129770146_tu_530769_k_r
_rhist_n 199.18.1.8 _nn_node1_req_s_rc_0_t_1129770146_tu_530778_k_w
_R 0 512 _NR 526
_L 0.0 1.0 _NL 501
_L 1.1 10.0 _NL 14
_L 10.1 30.0 _NL 2
_L 30.1 100.0 _NL 6
_L 100.1 200.0 _NL 3
_R 513 1024 _NR 74
_L 0.0 1.0 _NL 70
_L 1.1 10.0 _NL 4
_R 1025 2048 _NR 123
_L 0.0 1.0 _NL 117
_L 1.1 10.0 _NL 6
_R 2049 4096 _NR 91
_L 0.0 1.0 _NL 84
_L 1.1 10.0 _NL 7
_R 4097 0 _NR 87
_L 0.0 1.0 _NL 81
_L 1.1 10.0 _NL 6
_end
..... and so forth .....

```

If this command is issued with the same file contents:

```
mmpmon -i command.file -r 0 -d 5000 | tee output.file.english
```

The file **output.file.english** is similar to this:

```

mmpmon node 199.18.1.8 name node1 version 3.1.0
mmpmon node 199.18.1.8 name node1 reset OK
mmpmon node 199.18.1.8 name node1 rhist nr 512;1024;2048;4096 = OK
mmpmon node 199.18.1.8 name node1 rhist on OK
mmpmon node 199.18.1.8 name node1 fs_io_s OK

```

```

cluster:      node1.localdomain
filesystem:   gpfs1
disks:        1
timestamp:    1129770175/950895
bytes read:   0
bytes written: 0
opens:        0
closes:       0
reads:        0
writes:        0
readdir:      0
inode updates: 0

mmpmon node 199.18.1.8 name node1 fs_io_s OK
cluster:      node1.localdomain
filesystem:   gpfs2
disks:        2
timestamp:    1129770175/950895
bytes read:   0
bytes written:

opens:        0
closes:       0
reads:        0
writes:        0
readdir:      0
inode updates: 0
mmpmon node 199.18.1.8 name node1 rhist s OK read timestamp 1129770175/951117
mmpmon node 199.18.1.8 name node1 rhist s OK write timestamp 1129770175/951125
mmpmon node 199.18.1.8 name node1 fs_io_s OK
cluster:      node1.localdomain
filesystem:   gpfs1
disks:        1
timestamp:    1129770180/952462
bytes read:   0
bytes written: 0
opens:        0
closes:       0
reads:        0
writes:        0
readdir:      0
inode updates: 0

mmpmon node 199.18.1.8 name node1 fs_io_s OK
cluster:      node1.localdomain
filesystem:   gpfs2
disks:        2
timestamp:    1129770180/952462
bytes read:   0
bytes written: 491310
opens:        659
closes:       329
reads:        0
writes:       327
readdir:      0
inode updates: 74
mmpmon node 199.18.1.8 name node1 rhist s OK read timestamp 1129770180/952711
mmpmon node 199.18.1.8 name node1 rhist s OK write timestamp 1129770180/952720
size range    0 to      512 count      214
  latency range 0.0 to    1.0 count      187
  latency range 1.1 to   10.0 count      15
  latency range 10.1 to  30.0 count       6
  latency range 30.1 to 100.0 count       5
  latency range 100.1 to 200.0 count       1
size range    513 to   1024 count       27
  latency range 0.0 to    1.0 count      26
  latency range 100.1 to  200.0 count       1

```

```

size range      1025 to      2048 count      32
  latency range  0.0 to      1.0 count      29
  latency range  1.1 to     10.0 count       1
  latency range  30.1 to   100.0 count       2
size range      2049 to     4096 count      31
  latency range  0.0 to      1.0 count      30
  latency range  30.1 to   100.0 count       1
size range      4097 to      0 count      23
  latency range  0.0 to      1.0 count      23
mmpmon node 199.18.1.8 name node1 fs_io_s OK
cluster:        node1.localdomain
filesystem:     gpfs1
disks:          1
timestamp:      1129770185/954401
bytes read:     0
bytes written:  0
opens:          0
closes:         0
reads:          0
writes:         0
readdir:        0
inode updates:  0

mmpmon node 199.18.1.8 name node1 fs_io_s OK
cluster:        node1.localdomain
filesystem:     gpfs2
disks:          2
timestamp:      1129770185/954401
bytes read:     0
bytes written:  1641935
opens:          1062
closes:         531
reads:          0
writes:         529
readdir:        0
inode updates:  523
mmpmon node 199.18.1.8 name node1 rhist s OK read timestamp 1129770185/954658
mmpmon node 199.18.1.8 name node1 rhist s OK write timestamp 1129770185/954667
size range      0 to      512 count      305
  latency range  0.0 to      1.0 count      270
  latency range  1.1 to     10.0 count      21
  latency range  10.1 to    30.0 count       6
  latency range  30.1 to   100.0 count       6
  latency range  100.1 to  200.0 count       2
size range      513 to     1024 count      39
  latency range  0.0 to      1.0 count      36
  latency range  1.1 to     10.0 count       1
  latency range  30.1 to   100.0 count       1
  latency range  100.1 to  200.0 count       1
size range      1025 to     2048 count      89
  latency range  0.0 to      1.0 count      84
  latency range  1.1 to     10.0 count       2
  latency range  30.1 to   100.0 count       3
size range      2049 to     4096 count      56
  latency range  0.0 to      1.0 count      54
  latency range  1.1 to     10.0 count       1
  latency range  30.1 to   100.0 count       1
size range      4097 to      0 count      40
  latency range  0.0 to      1.0 count      39
  latency range  1.1 to     10.0 count       1
mmpmon node 199.18.1.8 name node1 fs_io_s OK
cluster:        node1.localdomain
filesystem:     gpfs1
disks:          1
timestamp:      1129770190/956480
bytes read:     0
bytes written:  0

```

```

opens:                0
closes:               0
reads:                0
writes:               0
readdir:              0
inode updates:        0

mmpmon node 199.18.1.8 name node1 fs_io_s OK
cluster:              node1.localdomain
filesystem:           gpfs2
disks:                2
timestamp:            1129770190/956480
bytes read:           0
bytes written:        3357414
opens:                1940
closes:               969
reads:                0
writes:               952
readdir:              0
inode updates:        1101
mmpmon node 199.18.1.8 name node1 rhist s OK read timestamp 1129770190/956723
mmpmon node 199.18.1.8 name node1 rhist s OK write timestamp 1129770190/956732
size range            0 to          512 count          539
  latency range      0.0 to          1.0 count          494
  latency range      1.1 to          10.0 count           29
  latency range      10.1 to         30.0 count            6
  latency range      30.1 to        100.0 count            8
  latency range      100.1 to       200.0 count            2
size range            513 to        1024 count           85
  latency range      0.0 to          1.0 count            81
  latency range      1.1 to          10.0 count            2
  latency range      30.1 to        100.0 count            1
  latency range      100.1 to       200.0 count            1
size range            1025 to       2048 count           133
  latency range      0.0 to          1.0 count           124
  latency range      1.1 to          10.0 count            5
  latency range      10.1 to         30.0 count            1
  latency range      30.1 to        100.0 count            3
size range            2049 to       4096 count            99
  latency range      0.0 to          1.0 count            91
  latency range      1.1 to          10.0 count            6
  latency range      10.1 to         30.0 count            1
  latency range      30.1 to        100.0 count            1
size range            4097 to         0 count            95
  latency range      0.0 to          1.0 count            90
  latency range      1.1 to          10.0 count            4
  latency range      10.1 to         30.0 count            1
mmpmon node 199.18.1.8 name node1 fs_io_s OK
cluster:              node1.localdomain
filesystem:           gpfs1
disks:                1
timestamp:            1129770195/958310
bytes read:           0
bytes written:        0
opens:                0
closes:               0
reads:                0
writes:               0
readdir:              0
inode updates:        0

mmpmon node 199.18.1.8 name node1 fs_io_s OK
cluster:              node1.localdomain
filesystem:           gpfs2
disks:                2
timestamp:            1129770195/958310
bytes read:           0

```

```

bytes written:      3428107
opens:             2046
closes:           1023
reads:             0
writes:           997
readdir:          0
inode updates:    1321
mmpmon node 199.18.1.8 name node1 rhist s OK read timestamp 1129770195/958568
mmpmon node 199.18.1.8 name node1 rhist s OK write timestamp 1129770195/958577
size range         0 to      512 count      555
  latency range    0.0 to    1.0 count      509
  latency range    1.1 to   10.0 count       30
  latency range   10.1 to   30.0 count        6
  latency range   30.1 to  100.0 count        8
  latency range  100.1 to  200.0 count        2
size range         513 to   1024 count       96
  latency range    0.0 to    1.0 count       92
  latency range    1.1 to   10.0 count        2
  latency range   30.1 to  100.0 count        1
  latency range  100.1 to  200.0 count        1
size range        1025 to   2048 count      143
  latency range    0.0 to    1.0 count     134
  latency range    1.1 to   10.0 count        5
  latency range   10.1 to   30.0 count        1
  latency range   30.1 to  100.0 count        3
size range        2049 to   4096 count     103
  latency range    0.0 to    1.0 count      95
  latency range    1.1 to   10.0 count        6
  latency range   10.1 to   30.0 count        1
  latency range   30.1 to  100.0 count        1
size range        4097 to      0 count     100
  latency range    0.0 to    1.0 count      95
latency range     1.1 to   10.0 count        4
latency range    10.1 to   30.0 count        1
..... and so forth .....

```

For information on interpreting **mmpmon** output results, see “Other information about mmpmon output.”

## Other information about mmpmon output

When interpreting the results from the **mmpmon** output there are several points to consider.

Consider these important points:

- On a node acting as a server of a GPFS file system to NFS clients, NFS I/O is accounted for in the statistics. However, the I/O is that which goes between GPFS and NFS. If NFS caches data, in order to achieve better performance, this activity is not recorded.
- I/O requests made at the application level may not be exactly what is reflected to GPFS. This is dependent on the operating system, and other factors. For example, an application read of 100 bytes may result in obtaining, and caching, a 1 MB block of data at a code level above GPFS (such as the libc I/O layer.) . Subsequent reads within this block result in no additional requests to GPFS.
- The counters kept by **mmpmon** are not atomic and may not be exact in cases of high parallelism or heavy system load. This design minimizes the performance impact associated with gathering statistical data.
- Reads from data cached by GPFS will be reflected in statistics and histogram data. Reads and writes to data cached in software layers above GPFS will be reflected in statistics and histogram data when those layers actually call GPFS for I/O.
- Activity from snapshots affects statistics. I/O activity necessary to maintain a snapshot is counted in the file system statistics.
- Some (generally minor) amount of activity in the root directory of a file system is reflected in the statistics of the file system manager node, and not the node which is running the activity.



- The open count also includes **creat()** call counts.

## Counter sizes and counter wrapping

The **mmpmon** command may be run continuously for extended periods of time. The user must be aware that counters may wrap.

This information applies to the counters involved:

- The statistical counters used for the **io\_s** and **fs\_io\_s** requests are maintained by GPFS at all times, even when **mmpmon** has not been invoked. It is suggested that you use the **reset** request prior to starting a sequence of **io\_s** or **fs\_io\_s** requests.
- The bytes read and bytes written counters are unsigned 64-bit integers. They are used in the **fs\_io\_s** and **io\_s** requests, as the **\_br\_** and **\_bw\_** fields.
- The counters associated with the **rhist** requests are updated only when the request histogram facility has been enabled.
- The counters used in the **rhist** requests are unsigned 64-bit integers.
- All other counters are unsigned 32-bit integers.

For more information, see “fs\_io\_s and io\_s output - how to aggregate and analyze the results” on page 32 and “Request histogram (rhist) output - how to aggregate and analyze the results” on page 34.

## Return codes from mmpmon

This topic provides the **mmpmom** return codes and explanations for the codes.

These are the return codes that can appear in the **\_rc\_** field:

- 0 Successful completion.
- 1 One of these has occurred:
  1. For the **fs\_io\_s** request, no file systems are mounted.
  2. For an **rhist** request, a request was issued that requires the request histogram facility to be enabled, but it is not. The facility is not enabled if:
    - Since the last **mmstartup** was issued, **rhist on** was never issued.
    - **rhist nr** was issued and **rhist on** was not issued afterwards.
- 2 For one of the **nlist** requests, the node name is not recognized.
- 13 For one of the **nlist** requests, the node name is a remote node, which is not allowed.
- 16 For one of the **rhist** requests, the histogram operations lock is busy. Retry the request.
- 17 For one of the **nlist** requests, the node name is already in the node list.
- 22 For one of the **rhist** requests, the size or latency range parameters were not in ascending order or were otherwise incorrect.
- 233 For one of the **nlist** requests, the specified node is not joined to the cluster.
- 668 For one of the **nlist** requests, quorum has been lost in the cluster.

---

## Performance monitoring tool overview

The performance monitoring tool collects metrics from GPFS and protocols and provides performance information.

The performance monitoring system is started by default and consists of three parts: Collectors, Sensors, and Proxies.

## Collector

In the previous release of IBM Spectrum Scale, the performance monitoring tool could be configured with a single collector only. From version 4.2, the performance monitoring tool can be configured with multiple collectors to increase scalability and fault-tolerance. This latter configuration is referred to as federation.

In a multi-collector federated configuration, the collectors need to be aware of each other, otherwise a collector would only return the data stored in its own measurement database. Once the collectors are aware of their peer collectors, they can collaborate with each other to collate measurement data for a given measurement query. All collectors that are part of the federation are specified in the `peers` configuration option in the collector's configuration file as shown in the following example:

```
peers = { host = "collector1.mydomain.com" port = "9085" },
        { host = "collector2.mydomain.com" port = "9085" }
```

The port number is the one specified by the `federationport` configuration option, typically set to 9085. You can also list the current host so that the same configuration file can be used for all the collector machines.

Once the peers have been specified, any query for measurement data might be directed to any of the collectors listed in the `peers` section and the collector collects and assembles a response based on all relevant data from all collectors. Hence, clients need to only contact a single collector instead of all of them in order to get all the measurements available in the system.

To distribute the measurement data reported by sensors over multiple collectors, multiple collectors might be specified when configuring the sensors.

If multiple collectors are specified, the sensors pick one to report their measurement data to. The sensors use stable hashes to pick the collector such that the sensor-collector relationship does not change too much if new collectors are added or if a collector is removed.

Additionally, sensors and collectors can be configured for high availability. In this setting, sensors report their measurement data to more than one collector such that the failure of a single collector would not lead to any data loss. For instance, if the collector redundancy is increased to two, every sensor reports to two collectors. As a side-effect of increasing the redundancy to two, the bandwidth consumed for reporting measurement data is duplicated. The collector redundancy has to be configured before the sensor configuration is stored in IBM Spectrum Scale by changing the `colRedundancy` option in `/opt/IBM/zimon/ZIMonSensors.cfg`.

## Sensor

A sensor is a component that collects performance data from a node. Typically there are multiple sensors run on any node that is required to collect metrics. By default, the sensors are started on every node.

Sensors identify the collector from the information present in the configuration file, `/opt/IBM/zimon/ZIMonSensors.cfg`.

## Proxy

A proxy is run for each of the protocols to collect the metrics for that protocol.

By default, the NFS and SMB proxies are started automatically with those protocols. They do not need to be started or stopped. However, to retrieve metrics for SMB, NFS or Object, these protocols have to be active on the specific node.

For information on enabling Object metrics, see the "Enabling protocol metrics" on page 62 topic.

For information on enabling Transparent Cloud Tiering metrics, see *Integrating Transparent Cloud Tiering metrics with performance monitoring tool* in *IBM Spectrum Scale: Administration Guide*.

## Configuring the performance monitoring tool

The performance monitoring tool, collector, sensors, and proxies, are a part of the IBM Spectrum Scale distribution. The tool is installed with the GPFS core packages on all nodes. The tools packages are small, approximately 300 KB for the sensors and 500 KB for the collector.

**Note:** The tool is supported on Linux nodes only.

For information on the usage of ports for the performance monitoring tool, see the *Firewall recommendations for Performance Monitoring tool* in *IBM Spectrum Scale: Administration Guide*.

## Configuring the sensor

Performance monitoring sensors can either be managed manually as individual files on each node or managed automatically by IBM Spectrum Scale.

### Identifying the type of configuration in use:

If the performance monitoring infrastructure was installed previously, you might need to identify the type of configuration the system is currently using.

If the sensor configuration is managed automatically, the configuration is stored within IBM Spectrum Scale. If it is managed automatically, it can be viewed with the `mmperfmon config show` command. The set of nodes where this configuration is enabled can be identified through the `mmiscluster` command. Those nodes where performance monitoring metrics collection is enabled are marked with the `perfmon` designation as shown in the following sample:

```
prompt# mmiscluster
GPFS cluster information
=====
GPFS cluster name:      s1.zimon.zc2.ibm.com
GPFS cluster id:       13860500485217864948
GPFS UID domain:       s1.zimon.zc2.ibm.com
Remote shell command:  /usr/bin/ssh
Remote file copy command: /usr/bin/scp
Repository type:       CCR
Node Daemon node name  IP address      Admin node name  Designation
-----
1      s1.zimon.zc2.ibm.com  9.4.134.196     s1.zimon.zc2.ibm.com  quorum-perfmon
2      s2.zimon.zc2.ibm.com  9.4.134.197     s2.zimon.zc2.ibm.com  quorum-perfmon
3      s3.zimon.zc2.ibm.com  9.4.134.198     s3.zimon.zc2.ibm.com  quorum-perfmon
4      s4.zimon.zc2.ibm.com  9.4.134.199     s4.zimon.zc2.ibm.com  quorum-perfmon
5      s5.zimon.zc2.ibm.com  9.4.134.2       s5.zimon.zc2.ibm.com  quorum-perfmon
```

If `mmperfmon config show` does not show any configuration and no nodes are designated `perfmon`, the configuration can be managed manually.

### Automated configuration:

Starting with version 4.2 of the performance monitoring tool, sensors can be configured on nodes that are part of an IBM Spectrum Scale cluster through an IBM Spectrum Scale based configuration mechanism. However, this requires the installation of IBM Spectrum Scale 4.2 or later versions on all the nodes where a sensor is running and where the sensors are to be configured. It also requires the entire cluster to be at least running IBM Spectrum Scale 4.1.1 or later version, and the execution of the `mmchconfig release=LATEST` command.

The automated configuration method allows the sensor configuration to be stored as part of the IBM Spectrum Scale configuration. Automated configuration is only available for the sensor configuration files (/opt/IBM/zimon/ZIMonSensors.cfg) but not for the collector configuration files (/opt/IBM/zimon/ZIMonCollector.cfg). In this setup, the /opt/IBM/zimon/ZIMonSensors.cfg configuration file on each IBM Spectrum Scale node is maintained by IBM Spectrum Scale. As a result, the file must not be edited manually because whenever IBM Spectrum Scale needs to update a configuration parameter, the file is regenerated and any manual modifications are overwritten. Before using the automated configuration, an initial configuration needs to be stored within IBM Spectrum Scale. You can store this initial configuration by using the **mmperfmon config generate** command as shown:

```
prompt# mmperfmon config generate \  
--collectors collector1.domain.com,collector2.domain.com,...
```

The **mmperfmon config generate** command will use a template configuration file for generating the automated configuration. The default location for that template configuration is /opt/IBM/zimon/defaults/ZIMonSensors.cfg.

The template configuration includes the initial settings for all the sensors and may be modified prior to invoking the **mmperfmon config generate** command. This file also includes a parameter called **colCandidates**. This parameter specifies the number of collectors that each sensor should report its data to. This may be of interest for high-availability setups, where each metric should be sent to two collectors in case one collector becomes unavailable.

Once the configuration file is stored within IBM Spectrum Scale, it can be activated as follows:

```
prompt# mmchnode --perfmon -N nodeclass1,nodeclass2,...
```

**Note:** Any previously existing configuration file is overwritten. Configuration changes result in a new version of the configuration file, which is then propagated through the IBM Spectrum Scale cluster at the file level.

To deactivate the performance monitoring tool, the same command is used but with the **--noperfmon** switch supplied instead. Configuration parameters can be changed with the following command where **param1** is of the form **sensorname.sensorattribute**:

```
prompt# mmperfmon config update param1=value1 param2=value2 ...
```

Some sensors such as the cluster export services sensors are run on a specific set of nodes. Other sensors such as the GPFSDiskCap sensor are run on a single node in the cluster since the data reported is the same, independent of the node the sensor is running on. For these types of sensors, the **restrict** function is especially intended. For example, to restrict a **NFSIO** sensor to a node class and change the reporting period to once every 10 hours, you can specify **NFSIO.period=36000 NFSIO.restrict=nodeclass1** as attribute value pairs in the update command.

Some sensors, such as VFS, are not enabled by default even though they have associated predefined queries with the **mmperfmon query** command. This is so because the collector might display performance issues of its own if it is required to collect more than 1000000 metrics per second. To enable VFS sensors, use the **mmfsadm vfsstats enable** command on the node. To enable a sensor, set the period value to an integer greater than 0 and restart the sensors on that node by using the **systemctl restart pmsensors** command.

## Removing an automated configuration

When upgrading the performance monitoring tool, it is important to note how the previous version was configured and if the configuration mechanism is to be changed. Before IBM Spectrum Scale 4.2, the system was configured using a file-based configuration where the configuration files were manually edited and propagated to the requisite nodes. If the configuration mechanism is to be changed, it is important to verify that the installed versions of both IBM Spectrum Scale and the performance

monitoring tool support the new configuration method. However, if you want to use the manual configuration method, then take care of the following:

1. None of the nodes in the cluster should be designated **perfmon** nodes. If the nodes in the cluster are designated as **perfmon** nodes then run `mmchnode --noperfmon -N all` command.
2. Delete the centrally stored configuration information by issuing `mmperfmon config delete --all` command.

The `/opt/IBM/zimon/ZIMonSensors.cfg` file is then maintained manually. This mode is useful if sensors are to be installed on non-Spectrum Scale nodes or if you want to have a cluster with multiple levels of IBM Spectrum Scale running.

### Manual configuration:

Performance monitoring tools can also be configured manually by the user.

**Important:** If you are using IBM Spectrum Scale 4.1.1 or later version, the performance monitoring tool gets automatically configured. This will automatically override any manual changes you try to make to the configuration. If you wish to change an automated configuration to a manual one, follow the steps given in *Removing an automated configuration* in the *Automated configuration* section in the *IBM Spectrum Scale: Administration Guide*.

When configuring the performance monitoring tool manually, the installation toolkit sets up a default set of sensors to monitor on each node. You can modify the sensors on each individual node.

The configuration file of the sensors, `ZimonSensors.cfg`, is located on each node in the `/opt/IBM/zimon` folder. The file lists all groups of sensors in it. The configuration file includes the parameter setting of the sensors, such as the reporting frequency, and controls the sensors that are active within the cluster. The file also contains the host name of the node where the collector is running that the sensor must be reporting to.

For example:

```
sensors =
{
    name = "CPU"
    period = 1
},
{
    name = "Load"
    period = 1
},
{
    name = "Memory"
    period = 1
},
{
    name = "Network"
    period = 1
    filter = "eth*"
    # filters are currently ignored.
},
{
    name = "Netstat"
    period = 1
},
```

The period in the example specifies the interval size in number of seconds when a sensor group gathers data. 0 means that the sensor group is disabled and 1 runs the sensor group every second. You can specify a higher value to decrease the frequency at which the data is collected.

Whenever the configuration file is changed, you must stop and restart the pmsensor daemon by using the following commands:

1. Issue the **systemctl stop pmsensors** command to stop (deactivate) the sensor.
2. Issue the **systemctl start pmsensors** command to restart (activate) the sensor.

Some sensors such as the cluster export services sensors run on a specific set of nodes. Other sensors such as the GPFSDiskCap sensor must run on a single node in the cluster since the data reported is the same, independent of the node the sensor is running on. For these types of sensors, the restrict function is especially intended. For example, to restrict a **NFSIO** sensor to a node class and change the reporting period to once every 10 hours, you can specify **NFSIO.period=36000 NFSIO.restrict=nodeclass1** as attribute value pairs in the update command.

Some sensors, such as VFS, are not enabled by default even though they have associated predefined queries with the **mmperfmon query** command. This is so because the collector might display performance issues of its own if it is required to collect more than 1000000 metrics per second. To enable VFS sensors, use the **mmfsadm vfsstats enable** command on the node. To enable a sensor, set the period value to an integer greater than 0 and restart the sensors on that node by using the **systemctl restart pmsensors** command.

### List of performance metrics:

The performance monitoring tool can report the following metrics:

#### Network and general

All network and general metrics are native. There are no computed metrics in this section.

#### CPU

This section lists information about CPU in the system

- **cpu\_contexts**: Number of context switches across all CPU cores.
- **cpu\_guest**: Percentage of total CPU spent running a guest OS. Included in **cpu\_user**.
- **cpu\_guest\_nice**: Percentage of total CPU spent running as niced guest OS. Included in **cpu\_nice**.
- **cpu\_hiq**: Percentage of total CPU spent serving hardware interrupts.
- **cpu\_idle**: Percentage of total CPU spent idling.
- **cpu\_interrupts**: Number of interrupts serviced.
- **cpu\_iowait**: Percentage of total CPU spent waiting for I/O to complete.
- **cpu\_nice**: Percentage of total CPU time spent in lowest-priority user processes.
- **cpu\_siq**: Percentage of total CPU spent serving software interrupts.
- **cpu\_steal**: Percentage of total CPU spent waiting for other OS when running in a virtualized environment.
- **cpu\_system**: Percentage of total CPU time spent in kernel mode.
- **cpu\_user**: Percentage of total CPU time spent in normal priority user processes.

#### DiskFree

This section contains information about the free disk. Each mounted directory will have a separate section, for example **DiskFree|/boot/df\_free**

- **df\_free**: Amount of free disk space on the file system
- **df\_total**: Amount of total disk space on the file system
- **df\_used**: Amount of used disk space on the file system

## Diskstat

This section contains Disk status information for each of the disks. For example, `Diskstat|sda|disk_active_ios`.

- **disk\_active\_ios**: Number of I/O operations currently in progress.
- **disk\_aveq**: Weighted number of milliseconds spent doing I/Os.
- **disk\_io\_time**: Number of milliseconds the system spent doing I/O operation.
- **disk\_read\_ios**: Total number of read operations completed successfully.
- **disk\_read\_merged**: Number of (small) read operations that have been merged into a larger read.
- **disk\_read\_sect**: Number of sectors read.
- **disk\_read\_time**: Amount of time in milliseconds spent reading.
- **disk\_write\_ios**: Number of write operations completed successfully.
- **disk\_write\_merged**: Number of (small) write operations that have been merged into a larger write.
- **disk\_write\_sect**: Number of sectors written.
- **disk\_write\_time**: Amount of time in milliseconds spent writing.

## Load

- **jobs**: The total number of jobs that currently exist in the system.
- **load1**: The average load (number of jobs in the run queue) over the last minute.
- **load15**: The average load (number of jobs in the run queue) over the last 15 minutes.
- **load5**: The average load (number of jobs in the run queue) over the five minutes.

## Memory

- **mem\_active**: Active memory that was recently accessed.
- **mem\_active\_anon**: Active memory with no file association, that is, heap and stack memory.
- **mem\_active\_file**: Active memory that is associated with a file, for example, page cache memory.
- **mem\_buffers**: Temporary storage used for raw disk blocks.
- **mem\_cached**: In-memory cache for files read from disk (the page cache). Does not include `mem_swapcached`.
- **mem\_dirty**: Memory which is waiting to get written back to the disk.
- **mem\_inactive**: Inactive memory that hasn't been accessed recently.
- **mem\_inactive\_anon**: Inactive memory with no file association, that is, inactive heap and stack memory.
- **mem\_inactive\_file**: Inactive memory that is associated with a file, for example, page cache memory.
- **mem\_memfree**: Total free RAM.
- **mem\_mtotal**: Total usable RAM.
- **mem\_mlocked**: Memory that is locked.
- **mem\_swapcached**: In-memory cache for pages that are swapped back in.
- **mem\_swapfree**: Amount of swap space that is currently unused.
- **mem\_swaptotal**: Total amount of swap space available.
- **mem\_unevictable**: Memory that cannot be paged out.

## Netstat

- **ns\_closewait**: Number of connections in state `TCP_CLOSE_WAIT`
- **ns\_established**: Number of connections in state `TCP_ESTABLISHED`
- **ns\_listen**: Number of connections in state `TCP_LISTEN`
- **ns\_local\_bytes\_r**: Number of bytes received (local -> local)
- **ns\_local\_bytes\_s**: Number of bytes sent (local -> local)

- **ns\_localconn**: Number of local connections (local -> local)
- **ns\_remote\_bytes\_r**: Number of bytes sent (local -> remote)
- **ns\_remote\_bytes\_s**: Number of bytes sent (remote -> local)
- **ns\_remoteconn**: Number of remote connections (local -> remote)
- **ns\_timewait**: Number of connections in state TCP\_TIME\_WAIT

## Network

- **netdev\_bytes\_r**: Number of bytes received.
- **netdev\_bytes\_s**: Number of bytes sent.
- **netdev\_carrier**: Number of carrier loss events.
- **netdev\_collisions**: Number of collisions.
- **netdev\_compressed\_r**: Number of compressed frames received.
- **netdev\_compressed\_s**: Number of compressed packets sent.
- **netdev\_drops\_r**: Number of packets dropped while receiving.
- **netdev\_drops\_s**: Number of packets dropped while sending.
- **netdev\_errors\_r**: Number of read errors.
- **netdev\_errors\_s**: Number of write errors.
- **netdev\_fifo\_r**: Number of FIFO buffer errors.
- **netdev\_fifo\_s**: Number of FIFO buffer errors while sending.
- **netdev\_frames\_r**: Number of frame errors while receiving.
- **netdev\_multicast\_r**: Number of multicast packets received.
- **netdev\_packets\_r**: Number of packets received.
- **netdev\_packets\_s**: Number of packets sent.

## GPFS

### GPFSDisk

For each NSD in the system, for example `GPFSdisk|myMachine|myFilesystem|myNSD|gpfs_ds_bytes_read`

- **gpfs\_ds\_bytes\_read**: Number of bytes read.
- **gpfs\_ds\_bytes\_written**: Number of bytes written.
- **gpfs\_ds\_max\_disk\_wait\_rd**: The longest time spent waiting for a disk read operation.
- **gpfs\_ds\_max\_disk\_wait\_wr**: The longest time spent waiting for a disk write operation.
- **gpfs\_ds\_max\_queue\_wait\_rd**: The longest time between being enqueued for a disk read operation and the completion of that operation.
- **gpfs\_ds\_max\_queue\_wait\_wr**: The longest time between being enqueued for a disk write operation and the completion of that operation.
- **gpfs\_ds\_min\_disk\_wait\_rd**: The shortest time spent waiting for a disk read operation.
- **gpfs\_ds\_min\_disk\_wait\_wr**: The shortest time spent waiting for a disk write operation.
- **gpfs\_ds\_min\_queue\_wait\_rd**: The shortest time between being enqueued for a disk read operation and the completion of that operation.
- **gpfs\_ds\_min\_queue\_wait\_wr**: The shortest time between being enqueued for a disk write operation and the completion of that operation.
- **gpfs\_ds\_read\_ops**: Number of read operations.
- **gpfs\_ds\_tot\_disk\_wait\_rd**: The total time in seconds spent waiting for disk read operations.
- **gpfs\_ds\_tot\_disk\_wait\_wr**: The total time in seconds spent waiting for disk write operations.



- **gpfs\_ds\_tot\_queue\_wait\_rd**: The total time spent between being enqueued for a read operation and the completion of that operation.
- **gpfs\_ds\_tot\_queue\_wait\_wr**: The total time spent between being enqueued for a write operation and the completion of that operation.
- **gpfs\_ds\_write\_ops**: Number of write operations.

### GPFSFileSystem

For each file system, for example GPFSFileSystem

- **gpfs\_fs\_bytes\_read**: Number of bytes read.
- **gpfs\_fs\_bytes\_written**: Number of bytes written.
- **gpfs\_fs\_disks**: Number of disks in the file system.
- **gpfs\_fs\_max\_disk\_wait\_rd**: The longest time spent waiting for a disk read operation.
- **gpfs\_fs\_max\_disk\_wait\_wr**: The longest time spent waiting for a disk write operation.
- **gpfs\_fs\_max\_queue\_wait\_rd**: The longest time between being enqueued for a disk read operation and the completion of that operation.
- **gpfs\_fs\_max\_queue\_wait\_wr**: The longest time between being enqueued for a disk write operation and the completion of that operation.
- **gpfs\_fs\_min\_disk\_wait\_rd**: The shortest time spent waiting for a disk read operation.
- **gpfs\_fs\_min\_disk\_wait\_wr**: The shortest time spent waiting for a disk write operation.
- **gpfs\_fs\_min\_queue\_wait\_rd**: The shortest time between being enqueued for a disk read operation and the completion of that operation.
- **gpfs\_fs\_min\_queue\_wait\_wr**: The shortest time between being enqueued for a disk write operation and the completion of that operation.
- **gpfs\_fs\_read\_ops**: Number of read operations
- **gpfs\_fs\_tot\_disk\_wait\_rd**: The total time in seconds spent waiting for disk read operations.
- **gpfs\_fs\_tot\_disk\_wait\_wr**: The total time in seconds spent waiting for disk write operations.
- **gpfs\_fs\_tot\_queue\_wait\_rd**: The total time spent between being enqueued for a read operation and the completion of that operation.
- **gpfs\_fs\_tot\_queue\_wait\_wr**: The total time spent between being enqueued for a write operation and the completion of that operation.
- **gpfs\_fs\_write\_ops**: Number of write operations.

### GPFSFileSystemAPI

- **gpfs\_fis\_bytes\_read**: Number of bytes read.
- **gpfs\_fis\_bytes\_written**: Number of bytes written.
- **gpfs\_fis\_close\_calls**: Number of close calls.
- **gpfs\_fis\_disks**: Number of disks in the file system.
- **gpfs\_fis\_inodes\_written**: Number of inode updates to disk.
- **gpfs\_fis\_open\_calls**: Number of open calls.
- **gpfs\_fis\_read\_calls**: Number of read calls.
- **gpfs\_fis\_readdir\_calls**: Number of readdir calls.
- **gpfs\_fis\_write\_calls**: Number of write calls.

### GPFSNSDDisk

- **gpfs\_nsdds\_bytes\_read**: Number of bytes read.
- **gpfs\_nsdds\_bytes\_written**: Number of bytes written.
- **gpfs\_nsdds\_max\_disk\_wait\_rd**: The longest time spent waiting for a disk read operation.

- **gpfs\_nsdds\_max\_disk\_wait\_wr**: The longest time spent waiting for a disk write operation.
- **gpfs\_nsdds\_max\_queue\_wait\_rd**: The longest time between being enqueued for a disk read operation and the completion of that operation.
- **gpfs\_nsdds\_max\_queue\_wait\_wr**: The longest time between being enqueued for a disk write operation and the completion of that operation.
- **gpfs\_nsdds\_min\_disk\_wait\_rd**: The shortest time spent waiting for a disk read operation.
- **gpfs\_nsdds\_min\_disk\_wait\_wr**: The shortest time spent waiting for a disk write operation.
- **gpfs\_nsdds\_min\_queue\_wait\_rd**: The shortest time between being enqueued for a disk read operation and the completion of that operation.
- **gpfs\_nsdds\_min\_queue\_wait\_wr**: The shortest time between being enqueued for a disk write operation and the completion of that operation.
- **gpfs\_nsdds\_read\_ops**: Number of read operations.
- **gpfs\_nsdds\_tot\_disk\_wait\_rd**: The total time in seconds spent waiting for disk read operations.
- **gpfs\_nsdds\_tot\_disk\_wait\_wr**: The total time in seconds spent waiting for disk write operations.
- **gpfs\_nsdds\_tot\_queue\_wait\_rd**: The total time spent between being enqueued for a read operation and the completion of that operation.
- **gpfs\_nsdds\_tot\_queue\_wait\_wr**: The total time spent between being enqueued for a write operation and the completion of that operation.
- **gpfs\_nsdds\_write\_ops**: Number of write operations.

#### **GPFSNSDFS**

- **gpfs\_nsdfs\_bytes\_read**: Number of NSD bytes read, aggregated to the file system.
- **gpfs\_nsdfs\_bytes\_written**: Number of NSD bytes written, aggregated to the file system.
- **gpfs\_nsdfs\_read\_ops**: Number of NSD read operations, aggregated to the file system.
- **gpfs\_nsdfs\_write\_ops**: Number of NSD write operations, aggregated to the file system.

#### **GPFSNSDPool**

- **gpfs\_nsdpool\_bytes\_read**: Number of NSD bytes read, aggregated to the file system.
- **gpfs\_nsdpool\_bytes\_written**: Number of NSD bytes written, aggregated to the file system.
- **gpfs\_nsdpool\_read\_ops**: Number of NSD read operations, aggregated to the file system.
- **gpfs\_nsdpool\_write\_ops**: Number of NSD write operations, aggregated to the file system.

#### **GPFSNode**

- **gpfs\_ns\_bytes\_read**: Number of bytes read.
- **gpfs\_ns\_bytes\_written**: Number of bytes written.
- **gpfs\_ns\_clusters**: Number of clusters participating
- **gpfs\_ns\_disks**: Number of disks in all mounted file systems
- **gpfs\_ns\_filesys**: Number of mounted file systems
- **gpfs\_ns\_max\_disk\_wait\_rd**: The longest time spent waiting for a disk read operation.
- **gpfs\_ns\_max\_disk\_wait\_wr**: The longest time spent waiting for a disk write operation.
- **gpfs\_ns\_max\_queue\_wait\_rd**: The longest time between being enqueued for a disk read operation and the completion of that operation.
- **gpfs\_ns\_max\_queue\_wait\_wr**: The longest time between being enqueued for a disk write operation and the completion of that operation.
- **gpfs\_ns\_min\_disk\_wait\_rd**: The shortest time spent waiting for a disk read operation.
- **gpfs\_ns\_min\_disk\_wait\_wr**: The shortest time spent waiting for a disk write operation.
- **gpfs\_ns\_min\_queue\_wait\_rd**: The shortest time between being enqueued for a disk read operation and the completion of that operation.

- **gpfs\_ns\_min\_queue\_wait\_wr**: The shortest time between being enqueued for a disk write operation and the completion of that operation.
- **gpfs\_ns\_read\_ops**: Number of read operations.
- **gpfs\_ns\_tot\_disk\_wait\_rd**: The total time in seconds spent waiting for disk read operations.
- **gpfs\_ns\_tot\_disk\_wait\_wr**: The total time in seconds spent waiting for disk write operations.
- **gpfs\_ns\_tot\_queue\_wait\_rd**: The total time spent between being enqueued for a read operation and the completion of that operation.
- **gpfs\_ns\_tot\_queue\_wait\_wr**: The total time spent between being enqueued for a write operation and the completion of that operation.
- **gpfs\_ns\_write\_ops**: Number of write operations.

#### GPFSNodeAPI

- **gpfs\_is\_bytes\_read**: Number of bytes read.
- **gpfs\_is\_bytes\_written**: Number of bytes written.
- **gpfs\_is\_close\_calls**: Number of close calls.
- **gpfs\_is\_inodes\_written**: Number of inode updates to disk.
- **gpfs\_is\_open\_calls**: Number of open calls.
- **gpfs\_is\_readDir\_calls**: Number of readdir calls.
- **gpfs\_is\_read\_calls**: Number of read calls.
- **gpfs\_is\_write\_calls**: Number of write calls.

#### GPFSPoolIO

For the system within each GPFS device:

- **gpfs\_pool\_bytes\_rd**: Total size of all disks for this usage type.
- **gpfs\_pool\_bytes\_wr**: Total available disk space in full blocks for this usage type.
- **gpfs\_pool\_free\_fragkb**: Total available space in fragments for this usage type.

#### GPFSVFS

- **gpfs\_vfs\_accesses**: Number of accesses operations.
- **gpfs\_vfs\_accesses\_t**: Amount of time in seconds spent in accesses operations.
- **gpfs\_vfs\_aioread**: Number of aioread operations.
- **gpfs\_vfs\_aioread\_t**: Amount of time in seconds spent in aioread operations.
- **gpfs\_vfs\_aiowrite**: Number of aiowrite operations.
- **gpfs\_vfs\_aiowrite\_t**: Amount of time in seconds spent in aiowrite operations.
- **gpfs\_vfs\_clear**: Number of clear operations.
- **gpfs\_vfs\_clear\_t**: Amount of time in seconds spent in clear operations.
- **gpfs\_vfs\_close**: Number of close operations.
- **gpfs\_vfs\_close\_t**: Amount of time in seconds spent in close operations.
- **gpfs\_vfs\_create**: Number of create operations.
- **gpfs\_vfs\_create\_t**: Amount of time in seconds spent in create operations.
- **gpfs\_vfs\_decodeFh**: Number of decodeFh operations.
- **gpfs\_vfs\_decodeFh\_t**: Amount of time in seconds spent in decodeFh operations.
- **gpfs\_vfs\_detDentry**: Number of detDentry operations.
- **gpfs\_vfs\_encodeFh**: Number of encodeFh operations.
- **gpfs\_vfs\_encodeFh\_t**: Amount of time in seconds spent in encodeFh operations.
- **gpfs\_vfs\_flock**: Number of flock operations.

- **gpfs\_vfs\_flock\_t**: Amount of time in seconds spent in flock operations.
- **gpfs\_vfs\_fsync**: Number of fsync operations.
- **gpfs\_vfs\_fsyncRange**: Number of fsyncRange operations.
- **gpfs\_vfs\_fsyncRange\_t**: Amount of time in seconds spent in fsyncRange operations.
- **gpfs\_vfs\_fsync\_t**: Amount of time in seconds spent in fsync operations.
- **gpfs\_vfs\_ftrunc**: Number of ftrunc operations.
- **gpfs\_vfs\_ftrunc\_t**: Amount of time in seconds spent in ftrunc operations.
- **gpfs\_vfs\_getDentry\_t**: Amount of time in seconds spent in getDentry operations.
- **gpfs\_vfs\_getParent**: Number of getParent operations.
- **gpfs\_vfs\_getParent\_t**: Amount of time in seconds spent in getParent operations.
- **gpfs\_vfs\_getattr**: Number of getattr operations.
- **gpfs\_vfs\_getattr\_t**: Amount of time in seconds spent in getattr operations.
- **gpfs\_vfs\_getxattr**: Number of getxattr operations.
- **gpfs\_vfs\_getxattr\_t**: Amount of time in seconds spent in getxattr operations.
- **gpfs\_vfs\_link**: Number of link operations.
- **gpfs\_vfs\_link\_t**: Amount of time in seconds spent in link operations.
- **gpfs\_vfs\_listxattr**: Number of listxattr operations.
- **gpfs\_vfs\_listxattr\_t**: Amount of time in seconds spent in listxattr operations.
- **gpfs\_vfs\_lockctl**: Number of lockctl operations.
- **gpfs\_vfs\_lockctl\_t**: Amount of time in seconds spent in lockctl operations.
- **gpfs\_vfs\_lookup**: Number of lookup operations.
- **gpfs\_vfs\_lookup\_t**: Amount of time in seconds spent in lookup operations.
- **gpfs\_vfs\_mapLloff**: Number of mapLloff operations.
- **gpfs\_vfs\_mapLloff\_t**: Amount of time in seconds spent in mapLloff operations.
- **gpfs\_vfs\_mkdir**: Number of mkdir operations.
- **gpfs\_vfs\_mkdir\_t**: Amount of time in seconds spent in mkdir operations.
- **gpfs\_vfs\_mknod**: Number of mknod operations.
- **gpfs\_vfs\_mknod\_t**: Amount of time in seconds spent in mknod operations.
- **gpfs\_vfs\_mmapread**: Number of mmapread operations.
- **gpfs\_vfs\_mmapread\_t**: Amount of time in seconds spent in mmapread operations.
- **gpfs\_vfs\_mmapwrite**: Number of mmapwrite operations.
- **gpfs\_vfs\_mmapwrite\_t**: Amount of time in seconds spent in mmapwrite operation.
- **gpfs\_vfs\_mount**: Number of mount operations.
- **gpfs\_vfs\_mount\_t**: Amount of time in seconds spent in mount operations.
- **gpfs\_vfs\_open**: Number of open operations.
- **gpfs\_vfs\_open\_t**: Amount of time in seconds spent in open operations.
- **gpfs\_vfs\_read**: Number of read operations.
- **gpfs\_vfs\_read\_t**: Amount of time in seconds spent in read operations.
- **gpfs\_vfs\_readdir**: Number of readdir operations.
- **gpfs\_vfs\_readdir\_t**: Amount of time in seconds spent in readdir operations.
- **gpfs\_vfs\_readlink**: Number of readlink operations.
- **gpfs\_vfs\_readlink\_t**: Amount of time in seconds spent in readlink operations
- **gpfs\_vfs\_readpage**: Number of readpage operations.
- **gpfs\_vfs\_readpage\_t**: Amount of time in seconds spent in readpage operations.
- **gpfs\_vfs\_remove**: Number of remove operations.

- **gpfs\_vfs\_remove\_t**: Amount of time in seconds spent in remove operations.
- **gpfs\_vfs\_removexattr**: Number of removexattr operations.
- **gpfs\_vfs\_removexattr\_t**: Amount of time in seconds spent in removexattr operations.
- **gpfs\_vfs\_rename**: Number of rename operations.
- **gpfs\_vfs\_rename\_t**: Amount of time in seconds spent in rename operations.
- **gpfs\_vfs\_rmdir**: Number of rmdir operations.
- **gpfs\_vfs\_rmdir\_t**: Amount of time in seconds spent in rmdir operations.
- **gpfs\_vfs\_setacl**: Number of setacl operations.
- **gpfs\_vfs\_setacl\_t**: Amount of time in seconds spent in setacl operations.
- **gpfs\_vfs\_setattr**: Number of setattr operations.
- **gpfs\_vfs\_setattr\_t**: Amount of time in seconds spent in setattr operations.
- **gpfs\_vfs\_setxattr**: Number of setxattr operations.
- **gpfs\_vfs\_setxattr\_t**: Amount of time in seconds spent in setxattr operations.
- **gpfs\_vfs\_statfs**: Number of statfs operations.
- **gpfs\_vfs\_statfs\_t**: Amount of time in seconds spent in statfs operations.
- **gpfs\_vfs\_symlink**: Number of symlink operations.
- **gpfs\_vfs\_symlink\_t**: Amount of time in seconds spent in symlink operations.
- **gpfs\_vfs\_sync**: Number of sync operations.
- **gpfs\_vfs\_sync\_t**: Amount of time in seconds spent in sync operations.
- **gpfs\_vfs\_tsfattr**: Number of tsfattr operation.
- **gpfs\_vfs\_tsfattr\_t**: Amount of time in seconds spent in tsfattr operations.
- **gpfs\_vfs\_tsfattr**: Number of tsfattr operations.
- **gpfs\_vfs\_tsfattr\_t**: Amount of time in seconds spent in tsfattr operations.
- **gpfs\_vfs\_unmap**: Number of unmap operations.
- **gpfs\_vfs\_unmap\_t**: Amount of time in seconds spent in unmap operations.
- **gpfs\_vfs\_vget**: Number of vget operations.
- **gpfs\_vfs\_vget\_t**: Amount of time in seconds spent in vget operations.
- **gpfs\_vfs\_write**: Number of write operations.
- **gpfs\_vfs\_write\_t**: Amount of time in seconds spent in write operations.
- **gpfs\_vfs\_writepage**: Number of writepage operations.
- **gpfs\_vfs\_writepage\_t**: Amount of time in seconds spent in writepage operations.

## Computed Metrics

The following metrics are computed for GPFS:

- **gpfs\_write\_avg\_lat (latency)**:  $\text{gpfs\_vfs\_write\_t} / \text{gpfs\_vfs\_write}$
- **gpfs\_read\_avg\_lat (latency)**:  $\text{gpfs\_vfs\_read\_t} / \text{gpfs\_vfs\_read}$
- **gpfs\_create\_avg\_lat (latency)**:  $\text{gpfs\_vfs\_create\_t} / \text{gpfs\_vfs\_create}$
- **gpfs\_remove\_avg\_lat (latency)**:  $\text{gpfs\_vfs\_remove\_t} / \text{gpfs\_vfs\_remove}$

## NFS

### Native Metrics

- **nfs\_read\_req**: Number of bytes requested for reading.
- **nfs\_write\_req**: Number of bytes requested for writing.
- **nfs\_read**: Number of bytes transferred for reading.

- **nfs\_write**: Number of bytes transferred for writing.
- **nfs\_read\_ops**: Number of total read operations.
- **nfs\_write\_ops**: Number of total write operations.
- **nfs\_read\_err**: Number of erroneous read operations.
- **nfs\_write\_err**: Number of erroneous write operations.
- **nfs\_read\_lat**: Time consumed by read operations (in ns).
- **nfs\_write\_lat**: Time consumed by write operations (in ns).
- **nfs\_read\_queue**: Time spent in the rpc wait queue.
- **nfs\_write\_queue**: Time spent in the rpc wait queue.

## Computed Metrics

The following metrics are computed for NFS:

- **nfs\_total\_ops**:  $\text{nfs\_read\_ops} + \text{nfs\_write\_ops}$
- **nfsIOlatencyRead**:  $(\text{nfs\_read\_lat} + \text{nfs\_read\_queue}) / \text{nfs\_read\_ops}$
- **nfsIOlatencyWrite**:  $(\text{nfs\_write\_lat} + \text{nfs\_write\_queue}) / \text{nfs\_write\_ops}$
- **nfsReadOpThroughput**:  $\text{nfs\_read} / \text{nfs\_read\_ops}$
- **nfsWriteOpThroughput**:  $\text{nfs\_write} / \text{nfs\_write\_ops}$

## Object

### Native Metrics

#### ObjectAccount

- **account\_auditor\_time**: Timing data for individual account database audits.
- **account\_reaper\_time**: Timing data for each reap\_account() call.
- **account\_replicator\_time**: Timing data for each database replication attempt not resulting in a failure.
- **account\_DEL\_time**: Timing data for each DELETE request not resulting in an error.
- **account\_DEL\_err\_time**: Timing data for each DELETE request resulting in an error: bad request, not mounted, missing timestamp.
- **account\_GET\_time**: Timing data for each GET request not resulting in an error.
- **account\_GET\_err\_time**: Timing data for each GET request resulting in an error: bad request, not mounted, bad delimiter, account listing limit too high, bad accept header.
- **account\_HEAD\_time**: Timing data for each HEAD request not resulting in an error.
- **account\_HEAD\_err\_time**: Timing data for each HEAD request resulting in an error: bad request, not mounted.
- **account\_POST\_time**: Timing data for each POST request not resulting in an error.
- **account\_POST\_err\_time**: Timing data for each POST request resulting in an error: bad request, bad or missing timestamp, not mounted.
- **account\_PUT\_time**: Timing data for each PUT request not resulting in an error.
- **account\_PUT\_err\_time**: Timing data for each PUT request resulting in an error: bad request, not mounted, conflict, recently-deleted.
- **account\_REPLICATE\_time**: Timing data for each REPLICATE request not resulting in an error.
- **account\_REPLICATE\_err\_time**: Timing data for each REPLICATE request resulting in an error: bad request, not mounted.

#### ObjectContainer

- **container\_auditor\_time**: Timing data for each container audit.

- **container\_replicator\_time**: Timing data for each database replication attempt not resulting in a failure.
- **container\_DEL\_time**: Timing data for each DELETE request not resulting in an error.
- **container\_DEL\_err\_time**: Timing data for DELETE request errors: bad request, not mounted, missing timestamp, conflict.
- **container\_GET\_time**: Timing data for each GET request not resulting in an error.
- **container\_GET\_err\_time**: Timing data for GET request errors: bad request, not mounted, parameters not utf8, bad accept header.
- **container\_HEAD\_time**: Timing data for each HEAD request not resulting in an error.
- **container\_HEAD\_err\_time**: Timing data for HEAD request errors: bad request, not mounted.
- **container\_POST\_time**: Timing data for each POST request not resulting in an error.
- **container\_POST\_err\_time**: Timing data for POST request errors: bad request, bad x-container-sync-to, not mounted.
- **container\_PUT\_time**: Timing data for each PUT request not resulting in an error.
- **container\_PUT\_err\_time**: Timing data for PUT request errors: bad request, missing timestamp, not mounted, conflict.
- **container\_REPLICATE\_time**: Timing data for each REPLICATE request not resulting in an error.
- **container\_REPLICATE\_err\_time**: Timing data for REPLICATE request errors: bad request, not mounted.
- **container\_sync\_deletes\_time**: Timing data for each container database row synchronization via deletion.
- **container\_sync\_puts\_time**: Timing data for each container database row synchronization via PUTing.
- **container\_updater\_time**: Timing data for processing a container; only includes timing for containers which needed to update their accounts.

## ObjectObject

- **object\_auditor\_time**: Timing data for each object audit (does not include any rate-limiting sleep time for max\_files\_per\_second, but does include rate-limiting sleep time for max\_bytes\_per\_second).
- **object\_expirer\_time**: Timing data for each object expiration attempt, including ones resulting in an error.
- **object\_replicator\_partition\_delete\_time**: Timing data for partitions replicated to another node because they didn't belong on this node. This metric is not tracked per device.
- **object\_replicator\_partition\_update\_time**: Timing data for partitions replicated which also belong on this node. This metric is not tracked per-device.
- **object\_DEL\_time**: Timing data for each DELETE request not resulting in an error.
- **object\_DEL\_err\_time**: Timing data for DELETE request errors: bad request, missing timestamp, not mounted, precondition failed. Includes requests which couldn't find or match the object.
- **object\_GET\_time**: Timing data for each GET request not resulting in an error. Includes requests which couldn't find the object (including disk errors resulting in file quarantine).
- **object\_GET\_err\_time**: Timing data for GET request errors: bad request, not mounted, header timestamps before the epoch, precondition failed. File errors resulting in a quarantine are not counted here.
- **object\_HEAD\_time**: Timing data for each HEAD request not resulting in an error. Includes requests which couldn't find the object (including disk errors resulting in file quarantine).
- **object\_HEAD\_err\_time**: Timing data for HEAD request errors: bad request, not mounted.
- **object\_POST\_time**: Timing data for each POST request not resulting in an error.
- **object\_POST\_err\_time**: Timing data for POST request errors: bad request, missing timestamp, delete-at in past, not mounted.
- **object\_PUT\_time**: Timing data for each PUT request not resulting in an error.

- **object\_PUT\_err\_time:** Timing data for PUT request errors: bad request, not mounted, missing timestamp, object creation constraint violation, delete-at in past.
- **object\_REPLICATE\_time:** Timing data for each REPLICATE request not resulting in an error.
- **object\_REPLICATE\_err\_time:** Timing data for REPLICATE request errors: bad request, not mounted.
- **object\_updater\_time:** Timing data for object sweeps to flush async\_pending container updates. Does not include object sweeps which did not find an existing async\_pending storage directory.

## ObjectProxy

- **proxy\_account\_latency:** Timing data up to completion of sending the response headers, 200: standard response for successful HTTP requests.
- **proxy\_container\_latency:** Timing data up to completion of sending the response headers, 200: standard response for successful HTTP requests.
- **proxy\_object\_latency:** Timing data up to completion of sending the response headers, 200: standard response for successful HTTP requests.
- **proxy\_account\_GET\_time:** Timing data for GET request, start to finish, 200: standard response for successful HTTP requests
- **proxy\_account\_GET\_bytes:** The sum of bytes transferred in (from clients) and out (to clients) for requests, 200: standard response for successful HTTP requests.
- **proxy\_account\_HEAD\_time:** Timing data for HEAD request, start to finish, 204: request processed, no content returned.
- **proxy\_account\_HEAD\_bytes:** The sum of bytes transferred in (from clients) and out (to clients) for requests, 204: request processed, no content returned.
- **proxy\_container\_DEL\_time:** Timing data for DELETE request, start to finish, 204: request processed, no content returned.
- **proxy\_container\_DEL\_bytes:** The sum of bytes transferred in (from clients) and out (to clients) for requests, 204: request processed, no content returned.
- **proxy\_container\_GET\_time:** Timing data for GET request, start to finish, 200: standard response for successful HTTP requests.
- **proxy\_container\_GET\_bytes:** The sum of bytes transferred in (from clients) and out (to clients) for requests, 200: standard response for successful HTTP requests.
- **proxy\_container\_HEAD\_time:** Timing data for HEAD request, start to finish, 204: request processed, no content returned.
- **proxy\_container\_HEAD\_bytes:** The sum of bytes transferred in (from clients) and out (to clients) for requests, 204: request processed, no content returned. 1
- **proxy\_container\_PUT\_time:** Timing data for each PUT request not resulting in an error, 201: request has been fulfilled; new resource created.
- **proxy\_container\_PUT\_bytes:** The sum of bytes transferred in (from clients) and out (to clients) for requests, 201: request has been fulfilled; new resource created.
- **proxy\_object\_DEL\_time:** Timing data for DELETE request, start to finish, 204: request processed, no content returned.
- **proxy\_object\_DEL\_bytes:** The sum of bytes transferred in (from clients) and out (to clients) for requests, 204: request processed, no content returned.
- **proxy\_object\_GET\_time:** Timing data for GET request, start to finish, 200: standard response for successful HTTP requests.
- **proxy\_object\_GET\_bytes:** The sum of bytes transferred in (from clients) and out (to clients) for requests, 200: standard response for successful HTTP requests.
- **proxy\_object\_HEAD\_time:** Timing data for HEAD request, start to finish, 200: request processed, no content returned.
- **proxy\_object\_HEAD\_bytes:** The sum of bytes transferred in (from clients) and out (to clients) for requests, , 200: request processed, no content returned.



- **proxy\_object\_PUT\_time**: Timing data for each PUT request not resulting in an error, 201: request has been fulfilled; new resource created.
- **proxy\_object\_PUT\_bytes**: The sum of bytes transferred in (from clients) and out (to clients) for requests, 201: request has been fulfilled; new resource created.

**Note:** There are no computed metrics for Object protocol.

## SMB

### SMBGlobalStats

- **connect count**: Number of connections since startup of parent smbd process
- **disconnect count**: Number of connections closed since startup
- **idle**: Describes idling behavior of smbds
  - **count**: Number of times the smbd processes are waiting for events in epoll
  - **time**: Times the smbd process spend in epoll waiting for events
- **cpu\_user time**: The user time determined by the get\_rusage system call in seconds
- **cpu\_system time**: The system time determined by the get\_rusage system call in seconds
- **request count**: Number of SMB requests since startup
- **push\_sec\_ctx**: Smbds switch between the user and the root security context; push allows to put the current context onto a stack
  - **count**: Number of time the current security context is pushed onto the stack
  - **time**: The time it takes to put the current security context; this includes all syscalls required to save the current context on the stack
- **pop\_sec\_ctx**: Getting the last security context from the stack and restore it
  - **count**: Number of times the current security context is restored from the stack
  - **time**: The time it takes to put the restore the security context from the stack; this includes all syscalls required to get restore the security context from the stack
- **set\_sec\_ctx**:
  - **count**: Number of times the security context is set for user
  - **time**: The time it takes to set the security context for user
- **set\_root\_sec\_ctx**:
  - **count**: Number of times the security context is set for user
  - **time**: The time it takes to set the security context for user

### SMB2 metrics

These metrics are available for all of the following areas:

- **op\_count**: Number of times the corresponding SMB request has been called.
- **op\_idle**
  - **for notify**: Time between notification request and a corresponding notification being sent
  - **for oplock breaks**: Time waiting until an oplock is broken
  - for all others the value is always zero
- **op\_inbytes**: Number of bytes received for the corresponding request including protocol headers
- **op\_outbytes**: Number of bytes sent for the corresponding request including protocol headers.
- **op\_time**: The total amount of time spent for all corresponding SMB2 requests.

### CTDB

- **CTDB version**: Version of the CTDB protocol used by the node.

- **Current time of statistics:** Time when the statistics are generated. This is useful when collecting statistics output periodically for post-processing.
- **Statistics collected since:** Time when CTDB was started or the last time statistics was reset. The output shows the duration and the timestamp.
- **num\_clients:** Number of processes currently connected to CTDB's UNIX socket. This includes recovery daemon, CTDB tool and SMB processes (smbd, winbindd).
- **frozen:** 1 if the databases are currently frozen, 0 if otherwise.
- **recovering:** 1 if recovery is active, 0 if otherwise.
- **num\_recoveries:** Number of recoveries since the start of CTDB or since the last statistics reset.
- **client\_packets\_sent:** Number of packets sent to client processes via UNIX domain socket.
- **client\_packets\_recv:** Number of packets received from client processes via UNIX domain socket.
- **node\_packets\_sent:** Number of packets sent to the other nodes in the cluster via TCP.
- **node\_packets\_recv:** Number of packets received from the other nodes in the cluster via TCP.
- **keepalive\_packets\_sent:** Number of keepalive messages sent to other nodes. CTDB periodically sends keepalive messages to other nodes. For more information, see the `KeepAliveInterval` tunable in `CTDB-tunables(7)` on the CTDB documentation website.
- **keepalive\_packets\_recv:** Number of keepalive messages received from other nodes.
- **node:** This section lists various types of messages processed which originated from other nodes via TCP.
  - **req\_call:** Number of `REQ_CALL` messages from the other nodes.
  - **reply\_call:** Number of `REPLY_CALL` messages from the other nodes.
  - **req\_dmaster:** Number of `REQ_DMASTER` messages from the other nodes.
  - **reply\_dmaster:** Number of `REPLY_DMASTER` messages from the other nodes.
  - **reply\_error:** Number of `REPLY_ERROR` messages from the other nodes.
  - **req\_message:** Number of `REQ_MESSAGE` messages from the other nodes.
  - **req\_control:** Number of `REQ_CONTROL` messages from the other nodes.
  - **reply\_control:** Number of `REPLY_CONTROL` messages from the other nodes.
- **client:** This section lists various types of messages processed which originated from clients via UNIX domain socket.
  - **req\_call:** Number of `REQ_CALL` messages from the clients.
  - **req\_message:** Number of `REQ_MESSAGE` messages from the clients.
  - **req\_control:** Number of `REQ_CONTROL` messages from the clients.
- **timeouts:** This section lists timeouts occurred when sending various messages.
  - **call:** Number of timeouts for `REQ_CALL` messages.
  - **control:** Number of timeouts for `REQ_CONTROL` messages.
  - **traverse:** Number of timeouts for database traverse operations.
- **locks:** This section lists locking statistics.
  - **num\_calls:** Number of completed lock calls. This includes database locks and record locks.
  - **num\_current:** Number of scheduled lock calls. This includes database locks and record locks.
  - **num\_pending:** Number of queued lock calls. This includes database locks and record locks.
  - **num\_failed:** Number of failed lock calls. This includes database locks and record locks.
- **total\_calls:** Number of `req_call` messages processed from clients. This number should be same as `client --> req_call`.
- **pending\_calls:** Number of `req_call` messages which are currently being processed. This number indicates the number of record migrations in flight.
- **childwrite\_calls:** Number of record update calls. Record update calls are used to update a record under a transaction.

- **pending\_childwrite\_calls**: Number of record update calls currently active.
- **memory\_used**: The amount of memory in bytes currently used by CTDB using `talloc`. This includes all the memory used for CTDB's internal data structures. This does not include the memory mapped TDB databases.
- **max\_hop\_count**: The maximum number of hops required for a record migration request to obtain the record. High numbers indicate record contention.
- **total\_ro\_delegations**: Number of read-only delegations created.
- **total\_ro\_revokes**: Number of read-only delegations that were revoked. The difference between `total_ro_revokes` and `total_ro_delegations` gives the number of currently active read-only delegations.
- **hop\_count\_buckets**: Distribution of migration requests based on hop counts values.
- **lock\_buckets**: Distribution of record lock requests based on time required to obtain locks. Buckets are `< 1ms`, `< 10ms`, `< 100ms`, `< 1s`, `< 2s`, `< 4s`, `< 8s`, `< 16s`, `< 32s`, `< 64s`, `> 64s`.
- **locks\_latency**: The minimum, the average and the maximum time (in seconds) required to obtain record locks.
- **relock\_ctdbd**: The minimum, the average and the maximum time (in seconds) required to check if recovery lock is still held by recovery daemon when recovery mode is changed. This check is done in `ctdbd` daemon.
- **relock\_recd**: The minimum, the average and the maximum time (in seconds) required to check if recovery lock is still held by recovery daemon during recovery. This check is done in `recovery` daemon.
- **call\_latency**: The minimum, the average and the maximum time (in seconds) required to process a `REQ_CALL` message from client. This includes the time required to migrate a record from remote node, if the record is not available on the local node.
- **childwrite\_latency**: The minimum, the average and the maximum time (in seconds) required to update records under a transaction.

#### Cross Protocol

- **nfs\_iorate\_read\_perc**:  $\text{nfs\_read\_ops} / (\text{op\_count} + \text{nfs\_read\_ops})$
- **nfs\_iorate\_read\_perc\_exports**:  $1.0 * \text{nfs\_read\_ops} / (\text{op\_count} + \text{nfs\_read\_ops})$
- **nfs\_iorate\_write\_perc**:  $\text{nfs\_write\_ops} / (\text{write} | \text{op\_count} + \text{nfs\_write\_ops})$
- **nfs\_iorate\_write\_perc\_exports**:  $1.0 * \text{nfs\_write\_ops} / (\text{op\_count} + \text{nfs\_write\_ops})$
- **nfs\_read\_throughput\_perc**:  $\text{nfs\_read} / (\text{read} | \text{op\_outbytes} + \text{nfs\_read})$
- **nfs\_write\_throughput\_perc**:  $\text{nfs\_write} / (\text{write} | \text{op\_outbytes} + \text{nfs\_write})$
- **smb\_iorate\_read\_perc**:  $\text{op\_count} / (\text{op\_count} + \text{nfs\_read\_ops})$
- **smb\_iorate\_write\_perc**:  $\text{op\_count} / (\text{op\_count} + \text{nfs\_write\_ops})$
- **smb\_latency\_read**:  $\text{read} | \text{op\_time} / \text{read} | \text{op\_count}$
- **smb\_latency\_write**:  $\text{write} | \text{op\_time} / \text{write} | \text{op\_count}$
- **smb\_read\_throughput\_perc**:  $\text{read} | \text{op\_outbytes} / (\text{read} | \text{op\_outbytes} + \text{nfs\_read})$
- **smb\_total\_cnt**:  $\text{write} | \text{op\_count} + \text{close} | \text{op\_count}$
- **smb\_tp**:  $\text{op\_inbytes} + \text{op\_outbytes}$
- **smb\_write\_throughput\_perc**:  $\text{write} | \text{op\_outbytes} / (\text{write} | \text{op\_outbytes} + \text{nfs\_write})$
- **total\_read\_throughput**:  $\text{nfs\_read} + \text{read} | \text{op\_outbytes}$
- **total\_write\_throughput**:  $\text{nfs\_write} + \text{write} | \text{op\_inbytes}$

#### Transparent Cloud Tiering

- **mcs\_total\_bytes**: Total number of bytes uploaded to or downloaded from the cloud storage tier
- **mcs\_total\_requests**: Total number of migration, recall, or remove requests
- **mcs\_total\_request\_time**: Time (in second) taken for all migration, recall, or remove requests
- **mcs\_total\_failed\_requests**: Total number of failed migration, recall, or remove requests

- **mcs\_total\_failed\_requests\_time:** The total time (msec) spent in failed migration, recall, or remove requests
- **mcs\_total\_persisted\_bytes:** The total number of transferred bytes that are successfully persisted on the cloud provider. This is used for both migrate and recall operations.
- **mcs\_total\_retried\_operations:** The total number of retry PUT operations. This is used for both migrate and recall operations.
- **mcs\_total\_operation\_errors:** The total number of error PUT/GET operations based on the operation specified in the mcs\_operation key

## Configuring the collector

The following section describes how to configure the collector in a performance monitoring tool.

The most important configuration options are the domains and the peers configuration options. All other configuration options are best left at their defaults and are explained within the default configuration file shipped with ZIMon.

The configuration file of the collector, ZIMonCollector.cfg, is located in the /opt/IBM/zimon/ folder.

## Metric Domain Configuration

The domains configuration indicates the number of metrics to be collected and how long they must be retained and in what granularity. Multiple domains might be specified. If data no longer fits into the current domain, data is spilled over into the next domain and re-sampled.

A simple configuration is:

```
domains = {
# this is the raw domain, aggregation factor for the raw domain is always 0
aggregation = 0
ram = "500m" # amount of RAM to be used
duration = "12h"
filesize = "1g" # maximum file size
files = 16 # number of files.
}

,
{
# this is the second domain that aggregates to 60 seconds
aggregation = 60
ram = "500m" # amount of RAM to be used
duration = "4w"
filesize = "500m" # maximum file size
files = 4 # number of files.
}

,
{
# this is the third domain that aggregates to 30*60 seconds == 30 minutes
aggregation = 30
ram = "500m" # amount of RAM to be used
duration = "1y"
filesize = "500m" # maximum file size
files = 4 # number of files.
}
```

The configuration file lists several data domains. At least one domain must be present and the first domain represents the raw data collection as the data is collected by sensors. The aggregation parameter for this first domain must be set to 0.

Each domain specifies the following parameters:

- The **duration** parameter indicates the time period until the collected metrics are pushed into the next (coarser-grained) domain. If this option is left out, no limit on the duration is imposed.
- The **ram** parameter indicates the amount of RAM to be allocated for the domain. Once that amount of RAM is filled up, collected metrics are pushed into the next (coarser-grained) domain. If this option is left out, no limit on the amount of RAM available is imposed.
- The **filesize** and **files** parameter indicates how much space is allocated on disk for a given domain. While storing metrics in memory, there is a persistence mechanism in place that also stores the metrics on disk in files of size **filesize**. Once the number of files is reached and a new file is to be allocated, the oldest file is removed from the disk. The persistent storage must be at least as large as the amount of main memory to be allocated for a domain because when the collector is restarted, the in-memory database is re-created from these files. Queries can also be served from these files if they are executed in archive mode (-a).

The aggregation value, which is used for the second and following domains, indicates the resampling to be performed. Once data is spilled into this domain, the data is resampled to be no better than indicated by the aggregation factor. The value for the second domain is in seconds, the value for domain n (n>2) is the value of domain n-1 multiplied by the aggregation value of domain n.

The collector collects the metrics from the sensors. For example, in a five-node cluster where only the load values (load1, load5, load15) are reported, the collector will maintain 15 metrics (3 metrics times 5 nodes). Depending on the number of metrics that are collected, the collector requires a different amount of main memory to store the collected metrics in memory. Assuming 500000 metrics are collected, the following configurations are possible. Depending on the amount of data to be collected, 500000 metrics corresponds to about 1000 nodes.

Configuration 1 (4GB of RAM). Domain one configured at one second granularity for a period of six hours, domain 2 configured at 30 seconds granularity for the next two days, domain 3 configured at 15 minutes granularity for the next two weeks and domain 4 configured at 6-hour granularity for the next 2 months.

Configuration 2 (16GB of RAM). Domain one configured at 1 second granularity for a period of one day, domain 2 configured at 30 sec granularity for the next week, domain 3 configured at 15 minute granularity for the next two months and domain 4 configured at 6-hour granularity for the next year.

**Note:** The above computation only gives the memory required for the in-memory database, not including the indices necessary for the persistent storage or for the collector program itself.

The collectors can be stopped (deactivated) using the **systemctl stop pmcollector** command.

The collectors can be started (activated) using the **systemctl start pmcollector** command.

### Configuring multiple collectors:

The performance monitoring tool installation can have a single collector, or can consist of multiple collectors to increase the scalability or the fault-tolerance of the performance monitoring system. This latter configuration is referred to as “federation”.

**Note:** For federation to work, all the collectors need to have the same version number.

In a multi-collector federated configuration, the collectors need to know about each other, else a collector would only return the data stored in its own measurement database. Once the collectors know the peer collectors, they will collaborate with each other to collect data for a given measurement query. All collectors that are part of the federation are specified in the peers configuration option in the collector’s configuration file as shown below:

```
peers = {
host = "collector1.mydomain.com"
port = "9085"
}, {
host = "collector2.mydomain.com"
port = "9085"
}
```

The port number is the one specified by the `federationport` configuration option, typically set to 9085. It is acceptable to list the current host as well so that the same configuration file can be used for all the collector machines.

Once the peers have been specified, a query for measurement data can be directed to any of the collectors listed in the peers section, and the collector will collect and assemble a response based on all relevant data from all collectors. Hence, clients only need to contact a single collector in order to get all the measurements available in the system.

To distribute the measurement data reported by sensors over multiple collectors, multiple collectors may be specified when automatically configuring the sensors, as shown in the following sample:

```
prompt# mmperfmon config generate \
--collectors collector1.domain.com,collector2.domain.com,...
```

If multiple collectors are specified, the sensors will pick one of the many collectors to report their measurement data to. The sensors use stable hashes to pick the collector such that the sensor-collector relationship does not change too much if new collectors are added or if a collector is removed.

Additionally, sensors and collectors can be configured for high availability. To maintain high availability each metric should be sent to two collectors in case one collector becomes unavailable. In this setting, sensors report their measurement data to more than one collector, so that the failure of a single collector would not lead to any data loss. For instance, if the collector redundancy is increased to two, every sensor will report to two collectors. As a side-effect of increasing the redundancy to two, the bandwidth consumed for reporting measurement data will be duplicated. The collector redundancy has to be configured before the sensor configuration is stored in GPFS by changing the `colRedundancy` option in `/opt/IBM/zimon/defaults/ZIMonSensors.cfg` as explained in the sensor configuration section.

## Enabling protocol metrics

The type of information that is collected for NFS, SMB and Object protocols are configurable. This section describes the location of the configuration data for these protocols.

Configuration information for SMB and NFS in the `ZimonSensors.cfg` file references the sensor definition files in the `/opt/IBM/zimon` folder. For example:

- The `CTDBDBStats.cfg` file is referred in:

```
{
    name = "CTDBDBStats"
    period = 1
    type = "Generic"
},
```

- The `CTDBStats.cfg` file is referred in:

```
{
    name = "CTDBStats"
    period = 1
    type = "Generic"
},
```

- The `NFSIO.cfg` file is referred in:

```
{
    # NFS Ganesha statistics
    name = "NFSIO"
    period = 1
    type = "Generic"
},
```

- The `SMBGlobalStats.cfg` file is referred in:

```
{
    name = "SMBGlobalStats"
    period = 1
    type = "Generic"
},
```

- The `SMBStats.cfg` file is referred in:

```
{
    name = "SMBStats"
    period = 1
    type = "Generic"
},
```

At the time of installation, the object metrics proxy is configured to start by default on each Object protocol node.

The object metrics proxy server, `pmswiftd` is controlled by the corresponding service script called `pmswiftd`, located at `/etc/rc.d/init.d/pmswiftd.service`. You can start and stop the `pmswiftd` service script using the `systemctl start pmswiftd` and `systemctl stop pmswiftd` commands respectively. You can also view the status of the `pmswiftd` service script by using the `systemctl status pmswiftd` command.

In a system restart, the object metrics proxy server restarts automatically. In case of a failover, the server will start automatically. If for some reason this does not occur, the server must be started manually using the `systemctl start pmswiftd` command.

## Restarting the performance monitoring tool

You must manually restart the performance monitoring tool after you install a later version or reconfigure it.

Use the `systemctl start pmsensors` command to start monitoring on that node.

Use the `systemctl start pmcollector` command on the node running the collector.

When restarting the tool, the sensors identify the collector from the information given in the `ZimonSensors.cfg` file that is updated on every node.

The Performance monitoring tool also needs to restart if you make changes in the configuration files of either sensors, `ZimonSensors.cfg`, or the collector, `ZIMonCollector.cfg`, located in the `/opt/IBM/zimon/` folder.

To restart the sensor on a particular node, use the `systemctl restart pmsensors` command.

To restart the collector, use the `systemctl restart pmcollector` command.

For information on restarting the sensors and collectors for Transparent Cloud Tiering, see *Integrating Transparent Cloud Tiering metrics with performance monitoring tool* in *IBM Spectrum Scale: Administration Guide*.

## Configuring the metrics to collect performance data

For performance reasons, the performance monitoring tool by default does not collect all the available metrics. You can add other metrics to focus on particular performance problems.

For the available metrics, see “List of performance metrics” on page 46.

After you define the required metrics in the configuration file, you must restart all the sensors. To restart the sensors, run the following command:

```
systemctl restart pmsensors
```

## Viewing performance data with `mmperfmon`

To view the metrics associated with GPFS and the associated protocols, run the `mmperfmon` command with the `query` option. You can also use the `mmperfmon` command with the `query` option to detect performance issues and problems. You can collect metrics for all nodes or for a particular node.

- Problem: System slowing down

Use `mmperfmon query compareNodes cpu_user` or `mmperfmon query compareNodes cpu_system` command to compare CPU metrics for all the nodes in your system.

1. Check if there is a node that has a significantly higher CPU utilization for the entire time period. If so, see if this trend continues. You might need to investigate further on this node.
2. Check if there is a node that has significantly lower CPU utilization over the entire period. If so, check if that node has a health problem.
3. Use `mmperfmon query compareNodes protocolThroughput` to look at the throughput for each of the nodes for the different protocols.

**Note:** Note that the metrics of each individual protocol cannot always include exact I/O figures.

4. Use `mmperfmon query compareNodes protocolIORate` to look at the I/O performance for each of the nodes in your system.

- Problem: A particular node is causing problems

Use `mmperfmon query usage` to show the CPU, memory, storage, and network usage.

- Problem: A particular protocol is causing problems

Use `mmperfmon query` to investigate problems with your specific protocol. You can compare cross-node metrics using `mmperfmon query compareNodes`.

For example, `mmperfmon query compareNodes nfs_read_ops`.

Compare the NFS read operations on all the nodes that are using NFS. By comparing the different NFS metrics, you can identify which node is causing the problems. The problem might either manifest itself as running with much higher values than the other nodes, or much lower (depending on the issue) when considered over several buckets of time.

- Problem: A particular protocol is causing problems on a particular node.

Use `mmperfmon query` on the particular node to look deeper into the protocol performance on that node.

For example, if there is a problem with NFS:

- `mmperfmon query nfsIOlatency` - To get details of the nfsIOlatency.
- `mmperfmon query nfsIORate` - To get details of the NFS I/O rate.
- `mmperfmon query nfsThroughput` - To get details of the NFS throughput.

For more information on `mmperfmon`, see `mmperform` in *IBM Spectrum Scale: Command and Programming Reference*

## List of queries

You can make the following predefined queries with `query` option of the `mmperfmon` command.

### General and network

- **usage:** Retrieves details about the CPU, memory, storage and network usage
- **cpu:** Retrieves details of the CPU utilization in system and user space, and context switches.



- | • **netDetails**: Retrieves details about the network.
- | • **NetErrors**: Retrieves details about network problems, such as collisions, drops, and errors, for all available networks.
- | • **compareNodes**: Compares a single metric across all nodes running sensors

## | **GPFS**

| GPFS metric queries gives an overall view of the GPFS without considering the protocols.

- | • **gpfsCRUDopsLatency**: Retrieves information about the GPFS CRUD operations latency
- | • **gpfsFSWaits**: Retrieves information on the maximum waits for read and write operations for all file systems.
- | • **gpfsNSDWaits**: Retrieves information on the maximum waits for read and write operations for all disks.
- | • **gpfsNumberOperations**: Retrieves the number of operations to the GPFS file system.
- | • **gpfsVFSOpCounts**: Retrieves VFS operation counts.

## | **Cross protocol**

| These queries retrieve information after comparing metrics between different protocols on a particular node.

- | • **protocolIOLatency**: Compares latency per protocol (SMB, NFS, Object).
- | • **protocolIORate**: Retrieves the percentage of total I/O rate per protocol (SMB, NFS, Object).
- | • **protocolThroughput**: Retrieves the percentage of total throughput per protocol (SMB, NFS, Object).

## | **NFS**

| These queries retrieve metrics associated with the NFS protocol.

- | • **nfsIOLatency**: Retrieves the NFS I/O Latency in nanoseconds.
- | • **nfsIORate**: Retrieves the NFS I/O operations per second (NFS IOPS).
- | • **nfsThroughput**: Retrieves the NFS Throughput in bytes per second.
- | • **nfsErrors**: Retrieves the NFS error count for read and write operations.
- | • **nfsQueue**: Retrieves the NFS read and write queue latency in nanoseconds.
- | • **nfsThroughputPerOp**: Retrieves the NFS read and write throughput per op in bytes

## | **Object**

- | • **objAcc**: Details on the Object Account performance

| Retrieved metrics:

- | – **account\_auditor\_time**
- | – **account\_reaper\_time**
- | – **account\_replicator\_time**
- | – **account\_DEL\_time**
- | – **account\_DEL\_err\_time**
- | – **account\_GET\_time**
- | – **account\_GET\_err\_time**
- | – **account\_HEAD\_time**
- | – **account\_HEAD\_err\_time**
- | – **account\_POST\_time**
- | – **account\_POST\_err\_time**
- | – **account\_PUT\_time**

- |   – **account\_PUT\_err\_time**
- |   – **account\_REPLICATE\_time**
- |   – **account\_REPLICATE\_err\_time**
- | • **objCon:** Details on the Object Container performance
- |   Retrieved metrics:
- |   – **container\_auditor\_time**
- |   – **container\_replicator\_time**
- |   – **container\_DEL\_time**
- |   – **container\_DEL\_err\_time**
- |   – **container\_GET\_time**
- |   – **container\_GET\_err\_time**
- |   – **container\_HEAD\_time**
- |   – **container\_HEAD\_err\_time**
- |   – **container\_POST\_time**
- |   – **container\_POST\_err\_time**
- |   – **container\_PUT\_time**
- |   – **container\_PUT\_err\_time**
- |   – **container\_REPLICATE\_time**
- |   – **container\_REPLICATE\_err\_time**
- |   – **container\_sync\_deletes\_time**
- |   – **container\_sync\_puts\_time**
- |   – **container\_updater\_time**
- | • **objObj:** Details on the Object performance
- |   Retrieved metrics:
- |   – **object\_auditor\_time**
- |   – **object\_expirer\_time**
- |   – **object\_replicator\_partition\_delete\_time**
- |   – **object\_replicator\_partition\_update\_time**
- |   – **object\_DEL\_time**
- |   – **object\_DEL\_err\_time**
- |   – **object\_GET\_time**
- |   – **object\_GET\_err\_time**
- |   – **object\_HEAD\_time**
- |   – **object\_HEAD\_err\_time**
- |   – **object\_POST\_time**
- |   – **object\_POST\_err\_time**
- |   – **object\_PUT\_time**
- |   – **object\_PUT\_err\_time**
- |   – **object\_REPLICATE\_err\_time**
- |   – **object\_REPLICATE\_time**
- |   – **object\_updater\_time**
- | • **objPro:** Details on the Object Proxy performance
- |   Retrieved metrics:
- |   – **proxy\_account\_latency**
- |   – **proxy\_container\_latency**

```

| - proxy_object_latency
| - proxy_account_GET_time
| - proxy_account_GET_bytes
| - proxy_account_HEAD_time
| - proxy_account_HEAD_bytes
| - proxy_container_DEL_time
| - proxy_container_DEL_bytes
| - proxy_container_GET_time
| - proxy_container_GET_bytes
| - proxy_container_HEAD_time
| - proxy_container_HEAD_bytes
| - proxy_container_PUT_time
| - proxy_container_PUT_bytes
| - proxy_object_DEL_time
| - proxy_object_DEL_bytes
| - proxy_object_GET_time
| - proxy_object_GET_bytes
| - proxy_object_HEAD_time
| - proxy_object_HEAD_bytes
| - proxy_object_PUT_time
| - proxy_object_PUT_bytes
| • objAccIO: Information on the Object Account IO rate
| Retrieved metrics:
| - account_GET_time
| - account_GET_err_time
| - account_HEAD_time
| - account_HEAD_err_time
| - account_POST_time
| - account_POST_err_time
| - account_PUT_time
| - account_PUT_err_time
| • objConIO: Information on the Object Container IO rate
| Retrieved metrics:
| - container_GET_time
| - container_GET_err_time
| - container_HEAD_time
| - container_HEAD_err_time
| - container_POST_time
| - container_POST_err_time
| - container_PUT_time
| - container_PUT_err_time
| • objObjIO: Information on the Object Object IO rate
| Retrieved metrics:
| - object_GET_time
| - object_GET_err_time

```

- |   – **object\_HEAD\_time**
- |   – **object\_HEAD\_err\_time**
- |   – **object\_POST\_time**
- |   – **object\_POST\_err\_time**
- |   – **object\_PUT\_time**
- |   – **object\_PUT\_err\_time**
- | • **objProIO**: Information on the Object Proxy IO rate
- |   Retrieved metrics:
- |   – **proxy\_account\_GET\_time**
- |   – **proxy\_account\_GET\_bytes**
- |   – **proxy\_container\_GET\_time**
- |   – **proxy\_container\_GET\_bytes**
- |   – **proxy\_container\_PUT\_time**
- |   – **proxy\_container\_PUT\_bytes**
- |   – **proxy\_object\_GET\_time**
- |   – **proxy\_object\_GET\_bytes**
- |   – **proxy\_object\_PUT\_time**
- |   – **proxy\_object\_PUT\_bytes**
- | • **objAccThroughput**: Information on the Object Account Throughput
- |   Retrieved metrics:
- |   – **account\_GET\_time**
- |   – **account\_PUT\_time**
- | • **objConThroughput**: Information on the Object Container Throughput
- |   Retrieved metrics:
- |   – **container\_GET\_time**
- |   – **container\_PUT\_time**
- | • **objObjThroughput**: Information on the Object Throughput
- |   Retrieved metrics:
- |   – **object\_GET\_time**
- |   – **object\_PUT\_time**
- |   –
- | • **objProThroughput**: Information on the Object Proxy Throughput
- |   Retrieved metrics:
- |   – **proxy\_account\_GET\_time**
- |   – **proxy\_account\_GET\_bytes**
- |   – **proxy\_container\_GET\_time**
- |   – **proxy\_container\_GET\_bytes**
- |   – **proxy\_container\_PUT\_time**
- |   – **proxy\_container\_PUT\_bytes**
- |   – **proxy\_object\_GET\_time**
- |   – **proxy\_object\_GET\_bytes**
- |   – **proxy\_object\_PUT\_time**
- |   – **proxy\_object\_PUT\_bytes**
- | • **objAccLatency**: Information on the Object Account Latency
- |   Retrieved metric:

- |    – **proxy\_account\_latency**
- | • **objConLatency**: Information on the Object Container Latency
- |    Retrieved metric:
- |    – **proxy\_container\_latency**
- | • **objObjLatency**: Information on the Object Latency
- |    Retrieved metric:
- |    – **proxy\_object\_latency**

## | **SMB**

- | These queries retrieve metrics associated with SMB.
- | • **smb2IOLatency**: Retrieves the SMB2 I/O latencies per bucket size (default 1 sec).
  - | • **smb2IORate**: Retrieves the SMB2 I/O rate in number of operations per bucket size (default 1 sec).
  - | • **smb2Throughput**: Retrieves the SMB2 Throughput in bytes per bucket size (default 1 sec).
  - | • **smb2Writes** : Retrieves count, # of idle calls, bytes in and out, and operation time for SMB2 writes.
  - | • **smbConnections**: - Retrieves the number of SMB connections.

## | **CTDB**

- | These queries retrieve metrics associated with CTDB.
- | • **ctdbCallLatency**: Retrieves information on the CTDB call latency.
  - | • **ctdbHopCountDetails**: Retrieves information on the CTDB hop count buckets 0 to 5 for one database.
  - | • **ctdbHopCounts** :Retrieves information on the CTDB hop counts (bucket 00 = 1-3 hops) for all databases.

---

## **Performance monitoring through IBM Spectrum Scale GUI**

The IBM Spectrum Scale GUI provides a graphical representation of the status and historical trends of the key performance indicators. This helps the users to make decisions easily without wasting time. Use **Monitoring > Dashboard** and **Monitoring > Statistics** pages in the GUI to monitor the performance of the system based on various aspects.

The **Statistics** page is primarily used for selecting the attributes based on which the performance of the system needs to be monitored and also for monitoring and comparing the performance. Whereas, the **Dashboard** page is just used for monitoring the performance data collected at the GUI server.

You can configure the system to monitor the performance of the following functional areas in the system:

- Network
- System resources
- NSD server
- IBM Spectrum Scale client
- NFS
- SMB
- Object
- CTDB
- Transparent cloud tiering. This option is available only when the cluster is configured to work with the transparent cloud tiering service.

The performance data is collected with the help of the following two components:

- **Sensor:** The sensors are placed on all the nodes and they share the data with the collector. The sensors run on any node that is required to collect metrics. Sensors are started by default only on the protocol nodes.
- **Collector:** Collects data from the sensors. The metric collector runs on a single node and gathers metrics from all the nodes that are running the associated sensors. The metrics are stored in a database on the collector node. The collector ensures aggregation of data once data gets older. The collector can run on any node in the system. By default, the collector runs on the management node, which is usually the installation node. To configure performance monitoring through GUI, it is mandatory to configure collector on each GUI node.

The following picture provides a graphical representation of the performance monitoring configuration for GUI.

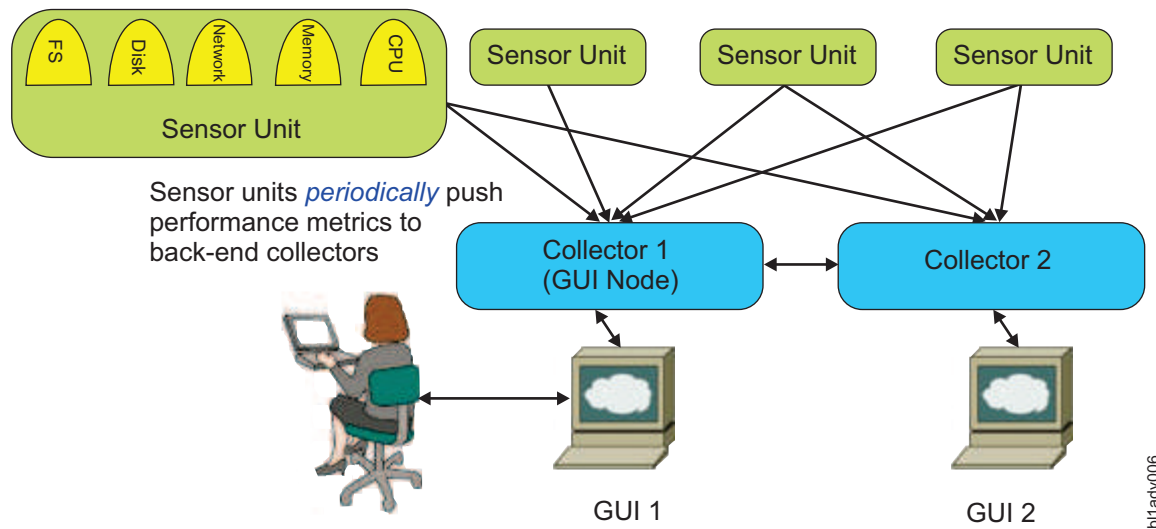


Figure 2. Performance monitoring configuration for GUI

The `mmperfmon` command can be used to query performance data through CLI, and configure the performance data collection. The GUI displays a subset of the available metrics.

## Configuring performance monitoring options in GUI

You need to configure and enable the performance monitoring for GUI to view the performance data in the GUI.

### Enabling performance tools in management GUI

You need to enable performance tools in the management GUI to display performance data in the management GUI. For more information on how to enable performance tools in GUI, see *Enabling performance tools in management GUI* section in the *IBM Spectrum Scale: Administration Guide*.

### Verifying sensor and collector configurations

Do the following to verify whether collectors are working properly:

1. Issue `systemctl status pmcollector` on the GUI node to confirm that the collector is running. Start collector if it is not started already.
2. If you cannot start the service, verify the log file that is located at the following location to fix the issue: `/var/log/zimon/ZIMonCollector.log`.
3. Use a sample CLI query to test if data collection works properly. For example:

```
mmpfmon query cpu_user
```

Do the following to verify whether sensors are working properly:

1. Confirm that the sensor is configured correctly by issuing the **mmpfmon config show** command. This command lists the content of the sensor configuration that is located at the following location: `/opt/IBM/zimon/ZIMonSensors.cfg`. The configuration must point to the node where the collector is running and all the expected sensors must be enabled. An enabled sensor has a period greater than 0 in the same config file.
2. After updating the configuration file, issue **systemctl restart pmsensors** command to restart the sensors:
3. Issue **systemctl status pmsensors** to verify the status of the sensors.

## Configuring performance metrics and display options in the Statistics page of the GUI

Use the **Monitoring > Statistics** page to monitor the performance of system resources and file and object storage. Performance of the system can be monitored by using various pre-defined charts. You can select the required charts and monitor the performance based on the filter criteria.

The pre-defined performance charts and metrics help in investigating every node or any particular node that is collecting the metrics. The following figure shows various configuration options that are available in the Statistics page of the management GUI.

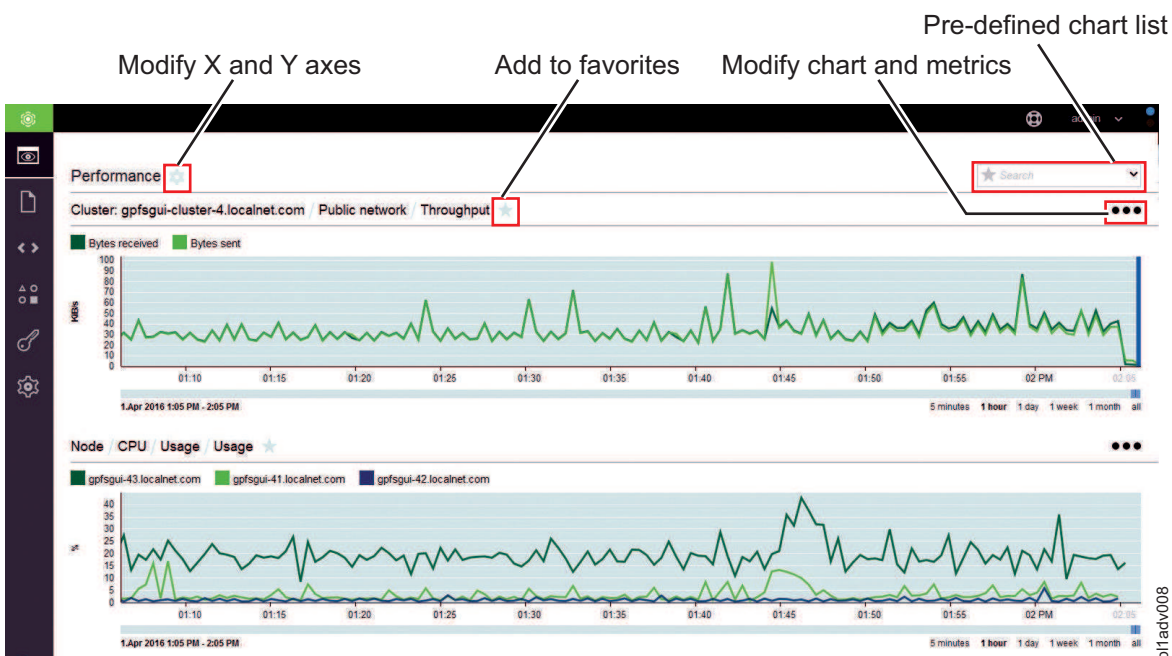


Figure 3. Statistics page in the IBM Spectrum Scale management GUI

You can select pre-defined charts that are available for selection from pre-defined chart list. You can display up to two charts at a time.

## Display options in performance charts

The charting section displays the performance details based on various aspects. The GUI provides a rich set of controls to view performance charts. You can use these controls to perform the following actions on the charts that are displayed in the page:

- Zoom the chart by using the mouse wheel or resizing the timeline control. Y-axis can be automatically adjusted during zooming.
- Click and drag the chart or the timeline control at the bottom. Y-axis can be automatically adjusted during panning.
- Compare charts side by side. You can synchronize y-axis and bind x-axis in the dual chart mode. If only one chart is displayed in the page, click the configuration icon that is placed in the upper right corner of the performance chart and select **Dual chart mode** to switch to the dual chart mode. To modify the x and y axes of the chart in the dual chart mode, click the configuration symbol next to the title *Performance* and select the required options.
- Link the timelines of the two charts together by using the display options that are available.
- The Dashboard helps to access all single graph charts, which are either predefined or custom created favorites.

## Selecting performance and capacity metrics

To monitor the performance of the system, you need to select the appropriate metrics to be displayed in the performance charts. Metrics are grouped under the combination of resource types and aggregation levels. The resource types determine the area from which the data is taken to create the performance analysis and aggregation level determines the level at which the data is aggregated. The aggregation levels that are available for selection varies based on the resource type.

Sensors are configured against each resource type. The following table provides a mapping between resource types and sensors under the Performance category.

*Table 26. Sensors available for each resource type*

Resource type	Sensor name	Candidate nodes
Network	Network	All
System Resources	CPU	All
	Load	
	Memory	
NSD Server	GPFSNSDDisk	NSD Server nodes
IBM Spectrum Scale Client	GPFSFilesystem	IBM Spectrum Scale Client nodes
	GPFSVFS	
	GPFSFilesystemAPI	
NFS	NFSIO	Protocol nodes running NFS service
SMB	SMBStats	Protocol nodes running SMB service
	SMBGlobalStats	
Waiters	GPFSWaiters	All nodes
CTDB	CTDBStats	Protocol nodes running SMB service
Object	SwiftAccount	Protocol nodes running Object service
	SwiftContainer	
	SwiftObject	
	SwiftProxy	
Transparent Cloud Tiering	MCStoreGPFSStats	Cloud gateway nodes
	MCStoreIcstoreStats	
	MCStoreLWESStats	



The resource type *Waiters* are used to monitor the long running file system threads. Waiters are characterized by the purpose of the corresponding file system threads. For example, an RPC call waiter that is waiting for Network I/O threads or a waiter that is waiting for a local disk I/O file system operation. Each waiter has a wait time associated with it and it defines how long the waiter is already waiting. With some exceptions, long waiters typically indicate that something in the system is not healthy.

The *Waiters* performance chart shows the aggregation of the total count of waiters of all nodes in the cluster above a certain threshold. Different thresholds from 100 milliseconds to 60 seconds can be selected in the list below the aggregation level. By default, the value shown in the graph is the sum of the number of waiters that exceed threshold in all nodes of the cluster at that point in time. The filter functionality can be used to display waiters data only for some selected nodes or file systems. Furthermore, there are separate metrics for different waiter types such as Local Disk I/O, Network I/O, ThCond, ThMutex, Delay, and Syscall.

You can also monitor the capacity details that are aggregated at the following levels:

- NSD
- Node
- File system
- Pool
- Fileset
- Cluster

The following table lists the sensors that are used for capturing the capacity details.

*Table 27. Sensors available to capture capacity details*

Sensor name	Candidate nodes
DiskFree	All nodes
GPFSSetQuota	Only a single node
GPFSDiskCap	Only a single node

You can edit an existing chart by clicking the icon that is available on the upper right corner of the performance chart and select Edit to modify the metrics selections. Do the following to drill down to the metric you are interested in:

1. Select **Resource type**. This is the area from which the data is taken to create the performance analysis.
2. Select **Aggregation level**. The aggregation level determines the level at which the data is aggregated.
3. Select the entities that need to be graphed. The table lists all entities that are available for the chosen resource type and aggregation level. When a metric is selected, you can also see the selected metrics in the same grid and use methods like sorting, filtering, or adjusting the time frame to select the entities that you want to select.
4. Select **Metrics**. Metrics is the type of data that needs to be included in the performance chart. The aggregation levels that are available for selection varies based on the resource type.
5. Use the filter option to further narrow down in addition to the objects and metrics selection by using filters. Depending on the selected object category and aggregation level, the "Filter" section can be displayed underneath the aggregation level, allowing one or more filters to be set. Filters are specified as regular expressions as shown in the following examples:
  - As a single entity:
    - node1*
    - eth0*
  - Filter metrics applicable to multiple nodes as shown in the following examples:

- To select a range of nodes such as node1, node2 and node3:  
*node1 | node2 | node3*  
*node[1-3]*
- To filter based on a string of text. For example, all nodes starting with 'nod' or ending with 'int':  
*nod.+ | .+int*
- To filter network interfaces eth0 through eth6, bond0 and eno0 through eno6:  
*eth[0-6] | bond0 | eno[0-6]*
- To filter nodes starting with 'strg' or 'int' and ending with 'nx':  
*(strg) | (int).+nx*

## Creating favorite charts

Favorite charts are nothing but customized predefined charts. A favorite chart with 'single chart' configuration can be displayed in the Dashboards page. Favorite charts along with the predefined charts are available for selection when you add widgets in the dashboard.

To create favorite charts, click the 'star' symbol that is placed next to the chart title and enter the label and select whether you need to display the favorite chart in single chart or dual chart mode.

## Configuring the dashboard to view performance charts

The **Monitoring > Dashboard** page provides an easy to read, single page, and real-time user interface that provides a quick overview of the system performance.

The dashboard consists of several dashboard widgets and the associated favorite charts that can be displayed within a chosen layout. Currently, the following important widget types are available in the dashboard:

- Performance
- File system capacity by fileset
- System health events
- System overview
- Filesets with the largest growth rate in last week
- Timeline

The following picture highlights the configuration options that are available in the edit mode of the dashboard.

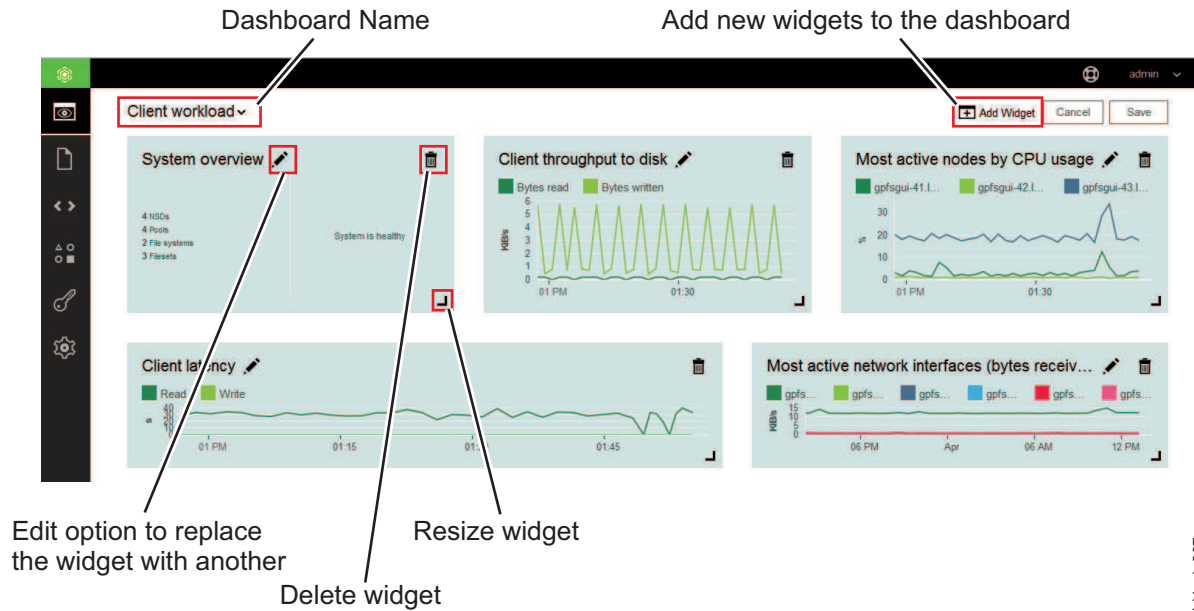


Figure 4. Dashboard page in the edit mode

## Layout options

The highly customizable dashboard layout options helps to add or remove widgets and change its display options. Select **Layout Options** option from the menu that is available in the upper right corner of the Dashboard GUI page to change the layout options. While selecting the layout options, you can either select the basic layouts that are available for selection or create a new layout by selecting an empty layout as the starting point.

You can also save the dashboard so that it can be used by other users. Select **Create Dashboard** and **Delete Dashboard** options from the menu that is available in the upper right corner of the Dashboard page to create and delete dashboards respectively. If several GUIs are running by using CCR, saved dashboards are available on all nodes.

When you open the IBM Spectrum Scale GUI after the installation or upgrade, you can see the default dashboards that are shipped with the product. You can further modify or delete the default dashboards to suit your requirements.

## Widget options

Several dashboard widgets can be added in the selected dashboard layout. Select **Edit Widgets** option from the menu that is available in the upper right corner of the Dashboard GUI page to edit or remove widgets in the dashboard. You can also modify the size of the widget in the edit mode. Use the **Add Widget** option that is available in the edit mode to add widgets in the dashboard.

The widgets with type *Performance* lists the charts that are marked as favorite charts with 'single chart' configuration in the Performance page of the GUI. Favorite charts along with the predefined charts are available for selection when you add widgets in the dashboard.

To create favorite charts, click the 'star' symbol that is placed next to the chart title in the **Monitoring > Performance** page and enter the label and select **Single chart** as the chart configuration.

bl1adv007

## Querying performance data shown in the GUI through CLI

You can query the performance data that is displayed in the GUI through the CLI. This is usually used for external system integration or to troubleshoot any issues with the performance data displayed in the GUI.

The following example shows how to query the performance data through CLI:

```
# mmpfmon query "sum(netdev_bytes_r)"
```

This query displays the following output:

Legend:

```
1: mr-31.localnet.com|Network|eth0|netdev_bytes_r
2: mr-31.localnet.com|Network|eth1|netdev_bytes_r
3: mr-31.localnet.com|Network|lo|netdev_bytes_r
```

Row	Timestamp	netdev_bytes_r	netdev_bytes_r	netdev_bytes_r
1	2016-03-15-14:52:09	10024		
2	2016-03-15-14:52:10	9456		
3	2016-03-15-14:52:11	9456		
4	2016-03-15-14:52:12	9456		
5	2016-03-15-14:52:13	9456		
6	2016-03-15-14:52:14	9456		
7	2016-03-15-14:52:15	27320		
8	2016-03-15-14:52:16	9456		
9	2016-03-15-14:52:17	9456		
10	2016-03-15-14:52:18	11387		

The sensor gets the performance data for the collector and the collector passes it to the performance monitoring tool to display it in the CLI and GUI. If sensors and collectors are not enabled in the system, the system does not display the performance data and when you try to query data from a system resource, it returns an error message. For example, if performance monitoring tools are not configured properly for the resource type *Transparent Cloud Tiering*, the system displays the following output while querying the performance data:

```
mmpfmon query "sum(mcs_total_requests)" number_buckets 1
Error: No data available for query: 3169
```

mmpfmon: Command failed. Examine previous error messages to determine cause.

For more information on how to troubleshoot the performance data issues, see Chapter 20, “Performance issues,” on page 327.

## Monitoring performance of nodes

The **Monitoring > Nodes** page provides an easy way to monitor the performance, health status, and configuration aspects of all available nodes in the IBM Spectrum Scale cluster.

The Nodes page provides the following options to analyze performance of nodes:

1. A quick view that gives the number of nodes in the system, and the overall performance of nodes based on CPU and memory usages.

You can access this view by selecting the expand button that is placed next to the title of the page. You can close this view if not required.

The graphs in the overview show the nodes that have the highest average performance metric over a past period. These graphs are refreshed regularly.

2. A nodes table that displays many different performance metrics.

To find nodes with extreme values, you can sort the values displayed in the nodes table by different performance metrics. Click the performance metric in the table header to sort the data based on that metric.

You can select the time range that determines the averaging of the values that are displayed in the table and the time range of the charts in the overview from the time range selector, which is placed in the upper right corner. The metrics in the table do not update automatically. The refresh button above the table allows to refresh the table content with more recent data.

You can group the nodes to be monitored based on the following criteria:

- All nodes
  - NSD server nodes
  - Protocol nodes
3. A detailed view of the performance and health aspects of individual nodes that are listed in the Nodes page.

Select the node for which you need to view the performance details and select **View Details**. The system displays various performance charts on the right pane.

The detailed performance view helps to drill-down to various performance aspects. The following list provides the performance details that can be obtained from each tab of the performance view:

- **Overview** tab provides performance chart for the following:
  - Client IOPS
  - Client data rate
  - Server data rate
  - Server IOPS
  - Network
  - CPU
  - Load
  - Memory
- **Events** tab helps to monitor the events that are reported in the node. Three filter options are available to filter the events by their status; such as **Current Issues**, **Unread Messages**, and **All Events** displays every event, no matter if it is fixed or marked as read. Similar to the Events page, you can also perform the operations like marking events as read and running fix procedure from this events view.
- **File Systems** tab provides performance details of the file systems mounted on the node. You can view the file system read or write throughput, average read or write transactions size, and file system read or write latency.
- **NSDs** tab gives status of the disks that are attached to the node. The NSD tab appears only if the node is configured as an NSD server.
- **SMB** and **NFS** tabs provide the performance details of the SMB and NFS services hosted on the node. These tabs appear in the chart only if the node is configured as a protocol node.
- **Network** tab displays the network performance details.

---

## Performance monitoring limitations

The following section lists the limitations of the performance monitoring tool:

- Performance monitoring is not supported by the following operating systems:
  - x86\_64/sles11
  - ppc64le/Ubuntu14.04
  - ppc64/aix
- If federation is used for performance monitoring, the collectors participating in a federation must be of the same version number.



---

## Chapter 2. Monitoring system health through the IBM Spectrum Scale GUI

Use **Monitoring > Events** page to monitor the events that are reported in the system. The Events page displays events and you can monitor and troubleshoot errors on your system.

There are three options to filter events by their status:

- **Current Issues** displays all unfixed errors and warnings.
- **Unread Messages** displays all unfixed errors and warnings and information messages that are not marked as read.
- **All Events** displays every event, no matter if it is fixed or marked as read.

The status icons help to quickly determine whether the event is informational, a warning, or an error. Click an event and select **Properties** from the **Action** menu to see detailed information on the event. The event table displays the most recent events first.

### Marking events as Read

You can mark certain events as read to change the status of the event in the events view. The status icons become gray in case an error or warning is fixed or if it is marked as read.

There are events on states that start with "MS\*". These events can be errors, warnings, or information messages that cannot be marked as read and these events automatically changes the status from current to historic when the problem is resolved or information condition changes. The user must either fix the problem or change the state of some component to make the current event a historical event. There are also message events that start with MM\*. These events never become historic by itself. The user must use the action **Mark as Read** on those events to make them historical because the system cannot detect itself even if the problem or information is not valid anymore.

### Running fix procedure

Some issues can be resolved by running a fix procedure. Use action **Run Fix Procedure** to do so. The Events page provides a recommendation for which fix procedure to run next.

For more information on how to set up event notifications, see "Set up event notifications."

---

## Set up event notifications

The system can use Simple Network Management Protocol (SNMP) traps and emails to notify you when significant events are detected. Any combination of these notification methods can be used simultaneously. Use **Settings > Event Notifications** page in the GUI to configure event notifications.

Notifications are normally sent immediately after an event is raised.

In email notification method, you can also define whether a recipient needs to get a report of events that are reported in the system. These reports are sent only once in a day. Based on the seriousness of the issue, each event that is reported in the system gets a severity level associated with it.

The following table describes the severity levels of event notifications.

Table 28. Notification levels

Notification level	Description
Critical	<p>Critical notification is sent to indicate a problem that must be corrected as soon as possible.</p> <p>This notification indicates a serious problem with the system. For example, the event that is being reported might indicate a loss of redundancy in the system, and it is possible that another failure might result in loss of access to data. The most typical reason that this type of notification is because of a hardware failure, but some configuration errors or fabric errors also are included in this notification level.</p>
Warning	<p>A warning notification is sent to indicate a problem or unexpected condition with the system. Always immediately investigate this type of notification to determine the effect that it might have on your operation, and make any necessary corrections.</p> <p>Therefore, a warning notification does not require any replacement parts and it does not require IBM Support Center involvement.</p>
Information	<p>An informational notification is sent to indicate that an expected event is occurred. For example, a NAS service is started. No remedial action is required when these notifications are sent.</p>

## Configuring email notifications

The email feature transmits operational and error-related data in the form of an event notification email.

To configure an email server, from the Event Notifications page, select **Email Server**. Select **Edit** and then click **Enable email notifications**. Enter required details and when you are ready, click **OK**.

Emails are sent to the recipients to intimate them about the quota reports or if there are issues or events in the clustered file system.

To create email recipients, select **Email Recipients** from the **Event Notifications** page, and then click **Create Recipient**.

**Note:** You can change the email notification configuration or disable the email service at any time.

## Configuring SNMP manager

Simple Network Management Protocol (SNMP) is a standard protocol for managing networks and exchanging messages. The system can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that the system sends.

With an SNMP manager, such as IBM Systems Director, you can view, and act on the messages that the SNMP agent sends. The SNMP manager can send SNMP notificationS, which are also known as traps, when an event occurs in the system. Select **Settings > Event Notifications > SNMP Manager** to configure SNMP managers for event notifications. You can specify up to a maximum of six SNMP managers.

In the SNMP mode of event notification, one SNMP notification (trap) with object identifiers (OID) .1.3.6.1.4.1.2.6.212.10.0.1 is sent by the GUI for each event. The following table provides the SNMP objects included in the event notifications.

Table 29. SNMP objects included in event notifications

OID	Description	Examples
.1.3.6.1.4.1.2.6.212.10.1.1	Cluster ID	317908494245422510
.1.3.6.1.4.1.2.6.212.10.1.2	Entity type	SERVER, FILESYSTEM



Table 29. SNMP objects included in event notifications (continued)

OID	Description	Examples
.1.3.6.1.4.1.2.6.212.10.1.3	Entity name	gss-11, fs01
.1.3.6.1.4.1.2.6.212.10.1.4	Component	SMB, AUTHENTICATION
.1.3.6.1.4.1.2.6.212.10.1.5	Log level	SEVERE, WARN, INFO
.1.3.6.1.4.1.2.6.212.10.1.6	Date and time	17.02.2016 13:27:42.516
.1.3.6.1.4.1.2.6.212.10.1.7	Event ID	MS1014
.1.3.6.1.4.1.2.6.212.10.1.8	Description	At least one CPU of "gss-11" is failed.

**Note:** The Component field in the SNMP traps contains the value "UNKNOWN" now. This will be resolved in a later release of IBM Spectrum Scale.

## Understanding the SNMP OID ranges

The following table gives the description of the SNMP OID ranges.

Table 30. SNMP OID ranges

OID range	Description
.1.3.6.1.4.1.2.6.212	IBM Spectrum Scale
.1.3.6.1.4.1.2.6.212.10	IBM Spectrum Scale GUI
.1.3.6.1.4.1.2.6.212.10.0.1	IBM Spectrum Scale GUI event notification (trap)
.1.3.6.1.4.1.2.6.212.10.1.x	IBM Spectrum Scale GUI event notification parameters (objects)

The traps for the core IBM Spectrum Scale and those trap objects are not included in the SNMP notifications that are configured through the IBM Spectrum Scale management GUI. For more information on SNMP traps from the core IBM Spectrum Scale, see Chapter 6, "GPFS SNMP support," on page 105

## Example for SNMP traps

The following example shows the SNMP event notification that is sent when CPU of a server is failed:

```
2016-02-18 12:46:50 localhost [UDP: [127.0.0.1]:54780->[127.0.0.1]:162]:
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (1) 0:00:00.01
SNMPv2-MIB::snmpTrapOID.0 = OID: SNMPv2-SMI::enterprises.2.6.212.10.0.1
SNMPv2-SMI::enterprises.2.6.212.10.1.1 = STRING: "317908494245422510"
SNMPv2-SMI::enterprises.2.6.212.10.1.2 = STRING: "SERVER"
SNMPv2-SMI::enterprises.2.6.212.10.1.3 = STRING: "gss-11"
SNMPv2-SMI::enterprises.2.6.212.10.1.4 = STRING: "UNKNOWN"
SNMPv2-SMI::enterprises.2.6.212.10.1.5 = STRING: "SEVERE"
SNMPv2-SMI::enterprises.2.6.212.10.1.6 = STRING: "18.02.2016 12:46:44.839"
SNMPv2-SMI::enterprises.2.6.212.10.1.7 = STRING: "MS1014"
SNMPv2-SMI::enterprises.2.6.212.10.1.8 = STRING: "At least one CPU of \"gss-11\" failed."
```

The following example shows the SNMP event notification that is sent for an SNMP test message:

```
2016-02-18 12:47:10 localhost [UDP: [127.0.0.1]:45933->[127.0.0.1]:162]:
DISMAN-EVENT-MIB::sysUpTimeInstance = Timeticks: (2062) 0:00:20.62
SNMPv2-MIB::snmpTrapOID.0 = OID: SNMPv2-SMI::enterprises.2.6.212.10.0.1
SNMPv2-SMI::enterprises.2.6.212.10.1.1 = STRING: "317908494245422510"
SNMPv2-SMI::enterprises.2.6.212.10.1.2 = STRING: "CLUSTER"
SNMPv2-SMI::enterprises.2.6.212.10.1.3 = STRING: "UNKNOWN"
SNMPv2-SMI::enterprises.2.6.212.10.1.4 = STRING: "UNKNOWN"
SNMPv2-SMI::enterprises.2.6.212.10.1.5 = STRING: "INFO"
```

```
SNMPv2-SMI::enterprises.2.6.212.10.1.6 = STRING: "18.02.2016 12:47:10.851"
SNMPv2-SMI::enterprises.2.6.212.10.1.7 = STRING: "MM0025"
SNMPv2-SMI::enterprises.2.6.212.10.1.8 = STRING: "This a SNMP test message."
```

## SNMP MIBs

The SNMP Management Information Base (MIB) is a collection of definitions that define the properties of the managed objects.

The IBM Spectrum Scale GUI MIB OID range starts with 1.3.6.1.4.1.2.6.212.10. The OID range 1.3.6.1.4.1.2.6.212.10.0.1 denotes IBM Spectrum Scale GUI event notification (trap) and .1.3.6.1.4.1.2.6.212.10.1.x denotes IBM Spectrum Scale GUI event notification parameters (objects). Use the following text to configure IBM Spectrum Scale GUI MIB:

```
IBM-SPECTRUM-SCALE-GUI-MIB DEFINITIONS ::= BEGIN
IMPORTS
MODULE-IDENTITY,
OBJECT-TYPE,
NOTIFICATION-TYPE,
Counter64,
enterprises
FROM SNMPv2-SMI
DisplayString
FROM RFC1213-MIB;
ibmSpectrumScaleGUI MODULE-IDENTITY
LAST-UPDATED "201602170000Z" -- February 17, 2016
ORGANIZATION "International Business Machines Corp."
CONTACT-INFO ""
DESCRIPTION "Definition of Specrum Scale GUI Notifications for Spectrum Scale product.
These objects are subject to modification by IBM as product specifications require."
-- Revision log, in reverse chronological order
REVISION "201602170000Z" -- February 17, 2016
DESCRIPTION "Version 0.1."
::= { ibmGPFS 10 }
-- ibmGPFS is copied from GPFS MIB (/usr/lpp/mmfs/data/GPFS-MIB.txt)
ibm OBJECT IDENTIFIER ::= { enterprises 2 }
ibmProd OBJECT IDENTIFIER ::= { ibm 6 }
ibmGPFS OBJECT IDENTIFIER ::= { ibmProd 212 }
ibmSpectrumScaleGuiNotification OBJECT IDENTIFIER ::= { ibmSpectrumScaleGUI 0}
ibmSpectrumScaleGuiEventObject OBJECT IDENTIFIER ::= { ibmSpectrumScaleGUI 1}
--- *****
-- IBM Spectrum Scale GUI Scalar object declarations - accessible for notifications
--- *****
ibmSpectrumScaleGuiEventCluster OBJECT-TYPE
SYNTAX DisplayString (SIZE (0..255))
MAX-ACCESS accessible-for-notify
STATUS current
DESCRIPTION
"The cluster where the notification occurred."
::= { ibmSpectrumScaleGuiEventObject 1 }
ibmSpectrumScaleGuiEventEntityType OBJECT-TYPE
SYNTAX DisplayString (SIZE (0..255))
MAX-ACCESS accessible-for-notify
STATUS current
DESCRIPTION
"The type of entity for which the notification occurred."
::= { ibmSpectrumScaleGuiEventObject 2 }
ibmSpectrumScaleGuiEventEntityName OBJECT-TYPE
SYNTAX DisplayString (SIZE (0..255))
MAX-ACCESS accessible-for-notify
STATUS current
DESCRIPTION
"The name of the entity for which the notification occurred."
::= { ibmSpectrumScaleGuiEventObject 3 }
ibmSpectrumScaleGuiEventComponent OBJECT-TYPE
SYNTAX DisplayString (SIZE (0..255))
```

```

MAX-ACCESS accessible-for-notify
STATUS current
DESCRIPTION
"The component for which the notification occurred."
 ::= { ibmSpectrumScaleGuiEventObject 4 }
ibmSpectrumScaleGuiEventLogLevel OBJECT-TYPE
SYNTAX DisplayString (SIZE (0..255))
MAX-ACCESS accessible-for-notify
STATUS current
DESCRIPTION
"The log level."
 ::= { ibmSpectrumScaleGuiEventObject 5 }
ibmSpectrumScaleGuiEventTime OBJECT-TYPE
SYNTAX DisplayString (SIZE (0..255))
MAX-ACCESS accessible-for-notify
STATUS current
DESCRIPTION
"A representation of the date and time when the notification occurred."
 ::= { ibmSpectrumScaleGuiEventObject 6 }
ibmSpectrumScaleGuiEventMessageId OBJECT-TYPE
SYNTAX DisplayString (SIZE (0..255))
MAX-ACCESS accessible-for-notify
STATUS current
DESCRIPTION
"The message ID of the notification."
 ::= { ibmSpectrumScaleGuiEventObject 7 }
ibmSpectrumScaleGuiEventMessageText OBJECT-TYPE
SYNTAX DisplayString (SIZE (0..1492))
MAX-ACCESS accessible-for-notify
STATUS current
DESCRIPTION
"The human readable message of the notification."
 ::= { ibmSpectrumScaleGuiEventObject 8 }
ibmSpectrumScaleGuiNotificationEvent NOTIFICATION-TYPE
OBJECTS {
  ibmSpectrumScaleGuiEventCluster,
  ibmSpectrumScaleGuiEventEntityType,
  ibmSpectrumScaleGuiEventEntityName,
  ibmSpectrumScaleGuiEventComponent,
  ibmSpectrumScaleGuiEventLogLevel,
  ibmSpectrumScaleGuiEventTime,
  ibmSpectrumScaleGuiEventMessageId,
  ibmSpectrumScaleGuiEventMessageText
}
STATUS current
DESCRIPTION
"This notification indicates a Health event as reported by the Spectrum Scale GUI."
 ::= { ibmSpectrumScaleGuiNotification 1 }
END

```

### Related concepts:

Chapter 6, "GPFS SNMP support," on page 105

GPFS supports the use of the SNMP protocol for monitoring the status and configuration of the GPFS cluster. Using an SNMP application, the system administrator can get a detailed view of the system and be instantly notified of important events, such as a node or disk failure.



---

## Chapter 3. Monitoring system health by using the mmhealth command

The **mmhealth** command helps to monitor health status of nodes and services hosted on the node. You can also use the **mmhealth** command to get a list of events that are reported in the system. Based on the details that are specified in these events, the system administrator can perform the corresponding troubleshooting steps to resolve the issue.

The following are the possible status of nodes and services:

- Unknown - Status of the node or the service hosted on the node is not known
- Healthy - The node or the service hosted on the node is working as expected. There are no active error events.
- Checking - The monitoring of a service or a component hosted on the node is starting at the moment. This state is a transient state and is updated when the startup is completed.
- Degraded - The node or the service hosted on the node is not working as expected. That is, a problem occurred with the component but it did not result in a complete failure.
- Failed - The node or the service hosted on the node failed due to errors or cannot be reached anymore.
- Dependent\_failed - The node or the services hosted on the node are failed due of the failure of some sub-components. For example, NFS or SMB service shows this status if authentication is failed.

The following two types of events are reported in the system:

- State-changing events
- Information events

The events are raised at an interval of 15 seconds and you can view them either through **Monitoring > Events** page in the IBM Spectrum Scale GUI or by issuing the **mmhealth node eventlog** command.

The status of the services hosted on a node becomes FAILED if there is at least one active event in the FAILED status for that corresponding service. The FAILED status gets more priority than the DEGRADED status while setting the status of the service. That is, if a service has an active event with DEGRADED status and another active with FAILED status, then the system sets the status of the service as FAILED.

### System health monitoring framework

Every service hosted on the node of IBM Spectrum Scale has its own health monitoring service. All the subcomponents like file systems or network interfaces are monitored through the monitoring service of their main component. Only the subcomponents of CES service such as NFS, SMB, Object, and authentication are having their own health monitors. The **mmhealth** command gets the health details from these monitoring services. The role of a node in monitoring determines the components that need to be monitored. The node roles are hardcoded. It also determines the monitoring service that is required on a specific node. For example, you do not need a CES monitoring on a non-CES node. The monitoring services are only started if the specific node role is assigned to the node. Every monitoring service includes at least one monitor.

The following list provides the details of the monitoring services available in the IBM Spectrum Scale system:

#### 1. GPFS

- Node role: GPFS daemon must be running on every node that has the **mmhealth** monitoring framework.

- Tasks: Monitors all GPFS daemon-related functionalities. For example, *mmfsd* process and *gpfs port accessibility*.
2. Network
    - Node role: Every IBM Spectrum Scale node has the node role accompanying the NETWORK monitor service.
  3. CES
    - Node role: Nodes with node class *cesNode* performs the monitoring services for all CES services. The CES service does not have its own monitoring service or events. The status of the CES is an aggregation of the status of the subservices.
  4. AUTH
    - Tasks: Monitors LDAP, AD and or NIS-based authentication services.
  5. AUTH\_OBJ
    - Tasks: Monitoring the *OpenStack* identity service functionalities.
  6. BLOCK
    - Tasks: Checks whether the iSCSI daemon is functioning properly.
  7. CESNETWORK
    - Tasks: Monitoring CES network-related adapters and IP addresses.
  8. NFS
    - Tasks: Monitoring NFS-related functionalities.
  9. OBJECT
    - Tasks: Monitors the IBM Spectrum Scale for object functionalities. Especially, the status of relevant system services and accessibilities to ports are checked.
  10. SMB
    - Tasks: Monitoring SMB-related functionalities like the *smbd* process, the ports and *ctdb* processes.
  11. CLOUDGATEWAY
    - Node role: A node gets the cloud gateway node role if it is identified as a clog gateway node. These nodes are listed when you issue the **mmcloudgateway nodelist** command.
    - Tasks: Check if the cloud gateway service functions as expected.
  12. DISK
    - Node role: This node role is always active on all IBM Spectrum Scale nodes.
    - Tasks: Checking, if IBM Spectrum Scale disks are available and running.
  13. FILESYSTEM
    - Node role: This node role is active on all IBM Spectrum Scale nodes.
    - Tasks: Monitors different aspects of IBM Spectrum Scale file systems
  14. GUI
    - Node role: Nodes with node class *GUI\_MGMT\_SERVERS* can monitor the GUI service.
    - Tasks: Verifies whether the GUI services are functioning properly.
  15. HADOOPCONNECTOR
    - Node role: Nodes where the Hadoop service is configured get the Hadoop connector node role.
    - Tasks: Monitors the Hadoop data node and name node services.
  16. PERFMON
    - Node role: Nodes where *PerfmonSensors* or *PerfmonCollectorservices* are installed get the *perfmon* node role. *PerfmonSensors* are determined through the *perfmon* designation in **mm1scluster**. *PerfmonCollector* are determined through the *colCandidates* line in the configuration file.
    - Tasks: Monitors whether *PerfmonSensors* and *PerfmonSensors* are running as expected.

| For more details on different events, their causes and possible user actions to resolves them, see “Events”  
| on page 371.

### | **Requirements/Prerequisites:**

| The following criteria must be met to use the health monitoring functionality on your GPFS cluster:

- | • Only Linux and AIX nodes are supported.
- | • Only GPFS monitoring is supported on AIX.
- | • The AIX nodes must have the Python 2.7.5 installed.

### | **Examples**

| You can check the health status of the nodes and their corresponding services by using the following  
| commands:

- | 1. To show the health status of the current node:

```
| mmhealth node show
```

| The system displays output similar to this:

```
| Node name:      test_node  
| Node status:   DEGRADED  
|  
| Component      Status          Reasons  
| -----  
| GPFS           HEALTHY        -  
| CES            FAILED         smbd_down  
| FileSystem     HEALTHY        -
```

- | 2. To view the health status of a specific node, issue this command:

```
| mmhealth node show -N node_test2
```

| The system displays output similar to this:

```
| Node name:      test_node2  
| Node status:   HEALTHY  
|  
| Component      Status          Reasons  
| -----  
| GPFS           HEALTHY        -  
| CES            HEALTHY        -  
| FileSystem     HEALTHY        -
```

- | 3. To view the health status of all the nodes, issue this command:

```
| mmhealth node show -N all
```

| The system displays output similar to this:

```
| Node name:      test_node  
| Node status:   DEGRADED  
|  
| Component      Status          Reasons  
| -----  
| GPFS           HEALTHY        -  
| CES            FAILED         smbd_down  
| FileSystem     HEALTHY        -  
|  
| Node name:      test_node2  
| Node status:   HEALTHY  
|  
| Component      Status          Reasons
```

```

| -----
| GPFS                HEALTHY    -
| CES                 HEALTHY    -
| FileSystem          HEALTHY    -

```

4. To view the detailed health status of the component and its subcomponents, issue this command:

```
mmhealth node show ces
```

The system displays output similar to this:

```

| Node name:          test_node
|
| Component           Status      Reasons
| -----
| CES                 FAILED    smbd_down
|   AUTH              HEALTHY   -
|   AUTH_OBJ          HEALTHY   -
|   NFS               HEALTHY   -
|   OBJ               HEALTHY   -
|   SMB               FAILED    smbd_down
|
| Event               Parameter  Severity  Description
| -----
| smbd_down           SMB        ERROR      SMBD process not running

```

5. To view the health status of only unhealthy components, issue this command:

```
mmhealth node show --unhealthy
```

The system displays output similar to this:

```

| Node name:          test_node
| Node status:        DEGRADED
|
| Component           Status      Reasons
| -----
| CES                 FAILED    smbd_down

```

6. To view the health status of subcomponents of a node's component, issue this command:

```
mmhealth node show --verbose
```

The system displays output similar to this:

```

| Node name:          test_node
| Node status:        DEGRADED
|
| Component           Status      Reasons
| -----
| GPFS                HEALTHY    -
| CES                 FAILED    smbd_down
|   AUTH              HEALTHY   -
|   AUTH_OBJ          HEALTHY   -
|   NFS               HEALTHY   -
|   OBJ               HEALTHY   -
|   SMB               FAILED    smbd_down
|
| FILESYSTEM          HEALTHY    -
|   gpfs0             HEALTHY   -
|   FSII              HEALTHY   -

```

7. To view the eventlog history of the node for the last hour, issue this command:

```
mmhealth node eventlog --hour
```

The system displays output similar to this:

```

| Timestamp           Event Name  Severity  Details
| 2016-04-07 10:12:24.394569--1:00CEST quorum_warn WARNING  GPFS quorum monitoring returned unknown result
| 2016-04-07 10:12:39.366279--1:00CEST quorum_warn WARNING  GPFS quorum monitoring returned unknown result
| 2016-04-07 10:12:54.356577--1:00CEST quorum_warn WARNING  GPFS quorum monitoring returned unknown result

```



| 8. To view more details of the last hour in the eventlog history, issue this command:

| `mmhealth node eventlog --hour --verbose`

| The system displays output similar to this:

Timestamp	Component	Event Name	Event ID	Severity	Details
2016-04-07 10:12:54.356577--1:00CEST	gpfs	quorum_warn	999291	WARNING	GPFS quorum monitoring returned unknown result
2016-04-07 10:13:09.359602--1:00CEST	gpfs	quorum_warn	999291	WARNING	GPFS quorum monitoring returned unknown result
2016-04-07 10:13:24.425680--1:00CEST	gpfs	quorum_warn	999291	WARNING	GPFS quorum monitoring returned unknown result

| **Related concepts:**

| Chapter 2, “Monitoring system health through the IBM Spectrum Scale GUI,” on page 79

| Use **Monitoring > Events** page to monitor the events that are reported in the system. The Events page displays events and you can monitor and troubleshoot errors on your system.



---

## Chapter 4. Monitoring events through callbacks

You can configure the callback feature to provide notifications when node and cluster events occur. Starting complex or long-running commands, or commands that involve GPFS files, might cause unexpected and undesired results, including loss of file system availability. Use the **mmaddcallback** command to configure the callback feature.

For more information on how to configure and manage callbacks, see the man page of the following commands in *IBM Spectrum Scale: Command and Programming Reference*:

- **mmaddcallback**
- **mmdelcallback**
- **mmiscallback**



---

## Chapter 5. Monitoring AFM

Active File Management (AFM) enables sharing of data across clusters, even if the networks are unreliable or have high latency.

The following sections inform you how to monitor AFM filesets.

---

### Fileset states for AFM

AFM fileset can have different states depending on the mode and queue states.

To view the current cache state, run the

```
mmafctl filesystem getstate
```

command, or the

```
mmafctl filesystem getstate -j cache_fileset
```

command.

See the following table for the explanation of the cache state:

*Table 31. AFM states and their description*

AFM fileset state	Condition	Description	Healthy or Unhealthy	Administrator's action
Inactive	The AFM cache is created	Operations were not initiated on the cache cluster after the last daemon restart.	Healthy	None
FlushOnly	Operations are queued	Operations have not started to flush.	Healthy	This is a temporary state and should move to Active when a write is initiated.
Active	The AFM cache is active	The cache cluster is ready for an operation.	Healthy	None
Dirty	The AFM is active	The pending changes in the cache cluster are not played at the home cluster. This state does not hamper the normal activity.	Healthy	None
Recovery	The cache is accessed after MDS failure	A new gateway is taking over a fileset as MDS after the old MDS failed.	Healthy	None
QueueOnly	The cache is running some operation.	Operations such as recovery, resync, failover are being executed, and operations are being queued and not flushed.	Healthy	This is a temporary state.

Table 31. AFM states and their description (continued)

AFM fileset state	Condition	Description	Healthy or Unhealthy	Administrator's action
Disconnected	MDS cannot connect to the NFS server at the home cluster.	Occurs only in a cache cluster that is created over an NFS export. When parallel I/O is configured, this state shows the connectivity between the MDS and the mapped home server, irrespective of other gateway nodes.	Unhealthy	Correct the errant NFS servers on the home cluster.
Unmounted	The cache that is using NFS has detected a change in the home cluster - sometimes during creation or in the middle of an operation if home exports are meddled with.	<ul style="list-style-type: none"> <li>The home NFS is not accessible</li> <li>The home exports are not exported properly</li> <li>The home export does not exist</li> </ul>	Unhealthy	<ol style="list-style-type: none"> <li>Fix the NFS export issue in the Home setup section and retry for access.</li> <li>Relink the cache cluster if the cache cluster does not recover.</li> </ol> <p>After <b>mountRetryInterval</b> of the MDS, the cache cluster retries connecting with home.</p>
Unmounted	The cache that is using the GPFS protocol detects a change in the home cluster, sometimes during creation or in the middle of an operation.	There are problems accessing the local mount of the remote file system.	Unhealthy	Check remote filesystem mount on the cache cluster and remount if necessary.
Dropped	Recovery failed.	The local file system is full, space is not available on the cache or the primary cluster, or case of a policy failure during recovery.	Unhealthy	Fix the issue and access the fileset to retry recovery.
Dropped	IW Failback failed.	The local file system is full, space is not available on the cache or the primary cluster, or there is a policy failure during recovery.	Unhealthy	Fix the issue and access the fileset to retry failback.

Table 31. AFM states and their description (continued)

AFM fileset state	Condition	Description	Healthy or Unhealthy	Administrator's action
Dropped	A cache with active queue operations is forcibly unlinked.	All queued operations are being de-queued, and the fileset remains in the Dropped state and moves to the Inactive state when the unlinking is complete.	Healthy	This is a temporary state.
Dropped	The old GW node starts functioning properly after a failure	AFM internally performs queue transfers from one gateway to another to handle gateway node failures.	Healthy	The system resolves this state on the next access.
Dropped	Cache creation or in the middle of an operation if the home exports changed.	Export problems at home such as following: <ul style="list-style-type: none"> <li>The home path is not exported on all NFS server nodes that are interacting with the cache clusters.</li> <li>The home cluster is exported after the operations have started on the fileset. Changing fsid on the home cluster after the fileset operations have begun.</li> <li>All home NFS servers do not have the same fsid for the same export path.</li> </ul>	Unhealthy	<ol style="list-style-type: none"> <li>Fix the NFS export issue in the Home setup section and retry for access.</li> <li>Relink the cache cluster if the cache cluster does not recover.</li> </ol> <p>After <b>mountRetryInterval</b> the MDS retries connecting with home cluster.</p>
Dropped	During recovery or normal operation	If gateway queue memory is exceeded, the queue can get dropped. The memory has to be increased to accommodate all requests and bring the queue back to the Active state.	Unhealthy	Increase <b>afmHardMemThreshold</b> .
Expired	The RO cache that is configured to expire	An event that occurs automatically after prolonged disconnection when the cached contents are not accessible.	Unhealthy	Fix the errant NFS servers on the home cluster

Table 31. AFM states and their description (continued)

AFM fileset state	Condition	Description	Healthy or Unhealthy	Administrator's action
NeedsFailback	The IW cache that needs to complete failback	A failback initiated on an IW cache cluster is interrupted and is incomplete.	Unhealthy	Failback is automatically triggered on the fileset, or the administrator can run failback again.
FailbackInProgress	Failback initiated on IW cache	Failback is in progress and automatically moves to failbackCompleted	Healthy	None
FailbackCompleted	The IW cache after failback	Failback successfully completes on the IW cache cluster.	Healthy	Run <b>mmafmctl failback --stop</b> on the cache cluster.
NeedsResync	The SW cache cluster during home corruption	Occurs when the home cluster is accidentally corrupted	Unhealthy	Run <b>mmafmctl resync</b> on the cache.
NeedsResync	Recovery on the SW cache	A rare state possible only under error conditions during recovery	Unhealthy	No administrator action required. The system would fix this in the subsequent recovery.

## Fileset states for AFM DR

AFM DR fileset can have different states depending on the mode and queue states.

Run the **mmafmctl getstate** command to view the current cache state.

See the following table:

Table 32. AFM DR states and their description

AFM fileset state	Condition	Description	Healthy or Unhealthy	Administrator's action
Inactive	AFM primary is created	Operations have not been initiated on the primary after last daemon restart.	Healthy	None
FlushOnly	Operations are queued	Operations have not started to flush. This is a temporary state and moves to Active when a write is initiated.	Healthy	
Active	AFM primary is active	Primary is ready for operation	Healthy	None



Table 32. AFM DR states and their description (continued)

AFM fileset state	Condition	Description	Healthy or Unhealthy	Administrator's action
Dirty	AFM primary is active	Indicates there are pending changes in primary not yet played at secondary. Does not hamper normal activity.	Healthy	None
Recovery	The primary is accessed after MDS failure	Can occur when a new gateway is taking over a fileset as MDS after the old MDS failed.	Healthy	None
QueueOnly	The primary is running some operation	Can occur when operations such as recovery are being executed and operations are being queued and are not yet flushed.	Healthy	This is a temporary state.
Disconnected	It occurs when the MDS cannot connect to the NFS server at secondary	Occurs only in a cache cluster that is created over NFS export. When parallel I/O is configured, this state shows the connectivity between the MDS and the mapped home server, irrespective of other gateway nodes.	Unhealthy	Correct the errant NFS servers on the secondary cluster.
Unmounted	Primary using NFS detects a change in secondary - sometimes during creation or in the middle of operation if secondary exports are interfered	This can occur if: <ul style="list-style-type: none"> <li>• Secondary NFS is not accessible</li> <li>• Secondary exports are not exported properly</li> <li>• Secondary export does not exist</li> </ul>	Unhealthy	<ol style="list-style-type: none"> <li>1. Rectify the NFS export issue as in secondary setup section and retry access</li> <li>2. Relink primary if it does not recover.</li> </ol> <p>After <b>mountRetryInterval</b> of the MDS, the primary retries connecting with secondary</p>
Unmounted	The primary that is using the GPFS protocol detects a change in the secondary cluster, sometimes during creation or in the middle of an operation	Occurs when there are problems accessing the local mount of the remote file system.	Unhealthy	Check remote filesystem mount on the primary cluster and remount if necessary.

Table 32. AFM DR states and their description (continued)

AFM fileset state	Condition	Description	Healthy or Unhealthy	Administrator's action
Dropped	Recovery failed.	Occurs when the local file system is full, space is not available on the primary, or a policy failure during recovery.	Unhealthy	Fix the issue and access the fileset to retry recovery.
Dropped	A primary with active queue operations is forcibly unlinked	All queued operations are being de-queued, and the fileset remains in the Dropped state and moves to the Inactive state when the unlinking is complete.	Healthy	This is a temporary state.
Dropped	Old GW node starts functioning properly after a failure	AFM internally performs queue transfers from one gateway to another to handle gateway node failures.	Healthy	The system resolves this state on the next access.
Dropped	Primary creation or in the middle of an operation if the home exports changed.	Export problems at secondary such as: <ul style="list-style-type: none"> <li>The home path is not exported on all NFS server nodes that are interacting with the cache clusters. Even if the home cluster is exported after the operations have started on the fileset, problems might persist.</li> <li>Changing fsid on the home cluster after the fileset operations have begun.</li> <li>All home nfs servers do not have the same fsid for the same export path.</li> </ul>	Unhealthy	<ol style="list-style-type: none"> <li>Fix the NFS export issue in the secondary setup section and retry for access.</li> <li>Relink the primary if the cache cluster does not recover.</li> </ol> <p>After <b>mountRetryInterval</b> the MDS retries connecting with the secondary.</p>

Table 32. AFM DR states and their description (continued)

AFM fileset state	Condition	Description	Healthy or Unhealthy	Administrator's action
Dropped	During recovery or normal operation	If gateway queue memory is exceeded, the queue can get dropped. The memory has to be increased to accommodate all requests and bring the queue back to the Active state.	Unhealthy	Increase <b>afmHardMemThreshold</b> .
NeedsResync	Recovery on primary	This is a rare state and is possible only under error conditions during recovery.	Unhealthy	The problem should get automatically fixed in the subsequent recovery.
NeedsResync	Failback on primary or conversion from GPFS/SW to primary	This is a rare state and is possible only under error conditions during failback or conversion.	Unhealthy	Rerun failback or conversion.
PrimInitProg	Creating a primary fileset	This state is used while primary and secondary are in the process of establishing a relationship while the psnap0 is in progress. All operations are disallowed till psnap0 is taken locally. This should move to active when psnap0 is queued and played on the secondary side.	Healthy	Should look into errors on psnap0 failure if fileset state did not become active.
PrimInitFail	Converting a GPFS or AFM fileset to primary using <b>mmafmctl convertToPrimary</b> command.	This is a rare failure state when the psnap0 has not been created at the primary. In this state no data is moved from the primary to the secondary. The administrator should check that the gateway nodes are up and file system is mounted on them on the primary. The secondary fileset should also be setup correctly and available for use.	Unhealthy	Re-running the <b>mmafmctl convertToPrimary</b> command without any parameters helps to come out of this state.

Table 32. AFM DR states and their description (continued)

AFM fileset state	Condition	Description	Healthy or Unhealthy	Administrator's action
FailbackInProgress	Primary failback started	This is the state when failback is initiated on the primary.	Healthy	None

## Callback events for AFM and AFM DR

You can use events to monitor AFM and AFM DR fileset.

All events are at the fileset level. To add the events, run the **mmaddcallback** command.

An example of the command is

```
#mmdelcallback callback3
mmaddcallback callback3 --command /tmp/recovery_events.sh --event
afmRecoveryStart --parms "%eventName %homeServer %fsName %filesetName
%reason"
```

Table 33. List of events that can be added using **mmaddcallback**

Events	Applicable to..	Description
afmprepend	All AFM filesets	Completion of the prefetch task.
afmRecoveryStart	SW, IW	Beginning of the recovery process.
afmRecoveryEnd	SW, IW	End of the recovery process.
afmHomeDisconnected	All AFM filesets	For NFS target: The AFM home is not reachable.
afmHomeConnected	All AFM filesets	For NFS target: The AFM home is reachable.
afmFilesetExpired	RO fileset	For RO fileset: Fileset has expired
afmFilesetUnexpired	RO fileset	For RO fileset: Fileset is back to Active after expiration.
afmManualResyncComplete	SW, IW	The SW resync or failover process is complete.
afmQueueDropped	All AFM filesets	The queue is dropped.
afmfilesetunmounted	All AFM filesets	The fileset is in the Unmounted state.

## Monitoring with **mmpmon**

Following commands can be used to monitor AFM and AFM DR.

1. To reset some statistics on a gateway node, run the following commands:

```
echo "afm_s reset" | mmpmon
echo "afm_s fset all" | mmpmon
```

2. To reset all statistics, run the following command:

```
mmfsadm afm resetall
```

3. To view the statistics, run the following command:

```
echo afm_s | mmpmon -s -r 0 -d 2000
```

This command shows statistics from the time the Gateway is functioning. Every gateway recycle resets the statistics.

The following example is from an AFM Gateway node. The example shows how many operations of each type were executed on the gateway node.

```
c2m3n10 Tue May 10 09:55:59 0]~# echo afm_s | mmpmon
|
| mmpmon> mmpmon node 192.168.2.20 name c2m3n10 afm_s s OK
| Name           Queued      Inflight  Completed  Errors    Filtered  ENOENT
| lookup         0           0         1           0         0         0
| create         0           0         20          0         10        0
| remove         0           0         0           0         10        0
| open           0           0         2           0         0         0
| read           0           0         0           0         1         0
| write          0           0         20          0         650       0
| BytesWritten = 53320860 (50.85 MB) (26035.58 KB/s) BytesToWrite = 0 (0.00 KB)
| Queue Delay (s) (min:0 max:19 avg:18)
| Async Msgs (expire:50 force:0 sync:4 revoke:0)
| NumMsgExecuted = 715
| NumHomeconn   = 292
| NumHomedisc   = 292
| NumRPOMisses  = 1
```

The fields are as under -

Table 34. Field description of the example

Field name	Description
<b>BytesWritten</b>	The amount of data synchronized to home.
<b>BytesToWrite</b>	The amount of data in queue.
<b>QueueDelay</b>	The maximum delay experienced by operations before sync to home.
<b>NumMsgExecuted</b>	The number of operations executed at home.
<b>NumHomeconn</b>	The number of times home reconnected after disconnection.
<b>NumHomedisc</b>	The number of times home disconnected.
<b>NumRPOMisses</b>	Related to RPOs for AFM primary fileset.

## Monitor prefetch

You can display the status of an AFM prefetch request by running the **mmafmctl prefetch** command without the **list-file** option.

For example, for file system **gpfs1** and fileset **iw\_1**, run the following command:

```
# mmafmctl gpfs1 prefetch -j iw_1
|
| Fileset Name   Async Read (Pending) Async Read (Failed) Async Read (Already Cached) Async Read(Total)
| Async Read (Data in Bytes)
| -----
| iw_1          11           0           0           11           0
```

This output displays that there are 11 inodes that must be prefetched Async Read (Pending). When the job has completed, the status command displays:

```
# mmafmctl gpfs1 prefetch -j iw_1
| Fileset Name   Async Read (Pending) Async Read (Failed) Async Read (Already Cached) Async Read(Total)
| Async Read (Data in Bytes)
| -----
| iw_1          0           0           10           11
```

## Policies used for monitoring AFM and AFM DR

You can monitor AFM and AFM DR using some policies and commands.

Following are the policies used for monitoring:

- The following file attributes are available through the policy engine:

Table 35. Attributes with their description

Attribute	Description
P	The file is managed by AFM and AFM DR.
u	The file is managed by AFM and AFM DR, and the file is fully cached. When a file originates at the home, it indicates that the entire file is copied from the home cluster.
v	A file or a soft link is newly created, but not copied to the home cluster.
w	The file has outstanding data updates.
x	A hard link is newly created, but not copied to the home cluster.
y	A file metadata was changed and the change not copied to the home cluster.
z	A file is local to the cache and is not queued at the home cluster.
j	A file is appended, but not copied to the home cluster. This attribute also indicates complete directories.
k	All files and directories that are not orphan and are repaired.

- A list of dirty files in the cache cluster:

There is an example of a LIST policy that generates a list of files in the cache with pending changes that have not been copied to the home cluster.

```
RULE 'listall' list 'all-files' SHOW( varchar(kb_allocated) || ' ' || varchar(file_size) || ' ' ||  
    varchar(misc_attributes) || ' ' || fileset_name) WHERE REGEX(misc_attributes,'[P]') AND  
    REGEX(misc_attributes,'[w|v|x|y|j]')
```

If there are no outstanding updates, an output file is not created.

- A list of partially cached files:

The following example is that of a LIST policy that generates a list of partially-cached files. If the file is in progress, partial caching is enabled or the home cluster becomes unavailable before the file is completely copied.

```
RULE 'listall' list 'all-files'  
    SHOW(varchar(kb_allocated) || ' ' || varchar(file_size) || ' ' ||  
    varchar(misc_attributes) || ' ' || fileset_name )  
    WHERE REGEX(misc_attributes,'[P]') AND NOT_REGEX(misc_attributes,'[u]') AND kb_allocated > 0
```

This list does not include files that are not cached. If partially-cached files do not exist, an output file is not created.

- The custom eviction policy:

The steps to use policies for AFM file eviction are - generate a list of files and run the eviction. This policy lists all the files that are managed by AFM are not accessed in the last seven days.

```

| RULE 'prefetch-list'
|     LIST 'toevict'
|     WHERE CURRENT_TIMESTAMP - ACCESS_TIME > INTERVAL '7' DAYS
|     AND REGEX(misc_attributes, '[P]') /* only list AFM managed files */

```

To limit the scope of the policy or to use it on different filesets run **mmapplypolicy** by using a directory path instead of a file system name. `/usr/lpp/mmfs/bin/mmapplypolicy $path -f $localworkdir -s $localworkdir -P $sharedworkdir/${policy} -I defer`

Use **mmafmctl** to evict the files: **mmafmctl datafs evict --list-file \$localworkdir/list.evict**

- A policy of uncached files:

The following example is of a LIST policy that generates a list of uncached files in the cache directory:

```

| RULE EXTERNAL LIST 'u_list'
|     RULE 'u_Rule' LIST 'u_list' DIRECTORIES_PLUS FOR FILESET ('sw1') WHERE NOT
|     REGEX(misc_attributes, '[u]')

```

| .





---

## Chapter 6. GPFS SNMP support

GPFS supports the use of the SNMP protocol for monitoring the status and configuration of the GPFS cluster. Using an SNMP application, the system administrator can get a detailed view of the system and be instantly notified of important events, such as a node or disk failure.

The Simple Network Management Protocol (SNMP) is an application-layer protocol that facilitates the exchange of management information between network devices. It is part of the Transmission Control Protocol/Internet Protocol (TCP/IP) protocol suite. SNMP enables network administrators to manage network performance, find and solve network problems, and plan for network growth.

SNMP consists of commands to enumerate, read, and write managed variables that are defined for a particular device. It also has a **trap** command, for communicating events asynchronously.

The variables are organized as instances of objects, known as management information bases (MIBs). MIBs are organized in a hierarchical tree by organization (for example, IBM). A GPFS MIB is defined for monitoring many aspects of GPFS.

An SNMP agent software architecture typically consists of a master agent and a set of subagents, which communicate with the master agent through a specific agent/subagent protocol (the AgentX protocol in this case). Each subagent handles a particular system or type of device. A GPFS SNMP subagent is provided, which maps the SNMP objects and their values.

---

### Installing Net-SNMP

The SNMP subagent runs on the collector node of the GPFS cluster. The collector node is designated by the system administrator.

For more information, see “Collector node administration” on page 107.

The Net-SNMP master agent (also called the SNMP daemon, or **snmpd**) must be installed on the collector node to communicate with the GPFS subagent and with your SNMP management application. Net-SNMP is included in most Linux distributions and should be supported by your Linux vendor. Source and binaries for several platforms are available from the download section of the Net-SNMP website ([www.net-snmp.org/download.html](http://www.net-snmp.org/download.html)).

**Note:** Currently, the collector node must run on the Linux operating system. For an up-to-date list of supported operating systems, specific distributions, and other dependencies, refer to the IBM Spectrum Scale FAQ in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html](http://www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html)).

The GPFS subagent expects to find the following shared object libraries:

```
libnetsnmpagent.so    -- from Net-SNMP
libnetsnmphelpers.so -- from Net-SNMP
libnetsnmpmibs.so    -- from Net-SNMP
libnetsnmp.so        -- from Net-SNMP
libwrap.so            -- from TCP Wrappers
libcrypto.so          -- from OpenSSL
```

**Note:** TCP Wrappers and OpenSSL are prerequisites and should have been installed when you installed Net-SNMP.

The installed libraries will be found in `/lib64` or `/usr/lib64` or `/usr/local/lib64`. They may be installed under names like `libnetsnmp.so.5.1.2`. The GPFS subagent expects to find them without the appended version information in the name. Library installation should create these symbolic links for you, so you will rarely need to create them yourself. You can ensure that symbolic links exist to the versioned name from the plain name. For example,

```
# cd /usr/lib64
# ln -s libnetsnmpmibs.so.5.1.2 libnetsnmpmibs.so
```

Repeat this process for all the libraries listed in this topic.

**Note:** For possible Linux platform and Net-SNMP version compatibility restrictions, see the IBM Spectrum Scale FAQ in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html](http://www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html)).

---

## Configuring Net-SNMP

The GPFS subagent process connects to the Net-SNMP master agent, `snmpd`.

The following entries are required in the `snmpd` configuration file on the collector node (usually, `/etc/snmp/snmpd.conf`):

```
master agentx
AgentXSocket tcp:localhost:705
trap2sink managementhost
```

where:

### *managementhost*

Is the host name or IP address of the host to which you want SNMP traps sent.

If your GPFS cluster has a large number of nodes or a large number of file systems for which information must be collected, you must increase the timeout and retry parameters for communication between the SNMP master agent and the GPFS subagent to allow time for the volume of information to be transmitted. The `snmpd` configuration file entries for this are:

```
agentXTimeout 60
agentXRetries 10
```

where:

### `agentXTimeout`

Is set to 60 seconds for subagent to master agent communication.

### `agentXRetries`

Is set to 10 for the number of communication retries.

**Note:** Other values may be appropriate depending on the number of nodes and file systems in your GPFS cluster.

After modifying the configuration file, restart the SNMP daemon.

---

## Configuring management applications

To configure any SNMP-based management applications you might be using (such as Tivoli® NetView® or Tivoli Netcool®, or others), you must make the GPFS MIB file available on the processor on which the management application runs.

You must also supply the management application with the host name or IP address of the collector node to be able to extract GPFS monitoring information through SNMP. To do this, you must be familiar with your SNMP-based management applications.

For more information about Tivoli NetView or Tivoli Netcool, see IBM Knowledge Center ([www.ibm.com/support/knowledgecenter](http://www.ibm.com/support/knowledgecenter)).

---

## Installing MIB files on the collector node and management node

The GPFS management information base (MIB) file is found on the collector node in the `/usr/lpp/mmfs/data` directory with the name `GPFS-MIB.txt`.

To install this file on the collector node, do the following:

1. Copy or link the `/usr/lpp/mmfs/data/GPFS-MIB.txt` MIB file into the **SNMP MIB** directory (usually, `/usr/share/snmp/mibs`).

Alternatively, you could add the following line to the `snmp.conf` file (usually found in the directory `/etc/snmp`):

```
mibdirs +/usr/lpp/mmfs/data
```

2. Add the following entry to the `snmp.conf` file (usually found in the directory `/etc/snmp`):

```
mibs +GPFS-MIB
```

3. Restart the SNMP daemon.

Different management applications have different locations and ways for installing and loading a new MIB file. The following steps for installing the GPFS MIB file apply only to Net-SNMP. If you are using other management applications, such as NetView and NetCool, refer to corresponding product manuals (listed in “Configuring management applications” on page 106) for the procedure of MIB file installation and loading.

1. Remotely copy the `/usr/lpp/mmfs/data/GPFS-MIB.txt` MIB file from the collector node into the **SNMP MIB** directory (usually, `/usr/share/snmp/mibs`).

2. Add the following entry to the `snmp.conf` file (usually found in the directory `/etc/snmp`):

```
mibs +GPFS-MIB
```

3. You might need to restart the SNMP management application. Other steps might be necessary to make the GPFS MIB available to your management application.

---

## Collector node administration

Collector node administration includes: assigning, unassigning, and changing collector nodes. You can also see if a collector node is defined.

To assign a collector node and start the SNMP agent, enter:

```
mmchnode --snmp-agent -N NodeName
```

To unassign a collector node and stop the SNMP agent, enter:

```
mmchnode --nosnmp-agent -N NodeName
```

To see if there is a GPFS SNMP subagent collector node defined, enter:

```
mmiscluster | grep snmp
```

To change the collector node, issue the following two commands:

```
mmchnode --nosnmp-agent -N OldNodeName
```

```
mmchnode --snmp-agent -N NewNodeName
```

---

## Starting and stopping the SNMP subagent

The SNMP subagent is started and stopped automatically.

The SNMP subagent is started automatically when GPFS is started on the collector node. If GPFS is already running when the collector node is assigned, the **mmchnode** command will automatically start the SNMP subagent.

The SNMP subagent is stopped automatically when GPFS is stopped on the node (**mmshutdown**) or when the SNMP collector node is unassigned (**mmchnode**).

---

## The management and monitoring subagent

The GPFS SNMP management and monitoring subagent runs under an SNMP master agent such as Net-SNMP. It handles a portion of the SNMP OID space.

The management and monitoring subagent connects to the GPFS daemon on the collector node to retrieve updated information about the status of the GPFS cluster.

SNMP data can be retrieved using an SNMP application such as Tivoli NetView. NetView provides a MIB browser for retrieving user-requested data, as well as an event viewer for displaying asynchronous events.

Information that is collected includes status, configuration, and performance data about GPFS clusters, nodes, disks, file systems, storage pools, and asynchronous events. The following is a sample of the data that is collected for each of the following categories:

- Cluster status and configuration (see “Cluster status information” on page 109 and “Cluster configuration information” on page 109)
  - Name
  - Number of nodes
  - Primary and secondary servers
- Node status and configuration (see “Node status information” on page 110 and “Node configuration information” on page 110)
  - Name
  - Current status
  - Type
  - Platform
- File system status and performance (see “File system status information” on page 111 and “File system performance information” on page 112)
  - Name
  - Status
  - Total space
  - Free space
  - Accumulated statistics
- Storage pools (see “Storage pool information” on page 112)
  - Name
  - File system to which the storage pool belongs
  - Total storage pool space
  - Free storage pool space
  - Number of disks in the storage pool
- Disk status, configuration, and performance (see “Disk status information” on page 113, “Disk configuration information” on page 113, and “Disk performance information” on page 114)
  - Name
  - Status
  - Total space

- Free space
- Usage (metadata/data)
- Availability
- Statistics
- Asynchronous events (traps) (see “Net-SNMP traps” on page 115)
  - File system mounted or unmounted
  - Disks added, deleted, or changed
  - Node failure or recovery
  - File system creation, deletion, or state change
  - Storage pool is full or nearly full

**Note:** If file systems are not mounted on the collector node at the time that an SNMP request is received, the subagent can still obtain a list of file systems, storage pools, and disks, but some information, such as performance statistics, will be missing.

## SNMP object IDs

This topic defines the SNMP object IDs.

The management and monitoring SNMP subagent serves the OID space defined as **ibm.ibmProd.ibmGPFS**, which is the numerical **enterprises.2.6.212** OID space.

Underneath this top-level space are the following:

- **gpfsTraps** at **ibmGPFS.0**
- **gpfsMIBObjects** at **ibmGPFS.1**

## MIB objects

**gpfsMIBObjects** provides a space of objects that can be retrieved using a MIB browser application. Net-SNMP provides the **snmpget**, **snmpgetnext**, **snmptable**, and **snmpwalk** commands, which can be used to retrieve the contents of these fields.

## Cluster status information

This topic provides the values and descriptions for the GPFS cluster.

Table 36 shows the current status information for the GPFS cluster:

*Table 36. gpfsClusterStatusTable: Cluster status information*

Value	Description
<b>gpfsClusterName</b>	The cluster name.
<b>gpfsClusterId</b>	The cluster ID.
<b>gpfsClusterMinReleaseLevel</b>	The currently enabled cluster functionality level.
<b>gpfsClusterNumNodes</b>	The number of nodes that belong to the cluster.
<b>gpfsClusterNumFileSystems</b>	The number of file systems that belong to the cluster.

## Cluster configuration information

This topic shows the values and descriptions for the GPFS cluster configuration.

Table 37 on page 110 shows the GPFS cluster configuration information:

Table 37. *gpfsClusterConfigTable*: Cluster configuration information

Value	Description
<b>gpfsClusterConfigName</b>	The cluster name.
<b>gpfsClusterUidDomain</b>	The UID domain name for the cluster.
<b>gpfsClusterRemoteShellCommand</b>	The remote shell command being used.
<b>gpfsClusterRemoteFileCopyCommand</b>	The remote file copy command being used.
<b>gpfsClusterPrimaryServer</b>	The primary GPFS cluster configuration server.
<b>gpfsClusterSecondaryServer</b>	The secondary GPFS cluster configuration server.
<b>gpfsClusterMaxBlockSize</b>	The maximum file system block size.
<b>gpfsClusterDistributedTokenServer</b>	Indicates whether the distributed token server is enabled.
<b>gpfsClusterFailureDetectionTime</b>	The desired time for GPFS to react to a node failure.
<b>gpfsClusterTCPPort</b>	The TCP port number.
<b>gpfsClusterMinMissedPingTimeout</b>	The lower bound on a missed <b>ping</b> timeout (seconds).
<b>gpfsClusterMaxMissedPingTimeout</b>	The upper bound on missed <b>ping</b> timeout (seconds).

## Node status information

This topic provides a description for each GPFS node.

Table 38 shows the collected status data for each node:

Table 38. *gpfsNodeStatusTable*: Node status information

Node	Description
<b>gpfsNodeName</b>	The node name used by the GPFS daemon.
<b>gpfsNodeIp</b>	The node IP address.
<b>gpfsNodePlatform</b>	The operating system being used.
<b>gpfsNodeStatus</b>	The node status (for example, up or down).
<b>gpfsNodeFailureCount</b>	The number of node failures.
<b>gpfsNodeThreadWait</b>	The longest hung thread's wait time (milliseconds).
<b>gpfsNodeHealthy</b>	Indicates whether the node is healthy in terms of hung threads. If there are hung threads, the value is no.
<b>gpfsNodeDiagnosis</b>	Shows the number of hung threads and detail on the longest hung thread.
<b>gpfsNodeVersion</b>	The GPFS product version of the currently running daemon.

## Node configuration information

This topic shows the collected configuration data for each GPFS node.

Table 39 shows the collected configuration data for each node:

Table 39. *gpfsNodeConfigTable*: Node configuration information

Node	Description
<b>gpfsNodeConfigName</b>	The node name used by the GPFS daemon.
<b>gpfsNodeType</b>	The node type (for example, manager/client or quorum/nonquorum).

Table 39. *gpfsNodeConfigTable*: Node configuration information (continued)

Node	Description
<b>gpfsNodeAdmin</b>	Indicates whether the node is one of the preferred admin nodes.
<b>gpfsNodePagePoolL</b>	The size of the cache (low 32 bits).
<b>gpfsNodePagePoolH</b>	The size of the cache (high 32 bits).
<b>gpfsNodePrefetchThreads</b>	The number of prefetch threads.
<b>gpfsNodeMaxMbps</b>	An estimate of how many megabytes of data can be transferred per second.
<b>gpfsNodeMaxFilesToCache</b>	The number of inodes to cache for recently-used files that have been closed.
<b>gpfsNodeMaxStatCache</b>	The number of inodes to keep in the stat cache.
<b>gpfsNodeWorker1Threads</b>	The maximum number of worker threads that can be started.
<b>gpfsNodeDmapiEventTimeout</b>	The maximum time the file operation threads will block while waiting for a DMAPI synchronous event (milliseconds).
<b>gpfsNodeDmapiMountTimeout</b>	The maximum time that the mount operation will wait for a disposition for the mount event to be set (seconds).
<b>gpfsNodeDmapiSessFailureTimeout</b>	The maximum time the file operation threads will wait for the recovery of the failed DMAPI session (seconds).
<b>gpfsNodeNsdServerWaitTimeWindowOnMount</b>	Specifies a window of time during which a mount can wait for NSD servers to come up (seconds).
<b>gpfsNodeNsdServerWaitTimeForMount</b>	The maximum time that the mount operation will wait for NSD servers to come up (seconds).
<b>gpfsNodeUnmountOnDiskFail</b>	Indicates how the GPFS daemon will respond when a disk failure is detected. If it is "true", any disk failure will cause only the local node to forcibly unmount the file system that contains the failed disk.

## File system status information

This topic shows the collected status information for each GPFS file system.

Table 40 shows the collected status information for each file system:

Table 40. *gpfsFileSystemStatusTable*: File system status information

Value	Description
<b>gpfsFileSystemName</b>	The file system name.
<b>gpfsFileSystemStatus</b>	The status of the file system.
<b>gpfsFileSystemXstatus</b>	The executable status of the file system.
<b>gpfsFileSystemTotalSpaceL</b>	The total disk space of the file system in kilobytes (low 32 bits).
<b>gpfsFileSystemTotalSpaceH</b>	The total disk space of the file system in kilobytes (high 32 bits).
<b>gpfsFileSystemNumTotalInodesL</b>	The total number of file system inodes (low 32 bits).
<b>gpfsFileSystemNumTotalInodesH</b>	The total number of file system inodes (high 32 bits).
<b>gpfsFileSystemFreeSpaceL</b>	The free disk space of the file system in kilobytes (low 32 bits).

Table 40. *gpfsFileSystemStatusTable: File system status information (continued)*

Value	Description
<b>gpfsFileSystemFreeSpaceH</b>	The free disk space of the file system in kilobytes (high 32 bits).
<b>gpfsFileSystemNumFreeInodesL</b>	The number of free file system inodes (low 32 bits).
<b>gpfsFileSystemNumFreeInodesH</b>	The number of free file system inodes (high 32 bits).

## File system performance information

This topic provides the GPFS file system performance information.

Table 41 shows the file system performance information:

Table 41. *gpfsFileSystemPerfTable: File system performance information*

Value	Description
<b>gpfsFileSystemPerfName</b>	The file system name.
<b>gpfsFileSystemBytesReadL</b>	The number of bytes read from disk, not counting those read from cache (low 32 bits).
<b>gpfsFileSystemBytesReadH</b>	The number of bytes read from disk, not counting those read from cache (high 32 bits).
<b>gpfsFileSystemBytesCacheL</b>	The number of bytes read from the cache (low 32 bits).
<b>gpfsFileSystemBytesCacheH</b>	The number of bytes read from the cache (high 32 bits).
<b>gpfsFileSystemBytesWrittenL</b>	The number of bytes written, to both disk and cache (low 32 bits).
<b>gpfsFileSystemBytesWrittenH</b>	The number of bytes written, to both disk and cache (high 32 bits).
<b>gpfsFileSystemReads</b>	The number of read operations supplied from disk.
<b>gpfsFileSystemCaches</b>	The number of read operations supplied from cache.
<b>gpfsFileSystemWrites</b>	The number of write operations to both disk and cache.
<b>gpfsFileSystemOpenCalls</b>	The number of file system open calls.
<b>gpfsFileSystemCloseCalls</b>	The number of file system close calls.
<b>gpfsFileSystemReadCalls</b>	The number of file system read calls.
<b>gpfsFileSystemWriteCalls</b>	The number of file system write calls.
<b>gpfsFileSystemReaddirCalls</b>	The number of file system readdir calls.
<b>gpfsFileSystemInodesWritten</b>	The number of inode updates to disk.
<b>gpfsFileSystemInodesRead</b>	The number of inode reads.
<b>gpfsFileSystemInodesDeleted</b>	The number of inode deletions.
<b>gpfsFileSystemInodesCreated</b>	The number of inode creations.
<b>gpfsFileSystemStatCacheHit</b>	The number of stat cache hits.
<b>gpfsFileSystemStatCacheMiss</b>	The number of stat cache misses.

## Storage pool information

This topic provides the collected information for each GPFS storage pool.

Table 42 on page 113 shows the collected information for each storage pool:



Table 42. *gpfsStgPoolTable*: Storage pool information

Value	Description
<b>gpfsStgPoolName</b>	The name of the storage pool.
<b>gpfsStgPoolFSName</b>	The name of the file system to which the storage pool belongs.
<b>gpfsStgPoolTotalSpaceL</b>	The total disk space in the storage pool in kilobytes (low 32 bits).
<b>gpfsStgPoolTotalSpaceH</b>	The total disk space in the storage pool in kilobytes (high 32 bits).
<b>gpfsStgPoolFreeSpaceL</b>	The free disk space in the storage pool in kilobytes (low 32 bits).
<b>gpfsStgPoolFreeSpaceH</b>	The free disk space in the storage pool in kilobytes (high 32 bits).
<b>gpfsStgPoolNumDisks</b>	The number of disks in the storage pool.

## Disk status information

This topic provides the collected status information for each GPFS disk.

Table 43 shows the collected status information for each disk:

Table 43. *gpfsDiskStatusTable*: Disk status information

Value	Description
<b>gpfsDiskName</b>	The disk name.
<b>gpfsDiskFSName</b>	The name of the file system to which the disk belongs.
<b>gpfsDiskStgPoolName</b>	The name of the storage pool to which the disk belongs.
<b>gpfsDiskStatus</b>	The status of a disk (values: NotInUse, InUse, Suspended, BeingFormatted, BeingAdded, To Be Emptied, Being Emptied, Emptied, BeingDeleted, BeingDeleted-p, ReferencesBeingRemoved, BeingReplaced or Replacement).
<b>gpfsDiskAvailability</b>	The availability of the disk (Unchanged, OK, Unavailable, Recovering).
<b>gpfsDiskTotalSpaceL</b>	The total disk space in kilobytes (low 32 bits).
<b>gpfsDiskTotalSpaceH</b>	The total disk space in kilobytes (high 32 bits).
<b>gpfsDiskFullBlockFreeSpaceL</b>	The full block (unfragmented) free space in kilobytes (low 32 bits).
<b>gpfsDiskFullBlockFreeSpaceH</b>	The full block (unfragmented) free space in kilobytes (high 32 bits).
<b>gpfsDiskSubBlockFreeSpaceL</b>	The sub-block (fragmented) free space in kilobytes (low 32 bits).
<b>gpfsDiskSubBlockFreeSpaceH</b>	The sub-block (fragmented) free space in kilobytes (high 32 bits).

## Disk configuration information

This topic provides the collected configuration information for each GPFS disk.

Table 44 on page 114 shows the collected disk configuration information for each disk:

Table 44. *gpfsDiskConfigTable*: Disk configuration information

Value	Description
<b>gpfsDiskConfigName</b>	The disk name.
<b>gpfsDiskConfigFSName</b>	The name of the file system to which the disk belongs.
<b>gpfsDiskConfigStgPoolName</b>	The name of the storage pool to which the disk belongs.
<b>gpfsDiskMetadata</b>	Indicates whether the disk holds metadata.
<b>gpfsDiskData</b>	Indicates whether the disk holds data.

## Disk performance information

This topic provides the collected configuration information for each disk.

Table 45 shows the collected disk performance information for each disk:

Table 45. *gpfsDiskPerfTable*: Disk performance information

Value	Description
<b>gpfsDiskPerfName</b>	The disk name.
<b>gpfsDiskPerfFSName</b>	The name of the file system to which the disk belongs.
<b>gpfsDiskPerfStgPoolName</b>	The name of the storage pool to which the disk belongs.
<b>gpfsDiskReadTimeL</b>	The total time spent waiting for disk read operations (low 32 bits).
<b>gpfsDiskReadTimeH</b>	The total time spent waiting for disk read operations (high 32 bits).
<b>gpfsDiskWriteTimeL</b>	The total time spent waiting for disk write operations in microseconds (low 32 bits).
<b>gpfsDiskWriteTimeH</b>	The total time spent waiting for disk write operations in microseconds (high 32 bits).
<b>gpfsDiskLongestReadTimeL</b>	The longest disk read time in microseconds (low 32 bits).
<b>gpfsDiskLongestReadTimeH</b>	The longest disk read time in microseconds (high 32 bits).
<b>gpfsDiskLongestWriteTimeL</b>	The longest disk write time in microseconds (low 32 bits).
<b>gpfsDiskLongestWriteTimeH</b>	The longest disk write time in microseconds (high 32 bits).
<b>gpfsDiskShortestReadTimeL</b>	The shortest disk read time in microseconds (low 32 bits).
<b>gpfsDiskShortestReadTimeH</b>	The shortest disk read time in microseconds (high 32 bits).
<b>gpfsDiskShortestWriteTimeL</b>	The shortest disk write time in microseconds (low 32 bits).
<b>gpfsDiskShortestWriteTimeH</b>	The shortest disk write time in microseconds (high 32 bits).
<b>gpfsDiskReadBytesL</b>	The number of bytes read from the disk (low 32 bits).
<b>gpfsDiskReadBytesH</b>	The number of bytes read from the disk (high 32 bits).
<b>gpfsDiskWriteBytesL</b>	The number of bytes written to the disk (low 32 bits).
<b>gpfsDiskWriteBytesH</b>	The number of bytes written to the disk (high 32 bits).
<b>gpfsDiskReadOps</b>	The number of disk read operations.

Table 45. *gpfsDiskPerfTable*: Disk performance information (continued)

Value	Description
gpfsDiskWriteOps	The number of disk write operations.

## Net-SNMP traps

Traps provide asynchronous notification to the SNMP application when a particular event has been triggered in GPFS.

Table 46 shows the trap types that are defined:

Table 46. *Net-SNMP traps*

Net-SNMP trap type	This event is triggered by:
<b>Mount</b>	By the mounting node when the file system is mounted on a node.
<b>Unmount</b>	By the unmounting node when the file system is unmounted on a node.
<b>Add Disk</b>	By the file system manager when a disk is added to a file system on a node.
<b>Delete Disk</b>	By the file system manager when a disk is deleted from a file system.
<b>Change Disk</b>	By the file system manager when the status of a disk or the availability of a disk is changed within the file system.
<b>SGMGR Takeover</b>	By the cluster manager when a file system manager takeover is successfully completed for the file system.
<b>Node Failure</b>	By the cluster manager when a node fails.
<b>Node Recovery</b>	By the cluster manager when a node recovers normally.
<b>File System Creation</b>	By the file system manager when a file system is successfully created.
<b>File System Deletion</b>	By the file system manager when a file system is deleted.
<b>File System State Change</b>	By the file system manager when the state of a file system changes.
<b>New Connection</b>	When a new connection thread is established between the events exporter and the management application.
<b>Event Collection Buffer Overflow</b>	By the collector node when the internal event collection buffer in the GPFS daemon overflows.
<b>Hung Thread</b>	By the affected node when a hung thread is detected. The GPFS Events Exporter Watchdog thread periodically checks for threads that have been waiting for longer than a threshold amount of time.
<b>Storage Pool Utilization</b>	By the file system manager when the utilization of a storage pool becomes full or almost full.



---

## Chapter 7. Monitoring the IBM Spectrum Scale system by using call home

The call home feature collects files, logs, traces, and details of certain system health events from different nodes and services. These details are shared with the IBM support center for monitoring and problem determination. After analyzing this information, the IBM Support can contact the system administrator if there are any issues. Configuring call home reduces the response time for IBM Support to address the issues.

The call home feature is disabled by default. Use the `mmcallhome` command to enable and configure the call home feature.

---

### Collecting details that need to be shared with IBM Support

The details are collected from individual nodes that are marked as call home child nodes in the cluster. The details from each child node are collected by the call home node. You need to create a call home group by grouping call home child nodes. One of the nodes in the group is configured as the call home node and it performs data collection and upload.

The data gathering and upload can be configured individually on each group. Use the groups to reflect logical units in the cluster. For example, it is easier to manage when you create a group for all CES nodes and another group for all non-CES nodes. Call home groups help to distribute the data-gather and data-upload workload to prevent bottlenecks. It is recommended to limit the group size as 32. Larger groups are also possible but it might result in performance issues.

The following figure provides basic structure of call home functions.

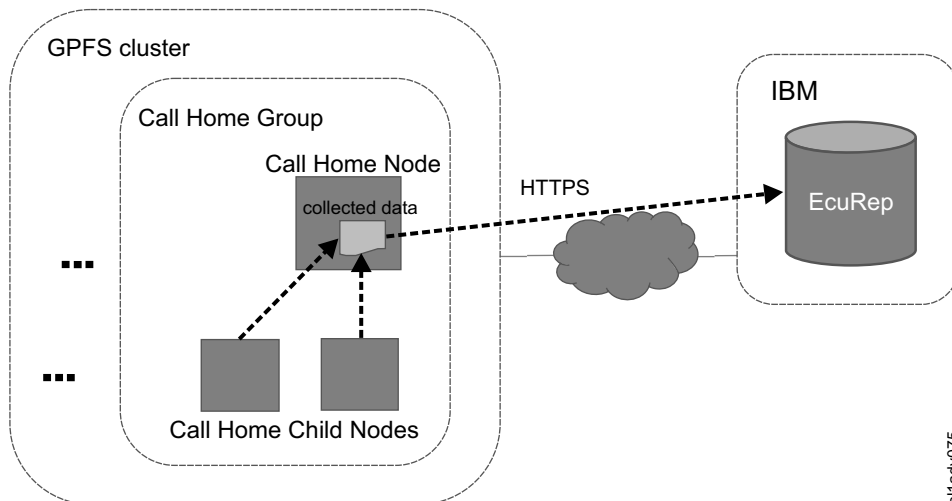


Figure 5. Call home architecture

The following are the descriptions of the important terminologies associated with the call home feature:

#### Call home group

A group of nodes configured by using the `mmcallhome group` command. A call home group consists of one call home node and multiple call home child nodes. Multiple call home groups can be configured within a GPFS cluster. You can automate the call home group creation by using the `mmcallhome group auto [options]` command.

## Call home node

This node initiates the data collection within the call home group and uploads the data package to IBM support center. A gather-send task process that runs on the call home node collects data from the child nodes and upload the data to the configured server. The gather-send configuration file includes information about what needs to be collected from the child nodes.

## Call home child node

A node that the call home node collects data from. The call home node is also determined as its child node.

The call home functionality provides the following data upload schemes to collect and upload the data:

1. **File upload:** Any file can be specified for upload.
2. **Package upload:** Collects predefined data package regularly. The call home feature provides *weekly.conf* schema to collect the package weekly and *daily.conf* schema to collect the package daily. These gather schemas are located at: `/usr/lpp/mmfs/data/callhome/gather.d`. After the upload, the data packages are stored in the data package directory for backup.

If the data collection is specified weekly, Cron is started once a week and data from call home child node is gathered by the call home node as specified in the *weekly.conf* file. When the gather task is finished, the data is uploaded from the call home node to the IBM Support. The following commands are issued internally to generate the data that needs to be shared with IBM Support:

- `tail -n 10000 /var/log/messages`
- `mmdiag --version`
- `mmdiag --config`
- `mmdiag --rpc`
- `mmlsnsd -L`
- `mmlsnsd -X`
- `mmgetstate -a`
- For each file system: `mmlsdisk <fs> -L`
- For each file system: `mmdf <fs>`
- `callhomeutils.pl info list`
- `mmces service list -a`
- `curl -i http://localhost:8080/info -X GET`
- `mmuserauth service list`
- `/usr/bin/curl -g -i -X GET http://localhost:35357/v3/projects`

If the data collection is specified daily, Cron is started once every day and data from call home child node is gathered by the call home node as specified in the *daily.conf* file. When the gather task is finished, the data is uploaded from the call home node to the IBM Support. The following commands are issued internally to generate the data that needs to be shared with IBM Support:

- `uname -a`
- `hostname`
- `sysctl -a`
- `uptime`
- `fdisk -l`
- `df, df -i, and df -k`
- `ifconfig -a`
- `lspci`
- `lsmod`
- `dmidecode`

- `cat /proc/device-tree/system-id`
- `ppc64_cpu --smt;ppc64_cpu --cores-present;ppc64_cpu --cores-on`
- `cat /proc/cpuinfo`
- `cat /proc/meminfo`
- `tail -n 10000 /var/adm/ras/mmfs.log.latest`
- `tail -n 10000 /var/log/messages`
- `mmdiag --rpc`
- `mmlscluster`
- `mmlsnode -a`
- `mmlsmgr`
- `mmlsnsd -L`
- `mmlscallback`
- `mmgetstate -a`
- `mmremoteccluster show all`
- `mmremotefs show all`
- `mmauth show`
- `tsstatus`
- `mmlsfs all`
- For each file system: `mmlsfs <fs>`
- For each file system: `mmlsfs <fs> --snc`
- For each file system: `mmlsdisk <fs> -L`
- For each file system: `mmlspolicy <fs> -L`
- For each file system: `mmlsfileset <fs> -L`
- For each file system: `mmlsfileset <fs> --afm -L`
- For each file system: `mmlssnapshot <fs>`
- For each file system: `mmdf <fs>`
- `mmlsnode -N waiters -L`
- `callhomeutils.pl info list`

When an event occurs and if the system is configured to upload data to IBM Support based on that event, the call home node that is configured for capturing the details on events collects the data, creates a file, and uses the `mmcallhome` command to upload the data. The following events can be configured to start the relevant data collection and upload:

- `nfsd_down`
- `ctdb_down`
- `ctdb_state_down`
- `smbd_down`

---

## Configuring call home

Perform the following steps to configure the call home functions:

1. Set the customer name, customer ID, and the country code. Optional to define a proxy.
2. Configure the call home groups
3. Enable the call home capability
4. Add the tasks required for periodic data collection
5. If a specific file needs to be uploaded, issue the `mmcallhome` run command

The **mmcallhome** command provides options to configure, enable, run, schedule, and monitor call home related tasks in the GPFS cluster. For more information, see **mmcallhome command** in *IBM Spectrum Scale: Command and Programming Reference*.

The following examples show how to perform the basic configuration steps that are used for configuring the call home functions:

1. To change customer information such as customer name, customer ID, and the country code, issue the **mmcallhome info change** command as shown in the following example:

```
mmcallhome info change --customer-name "SpectrumScaleTest" --customerid
"1234" --country-code "JP"
```

The system displays output similar to this:

Success

To set the parameters of proxy server, issue the **mmcallhome proxy change** command as shown in the following example:

```
mmcallhome proxy change --proxy-location okapi --proxy-port 80 --proxyusername
root --proxy-password <password>
```

The system displays output similar to this:

Success

2. To configure a call home group, issue the **mmcallhome group add** command as shown in the following example:

```
mmcallhome group add group1 themisto0 -N themisto0,themisto1,themisto2
```

The system displays output similar to this:

Call home group group1 has been created

You can automate the call home group creation by using the **mmcallhome group auto [options]** command as shown in the following examples:

- a. To create a call home group automatically, issue this command:

```
mmcallhome group auto
```

The system displays output similar to this:

```
mmcallhome group auto: [I] In progress: Create 1 new call home groups.
mmcallhome group auto: [I] In progress: Nodes without call home: 1
See /var/adm/ras/mmfs.log.latest for details.
group: autoGroup_1 successfully added
```

- b. To create a call home group automatically and enable the cluster for call home functionality by displaying options for acceptance, issue this command:

```
mmcallhome group auto --enable
```

The system displays output similar to this:

```
By accepting this request, you agree to allow IBM and its subsidiaries to store and use
your contact information and your support information anywhere they do business worldwide.
For more information, please refer to the Program license agreement and documentation.
If you agree, please respond with "accept" for acceptance, else with "not accepted" to decline.
(accept / not accepted)
accept
mmcallhome group auto: [I] In progress: Create 1 new call home groups.
mmcallhome group auto: [I] In progress: Nodes without call home: 1 See
/var/adm/ras/mmfs.log.latest for details.
group: autoGroup_1 successfully added
```

**Note:** To accept the call home functionality, type **accept** manually. Type **mmcallhome group auto --enable accept** to avoid the explicit acceptance from the user.

- c. To use create new group after deleting the existing group, issue this command:



```
| mmcallhome group auto --force
```

| The system displays output similar to this:

```
| mmcallhome group auto: [I] In progress: Create 1 new call home groups.  
| mmcallhome group auto: [I] In progress: Nodes without call home: 1  
| See /var/adm/ras/mmfs.log.latest for details.  
| mmcallhome group auto: [I] In progress: Delete existing groups.  
| Call home group autoGroup_1 has been deleted  
| group: autoGroup_2 successfully added
```

The **mmcallhome group auto --force --enable accept** performs the following functions:

- Deletes all existing groups
  - Creates new groups with optimized group size for all nodes where call home is installed
  - Enable all groups and assume the user accept the license. Without specifying the parameter, **accept** for the **--enable** option, the system prompts the user to accept or reject the displayed license.
3. To view the customer information, issue the **mmcallhome info list** command as shown in the following example:

```
mmcallhome info list
```

The system displays output similar to this:

```
| Parameter                Value  
| customer-name            SpectrumScaleTest  
| customer-id              1234  
| callhome-method          ethernet  
| country-code              JP
```

4. To enable the call home service, issue the **mmcallhome capability enable** command as shown in the following example:

```
mmcallhome capability enable
```

The system displays output similar to this:

```
Call home node: themisto0  
Call home child nodes to collect data: themisto0 themisto1 themisto2 (total 3 nodes)  
Excluded nodes:  
SSH Access Check: OK  
Data package directory: /tmp/mmfs/callhome  
Success
```

The **mmcallhome capability enable** command can enable or disable call home only in a single group. Automatic group generation is possible by using **mmcallhome group auto** command. This command can also enable and disable call home on all call home groups on the cluster. For more information on the automatic option, issue **mmcallhome group auto -h**.

5. To register a daily task with cron, issue the **mmcallhome schedule add** command as shown in the following example:

```
mmcallhome schedule add --task daily
```

The system displays output similar to this:

```
/etc/cron.d/gpfsallhome_gatherSend_daily.conf registered  
41 command entries are defined for this task
```

6. To register a weekly task with cron, issue the **mmcallhome schedule add** command as shown in the following example:

```
mmcallhome schedule add --task weekly
```

The system displays output similar to this:

```
/etc/cron.d/gpfsallhome_gatherSend_weekly.conf registered  
14 command entries are defined for this task
```

- To list the registered tasks for gather-send, issue the **mmcallhome schedule list** command as shown in the following example:

```
mmcallhome schedule list
```

The system displays output similar to this:

```
Registered Tasks for gatherSend:
ConfFile          CronParameters
daily.conf        3 2 * * *
weekly.conf       54 3 * * sun
```

**Note:** The CronParameter indicates the date and time settings for the execution of the command. It displays the values for minutes (0-59), hours (0-23), day of month (1-31), month (1-12 or jan-dec), and day of week (0-6, where sun=0 or sun-sat). For example, CronParameter 54 3 \* \* sun indicates that the command runs on every Sunday at 3:54 AM. For more details, see `crontab(5)` - Linux man page.

- To monitor the call home tasks, issue the **mmcallhome status list** command as shown in the following example:

```
mmcallhome status list
```

The system displays output similar to this:

```
Task   Start time      Status      Package file name
daily  20150930132656.582 success    ...aultDaily.g_daily.20150930132656582.c10.DC
daily  20150930133134.802 success    ...aultDaily.g_daily.20150930133134802.c10.DC
daily  20150930133537.509 success    ...aultDaily.g_daily.20150930133537509.c10.DC
daily  20150930133923.063 success    ...aultDaily.g_daily.20150930133923063.c10.DC
RunSendFile 20150930133422.843 success
...group2.MyTestData.s_file.20150930133422843.c10.DC
```

- To view the status of the currently running and the already completed call home tasks, issue the **mmcallhome status list** command as shown in the following example:

```
mmcallhome status list --verbose
```

The system displays output similar to this:

```
Task   Start time      Updated time      Status      RC or Step
      Package file name
      [ additional info: value ]
-----
      31790849425327.4_2_1_0.x.abc.autoGroup_1.gat_weekly.g_weekly.20160412160447854.c10.DC
-----
      31790849425327.4_2_1_0.x.abc.autoGroup_1.gat_weekly.g_weekly.20160412173941161.c10.DC
-----
weekly 20160412174030.803 20160412174034    failed      RC=6 (lock err) NA
-----
      31790849425327.4_2_1_0.x.abc.autoGroup_1.gat_weekly.g_weekly.20160412175159390.c10.DC
```

**Note:** Sometimes, the output of `mmcallhome status list --verbose` displays a single line without detailed information about RC indicating successful completion of call home tasks. The failed status indicates an issue with the call home task and the RC numeral indicates the respective issue. If the value of RC is zero, it indicates that the upload procedure is successful, but some automatically resolvable issue occurred while uploading the data. The value, `RC != 0`, indicates that the upload procedure is not successful. The detailed information about the upload procedure is available in the logs.

## Example for configuring call home based on a use case

Each call home group can be configured independently. The following example shows how to address the following use case.

There are three call home groups such as `group_1`, `group_2`, and `group_3`.

## group\_1

The group that is named *group\_1* has weekly and daily data upload scheduled but call home is disabled. The following example shows how to configure this scenario:

```
mmcallhome group add group_1 callhome_1_Node --node childe_1_Group_1,child_2_Group_1
```

The call home feature is disabled by default for this group.

Issue the following command on one of the nodes of the call home group (callhome\_1\_Node, childe\_1\_Group\_1, child\_2\_Group\_1)

```
mcallhome shedule add --task DAILY
```

Issue the following command on one of the nodes of the call home group:

```
mmcallhome shedule add --task WEEKLY
```

Although daily and weekly gather task are scheduled, no data is collected and uploaded because the call home feature is still disabled.

## group\_2

Weekly upload is scheduled and call home feature is enabled for the group named *group\_2*. The following example shows how to configure this scenario:

```
mmcallhome group add group_2 callhome_2_Node --node childe_1_Group_2,child_2_Group_2
```

The call home feature is disabled by default for this group.

Issue the following command on one of the nodes of the call home group (nodes callhome\_2\_Node, childe\_1\_Group\_2, child\_2\_Group\_2):

```
mmcallhome capability enable
```

Call home feature is enabled on group\_2.

Issue the following command on one of the nodes of the call home group (nodes callhome\_2\_Node, childe\_1\_Group\_2, child\_2\_Group\_2):

```
mmcallhome shedule add --task WEEKLY
```

You can issue the **mmcallhome schedule** command before the **mmcallhome capability** command. However, no data upload is possible until the **mmcallhome capability enable** command is issued.

After performing the above steps, data is gathered and uploaded weekly. Apart from the scheduled data collection and upload, an admin or a system health event can also do data collection and upload from the group.

## group\_3

Call home is enabled but nothing is scheduled on the group named *group\_3*. The following example shows how to configure this scenario:

```
mmcallhome group add group_3 callhome_3_Node --node childe_1_Group_3,child_2_Group_3
```

The call home feature is disabled by default for this group.

Issue the following command on one of the nodes of the call home group (callhome\_3\_Node, childe\_1\_Group\_3, child\_2\_Group\_3):

```
mmcallhome capability enable
```

Call home feature is on *group\_3*. That is, an admin or a system health event can do data collection and upload from the group. No data is uploaded periodically.

---

## Sharing collected details with IBM Support

The send file task that runs on the call home node uploads files and packages that are collected at the call home node.

The call home feature allows data upload in following two ways:

1. **Manual upload:** The call home feature provides manual upload option to upload the files or packages manually to the IBM server. To upload any data manually, issue **mmcallhome run [option]**.
2. **Automatic upload:** Use the **mmcallhome schedule [option]** to schedule a weekly or daily schema to upload the predefined data. If system health detects an event that requires a specific data to be uploaded, then the **mmcallhome run [option]** command is automatically run to upload the specified file, which is required for problem determination. This file contains previously collected system health relevant data.

The manual and automatic upload options can upload the data to the IBM server. This data is not analyzed and gets deleted after a specified time period (generally, 2 weeks). The data that is collected in EcuRep database are accessed and analyzed only if any specific event, relevant to problem determination, occurs.

Every time call home collects data or the call home command is started to upload a specific file, call home first creates a data package file. The data package file is a tar file. The successfully uploaded files are saved on the call home node under `<gpfs data dump dir>/callhome/rsENUploaded(<gpfs data dump dir>` is a GPFS global define) for backup.

---

## Prerequisites for configuring call home

The following criteria must be met to be able to use the call home functionality on your GPFS cluster:

- Only Linux nodes are supported.
- Only x86 or Power® PC big endian platform is supported.
- Only supported on RHEL 7 and SLES 12.
- At least one call home group needs to be defined.
- GPFS CCR (Clustered Configuration Repository) needs to be enabled
- The call home node needs to be able to access the following IP addresses and ports:
  - Host name: `esupport.ibm.com`
  - IP address: `129.42.56.189`, `129.42.60.189`, `129.42.52.189`
  - Port number: `443`

The recommendation is to open `129.42.0.0/18`

- The call home node needs to have 1 GB free space under the data package directory (2 - 3 GB preferred)
- There must not be multiple nodes with the same short host name within the GPFS cluster.
- Only GPFS cluster with `/usr/bin/scp` (or any other path to `scp`) configured for the file transfer method is supported.
- Install the call home RPMs in the following order:
  1. `gpfs.callhome-ecc-client-4.2.1-0.001.noarch.rpm`
  2. `gpfs.callhome-4.2-1`
    - a. RHEL: `gpfs.callhome-4.2.1-0.007.el7.noarch.rpm`
    - b. SLES: `gpfs.callhome-4.2.1-0.007.sles12.noarch.rpm`

- Ensure that the following dependencies are resolved when you install call home:
  - Dependencies for gpfs.callhome-4.2 and above:
    - perl-List-MoreUtils
    - perl-Module-Runtime
    - perl-Try-Tiny
    - perl-Module-Implementation
    - perl-Params-Validate
    - perl-DateTime-Locale
    - perl-Sub-Install
    - perl-Params-Util
    - perl-Data-OptList
    - perl-Package-DeprecationManager
    - perl-Package-Stash-XS
    - perl-Package-Stash
    - perl-Class-Load
    - perl-Class-Singleton
    - perl-TimeDate
    - perl-DateTime
    - perl-DateTime-TimeZone
    - perl-Data-Dumper package



---

## Chapter 8. Monitoring the health of Transparent Cloud Tiering service

This topic describes how to check the health of Transparent Cloud Tiering service in IBM Spectrum Scale.

To monitor the health of the Transparent Cloud Tiering service, issue a command according to this syntax:

```
mmcloudgateway service status [-N { Node [,Node ...]  
| NodeFile | NodeClass }]
```

For example, to check the service status of all the nodes, issue this command:

```
mmcloudgateway service status
```

The system displays output similar to this:

Node	Daemon node name	TCT Server Status	TCT Filesystem Status	TCT Account Status
1	jupiter-vm581.pok.stglabs.ibm.com	Started	Configured	Active

For more information on various options available for checking the health status, see **mmcloudgateway command** in *IBM Spectrum Scale: Command and Programming Reference*.

**Note:** You must run this command on the Transparent Cloud Tiering node.





E!

---

## E! Chapter 9. Best practices for troubleshooting

E! Following certain best practices helps you to make the troubleshooting process easier.

E!

### E! How to get started with troubleshooting

E! Troubleshooting the issues reported in the system is easier when you follow the process step-by-step.

E! When you experience some issues with the system, go through the following steps to get started with the E! troubleshooting:

- E! 1. Check the events that are reported in various nodes of the cluster by using the `mmhealth node show` command.
- E! 2. Check the user action corresponding to the active events and take the appropriate action. For more information on the events and corresponding user action, see “Events” on page 371.
- E! 3. If you are facing a deadlock issue, see Chapter 12, “Managing deadlocks,” on page 207 to know how to resolve the issue.
- E! 4. Collect the details of the issues through logs, dumps, and traces. You can use various CLI commands and **Settings > Diagnostic Data** GUI page to collect the details of the issues reported in the system. For more information, see Chapter 11, “Collecting details of the issues,” on page 139.
- E! 5. Based on the type of issue, browse through the various topics that are listed in the troubleshooting section and try to resolve the issue.
- E! 6. If you cannot resolve the issue by yourself, contact IBM Support. For more information on how to contact IBM Support, see Chapter 26, “Support for troubleshooting,” on page 367.

E!

### E! Set up event notifications

E! You can configure the system to send notifications when a new event is reported. It helps to correct any E! issues reported in the system as soon as possible. You can select the type of event that you want to be E! notified about. For example, restrict notifications to just events that require immediate action.

E! The following types of event notifications features are available in the IBM Spectrum Scale system:

- E! 1. **Email:** Use **Settings > Event Notifications** page in the GUI to configure email notifications. In this E! mode of event notification, the operational and error-related data is sent to the recipients in an email. E! An event notification can be sent to one or more email addresses.
- E! 2. **Simple Network Management Protocol (SNMP):** SNMP is a standard protocol for managing E! networks and exchanging messages. The system can send SNMP messages that notify personnel about E! an event. You can use an SNMP manager to view the SNMP messages that the system sends. Select E! **Settings > Event Notifications > SNMP Manager** in the GUI to configure SNMP managers for event E! notifications. You can specify up to a maximum of six SNMP managers.
- E! 3. **Call home:** This feature helps the IBM Support to monitor and troubleshoot the system. You can E! configure the IBM Spectrum Scale system to send certain details to the IBM Support. The IBM E! Support constantly monitors the details that are received from the system and if appropriate, contacts E! the IBM Spectrum Scale system administrator to take certain actions to fix the issues.
- E! 4. **Callback:** Use the `mmaddcallback` command to configure the system to provide notifications when E! node and cluster events occur.

E! **Related concepts:**

E! “Configuring email notifications” on page 80

E! The email feature transmits operational and error-related data in the form of an event notification email.

E! “Configuring SNMP manager” on page 80

E! Simple Network Management Protocol (SNMP) is a standard protocol for managing networks and exchanging messages. The system can send SNMP messages that notify personnel about an event. You can use an SNMP manager to view the SNMP messages that the system sends.

E! **Related reference:**

E! “Call home” on page 370

E! The call home feature automatically notifies IBM Support if certain types of events occur in the system.

E! Using this information, IBM Support can contact the system administrator in case of any issues.

E! Configuring call home reduces the response time for IBM Support to address the issues.

E! \_\_\_\_\_

## E! **Back up your data**

E! You need to back up data regularly to avoid data loss. It is also recommended to take backups before you start troubleshooting. The IBM Spectrum Scale provides various options to create data backups.

E! Follow the guidelines in the following sections to avoid any issues while creating backup:

E! • *GPFS(tm) backup data in IBM Spectrum Scale: Concepts, Planning, and Installation Guide*

E! • *Backup considerations for using IBM Spectrum Protect in IBM Spectrum Scale: Concepts, Planning, and Installation Guide*

E! • *Configuration reference for using IBM Spectrum Protect with IBM Spectrum Scale(tm) in IBM Spectrum Scale: Administration Guide*

E! • *Protecting data in a file system using backup in IBM Spectrum Scale: Administration Guide*

E! • *Backup procedure with SOBAR in IBM Spectrum Scale: Administration Guide*

E! The following best practices help you to troubleshoot the issues that might arise in the data backup process:

E! 1. Enable the most useful messages in **mmbackup** command by setting the **MMBACKUP\_PROGRESS\_CONTENT** and **MMBACKUP\_PROGRESS\_INTERVAL** environment variables in the command environment prior to issuing the **mmbackup** command. Setting **MMBACKUP\_PROGRESS\_CONTENT=7** provides the most useful messages. For more information on these variables, see *mmbackup command* in *IBM Spectrum Scale: Command and Programming Reference*.

E! 2. If the **mmbackup** process is failing regularly, enable debug options in the backup process:

E! Use the **DEBUGmmbackup** environment variable or the **-d** option that is available in the **mmbackup** command to enable debugging features. This variable controls what debugging features are enabled. It is interpreted as a bitmask with the following bit meanings:

E! **0x001** Specifies that basic debug messages are printed to STDOUT. There are multiple components that comprise **mmbackup**, so the debug message prefixes can vary. Some examples include:

E! `mmbackup:mbackup.sh`  
E! `DEBUGtsbackup33:`

E! **0x002** Specifies that temporary files are to be preserved for later analysis.

E! **0x004** Specifies that all **dsmc** command output is to be mirrored to STDOUT.

E! The **-d** option in the **mmbackup** command line is equivalent to **DEBUGmmbackup = 1**.

E! 3. To troubleshoot problems with backup subtask execution, enable debugging in the **tsbuhelper** program.

E! Use the **DEBUGtsbuhelper** environment variable to enable debugging features in the **mmbackup** helper program **tsbuhelper**.

E!

---

## E! **Resolve events in a timely manner**

E! Resolving the issues in a timely manner helps to get attention on the new and most critical events. If E! there are a number of unfixed alerts, fixing any one event might become more difficult because of the E! effects of the other events. You can use either CLI or GUI to view the list of issues that are reported in E! the system.

E! You can use the `mmhealth node eventlog` to list the events that are reported in the system.

E! The **Monitoring > Events** GUI page lists all events reported in the system. You can also mark certain E! events as read to change the status of the event in the events view. The status icons become gray in case E! an error or warning is fixed or if it is marked as read. Some issues can be resolved by running a fix E! procedure. Use the action **Run Fix Procedure** to do so. The Events page provides a recommendation for E! which fix procedure to run next.

E!

---

## E! **Keep your software up to date**

E! Check for new code releases and update your code on a regular basis.

E! This can be done by checking the IBM support website to see if new code releases are available:  
E! [http://www.ibm.com/support/entry/portal/product/system\\_storage/storage\\_software/  
E! software\\_defined\\_storage/ibm\\_spectrum\\_scale](http://www.ibm.com/support/entry/portal/product/system_storage/storage_software/software_defined_storage/ibm_spectrum_scale). The release notes provide information about new function E! in a release plus any issues that are resolved with the new release. Update your code regularly if the E! release notes indicate a potential issue.

E! **Note:** If a critical problem is detected on the field, IBM may send a flash, advising the user to contact E! IBM for an efix. The efix when applied might resolve the issue.

E!

---

## E! **Subscribe to the support notification**

E! Subscribe to support notifications so that you are aware of best practices and issues that might affect your E! system.

E! Subscribe to support notifications by visiting the IBM support page on the following IBM website:  
E! <http://www.ibm.com/support/mynotifications>.

E! By subscribing, you are informed of new and updated support site information, such as publications, E! hints and tips, technical notes, product flashes (alerts), and downloads.

E!

---

## E! **Know your IBM warranty and maintenance agreement details**

E! If you have a warranty or maintenance agreement with IBM, know the details that must be supplied E! when you call for support.

E! For more information on the IBM Warranty and maintenance details, see Warranties, licenses and E! maintenance.

E!

---

## E! **Know how to report a problem**

E! If you need help, service, technical assistance, or want more information about IBM products, you find a E! wide variety of sources available from IBM to assist you.

E! IBM maintains pages on the web where you can get information about IBM products and fee services, E! product implementation and usage assistance, break and fix service support, and the latest technical E! information. The following table provides the URLs of the IBM websites where you can find the support E! information.

E! *Table 47. IBM websites for help, services, and information*

E! Website	E! Address
E! IBM home page	http://www.ibm.com
E! Directory of worldwide contacts	http://www.ibm.com/planetwide
E! Support for IBM Spectrum Scale E! E!	http://www.ibm.com/support/entry/portal/product/system_storage/storage_software/software_defined_storage/ibm_spectrum_scale
E! Support for IBM System Storage® and IBM Total Storage products E! E!	http://www.ibm.com/support/entry/portal/product/system_storage/

E! **Note:** Available services, telephone numbers, and web links are subject to change without notice.

E! **Before you call**

E! Make sure that you have taken steps to try to solve the problem yourself before you call. Some E! suggestions for resolving the problem before calling IBM Support include:

- E! • Check all hardware for issues beforehand.
- E! • Use the troubleshooting information in your system documentation. The troubleshooting section of the E! IBM Knowledge Center contains procedures to help you diagnose problems.

E! To check for technical information, hints, tips, and new device drivers or to submit a request for E! information, Go to the IBM Support website at: [http://www.ibm.com/support/entry/portal/product/system\\_storage/storage\\_software/software\\_defined\\_storage/ibm\\_spectrum\\_scale](http://www.ibm.com/support/entry/portal/product/system_storage/storage_software/software_defined_storage/ibm_spectrum_scale).

E! **Using the documentation**

E! Information about your IBM storage system is available in the documentation that comes with the E! product. That documentation includes printed documents, online documents, readme files, and help files E! in addition to the IBM Knowledge Center. See the troubleshooting information for diagnostic instructions. E! To access this information, go to [http://www.ibm.com/support/entry/portal/product/system\\_storage/storage\\_software/software\\_defined\\_storage/ibm\\_spectrum\\_scale](http://www.ibm.com/support/entry/portal/product/system_storage/storage_software/software_defined_storage/ibm_spectrum_scale) and follow the instructions. The entire E! product documentation is available at: [https://www.ibm.com/support/knowledgecenter/STXKQY/ibmspectrumscale\\_welcome.html?lang=en](https://www.ibm.com/support/knowledgecenter/STXKQY/ibmspectrumscale_welcome.html?lang=en).

---

E! **Other problem determination hints and tips**

E! These hints and tips might be helpful when investigating problems related to logical volumes, quorum E! nodes, or system performance that can be encountered while using GPFS.

E! See these topics for more information:

- E! • “Which physical disk is associated with a logical volume in AIX systems?” on page 133
- E! • “Which nodes in my cluster are quorum nodes?” on page 133
- E! • “What is stored in the /tmp/mmfs directory and why does it sometimes disappear?” on page 134
- E! • “Why does my system load increase significantly during the night?” on page 134
- E! • “What do I do if I receive message 6027-648?” on page 134
- E! • “Why can't I see my newly mounted Windows file system?” on page 134
- E! • “Why is the file system mounted on the wrong drive letter?” on page 134
- E! • “Why does the offline mmfsck command fail with "Error creating internal storage"?” on page 135
- E! • “Questions related to active file management” on page 135

## E! Which physical disk is associated with a logical volume in AIX systems?

E! Earlier releases of GPFS allowed AIX logical volumes to be used in GPFS file systems. Their use is now discouraged because they are limited with regard to their clustering ability and cross platform support.

E! Existing file systems using AIX logical volumes are, however, still supported. This information might be of use when working with those configurations.

E! If an error report contains a reference to a logical volume pertaining to GPFS, you can use the `lslv -l` command to list the physical volume name. For example, if you want to find the physical disk associated with logical volume `gpfs7lv`, issue:

```
E! lslv -l gpfs44lv
```

E! Output is similar to this, with the physical volume name in column one.

```
E! gpfs44lv:N/A
E! PV          COPIES      IN BAND    DISTRIBUTION
E! hdisk8      537:000:000  100%      108:107:107:107:108
```

## E! Which nodes in my cluster are quorum nodes?

E! Use the `mmlscluster` command to determine which nodes in your cluster are quorum nodes.

E! Output is similar to this:

```
E! GPFS cluster information
```

```
E! =====
```

```
E! GPFS cluster name:      cluster1.kgn.ibm.com
E! GPFS cluster id:       680681562214606028
E! GPFS UID domain:       cluster1.kgn.ibm.com
E! Remote shell command:  /usr/bin/rsh
E! Remote file copy command: /usr/bin/rcp
E! Repository type:       server-based
E!
```

```
E! GPFS cluster configuration servers:
```

```
E! -----
```

```
E! Primary server:      k164n06.kgn.ibm.com
```

```
E! Secondary server:    k164n05.kgn.ibm.com
```

```
E!
```

```
E! Node  Daemon node name      IP address      Admin node name  Designation
E! -----
E!  1    k164n04.kgn.ibm.com      198.117.68.68  k164n04.kgn.ibm.com  quorum
E!  2    k164n05.kgn.ibm.com      198.117.68.71  k164n05.kgn.ibm.com  quorum
E!  3    k164n06.kgn.ibm.com      198.117.68.70  k164n06.kgn.ibm.com
```

E! In this example, `k164n04` and `k164n05` are quorum nodes and `k164n06` is a nonquorum node.

E! To change the quorum status of a node, use the `mmchnode` command. To change one quorum node to nonquorum, GPFS does not have to be stopped. If you are changing more than one node at the same time, GPFS needs to be down on all the affected nodes. GPFS does not have to be stopped when changing nonquorum nodes to quorum nodes, nor does it need to be stopped on nodes that are not affected.

E! For example, to make `k164n05` a nonquorum node, and `k164n06` a quorum node, issue these commands:

```
E! mmchnode --nonquorum -N k164n05
```

```
E! mmchnode --quorum -N k164n06
```

E! To set a node's quorum designation at the time that it is added to the cluster, see `mmcrcluster` or `mmaddnode` command in *IBM Spectrum Scale: Command and Programming Reference*.

## E! **What is stored in the /tmp/mmfs directory and why does it sometimes disappear?**

E! When GPFS encounters an internal problem, certain state information is saved in the GPFS dump E! directory for later analysis by the IBM Support Center.

E! The default dump directory for GPFS is **/tmp/mmfs**. This directory might disappear on Linux if cron is E! set to run the **/etc/cron.daily/tmpwatch** script. The **tmpwatch** script removes files and directories in **/tmp** E! that have not been accessed recently. Administrators who want to use a different directory for GPFS E! dumps can change the directory by issuing this command:

```
E! mmchconfig dataStructureDump=/name_of_some_other_big_file_system
```

E! **Note:** This state information (possibly large amounts of data in the form of GPFS dumps and traces) can E! be dumped automatically as part of the first failure data capture mechanisms of GPFS, and can accumulate E! in the (default **/tmp/mmfs**) directory that is defined by the **dataStructureDump** configuration parameter. It E! is recommended that a cron job (such as **/etc/cron.daily/tmpwatch**) be used to remove E! **dataStructureDump** directory data that is older than two weeks, and that such data is collected (for E! example, via **gpfs.snap**) within two weeks of encountering any problem that requires investigation.

## E! **Why does my system load increase significantly during the night?**

E! On some Linux distributions, cron runs the **/etc/cron.daily/slocate.cron** job every night. This will try to E! index all the files in GPFS. This will put a very large load on the GPFS token manager.

E! You can exclude all GPFS file systems by adding **gpfs** to the **excludeFileSytemType** list in this script, or E! exclude specific GPFS file systems in the **excludeFileSytemType** list.

```
E! /usr/bin/updatedb -f "excludeFileSystemType" -e "excludeFileSystem"
```

E! If indexing GPFS file systems is desired, only one node should run the **updatedb** command and build the E! database in a GPFS file system. If the database is built within a GPFS file system it will be visible on all E! nodes after one node finishes building it.

## E! **What do I do if I receive message 6027-648?**

E! The **mmedquota** or **mmdefedquota** commands can fail with message **6027-648: EDITOR environment** E! **variable must be full path name**.

E! To resolve this error, do the following:

E! 1. Change the value of the EDITOR environment variable to an absolute path name.

E! 2. Check to see if the EDITOR variable is set in the **\$HOME/.kshrc** file. If it is set, check to see if it is an E! absolute path name because the **mmedquota** or **mmdefedquota** command could retrieve the EDITOR E! environment variable from that file.

## E! **Why can't I see my newly mounted Windows file system?**

E! On Windows, a newly mounted file system might not be visible to you if you are currently logged on to E! a system. This can happen if you have mapped a network share to the same drive letter as GPFS.

E! Once you start a new session (by logging out and logging back in), the use of the GPFS drive letter will E! supersede any of your settings for the same drive letter. This is standard behavior for all local file E! systems on Windows.

## E! **Why is the file system mounted on the wrong drive letter?**

E! Before mounting a GPFS file system, you must be certain that the drive letter required for GPFS is freely E! available and is not being used by a local disk or a network-mounted file system on *all* computation E! nodes where the GPFS file system will be mounted.

## E! **Why does the offline mmfsck command fail with "Error creating internal storage"?**

E! Use **mmfsck** command on the file system manager for storing internal data during a file system scan.  
E! The command fails if the GPFS fails to provide a temporary file of the required size.

E! The **mmfsck** command requires some temporary space on the file system manager for storing internal data during a file system scan. The internal data will be placed in the directory specified by the **mmfsck** **-t** command line parameter (**/tmp** by default). The amount of temporary space that is needed is proportional to the number of inodes (used and unused) in the file system that is being scanned. If GPFS is unable to create a temporary file of the required size, the **mmfsck** command will fail with the following error message:

E! Error creating internal storage

E! This failure could be caused by:

- E! • The lack of sufficient disk space in the temporary directory on the file system manager
- E! • The lack of sufficient pagepool on the file system manager as shown in **mmlsconfig pagepool** output
- E! • Insufficiently high **filesize** limit set for the **root** user by the operating system
- E! • The lack of support for large files in the file system that is being used for temporary storage. Some file systems limit the maximum file size because of architectural constraints. For example, JFS on AIX does not support files larger than 2 GB, unless the **Large file support** option has been specified when the file system was created. Check local operating system documentation for maximum file size limitations.

## E! **Why do I get timeout executing function error message?**

E! If any of the commands fails due to timeout while executing **mmccr**, rerun the command to fix the issue.  
E! This timeout issue is likely related to an increased workload of the system.

## E! **Questions related to active file management**

E! Issues and explanations pertaining to active file management.

E! The following questions are related to active file management (AFM).

### E! **How can I change the mode of a fileset?**

E! The mode of an AFM client cache fileset cannot be changed from local-update mode to any other mode; however, it can be changed from read-only to single-writer (and vice versa), and from either read-only or single-writer to local-update.

E! To change the mode, do the following:

- E! 1. Ensure that fileset status is active and that the gateway is available.
- E! 2. Umount the file system.
- E! 3. Unlink the fileset.
- E! 4. Run the **mmchfileset** command to change the mode.
- E! 5. Mount the file system again.
- E! 6. Link the fileset again.

### E! **Why are setuid/setgid bits in a single-writer cache reset at home after data is appended?**

E! The setuid/setgid bits in a single-writer cache are reset at home after data is appended to files on which those bits were previously set and synced. This is because over NFS, a write operation to a setuid file resets the setuid bit.

### **E! How can I traverse a directory that has not been cached?**

E! On a fileset whose metadata in all subdirectories is not cached, any application that optimizes by assuming that directories contain two fewer subdirectories than their hard link count will not traverse the last subdirectory. One such example is **find**; on Linux, a workaround for this is to use **find -noleaf** to correctly traverse a directory that has not been cached.

### **E! What extended attribute size is supported?**

E! For an operating system in the gateway whose Linux kernel version is below 2.6.32, the NFS max rsize is 32K, so AFM would not support an extended attribute size of more than 32K on that gateway.

### **E! What should I do when my file system or fileset is getting full?**

E! The **.ptrash** directory is present in cache and home. In some cases, where there is a conflict that AFM cannot resolve automatically, the file is moved to **.ptrash** at cache or home. In cache the **.ptrash** gets cleaned up when eviction is triggered. At home, it is not cleared automatically. When the administrator is looking to clear some space, the **.ptrash** should be cleaned up first.



---

## Chapter 10. Understanding the system limitations

It is important to understand the system limitations to analyze whether you are facing a real issue in the IBM Spectrum Scale system.

The following topics list the IBM Spectrum Scale system limitations:

### **AFM limitations**

See *AFM limitations* in *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.

### **AFM-based DR limitations**

See *AFM-based DR limitations* in *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.

### **Authentication limitations**

See *Authentication limitations* in *IBM Spectrum Scale: Administration Guide*.

### **File authorization limitations**

See *Authorization limitations* in *IBM Spectrum Scale: Administration Guide*.

### **File compression limitations**

See *File compression* in *IBM Spectrum Scale: Administration Guide*.

### **FPO limitations**

See *Restrictions* in *IBM Spectrum Scale: Administration Guide*.

### **General NFS V4 Linux Exceptions and Limitations**

See *General NFS V4 Linux exceptions and limitations* in *IBM Spectrum Scale: Administration Guide*.

### **GPFS exceptions and limitations to NFSv4 ACLs**

See *GPFS exceptions and limitations to NFS V4 ACLs* in *IBM Spectrum Scale: Administration Guide*.

### **GUI limitations**

See *GUI limitations* in *IBM Spectrum Scale: Administration Guide*.

### **HDFS transparency limitations**

See *Limitation and difference from HDFS* in *IBM Spectrum Scale: Administration Guide*.

### **HDFS transparency federation limitations**

See *Known limitations* in *IBM Spectrum Scale: Administration Guide*.

### **mmuserauth service create command limitations**

See *Limitations of the mmuserauth service create command while configuring AD with RFC2307* in *IBM Spectrum Scale: Administration Guide*.

### **Multiprotocol export limitations**

See *Multiprotocol export considerations* in *IBM Spectrum Scale: Administration Guide*.

### **Performance monitoring limitations**

See *Performance monitoring limitations* in *IBM Spectrum Scale: Administration Guide*.

### **Protocol cluster disaster recovery limitations**

See *Protocols cluster disaster recovery limitations* in *IBM Spectrum Scale: Administration Guide*.

### **Protocol data security limitations**

See *Data security limitations* in *IBM Spectrum Scale: Administration Guide*.

### **S3 API support limitations**

See *Managing OpenStack access control lists using S3 API* in *IBM Spectrum Scale: Administration Guide*.

### **SMB share limitations**

See *SMB share limitations* in *IBM Spectrum Scale: Administration Guide*.

**Transparent Cloud Tiering limitations**

*See Known limitations of Transparent Cloud Tiering in IBM Spectrum Scale: Administration Guide.*

**Unified file and object access limitations**

*See Limitations of unified file and object access in IBM Spectrum Scale: Administration Guide.*

---

## Chapter 11. Collecting details of the issues

You need to collect the details of the issues that are reported in the system to start the troubleshooting process.

The IBM Spectrum Scale system provides the following options to collect the details of the issues reported in the system:

- Logs
- Dumps
- Traces
- Diagnostic data collection through CLI
- Diagnostic data collection through GUI

---

### E! Collecting details of issues by using logs, dumps, and traces

The problem determination tools that are provided with IBM Spectrum Scale are intended to be used by experienced system administrators who know how to collect data and run debugging routines.

You can collect various types of logs such as GPFS logs, protocol service logs, operating system logs, and transparent cloud tiering logs. The GPFS™ log is a repository of error conditions that are detected on each node, as well as operational events such as file system mounts. The operating system error log is also useful because it contains information about hardware failures and operating system or other software failures that can affect the IBM Spectrum Scale system.

**Note:** The GPFS error logs and messages contain the MMFS prefix to distinguish it from the components of the IBM Multi-Media LAN Server, a related licensed program.

The IBM Spectrum Scale system also provides a system snapshot dump, trace, and other utilities that can be used to obtain detailed information about specific problems.

The information is organized as follows:

- “GPFS logs” on page 140
- E! • “Operating system error logs” on page 157
- “Using the gpfs.snap command” on page 176
- “mmdumpperfdata command” on page 184
- “mmfsadm command” on page 186
- “Trace facility” on page 162

### Time stamp in GPFS log entries

The time stamp in a GPFS log entry indicates the time of an event.

#### Time stamp format

The time stamp in the GPFS log has the following format:

```
Www Mmm DD hh:mm:ss.sss YYYY
```

where

*Www* is a three-character abbreviation for the name of the day of the week.

*Mmm* is a three-character abbreviation for the name of the month.  
*DD* is the day of the month.  
*hh:mm:sec* is the hours (24-hour clock), minutes, seconds, and milliseconds.  
*YYYY* is the year

The following examples show the earlier time stamp format:

Mon May 09 15:12:20.603 2016

Sun Aug 15 07:04:33.078 2016

## E! Logs

E! This topic describes various logs that are generated in the IBM Spectrum Scale.

### E! GPFS logs

E! The GPFS log is a repository of error conditions that are detected on each node, as well as operational events such as file system mounts. The GPFS log is the first place to look when you start debugging the abnormal events. As GPFS is a cluster file system, events that occur on one node might affect system behavior on other nodes, and all GPFS logs can have relevant data.

E! The GPFS log can be found in the */var/adm/ras* directory on each node. The GPFS log file is named *mmfs.log.date.nodeName*, where *date* is the time stamp when the instance of GPFS started on the node and *nodeName* is the name of the node. The latest GPFS log file can be found by using the symbolic file name */var/adm/ras/mmfs.log.latest*.

E! The GPFS log from the prior startup of GPFS can be found by using the symbolic file name */var/adm/ras/mmfs.log.previous*. All other files have a time stamp and node name appended to the file name.

E! At GPFS startup, log files that are not accessed during the last 10 days are deleted. If you want to save old log files, copy them elsewhere.

E! Many GPFS log messages can be sent to **syslog** on Linux. The **systemLogLevel** attribute of the **mmchconfig** command determines the GPFS log messages to be sent to the **syslog**. For more information, see the **mmchconfig** command in the *IBM Spectrum Scale: Command and Programming Reference*.

E! This example shows normal operational messages that appear in the GPFS log file on Linux node:

```
E! Removing old /var/adm/ras/mmfs.log.* files:
E! Unloading modules from /lib/modules/3.0.13-0.27-default/extra
E! Unloading module tracedev
E! Loading modules from /lib/modules/3.0.13-0.27-default/extra
E! Module          Size Used by
E! mmfs26          2155186 0
E! mmfslinux       379348 1 mmfs26
E! tracedev        48513 2 mmfs26,mmfslinux
E! Tue Oct 27 11:45:47.149 2015: [I] mmfsd initializing. {Version: 4.2.0.0 Built: Oct 26 2015 15:19:01} ...
E! Tue Oct 27 11:45:47.150 2015: [I] Tracing in blocking mode
E! Tue Oct 27 11:45:47.151 2015: [I] Cleaning old shared memory ...
E! Tue Oct 27 11:45:47.152 2015: [I] First pass parsing mmfs.cfg ...
E! Tue Oct 27 11:45:47.153 2015: [I] Enabled automated deadlock detection.
E! Tue Oct 27 11:45:47.154 2015: [I] Enabled automated deadlock debug data collection.
E! Tue Oct 27 11:45:47.155 2015: [I] Enabled automated expel debug data collection.
E! Tue Oct 27 11:45:47.156 2015: [I] Initializing the main process ...
E! Tue Oct 27 11:45:47.169 2015: [I] Second pass parsing mmfs.cfg ...
E! Tue Oct 27 11:45:47.170 2015: [I] Initializing the page pool ...
E! Tue Oct 27 11:45:47.500 2015: [I] Initializing the mailbox message system ...
E! Tue Oct 27 11:45:47.521 2015: [I] Initializing encryption ...
E! Tue Oct 27 11:45:47.522 2015: [I] Encryption: loaded crypto library: IBM
E! CryptoLite for C v4.10.1.5600 (c4T3/GPFSLNXPPC64).
E! Tue Oct 27 11:45:47.523 2015: [I] Initializing the thread system ...
E! Tue Oct 27 11:45:47.524 2015: [I] Creating threads ...
E! Tue Oct 27 11:45:47.529 2015: [I] Initializing inter-node communication ...
```

```

E! Tue Oct 27 11:45:47.530 2015: [I] Creating the main SDR server object ...
E! Tue Oct 27 11:45:47.531 2015: [I] Initializing the sdrServ library ...
E! Tue Oct 27 11:45:47.532 2015: [I] Initializing the ccrServ library ...
E! Tue Oct 27 11:45:47.538 2015: [I] Initializing the cluster manager ...
E! Tue Oct 27 11:45:48.813 2015: [I] Initializing the token manager ...
E! Tue Oct 27 11:45:48.819 2015: [I] Initializing network shared disks ...
E! Tue Oct 27 11:45:51.126 2015: [I] Start the ccrServ ...
E! Tue Oct 27 11:45:51.879 2015: [N] Connecting to 192.168.115.171 js21n07 <c0p1>
E! Tue Oct 27 11:45:51.880 2015: [I] Connected to 192.168.115.171 js21n07 <c0p1>
E! Tue Oct 27 11:45:51.897 2015: [I] Node 192.168.115.171 (js21n07) is now the Group Leader.
E! Tue Oct 27 11:45:51.911 2015: [N] mmfsd ready
E! Tue Oct 27 11:45:52 EDT 2015: mmcommon mmfsup invoked.
E! Parameters: 192.168.115.220 192.168.115.171 all

```

E! The **mmcommon logRotate** command can be used to rotate the GPFS log without shutting down and restarting the daemon. After the **mmcommon logRotate** command is issued, **/var/adm/ras/mmfs.log.previous** will contain the messages that occurred since the previous startup of GPFS or the last run of **mmcommon logRotate**. The **/var/adm/ras/mmfs.log.latest** file starts over at the point in time that **mmcommon logRotate** was run.

E! Depending on the size and complexity of your system configuration, the amount of time to start GPFS varies. If you cannot access a file system that is mounted, examine the log file for error messages.

#### E! **Creating a master GPFS log file:**

E! The GPFS log frequently shows problems on one node that actually originated on another node.

E! GPFS is a file system that runs on multiple nodes of a cluster. This means that problems originating on one node of a cluster often have effects that are visible on other nodes. It is often valuable to merge the GPFS logs in pursuit of a problem. Having accurate time stamps aids the analysis of the sequence of events.

E! Before following any of the debug steps, IBM suggests that you:

E! 1. Synchronize all clocks of all nodes in the GPFS cluster. If this is not done, and clocks on different nodes are out of sync, there is no way to establish the real time line of events occurring on multiple nodes. Therefore, a merged error log is less useful for determining the origin of a problem and tracking its effects.

E! 2. Merge and chronologically sort all of the GPFS log entries from each node in the cluster. The **--gather-logs** option of the **gpfs.snap** command can be used to achieve this:

```
E! gpfs.snap --gather-logs -d /tmp/logs -N all
```

E! The system displays information similar to:

```
E! gpfs.snap: Gathering mmfs logs ...
```

```
E! gpfs.snap: The sorted and unsorted mmfs.log files are in /tmp/logs
```

E! If the **--gather-logs** option is not available on your system, you can create your own script to achieve the same task; use **/usr/lpp/mmfs/samples/gatherlogs.samples.sh** as an example.

#### E! **Audit messages for cluster configuration changes**

E! As an aid to troubleshooting and to improve cluster security, IBM Spectrum Scale can send an audit message to syslog and the GPFS log whenever a GPFS command changes the configuration of the cluster.

E! You can use the features of syslog to mine, process, or redirect the audit messages.

E! **Restriction:** Audit messages are not available on Windows operating systems.

## E! **Configuring syslog**

E! On Linux operating systems, syslog typically is enabled by default. On AIX, syslog must be set up and E! configured. Check your operating system and its documentation for details.

## E! **Configuring audit messages**

E! By default, audit messages are enabled and messages are sent to syslog but not to the GPFS log. You can E! control audit messages with the **commandAudit** attribute of the **mmchconfig** command. For more E! information, see the topic *mmchconfig command* in the *IBM Spectrum Scale: Command and Programming E! Reference* guide.

E! Audit messages are not affected by the **systemLogLevel** attribute of the **mmchconfig** command.

E! If audit logs are enabled, the GUI receives the updates on configuration changes that you made through E! CLI and updates its configuration cache to reflect the changes in the GUI. You can also disable audit E! logging with the **mmchconfig** command. If the audit logs are disabled, the GUI does not show the E! configuration changes immediately and might be as much as an hour late in reflecting configuration E! changes that are made through the CLI .

## E! **Message format**

E! For security, the text of sensitive information such as a password is replaced with asterisks (\*) in the audit E! message.

E! Audit messages are sent to syslog with an identity of **mmfs**, a facility code of **user**, and a severity level of E! **informational**.

E! The format of the message depends on the source of the GPFS command:

E! •

E! Messages about GPFS commands that are entered at the command line have the following format:

E! **CLI** *user\_name user\_name* [*AUDIT\_TYPE1,AUDIT\_TYPE2*] '*command*' **RC**=*return\_code*

E! where:

E! **CLI** The source of the command. Indicates that the command was entered from the command line.

E! *user\_name user\_name*

E! The name of the user who entered the command, such as root. The same name appears twice.

E! *AUDIT\_TYPE1*

E! The point in the command when the message was sent to syslog. Always **EXIT**.

E! *AUDIT\_TYPE2*

E! The action taken by the command. Always **CHANGE**.

E! *command*

E! The text of the command.

E! *return\_code*

E! The return code of the GPFS command.

E! •

E! Messages about GPFS commands that are issued by GUI commands have a similar format:

E! **GUI-CLI** *user\_name GUI\_user\_name* [*AUDIT\_TYPE1,AUDIT\_TYPE2*] '*command*' **RC**=*return\_code*

E! where:

E! **GUI-CLI**  
E! The source of the command. Indicates that the command was called by a GUI command.  
E! *user\_name*  
E! The name of the user, such as root.  
E! *GUI\_user\_name*  
E! The name of the user who logged on to the GUI.  
E! The remaining fields are the same as in the **CLI** message.

E! The following lines are examples from a syslog:

```
E! Apr 24 13:56:26 c12c3apv12 mmfs[63655]: CLI root root [EXIT, CHANGE] 'mmchconfig  
E! autoload=yes' RC=0  
E! Apr 24 13:58:42 c12c3apv12 mmfs[65315]: CLI root root [EXIT, CHANGE] 'mmchconfig  
E! deadlockBreakupDelay=300' RC=0  
E! Apr 24 14:04:47 c12c3apv12 mmfs[67384]: CLI root root [EXIT, CHANGE] 'mmchconfig  
E! FIPS1402mode=no' RC=0
```

E! The following lines are examples from a syslog where GUI is the originator:

```
E! Apr 24 13:56:26 c12c3apv12 mmfs[63655]: GUI-CLI root admin [EXIT, CHANGE] 'mmchconfig autoload=yes' RC=0
```

## E! **Commands**

E! The product sends audit messages to syslog for the following commands and options:

- E! **mmaddcallback**
- E! **mmadddisk**
- E! **mmaddnode**
- E! **mmafmconfig add**
- E! **mmafmconfig delete**
- E! **mmafmconfig disable**
- E! **mmafmconfig enable**
- E! **mmafmconfig update**
- E! **mmafmctl**
- E! **mmapplypolicy**
- E! **mmauth add**
- E! **mmauth delete**
- E! **mmauth deny**
- E! **mmauth gencert**
- E! **mmauth genkey**
- E! **mmauth grant**
- E! **mmauth update**
- E! **mmbackup**
- E! **mmbackupconfig**
- E! **mmces address add**
- E! **mmces address change**
- E! **mmces address move**
- E! **mmces address remove**
- E! **mmces log**
- E! **mmces node resume**
- E! **mmces node suspent**

E! mmces service disable  
E! mmces service enable  
E! mmces service start  
E! mmces service stop  
E! mmcesdr  
E! mmcesmonitor  
E! mmchcluster  
E! mmchconfig  
E! mmchdisk  
E! mmchfileset  
E! mmchfs  
E! mmchlicense  
E! mmchmgr  
E! mmchnode  
E! mmchnodeclass  
E! mmchnsd  
E! mmchpolicy  
E! mmchpool  
E! mmchqos  
E! mmcloudgateway account create  
E! mmcloudgateway account delete  
E! mmcloudgateway account update  
E! mmcloudgateway config set  
E! mmcloudgateway config unset  
E! mmcloudgateway files delete  
E! mmcloudgateway files migrate  
E! mmcloudgateway files recall  
E! mmcloudgateway files reconcile  
E! mmcloudgateway files restore  
E! mmcloudgateway filesystem create  
E! mmcloudgateway filesystem delete  
E! mmcloudgateway service start  
E! mmcloudgateway service stop  
E! mmcrcluster  
E! mmcrfileset  
E! mmcrfs  
E! mmcrnodeclass  
E! mmcrnsd  
E! mmcrsnapshot  
E! mmdefedquota  
E! mmdefquotaoff  
E! mmdefquotaon  
E! mmdefragfs  
E! mmdelcallback  
E! mmdeldisk



E! mmdelfileset  
E! mmdelfs  
E! mmdelnode  
E! mmdelnodeclass  
E! mmdelnsd  
E! mmdelsnapshot  
E! mmedquota  
E! mmexpelnode  
E! mmexportfs  
E! mmfsctl  
E! mmimgbackup  
E! mmimgrestore  
E! mmimportfs  
| mmkeyserv  
| mmlinkfileset  
| mmmigratefs  
| mmnfs configuration change  
| mmnfs export add  
| mmnfs export change  
| mmnfs export load  
| mmnfs export remove  
| mmnsddiscover  
| mmobj config change  
| mmobj file access  
| mmobj multiregion enable  
| mmobj multiregion export  
| mmobj multiregion import  
| mmobj multiregion remove  
| mmobj policy change  
| mmobj policy create  
| mmobj policy deprecate  
| mmobj swift base  
| mmperfmon config add  
| mmperfmon config delete  
| mmperfmon config generate  
| mmperfmon config update  
| mmpsnap create  
| mmpsnap delete  
| mmquotaoff  
| mmquotaon  
| mmremoteclass add  
| mmremoteclass delete  
| mmremoteclass update  
| mmremotefs add  
| mmremotefs delete

- | **mmremotefs update**
- | **mmrestoreconfig**
- | **mmrestorefs**
- | **mmrestripefile**
- | **mmrestripefs**
- | **mmrpldisk**
- | **mmsdrrestore**
- | **mmsetquota**
- | **mmshutdown**
- | **mmsmb config change**
- | **mmsmb export add**
- | **mmsmb export change**
- | **mmsmb export remove**
- | **mmsmb exportacl add**
- | **mmsmb exportacl change**
- | **mmsmb exportacl delete**
- | **mmsmb exportacl remove**
- | **mmsmb exportacl replace**
- | **mmsnapdir**
- | **mmstartup**
- | **mmumount**
- | **mmumount**
- | **mmunlinkfileset**
- | **mmuserauth service create**
- | **mmuserauth service remove**
- | **mmwinservctl**

## | **Protocol services logs**

E! The protocol service logs contains the information that helps you to troubleshoot the issues related to the E! NFS, SMB, and Object services.

E! By default, the NFS, SMB, and Object protocol logs are stored at: `/var/log/messages`.

E! For more information on logs of the **spectrumscale** installation toolkit, see *Logging and debugging* in *IBM E! Spectrum Scale: Concepts, Planning, and Installation Guide*.

### E! **SMB logs:**

E! The SMB services write the most important messages to `syslog`.

E! With the standard syslog configuration, you can search for the terms such as `ctdbd` or `smbd` in the E! `/var/log/messages` file to see the relevant logs. For example:

E! **grep ctdbd /var/log/messages**

E! The system displays output similar to the following example:

```
E! May 31 09:11:23 prt002st001 ctdbd: Updated hot key database=locking.tdb key=0x2795c3b1 id=0 hop_count=1
E! May 31 09:27:33 prt002st001 ctdbd: Updated hot key database=smbXsrv_open_global.tdb key=0x0d0d4abe id=0 hop_count=1
E! May 31 09:37:17 prt002st001 ctdbd: Updated hot key database=brlock.tdb key=0xc37fe57c id=0 hop_count=1
```

E! **grep smbd /var/log/messages**



## E! NFS logs:

E! The clustered export services (CES) NFS server writes log messages in the `/var/log/ganesha.log` file at E! runtime.

E! Operating system's log rotation facility is used to manage NFS logs. The NFS logs are configured and E! enabled during the NFS server packages installation.

E! The following example shows a sample log file:

```
E! # tail -f /var/log/ganesha.log
E! 2015-05-31 17:08:04 : epoch 556b23d2 : cluster1.ibm.com : ganesha.nfsd-27204[main] nfs_rpc_cb_init_ccache
E! :NFS STARTUP :WARN :gssd_refresh_krb5_machine_credential failed (-1765328160:0)
E! 2015-05-31 17:08:04 : epoch 556b23d2 : cluster1.ibm.com : ganesha.nfsd-27204[main] nfs_Start_threads
E! :THREAD :EVENT :Starting delayed executor.
E! 2015-05-31 17:08:04 : epoch 556b23d2 : cluster1.ibm.com : ganesha.nfsd-27204[main] nfs_Start_threads
E! :THREAD :EVENT :gsh_dbusthread was started successfully
E! 2015-05-31 17:08:04 : epoch 556b23d2 : cluster1.ibm.com : ganesha.nfsd-27204[main] nfs_Start_threads
E! :THREAD :EVENT :admin thread was started successfully
E! 2015-05-31 17:08:04 : epoch 556b23d2 : cluster1.ibm.com : ganesha.nfsd-27204[main] nfs_Start_threads
E! :THREAD :EVENT :reaper thread was started successfully
E! 2015-05-31 17:08:04 : epoch 556b23d2 : cluster1.ibm.com : ganesha.nfsd-27204[main] nfs_Start_threads
E! :THREAD :EVENT :General fridge was started successfully
E! 2015-05-31 17:08:04 : epoch 556b23d2 : cluster1.ibm.com : ganesha.nfsd-27204[reaper] nfs_in_grace
E! :STATE :EVENT :NFS Server Now IN GRACE
E! 2015-05-31 17:08:04 : epoch 556b23d2 : cluster1.ibm.com : ganesha.nfsd-27204[main] nfs_start
E! :NFS STARTUP :EVENT :-----
E! 2015-05-31 17:08:04 : epoch 556b23d2 : cluster1.ibm.com : ganesha.nfsd-27204[main] nfs_start
E! :NFS STARTUP :EVENT :           NFS SERVER INITIALIZED
E! 2015-05-31 17:08:04 : epoch 556b23d2 : cluster1.ibm.com : ganesha.nfsd-27204[main] nfs_start
E! :NFS STARTUP :EVENT :-----
E! 2015-05-31 17:09:04 : epoch 556b23d2 : cluster1.ibm.com : ganesha.nfsd-27204[reaper] nfs_in_grace
E! :STATE :EVENT :NFS Server Now NOT IN GRACE
```

E! Log levels can be displayed by using the `mmnfs configuration list | grep LOG_LEVEL` command. For E! example:

```
E! mmnfs configuration list | grep LOG_LEVEL
```

E! The system displays output similar to the following example:

```
E! LOG_LEVEL: EVENT
```

E! By default, the log level is EVENT. Additionally, the following NFS log levels can also be used; starting E! from lowest to highest verbosity:

E! • FATAL

E! • MAJ

E! • CRIT

E! • WARN

E! • INFO

E! • DEBUG

E! • MID\_DEBUG

E! • FULL\_DEBUG

E! **Note:** The FULL\_DEBUG level increases the size of the log file. Use it in the production mode only if E! instructed by the IBM Support.

E! Increasing the verbosity of the NFS server log impacts the overall NFS I/O performance.

E! To change the logging to the verbose log level INFO, use the following command:

## E! **mmnfs configuration change LOG\_LEVEL=INFO**

E! The system displays output similar to the following example:

E! NFS Configuration successfully changed. NFS server restarted on all NFS nodes on which NFS server is running.

E! This change is cluster-wide and restarts all NFS instances to activate this setting. The log file now displays more informational messages, for example:

```
E! 2015-06-03 12:49:31 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[main] nfs_rpc_dispatch_threads
E! :THREAD :INFO :5 rpc dispatcher threads were started successfully
E! 2015-06-03 12:49:31 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[disp] rpc_dispatcher_thread
E! :DISP :INFO :Entering nfs/rpc dispatcher
E! 2015-06-03 12:49:31 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[disp] rpc_dispatcher_thread
E! :DISP :INFO :Entering nfs/rpc dispatcher
E! 2015-06-03 12:49:31 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[disp] rpc_dispatcher_thread
E! :DISP :INFO :Entering nfs/rpc dispatcher
E! 2015-06-03 12:49:31 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[disp] rpc_dispatcher_thread
E! :DISP :INFO :Entering nfs/rpc dispatcher
E! 2015-06-03 12:49:31 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[main] nfs_Start_threads
E! :THREAD :EVENT :gsh_dbusthread was started successfully
E! 2015-06-03 12:49:31 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[main] nfs_Start_threads
E! :THREAD :EVENT :admin thread was started successfully
E! 2015-06-03 12:49:31 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[main] nfs_Start_threads
E! :THREAD :EVENT :reaper thread was started successfully
E! 2015-06-03 12:49:31 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[main] nfs_Start_threads
E! :THREAD :EVENT :General fridge was started successfully
E! 2015-06-03 12:49:31 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[reaper] nfs_in_grace
E! :STATE :EVENT :NFS Server Now IN GRACE
E! 2015-06-03 12:49:32 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[main] nfs_start
E! :NFS STARTUP :EVENT :-----
E! 2015-06-03 12:49:32 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[main] nfs_start
E! :NFS STARTUP :EVENT : NFS SERVER INITIALIZED
E! 2015-06-03 12:49:32 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[main] nfs_start
E! :NFS STARTUP :EVENT :-----
E! 2015-06-03 12:50:32 : epoch 556edba9 : cluster1.ibm.com : ganesha.nfsd-21582[reaper] nfs_in_grace
E! :STATE :EVENT :NFS Server Now NOT IN GRACE
```

E! To display the currently configured CES log level, use the following command:

## E! **mmces log level**

E! The system displays output similar to the following example:

E! CES log level is currently set to 0

E! The log file is `/var/adm/ras/mnfs.log.latest`. By default, the log level is 0 and other possible values are 1, 2, and 3. To increase the log level, use the following command:

## E! **mmces log level 1**

E! NFS-related log information is written to the standard GPFS log files as part of the overall CES infrastructure. This information relates to the NFS service management and recovery orchestration within CES.

## E! **Object logs:**

E! There are a number of locations where messages are logged with the object protocol.

E! The core object services, proxy, account, container, and object server have their own logging level sets in their respective configuration files. By default, unified file and object access logging is set to show messages at or above the ERROR level, but can be changed to INFO or DEBUG levels if more detailed logging information is required.

E! By default, the messages logged by these services are saved in the `/var/log/swift` directory.

E! You can also configure these services to use separate syslog facilities by the **log\_facility** parameter in E! one or all of the object service configuration files and by updating the rsyslog configuration. These E! parameters are described in the Swift Deployment Guide([docs.openstack.org/developer/swift/deployment\\_guide.html](https://docs.openstack.org/developer/swift/deployment_guide.html)) that is available in the OpenStack documentation.

E! An example of how to set up this configuration can be found in the SAIO - Swift All In One E! documentation([docs.openstack.org/developer/swift/development\\_saio.html#optional-setting-up-rsyslog-for-individual-logging](https://docs.openstack.org/developer/swift/development_saio.html#optional-setting-up-rsyslog-for-individual-logging)) that is available in the OpenStack documentation.

E! **Note:** To configure rsyslog for unique log facilities in the protocol nodes, the administrator needs to E! ensure that the manual steps mentioned in the preceding link are carried out on each of those protocol E! nodes.

E! The Keystone authentication service writes its logging messages to `/var/log/keystone/keystone.log` file. E! By default, Keystone logging is set to show messages at or above the WARNING level.

E! For information on how to view or change log levels on any of the object related services, see *CES tracing and debug data collection* in *IBM Spectrum Scale: Problem Determination Guide*.

E! The following commands can be used to determine the health of object services:

E! • To see whether there are any nodes in an active (failed) state, run the following command:

E! **mmces state cluster OBJ**

E! The system displays output similar to this:

E! NODE	E! COMPONENT	E! STATE	E! EVENTS
E! prt001st001	E! OBJECT	E! HEALTHY	
E! prt002st001	E! OBJECT	E! HEALTHY	
E! prt003st001	E! OBJECT	E! HEALTHY	
E! prt004st001	E! OBJECT	E! HEALTHY	
E! prt005st001	E! OBJECT	E! HEALTHY	
E! prt006st001	E! OBJECT	E! HEALTHY	
E! prt007st001	E! OBJECT	E! HEALTHY	

E! In this example, all nodes are healthy so no active events are shown.

E! • To display the history of events generated by the monitoring framework, run the following command:

E! **mmces events list OBJ**

E! The system displays output similar to this:

E! Node	E! Timestamp	E! Event Name	E! Severity	E! Details
E! node1	E! 2015-06-03 13:30:27.478725+08:08PDT	E! proxy-server_ok	E! INFO	E! proxy process as expected
E! node1	E! 2015-06-03 14:26:30.567245+08:08PDT	E! object-server_ok	E! INFO	E! object process as expected
E! node1	E! 2015-06-03 14:26:30.720534+08:08PDT	E! proxy-server_ok	E! INFO	E! proxy process as expected
E! node1	E! 2015-06-03 14:28:30.689257+08:08PDT	E! account-server_ok	E! INFO	E! account process as expected
E! node1	E! 2015-06-03 14:28:30.853518+08:08PDT	E! container-server_ok	E! INFO	E! container process as expected
E! node1	E! 2015-06-03 14:28:31.015307+08:08PDT	E! object-server_ok	E! INFO	E! object process as expected
E! node1	E! 2015-06-03 14:28:31.177589+08:08PDT	E! proxy-server_ok	E! INFO	E! proxy process as expected
E! node1	E! 2015-06-03 14:28:49.025021+08:08PDT	E! postIpChange_info	E! INFO	E! IP addresses modified 192.167.12.21_0_1.
E! node1	E! 2015-06-03 14:28:49.194499+08:08PDT	E! enable_Address_database_node	E! INFO	E! Enable Address Database Node
E! node1	E! 2015-06-03 14:29:16.483623+08:08PDT	E! postIpChange_info	E! INFO	E! IP addresses modified 192.167.12.22_0_2.
E! node1	E! 2015-06-03 14:29:25.274924+08:08PDT	E! postIpChange_info	E! INFO	E! IP addresses modified 192.167.12.23_0_3.
E! node1	E! 2015-06-03 14:29:30.844626+08:08PDT	E! postIpChange_info	E! INFO	E! IP addresses modified 192.167.12.24_0_4.

E! • To retrieve the OBJ related log entries, query the monitor client and grep for the name of the E! component you want to filter on, either object, proxy, account, container, keystone or postgres. For E! example, to see proxy-server related events, run the following command:

E! **mmces events list | grep proxy**

E! The system displays output similar to this:

E! node1	E! 2015-06-01 14:39:49.120912+08:08PDT	E! proxy-server_failed	E! ERROR	E! proxy process should be started but is stopped
E! node1	E! 2015-06-01 14:44:49.277940+08:08PDT	E! proxy-server_ok	E! INFO	E! proxy process as expected
E! node1	E! 2015-06-01 16:27:37.923696+08:08PDT	E! proxy-server_failed	E! ERROR	E! proxy process should be started but is stopped
E! node1	E! 2015-06-01 16:40:39.789920+08:08PDT	E! proxy-server_ok	E! INFO	E! proxy process as expected
E! node1	E! 2015-06-03 13:28:18.875566+08:08PDT	E! proxy-server_failed	E! ERROR	E! proxy process should be started but is stopped
E! node1	E! 2015-06-03 13:30:27.478725+08:08PDT	E! proxy-server_ok	E! INFO	E! proxy process as expected
E! node1	E! 2015-06-03 13:30:57.482977+08:08PDT	E! proxy-server_failed	E! ERROR	E! proxy process should be started but is stopped

```

E! node1 2015-06-03 14:26:30.720534+08:08PDT proxy-server_ok INFO proxy process as expected
E! node1 2015-06-03 14:27:00.759696+08:08PDT proxy-server_failed ERROR proxy process should be started but is stopped
E! node1 2015-06-03 14:28:31.177589+08:08PDT proxy-server_ok INFO proxy process as expected

```

E! • To check the monitor log, grep for the component you want to filter on, either object, proxy, account, container, keystone or postgres. For example, to see object-server related log messages:

```
E! grep object /var/adm/ras/mmcesmonitor.log | head -n 10
```

E! The system displays output similar to this:

```

E! 2015-06-03T13:59:28.805-08:00 util5.sonasad.almaden.ibm.com D:522632:Thread-9:object:OBJ running command
E! 'systemctl status openstack-swift-proxy'
E! 2015-06-03T13:59:28.916-08:00 util5.sonasad.almaden.ibm.com D:522632:Thread-9:object:OBJ command result
E! ret:3 sout:openstack-swift-proxy.service - OpenStack Object Storage (swift) - Proxy Server
E! 2015-06-03T13:59:28.916-08:00 util5.sonasad.almaden.ibm.com I:522632:Thread-9:object:OBJ openstack-swift-proxy is not started, ret3
E! 2015-06-03T13:59:28.916-08:00 util5.sonasad.almaden.ibm.com D:522632:Thread-9:object:OBJProcessMonitor openstack-swift-proxy failed:
E! 2015-06-03T13:59:28.916-08:00 util5.sonasad.almaden.ibm.com D:522632:Thread-9:object:OBJProcessMonitor memcached started
E! 2015-06-03T13:59:28.917-08:00 util5.sonasad.almaden.ibm.com D:522632:Thread-9:object:OBJ running command
E! 'systemctl status memcached'
E! 2015-06-03T13:59:29.018-08:00 util5.sonasad.almaden.ibm.com D:522632:Thread-9:object:OBJ command result
E! ret:0 sout:memcached.service - Memcached
E! 2015-06-03T13:59:29.018-08:00 util5.sonasad.almaden.ibm.com I:522632:Thread-9:object:OBJ memcached is started and active running
E! 2015-06-03T13:59:29.018-08:00 util5.sonasad.almaden.ibm.com D:522632:Thread-9:object:OBJProcessMonitor memcached succeeded
E! 2015-06-03T13:59:29.018-08:00 util5.sonasad.almaden.ibm.com I:522632:Thread-9:object:OBJ service started checks
E! after monitor loop, event count:6

```

E! The following tables list the IBM Spectrum Scale for object storage log files.

E! *Table 48. Core object log files in /var/log/swift.* Core object log files in /var/log/swift

Log file	Component	Configuration file
account-auditor.log	Account auditor Swift service	account-server.conf
account-auditor.error		
account-reaper.log	Account reaper Swift service	account-server.conf
account-reaper.error		
account-replicator.log	Account replicator Swift service	account-server.conf
account-replicator.error		
account-server.log	Account server Swift service	account-server.conf
account-server.error		
container-auditor.log	Container auditor Swift service	container-server.conf
container-auditor.error		
container-replicator.log	Container replicator Swift service	container-server.conf
container-replicator.error		
container-server.log	Container server Swift service	container-server.conf
container-server.error		
container-updater.log	Container updater Swift service	container-server.conf
container-updater.error		
object-auditor.log	Object auditor Swift service	object-server.conf
object-auditor.error		
object-expirer.log	Object expirer Swift service	object-expirer.conf
object-expirer.error		
object-replicator.log	Object replicator Swift service	object-server.conf
object-replicator.error		

E! Table 48. Core object log files in /var/log/swift (continued). Core object log files in /var/log/swift

Log file	Component	Configuration file
object-server.log	Object server Swift service	object-server.conf
object-server.error		object-server-sof.conf
object-updater.log	Object updater Swift service	object-server.conf
object-updater.error		
proxy-server.log	Proxy server Swift service	proxy-server.conf
proxy-server.error		

E! Table 49. Additional object log files in /var/log/swift. Additional object log files in /var/log/swift

Log file	Component	Configuration file
ibmobjectizer.log	Unified file and object access objectizer service	spectrum-scale-objectizer.conf
ibmobjectizer.error		spectrum-scale-object.conf
policyscheduler.log	Object storage policies	spectrum-scale-object-policies.conf
policyscheduler.error		
swift.log	Performance metric collector (pmswift)	
swift.error		

E! Table 50. General system log files in /var/adm/ras. General system log files in /var/adm/ras

Log file	Component
mmsysmonitor.log	Includes everything that is monitored in the monitoring framework
mmfs.log	Various IBM Spectrum Scale command logging

### E! The IBM Spectrum Scale HDFS transparency log:

E! In IBM Spectrum Scale HDFS transparency, all logs are recorded using log4j. The **log4j.properties** file is under the /usr/lpp/mmfs/hadoop/etc/hadoop directory.

E! By default, the logs are written under the /usr/lpp/mmfs/hadoop/logs directory.

E! The following entries can be added into the **log4j.properties** file to turn on the debugging information:

```
E! log4j.logger.org.apache.hadoop.yarn=DEBUG
E! log4j.logger.org.apache.hadoop.hdfs=DEBUG
E! log4j.logger.org.apache.hadoop.gpfs=DEBUG
E! log4j.logger.org.apache.hadoop.security=DEBUG
```



**E! Protocol authentication log files:**

E! The log files pertaining to protocol authentication are described here.

E! *Table 51. Authentication log files*

Service name	Log configuration file	Log files	Logging levels
Keystone	/etc/keystone/ keystone.conf  /etc/keystone/ logging.conf	/var/log/keystone/keystone.log  /var/log/keystone/httpd- error.log  /var/log/keystone/httpd- access.log	In keystone.conf change  1. debug = true- for getting debugging information in log file.  2. verbose = true - for getting Info messages in log file .  By default, these values are false and only warning messages are logged.  Finer grained control of keystone logging levels can be specified by updating the keystones logging.conf file. For information on the logging levels in the logging.conf file, see OpenStack logging.conf documentation (docs.openstack.org/kilo/config- reference/content/ section_keystone- logging.conf.html).

E! Table 51. Authentication log files (continued)

E! Service name	E! Log configuration file	E! Log files	E! Logging levels
E! SSSD	E! /etc/sss/ E! sssd.conf	E! /var/log/sss/sss.log E! /var/log/sss/sss_nss.log E! /var/log/sss/ E! sssd_LDAPDOMAIN.log (depends E! upon configuration) E! /var/log/sss/ E! sssd_NISDOMAIN.log (depends E! upon configuration) E! <b>Note:</b> For more information on E! SSSD log files, see Red Hat Linux E! documentation.	E! 0x0010: Fatal failures. Issue with E! invoking or running SSSD. E! E! 0x0020: Critical failures. SSSD does E! not stop functioning. However, this E! error indicates that at least one E! major feature of SSSD is not to E! work properly. E! E! 0x0040: Serious failures. A E! particular request or operation has E! failed. E! E! 0x0080: Minor failures. These are E! the errors that would percolate E! down to cause the operation failure E! of 2. E! E! 0x0100: Configuration settings. E! E! 0x0200: Function data. E! E! 0x0400: Trace messages for E! operation functions. E! E! 0x1000: Trace messages for internal E! control functions. E! E! 0x2000: Contents of E! function-internal variables that E! might be interesting. E! E! 0x4000: Extremely low-level tracing E! information. E! <b>Note:</b> For more information on E! SSSD log levels, see E! Troubleshooting SSSD in RedHat E! Enterprise Linux documentation.
E! Winbind	E! /var/mmfs E! /ces/smb.conf	E! /var/adm/ras/log.wb-<DOMAIN> E! [Depends upon available E! domains] E! /var/adm/ras/log.winbindd-dc- E! connect E! /var/adm/ras/log.winbindd-idmap E! /var/adm/ras/log.winbindd	E! Log level is an integer. The value E! can be from 0-10. E! E! The default value for log level is 1.

E! **Note:** Some of the authentication modules like keystone services log information also in  
E! /var/log/messages.

E! If you change the log levels, the respective authentication service must be restarted manually on each  
E! protocol node. Restarting authentication services might result in disruption of protocol I/O.

## E! CES monitoring and troubleshooting:

E! You can monitor system health, query events, and perform maintenance and troubleshooting tasks in E! Cluster Export Services (CES).

### E! System health monitoring

E! Each CES node runs a separate GPFS process that monitors the network address configuration of the E! node. If a conflict between the network interface configuration of the node and the current assignments of E! the CES address pool is found, corrective action is taken. If the node is unable to detect an address that is E! assigned to it, the address is reassigned to another node.

E! Additional monitors check the state of the services that are implementing the enabled protocols on the E! node. These monitors cover NFS, SMB, Object, and Authentication services that monitor, for example, E! daemon liveness and port responsiveness. If it is determined that any enabled service is not functioning E! correctly, the node is marked as failed and its CES addresses are reassigned. When the node returns to E! normal operation, it returns to the normal (healthy) state and is available to host addresses in the CES E! address pool.

E! An additional monitor runs on each protocol node if Microsoft Active Directory (AD), Lightweight E! Directory Access Protocol (LDAP), or Network Information Service (NIS) user authentication is E! configured. If a configured authentication server does not respond to test requests, GPFS marks the E! affected node as failed.

### E! Querying state and events

E! Aside from the automatic failover and recovery of CES addresses, two additional outputs are provided by E! the monitoring that can be queried: events and state.

E! State can be queried by entering the **mmces state show** command, which shows you the state of each of E! the CES components. The possible states for a component follow:

#### E! HEALTHY

E! The component is working as expected.

#### E! DISABLED

E! The component has not been enabled.

#### E! SUSPENDED

E! When a CES node is in the suspended state, most components also report suspended.

#### E! STARTING

E! The component (or monitor) recently started. This state is a transient state that is updated after E! the startup is complete.

#### E! UNKNOWN

E! Something is preventing the monitoring from determining the state of the component.

#### E! STOPPED

E! The component was intentionally stopped. This situation might happen briefly if a service is E! being restarted due to a configuration change. It might also happen because a user ran the **mmces** E! **service stop *protocol*** command for a node.

#### E! DEGRADED

E! There is a problem with the component but not a complete failure. This state does not cause the E! CES addresses to be reassigned.

#### E! FAILED

E! The monitoring detected a significant problem with the component that means it is unable to E! function correctly. This state causes the CES addresses of the node to be reassigned.

## E! **DEPENDENCY\_FAILED**

E! This state implies that a component has a dependency that is in a failed state. An example would be NFS or SMB reporting **DEPENDENCY\_FAILED** because the authentication failed.

E! Looking at the states themselves can be useful to find out which component is causing a node to fail and E! have its CES addresses reassigned. To find out why the component is being reported as failed, you can E! look at events.

E! The **mmces events** command can be used to show you either events that are currently causing a E! component to be unhealthy or a list of historical events for the node. If you want to know why a E! component on a node is in a failed state, use the **mmces events active** invocation. This command gives E! you a list of any currently active events that are affecting the state of a component, along with a message E! that describes the problem. This information should provide a place to start when you are trying to find E! and fix the problem that is causing the failure.

E! If you want to get a complete idea of what is happening with a node over a longer time period, use the E! **mmces events list** invocation. By default, this command prints a list of all events that occurred on this E! node, with a time stamp. This information can be narrowed down by component, time period, and E! severity. As well as being viewable with the command, all events are also pushed to the syslog.

## E! **Maintenance and troubleshooting**

E! A CES node can be marked as unavailable by the monitoring process. The command **mmces node list** E! can be used to show the nodes and the current state flags that are associated with it. When unavailable E! (one of the following node flags are set), the node does not accept CES address assignments. The E! following possible node states can be displayed:

### E! **Suspended**

E! Indicates that the node is suspended with the **mmces node suspend** command. When suspended, E! health monitoring on the node is discontinued. The node remains in the suspended state until it E! is resumed with the **mmces node resume** command.

### E! **Network-down**

E! Indicates that monitoring found a problem that prevents the node from bringing up the CES E! addresses in the address pool. The state reverts to normal when the problem is corrected. Possible E! causes for this state are missing or non-functioning network interfaces and network interfaces E! that are reconfigured so that the node can no longer host the addresses in the CES address pool.

### E! **No-shared-root**

E! Indicates that the CES shared root directory cannot be accessed by the node. The state reverts to E! normal when the shared root directory becomes available. Possible cause for this state is that the E! file system that contains the CES shared root directory is not mounted.

E! **Failed** Indicates that monitoring found a problem with one of the enabled protocol servers. The state E! reverts to normal when the server returns to normal operation or when the service is disabled.

### E! **Starting up**

E! Indicates that the node is starting the processes that are required to implement the CES services E! that are enabled in the cluster. The state reverts to normal when the protocol servers are E! functioning.

E! Additionally, events that affect the availability and configuration of CES nodes are logged in the GPFS E! log file `/var/adm/ras/mmfs.log.latest`. The verbosity of the CES logging can be changed with the **mmces** E! **log level n** command, where *n* is a number from 0 (less logging) to 4 (more logging). The current log E! level can be viewed with the **mmfscluster --ces** command.

E! For more information about CES troubleshooting, see the IBM Spectrum Scale Wiki ([www.ibm.com/developerworks/community/wikis/home/wiki/General Parallel File System \(GPFS\)](http://www.ibm.com/developerworks/community/wikis/home/wiki/General%20Parallel%20File%20System%20(GPFS))).

## E! **Operating system error logs**

E! GPFS records file system or disk failures using the error logging facility provided by the operating system: **syslog** facility on Linux, **errpt** facility on AIX, and Event Viewer on Windows.

E! The error logging facility is referred to as *the error log* regardless of operating-system specific error log facility naming conventions.

E! Failures in the error log can be viewed by issuing this command on an AIX node:

```
E! errpt -a
```

E! and this command on a Linux node:

```
E! grep "mmfs:" /var/log/messages
```

E! You can also grep the appropriate filename where syslog messages are redirected to. For example, in Ubuntu, after the Natty release, this file will be at `/var/log/syslog`

E! On Windows, use the Event Viewer and look for events with a source label of **GPFS** in the **Application** event category.

E! On Linux, **syslog** may include GPFS log messages and the error logs described in this section. The **systemLogLevel** attribute of the **mmchconfig** command controls which GPFS log messages are sent to **syslog**. It is recommended that some kind of monitoring for GPFS log messages be implemented, particularly MMFS\_FSSTRUCT errors. For more information, see the **mmchconfig** command in the *IBM Spectrum Scale: Command and Programming Reference*.

E! The error log contains information about several classes of events or errors. These classes are:

E! • “MMFS\_ABNORMAL\_SHUTDOWN”

E! • “MMFS\_DISKFAIL”

E! • “MMFS\_ENVIRON”

E! • “MMFS\_FSSTRUCT” on page 158

E! • “MMFS\_GENERIC” on page 158

E! • “MMFS\_LONGDISKIO” on page 158

E! • “MMFS\_QUOTA” on page 159

E! • “MMFS\_SYSTEM\_UNMOUNT” on page 159

E! • “MMFS\_SYSTEM\_WARNING” on page 160

E! **MMFS\_ABNORMAL\_SHUTDOWN:** The **MMFS\_ABNORMAL\_SHUTDOWN** error log entry means that GPFS has determined that it must shutdown all operations on this node because of a problem. Insufficient memory on the node to handle critical recovery situations can cause this error. In general there will be other error log entries from GPFS or some other component associated with this error log entry.

### E! **MMFS\_DISKFAIL:**

E! This topic describes the **MMFS\_DISKFAIL** error log available in IBM Spectrum Scale.

E! The **MMFS\_DISKFAIL** error log entry indicates that GPFS has detected the failure of a disk and forced the disk to the stopped state. This is ordinarily not a GPFS error but a failure in the disk subsystem or the path to the disk subsystem.

### E! **MMFS\_ENVIRON:**

E! This topic describes the **MMFS\_ENVIRON** error log available in IBM Spectrum Scale.

E! **MMFS\_ENVIRON** error log entry records are associated with other records of the **MMFS\_GENERIC** or **MMFS\_SYSTEM\_UNMOUNT** types. They indicate that the root cause of the error is external to GPFS and usually in the network that supports GPFS. Check the network and its physical connections. The data portion of this record supplies the return code provided by the communications code.

#### E! **MMFS\_FSSTRUCT**:

E! This topic describes the **MMFS\_FSSTRUCT** error log available in IBM Spectrum Scale.

E! The **MMFS\_FSSTRUCT** error log entry indicates that GPFS has detected a problem with the on-disk structure of the file system. The severity of these errors depends on the exact nature of the inconsistent data structure. If it is limited to a single file, **EIO** errors will be reported to the application and operation will continue. If the inconsistency affects vital metadata structures, operation will cease on this file system. These errors are often associated with an **MMFS\_SYSTEM\_UNMOUNT** error log entry and will probably occur on all nodes. If the error occurs on all nodes, some critical piece of the file system is inconsistent. This can occur as a result of a GPFS error or an error in the disk system.

E! If the file system is severely damaged, the best course of action is to follow the procedures in “Additional information to collect for file system corruption or **MMFS\_FSSTRUCT** errors” on page 368, and then contact the IBM Support Center.

#### E! **MMFS\_GENERIC**:

E! This topic describes **MMFS\_GENERIC** error logs available in IBM Spectrum Scale.

E! The **MMFS\_GENERIC** error log entry means that GPFS self diagnostics have detected an internal error, or that additional information is being provided with an **MMFS\_SYSTEM\_UNMOUNT** report. If the record is associated with an **MMFS\_SYSTEM\_UNMOUNT** report, the event code fields in the records will be the same. The error code and return code fields might describe the error. See “Messages” on page 443 for a listing of codes generated by GPFS.

E! If the error is generated by the self diagnostic routines, service personnel should interpret the return and error code fields since the use of these fields varies by the specific error. Errors caused by the self checking logic will result in the shutdown of GPFS on this node.

E! **MMFS\_GENERIC** errors can result from an inability to reach a critical disk resource. These errors might look different depending on the specific disk resource that has become unavailable, like logs and allocation maps. This type of error will usually be associated with other error indications. Other errors generated by disk subsystems, high availability components, and communications components at the same time as, or immediately preceding, the GPFS error should be pursued first because they might be the cause of these errors. **MMFS\_GENERIC** error indications without an associated error of those types represent a GPFS problem that requires the IBM Support Center.

E! Before you contact IBM support center, see “Information to be collected before contacting the IBM Support Center” on page 367.

#### E! **MMFS\_LONGDISKIO**:

E! This topic describes the **MMFS\_LONGDISKIO** error log available in IBM Spectrum Scale.

E! The **MMFS\_LONGDISKIO** error log entry indicates that GPFS is experiencing very long response time for disk requests. This is a warning message and can indicate that your disk system is overloaded or that a failing disk is requiring many I/O retries. Follow your operating system's instructions for monitoring the performance of your I/O subsystem on this node and on any disk server nodes that might be involved. The data portion of this error record specifies the disk involved. There might be related error log entries from the disk subsystems that will pinpoint the actual cause of the problem. If the disk is

E! attached to an AIX node, refer to AIX in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/ssw\\_aix/welcome](http://www.ibm.com/support/knowledgecenter/ssw_aix/welcome)) and search for *performance management*. To enable or disable, use the E! **mmchfs -w** command. For more details, contact the IBM Support Center.

E! The **mmpmon** command can be used to analyze I/O performance on a per-node basis. For more E! information, see “Monitoring GPFS I/O performance with the mmpmon command” on page 3 and E! “Failures using the mmpmon command” on page 271.

#### E! MMFS\_QUOTA:

E! This topic describes the MMFS\_QUOTA error log available in IBM Spectrum Scale.

E! The **MMFS\_QUOTA** error log entry is used when GPFS detects a problem in the handling of quota E! information. This entry is created when the quota manager has a problem reading or writing the quota E! file. If the quota manager cannot read all entries in the quota file when mounting a file system with E! quotas enabled, the quota manager shuts down but file system manager initialization continues. Mounts E! will not succeed and will return an appropriate error message (see “File system forced unmount” on page E! 250).

E! Quota accounting depends on a consistent mapping between user names and their numeric identifiers. E! This means that a single user accessing a quota enabled file system from different nodes should map to E! the same numeric user identifier from each node. Within a local cluster this is usually achieved by E! ensuring that **/etc/passwd** and **/etc/group** are identical across the cluster.

E! When accessing quota enabled file systems from other clusters, you need to either ensure individual E! accessing users have equivalent entries in **/etc/passwd** and **/etc/group**, or use the user identity mapping E! facility as outlined in the IBM white paper entitled *UID Mapping for GPFS in a Multi-cluster Environment* E! in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/SSFKCN/com.ibm.cluster.gpfs.doc/gpfs\\_uid/uid\\_gpfs.html](http://www.ibm.com/support/knowledgecenter/SSFKCN/com.ibm.cluster.gpfs.doc/gpfs_uid/uid_gpfs.html)).

E! It might be necessary to run an offline quota check (**mmcheckquota**) to repair or recreate the quota file. If E! the quota file is corrupted, **mmcheckquota** will not restore it. The file must be restored from the backup E! copy. If there is no backup copy, an empty file can be set as the new quota file. This is equivalent to E! recreating the quota file. To set an empty file or use the backup file, issue the **mmcheckquota** command E! with the appropriate operand:

- E! • **-u** *UserQuotaFilename* for the user quota file
- E! • **-g** *GroupQuotaFilename* for the group quota file
- E! • **-j** *FilesetQuotaFilename* for the fileset quota file

E! After replacing the appropriate quota file, reissue the **mmcheckquota** command to check the file system E! inode and space usage.

E! For information about running the **mmcheckquota** command, see “The mmcheckquota command” on E! page 199.

#### E! MMFS\_SYSTEM\_UNMOUNT:

E! This topic describes the MMFS\_SYSTEM\_UNMOUNT error log available in IBM Spectrum Scale.

E! The **MMFS\_SYSTEM\_UNMOUNT** error log entry means that GPFS has discovered a condition that E! might result in data corruption if operation with this file system continues from this node. GPFS has E! marked the file system as disconnected and applications accessing files within the file system will receive E! **ESTALE** errors. This can be the result of:

- E! • The loss of a path to all disks containing a critical data structure.

E! If you are using SAN attachment of your storage, consult the problem determination guides provided by your SAN switch vendor and your storage subsystem vendor.  
E! • An internal processing error within the file system.

E! See “File system forced unmount” on page 250. Follow the problem determination and repair actions specified.

### E! **MMFS\_SYSTEM\_WARNING:**

E! This topic describes the MMFS\_SYSTEM\_WARNING error log available in IBM Spectrum Scale.

E! The **MMFS\_SYSTEM\_WARNING** error log entry means that GPFS has detected a system level value approaching its maximum limit. This might occur as a result of the number of inodes (files) reaching its limit. If so, issue the **mmchfs** command to increase the number of inodes for the file system so there is at least a minimum of 5% free.

### E! **Error log entry example:**

E! This topic describes an example of an error log entry in IBM Spectrum Scale.

E! This is an example of an error log entry that indicates a failure in either the storage subsystem or communication subsystem:

```
E! LABEL: MMFS_SYSTEM_UNMOUNT
E! IDENTIFIER: C954F85D
E!
E! Date/Time: Thu Jul 8 10:17:10 CDT
E! Sequence Number: 25426
E! Machine Id: 000024994C00
E! Node Id: nos6
E! Class: S
E! Type: PERM
E! Resource Name: mmfs
E!
E! Description
E! STORAGE SUBSYSTEM FAILURE
E!
E! Probable Causes
E! STORAGE SUBSYSTEM
E! COMMUNICATIONS SUBSYSTEM
E!
E! Failure Causes
E! STORAGE SUBSYSTEM
E! COMMUNICATIONS SUBSYSTEM
E!
E! Recommended Actions
E! CONTACT APPROPRIATE SERVICE REPRESENTATIVE
E!
E! Detail Data
E! EVENT CODE
E! 15558007
E! STATUS CODE
E! 212
E! VOLUME
E! gpfsd
```

### E! **Transparent Cloud Tiering logs**

E! This topic describes how to collect logs that are associated with Transparent Cloud Tiering.

E! To collect details of issues specific to Transparent Cloud Tiering, issue this command:

```
E! gpfs.snap [--cloud-gateway {BASIC | FULL}]
```



E! With the BASIC option, the Transparent Cloud Tiering service debugs information such as logs, traces, E! Java™ cores, along with minimal system and IBM Spectrum Scale cluster information is collected. No E! customer sensitive information is collected.

E! With the FULL option, extra details such as Java Heap dump are collected, along with the information E! captured with the BASIC option.

E! Successful invocation of this command generates a new .tar file at a specified location, and the file can be E! shared with IBM support team to debug a field issue.

### E! **Performance monitoring tool logs**

E! The performance monitoring tool logs can be found in the /var/log/zimon directory on each node E! configured for performance monitoring.

E! The nodes that are configured as **Collector** have two files in this directory: ZIMonCollector.log and E! ZIMonSensors.log. For nodes configured as **Sensor**, only the ZIMonSensors.log file is present. These log E! files contain information, warning, and error messages for the collector service **pmcollector**, and the E! sensor service **pmsensors**.

E! Both log files are rotated every day. The previous logs are compressed and saved in the same E! /var/log/zimon directory.

E! During installation, the log level is set to info. Issue the **mmpperfmon config show** command to see the E! current log level as shown in the following sample output:

```
E! # mmpperfmon config show
E! cephMon = "/opt/IBM/zimon/CephMonProxy"
E! cephRados = "/opt/IBM/zimon/CephRadosProxy"
E! colCandidates = "nsd003st001", "nsd004st001"
E! colRedundancy = 2
E! collectors = {
E! host = ""
E! port = "4739"
E! }
E! config = "/opt/IBM/zimon/ZIMonSensors.cfg"
E! ctdbstat = ""
E! daemonize = T
E! hostname = ""
E! ipfixinterface = "0.0.0.0"
E! logfile = "/var/log/zimon/ZIMonSensors.log"
E! loglevel = "info"
```

### E! **Setting up core dumps on a client system**

No core dump configuration is set up by IBM Spectrum Scale by default. Core dumps can be configured in a few ways.

#### **core\_pattern + ulimit**

The simplest way is to change the core\_pattern file at /proc/sys/kernel/core\_pattern and to enable core dumps using the command 'ulimit -c unlimited'. Setting it to something like /var/log/cores/core.%e.%t.%h.%p will produce core dumps similar to core.bash.1236975953.node01.2344 in /var/log/cores. This will create core dumps for Linux binaries but will not produce information for Java or Python exceptions.

#### **ABRT**

ABRT can be used to produce more detailed output as well as output for Java and Python exceptions.

The following packages should be installed to configure abrt:

- abrt (Core package)
- abrt-cli (CLI tools)
- abrt-libs (Libraries)
- abrt-addon-ccpp (C/C++ crash handler)
- abrt-addon-python (Python unhandled exception handler)
- abrt-java-connector (Java crash handler)

This overwrites the values stored in `core_pattern` to pass core dumps to abrt. It then writes this information to the abrt directory configured in `/etc/abrt/abrt.conf`. Python exceptions is caught by the python interpreter automatically importing the `abrt.pth` file installed in `/usr/lib64/python2.7/site-packages/`. If some custom configuration has changed this behavior, Python dumps may not be created.

To get Java runtimes to report unhandled exceptions through abrt, they must be executed with the command line argument `'-agentpath=/usr/lib64/libabrt-java-connector.so'`.

**Note:** Passing exception information to ABRT by using the ABRT library will cause a decrease in the performance of the application.

### ABRT Config files

The ability to collect core dumps has been added to `gpfs.snap` using the `'--protocol core'` option.

This attempts to gather core dumps from a number of locations:

- If `core_pattern` is set to dump to a file it will attempt to get dumps from the absolute path or from the root directory (the CWD for all IBM Spectrum Scale processes)
- If `core_pattern` is set to redirect to abrt it will try to read the `/etc/abrt/abrt.conf` file and read the 'DumpLocation' variable. All files and folders under this directory will be gathered.
- If the 'DumpLocation' value cannot be read then a default of `'/var/tmp/abrt'` is used.
- If `core_pattern` is set to use something other than abrt or a file path, core dumps will not be collected for the OS.

Samba can dump to the directory `'/var/adm/ras/cores/'`. Any files in this directory will be gathered.

## Trace facility

The IBM Spectrum Scale system includes many different trace points to facilitate rapid problem determination of failures.

IBM Spectrum Scale tracing is based on the kernel trace facility on AIX, embedded GPFS trace subsystem on Linux, and the Windows ETL subsystem on Windows. The level of detail that is gathered by the trace facility is controlled by setting the trace levels using the `mmtracectl` command.

The `mmtracectl` command sets up and enables tracing using default settings for various common problem situations. Using this command improves the probability of gathering accurate and reliable problem determination information. For more information about the `mmtracectl` command, see the *IBM Spectrum Scale: Command and Programming Reference*.

### Generating GPFS trace reports

Use the `mmtracectl` command to configure trace-related configuration variables and to start and stop the trace facility on any range of nodes in the GPFS cluster.

To configure and use the trace properly:

1. Issue the `mmmlsconfig dataStructureDump` command to verify that a directory for dumps was created when the cluster was configured. The default location for trace and problem determination data is

**/tmp/mmfs**. Use **mmtracectl**, as instructed by the IBM Support Center, to set trace configuration parameters as required if the default parameters are insufficient. For example, if the problem results in GPFS shutting down, set the *traceRecycle* variable with **--trace-recycle** as described in the **mmtracectl** command in order to ensure that GPFS traces are performed at the time the error occurs.

If desired, specify another location for trace and problem determination data by issuing this command:

```
mmchconfig dataStructureDump=path_for_storage_of_dumps
```

2. To start the tracing facility on all nodes, issue this command:

```
mmtracectl --start
```

3. Re-create the problem.

4. When the event to be captured occurs, stop the trace as soon as possible by issuing this command:

```
mmtracectl --stop
```

5. The output of the GPFS trace facility is stored in **/tmp/mmfs**, unless the location was changed using the **mmchconfig** command in Step 1 on page 162. Save this output.

6. If the problem results in a shutdown and restart of the GPFS daemon, set the *traceRecycle* variable as necessary to start tracing automatically on daemon startup and stop the trace automatically on daemon shutdown.

If the problem requires more detailed tracing, the IBM Support Center might ask you to modify the GPFS trace levels. Use the **mmtracectl** command to establish the required trace classes and levels of tracing. The syntax to modify trace classes and levels is as follows:

```
mmtracectl --set --trace={io | all | def | "Class Level [Class Level ...]}
```

For example, to tailor the trace level for I/O, issue the following command:

```
mmtracectl --set --trace=io
```

Once the trace levels are established, start the tracing by issuing:

```
mmtracectl --start
```

After the trace data has been gathered, stop the tracing by issuing:

```
mmtracectl --stop
```

To clear the trace settings and make sure tracing is turned off, issue:

```
mmtracectl --off
```

Other possible values that can be specified for the trace *Class* include:

**afm**

active file management

**alloc**

disk space allocation

**allocmgr**

allocation manager

**basic**

'basic' classes

**brl**

byte range locks

**cksum**

checksum services

**cleanup**

cleanup routines

**cmd**  
ts commands

**defrag**  
defragmentation

**dentry**  
dentry operations

**dentryexit**  
daemon routine entry/exit

**disk**  
physical disk I/O

**disklease**  
disk lease

**dmapi**  
Data Management API

**ds** data shipping

**errlog**  
error logging

**eventsExporter**  
events exporter

**file**  
file operations

**fs** file system

**fsck**  
online multinode fsck

**ialloc**  
inode allocation

**io** physical I/O

**kentryexit**  
kernel routine entry/exit

**kernel**  
kernel operations

**klockl**  
low-level vfs locking

**ksvfs**  
generic kernel vfs information

**lock**  
interprocess locking

**log**  
recovery log

**malloc**  
malloc and free in shared segment

**mb** mailbox message handling

**mmpmon**  
**mmpmon** command

**mnode**  
mnode operations

**msg**  
call to routines in SharkMsg.h

**mutex**  
mutexes and condition variables

**nsd**  
network shared disk

**perfmon**  
performance monitors

**pgalloc**  
page allocator tracing

**pin**  
pinning to real memory

**pit**  
parallel inode tracing

**quota**  
quota management

**rdma**  
rdma

**sanergy**  
SANergy

**scsi**  
scsi services

**sec**  
cluster security

**shared**  
shared segments

**smb**  
SMB locks

**sp** SP message handling

**super**  
super\_operations

**tasking**  
tasking system but not Thread operations

**thread**  
operations in Thread class

**tm** token manager

**ts** daemon specific code

**user1**  
miscellaneous tracing and debugging

**user2**  
miscellaneous tracing and debugging

**vbhv1**  
behaviorals

**vnode**  
vnode layer of VFS kernel support

**vnop**  
one line per VNOP with all important information

Values that can be specified for the trace *Class*, relating to vdisks, include:

**vdb**  
vdisk debugger

**vdisk**  
vdisk

**vhosp**  
vdisk hospital

For more information about vdisks and IBM Spectrum Scale RAID, see *IBM Spectrum Scale RAID: Administration*.

The trace *Level* can be set to a value from 0 through 14, which represents an increasing level of detail. A value of 0 turns tracing off. To display the trace level in use, issue the **mmfsadm showtrace** command.

On AIX, the **-aix-trace-buffer-size** option can be used to control the size of the trace buffer in memory.

On Linux nodes only, use the **mmtracectl** command to change the following:

- The trace buffer size in blocking mode.

For example, to set the trace buffer size in blocking mode to 8K, issue:

```
mmtracectl --set --tracedev-buffer-size=8K
```

- The raw data compression level.

For example, to set the trace raw data compression level to the best ratio, issue:

```
mmtracectl --set --tracedev-compression-level=9
```

- The trace buffer size in overwrite mode.

For example, to set the trace buffer size in overwrite mode to 500M, issue:

```
mmtracectl --set --tracedev-overwrite-buffer-size=500M
```

- When to overwrite the old data.

For example, to wait to overwrite the data until the trace data is written to the local disk and the buffer is available again, issue:

```
mmtracectl --set --tracedev-write-mode=blocking
```

**--tracedev-write-mode=blockingspecifies** that if the trace buffer is full, wait until the trace data is written to the local disk and the buffer becomes available again to overwrite the old data. This is the default. **--tracedev-write-mode=overwrite** specifies that if the trace buffer is full, overwrite the old data.

**Note:** Before switching between **--tracedev-write-mode=overwrite** and **--tracedev-write-mode=blocking**, or vice versa, run the **mmtracectl --stop** command first. Next, run the **mmtracectl --set --tracedev-write-mode** command to switch to the desired mode. Finally, restart tracing with the **mmtracectl --start** command.

For more information about the **mmtracectl** command, see the *IBM Spectrum Scale: Command and Programming Reference*.

## CES tracing and debug data collection

You can collect debugging information in Cluster Export Services.

**Data collection (FTDC):** To diagnose the cause of an issue, it might be necessary to gather some extra information from the cluster. This information can then be used to determine the root cause of an issue.

Collection of debugging information, such as configuration files and logs, can be gathered by using the **gpfs.snap** command. This command gathers data about GPFS, operating system information, and information for each of the protocols.

### GPFS + OS

GPFS configuration and logs plus operating system information such as network configuration or connected drives.

**CES** Generic protocol information such as configured CES nodes.

**NFS** NFS Ganesha configuration and logs.

**SMB** SMB and CTDB configuration and logs.

### OBJECT

Openstack Swift and Keystone configuration and logs.

### AUTHENTICATION

Authentication configuration and logs.

### PERFORMANCE

Dump of the performance monitor database.

Information for each of the enabled protocols is gathered automatically when the **gpfs.snap** command is run. If any protocol is enabled, then information for CES and authentication is gathered.

To gather performance data, add the **--performance** option. The **--performance** option causes **gpfs.snap** to try to collect performance information.

**Note:** Because this process can take up to 30 minutes to run, gather performance data only if necessary.

If data is only required for one protocol or area, the automatic collection can be bypassed. Provided one or more of the following options to the **--protocol** argument: **smb,nfs,object,ces,auth,none**

If the **--protocol** command is provided, automatic data collection is disabled. If **--protocol smb,nfs** is provided to **gpfs.snap**, only NFS and SMB information is gathered and no CES or Authentication data is collected. To disable all protocol data collection, use the argument **--protocol none**.

### Types of tracing:

Tracing is logging at a high level. The command for starting and stopping tracing (**mmprotocoltrace**) supports SMB tracing. NFS and Object tracing can be done with a combination of commands.

| **SMB** To start SMB tracing, use the **mmprotocoltrace start smb -c <clientIP>** command. The output looks similar to this example:

| Trace 'fcb7cb07-c45e-43f8-8f1f-2de50cf15062' created successfully for 'smb'

To see the status of the trace command, use the **mmprotocoltrace status smb** command. The output looks similar to this example:

```
| Trace ID:    fcb7cb07-c45e-43f8-8f1f-2de50cf15062
| State:      Active
| User ID:    root
| Protocol:   smb
| Start Time: 10:57:43 04/03/2016
| End Time:   11:07:43 04/03/2016
| Client IPs: 10.0.100.42, 10.0.100.43
```

```

| Origin Node: ch-42.localnet.com
| Syscall:      False
| Syscall Only: False
| Nodes:
|   Node Name:  ch-41.localnet.com
|   State:      ACTIVE
|   Trace Location: /tmp/mmfs/smb.20160304_105742.trc
|
|   Node Name:  ch-42.localnet.com
|   State:      ACTIVE
|   Trace Location: /tmp/mmfs/smb.20160304_105742.trc
|
|   Node Name:  ch-43.localnet.com
|   State:      ACTIVE
|   Trace Location: /tmp/mmfs/smb.20160304_105742.trc

```

To stop the trace the command, use the **mmprotocoltrace stop smb** command:

```

| Stopping traces
| Trace 'fcb7cb07-c45e-43f8-8f1f-2de50cf15062' stopped for smb
| Waiting for traces to complete
| Waiting for node 'node1'
| Waiting for node 'node2'
| Waiting for node 'node3'
| Finishing trace 'fcb7cb07-c45e-43f8-8f1f-2de50cf15062'
| Successfully copied file from 'node1:/tmp/mmfs/smb.20160304_105742.trc'
| Successfully copied file from 'node2:/tmp/mmfs/smb.20160304_105742.trc'
| Successfully copied file from 'node3:/tmp/mmfs/smb.20160304_105742.trc'
| Trace tar file has been written to '/tmp/mmfs/smb.trace.20160304_105845.tar.gz'

```

The tar file then includes the log files that contain top-level logs and configuration details of SMB for each node and every connected client for the time period the trace was running for.

Traces time out after a certain amount of time. By default, this time is 10 minutes. The timeout can be changed by using the **-d** argument when you start the trace. When a trace times out, the first node with the timeout ends the trace and writes the location of the collected data into the **mmprotocoltrace** logs. Each other node writes an information message that states that another node ended the trace.

A full usage message for the **mmprotocoltrace** command is printable by using the **-h** argument.

**NFS** NFS tracing is achieved by increasing the log level, repeating the issue, capturing the log file, and then restoring the log level.

To increase the log level, use the command **mmnfs** configuration change **LOG\_LEVEL=FULL\_DEBUG**.

You can set the log level to the following values: NULL, FATAL, MAJ, CRIT, WARN, EVENT, INFO, DEBUG, MID\_DEBUG, and FULL\_DEBUG.

FULL\_DEBUG is the most useful for debugging purposes.

After the issue is recreated by running the **gpfs.snap** command either with no arguments or with the **--protocol nfs** argument, the NFS logs are captured. The logs can then be used to diagnose any issues.

To return the log level to normal, use the same command but with a lower logging level (the default is EVENT).

## Object

The process for tracing the object protocol is similar to NFS. The Object service consists of multiple processes that can be controlled individually.

The Object services use these logging levels, at increasing severity: DEBUG, INFO, AUDIT, WARNING, ERROR, CRITICAL, and TRACE.



### Keystone and Authentication

```
mmobj config change --ccrfile keystone.conf --section DEFAULT --property debug
--value True
```

Finer grained control of Keystone logging levels can be specified by updating the Keystone's logging.conf file. For information on the logging levels in the logging.conf file, see the OpenStack logging.conf documentation ([docs.openstack.org/kilo/config-reference/content/section\\_keystone-logging.conf.html](https://docs.openstack.org/kilo/config-reference/content/section_keystone-logging.conf.html)).

### Swift Proxy Server

```
mmobj config change --ccrfile proxy-server.conf --section DEFAULT --property
log_level --value DEBUG
```

### Swift Account Server

```
mmobj config change --ccrfile account-server.conf --section DEFAULT --property
log_level --value DEBUG
```

### Swift Container Server

```
mmobj config change --ccrfile container-server.conf --section DEFAULT --property
log_level --value DEBUG
```

### Swift Object Server

```
mmobj config change --ccrfile object-server.conf --section DEFAULT --property
log_level --value DEBUG
```

These commands increase the log level for the particular process to the debug level. After you have re-created the problem, run the **gpfs.snap** command with no arguments or with the **--protocol object** argument.

Then, decrease the log levels again by using the commands that are shown previously but with **--value ERROR** instead of **--value DEBUG**.

**Collecting trace information:** Use the **mmprotocoltrace** command to collect trace information for debugging system problems or performance issues. For more information, see the **mmprotocoltrace** command in the *IBM Spectrum Scale: Command and Programming Reference*. This section is divided into the following subsections:

- “Running a typical trace”
- “Trace timeout” on page 171
- “Trace log files” on page 171
- “Trace configuration file” on page 171
- “Resetting the trace system ” on page 173
- “Using advanced options” on page 173

### Running a typical trace

The following steps describe how to run a typical trace. It is assumed that the trace system is reset for the type of trace that you want to run: SMB, Network, or Object. The examples use the SMB trace.

1. Before you start the trace, you can check the configuration settings for the type of trace that you plan to run:

```
mmprotocoltrace config smb
```

The response to this command displays the current settings from the trace configuration file. For more information about this file, see the “Trace configuration file” on page 171 subtopic.

2. Clear the trace records from the previous trace of the same type:

```
mmprotocoltrace clear smb
```

This command responds with an error message if the previous state of a trace node is something other than **DONE** or **FAILED**. If this error occurs, follow the instructions in the “Resetting the trace system ” on page 173 subtopic.

3. Start the new trace:

```
| ]# mmprotocoltrace start smb -c <clientIP>
```

The following response is typical:

```
Trace '3f36dbed-b567-4566-9beb-63b6420bbb2d' created successfully for 'smb'
```

4. Check the status of the trace to verify that tracing is active on all the configured nodes:

```
mmprotocoltrace status smb
```

The following response is typical:

```
| Trace ID:    fcb7cb07-c45e-43f8-8f1f-2de50cf15062
| State:      Active
| User ID:    root
| Protocol:   smb
| Start Time: 10:57:43 04/03/2016
| End Time:   11:07:43 04/03/2016
| Client IPs: 10.0.100.42, 10.0.100.43
| Origin Node: ch-42.localnet.com
| Syscall:    False
| Syscall Only:False
| Nodes:
|   Node Name:  ch-41.localnet.com
|   State:      ACTIVE
|   Trace Location: /tmp/mmfs/smb.20160304_105742.trc
|
|   Node Name:  ch-42.localnet.com
|   State:      ACTIVE
|   Trace Location: /tmp/mmfs/smb.20160304_105742.trc
|
|   Node Name:  ch-43.localnet.com
|   State:      ACTIVE
|   Trace Location: /tmp/mmfs/smb.20160304_105742.trc
```

To display more status information, add the **-v** (verbose) option:

```
mmprotocoltrace -v status smb
```

If the status of a node is **FAILED**, the node did not start successfully. Look at the logs for the node to determine the problem. After you fix the problem, reset the trace system by following the steps in the “Resetting the trace system ” on page 173 subtopic.

5. If all the nodes started successfully, perform the actions that you want to trace. For example, if you are tracing a client IP address, enter commands that create traffic on that client.
6. Stop the trace:

```
mmprotocoltrace stop smb
```

The following response is typical. The last line gives the location of the trace log file:

```
Stopping traces
Trace '01239483-be84-wev9-a2d390i9ow02' stopped for smb
Waiting for traces to complete
Waiting for node 'node1'
Waiting for node 'node2'
Finishing trace '01239483-be84-wev9-a2d390i9ow02'
Trace tar file has been written to '/tmp/mmfs/smb.20150513_162322.trc/smb.trace.20150513_162542.tar.gz'
```

If you do not stop the trace, it continues until the trace duration expires. For more information, see the “Trace timeout” on page 171 subtopic.

7. Look in the trace log files for the results of the trace. For more information, see the “Trace log files” on page 171 subtopic.

## Trace timeout

If you do not stop a trace manually, the trace runs until its trace duration expires. The default trace duration is 10 minutes, but you can set a different value in the **mmprotocoltrace** command. Each node that participates in a trace starts a timeout process that is set to the trace duration. When a timeout occurs, the process checks the trace status. If the trace is active, the process stops the trace, writes the file location to the log file, and exits. If the trace is not active, the timeout process exits.

If a trace stops because of a timeout, look in the log file of each node to find the location of the trace log file. The log entry is similar to the following entry:

```
2015-08-26T16:53:35.885 W:14150:MainThread:TIMEOUT:
    Trace 'd4643ccf-96c1-467d-93f8-9c71db7333b2' tar file located at
    '/tmp/mmfs/smb.20150826_164328.trc/smb.trace.20150826_165334.tar.gz'
```

## Trace log files

Trace log files are compressed files in the `/var/adm/ras` directory. The contents of a trace log file depends on the type of trace. The product supports three types of tracing: SMB, Network, Object, and Windbind.

**SMB** SMB tracing captures System Message Block information. The resulting trace log file contains an `smbd.log` file for each node for which information has been collected and for each client that is connected to this node. A trace captures information for all clients with the specified IP address.

### Network

Network tracing calls Wireshark's `dumppcap` utility to capture network packets. The resulting trace log file contains a `pcappng` file that is readable by Wireshark and other programs. The file name is similar to `bfn22-10g_all_00001_20150907125015.pcap`.

If the **mmprotocoltrace** command specifies a client IP address, the trace captures traffic between that client and the server. If no IP address is specified, the trace captures traffic across all network interfaces of each participating node.

### Object

The trace log file contains log files for each node, one for each of the object services.

Object tracing sets the log location in the `rsyslog` configuration file. For more information about this file, see the description of the *rsyslogconflocation* configuration parameter in the “Trace configuration file” subtopic.

It is not possible to configure an Object trace by clients so that information for all connections is recorded.

## Trace configuration file

Each node in the cluster has its own trace configuration file, which is stored in the `/var/mmfs/ces` directory. The configuration file contains settings for logging and for each type of tracing:

### [logging]

#### filename

The name of the log file.

**level** The current logging level, which can be `debug`, `info`, `warning`, `error`, or `critical`.

### [smb]

#### defaultloglocation

The default log location that is used by the `reset` command or when current information is not retrievable.

**defaultloglevel**

The default log level that is used by the reset command or when current information is not retrievable.

**traceloglevel**

The log level for tracing.

**maxlogsize**

The maximum size of the log file in kilobytes.

**esttracesize**

The estimated trace size in kilobytes.

**[network]****numoflogfiles**

The maximum number of log files.

**logfilesize**

The maximum size of the log file in kilobytes.

**esttracesize**

The estimated trace size in kilobytes.

**[object]****defaultloglocation**

The default log location that is used by the reset command or when current information is not retrievable.

**defaultloglevel**

The default log level that is used by the reset command or when current information is not retrievable.

**traceloglevel**

The log level for tracing.

**rsyslogconflocation**

The location of the rsyslog configuration file. Rsyslog is a service that is provided by Red Hat, Inc. that redirects log output. The default location is /etc/rsyslog.d/00-swift.conf..

**esttracesize**

The estimated trace size in kilobytes.

**[winbind]****defaultlogfiles**

The location of the winbind log files. The default location is /var/adm/ras/log.w\*.

**defaultloglevel**

The default log level that is used by the reset command or when current information is not retrievable. The value of defaultloglevel is set to 1.

**traceloglevel**

The log level for tracing. The value for traceloglevel is set to 10.

**esttracesize**

The estimated trace size in kilobytes. The value of esttracesize is set to 500000.

**[syscalls]**

**args** The CLI arguments, used while executing the strace\_executable. By default: -T -tt -C.

## Resetting the trace system

Before you run a new trace, verify that the trace system is reset for the type of trace that you want to run: SMB, Network, or Object. The examples in the following instructions use the SMB trace system. To reset the trace system, follow these steps:

1. Stop the trace if it is still running.
  - a. Check the trace status to see whether the current trace is stopped on all the nodes:

```
mmprotocoltrace status smb
```

If the trace is still running, stop it:

```
mmprotocoltrace stop smb
```

2. Clear the trace records:

```
mmprotocoltrace clear smb
```

If the command is successful, then you have successfully reset the trace system. Skip to the last step in these instructions.

If the command returns an error message, go to the next step.

**Note:** The command responds with an error message if the trace state of a node is something other than **DONE** or **FAILED**. You can verify the trace state of the nodes by running the **status** command:

```
mmprotocoltrace status smb
```

3. Run the clear command again with the **-f** (force) option.

```
mmprotocoltrace -f clear smb
```

4. After a forced clear, the trace system might still be in an invalid state. Run the reset command. For more information about the command, see the “Using advanced options.”

```
mmprotocoltrace reset smb
```

5. Check the default values in the trace configuration file to verify that they are correct. To display the values in the trace configuration file, run the config command. For more information about the file, see the “Trace configuration file” on page 171 subtopic.

```
mmprotocoltrace config smb
```

6. The trace system is ready. You can now start a new trace.

## Using advanced options

The **reset** command restores the trace system to the default values that are set in the trace configuration file. The command also performs special actions for each type of trace:

- For an SMB trace, the reset removes any IP-specific configuration files and sets the log level and log location to the default values.
- For a Network trace, the reset stops all dumpcap processes.
- For an Object trace, the reset sets the log level to the default value. It then sets the log location to the default location in the rsyslog configuration file, and restarts the rsyslog service.

The following command resets the SMB trace:

```
mmprotocoltrace reset smb
```

The **status** command with the **-v** (verbose) option provides more trace information, including the values of trace variables. The following command returns verbose trace information for the SMB trace:

```
mmprotocoltrace -v status smb
```

## Tips for using mmprotocoltrace

Follow these tips for **mmprotocoltrace**.

## Specifying nodes with the -N and -c parameters.

It is important to understand the difference between the -N and -c parameters of the **mmprotocoltrace** command:

- The -N parameter specifies the CES nodes where you want tracing to be done. The default value is all CES nodes.
- The -c parameter specifies the IP addresses of clients whose incoming connections are to be traced. Where these clients are connected to the CES nodes that are specified in the -N parameter, those CES nodes trace the connections with the clients.

For example, in the SMB trace started by the following command, the CES node 10.40.72.105 traces incoming connections from clients 192.168.4.1, 192.168.4.26, and 192.168.4.22. The command is all on one line:

```
mmprotocoltrace start smb -c 192.168.4.1,192.168.4.26,192.168.4.22  
-N 10.40.72.105
```

## Discovering client IP addresses for an smb trace

If you have only a few clients that you want to trace, you can list their IP addresses by running the system command **smbstatus** on a CES node. This command lists the IP addresses of all smb clients that are connected to the node.

However, if many clients are connected to the CES node, running **smbstatus** on the node to discover client IP addresses might not be practical. The command sets a global lock on the node for the entire duration of the command, which might be a long time if many clients are connected.

Instead, run the system command **ip** on each client that you are interested in and filter the results according to the type of device that you are looking for. In the following example, the command is run on client ch-41 and lists the IP address 10.0.100.41 for that client:

```
[root@ch-41 ~]# ip a | grep "inet "  
inet 127.0.0.1/8 scope host lo  
inet 10.0.100.41/24 brd 10.255.255.255 scope global eth0
```

A client might have more than one IP address, as in the following example where the command **ip** is run on client ch-44:

```
[root@ch-44 ~]# ip a | grep "inet "  
inet 127.0.0.1/8 scope host lo  
inet 10.0.100.44/24 brd 10.255.255.255 scope global eth0  
inet 192.168.4.1/16 brd 192.168.255.255 scope global eth1  
inet 192.168.4.26/16 brd 192.168.255.255 scope global secondary eth1:0  
inet 192.168.4.22/16 brd 192.168.255.255 scope global secondary eth1:1
```

In such a case, specify all the possible IP addresses in the **mmprotocoltrace** command because you cannot be sure which IP address the client will use. The following example specifies all the IP addresses that the previous example listed for client ch-44, and by default all CES nodes will trace incoming connections from any of these IP addresses:

```
mmprotocoltrace start smb -c 10.0.100.44,192.168.4.1,192.168.4.26,192.168.4.22
```

---

## E! Collecting diagnostic data through GUI

IBM Support might ask you to collect logs, trace files, and dump files from the system to help them resolve a problem. You can perform this task from the management GUI or by using the **gpfs.snap** command. Use the **Settings > Diagnostic Data** page in the IBM Spectrum Scale GUI to collect details of the issues reported in the system.

The entire set of diagnostic data available in the system helps to analyze all kinds of IBM Spectrum Scale issues. Depending on the data selection criteria, these files can be large (gigabytes) and might take an

hour to download. The diagnostic data is collected from each individual node in a cluster. In a cluster with hundreds of nodes, downloading the diagnostic data might take a long time and the downloaded file can be large in size.

It is always better to reduce the size of the log file as you might need to send it to IBM Support to help fix the issues. You can reduce the size of the diagnostic data file by reducing the scope. The following options are available to reduce the scope of the diagnostic data:

- Include only affected functional areas
- Include only affected nodes
- Reduce the number of days for which the diagnostic data needs to be collected

The following three modes are available in the GUI to select the functional areas of the diagnostic data:

#### 1. Standard diagnostics

The data that is collected in the standard diagnostics consists of the configuration, status, log files, dumps, and traces in the following functional areas:

- Core IBM Spectrum Scale
- Network
- GUI
- NFS
- SMB
- Object
- Authentication
- Cluster export services (CES)
- Crash dumps

You can download the diagnostic data for the above functional areas at the following levels:

- All nodes
- Specific nodes
- All nodes within one or more node classes

#### 2. Deadlock diagnostics

The data that is collected in this category consists of the minimum amount of data that is needed to investigate a deadlock problem.

#### 3. Performance diagnostics

The data that is collected in this category consists of the system performance details collected from performance monitoring tools. You can only use this option if it is requested by the IBM Support.

The GUI log files contain the issues that are related to GUI and it is smaller in size as well. The GUI log consists of the following types of information:

- Traces from the GUI that contain the information about errors occurred inside GUI code
- Several configuration files of GUI and postgresSQL
- Dump of postgresSQL database that contains IBM Spectrum Scale configuration data and events
- Output of most `mm1s*` commands
- Logs from the performance collector

**Note:** Instead of collecting the diagnostic data again, you can also utilize the diagnostic data that is collected in the past. You can analyze the relevance of the historic data based on the date on which the issue is reported in the system. Ensure to delete the diagnostic data that is no longer needed to save disk space.

---

## E! CLI commands for collecting issue details

E! You can issue several CLI commands to collect details of the issues that you might encounter while using E! IBM Spectrum Scale.

## E! Using the `gpfs.snap` command

E! This topic describes the usage of `gpfs.snap` command in IBM Spectrum Scale.

E! Running the `gpfs.snap` command with no options is similar to running `gpfs.snap -a`. It collects data from E! all nodes in the cluster. This invocation creates a file that is made up of multiple `gpfs.snap` snapshots. E! The file that is created includes a master snapshot of the node from which the `gpfs.snap` command was E! invoked and non-master snapshots of each of other nodes in the cluster.

E! If the node on which the `gpfs.snap` command is run is not a file system manager node, `gpfs.snap` creates E! a non-master snapshot on the file system manager nodes.

E! The difference between a master snapshot and a non-master snapshot is the data that is gathered. A E! master snapshot gathers information from nodes in the cluster. A master snapshot contains all data that a E! non-master snapshot has. There are two categories of data that is collected:

E! 1. Data that is always gathered by `gpfs.snap` (for master snapshots and non-master snapshots):

- E! • “Data gathered by `gpfs.snap` on all platforms”
- E! • “Data gathered by `gpfs.snap` on AIX” on page 177
- E! • “Data gathered by `gpfs.snap` on Linux” on page 178
- E! • “Data gathered by `gpfs.snap` on Windows” on page 179

E! 2. Data that is gathered by `gpfs.snap` only in the case of a master snapshot. See “Data gathered by E! `gpfs.snap` for a master snapshot” on page 179.

E! When the `gpfs.snap` command runs with no options, data is collected for each of the enabled protocols. E! You can turn off the collection of all protocol data and specify the type of protocol information to be E! collected using the `--protocol` option. For more information, see `gpfs.snap command` in *IBM Spectrum E! Scale: Command and Programming Reference*.

E! The following categories of data is collected:

E! • Data that is always gathered by `gpfs.snap` on Linux for protocols:

- E! – “Data gathered for SMB on Linux” on page 180
- E! – “Data gathered for NFS on Linux” on page 180
- E! – “Data gathered for Object on Linux” on page 181
- E! – “Data gathered for CES on Linux” on page 182
- E! – “Data gathered for authentication on Linux” on page 182
- E! – “Data gathered for performance on Linux” on page 183

## E! Data gathered by `gpfs.snap` on all platforms

E! These items are always obtained by the `gpfs.snap` command when gathering data for an AIX, Linux, or E! Windows node:

E! 1. The output of these commands:

- E! • `ls -l /user/lpp/mmfs/bin`
- E! • `mmdevdiscover`
- E! • `tspreparedisk -S`
- E! • `mmfsadm dump malloc`
- E! • `mmfsadm dump fs`



- E! • `df -k`
  - E! • `ifconfig interface`
  - E! • `ipcs -a`
  - E! • `ls -l /dev`
  - E! • `mmfsadm dump alloc hist`
  - E! • `mmfsadm dump alloc stats`
  - E! • `mmfsadm dump allocmgr`
  - E! • `mmfsadm dump allocmgr hist`
  - E! • `mmfsadm dump allocmgr stats`
  - E! • `mmfsadm dump cfgmgr`
  - E! • `mmfsadm dump config`
  - E! • `mmfsadm dump dealloc stats`
  - E! • `mmfsadm dump disk`
  - E! • `mmfsadm dump mmap`
  - E! • `mmfsadm dump mutex`
  - E! • `mmfsadm dump nsd`
  - E! • `mmfsadm dump rpc`
  - E! • `mmfsadm dump sgmgr`
  - E! • `mmfsadm dump stripe`
  - E! • `mmfsadm dump tscomm`
  - E! • `mmfsadm dump version`
  - E! • `mmfsadm dump waiters`
  - E! • `netstat` with the `-i`, `-r`, `-rn`, `-s`, and `-v` options
  - E! • `ps -edf`
  - E! • `vmstat`
- E! 2. The contents of these files:
- E! • `/etc/syslog.conf` or `/etc/syslog-ng.conf`
  - E! • `/tmp/mmfs/internal*`
  - E! • `/tmp/mmfs/trcrpt*`
  - E! • `/var/adm/ras/mmfs.log.*`
  - E! • `/var/mmfs/gen/*`
  - E! • `/var/mmfs/etc/*`
  - E! • `/var/mmfs/tmp/*`
  - E! • `/var/mmfs/ssl/*` except for `complete.map` and `id_rsa` files

### E! **Data gathered by gpfs.snap on AIX**

E! This topic describes the type of data that is always gathered by the `gpfs.snap` command on the AIX platform.

E! These items are always obtained by the `gpfs.snap` command when gathering data for an AIX node:

E! 1. The output of these commands:

- E! • `errpt -a`
- E! • `lssrc -a`
- E! • `lslpp -hac`
- E! • `no -a`

E! 2. The contents of these files:

- E! • /etc/filesystems
- E! • /etc/trcfmt

## E! **Data gathered by gpfs.snap on Linux**

E! This topic describes the type of data that is always gathered by the **gpfs.snap** command on the Linux platform.

E! These items are always obtained by the **gpfs.snap** command when gathering data for a Linux node:

E! 1. The output of these commands:

- E! • **dmesg**
- E! • **fdisk -l**
- E! • **lsmod**
- E! • **lspci**
- E! • **rpm -qa**
- E! • **rpm --verify gpfs.base**
- E! • **rpm --verify gpfs.docs**
- E! • **rpm --verify gpfs.gpl**
- E! • **rpm --verify gpfs.msg.en\_US**

E! 2. The contents of these files:

- E! • /etc/filesystems
- E! • /etc/fstab
- E! • /etc/\*release
- E! • /proc/cpuinfo
- E! • /proc/version
- E! • /usr/lpp/mmfs/src/config/site.mcr
- E! • /var/log/messages\*

### | **Performance monitoring data**

E! The following data is collected to enable performance monitoring diagnosis:

E! 1. The output of these commands:

- E! • **mmperfmon config show**
- E! • **ps auxw | grep ZIMon**
- E! • **service pmsensors status**
- E! • **service pmcollector status**
- E! • **du -h /opt/IBM/zimon**
- E! • **ls -laR /opt/IBM/zimon/data**

E! 2. The contents of these files:

- E! • /var/log/zimon/\*
- E! • /opt/IBM/zimon/\*.cfg

| The following data is also collected on Linux on z Systems:

E! 1. The output of the **dbginfo.sh** tool.

E! If s390-tools are installed, then the output of **dbginfo.sh** is captured.

E! 2. The content of these files:

- E! • /boot/config-\$(**active-kernel**) (for example: /boot/config-3.10.0-123.6.3.el7.s390x)

## E! **Data gathered by gpfs.snap on Windows**

E! This topic describes the type of data that is always gathered by the **gpfs.snap** command on the Windows E! platform.

E! These items are always obtained by the **gpfs.snap** command when gathering data for a Windows node:

- E! 1. The output from **systeminfo.exe**
- E! 2. Any raw trace files **\*.tmf** and **mmfs.trc\***
- E! 3. The **\*.pdb** symbols from **/usr/lpp/mmfs/bin/symbols**

## E! **Data gathered by gpfs.snap for a master snapshot**

E! This topic describes the type of data that is always gathered by the **gpfs.snap** command for a master E! snapshot.

E! When the **gpfs.snap** command is specified with no options, a master snapshot is taken on the node E! where the command was issued. All of the information from “Data gathered by gpfs.snap on all E! platforms” on page 176, “Data gathered by gpfs.snap on AIX” on page 177, “Data gathered by gpfs.snap E! on Linux” on page 178, and “Data gathered by gpfs.snap on Windows” is obtained, as well as this data:

E! 1. The output of these commands:

- E! • **mmauth**
- E! • **mmgetstate -a**
- E! • **mmlscluster**
- E! • **mmlsconfig**
- E! • **mmlsdisk**
- E! • **mmlsfileset**
- E! • **mmlsfs**
- E! • **mmlspolicy**
- E! • **mmlsmgr**
- E! • **mmlsnode -a**
- E! • **mmlsnsd**
- E! • **mmlssnapshot**
- E! • **mmremotefilesystem**
- E! • **mmremotefilesystems**
- E! • **tsstatus**

E! 2. The contents of the **/var/adm/ras/mmfs.log.\*** file (on all nodes in the cluster)

### | **Performance monitoring data**

E! The master snapshot, when taken on a Linux node, collects the following data:

E! 1. The output of these commands:

- E! • **mmlscluster**
- E! • **mmdiag --waiters --iohist --threads --stats --memory**
- E! • **mmfsadm eventsExporter mmpmon chms**
- E! • **mmfsadm dump nsd**
- E! • **mmfsadm dump mb**

E! **Note:** The performance monitoring data is only collected if the master node is a Linux node.

## E! **Data gathered by gpfs.snap on Linux for protocols**

E! When the **gpfs.snap** command runs with no options, data is collected for each of the enabled protocols.

E! You can turn off the collection of all protocol data and specify the type of protocol information to be E! collected using the **--protocol** option..

#### E! **Data gathered for SMB on Linux:**

E! The following data is always obtained by the `gpfs.snap` command for the server message block (SMB).

E! 1. The output of these commands:

- E! • **ctdb status**
- E! • **ctdb scriptstatus**
- E! • **ctdb ip**
- E! • **ctdb statistics**
- E! • **ctdb uptime**
- E! • **smbstatus**
- E! • **wbinfo -t**
- E! • **rpm -q gpfs.smb**
- E! • **rpm -q samba**
- E! • **net conf list**
- E! • **sharesec --view-all**
- E! • **mmlsperfddata smb2Throughput -n 1440 -b 60**
- E! • **mmlsperfddata smb2IORate -n 1440 -b 60**
- E! • **mmlsperfddata smb2IOLatency -n 1440 -b 60**
- E! • **ls -l /var/ctdb**
- E! • **ls -l /var/ctdb/persistent**
- E! • **tdbtool info** for all **.tdb** files in **/var/ctdb/\***
- E! • **tdbtool check** for all **.tdb** files in **/var/ctdb/persistent**

E! 2. The content of these files:

- E! • **/var/adm/ras/log.smbd**
- E! • **/var/lib/samba/\***
- E! • **/var/lib/ctdb/\***
- E! • **/etc/sysconfig/gpfs-ctdb**
- E! • **/var/mmfs/ces/smb.conf**
- E! • **/var/mmfs/ces/smb.ctdb.nodes**

#### E! **Data gathered for NFS on Linux:**

E! The following data is always obtained by the `gpfs.snap` command for NFS.

E! 1. The output of these commands:

- E! • **mmnfs export list**
- E! • **mmnfs configuration list**
- E! • **rpm -qi - for all installed ganesha packages**

E! 2. The content of these files:

- E! • **/var/mmfs/ces/nfs-config/\***
- E! • **/var/log/ganesha.log**
- E! • **/var/tmp/abrt/\***

E! Files stored in the CCR:

- E! • **gpfs.ganesha.exports.conf**
- E! • **gpfs.ganesha.main.conf**

- E! • gpfs.ganesha.nfsd.conf
- E! • gpfs.ganesha.log.conf

### E! Data gathered for Object on Linux:

E! The following data is always obtained by the gpfs.snap command for Object protocol.

E! 1. The output of these commands:

- E! • **swift info**
- E! • **rpm -qi - for all installed openstack rpms**

E! 2. The content of these files:

- E! • /var/log/swift/\*
- E! • /var/log/keystone/\*
- E! • /var/log/httpd/\*
- E! • /var/log/messages
- E! • /etc/httpd/conf/httpd.conf
- E! • /etc/httpd/conf.d/ssl.conf
- E! • /etc/httpd/conf.d/wsgi-keystone.conf
- E! • All files stored in the directory specified in the spectrum-scale-objectizer.conf CCR file in the **objectization\_tmp\_dir** parameter.

E! Files stored in the CCR:

- E! • account-server.conf
- E! • account.builder
- E! • account.ring.gz
- E! • container-server.conf
- E! • container.builder
- E! • container.ring.gz
- E! • object-server.conf
- E! • object\*.builder
- E! • object\*.ring.gz
- E! • container-reconciler.conf
- E! • swift.conf
- E! • spectrum-scale-compression-scheduler.conf
- E! • spectrum-scale-object-policies.conf
- E! • spectrum-scale-objectizer.conf
- E! • spectrum-scale-object.conf
- E! • object-server-sof.conf
- E! • object-expirer.conf
- E! • keystone-paste.ini
- E! • policy\*.json
- E! • sso/certs/ldap\_cacert.pem
- E! • object-expirer.conf
- E! • object-server-sof.conf
- E! • spectrum-scale-compression-scheduler.conf
- E! • spectrum-scale-compression-status.stat
- E! • spectrum-scale-object.conf
- E! • spectrum-scale-object-policies.conf

## E! Data gathered for CES on Linux:

E! The following data is always obtained by the `gpfs.snap` command for any enabled protocols.

E! 1. The output of these commands:

- E! • `mmlscluster --ces`
- E! • `mmces node list`
- E! • `mmces address list`
- E! • `mmces service list -a`
- E! • `mmccr flist`

E! 2. The content of these files:

- E! • `/var/adm/ras/mmcesmonitor.log`
  - E! • `/var/adm/ras/mmcesmonitor.log.*`
  - E! • `/var/adm/ras/ras.db` (Contents exported as csv file)
  - E! • All files stored at the `cesSharedRoot + /ces/connections/`
  - E! • All files stored at the `cesSharedRoot + /ces/addr/`
- E! Files stored in the CCR:
- E! • `cesiplist`
  - E! • `ccr.nodes`
  - E! • `ccr.disks`

## E! Data gathered for authentication on Linux:

E! The following data is always obtained by the `gpfs.snap` command for any enabled protocol.

E! 1. The output of these commands:

- E! • `mmcesuserauthlsservice`
- E! • `mmcesuserauthckservice --data-access-method all --nodes cesNodes`
- E! • `mmcesuserauthckservice --data-access-method all --nodes cesNodes --server-reachability`
- E! • `systemctl status ypbind`
- E! • `systemctl status sssd`
- E! • `ps aux | grep keystone`
- E! • `lsuf -i`
- E! • `sestatus`
- E! • `systemctl status firewalld`
- E! • `systemstl status iptables`

E! 2. The content of these files:

- E! • `/etc/nsswitch.conf`
- E! • `/etc/ypbind.conf`
- E! • `/etc/idmapd.conf`
- E! • `/etc/sss/*`
- E! • `/etc/krb5.conf`
- E! • `/etc/krb5.keytab`
- E! • `/etc/firewalld/*`
- E! • `/etc/openldap/certs/*`
- E! • `/etc/keystone/keystone-paste.ini`
- E! • `/etc/keystone/logging.conf`
- E! • `/etc/keystone/policy.json`

- E! • /etc/keystone/ssl/certs/\*
- E! • /var/log/keystone/\*
- E! • /var/log/sss/\*
- E! • /var/log/secure/\*
- E! • /var/log/httpd/\*
- E! • /etc/httpd/conf/httpd.conf
- E! • /etc/httpd/conf.d/ssl.conf
- E! • /etc/httpd/conf.d/wsgi-keystone.conf
- E! Files stored in the CCR:
- E! • NSSWITCH\_CONF
- E! • KEYSTONE\_CONF
- E! • YP\_CONF
- E! • SSSD\_CONF
- E! • LDAP\_TLS\_CACERT
- E! • KS\_SIGNING\_CERT
- E! • KS\_SIGNING\_KEY
- E! • KS\_SIGNING\_CACERT
- E! • KS\_SSL\_CERT
- E! • KS\_SSL\_CACERT
- E! • KS\_LDAP\_CACERT
- E! • authccr

#### E! Data gathered for performance on Linux:

E! The following data is always obtained by the `gpfs.snap` command for any enabled protocols.

E! 1. The output of these commands:

- E! • **top -n 1 -b**
- E! • **mmdiag --waiters --iohist --threads --stats --memory**
- E! • **mmfsadm eventsExporter**
- E! • **mmpmon chms**
- E! • **mmfsadm dump nsd**
- E! • **mmfsadm dump mb**
- E! • **mmdumpperfdata -r 86400**

E! 2. The content of these files:

- E! • /opt/IBM/zimon/\*
- E! • /var/log/cnlog/zimon/\*

#### E! Data gathered for core dumps on Linux:

E! The following data is gathered when running `gpfs.snap` with the `--protocol core` argument:

- E! • If `core_pattern` is set to dump to a file it will gather files matching that pattern.
  - E! • If `core_pattern` is set to redirect to `abrt` then everything is gathered from the directory specified in the `abrt.conf` file under `DumpLocation`. If this is not set then `'/var/tmp/abrt'` is used.
  - E! • Other core dump mechanisms are not supported by the script.
  - E! • Any files in the directory `'/var/adm/ras/cores/'` will also be gathered.
- E!

## E! **mmdumpperfdata** command

E! Collects and archives the performance metric information.

### E! **Synopsis**

E! **mmdumpperfdata** [--remove-tree] [*StartTime* *EndTime* | *Duration*]

### E! **Availability**

E! Available with IBM Spectrum Scale Standard Edition or higher.

### E! **Description**

E! The **mmdumpperfdata** command runs all named queries and computed metrics used in the **mmperfmon** **query** command for each cluster node, writes the output into CSV files, and archives all the files in a single .tgz file. The file name is in the `iss_perfdump_YYYYMMDD_hhmmss.tgz` format.

E! The tar archive file contains a folder for each cluster node and within that folder there is a text file with the output of each named query and computed metric.

E! If the start and end time, or duration are not given, then by default the last four hours of metrics information is collected and archived.

### E! **Parameters**

E! **--remove-tree** or **-r**

E! Removes the folder structure that was created for the TAR archive file.

E! *StartTime*

E! Specifies the start timestamp for query in the YYYY-MM-DD[-hh:mm:ss] format.

E! *EndTime*

E! Specifies the end timestamp for query in the YYYY-MM-DD[-hh:mm:ss] format.

E! *Duration*

E! Specifies the duration in seconds

### E! **Exit status**

E! **0** Successful completion.

E! **nonzero**

E! A failure has occurred.

### E! **Security**

E! You must have root authority to run the **mmdumpperfdata** command.

E! The node on which the command is issued must be able to execute remote shell commands on any other node in the cluster without the use of a password and without producing any extraneous messages. For more information, see *Requirements for administering a GPFS file system* in *IBM Spectrum Scale: Administration Guide*.

### E! **Examples**

E! 1. To archive the performance metric information collected for the default time period of last four hours and also delete the folder structure that the command creates, issue this command:

E! `mmdumpperfdata --remove-tree`

E! The system displays output similar to this:



```

E! Using the following options:
E! tstart :
E! tend :
E! duration: 14400
E! rem tree: True
E! Target folder: ./iss_perfdump_20150513_142420
E! [1/120] Dumping data for node=fsc-cc-hs21-22 and query q=swiftAccThroughput
E! file: ./iss_perfdump_20150513_142420/fsc-cc-hs21-22/swiftAccThroughput
E! [2/120] Dumping data for node=fsc-cc-hs21-22 and query q=NetDetails
E! file: ./iss_perfdump_20150513_142420/fsc-cc-hs21-22/NetDetails
E! [3/120] Dumping data for node=fsc-cc-hs21-22 and query q=ctdbCallLatency
E! file: ./iss_perfdump_20150513_142420/fsc-cc-hs21-22/ctdbCallLatency
E! [4/120] Dumping data for node=fsc-cc-hs21-22 and query q=usage
E! file: ./iss_perfdump_20150513_142420/fsc-cc-hs21-22/usage

E! 2. To archive the performance metric information collected for a specific time period, issue this
E! command:
E! mmdumperpdata --remove-tree 2015-01-25-04:04:04 2015-01-26-04:04:04

```

E! The system displays output similar to this:

```

E! Using the following options:
E! tstart : 2015-01-25 04:04:04
E! tend : 2015-01-26 04:04:04
E! duration:
E! rem tree: True
E! Target folder: ./iss_perfdump_20150513_144344
E! [1/120] Dumping data for node=fsc-cc-hs21-22 and query q=swiftAccThroughput
E! file: ./iss_perfdump_20150513_144344/fsc-cc-hs21-22/swiftAccThroughput
E! [2/120] Dumping data for node=fsc-cc-hs21-22 and query q=NetDetails
E! file: ./iss_perfdump_20150513_144344/fsc-cc-hs21-22/NetDetails

```

E! 3. To archive the performance metric information collected in the last 200 seconds, issue this command:

```

E! mmdumperpdata --remove-tree 200

```

E! The system displays output similar to this:

```

E! Using the following options:
E! tstart :
E! tend :
E! duration: 200
E! rem tree: True
E! Target folder: ./iss_perfdump_20150513_144426
E! [1/120] Dumping data for node=fsc-cc-hs21-22 and query q=swiftAccThroughput
E! file: ./iss_perfdump_20150513_144426/fsc-cc-hs21-22/swiftAccThroughput
E! [2/120] Dumping data for node=fsc-cc-hs21-22 and query q=NetDetails
E! file: ./iss_perfdump_20150513_144426/fsc-cc-hs21-22/NetDetails
E! [3/120] Dumping data for node=fsc-cc-hs21-22 and query q=ctdbCallLatency
E! file: ./iss_perfdump_20150513_144426/fsc-cc-hs21-22/ctdbCallLatency
E! [4/120] Dumping data for node=fsc-cc-hs21-22 and query q=usage
E! file: ./iss_perfdump_20150513_144426/fsc-cc-hs21-22/usage
E! [5/120] Dumping data for node=fsc-cc-hs21-22 and query q=smb2IORate
E! file: ./iss_perfdump_20150513_144426/fsc-cc-hs21-22/smb2IORate
E! [6/120] Dumping data for node=fsc-cc-hs21-22 and query q=swiftConLatency
E! file: ./iss_perfdump_20150513_144426/fsc-cc-hs21-22/swiftConLatency
E! [7/120] Dumping data for node=fsc-cc-hs21-22 and query q=swiftCon
E! file: ./iss_perfdump_20150513_144426/fsc-cc-hs21-22/swiftCon
E! [8/120] Dumping data for node=fsc-cc-hs21-22 and query q=gpfsNSDWaits
E! file: ./iss_perfdump_20150513_144426/fsc-cc-hs21-22/gpfsNSDWaits
E! [9/120] Dumping data for node=fsc-cc-hs21-22 and query q=smb2Throughput
E! file: ./iss_perfdump_20150513_144426/fsc-cc-hs21-22/smb2Throughput

```

## E! See also

E! For more information, see **mmpfmon** command in the *IBM Spectrum Scale: Command and Programming Reference*.

## E! Location

E! /usr/lpp/mmfs/bin

## E! mmfsadm command

E! The **mmfsadm** command is intended for use by trained service personnel. IBM suggests you do not run E! this command except under the direction of such personnel.

E! **Note:** The contents of **mmfsadm** output might vary from release to release, which could obsolete any E! user programs that depend on that output. Therefore, we suggest that you do not create user programs E! that invoke **mmfsadm**.

E! The **mmfsadm** command extracts data from GPFS without using locking, so that it can collect the data in E! the event of locking errors. In certain rare cases, this can cause GPFS or the node to fail. Several options E! of this command exist and might be required for use:

### E! cleanup

E! Delete shared segments left by a previously failed GPFS daemon without actually restarting the E! daemon.

### E! dump *what*

E! Dumps the state of a large number of internal state values that might be useful in determining E! the sequence of events. The *what* parameter can be set to **all**, indicating that all available data E! should be collected, or to another value, indicating more restricted collection of data. The output E! is presented to STDOUT and should be collected by redirecting STDOUT. For more information E! about internal GPFS™ states, see the **mmdiag** command in *IBM Spectrum Scale: Command and E! Programming Reference*.

### E! showtrace

E! Shows the current level for each subclass of tracing available in GPFS. Trace level 14 provides the E! highest level of tracing for the class and trace level 0 provides no tracing. Intermediate values E! exist for most classes. More tracing requires more storage and results in a higher probability of E! overlaying the required event.

### E! trace *class n*

E! Sets the trace class to the value specified by *n*. Actual trace gathering only occurs when the E! **mmtracectl** command has been issued.

E! Other options provide interactive GPFS debugging, but are not described here. Output from the E! **mmfsadm** command will be required in almost all cases where a GPFS problem is being reported. The E! **mmfsadm** command collects data only on the node where it is issued. Depending on the nature of the E! problem, **mmfsadm** output might be required from several or all nodes. The **mmfsadm** output from the E! file system manager is often required.

E! To determine where the file system manager is, issue the **mmlsmgr** command:

E! mmlsmgr

E! Output similar to this example is displayed:

```
E! file system      manager node
E! -----
E! fs3              9.114.94.65 (c154n01)
E! fs2              9.114.94.73 (c154n09)
E! fs1              9.114.94.81 (c155n01)
E!
E! Cluster manager node: 9.114.94.65 (c154n01)
```

## E! **Commands for GPFS cluster state information**

E! There are a number of GPFS commands used to obtain cluster state information.

E! The information is organized as follows:

E! • “The **mmafmctl Device getstate** command”

E! • “The **mmdiag** command”

E! • “The **mmgetstate** command”

E! • “The **mmlscluster** command” on page 188

E! • “The **mmlsconfig** command” on page 189

E! • “The **mmrefresh** command” on page 189

E! • “The **mmsdrrestore** command” on page 190

E! • “The **mmexpelnode** command” on page 190

### E! **The mmafmctl Device getstate command**

E! The **mmafmctl Device getstate** command displays the status of active file management cache filesets and gateway nodes.

E! When this command displays a NeedsResync target/fileset state, inconsistencies between home and cache are being fixed automatically; however, unmount and mount operations are required to return the state to Active.

E! The **mmafmctl Device getstate** command is fully described in the *Command reference* section in the *IBM Spectrum Scale: Command and Programming Reference*.

### E! **The mmhealth command**

E! The **mmhealth** command monitors and displays the health status of services hosted on nodes.

E! Use the **mmhealth** command to monitor the health of the node and services hosted on the node in IBM Spectrum Scale. If the status of a service hosted on any node is failed, the **mmhealth** command allows the user to view the event log to analyze and determine the problem. The **mmhealth** command provides list of events responsible for the failure of any service. On detailed analysis of these events a set of troubleshooting steps might be followed to resume the failed service.

E! The **mmhealth** command is fully described in the *mmhealth command* section in the *IBM Spectrum Scale: Command and Programming Reference* and Chapter 3, “Monitoring system health by using the mmhealth command,” on page 85.

### E! **The mmdiag command**

E! The **mmdiag** command displays diagnostic information about the internal GPFS state on the current node.

E! Use the **mmdiag** command to query various aspects of the GPFS internal state for troubleshooting and tuning purposes. The **mmdiag** command displays information about the state of GPFS on the node where it is executed. The command obtains the required information by querying the GPFS daemon process (**mmfsd**), and thus will only function when the GPFS daemon is running.

E! The **mmdiag** command is fully described in the *Command reference* section in *IBM Spectrum Scale: Command and Programming Reference*.

### E! **The mmgetstate command**

E! The **mmgetstate** command displays the state of the GPFS daemon on one or more nodes.

E! These flags are of interest for problem determination:

E! **-a** List all nodes in the GPFS cluster. The option does not display information for nodes that cannot be reached. You may obtain more information if you specify the **-v** option.

E! **-L** Additionally display quorum, number of nodes up, and total number of nodes.

E! The total number of nodes may sometimes be larger than the actual number of nodes in the cluster.  
E! This is the case when nodes from other clusters have established connections for the purposes of mounting a file system that belongs to your cluster.

E! **-s** Display summary information: number of local and remote nodes that have joined in the cluster, number of quorum nodes, and so forth.

E! **-v** Display intermediate error messages.

E! The remaining flags have the same meaning as in the **mmshutdown** command. They can be used to specify the nodes on which to get the state of the GPFS daemon.

E! The GPFS *states* recognized and displayed by this command are:

E! **active**

E! GPFS is ready for operations.

E! **arbitrating**

E! A node is trying to form quorum with the other available nodes.

E! **down**

E! GPFS daemon is not running on the node or is recovering from an internal error.

E! **unknown**

E! Unknown value. Node cannot be reached or some other error occurred.

E! For example, to display the quorum, the number of nodes up, and the total number of nodes, issue:

E! `mmgetstate -L -a`

E! The system displays output similar to:

E! Node number	E! Node name	E! Quorum	E! Nodes up	E! Total nodes	E! GPFS state	E! Remarks
E! 2	k154n06	E! 1*	E! 3	E! 7	E! active	E! quorum node
E! 3	k155n05	E! 1*	E! 3	E! 7	E! active	E! quorum node
E! 4	k155n06	E! 1*	E! 3	E! 7	E! active	E! quorum node
E! 5	k155n07	E! 1*	E! 3	E! 7	E! active	
E! 6	k155n08	E! 1*	E! 3	E! 7	E! active	
E! 9	k1561nx02	E! 1*	E! 3	E! 7	E! active	
E! 11	k155n09	E! 1*	E! 3	E! 7	E! active	

E! where \*, if present, indicates that tiebreaker disks are being used.

E! The **mmgetstate** command is fully described in the *Command reference* section in the *IBM Spectrum Scale: Command and Programming Reference*.

### E! **The mmlscluster command**

E! The **mmlscluster** command displays GPFS cluster configuration information.

E! The syntax of the **mmlscluster** command is:

E! `mmlscluster`

E! The system displays output similar to:

```
E! GPFS cluster information
E! =====
E! GPFS cluster name:      cluster1.kgn.ibm.com
E! GPFS cluster id:       680681562214606028
```

```

E! GPFS UID domain:          cluster1.kgn.ibm.com
E! Remote shell command:    /usr/bin/rsh
E! Remote file copy command: /usr/bin/rcp
E! Repository type:         server-based
E!
E! GPFS cluster configuration servers:
E! -----
E! Primary server:         k164n06.kgn.ibm.com
E! Secondary server:       k164n05.kgn.ibm.com
E!
E! Node  Daemon node name      IP address      Admin node name  Designation
E! -----
E!  1    k164n04.kgn.ibm.com      198.117.68.68  k164n04.kgn.ibm.com  quorum
E!  2    k164n05.kgn.ibm.com      198.117.68.71  k164n05.kgn.ibm.com  quorum
E!  3    k164n06.kgn.ibm.com      198.117.68.70  k164n06.kgn.ibm.com  quorum-manager

```

E! The **mmlscluster** command is fully described in the *Command reference* section in the *IBM Spectrum Scale: Command and Programming Reference*.

### E! The **mmlsconfig** command

E! The **mmlsconfig** command displays current configuration data for a GPFS cluster.

E! Depending on your configuration, additional information not documented in either the **mmcrcluster** command or the **mmchconfig** command may be displayed to assist in problem determination.

E! If a configuration parameter is not shown in the output of this command, the default value for that parameter, as documented in the **mmchconfig** command, is in effect.

E! The syntax of the **mmlsconfig** command is:

```
E! mmlsconfig
```

E! The system displays information similar to:

```

E! Configuration data for cluster cl1.cluster:
E! -----
E! clusterName cl1.cluster
E! clusterId 680752107138921233
E! autoload no
E! minReleaseLevel 4.1.0.0
E! pagepool 1G
E! maxblocksize 4m
E! [c5n97g]
E! pagepool 3500m
E! [common]
E! cipherList EXP-RC4-MD5
E!
E! File systems in cluster cl1.cluster:
E! -----
E! /dev/fs2

```

E! The **mmlsconfig** command is fully described in the *Command reference* section in the *IBM Spectrum Scale: Command and Programming Reference*.

### E! The **mmrefresh** command

E! The **mmrefresh** command is intended for use by experienced system administrators who know how to collect data and run debugging routines.

E! Use the **mmrefresh** command only when you suspect that something is not working as expected and the reason for the malfunction is a problem with the GPFS configuration data. For example, a **mount** command fails with a device not found error, and you know that the file system exists. Another example

E! is if any of the files in the `/var/mmfs/gen` directory were accidentally erased. Under normal E! circumstances, the GPFS command infrastructure maintains the cluster data files automatically and there E! is no need for user intervention.

E! The **mmrefresh** command places the most recent GPFS cluster configuration data files on the specified E! nodes. The syntax of this command is:

```
E! mmrefresh [-f] [-a | -N {Node[,Node...]} | NodeFile | NodeClass]
```

E! The **-f** flag can be used to force the GPFS cluster configuration data files to be rebuilt whether they E! appear to be at the most current level or not. If no other option is specified, the command affects only the E! node on which it is run. The remaining flags have the same meaning as in the **mmsshutdown** command, E! and are used to specify the nodes on which the refresh is to be performed.

E! For example, to place the GPFS cluster configuration data files at the latest level, on all nodes in the E! cluster, issue:

```
E! mmrefresh -a
```

### E! The **mmsdrrestore** command

E! The **mmsdrrestore** command is intended for use by experienced system administrators.

E! The **mmsdrrestore** command restores the latest GPFS system files on the specified nodes. If no nodes are E! specified, the command restores the configuration information only on the node where it is invoked. If E! the local GPFS configuration file is missing, the file specified with the **-F** option from the node specified E! with the **-p** option is used instead.

E! This command works best when used in conjunction with the **mmsdrbackup** user exit, which is E! described in the *GPFS user exits* topic in the *IBM Spectrum Scale: Command and Programming Reference*.

E! For more information, see **mmsdrrestore command** in *IBM Spectrum Scale: Command and Programming E! Reference*.

### E! The **mmexpelnode** command

E! The **mmexpelnode** command instructs the cluster manager to expel the target nodes and to run the E! normal recovery protocol.

E! The cluster manager keeps a list of the expelled nodes. Expelled nodes will not be allowed to rejoin the E! cluster until they are removed from the list using the **-r** or **--reset** option on the **mmexpelnode** command. E! The expelled nodes information will also be reset if the cluster manager node goes down or is changed E! with **mmchmgr -c**.

E! The syntax of the **mmexpelnode** command is:

```
E! mmexpelnode [-o | --once] [-f | --is-fenced] [-w | --wait] -N Node[,Node...]
```

E! Or,

```
E! mmexpelnode {-l | --list}
```

E! Or,

```
E! mmexpelnode {-r | --reset} -N {all | Node[,Node...]}
```

E! The flags used by this command are:

E! **-o | --once**

E! Specifies that the nodes should not be prevented from rejoining. After the recovery protocol E! completes, expelled nodes will be allowed to rejoin the cluster immediately, without the need to first E! invoke **mmexpelnode --reset**.

E! **-f | --is-fenced**

E! Specifies that the nodes are fenced out and precluded from accessing any GPFS disks without first rejoining the cluster (for example, the nodes were forced to reboot by turning off power). Using this flag allows GPFS to start log recovery immediately, skipping the normal 35-second wait.

E! **-w | --wait**

E! Instructs the **mmexpelnode** command to wait until GPFS recovery for the failed node has completed before it runs.

E! **-l | --list**

E! Lists all currently expelled nodes.

E! **-r | --reset**

E! Allows the specified nodes to rejoin the cluster (that is, resets the status of the nodes). To unexpel all of the expelled nodes, issue: **mmexpelnode -r -N all**.

E! **-N {all | Node[,Node...]}**

E! Specifies a list of host names or IP addresses that represent the nodes to be expelled or unexpelled.

E! Specify the daemon interface host names or IP addresses as shown by the **mmlscluster** command.

E! The **mmexpelnode** command does not support administration node names or node classes.

E! **Note:** **-N all** can *only* be used to unexpel nodes.

### E! **Examples of the mmexpelnode command**

E! 1. To expel node c100c1rp3, issue the command:

E! `mmexpelnode -N c100c1rp3`

E! 2. To show a list of expelled nodes, issue the command:

E! `mmexpelnode --list`

E! The system displays information similar to:

E! Node List

E! -----

E! 192.168.100.35 (c100c1rp3.ppd.pok.ibm.com)

E! 3. To allow node c100c1rp3 to rejoin the cluster, issue the command:

E! `mmexpelnode -r -N c100c1rp3`

### E! **GPFS file system and disk information commands**

E! The problem determination tools provided with GPFS for file system, disk and NSD problem determination are intended for use by experienced system administrators who know how to collect data and run debugging routines.

E! The information is organized as follows:

E! • “Restricted mode mount” on page 192

E! • “Read-only mode mount” on page 192

E! • “The lsof command” on page 192

E! • “The mmlsmount command” on page 192

E! • “The mmappypolicy -L command” on page 193

E! • “The mmcheckquota command” on page 199

E! • “The mmlsnsd command” on page 200

E! • “The mmwindisk command” on page 201

E! • “The mmfileid command” on page 201

E! • “The SHA digest” on page 204

## E! **Restricted mode mount**

E! GPFS provides a capability to mount a file system in a restricted mode when significant data structures E! have been destroyed by disk failures or other error conditions.

E! Restricted mode mount is not intended for normal operation, but may allow the recovery of some user E! data. Only data which is referenced by intact directories and metadata structures would be available.

### E! **Attention:**

E! 1. Follow the procedures in “Information to be collected before contacting the IBM Support Center” on E! page 367, and then contact the IBM Support Center before using this capability.

E! 2. Attempt this only after you have tried to repair the file system with the **mmfsck** command. (See E! “Why does the offline mmfsck command fail with "Error creating internal storage"?" on page 135.)

E! 3. Use this procedure only if the failing disk is attached to an AIX or Linux node.

E! Some disk failures can result in the loss of enough metadata to render the entire file system unable to E! mount. In that event it might be possible to preserve some user data through a *restricted mode mount*. This E! facility should only be used if a normal mount does not succeed, and should be considered a last resort E! to save some data after a fatal disk failure.

E! Restricted mode mount is invoked by using the **mmm mount** command with the **-o rs** flags. After a E! restricted mode mount is done, some data *may* be sufficiently accessible to allow copying to another file E! system. The success of this technique depends on the actual disk structures damaged.

## E! **Read-only mode mount**

E! Some disk failures can result in the loss of enough metadata to make the entire file system unable to E! mount. In that event, it might be possible to preserve some user data through a *read-only mode mount*.

E! **Attention:** Attempt this only after you have tried to repair the file system with the **mmfsck** command.

E! This facility should be used only if a normal mount does not succeed, and should be considered a last E! resort to save some data after a fatal disk failure.

E! Read-only mode mount is invoked by using the **mmm mount** command with the **-o ro** flags. After a E! read-only mode mount is done, some data *may* be sufficiently accessible to allow copying to another file E! system. The success of this technique depends on the actual disk structures damaged.

## E! **The lsof command**

E! The **lsof** (list open files) command returns the user processes that are actively using a file system. It is E! sometimes helpful in determining why a file system remains in use and cannot be unmounted.

E! The **lsof** command is available in Linux distributions or by using anonymous ftp from E! **lsof.itap.purdue.edu** (cd to **/pub/tools/unix/lsof**). The inventor of the **lsof** command is Victor A. Abell E! (abe@purdue.edu), Purdue University Computing Center.

## E! **The mmlsmount command**

E! The **mmlsmount** command lists the nodes that have a given GPFS file system mounted.

E! Use the **-L** option to see the node name and IP address of each node that has the file system in use. This E! command can be used for all file systems, all remotely mounted file systems, or file systems mounted on E! nodes of certain clusters.

E! While not specifically intended as a service aid, the **mmlsmount** command is useful in these situations:

E! 1. When writing and debugging new file system administrative procedures, to determine which nodes E! have a file system mounted and which do not.



- E! 2. When mounting a file system on multiple nodes, to determine which nodes have successfully completed the mount and which have not.
- E! 3. When a file system is mounted, but appears to be inaccessible to some nodes but accessible to others, to determine the extent of the problem.
- E! 4. When a normal (not force) unmount has not completed, to determine the affected nodes.
- E! 5. When a file system has force unmounted on some nodes but not others, to determine the affected nodes.

E! For example, to list the nodes having all file systems mounted:

```
E! mmlsmount all -L
```

E! The system displays output similar to:

```
E! File system fs2 is mounted on 7 nodes:
E! 192.168.3.53      c25m3n12      c34.cluster
E! 192.168.110.73   c34f2n01      c34.cluster
E! 192.168.110.74   c34f2n02      c34.cluster
E! 192.168.148.77   c12c4apv7     c34.cluster
E! 192.168.132.123  c20m2n03     c34.cluster      (internal mount)
E! 192.168.115.28   js21n92       c34.cluster      (internal mount)
E! 192.168.3.124    c3m3n14       c3.cluster
E!
E! File system fs3 is not mounted.
E!
E! File system fs3 (c3.cluster:fs3) is mounted on 7 nodes:
E! 192.168.2.11     c2m3n01       c3.cluster
E! 192.168.2.12     c2m3n02       c3.cluster
E! 192.168.2.13     c2m3n03       c3.cluster
E! 192.168.3.123    c3m3n13       c3.cluster
E! 192.168.3.124    c3m3n14       c3.cluster
E! 192.168.110.74   c34f2n02      c34.cluster
E! 192.168.80.20    c21f1n10     c21.cluster
```

E! The **mmlsmount** command is fully described in the *Command reference* section in the *IBM Spectrum Scale: Command and Programming Reference*.

### E! **The mmapplypolicy -L command**

E! Use the **-L** flag of the **mmapplypolicy** command when you are using policy files to manage storage resources and the data stored on those resources. This command has different levels of diagnostics to help debug and interpret the actions of a policy file.

E! The **-L** flag, used in conjunction with the **-I test** flag, allows you to display the actions that would be performed by a policy file without actually applying it. This way, potential errors and misunderstandings can be detected and corrected without actually making these mistakes.

E! These are the trace levels for the **mmapplypolicy -L** flag:

#### E! **Value Description**

- E! **0** Displays only serious errors.
- E! **1** Displays some information as the command runs, but not for each file.
- E! **2** Displays each chosen file and the scheduled action.
- E! **3** Displays the information for each of the preceding trace levels, plus each candidate file and the applicable rule.
- E! **4** Displays the information for each of the preceding trace levels, plus each explicitly excluded file, and the applicable rule.

- E! 5 Displays the information for each of the preceding trace levels, plus the attributes of candidate and excluded files.
- E! 6 Displays the information for each of the preceding trace levels, plus files that are not candidate files, and their attributes.

E! These terms are used:

E! **candidate file**

E! A file that matches a policy rule.

E! **chosen file**

E! A candidate file that has been scheduled for an action.

E! This policy file is used in the examples that follow:

```
E! /* Exclusion rule */
E! RULE 'exclude *.save files' EXCLUDE WHERE NAME LIKE '%.save'
E! /* Deletion rule */
E! RULE 'delete' DELETE FROM POOL 'sp1' WHERE NAME LIKE '%tmp%'
E! /* Migration rule */
E! RULE 'migration to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WHERE NAME LIKE '%file%'
E! /* Typo in rule : removed later */
E! RULE 'exclude 2' EXCULDE
E! /* List rule */
E! RULE EXTERNAL LIST 'tmpfiles' EXEC '/tmp/exec.list'
E! RULE 'all' LIST 'tmpfiles' where name like '%tmp%'
```

E! These are some of the files in file system **/fs1**:

```
E! . .. data1 file.tmp0 file.tmp1 file0 file1 file1.save file2.save
```

E! The **mmapplypolicy** command is fully described in the *Command reference* section in the *IBM Spectrum Scale: Command and Programming Reference*.

E! **mmapplypolicy -L 0:**

E! Use this option to display only serious errors.

E! In this example, there is an error in the policy file. This command:

```
E! mmapplypolicy fs1 -P policyfile -I test -L 0
```

E! produces output similar to this:

```
E! [E:-1] Error while loading policy rules.
E! PCSQLERR: Unexpected SQL identifier token - 'EXCULDE'.
E! PCSQLCTX: at line 8 of 8: RULE 'exclude 2' {{{EXCULDE}}}
E! mmapplypolicy: Command failed. Examine previous error messages to determine cause.
```

E! The error in the policy file is corrected by removing these lines:

```
E! /* Typo in rule */
E! RULE 'exclude 2' EXCULDE
```

E! Now rerun the command:

```
E! mmapplypolicy fs1 -P policyfile -I test -L 0
```

E! No messages are produced because no serious errors were detected.

## E! mmapplypolicy -L 1:

E! Use this option to display all of the information (if any) from the previous level, plus some information E! as the command runs, but not for each file. This option also displays total numbers for file migration and E! deletion.

E! This command:

```
E! mmapplypolicy fs1 -P policyfile -I test -L 1
```

E! produces output similar to this:

```
E! [I] GPFS Current Data Pool Utilization in KB and %
E! sp1      5120  19531264      0.026214%
E! system  102400 19531264      0.524288%
E! [I] Loaded policy rules from policyfile.
E! Evaluating MIGRATE/DELETE/EXCLUDE rules with CURRENT_TIMESTAMP = 2009-03-04@02:40:12 UTC
E! parsed 0 Placement Rules, 0 Restore Rules, 3 Migrate/Delete/Exclude Rules,
E!      1 List Rules, 1 External Pool/List Rules
E! /* Exclusion rule */
E! RULE 'exclude *.save files' EXCLUDE WHERE NAME LIKE '%.save'
E! /* Deletion rule */
E! RULE 'delete' DELETE FROM POOL 'sp1' WHERE NAME LIKE '%tmp%'
E! /* Migration rule */
E! RULE 'migration to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WHERE NAME LIKE '%file%'
E! /* List rule */
E! RULE EXTERNAL LIST 'tmpfiles' EXEC '/tmp/exec.list'
E! RULE 'all' LIST 'tmpfiles' where name like '%tmp%'
E! [I] Directories scan: 10 files, 1 directories, 0 other objects, 0 'skipped' files and/or errors.
E! [I] Inodes scan: 10 files, 1 directories, 0 other objects, 0 'skipped' files and/or errors.
E! [I] Summary of Rule Applicability and File Choices:
E! Rule#  Hit_Cnt  KB_Hit  Chosen  KB_Chosen  KB_Ill  Rule
E!  0      2        32      0        0          0      RULE 'exclude *.save files' EXCLUDE WHERE(.)
E!  1      2        16      2        16         0      RULE 'delete' DELETE FROM POOL 'sp1' WHERE(.)
E!  2      2        32      2        32         0      RULE 'migration to system pool' MIGRATE FROM POOL \
E!                                     'sp1' TO POOL 'system' WHERE(.)
E!  3      2        16      2        16         0      RULE 'all' LIST 'tmpfiles' WHERE(.)
E!
E! [I] Files with no applicable rules: 5.
E!
E! [I] GPFS Policy Decisions and File Choice Totals:
E! Chose to migrate 32KB: 2 of 2 candidates;
E! Chose to premigrate 0KB: 0 candidates;
E! Already co-managed 0KB: 0 candidates;
E! Chose to delete 16KB: 2 of 2 candidates;
E! Chose to list 16KB: 2 of 2 candidates;
E! 0KB of chosen data is illplaced or illreplicated;
E! Predicted Data Pool Utilization in KB and %:
E! sp1      5072  19531264      0.025969%
E! system  102432 19531264      0.524451%
```

## E! mmapplypolicy -L 2:

E! Use this option to display all of the information from the previous levels, plus each chosen file and the E! scheduled migration or deletion action.

E! This command:

```
E! mmapplypolicy fs1 -P policyfile -I test -L 2
```

E! produces output similar to this:

```
E! [I] GPFS Current Data Pool Utilization in KB and %
E! sp1      5120  19531264      0.026214%
E! system  102400 19531264      0.524288%
E! [I] Loaded policy rules from policyfile.
```

```

E! Evaluating MIGRATE/DELETE/EXCLUDE rules with CURRENT_TIMESTAMP = 2009-03-04@02:43:10 UTC
E! parsed 0 Placement Rules, 0 Restore Rules, 3 Migrate/Delete/Exclude Rules,
E!      1 List Rules, 1 External Pool/List Rules
E! /* Exclusion rule */
E! RULE 'exclude *.save files' EXCLUDE WHERE NAME LIKE '%.save'
E! /* Deletion rule */
E! RULE 'delete' DELETE FROM POOL 'sp1' WHERE NAME LIKE '%tmp%'
E! /* Migration rule */
E! RULE 'migration to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WHERE NAME LIKE '%file%'
E! /* List rule */
E! RULE EXTERNAL LIST 'tmpfiles' EXEC '/tmp/exec.list'
E! RULE 'all' LIST 'tmpfiles' where name like '%tmp%'
E! [I] Directories scan: 10 files, 1 directories, 0 other objects, 0 'skipped' files and/or errors.
E! [I] Inodes scan: 10 files, 1 directories, 0 other objects, 0 'skipped' files and/or errors.
E! WEIGHT(INF) LIST 'tmpfiles' /fs1/file.tmp1 SHOW()
E! WEIGHT(INF) LIST 'tmpfiles' /fs1/file.tmp0 SHOW()
E! WEIGHT(INF) DELETE /fs1/file.tmp1 SHOW()
E! WEIGHT(INF) DELETE /fs1/file.tmp0 SHOW()
E! WEIGHT(INF) MIGRATE /fs1/file1 TO POOL system SHOW()
E! WEIGHT(INF) MIGRATE /fs1/file0 TO POOL system SHOW()
E! [I] Summary of Rule Applicability and File Choices:
E! Rule#  Hit_Cnt KB_Hit  Chosen  KB_Chosen  KB_Ill  Rule
E!  0      2      32      0        0        0      RULE 'exclude *.save files' EXCLUDE WHERE(.)
E!  1      2      16      2        16       0      RULE 'delete' DELETE FROM POOL 'sp1' WHERE(.)
E!  2      2      32      2        32       0      RULE 'migration to system pool' MIGRATE FROM POOL \
E!                                     'sp1' TO POOL 'system' WHERE(.)
E!  3      2      16      2        16       0      RULE 'all' LIST 'tmpfiles' WHERE(.)
E!
E! [I] Files with no applicable rules: 5.
E!
E! [I] GPFS Policy Decisions and File Choice Totals:
E! Chose to migrate 32KB: 2 of 2 candidates;
E! Chose to premigrate 0KB: 0 candidates;
E! Already co-managed 0KB: 0 candidates;
E! Chose to delete 16KB: 2 of 2 candidates;
E! Chose to list 16KB: 2 of 2 candidates;
E! 0KB of chosen data is illplaced or illreplicated;
E! Predicted Data Pool Utilization in KB and %:
E! sp1      5072   19531264    0.025969%
E! system 102432 19531264    0.524451%

```

E! where the lines:

```

E! WEIGHT(INF) LIST 'tmpfiles' /fs1/file.tmp1 SHOW()
E! WEIGHT(INF) LIST 'tmpfiles' /fs1/file.tmp0 SHOW()
E! WEIGHT(INF) DELETE /fs1/file.tmp1 SHOW()
E! WEIGHT(INF) DELETE /fs1/file.tmp0 SHOW()
E! WEIGHT(INF) MIGRATE /fs1/file1 TO POOL system SHOW()
E! WEIGHT(INF) MIGRATE /fs1/file0 TO POOL system SHOW()

```

E! show the chosen files and the scheduled action.

### E! mmapplypolicy -L 3:

E! Use this option to display all of the information from the previous levels, plus each candidate file and the applicable rule.

E! This command:

```
E! mmapplypolicy fs1-P policyfile -I test -L 3
```

E! produces output similar to this:

```

E! [I] GPFS Current Data Pool Utilization in KB and %
E! sp1      5120   19531264    0.026214%
E! system 102400 19531264    0.524288%

```

```

E! [I] Loaded policy rules from policyfile.
E! Evaluating MIGRATE/DELETE/EXCLUDE rules with CURRENT_TIMESTAMP = 2009-03-04@02:32:16 UTC
E! parsed 0 Placement Rules, 0 Restore Rules, 3 Migrate/Delete/Exclude Rules,
E!      1 List Rules, 1 External Pool/List Rules
E! /* Exclusion rule */
E! RULE 'exclude *.save files' EXCLUDE WHERE NAME LIKE '%.save'
E! /* Deletion rule */
E! RULE 'delete' DELETE FROM POOL 'sp1' WHERE NAME LIKE '%tmp%'
E! /* Migration rule */
E! RULE 'migration to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WHERE NAME LIKE '%file%'
E! /* List rule */
E! RULE EXTERNAL LIST 'tmpfiles' EXEC '/tmp/exec.list'
E! RULE 'all' LIST 'tmpfiles' where name like '%tmp%'
E! [I] Directories scan: 10 files, 1 directories, 0 other objects, 0 'skipped' files and/or errors.
E! /fs1/file.tmp1 RULE 'delete' DELETE FROM POOL 'sp1' WEIGHT(INF)
E! /fs1/file.tmp1 RULE 'all' LIST 'tmpfiles' WEIGHT(INF)
E! /fs1/file.tmp0 RULE 'delete' DELETE FROM POOL 'sp1' WEIGHT(INF)
E! /fs1/file.tmp0 RULE 'all' LIST 'tmpfiles' WEIGHT(INF)
E! /fs1/file1 RULE 'migration to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WEIGHT(INF)
E! /fs1/file0 RULE 'migration to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WEIGHT(INF)
E! [I] Inodes scan: 10 files, 1 directories, 0 other objects, 0 'skipped' files and/or errors.
E! WEIGHT(INF) LIST 'tmpfiles' /fs1/file.tmp1 SHOW()
E! WEIGHT(INF) LIST 'tmpfiles' /fs1/file.tmp0 SHOW()
E! WEIGHT(INF) DELETE /fs1/file.tmp1 SHOW()
E! WEIGHT(INF) DELETE /fs1/file.tmp0 SHOW()
E! WEIGHT(INF) MIGRATE /fs1/file1 TO POOL system SHOW()
E! WEIGHT(INF) MIGRATE /fs1/file0 TO POOL system SHOW()
E! [I] Summary of Rule Applicability and File Choices:
E! Rule# Hit_Cnt KB_Hit Chosen KB_Chosen KB_Ill Rule
E! 0 2 32 0 0 0 RULE 'exclude *.save files' EXCLUDE WHERE(.)
E! 1 2 16 2 16 0 RULE 'delete' DELETE FROM POOL 'sp1' WHERE(.)
E! 2 2 32 2 32 0 RULE 'migration to system pool' MIGRATE FROM POOL \
E! 'sp1' TO POOL 'system' WHERE(.)
E! 3 2 16 2 16 0 RULE 'all' LIST 'tmpfiles' WHERE(.)
E!
E! [I] Files with no applicable rules: 5.
E!
E! [I] GPFS Policy Decisions and File Choice Totals:
E! Chose to migrate 32KB: 2 of 2 candidates;
E! Chose to premigrate 0KB: 0 candidates;
E! Already co-managed 0KB: 0 candidates;
E! Chose to delete 16KB: 2 of 2 candidates;
E! Chose to list 16KB: 2 of 2 candidates;
E! 0KB of chosen data is illplaced or illreplicated;
E! Predicted Data Pool Utilization in KB and %:
E! sp1 5072 19531264 0.025969%
E! system 102432 19531264 0.524451%

```

E! where the lines:

```

E! /fs1/file.tmp1 RULE 'delete' DELETE FROM POOL 'sp1' WEIGHT(INF)
E! /fs1/file.tmp1 RULE 'all' LIST 'tmpfiles' WEIGHT(INF)
E! /fs1/file.tmp0 RULE 'delete' DELETE FROM POOL 'sp1' WEIGHT(INF)
E! /fs1/file.tmp0 RULE 'all' LIST 'tmpfiles' WEIGHT(INF)
E! /fs1/file1 RULE 'migration to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WEIGHT(INF)
E! /fs1/file0 RULE 'migration to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WEIGHT(INF)

```

E! show the candidate files and the applicable rules.

#### E! mmapplypolicy -L 4:

E! Use this option to display all of the information from the previous levels, plus the name of each explicitly excluded file, and the applicable rule.

E! This command:

E! mmapplypolicy fs1 -P policyfile -I test -L 4

E! produces the following additional information:

```
E! [I] Directories scan: 10 files, 1 directories, 0 other objects, 0 'skipped' files and/or errors.
E! /fs1/file1.save  RULE 'exclude *.save files' EXCLUDE
E! /fs1/file2.save  RULE 'exclude *.save files' EXCLUDE
E! /fs1/file.tmp1   RULE 'delete' DELETE FROM POOL 'sp1' WEIGHT(INF)
E! /fs1/file.tmp1   RULE 'all' LIST 'tmpfiles' WEIGHT(INF)
E! /fs1/file.tmp0   RULE 'delete' DELETE FROM POOL 'sp1' WEIGHT(INF)
E! /fs1/file.tmp0   RULE 'all' LIST 'tmpfiles' WEIGHT(INF)
E! /fs1/file1      RULE 'migration to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WEIGHT(INF)
E! /fs1/file0      RULE 'migration to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WEIGHT(INF)
```

E! where the lines:

```
E! /fs1/file1.save  RULE 'exclude *.save files' EXCLUDE
E! /fs1/file2.save  RULE 'exclude *.save files' EXCLUDE
```

E! indicate that there are two excluded files, **/fs1/file1.save** and **/fs1/file2.save**.

E! **mmapplypolicy -L 5:**

E! Use this option to display all of the information from the previous levels, plus the attributes of candidate E! and excluded files.

E! These attributes include:

- E! • **MODIFICATION\_TIME**
- E! • **USER\_ID**
- E! • **GROUP\_ID**
- E! • **FILE\_SIZE**
- E! • **POOL\_NAME**
- E! • **ACCESS\_TIME**
- E! • **KB\_ALLOCATED**
- E! • **FILESET\_NAME**

E! This command:

E! mmapplypolicy fs1 -P policyfile -I test -L 5

E! produces the following additional information:

```
E! [I] Directories scan: 10 files, 1 directories, 0 other objects, 0 'skipped' files and/or errors.
E! /fs1/file1.save [2009-03-03@21:19:57 0 0 16384 sp1 2009-03-04@02:09:38 16 root] RULE 'exclude \
E! *.save files' EXCLUDE
E! /fs1/file2.save [2009-03-03@21:19:57 0 0 16384 sp1 2009-03-03@21:19:57 16 root] RULE 'exclude \
E! *.save files' EXCLUDE
E! /fs1/file.tmp1 [2009-03-04@02:09:31 0 0 0 sp1 2009-03-04@02:09:31 0 root] RULE 'delete' DELETE \
E! FROM POOL 'sp1' WEIGHT(INF)
E! /fs1/file.tmp1 [2009-03-04@02:09:31 0 0 0 sp1 2009-03-04@02:09:31 0 root] RULE 'all' LIST \
E! 'tmpfiles' WEIGHT(INF)
E! /fs1/file.tmp0 [2009-03-04@02:09:38 0 0 16384 sp1 2009-03-04@02:09:38 16 root] RULE 'delete' \
E! DELETE FROM POOL 'sp1' WEIGHT(INF)
E! /fs1/file.tmp0 [2009-03-04@02:09:38 0 0 16384 sp1 2009-03-04@02:09:38 16 root] RULE 'all' \
E! LIST 'tmpfiles' WEIGHT(INF)
E! /fs1/file1 [2009-03-03@21:32:41 0 0 16384 sp1 2009-03-03@21:32:41 16 root] RULE 'migration \
E! to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WEIGHT(INF)
E! /fs1/file0 [2009-03-03@21:21:11 0 0 16384 sp1 2009-03-03@21:32:41 16 root] RULE 'migration \
E! to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WEIGHT(INF)
```

E! where the lines:

```
E! /fs1/file1.save [2009-03-03@21:19:57 0 0 16384 sp1 2009-03-04@02:09:38 16 root] RULE 'exclude \  
E! *.save files' EXCLUDE  
E! /fs1/file2.save [2009-03-03@21:19:57 0 0 16384 sp1 2009-03-03@21:19:57 16 root] RULE 'exclude \  
E! *.save files' EXCLUDE
```

E! show the attributes of excluded files `/fs1/file1.save` and `/fs1/file2.save`.

E! **mmapplypolicy -L 6:**

E! Use this option to display all of the information from the previous levels, plus files that are not candidate E! files, and their attributes.

E! These attributes include:

- E! • **MODIFICATION\_TIME**
- E! • **USER\_ID**
- E! • **GROUP\_ID**
- E! • **FILE\_SIZE**
- E! • **POOL\_NAME**
- E! • **ACCESS\_TIME**
- E! • **KB\_ALLOCATED**
- E! • **FILESET\_NAME**

E! This command:

```
E! mmapplypolicy fs1 -P policyfile -I test -L 6
```

E! produces the following additional information:

```
E! [I] Directories scan: 10 files, 1 directories, 0 other objects, 0 'skipped' files and/or errors.  
E! /fs1/. [2009-03-04@02:10:43 0 0 8192 system 2009-03-04@02:17:43 8 root] NO RULE APPLIES  
E! /fs1/file1.save [2009-03-03@21:19:57 0 0 16384 sp1 2009-03-04@02:09:38 16 root] RULE \  
E! 'exclude *.save files' EXCLUDE  
E! /fs1/file2.save [2009-03-03@21:19:57 0 0 16384 sp1 2009-03-03@21:19:57 16 root] RULE \  
E! 'exclude *.save files' EXCLUDE  
E! /fs1/file.tmp1 [2009-03-04@02:09:31 0 0 0 sp1 2009-03-04@02:09:31 0 root] RULE 'delete' \  
E! DELETE FROM POOL 'sp1' WEIGHT(INF)  
E! /fs1/file.tmp1 [2009-03-04@02:09:31 0 0 0 sp1 2009-03-04@02:09:31 0 root] RULE 'all' LIST \  
E! 'tmpfiles' WEIGHT(INF)  
E! /fs1/data1 [2009-03-03@21:20:23 0 0 0 sp1 2009-03-04@02:09:31 0 root] NO RULE APPLIES  
E! /fs1/file.tmp0 [2009-03-04@02:09:38 0 0 16384 sp1 2009-03-04@02:09:38 16 root] RULE 'delete' \  
E! DELETE FROM POOL 'sp1' WEIGHT(INF)  
E! /fs1/file.tmp0 [2009-03-04@02:09:38 0 0 16384 sp1 2009-03-04@02:09:38 16 root] RULE 'all' LIST \  
E! 'tmpfiles' WEIGHT(INF)  
E! /fs1/file1 [2009-03-03@21:32:41 0 0 16384 sp1 2009-03-03@21:32:41 16 root] RULE 'migration \  
E! to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WEIGHT(INF)  
E! /fs1/file0 [2009-03-03@21:21:11 0 0 16384 sp1 2009-03-03@21:32:41 16 root] RULE 'migration \  
E! to system pool' MIGRATE FROM POOL 'sp1' TO POOL 'system' WEIGHT(INF)
```

E! where the line:

```
E! /fs1/data1 [2009-03-03@21:20:23 0 0 0 sp1 2009-03-04@02:09:31 0 root] NO RULE APPLIES
```

E! contains information about the **data1** file, which is not a candidate file.

### E! **The mmcheckquota command**

E! The **mmcheckquota** command counts inode and space usage for a file system and writes the collected E! data into quota files.

E! Indications leading you to the conclusion that you should run the **mmcheckquota** command include:

- E! • **MMFS\_QUOTA** error log entries. This error log entry is created when the quota manager has a  
E! problem reading or writing the quota file.

- E! • Quota information is lost due to node failure. Node failure could leave users unable to open files or deny them disk space that their quotas should allow.
- E! • The *in doubt* value is approaching the quota limit. The sum of the *in doubt* value and the current usage may not exceed the hard limit. Consequently, the actual block space and number of files available to the user of the group may be constrained by the *in doubt* value. Should the *in doubt* value approach a significant percentage of the quota, use the **mmcheckquota** command to account for the lost space and files.
- E! • User, group, or fileset quota files are corrupted.

E! During the normal operation of file systems with quotas enabled (not running **mmcheckquota** online), the usage data reflects the actual usage of the blocks and inodes in the sense that if you delete files you should see the usage amount decrease. The *in doubt* value does not reflect how much the user has used already, it is just the amount of quotas that the quota server has assigned to its clients. The quota server does not know whether the assigned amount has been used or not. The only situation where the *in doubt* value is important to the user is when the sum of the usage and the *in doubt* value is greater than the user's quota hard limit. In this case, the user is not allowed to allocate more blocks or inodes unless he brings the usage down.

E! The **mmcheckquota** command is fully described in the *Command reference* section in the *IBM Spectrum Scale: Administration Guide*.

## E! The **mmlsnsd** command

E! The **mmlsnsd** command displays information about the currently defined disks in the cluster.

E! For example, if you issue **mmlsnsd**, your output is similar to this:

```
E! File system  Disk name  NSD servers
E! -----
E! fs2          hd3n97    c5n97g.ppd.pok.ibm.com,c5n98g.ppd.pok.ibm.com,c5n99g.ppd.pok.ibm.com
E! fs2          hd4n97    c5n97g.ppd.pok.ibm.com,c5n98g.ppd.pok.ibm.com,c5n99g.ppd.pok.ibm.com
E! fs2          hd5n98    c5n98g.ppd.pok.ibm.com,c5n97g.ppd.pok.ibm.com,c5n99g.ppd.pok.ibm.com
E! fs2          hd6n98    c5n98g.ppd.pok.ibm.com,c5n97g.ppd.pok.ibm.com,c5n99g.ppd.pok.ibm.com
E! fs2          sdbnsd    c5n94g.ppd.pok.ibm.com,c5n96g.ppd.pok.ibm.com
E! fs2          sdcnsd    c5n94g.ppd.pok.ibm.com,c5n96g.ppd.pok.ibm.com
E! fs2          sddnsd    c5n94g.ppd.pok.ibm.com,c5n96g.ppd.pok.ibm.com
E! fs2          sdensd    c5n94g.ppd.pok.ibm.com,c5n96g.ppd.pok.ibm.com
E! fs2          sdgnsd    c5n94g.ppd.pok.ibm.com,c5n96g.ppd.pok.ibm.com
E! fs2          sdfnsd    c5n94g.ppd.pok.ibm.com,c5n96g.ppd.pok.ibm.com
E! fs2          sdhnsd    c5n94g.ppd.pok.ibm.com,c5n96g.ppd.pok.ibm.com
E! (free disk)  hd2n97    c5n97g.ppd.pok.ibm.com,c5n98g.ppd.pok.ibm.com
```

E! To find out the local device names for these disks, use the **mmlsnsd** command with the **-m** option. For example, issuing **mmlsnsd -m** produces output similar to this:

```
E! Disk name  NSD volume ID  Device  Node name  Remarks
E! -----
E! hd2n97     0972846145C8E924  /dev/hdisk2  c5n97g.ppd.pok.ibm.com  server node
E! hd2n97     0972846145C8E924  /dev/hdisk2  c5n98g.ppd.pok.ibm.com  server node
E! hd3n97     0972846145C8E927  /dev/hdisk3  c5n97g.ppd.pok.ibm.com  server node
E! hd3n97     0972846145C8E927  /dev/hdisk3  c5n98g.ppd.pok.ibm.com  server node
E! hd4n97     0972846145C8E92A  /dev/hdisk4  c5n97g.ppd.pok.ibm.com  server node
E! hd4n97     0972846145C8E92A  /dev/hdisk4  c5n98g.ppd.pok.ibm.com  server node
E! hd5n98     0972846245EB501C  /dev/hdisk5  c5n97g.ppd.pok.ibm.com  server node
E! hd5n98     0972846245EB501C  /dev/hdisk5  c5n98g.ppd.pok.ibm.com  server node
E! hd6n98     0972846245DB3AD8  /dev/hdisk6  c5n97g.ppd.pok.ibm.com  server node
E! hd6n98     0972846245DB3AD8  /dev/hdisk6  c5n98g.ppd.pok.ibm.com  server node
E! hd7n97     0972846145C8E934  /dev/hd7n97  c5n97g.ppd.pok.ibm.com  server node
```

E! To obtain extended information for NSDs, use the **mmlsnsd** command with the **-X** option. For example, issuing **mmlsnsd -X** produces output similar to this:



E!	Disk name	NSD volume ID	Device	Devtype	Node name	Remarks
E!	hd3n97	0972846145C8E927	/dev/hdisk3	hdisk	c5n97g.ppd.pok.ibm.com	server node,pr=no
E!	hd3n97	0972846145C8E927	/dev/hdisk3	hdisk	c5n98g.ppd.pok.ibm.com	server node,pr=no
E!	hd5n98	0972846245EB501C	/dev/hdisk5	hdisk	c5n97g.ppd.pok.ibm.com	server node,pr=no
E!	hd5n98	0972846245EB501C	/dev/hdisk5	hdisk	c5n98g.ppd.pok.ibm.com	server node,pr=no
E!	sdfnsd	0972845E45F02E81	/dev/sdf	generic	c5n94g.ppd.pok.ibm.com	server node
E!	sdfnsd	0972845E45F02E81	/dev/sdm	generic	c5n96g.ppd.pok.ibm.com	server node

E! The **mmlsnsd** command is fully described in the *Command reference* section in the *IBM Spectrum Scale: Administration Guide*.

### E! The **mmwindisk** command

E! On Windows nodes, use the **mmwindisk** command to view all disks known to the operating system E! along with partitioning information relevant to GPFS.

E! For example, if you issue **mmwindisk list**, your output is similar to this:

E!	Disk	Avail	Type	Status	Size	GPFS Partition ID
E!	0		BASIC	ONLINE	137 GiB	
E!	1		GPFS	ONLINE	55 GiB	362DD84E-3D2E-4A59-B96B-BDE64E31ACCF
E!	2		GPFS	ONLINE	200 GiB	BD5E64E4-32C8-44CE-8687-B14982848AD2
E!	3		GPFS	ONLINE	55 GiB	B3EC846C-9C41-4EFD-940D-1AFA6E2D08FB
E!	4		GPFS	ONLINE	55 GiB	6023455C-353D-40D1-BCEB-FF8E73BF6C0F
E!	5		GPFS	ONLINE	55 GiB	2886391A-BB2D-4BDF-BE59-F33860441262
E!	6		GPFS	ONLINE	55 GiB	00845DCC-058B-4DEB-BD0A-17BAD5A54530
E!	7		GPFS	ONLINE	55 GiB	260BCAEB-6E8A-4504-874D-7E07E02E1817
E!	8		GPFS	ONLINE	55 GiB	863B6D80-2E15-457E-B2D5-FEA0BC41A5AC
E!	9	YES	UNALLOC	OFFLINE	55 GiB	
E!	10	YES	UNALLOC	OFFLINE	200 GiB	

E! Where:

#### E! **Disk**

E! is the Windows disk number as shown in the Disk Management console and the DISKPART E! command-line utility.

#### E! **Avail**

E! shows the value **YES** when the disk is available and in a state suitable for creating an NSD.

#### E! **GPFS Partition ID**

E! is the unique ID for the GPFS partition on the disk.

E! The **mmwindisk** command does not provide the NSD volume ID. You can use **mmlsnsd -m** to find the E! relationship between NSDs and devices, which are disk numbers on Windows.

### E! The **mmfileid** command

E! The **mmfileid** command identifies files that are on areas of a disk that are damaged or suspect.

E! **Attention:** Use this command only when the IBM Support Center directs you to do so.

E! Before you run **mmfileid**, you must run a disk analysis utility and obtain the disk sector numbers that E! are damaged or suspect. These sectors are input to the **mmfileid** command.

E! The command syntax is as follows:

```
E! mmfileid Device
E! {-d DiskDesc | -F DescFile}
E! [-o OutputFile] [-f NumThreads] [-t Directory]
E! [-N {Node[,Node...]} | NodeFile | NodeClass}] [--qos QOSClass]
```

E! The input parameters are as follows:

E! *Device*

E! The device name for the file system.

E! **-d** *DiskDesc*

E! A descriptor that identifies the disk to be scanned. *DiskDesc* has the following format:

E! *NodeName:DiskName[:PhysAddr1[-PhysAddr2]]*

E! It has the following alternative format:

E! *:{NsdName|DiskNum|BROKEN}[:PhysAddr1[-PhysAddr2]]*

E! *NodeName*

E! Specifies a node in the GPFS cluster that has access to the disk to scan. You must specify this value if the disk is identified with its physical volume name. Do not specify this value if the disk is identified with its NSD name or its GPFS disk ID number, or if the keyword **BROKEN** is used.

E! *DiskName*

E! Specifies the physical volume name of the disk to scan as known on node *NodeName*.

E! *NsdName*

E! Specifies the GPFS NSD name of the disk to scan.

E! *DiskNum*

E! Specifies the GPFS disk ID number of the disk to scan as displayed by the **mmlsdisk -L** command.

E! **BROKEN**

E! Tells the command to scan all the disks in the file system for files with broken addresses that result in lost data.

E! *PhysAddr1[-PhysAddr2]*

E! Specifies the range of physical disk addresses to scan. The default value for *PhysAddr1* is zero. The default value for *PhysAddr2* is the value for *PhysAddr1*.

E! If both *PhysAddr1* and *PhysAddr2* are zero, the command searches the entire disk.

E! The following lines are examples of valid disk descriptors:

E! k148n07:hdisk9:2206310-2206810

E! :gpfs1008nsd:

E! :10:27645856

E! :BROKEN

E! **-F** *DescFile*

E! Specifies a file that contains a list of disk descriptors, one per line.

E! **-f** *NumThreads*

E! Specifies the number of worker threads to create. The default value is 16. The minimum value is 1.

E! The maximum value is the maximum number allowed by the operating system function

E! **pthread\_create** for a single process. A suggested value is twice the number of disks in the file system.

E! **-N** *{Node[,Node...] | NodeFile | NodeClass}*

E! Specifies the list of nodes that participate in determining the disk addresses. This command supports all defined node classes. The default is **all** or the current value of the **defaultHelperNodes** configuration parameter of the **mmchconfig** command.

E! For general information on how to specify node names, see *Specifying nodes as input to GPFS commands* in the *IBM Spectrum Scale: Administration Guide*.

E! **-o** *OutputFile*

E! The path name of a file to which the result from the **mmfileid** command is to be written. If not specified, the result is sent to standard output.

E! **-t** *Directory*

E! Specifies the directory to use for temporary storage during **mmfileid** command processing. The default directory is **/tmp**.

E! **--qos** *QoSClass*

E! Specifies the Quality of Service for I/O operations (QoS) class to which the instance of the command is assigned. If you do not specify this parameter, the instance of the command is assigned by default to the **maintenance** QoS class. This parameter has no effect unless the QoS service is enabled. For more information, see the help topic on the **mmchqos** command in the *IBM Spectrum Scale: Command and Programming Reference*. Specify one of the following QoS classes:

E! **maintenance**

E! This QoS class is typically configured to have a smaller share of file system IOPS. Use this class for I/O-intensive, potentially long-running GPFS commands, so that they contribute less to reducing overall file system performance.

E! **other** This QoS class is typically configured to have a larger share of file system IOPS. Use this class for administration commands that are not I/O-intensive.

E! For more information, see the help topic on *Setting the Quality of Service for I/O operations (QoS)* in the *IBM Spectrum Scale: Administration Guide*.

E! You can redirect the output to a file with the **-o** flag and sort the output on the inode number with the **sort** command.

E! The **mmfileid** command output contains one line for each inode found to be on a corrupted disk sector. Each line of the command output has this format:

E! *InodeNumber LogicalDiskAddress SnapshotId Filename*

E! *InodeNumber*

E! Indicates the inode number of the file identified by **mmfileid**.

E! *LogicalDiskAddress*

E! Indicates the disk block (disk sector) number of the file identified by **mmfileid**.

E! *SnapshotId*

E! Indicates the snapshot identifier for the file. A *SnapshotId* of 0 means that the file is not a snapshot file.

E! *Filename*

E! Indicates the name of the file identified by **mmfileid**. File names are relative to the root of the file system in which they reside.

E! Assume that a disk analysis tool reports that disks **hdisk6**, **hdisk7**, **hdisk8**, and **hdisk9** contain bad sectors, and that the file **addr.in** has the following contents:

E! k148n07:hdisk9:2206310-2206810

E! k148n07:hdisk8:2211038-2211042

E! k148n07:hdisk8:2201800-2202800

E! k148n01:hdisk6:2921879-2926880

E! k148n09:hdisk7:1076208-1076610

E! You run the following command:

E! **mmfileid /dev/gpfsB -F addr.in**

E! The command output might be similar to the following example:

E! Address 2201958 is contained in the Block allocation map (inode 1)

E! Address 2206688 is contained in the ACL Data file (inode 4, snapId 0)

E! Address 2211038 is contained in the Log File (inode 7, snapId 0)

E! 14336 1076256 0 /gpfsB/tesDir/testFile.out

E! 14344 2922528 1 /gpfsB/x.img

E! The lines that begin with the word Address represent GPFS system metadata files or reserved disk areas.  
E! If your output contains any lines like these, do not attempt to replace or repair the indicated files. If you  
E! suspect that any of the special files are damaged, call the IBM Support Center for assistance.

E! The following line of output indicates that inode number 14336, disk address 1072256 contains file  
E! **/gpfsB/tesDir/testFile.out**. The 0 to the left of the name indicates that the file does not belong to a  
E! snapshot. This file is on a potentially bad disk sector area:

```
E! 14336 1072256 0 /gpfsB/tesDir/testFile.out
```

E! The following line of output indicates that inode number 14344, disk address 2922528 contains file  
E! **/gpfsB/x.img**. The 1 to the left of the name indicates that the file belongs to snapshot number 1. This file  
E! is on a potentially bad disk sector area:

```
E! 14344 2922528 1 /gpfsB/x.img
```

### E! **The SHA digest**

E! The Secure Hash Algorithm (SHA) digest is relevant only when using GPFS in a multi-cluster  
E! environment.

E! The SHA digest is a short and convenient way to identify a key registered with either the **mmauth show**  
E! or **mmremotecoluster** command. In theory, two keys may have the same SHA digest. In practice, this is  
E! extremely unlikely. The SHA digest can be used by the administrators of two GPFS clusters to determine  
E! if they each have received (and registered) the right key file from the other administrator.

E! An example is the situation of two administrators named **Admin1** and **Admin2** who have registered the  
E! others' respective key file, but find that mount attempts by **Admin1** for file systems owned by **Admin2**  
E! fail with the error message: Authorization failed. To determine which administrator has registered the  
E! wrong key, they each run **mmauth show** and send the local clusters SHA digest to the other  
E! administrator. **Admin1** then runs the **mmremotecoluster** command and verifies that the SHA digest for  
E! **Admin2**'s cluster matches the SHA digest for the key that **Admin1** has registered. **Admin2** then runs the  
E! **mmauth show** command and verifies that the SHA digest for **Admin1**'s cluster matches the key that  
E! **Admin2** has authorized.

E! If **Admin1** finds that the SHA digests do not match, **Admin1** runs the **mmremotecoluster update**  
E! command, passing the correct key file as input.

E! If **Admin2** finds that the SHA digests do not match, **Admin2** runs the **mmauth update** command,  
E! passing the correct key file as input.

E! This is an example of the output produced by the **mmauth show all** command:

```
E! Cluster name: fksdcm.pok.ibm.com  
E! Cipher list: EXP1024-RC2-CBC-MD5  
E! SHA digest: d5eb5241eda7d3ec345ece906bfcfe0b6cd343bd  
E! File system access: fs1 (rw, root allowed)  
E!  
E!  
E! Cluster name: kremote.cluster  
E! Cipher list: EXP1024-RC4-SHA  
E! SHA digest: eb71a3aaa89c3979841b363fd6d0a36a2a460a8b  
E! File system access: fs1 (rw, root allowed)  
E!  
E! Cluster name: dkq.cluster (this cluster)  
E! Cipher list: AUTHONLY  
E! SHA digest: 090cd57a2e3b18ac163e5e9bd5f26ffabaa6aa25  
E! File system access: (all rw)
```

---

## E! Collecting details of the issues from performance monitoring tools

E! This topic describes how to collect details of issues that you might encounter in IBM Spectrum Scale by using performance monitoring tools.

E! With IBM Spectrum Scale, system administrators can monitor the performance of GPFS and the communications protocols that it uses. Issue the **mmperfmon query** command to query performance data.

E! **Note:** If you issue the **mmperfmon query** command without any additional parameters, you can see a list of options for querying performance-related information, as shown in the following sample output:

```
E! Usage:
E! mmperfmon query Metric[,Metric...] | Key[,Key...] | NamedQuery [StartTime EndTime | Duration] [Options]
E! OR
E! mmperfmon query compareNodes ComparisonMetric [StartTime EndTime | Duration] [Options]
E! where
E!     Metric                metric name
E!     Key                    a key consisting of node name, sensor group, optional additional filters,
E!                           metric name, separated by pipe symbol
E!                           e.g.: "cluster1.ibm.com|CTDBStats|locking|db_hop_count_bucket_00"
E!     NamedQuery            name of a pre-defined query
E!     ComparisonMetric      name of a metric to be compared if using CompareNodes
E!     StartTime             Start timestamp for query
E!                           Format: YYYY-MM-DD-hh:mm:ss
E!     EndTime               End timestamp for query. Omitted means: execution time
E!                           Format: YYYY-MM-DD-hh:mm:ss
E!     Duration              Number of seconds into the past from today or <EndTime>
E!
E! Options:
E! -h, --help                show this help message and exit
E! -N NodeName, --Node=NodeName
E!                           Defines the node that metrics should be retrieved from
E! -b BucketSize, --bucket-size=BucketSize
E!                           Defines a bucket size (number of seconds), default is
E!                           1
E! -n NumberBuckets, --number-buckets=NumberBuckets
E!                           Number of buckets ( records ) to show, default is 10
E! --filter=Filter           Filter criteria for the query to run
E! --format=Format          Common format for all columns
E! --csv                    Provides output in csv format.
E! --raw                    Provides output in raw format rather than a pretty
E!                           table format.
E! --nice                   Use colors and other text attributes for output.
E! --resolve                Resolve computed metrics, show metrics used
E! --short                  Shorten column names if there are too many to fit into
E!                           one row.
E! --list=List              Show list of specified values (overrides other
E!                           options). Values are all, metrics, computed, queries,
E!                           keys.
E! Possible named queries are:
E!     compareNodes - Compares a single metric across all nodes running sensors
E!     cpu          - Show CPU utilization in system and user space, and context switches
E!     ctdbCallLatency - Show CTDB call latency.
E!     ctdbHopCountDetails - Show CTDB hop count buckets 0 to 5 for one database.
E!     ctdbHopCounts - Show CTDB hop counts (bucket 00 = 1-3 hops) for all databases.
E!     gpfsCRUDopsLatency - Show GPFS CRUD operations latency
E!     gpfsFSWaits - Display max waits for read and write operations for all file systems
E!     gpfsNSDWaits - Display max waits for read and write operations for all disks
E!     gpfsNumberOperations - Get the number of operations to the GPFS file system.
E!     gpfsVFSOpCounts - Display VFS operation counts
E!     netDetails - Get details about the network.
E!     netErrors - Show network problems for all available networks: collisions, drops, errors
E!     nfsErrors - Get the NFS error count for read and write operations
E!     nfsIOLatency - Get the NFS IO Latency in nanoseconds per second
E!     nfsIORate - Get the NFS IOps per second
```

```

E!          nfsQueue - Get the NFS read and write queue size in bytes
E!          nfsThroughput - Get the NFS Throughput in bytes per second
E!    nfsThroughputPerOp - Get the NFS read and write throughput per op in bytes
E!          objAcc - Object account overall performance.
E!          objAccIO - Object account IO details.
E!          objAcclatency - Object proxy Latency.
E!    objAccThroughput - Object account overall Throughput.
E!          objCon - Object container overall performance.
E!          objConIO - Object container IO details.
E!          objConLatency - Object container Latency.
E!    objConThroughput - Object container overall Throughput.
E!          objObj - Object overall performance.
E!          objObjIO - Object overall IO details.
E!          objObjLatency - Object Latency.
E!    objObjThroughput - Object overall Throughput.
E!          objPro - Object proxy overall performance.
E!          objProIO - Object proxy IO details.
E!    objProThroughput - Object proxy overall Throughput.
E!    protocolIOLatency - Compare latency per protocol (smb, nfs, object).
E!    protocolIORate - Get the percentage of total I/O rate per protocol (smb, nfs, object).
E!    protocolThroughput - Get the percentage of total throughput per protocol (smb, nfs, object).
E!    smb2IOLatency - Get the SMB2 I/O latencies per bucket size ( default 1 sec )
E!    smb2IORate - Get the SMB2 I/O rate in number of operations per bucket size ( default 1 sec )
E!    smb2Throughput - Get the SMB2 Throughput in bytes per bucket size ( default 1 sec )
E!    smb2Writes - Count, # of idle calls, bytes in and out and operation time for smb2 writes
E!    smbConnections - Number of smb connections
E!    usage - Show CPU, memory, storage and network usage

```

E! For more information on monitoring performance and analyzing performance related issues, see  
E! “Performance monitoring tool overview” on page 41 and **mmpmon** command in the *IBM Spectrum Scale:  
E! Command and Programming Reference*

---

## Other problem determination tools

Other problem determination tools include the kernel debugging facilities and the **mmpmon** command.

If your problem occurs on the AIX operating system, see AIX in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/ssw\\_aix/welcome](http://www.ibm.com/support/knowledgecenter/ssw_aix/welcome)) and search for the appropriate kernel debugging documentation for information about the AIX **kdb** command.

If your problem occurs on the Linux operating system, see the documentation for your distribution vendor.

If your problem occurs on the Windows operating system, the following tools that are available from the Windows Sysinternals, might be useful in troubleshooting:

- Debugging Tools for Windows
- Process Monitor
- Process Explorer
- Microsoft Windows Driver Kit
- Microsoft Windows Software Development Kit

The **mmpmon** command is intended for system administrators to analyze their I/O on the node on which it is run. It is not primarily a diagnostic tool, but may be used as one for certain problems. For example, running **mmpmon** on several nodes may be used to detect nodes that are experiencing poor performance or connectivity problems.

The syntax of the **mmpmon** command is fully described in the *Command reference* section in the *IBM Spectrum Scale: Command and Programming Reference*. For details on the **mmpmon** command, see “Monitoring GPFS I/O performance with the mmpmon command” on page 3.

---

## Chapter 12. Managing deadlocks

| IBM Spectrum Scale provides functions for automatically detecting potential deadlocks, collecting deadlock debug data, and breaking up deadlocks.

The distributed nature of GPFS, the complexity of the locking infrastructure, the dependency on the proper operation of disks and networks, and the overall complexity of operating in a clustered environment all contribute to increasing the probability of a deadlock.

Deadlocks can be disruptive in certain situations, more so than other type of failure. A deadlock effectively represents a single point of failure that can render the entire cluster inoperable. When a deadlock is encountered on a production system, it can take a long time to debug. The typical approach to recovering from a deadlock involves rebooting all of the nodes in the cluster. Thus, deadlocks can lead to prolonged and complete outages of clusters.

To troubleshoot deadlocks, you must have specific types of debug data that must be collected while the deadlock is in progress. Data collection commands must be run manually before the deadlock is broken. Otherwise, determining the root cause of the deadlock after that is difficult. Also, deadlock detection requires some form of external action, for example, a complaint from a user. Waiting for a user complaint means that detecting a deadlock in progress might take many hours.

| In GPFS V4.1 and later, automated deadlock detection, automated deadlock data collection, and deadlock breakup options are provided to make it easier to handle a deadlock situation.

- | • “Debug data for deadlocks”
- | • “Automated deadlock detection” on page 208
- | • “Automated deadlock data collection” on page 209
- | • “Automated deadlock breakup” on page 210
- | • “Deadlock breakup on demand” on page 211

---

### | Debug data for deadlocks

| Debug data for potential deadlocks is automatically collected. System administrators must monitor and manage the file systems where debug data is stored.

| Automated deadlock detection and automated deadlock data collection are enabled by default.  
| Automated deadlock breakup is disabled by default.

| At the start of the GPFS daemon, the `mmfs.log` file shows entries like the following:

```
| Thu Jul 16 18:50:14.097 2015: [I] Enabled automated deadlock detection.  
| Thu Jul 16 18:50:14.098 2015: [I] Enabled automated deadlock debug data  
| collection.  
| Thu Jul 16 18:50:14.099 2015: [I] Enabled automated expel debug data collection.  
| Thu Jul 16 18:50:14.100 2015: [I] Please see https://ibm.biz/Bd4bNK for more  
| information on deadlock amelioration.
```

| The short URL points to this help topic to make it easier to find the information later.

| By default, debug data is put into the `/tmp/mmfs` directory, or the directory specified for the `dataStructureDump` configuration parameter, on each node. Plenty of disk space, typically many GBs, needs to be available. Debug data is not collected when the directory runs out of disk space.

| **Important:** Before you change the value of **dataStructureDump**, stop the GPFS trace. Otherwise you will lose GPFS trace data. Restart the GPFS trace afterwards.

| After a potential deadlock is detected and the relevant debug data is collected, IBM Service needs to be contacted to report the problem and to upload the debug data. Outdated debug data needs to be removed to make room for new debug data in case a new potential deadlock is detected.

| It is the responsibility of system administrators to manage the disk space under the /tmp/mmfs directory or **dataStructureDump**. They know which set of debug data is still useful.

| The "expel debug data" is similar to the "deadlock debug data", but it is collected when a node is expelled from a cluster for no apparent reasons.

---

## | Automated deadlock detection

| Automated deadlock detection flags unexpected long waiters as potential deadlocks. Effective deadlock detection thresholds are self-tuned to reduce false positive detection. You can register a user program for the **deadlockDetected** event to receive automatic notification.

| GPFS code uses waiters to track what a thread is waiting for and how long it is waiting. Many deadlocks involve long waiters. In a real deadlock, long waiters do not disappear naturally as the deadlock prevents the threads from getting what they are waiting for. With some exceptions, long waiters typically indicate that something in the system is not healthy. A deadlock might be in progress, some disk might be failing, or the entire system might be overloaded.

| Automated deadlock detection monitors waiters to detect potential deadlocks. Some waiters can become long legitimately under normal operating conditions and such waiters are ignored by automated deadlock detection. Such waiters appear in the **mmdiag --waiters** output but never in the **mmdiag --deadlock** output. From now on in this topic, the word *waiters* refers only to those waiters that are monitored by automated deadlock detection.

| Automated deadlock detection flags a waiter as a potential deadlock when the waiter length exceeds certain threshold for deadlock detection. For example, the following mmfs.log entry indicates that a waiter started on thread 8397 at 2015-07-18 09:36:58 passed 905 seconds at Jul 18 09:52:04.626 2015 and is suspected to be a deadlock waiter.

```
| Sat Jul 18 09:52:04.626 2015: [A] Unexpected long waiter detected: Waiting 905.9380 sec since
| 2015-07-18 09:36:58, on node c33f2in01,
| SharedHashTabFetchHandlerThread 8397: on MsgRecordCondvar,
| reason 'RPC wait' for tmMsgTellAcquire1
```

| The /var/log/messages file on Linux and the error log on AIX also log an entry for the deadlock detection, but the mmfs.log file has most details.

| The **deadlockDetected** event is triggered on "Unexpected long waiter detected" and any user program that is registered for the event is invoked. The user program can be made for recording and notification purposes. See /usr/lpp/mmfs/samples/deadlockdetected.sample for an example and more information.

| When the flagged waiter disappears, an entry like the following one might appear in the mmfs.log file:  
| Sat Jul 18 10:00:05.705 2015: [N] The unexpected long waiter on thread 8397 has disappeared in 1386 seconds.

| The **mmdiag --deadlock** command shows the flagged waiter and possibly other waiters closely behind which also passed the threshold for deadlock detection

| If the flagged waiter disappears on its own, without any deadlock breakup actions, then the flagged waiter is not a real deadlock, and the detection is a false positive. A reasonable threshold needs to be



| established to reduce false positive deadlock detection. It is a good practice to consider the trade-off  
| between waiting too long and not having a timely detection and not waiting long enough causing a  
| false-positive detection.

| A false positive deadlock detection and debug data collection are not necessarily a waste of resources. A  
| long waiter, even if it eventually disappears on its own, likely indicates that something is not working  
| well, and is worth looking into.

| The configuration parameter **deadlockDetectionThreshold** is used to specify the initial threshold for  
| deadlock detection. GPFS code adjusts the threshold on each node based on what's happening on the  
| node and cluster. The adjusted threshold is the effective threshold used in automated deadlock detection.

| An internal algorithm is used to evaluate whether a cluster is overloaded or not. Overload is a factor that  
| influences the adjustment of the effective deadlock detection threshold. The effective deadlock detection  
| threshold and the cluster overload index are shown in the output of the `mmdiag --deadlock`.

```
| Effective deadlock detection threshold on c37f2n04 is 1000 seconds  
| Effective deadlock detection threshold on c37f2n04 is 430 seconds for short waiters  
| Cluster my.cluster is overloaded. The overload index on c40bbc2xn2 is 1.14547
```

| If **deadlockDetectionThresholdForShortWaiters** is positive, and it is by default, certain waiters, including  
| most of the mutex waiters, are considered short waiters that should not be long. These short waiters have  
| a shorter effective deadlock detection threshold that is self-tuned separately.

| Certain waiters, including most of the mutex waiters, are considered short waiters that should not be  
| long. If **deadlockDetectionThresholdForShortWaiters** is positive, and it is by default, these short waiters  
| are monitored separately. Their effective deadlock detection threshold is also self-tuned separately.

| The overload index is the weighted average duration of all I/Os completed over a long time. Recent I/O  
| durations count more than the ones in the past. The cluster overload detection affects deadlock  
| amelioration functions only. The determination by GPFS that a cluster is overloaded is not necessarily the  
| same as the determination by a customer. But customers might use the determination by GPFS as a  
| reference and check the workload, hardware and network of the cluster to see whether anything needs  
| correction or adjustment. An overloaded cluster with a workload far exceeding its resource capability is  
| not healthy nor productive.

| If the existing effective deadlock detection threshold value is no longer appropriate for the workload, run  
| the **mmfsadm resetstats** command to restart the local adjustment.

| To view the current value of **deadlockDetectionThreshold** and  
| **deadlockDetectionThresholdForShortWaiters**, which are the initial thresholds for deadlock detection,  
| enter the following command:

```
| mmfsconfig deadlockDetectionThreshold  
| mmfsconfig deadlockDetectionThresholdForShortWaiters
```

| The system displays output similar to the following:

```
| deadlockDetectionThreshold 300  
| deadlockDetectionThresholdForShortWaiters 60
```

| To disable automated deadlock detection, specify a value of 0 for **deadlockDetectionThreshold**. All  
| deadlock amelioration functions, not just deadlock detection, are disabled by specifying 0 for  
| **deadlockDetectionThreshold**. A positive value must be specified for **deadlockDetectionThreshold** to  
| enable any part of the deadlock amelioration functions.

---

## | **Automated deadlock data collection**

| Automated deadlock data collection gathers crucial debug data when a potential deadlock is detected.

| Automated deadlock data collection helps gather crucial debug data on detection of a potential deadlock.  
| Messages similar to the following ones are written to the `mmfs.log` file:

```
| Sat Jul 18 09:52:04.626 2015: [A] Unexpected long waiter detected:  
| 2015-07-18 09:36:58: waiting 905.938 seconds on node c33f2in01:  
| SharedHashTabFetchHandlerThread 8397: on MsgRecordCondvar,  
| reason 'RPC wait' for tmMsgTellAcquire1  
| Sat Jul 18 09:52:04.627 2015: [I] Initiate debug data collection from  
| this node.  
| Sat Jul 18 09:52:04.628 2015: [I] Calling User Exit Script  
| gpfsDebugDataCollection: event deadlockDebugData,  
| Async command /usr/lpp/mmfs/bin/mmcommon.
```

| What debug data is collected depends on the value of the configuration parameter **debugDataControl**.  
| The default value is **light** and a minimum amount of debug data, the data that is most frequently needed  
| to debug a GPFS issue, is collected. The value **medium** gets more debug data collected. The value **heavy**  
| is meant to be used routinely by internal test teams only. The value **verbose** needed only for  
| troubleshooting special cases and can result in very large dumps. No debug data is collected when the  
| value **none** is specified. You can set different values for the **debugDataControl** parameter across nodes in  
| the cluster.

| Automated deadlock data collection is enabled by default and controlled by the configuration parameter  
| **deadlockDataCollectionDailyLimit**. This parameter specifies the maximum number of times debug data  
| can be collected in a 24-hour period by automated deadlock data collection

| To view the current value of **deadlockDataCollectionDailyLimit**, enter the following command:

```
| mmfsconfig deadlockDataCollectionDailyLimit
```

| The system displays output similar to the following:

```
| deadlockDataCollectionDailyLimit 3
```

| To disable automated deadlock data collection, specify a value of 0 for  
| **deadlockDataCollectionDailyLimit**.

| Another configuration parameter, **deadlockDataCollectionMinInterval**, is used to control the minimum  
| amount of time between consecutive debug data collections. The default is 3600 seconds or 1 hour.

---

## Automated deadlock breakup

Automated deadlock breakup helps resolve a deadlock situation without human intervention. To break up a deadlock, less disruptive actions are tried first; for example, causing a file system panic. If necessary, more disruptive actions are then taken; for example, shutting down a GPFS **mmfsd** daemon.

If a system administrator prefers to control the deadlock breakup process, the **deadlockDetected** callback can be used to notify system administrators that a potential deadlock was detected. The information from the **mmdiag --deadlock** section can then be used to help determine what steps to take to resolve the deadlock.

Automated deadlock breakup is disabled by default and controlled with the **mmchconfig** attribute **deadlockBreakupDelay**. The **deadlockBreakupDelay** attribute specifies how long to wait after a deadlock is detected before attempting to break up the deadlock. Enough time must be provided to allow the debug data collection to complete. To view the current breakup delay, enter the following command:

```
mmfsconfig deadlockBreakupDelay
```

The system displays output similar to the following:

```
deadlockBreakupDelay 0
```

The value of 0 shows that automated deadlock breakup is disabled. To enable automated deadlock breakup, specify a positive value for **deadlockBreakupDelay**. If automated deadlock breakup is to be enabled, a delay of 300 seconds or longer is recommended.

Automated deadlock breakup is done on a node-by-node basis. If automated deadlock breakup is enabled, the breakup process is started when the suspected deadlock waiter is detected on a node. The process first waits for the **deadlockBreakupDelay**, and then goes through various phases until the deadlock waiters disappear. There is no central coordination on the deadlock breakup, so the time to take deadlock breakup actions may be different on each node. Breaking up a deadlock waiter on one node can cause some deadlock waiters on other nodes to disappear, so no breakup actions need to be taken on those other nodes.

If a suspected deadlock waiter disappears while waiting for the **deadlockBreakupDelay**, the automated deadlock breakup process stops immediately without taking any further action. To lessen the number of breakup actions that are taken in response to detecting a false-positive deadlock, increase the **deadlockBreakupDelay**. If you decide to increase the **deadlockBreakupDelay**, a deadlock can potentially exist for a longer period.

If your goal is to break up a deadlock as soon as possible, and your workload can afford an interruption at any time, then enable automated deadlock breakup from the beginning. Otherwise, keep automated deadlock breakup disabled to avoid unexpected interruptions to your workload. In this case, you can choose to break the deadlock manually, or use the function that is described in the “Deadlock breakup on demand” topic.

Due to the complexity of the GPFS code, asserts or segmentation faults might happen during a deadlock breakup action. That might cause unwanted disruptions to a customer workload still running normally on the cluster. A good reason to use deadlock breakup on demand is to not disturb a partially working cluster until it is safe to do so. Try not to break up a suspected deadlock prematurely to avoid unnecessary disruptions. If automated deadlock breakup is enabled all of the time, it is good to set **deadlockBreakupDelay** to a large value such as 3600 seconds. If using **mmcommon breakDeadlock**, it is better to wait until the longest deadlock waiter is an hour or longer. Much shorter times can be used if a customer prefers fast action in breaking a deadlock over assurance that a deadlock is real.

The following messages, related to deadlock breakup, might be found in the **mmfs.log** files:

```
[I] Enabled automated deadlock breakup.  
[N] Deadlock breakup: starting in 300 seconds  
[N] Deadlock breakup: aborting RPC on 1 pending nodes.  
[N] Deadlock breakup: panicking fs fs1  
[N] Deadlock breakup: shutting down this node.  
[N] Deadlock breakup: the process has ended.
```

---

## Deadlock breakup on demand

Deadlocks can be broken up on demand, which allows a system administrator to choose the appropriate time to start the breakup actions.

A deadlock can be localized, for example, it might involve only one of many file systems in a cluster. The other file systems in the cluster can still be used, and a mission critical workload might need to continue uninterrupted. In these cases, the best time to break up the deadlock is after the mission critical workload ends.

The **mmcommon breakDeadlock** command can be used to break up an existing deadlock in a cluster when the deadlock was previously detected by deadlock amelioration. To start the breakup on demand, use the following syntax:

```
mmcommon breakDeadlock [-N {Node[,Node...] | NodeFile | NodeClass}]
```

If the **mmcommon breakDeadlock** command is issued without the **-N** parameter, then every node in the cluster receives a request to take action on any long waiter that is a suspected deadlock.

If the **mmcommon breakDeadlock** command is issued with the **-N** parameter, then only the nodes that are specified receive a request to take action on any long waiter that is a suspected deadlock. For example, assume that there are two nodes, called **node3** and **node6**, that require a deadlock breakup. To send the breakup request to just these nodes, issue the following command:

```
mmcommon breakDeadlock -N node3,node6
```

Shortly after running the **mmcommon breakDeadlock** command, issue the following command:

```
mmdsh -N all /usr/lpp/mmfs/bin/mmdiag --deadlock
```

The output of the **mmdsh** command can be used to determine if any deadlock waiters still exist and if any additional actions are needed.

The effect of the **mmcommon breakDeadlock** command only persists on a node until the longest deadlock waiter that was detected disappears. All actions that are taken by **mmcommon breakDeadlock** are recorded in the **mmfs.log** file. When **mmcommon breakDeadlock** is issued for a node that did not have a deadlock, no action is taken except for recording the following message in the **mmfs.log** file:

```
[N] Received deadlock breakup request from 192.168.40.72: No deadlock to break up.
```

The **mmcommon breakDeadlock** command provides more control over breaking up deadlocks, but multiple breakup requests might be required to achieve satisfactory results. All waiters that exceeded the **deadlockDetectionThreshold** might not disappear when **mmcommon breakDeadlock** completes on a node. In complicated deadlock scenarios, some long waiters can persist after the longest waiters disappear. Waiter length can grow to exceed the **deadlockDetectionThreshold** at any point, and waiters can disappear at any point as well. Examine the waiter situation after **mmcommon breakDeadlock** completes to determine whether the command must be repeated to break up the deadlock.

Another way to break up a deadlock on demand is to enable automated deadlock breakup by changing **deadlockBreakupDelay** to a positive value. By enabling automated deadlock breakup, breakup actions are initiated on existing deadlock waiters. The breakup actions repeat automatically if deadlock waiters are detected. Change **deadlockBreakupDelay** back to 0 when the results are satisfactory, or when you want to control the timing of deadlock breakup actions again. If automated deadlock breakup remains enabled, breakup actions start on any newly detected deadlocks without any intervention.

---

## Chapter 13. Installation and configuration issues

You might encounter errors with GPFS installation, configuration, and operation. Use the information in this topic to help you identify and correct errors.

An IBM Spectrum Scale installation problem should be suspected when GPFS modules are not loaded successfully, commands do not work, either on the node that you are working on or on other nodes, new command operands added with a new release of IBM Spectrum Scale are not recognized, or there are problems with the kernel extension.

A GPFS configuration problem should be suspected when the GPFS daemon will not activate, it will not remain active, or it fails on some nodes but not on others. Suspect a configuration problem also if quorum is lost, certain nodes appear to hang or do not communicate properly with GPFS, nodes cannot be added to the cluster or are expelled, or GPFS performance is very noticeably degraded once a new release of GPFS is installed or configuration parameters have been changed.

These are some of the errors encountered with GPFS installation, configuration and operation:

- “Post installation and configuration problems” on page 226
- “GPFS modules cannot be loaded on Linux” on page 232
- “GPFS daemon will not come up” on page 232
- “GPFS daemon went down” on page 236
- “IBM Spectrum Scale failures due to a network failure” on page 243
- “Kernel panics with a 'GPFS dead man switch timer has expired, and there's still outstanding I/O requests' message” on page 291
- “Quorum loss” on page 239
- “Delays and deadlocks” on page 344
- “Node cannot be added to the GPFS cluster” on page 227
- “Remote node expelled after remote file system successfully mounted” on page 250
- Chapter 19, “Disaster recovery issues,” on page 325
- “GPFS commands are unsuccessful” on page 237
- “Application program errors” on page 239
- E! • “Windows issues” on page 241
- “OpenSSH connection delays” on page 243

---

### **Resolving most frequent problems related to installation, deployment, and upgrade**

Use the following information to resolve the most frequent problems related to installation, deployment, and upgrade.

### **Finding deployment related error messages more easily and using them for failure analysis**

Use this information to find and analyze error messages related to installation, deployment, and upgrade from the respective logs when using the **spectrumscale** installation toolkit.

In case of any installation, deployment, and upgrade related error:

1. Go to the end of the corresponding log file and search upwards for the text FATAL.

2. Find the topmost occurrence of FATAL (or first FATAL error that occurred) and look above and below this error for further indications of the failure.

Error messages at the bottom of the installation, deployment, and upgrade related logs are specific to the Chef component which controls the entire activity and therefore they are not typically the first place to look during failure analysis. For more information, see the following examples:

- “Example 1 - Installation failed and the bottom of the log file contains the following Chef output, which is not indicative of the error”
- “Example 2 - Deployment failed and the bottom of the log file contains the following Chef output, which is not indicative of the error” on page 216

### Example 1 - Installation failed and the bottom of the log file contains the following Chef output, which is not indicative of the error

```
2016-01-28 09:29:21,839 [ TRACE ] Stopping chef zero
2016-01-28 09:29:21,839 [ ERROR ] The following error was encountered:
Traceback (most recent call last):
  File "/usr/lpp/mmfs/4.2.0.1/installer/espilib/reporting.py", line 193, in log_to_file
    yield handler
  File "/usr/lpp/mmfs/4.2.0.1/installer/espilib/install.py", line 152, in _install
    setup.install(config)
  File "/usr/lpp/mmfs/4.2.0.1/installer/espilib/setup/gpfs.py", line 481, in install
    self.deploy(config.admin_nodes[0], recipe, attributes)
  File "/usr/lpp/mmfs/4.2.0.1/installer/espilib/connectionmanager.py", line 52, in deploy
    ssh_identity=self._get_ssh_identity()
  File "/usr/lpp/mmfs/4.2.0.1/installer/espilib/deploy.py", line 108, in deploy_nodes
    raise DeployError()
DeployError: Installation failed on one or more nodes. Check the log for more details.
2016-01-28 09:29:21,927 [ INFO ] Detailed error log:
/usr/lpp/mmfs/4.2.0.1/installer/logs/INSTALL-28-01-2016_09:05:58.log
```

1. To find more details, go to the end of the log file and search upwards for the text FATAL.  
In this example, the first search hit is the last instance of the text FATAL in the log file that is being searched. The output typically shows what was printed to the screen and gives a general indication of where the failure occurred. It is also helpful to search for the terms fail and error. In this case, the failure occurred while creating the GPFS cluster with the default profile:

```
2016-01-28 09:28:52,994 [ FATAL ]
localhost.localdomain failure whilst: Creating GPFS cluster with default profile (SS04)
```

2. Search further upwards for the text FATAL to find its first occurrence in the log file.  
In this example, the text FATAL is found 3 times, wherein the following is its first occurrence in the log file:

**Note:** The following log text has been adjusted to fit in the PDF margin.

```
2016-01-28 09:28:52,787 [ TRACE ] localhost.localdomain
[2016-01-28T09:28:52+00:00] ERROR: Running exception handlers
2016-01-28 09:28:52,787 [ TRACE ] localhost.localdomain Running handlers complete
2016-01-28 09:28:52,788 [ TRACE ] localhost.localdomain
#[0m[2016-01-28T09:28:52+00:00] ERROR: Exception handlers complete
2016-01-28 09:28:52,788 [ TRACE ] localhost.localdomain
[2016-01-28T09:28:52+00:00] FATAL: Stacktrace dumped to /var/chef/cache/chef-stacktrace.out
2016-01-28 09:28:52,788 [ TRACE ] localhost.localdomain Chef Client failed.
3 resources updated in 14.197169001 seconds#[0m
2016-01-28 09:28:52,826 [ TRACE ] localhost.localdomain
[2016-01-28T09:28:52+00:00] ERROR: execute[create_gpfs_cluster_default_profile]
(gpfs::gpfs_cluster_create line 20) had an error: Mixlib::ShellOut::ShellCommandFailed:
Expected process to exit with [0], but received '1'
```

This log snippet mentions the exact Chef recipe (`gpfs::gpfs_cluster_create`) that failed during install.

3. To find more information, visually search upwards within the log file.  
Root cause output is typically close to this first occurrence (time-wise) of the text FATAL. Following is a snippet of the log text above the first occurrence of FATAL. It shows the start of the **Creating GPFS cluster** portion and then shows where the first error occurred: a stanza encapsulated in "=====" symbols. Immediately following this is the command executed by the **spectrumscale** installation toolkit:

```
/usr/lpp/mmfs/bin/mmcrccluster -N /tmp/NodesDesc -r /usr/bin/ssh -R /usr/bin/scp \
-C spectrumscale.example.com --profile gpfsprotocoldefaults
```

Following that is a STDERR: Warning when adding the host details to the list of known hosts. Because of this the installation has failed.

**Note:** The following log text has been adjusted to fit in the PDF margin.

```
2016-01-28 09:28:44,583 [ INFO ] [localhost.localdomain 28-01-2016 09:28:44]
IBM SPECTRUM SCALE: Creating GPFS cluster with default profile (SS04)
2016-01-28 09:28:44,583 [ TRACE ] localhost.localdomain
#[0m * log[IBM SPECTRUM SCALE: Creating GPFS cluster with default profile (SS04).] action write
2016-01-28 09:28:44,583 [ TRACE ] localhost.localdomain
2016-01-28 09:28:52,778 [ TRACE ] localhost.localdomain
#[0m * execute[create_gpfs_cluster_default_profile] action run
2016-01-28 09:28:52,778 [ TRACE ] localhost.localdomain #[0m
2016-01-28 09:28:52,779 [ TRACE ] localhost.localdomain =====#[0m
2016-01-28 09:28:52,779 [ TRACE ] localhost.localdomain #[31mError executing action `run` on resource
'execute[create_gpfs_cluster_default_profile]'#[0m
2016-01-28 09:28:52,779 [ TRACE ] localhost.localdomain =====#[0m
2016-01-28 09:28:52,779 [ TRACE ] localhost.localdomain
2016-01-28 09:28:52,779 [ TRACE ] localhost.localdomain #[0m Mixlib::ShellOut::ShellCommandFailed#[0m
2016-01-28 09:28:52,779 [ TRACE ] localhost.localdomain -----#[0m
2016-01-28 09:28:52,779 [ TRACE ] localhost.localdomain Expected process to exit with [0], but received '1'
2016-01-28 09:28:52,779 [ TRACE ] localhost.localdomain #[0m
---- Begin output of /usr/lpp/mmfs/bin/mmcrlcluster -N /tmp/NodesDesc -r /usr/bin/ssh -R /usr/bin/scp \
-C spectrumscale.example.com --profile gpfsprotocoldefaults ----
2016-01-28 09:28:52,780 [ TRACE ] localhost.localdomain #[0m
STDOUT: mmcrlcluster: Performing preliminary node verification ...
2016-01-28 09:28:52,780 [ TRACE ] localhost.localdomain #[0m
mmcrlcluster: Processing quorum and other critical nodes ...
2016-01-28 09:28:52,781 [ TRACE ] localhost.localdomain #[0m
STDERR: spectrumscale-102.example.com:
Warning: Permanently added 'spectrumscale-102.example.com,192.168.100.102' (ECDSA) to the list of known hosts.
2016-01-28 09:28:52,781 [ TRACE ] localhost.localdomain #[0m
spectrumscale-102.example.com:
checkNewClusterNode:success:%home%:20_MEMBER_NODE::0:1:localhost:%3A%3A1:
localhost:manager:::::localhost:localhost:1502:4.2.0.1:Linux:Q:::::
2016-01-28 09:28:52,781 [ TRACE ] localhost.localdomain #[0m
spectrumscale-103.example.com:
Warning: Permanently added 'spectrumscale-103.example.com,192.168.100.103'
(ECDSA) to the list of known hosts.
2016-01-28 09:28:52,781 [ TRACE ] localhost.localdomain #[0m
spectrumscale-103.example.com:
checkNewClusterNode:success:%home%:20_MEMBER_NODE::0:1:localhost:%3A%3A1:
localhost:manager:::::localhost:localhost:1502:4.2.0.1:Linux:Q:::::
2016-01-28 09:28:52,781 [ TRACE ] localhost.localdomain #[0m
mmcrlcluster: Removing GPFS cluster files from the nodes in the cluster . . .
2016-01-28 09:28:52,781 [ TRACE ] localhost.localdomain #[0m
mmcrlcluster: Command failed. Examine previous error messages to determine cause.
2016-01-28 09:28:52,782 [ TRACE ] localhost.localdomain #[0m
---- End output of /usr/lpp/mmfs/bin/mmcrlcluster -N /tmp/NodesDesc -r /usr/bin/ssh -R /usr/bin/scp \
-C spectrumscale.example.com --profile gpfsprotocoldefaults ----
2016-01-28 09:28:52,782 [ TRACE ] localhost.localdomain #[0m
Ran /usr/lpp/mmfs/bin/mmcrlcluster -N /tmp/NodesDesc -r /usr/bin/ssh -R /usr/bin/scp \
-C spectrumscale.example.com --profile gpfsprotocoldefaults returned 1#[0m
2016-01-28 09:28:52,782 [ TRACE ] localhost.localdomain
2016-01-28 09:28:52,782 [ TRACE ] localhost.localdomain #[0m Resource Declaration:#[0m
2016-01-28 09:28:52,782 [ TRACE ] localhost.localdomain -----#[0m
2016-01-28 09:28:52,782 [ TRACE ] localhost.localdomain #
In /var/chef/cache/cookbooks/gpfs/recipes/gpfs_cluster_create.rb
2016-01-28 09:28:52,782 [ TRACE ] localhost.localdomain #[0m
2016-01-28 09:28:52,782 [ TRACE ] localhost.localdomain #[0m
20: execute 'create_gpfs_cluster_default_profile' do
2016-01-28 09:28:52,782 [ TRACE ] localhost.localdomain #[0m
21: command "#{node['gpfs']['gpfs_path']}/mmcrlcluster -N /tmp/NodesDesc -r
#{node['gpfs']['RemoteShellCommand']} -R #{node['gpfs']['RemoteFileCopy']}"
-C #{node['gpfs']['cluster_name']} --profile gpfsprotocoldefaults "
2016-01-28 09:28:52,782 [ TRACE ] localhost.localdomain #[0m
22: not_if { node['gpfs']['profile'] == 'randomio' }
2016-01-28 09:28:52,783 [ TRACE ] localhost.localdomain #[0m
23: not_if "#{node['gpfs']['gpfs_path']}/mmfscscluster"
2016-01-28 09:28:52,783 [ TRACE ] localhost.localdomain #[0m
24: action :run
2016-01-28 09:28:52,783 [ TRACE ] localhost.localdomain #[0m
25: end
2016-01-28 09:28:52,783 [ TRACE ] localhost.localdomain #[0m
26:
2016-01-28 09:28:52,783 [ TRACE ] localhost.localdomain #[0m
2016-01-28 09:28:52,784 [ TRACE ] localhost.localdomain #[0m Compiled Resource:#[0m
2016-01-28 09:28:52,784 [ TRACE ] localhost.localdomain -----#[0m
2016-01-28 09:28:52,784 [ TRACE ] localhost.localdomain
# Declared in /var/chef/cache/cookbooks/gpfs/recipes/gpfs_cluster_create.rb:20:in `from_file'
2016-01-28 09:28:52,784 [ TRACE ] localhost.localdomain #[0m
2016-01-28 09:28:52,784 [ TRACE ] localhost.localdomain #[0m
execute("create_gpfs_cluster_default_profile") do
2016-01-28 09:28:52,784 [ TRACE ] localhost.localdomain #[0m
action [:run]
2016-01-28 09:28:52,784 [ TRACE ] localhost.localdomain #[0m
retries 0
2016-01-28 09:28:52,784 [ TRACE ] localhost.localdomain #[0m
```

```

retry_delay 2
2016-01-28 09:28:52,785 [ TRACE ] localhost.localdomain #[0m
default_guard_interpreter :execute
2016-01-28 09:28:52,785 [ TRACE ] localhost.localdomain #[0m
command "/usr/lpp/mmfs/bin/mmcrccluster -N /tmp/NodesDesc -r /usr/bin/ssh -R /usr/bin/scp \
-C spectrumscale.example.com --profile gpfsprotocoldefaults "
2016-01-28 09:28:52,785 [ TRACE ] localhost.localdomain #[0m
backup 5
2016-01-28 09:28:52,785 [ TRACE ] localhost.localdomain #[0m
returns 0
2016-01-28 09:28:52,785 [ TRACE ] localhost.localdomain #[0m
declared_type :execute
2016-01-28 09:28:52,785 [ TRACE ] localhost.localdomain #[0m
cookbook name "gpfs"
2016-01-28 09:28:52,786 [ TRACE ] localhost.localdomain #[0m
recipe_name "gpfs_cluster_create"
2016-01-28 09:28:52,786 [ TRACE ] localhost.localdomain #[0m
not_if { #code block }
2016-01-28 09:28:52,786 [ TRACE ] localhost.localdomain #[0m
not_if "/usr/lpp/mmfs/bin/mmlscluster"
2016-01-28 09:28:52,786 [ TRACE ] localhost.localdomain #[0m end
2016-01-28 09:28:52,787 [ TRACE ] localhost.localdomain #[0m
2016-01-28 09:28:52,787 [ TRACE ] localhost.localdomain #[0m#[0m
2016-01-28 09:28:52,787 [ TRACE ] localhost.localdomain Running handlers:#[0m
2016-01-28 09:28:52,787 [ TRACE ] localhost.localdomain [2016-01-28T09:28:52+00:00]
ERROR: Running exception handlers
2016-01-28 09:28:52,787 [ TRACE ] localhost.localdomain Running handlers complete
2016-01-28 09:28:52,788 [ TRACE ] localhost.localdomain #[0m[2016-01-28T09:28:52+00:00]
ERROR: Exception handlers complete
2016-01-28 09:28:52,788 [ TRACE ] localhost.localdomain [2016-01-28T09:28:52+00:00]
FATAL: Stacktrace dumped to /var/chef/cache/chef-stacktrace.out
2016-01-28 09:28:52,788 [ TRACE ] localhost.localdomain Chef Client failed.
3 resources updated in 14.197169001 seconds#[0m
2016-01-28 09:28:52,826 [ TRACE ] localhost.localdomain
[2016-01-28T09:28:52+00:00] ERROR: execute[create_GPFS_cluster_default_profile]
(gpfs::gpfs_cluster_create line 20) had an error:
Mixlib::ShellOut::ShellCommandFailed: Expected process to exit with [0], but received '1'

```

## Workaround

In this case, manually verify that prompt-less SSH is set up properly. Once prompt-less SSH is set up properly between all nodes, installation can be initiated again.

The warning that indicates that the host is added to the list of known hosts helped in determining that the prompt-less SSH setup is improper. If prompt-less SSH were completely set up before this installation, the host would already have existed within the known hosts file.

**Note:** IBM Spectrum Scale requires all admin nodes to have prompt-less SSH to and from all other nodes of the cluster.

## Example 2 - Deployment failed and the bottom of the log file contains the following Chef output, which is not indicative of the error

**Note:** The following log text has been adjusted to fit in the PDF margin.

```

2016-01-15 15:31:14,912 [ TRACE ] Stopping chef zero
2016-01-15 15:31:14,913 [ ERROR ] The following error was encountered:
Traceback (most recent call last):
  File "/usr/lpp/mmfs/4.2.0.0/installer/espyslib/reporting.py", line 222, in log_to_file
    yield handler
  File "/usr/lpp/mmfs/4.2.0.0/installer/espyslib/install.py", line 167, in _install
    setup.install(config)
  File "/usr/lpp/mmfs/4.2.0.0/installer/espyslib/setup/ces.py", line 325, in install
    self.deploy(config.protocol_nodes, options_fn)
  File "/usr/lpp/mmfs/4.2.0.0/installer/espyslib/deploy.py", line 133, in deploy_nodes
    raise DeployError()
DeployError: Installation failed on one or more nodes. Check the log for more details.
2016-01-15 15:31:14,957 [ INFO ] Detailed error log:
/usr/lpp/mmfs/4.2.0.0/installer/logs/DEPLOY-15-01-2016_15:29:59.log

```

1. To find more details, go to the end of the log file and search upwards for the text FATAL.

In this example, the first search hit is the last instance of the word FATAL in the log file that is being searched. The output typically shows what was printed to the screen and gives a general indication of where the failure occurred. It is also helpful to search for the terms 'fail' and 'error'. In this case, the failure occurred while installing object packages:



**Note:** The following log text has been adjusted to fit in the PDF margin.

```
2016-01-15 15:31:09,762 [ FATAL ]
objnode4 failure whilst: Installing Object packages (SS50)
2016-01-15 15:31:09,770 [ WARN ] SUGGESTED ACTION(S):
2016-01-15 15:31:09,770 [ WARN ]
Check Object dependencies are available via your package manager or are already met prior to installation.
2016-01-15 15:31:09,770 [ FATAL ]
objnode3 failure whilst: Installing Object packages (SS50)
```

2. Search upwards further for the text FATAL to find its first occurrence in the log file.

In this example, the text FATAL is found 8 times, wherein the following is its first occurrence in the log file:

**Note:** The following log text has been adjusted to fit in the PDF margin.

```
2016-01-15 15:31:09,447 [ TRACE ] objnode4 [2016-01-15T15:31:09+05:30] ERROR: Running exception handlers
2016-01-15 15:31:09,447 [ TRACE ] objnode4 Running handlers complete
2016-01-15 15:31:09,447 [ TRACE ] objnode4 [0m[2016-01-15T15:31:09+05:30] ERROR: Exception handlers complete
2016-01-15 15:31:09,448 [ TRACE ] objnode4 [2016-01-15T15:31:09+05:30]
FATAL: Stacktrace dumped to /var/chef/cache/chef-stacktrace.out
2016-01-15 15:31:09,448 [ TRACE ] objnode4 Chef Client failed. 32 resources updated in 46.185382251 seconds[0m
2016-01-15 15:31:09,474 [ TRACE ] objnode4 [2016-01-15T15:31:09+05:30]
ERROR: yum_package[spectrum-scale-object] (swift_on_gpfs::swift_node_install line 14) had an error:
Chef::Exceptions::Exec: yum -d0 -e0 -y install spectrum-scale-object-4.2.0-0 returned 1:
```

This log snippet mentions the exact Chef recipe (`swift_on_gpfs::swift_node_install`) that failed during deployment.

3. To find more information, visually search upwards within the log file.

Root cause output is typically close to the first occurrence of the text FATAL. Following is a snippet of the log text above the first occurrence of FATAL. It shows the start of the **Installing Object packages** portion of the deployment and then shows where the first error occurred: a stanza encapsulated in "=====" symbols. Immediately following this is the command executed by the installation toolkit for deployment:

```
yum -do -e0 -y install spectrum-scale-object-4.2.0.0
```

Following that is a STDERR: Error showing that a specific package, `libcap-ng` is already installed on this node with version `0.7.5-4`, yet this specific code level requires version `0.7.3-5` of `libcap-ng`. Because `libcap-ng` version `0.7.3-5` is a dependency for `spectrum-scale-object-4.2.0-0`, the deployment has failed.

**Note:** The following log text has been adjusted to fit in the PDF margin.

```
2016-01-15 15:30:51,858 [ INFO ] [objnode3 15-01-2016 15:30:51]
IBM SPECTRUM SCALE: Installing Object packages (SS50)
2016-01-15 15:30:51,858 [ TRACE ] objnode3 * log
[IBM SPECTRUM SCALE: Installing Object packages (SS50).] action write
2016-01-15 15:30:51,859 [ TRACE ] objnode3
2016-01-15 15:31:09,441 [ TRACE ] objnode4 [0m
* yum_package[spectrum-scale-object] action install
2016-01-15 15:31:09,441 [ TRACE ] objnode4 [0m
2016-01-15 15:31:09,441 [ TRACE ] objnode4
=====0m
2016-01-15 15:31:09,442 [ TRACE ] objnode4 [31mError executing action `install`
on resource 'yum_package[spectrum-scale-object]'[0m
=====0m
2016-01-15 15:31:09,442 [ TRACE ] objnode4
2016-01-15 15:31:09,442 [ TRACE ] objnode4 [0m Chef::Exceptions::Exec[0m
2016-01-15 15:31:09,442 [ TRACE ] objnode4 -----[0m
2016-01-15 15:31:09,442 [ TRACE ] objnode4
yum -d0 -e0 -y install spectrum-scale-object-4.2.0-0 returned 1:
2016-01-15 15:31:09,442 [ TRACE ] objnode4 [0m
STDOUT: You could try using --skip-broken to work around the problem
2016-01-15 15:31:09,443 [ TRACE ] objnode4 [0m
You could try running: rpm -Va --nofiles --nodigest
2016-01-15 15:31:09,443 [ TRACE ] objnode4 [0m
2016-01-15 15:31:09,443 [ TRACE ] objnode4 [0m
STDERR: Error: Package: libcap-ng-python-0.7.3-5.e17.x86_64 (ces_object)
2016-01-15 15:31:09,443 [ TRACE ] objnode4 [0m
Requires: libcap-ng = 0.7.3-5.e17
2016-01-15 15:31:09,443 [ TRACE ] objnode4 [0m
Installed: libcap-ng-0.7.5-4.e17.x86_64 (@RHEL7.1)
2016-01-15 15:31:09,443 [ TRACE ] objnode4 [0m
libcap-ng = 0.7.5-4.e17
2016-01-15 15:31:09,443 [ TRACE ] objnode4 [0m
2016-01-15 15:31:09,443 [ TRACE ] objnode4 [0m
Resource Declaration:[0m
2016-01-15 15:31:09,444 [ TRACE ] objnode4
-----[0m
```

```

2016-01-15 15:31:09,444 [ TRACE ] objnode4
# In /var/chef/cache/cookbooks/swift_on_gpfs/recipes/swift_node_install.rb
2016-01-15 15:31:09,444 [ TRACE ] objnode4 [0m
2016-01-15 15:31:09,444 [ TRACE ] objnode4 [0m
14: package pkg do
2016-01-15 15:31:09,444 [ TRACE ] objnode4 [0m
15:   retries 3
2016-01-15 15:31:09,444 [ TRACE ] objnode4 [0m
16:   retry_delay 3
2016-01-15 15:31:09,444 [ TRACE ] objnode4 [0m
17: end
2016-01-15 15:31:09,444 [ TRACE ] objnode4 [0m
18: end
2016-01-15 15:31:09,444 [ TRACE ] objnode4 [0m
2016-01-15 15:31:09,445 [ TRACE ] objnode4 [0m
Compiled Resource:[0m
2016-01-15 15:31:09,445 [ TRACE ] objnode4 -----[0m
2016-01-15 15:31:09,445 [ TRACE ] objnode4
# Declared in /var/chef/cache/cookbooks/swift_on_gpfs/recipes/swift_node_install.rb:14:in `block in from_file'
2016-01-15 15:31:09,445 [ TRACE ] objnode4 [0m
2016-01-15 15:31:09,445 [ TRACE ] objnode4 [0m
yum_package("spectrum-scale-object") do
2016-01-15 15:31:09,445 [ TRACE ] objnode4 [0m
action :install
2016-01-15 15:31:09,445 [ TRACE ] objnode4 [0m
retries 3
2016-01-15 15:31:09,445 [ TRACE ] objnode4 [0m
retry_delay 3
2016-01-15 15:31:09,446 [ TRACE ] objnode4 [0m
default_guard_interpreter :default
2016-01-15 15:31:09,446 [ TRACE ] objnode4 [0m
package_name "spectrum-scale-object"
2016-01-15 15:31:09,446 [ TRACE ] objnode4 [0m
version "4.2.0-0"
2016-01-15 15:31:09,446 [ TRACE ] objnode4 [0m
timeout 900
2016-01-15 15:31:09,446 [ TRACE ] objnode4 [0m
flush_cache {:before=>false, :after=>false}
2016-01-15 15:31:09,446 [ TRACE ] objnode4 [0m
declared_type :package
2016-01-15 15:31:09,446 [ TRACE ] objnode4 [0m
cookbook_name "swift_on_gpfs"
2016-01-15 15:31:09,446 [ TRACE ] objnode4 [0m
recipe_name "swift_node_install"
2016-01-15 15:31:09,447 [ TRACE ] objnode4 [0m end
2016-01-15 15:31:09,447 [ TRACE ] objnode4 [0m
2016-01-15 15:31:09,447 [ TRACE ] objnode4 [0m[0m
2016-01-15 15:31:09,447 [ TRACE ] objnode4 Running handlers:[0m
2016-01-15 15:31:09,447 [ TRACE ] objnode4
[2016-01-15T15:31:09+05:30] ERROR: Running exception handlers
2016-01-15 15:31:09,447 [ TRACE ] objnode4
Running handlers complete
2016-01-15 15:31:09,447 [ TRACE ] objnode4
[0m[2016-01-15T15:31:09+05:30] ERROR: Exception handlers complete
2016-01-15 15:31:09,448 [ TRACE ] objnode4
[2016-01-15T15:31:09+05:30] FATAL: Stacktrace dumped to
/var/chef/cache/chef-stacktrace.out
2016-01-15 15:31:09,448 [ TRACE ] objnode4 Chef Client failed.
32 resources updated in 46.185382251 seconds[0m
2016-01-15 15:31:09,474 [ TRACE ] objnode4
[2016-01-15T15:31:09+05:30] ERROR: yum_package[spectrum-scale-object]
(swift_on_gpfs::swift_node_install line 14) had an error:
Chef::Exceptions::Exec: yum -d0 -e0 -y install spectrum-scale-object-4.2.0-0 returned 1:

```

## Workaround

Manually remove the newer version of `libcapi-ng` installed on the node(s) and install the required version indicated in the error log. Once this is done on all nodes, deployment can be initiated again.

## Problems due to missing prerequisites

Use this information to ensure that prerequisites are met before using the `spectrumscale` installation toolkit for installation, deployment, and upgrade.

- “Prompt-less SSH setup” on page 219
- “Yum repository setup” on page 219
- “Firewall configuration” on page 219
- “CES IP address allocation” on page 220
- “Addition of CES IPs to `/etc/hosts`” on page 221

## | **Prompt-less SSH setup**

- | 1. Verify that prompt-less SSH is set up using the following commands.

```
| ssh <host name of the first node>  
| ssh <host name of the second node>
```

| Repeat this on all nodes.

| Verify that the user can log into the node successfully without being prompted for any input and that there are no warnings.

```
| ssh <FQDN of the first node>  
| ssh <FQDN of the second node>
```

| Repeat this on all nodes.

| Verify that the user can log into the node successfully without being prompted for any input and that there are no warnings.

```
| ssh <IP address of the first node>  
| ssh <IP address of the second node>
```

| Repeat this on all nodes.

| Verify that the user can log into the node successfully without being prompted for any input and that there are no warnings.

- | 2. If needed, set up prompt-less SSH using the following commands.

| **Note:** This is one of the several possible ways of setting up prompt-less SSH.

```
| ssh-keygen
```

| Repeat this on all cluster nodes.

```
| ssh-copy-id <host name of the first node>  
| ssh-copy-id <host name of the second node>
```

| Repeat this on all nodes.

```
| ssh-copy-id <FQDN of the first node>  
| ssh-copy-id <FQDN of the second node>
```

| Repeat this on all nodes.

## | **Yum repository setup**

- | • Verify that the Yum repository is set up using the following command on all cluster nodes.

```
| yum repolist
```

| This command should run clean with no errors if the Yum repository is set up.

## | **Firewall configuration**

| It is recommended that firewalls are in place to secure all nodes. For more information, see *Securing the IBM Spectrum Scale system using firewall* in *IBM Spectrum Scale: Administration Guide*.

- | • If you need to open specific ports, use the following steps.

- | 1. Check the firewall status.

```
| systemctl status firewalld
```

- | 2. Open ports required by the **spectrumscale** installation toolkit.

```
| firewall-cmd --permanent --add-port 8889/tcp  
| firewall-cmd --add-port 8889/tcp  
| firewall-cmd --permanent --add-port 10080/tcp  
| firewall-cmd --add-port 10080/tcp
```

## | CES IP address allocation

| As part of the deployment process, the IBM Spectrum Scale checks routing on the cluster and applies CES IPs as aliases on each protocol node. Furthermore, as service actions or failovers, nodes dynamically lose the alias IPs as they go down and other nodes gain additional aliases to hold all of the IPs passed to them from the down nodes.

### | Example - Before deployment

| The only address here is 192.168.251.161, which is the ssh address for the node. It is held by the eth0 adapter.

```
| # ifconfig -a
| eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>mtu 1500
|     inet 192.168.251.161 netmask 255.255.254.0 broadcast 192.168.251.255
|     inet6 2002:90b:e006:84:250:56ff:fea5:1d86 prefixlen 64 scopeid 0x0<global>
|     inet6 fe80::250:56ff:fea5:1d86 prefixlen 64 scopeid 0x20<link>
|     ether 00:50:56:a5:1d:86 txqueuelen 1000 (Ethernet)
|     RX packets 1978638 bytes 157199595 (149.9 MiB)
|     RX errors 0 dropped 2291 overruns 0 frame 0
|     TX packets 30884 bytes 3918216 (3.7 MiB)
|     TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0
|
| # ip addr
| 2: eth0:<BROADCAST,MULTICAST,UP,LOWER_UP>mtu 1500 qdisc mq state UP qlen 1000
|    link/ether 00:50:56:a5:1d:86 brd ff:ff:ff:ff:ff:ff
|    inet 192.168.251.161/23 brd 192.168.251.255 scope global eth0
|        valid_lft forever preferred_lft forever
|    inet6 2002:90b:e006:84:250:56ff:fea5:1d86/64 scope global dynamic
|        valid_lft 2591875sec preferred_lft 604675sec
|    inet6 fe80::250:56ff:fea5:1d86/64 scope link
|        valid_lft forever preferred_lft forever
```

### | Example - After deployment

| Now that the CES IP addresses exist, you can see that aliases called eth0:0 and eth0:1 have been created and the CES IP addresses specific to this node have been tagged to it. This allows the ssh IP of the node to exist at the same time as the CES IP address on the same adapter, if necessary. In this example, 192.168.251.161 is the initial ssh IP. The CES IP 192.168.251.165 is aliased onto eth0:0 and the CES IP 192.168.251.166 is aliased onto eth0:1. This occurs on all protocol nodes that are assigned a CES IP address. NSD server nodes or any client nodes that do not have protocols installed on them do not get a CES IP.

| Furthermore, as service actions or failovers, nodes dynamically lose the alias IPs as they go down and other nodes gain additional aliases such as eth0:1 and eth0:2 to hold all of the IPs passed to them from the down nodes.

```
| # ifconfig -a
| eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
|     inet 192.168.251.161 netmask 255.255.254.0 broadcast 192.168.251.255
|     inet6 2002:90b:e006:84:250:56ff:fea5:1d86 prefixlen 64 scopeid 0x0<global>
|     inet6 fe80::250:56ff:fea5:1d86 prefixlen 64 scopeid 0x20<link>
|     ether 00:50:56:a5:1d:86 txqueuelen 1000 (Ethernet)
|     RX packets 2909840 bytes 1022774886 (975.3 MiB)
|     RX errors 0 dropped 2349 overruns 0 frame 0
|     TX packets 712595 bytes 12619844288 (11.7 GiB)
|     TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0
| eth0:0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>mtu 1500
|     inet 192.168.251.165 netmask 255.255.254.0 broadcast 192.168.251.255
|     ether 00:50:56:a5:1d:86 txqueuelen 1000 (Ethernet)
| eth0:1: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>mtu 1500
|     inet 192.168.251.166 netmask 255.255.254.0 broadcast 192.168.251.255
|     ether 00:50:56:a5:1d:86 txqueuelen 1000 (Ethernet)
```

```

| # ip addr
| 2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP>mtu 1500 qdisc mq state UP qlen 1000
|   link/ether 00:50:56:a5:1d:86 brd ff:ff:ff:ff:ff:ff
|   inet 192.168.251.161/23 brd 9.11.85.255 scope global eth0
|     valid_lft forever preferred_lft forever
|   inet 192.168.251.165/23 brd 9.11.85.255 scope global secondary eth0:0
|     valid_lft forever preferred_lft forever
|   inet 192.168.251.166/23 brd 9.11.85.255 scope global secondary eth0:1
|     valid_lft forever preferred_lft forever
|   inet6 2002:90b:e006:84:250:56ff:fea5:1d86/64 scope global dynamic
|     valid_lft 2591838sec preferred_lft 604638sec
|   inet6 fe80::250:56ff:fea5:1d86/64 scope link
|     valid_lft forever preferred_lft forever

```

## Addition of CES IPs to /etc/hosts

Although it is highly recommended that all CES IPs are maintained in a central DNS and that they are accessible using both forward and reverse DNS lookup, there are times when this might not be possible. IBM Spectrum Scale always verify that forward or reverse DNS lookup is possible. To satisfy this check without a central DNS server containing the CES IPs, you must add the CES IPs to /etc/hosts and create a host name for them within /etc/hosts. The following example shows how a cluster might have multiple networks, nodes, and IPs defined.

For example:

```

| # cat /etc/hosts
| 127.0.0.1 localhost localhost.localdomain localhost4 localhost4.localdomain4
| ::1 localhost localhost.localdomain localhost6 localhost6.localdomain6
|
| # These are external addresses for GPFS
| # Use these for ssh in. You can also use these to form your GPFS cluster if you choose
| 198.51.100.2 ss-deploy-cluster3-1 ss-deploy-cluster3-1.example.com
| 198.51.100.4 ss-deploy-cluster3-2 ss-deploy-cluster3-2.example.com
| 198.51.100.6 ss-deploy-cluster3-3 ss-deploy-cluster3-3.example.com
| 198.51.100.9 ss-deploy-cluster3-4 ss-deploy-cluster3-4.example.com
|
| # These are addresses for the base adapter used to alias CES-IPs to.
| # Do not use these as CES-IPs.
| # You could use these for a gpfs cluster if you choose
| # Or you could leave these unused as placeholders
| 203.0.113.7 ss-deploy-cluster3-1_ces ss-deploy-cluster3-1_ces.example.com
| 203.0.113.10 ss-deploy-cluster3-2_ces ss-deploy-cluster3-2_ces.example.com
| 203.0.113.12 ss-deploy-cluster3-3_ces ss-deploy-cluster3-3_ces.example.com
| 203.0.113.14 ss-deploy-cluster3-4_ces ss-deploy-cluster3-4_ces.example.com
|
| # These are addresses to use for CES-IPs
| 203.0.113.17 ss-deploy-cluster3-ces ss-deploy-cluster3-ces.example.com
| 203.0.113.20 ss-deploy-cluster3-ces ss-deploy-cluster3-ces.example.com
| 203.0.113.21 ss-deploy-cluster3-ces ss-deploy-cluster3-ces.example.com
| 203.0.113.23 ss-deploy-cluster3-ces ss-deploy-cluster3-ces.example.com

```

In this example, the first two sets of addresses have unique host names and the third set of addresses that are associated with CES IPs are not unique. Alternatively, you could give each CES IP a unique host name but this is an arbitrary decision because only the node itself can see its own /etc/hosts file. Therefore, these host names are not visible to external clients/nodes unless they too contain a mirror copy of the /etc/hosts file. The reason for containing the CES IPs within the /etc/hosts file is solely to satisfy the IBM Spectrum Scale CES network verification checks. Without this, in cases with no DNS server, CES IPs cannot be added to a cluster.

## Problems due to mixed operating system levels in the cluster

Use the following guidelines to avoid problems due to mixed operating system levels in an IBM Spectrum Scale cluster.

| Many existing IBM Spectrum Scale clusters have nodes that are running on operating systems such as  
| Red Hat Enterprise Linux (RHEL) 6.x, SUSE Linux Enterprise Edition, Microsoft Windows, Ubuntu, IBM  
| AIX, Debian, PowerKVM, and other operating systems. The **spectrumscale** installation toolkit is  
| supported only on Red Hat Enterprise Linux 7.x for the IBM Spectrum Scale 4.1.1.x, 4.2.0.x, and 4.2.1.x  
| releases. Therefore, be aware of all operating systems in a cluster in which the installation toolkit is to be  
| used. In a mixed OS cluster, the installation toolkit has a limited functionality but it can be useful if it is  
| configured to operate only on nodes that have Red Hat Enterprise Linux 7.x installed.

| For latest information about supported operating systems, see IBM Spectrum Scale FAQ in IBM  
| Knowledge Center([www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html](http://www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html)).

| Verify that the installation toolkit is configured to operate only on supported nodes by using the  
| following command:

| `./spectrumscale node list`

| If any of the listed nodes are of an unsupported OS type, then they need to be removed by using the  
| following command:

| `./spectrumscale node delete node`

| If the node to be removed is an NSD node, then you might have to manually create NSDs and file  
| systems before using the installation toolkit.

| The installation toolkit does not need to be made aware of preexisting file systems and NSDs that are  
| present on unsupported node types. Ensure that the file systems are mounted before running the  
| installation toolkit and that they point at the mount points or directory structures.

## | **Upgrades in a mixed OS cluster**

| Upgrades in a mixed OS cluster need to be performed carefully due to a mix of manual and automated  
| steps. In this case, the installation toolkit can be made aware of a list of nodes that are running on  
| supported OS that are to be upgraded. It can then upgrade these nodes. However, the remaining nodes  
| need to be upgraded manually.

## | **Problems due to lack of clarity on using the spectrumscale installation toolkit for functions or configurations not fully supported**

| Use this information to determine node types, setups, and functions supported with the **spectrumscale**  
| installation toolkit, and to understand how to use the toolkit if a setup is not fully supported.

- | • “Support for mixed mode of install, deploy, or upgrade”
- | • “Support for advanced RPMs” on page 224
- | • “Support for DMAPi enabled nodes” on page 224
- | • “Support for ESS cluster” on page 225

## | **Support for mixed mode of install, deploy, or upgrade**

| **I want to use the installation toolkit but I already have an existing cluster. Can the installation toolkit  
| auto-detect my cluster or do I have to manually configure the toolkit?**

| The installation toolkit is stateless and it does not import an existing cluster configuration into its  
| cluster definition file. As a workaround to this scenario, use the steps in these topics of *IBM  
| Spectrum Scale: Concepts, Planning, and Installation Guide*.

- | • *Deploying protocols on an existing cluster*
- | • *Deploying protocols authentication on an existing cluster*
- | • *Adding nodes, NSDs, or file systems to an existing installation*
- | • *Enabling another protocol on an existing cluster that has protocols enabled*

If NSDs and file systems already exist, you do not need to provide that information to the installation toolkit.

**What are valid starting scenarios for which the installation toolkit can be used for an installation or a deployment or an upgrade?**

Scenario	Installation toolkit support
No cluster exists and no GPFS RPMs exist on any nodes.	The installation toolkit can be used to install GPFS and create a cluster.
No cluster exists and GPFS RPMs are already installed on nodes.	The installation toolkit can be used to install GPFS and create a cluster.
No cluster exists	The installation toolkit can be used to configure NTP during GPFS installation and cluster configuration.
No cluster exists	The installation GUI can be used to create a cluster.
A cluster exists	The installation toolkit can be used to add NSDs.
A cluster exists	The installation toolkit can be used to add nodes (manager, quorum, admin, nsd, protocol, gui).
A cluster exists and NSDs exist	The installation toolkit can be used to add file systems.
A cluster exists and some NSDs exist	The installation toolkit can be used to add more NSDs.
A cluster exists and some protocols are enabled	The installation toolkit can be used to enable more protocols.
A cluster exists and performance monitoring is enabled	The installation toolkit can be used to reconfigure performance monitoring.
An ESS cluster exists and protocol nodes have been added	The installation toolkit can be used to add protocols to protocol nodes.
SLES, RHEL 6, Windows, Debian, Ubuntu, and AIX nodes exist along with RHEL 7.x nodes	The installation toolkit can be used only on RHEL 7.x nodes.
A cluster is at mixed levels of 4.2.0.x	The installation toolkit can be used to upgrade all nodes or a subset of nodes to a common code level.

**What are invalid starting scenarios for the installation toolkit?**

- NSDs were not cleaned up or deleted prior to a cluster deletion.
- Unsupported node types were added to the installation toolkit.
- File systems or NSDs are served by unsupported node types.  
The installation toolkit cannot add or change these. It can only use file system paths for protocol configuration.
- An ESS cluster exists and protocol nodes have not yet been added to the cluster.  
Protocol nodes must first be added to the ESS cluster before the installation toolkit can install the protocols.

**Does the installation toolkit need to have my entire cluster information?**

No, but this depends on the use case. Here are some examples in which the installation toolkit does not need to be made aware of the configuration information of an existing cluster:

- **Deploying protocols on protocol nodes:** The installation toolkit needs only the protocol nodes information and that they are configured to point to cesSharedRoot.
- **Upgrading protocol nodes:** The installation toolkit can upgrade a portion of the cluster such as all protocol nodes. In this case, it does not need to be made aware of other NSD or client/server nodes within the cluster.
- **Adding protocols to an ESS cluster:** The installation toolkit does not need to be made aware of the EMS or I/O nodes. The installation toolkit needs only the protocol nodes information and that they are configured to point to cesSharedRoot.

- **Adding protocols to a cluster with AIX, SLES, Debian, Ubuntu, RHEL6, and Windows nodes:** The installation toolkit does not need to be made aware of any nodes except for the RHEL 7.x protocol nodes. The installation toolkit needs only the protocol nodes information and that they are configured to point to cesSharedRoot.

### Can the installation toolkit act on some protocol nodes but not all?

Protocol nodes must always be treated as a group of nodes. Therefore, do not use the installation toolkit to run install, deploy, or upgrade commands on a subset of protocol nodes.

## Support for advanced RPMs

The **spectrumscale** installation toolkit currently does not handle the AFM-based Async DR (AFM DR) and Encryption RPMs.

If performing an initial installation using the installation toolkit, allow the toolkit to fully complete the installation. Once the installation is complete, manually install the desired advanced RPMs as follows:

1. Go to the directory where the installation image is extracted:
 

```
cd /usr/lpp/mmfs/4.2.1.0/gpfs_rpms
```
2. Install the advanced RPM (that contains the AFM DR function):
 

```
yum localinstall gpfs.adv-*
```
3. Install the crypto RPM (that contains the Encryption function):
 

```
yum localinstall gpfs.crypto-*
```
4. Repeat steps 1 to 3 on all nodes.
5. Verify that the license has changed to Advanced by issuing the **mmfslicense** command.

## Support for DMAPi enabled nodes

On nodes with DMAPi enabled, the **spectrumscale** installation toolkit does not provide much help to users in case of an error including whether a DMAPi related function is supported or unsupported.

Use the following steps to verify whether DMAPi is enabled on your nodes and to use the **spectrumscale** installation toolkit on DMAPi enabled nodes.

1. Verify that DMAPi is enabled on a file system using the following command:

```
# mmfsfs all -z
File system attributes for /dev/fs1:
=====
flag                value                description
-----
-z                  yes                  Is DMAPi enabled?
```

2. Shut down all functions that are using DMAPi and unmount DMAPi using the following steps:
  - a. Shut down all functions that are using DMAPi. This includes HSM policies and IBM Spectrum Archive™.
  - b. Unmount the DMAPi file system from all nodes using the following command:

```
# mmumount fs1 -a
```

**Note:** If the DMAPi file system is also the CES shared root file system, then you must first shut down GPFS on all protocol nodes before unmounting the file system.

- 1) Check if the DMAPi file system is also the CES shared root file system, use the following command:

```
# mmfsconfig | grep cesSharedRoot
```

- 2) Compare the output of this command with that of Step 1 to determine if the CES shared root file system has DMAPi enabled.



- 3) Shut down GPFS on all protocol nodes using the following command:
  - # mmshutdown -N cesNodes
- c. Disable DMAPi using the following command:
  - # mmchfs fs1 -z no
- 3. If GPFS was shut down on the protocol nodes in one of the preceding steps, start GPFS on the protocol nodes using the following command:
  - # mmstartup -N cesNodes
- 4. Remount the file system on all nodes using the following command:
  - # mmmount fs1 -a
- 5. Proceed with using the **spectrumscale** installation toolkit as now it can be used on all file systems.
- 6. After the task being done using the **spectrumscale** installation toolkit is completed, enable DMAPi using the following steps:
  - a. Unmount the DMAPi file system from all nodes.
    - Note:** If the DMAPi file system is also the CES shared root file system, shut down GPFS on all protocol nodes before unmounting the file system.
  - b. Enable DMAPi using the following command:
    - # mmchfs fs1 -z yes
  - c. Start GPFS on all protocol nodes.
  - d. Remount the file system on all nodes.

## Support for ESS cluster

- For information on using the **spectrumscale** installation toolkit with a cluster containing ESS, see the following topics in *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*:
  - *Preparing a cluster that contains ESS for adding protocols*
  - *Deploying protocols on an existing cluster*

## Problems due to lack of clarity on supported upgrade functions with spectrumscale installation toolkit

- Use this information to understand the setups in which upgrade can be done using the **spectrumscale** installation toolkit.
  - “Scope of the upgrade process”
  - “Understanding implications of a failed upgrade” on page 226

## Scope of the upgrade process

- The upgrade process using the **spectrumscale** installation toolkit can be summarized as follows:
  - The upgrade process acts upon all nodes specified in the cluster definition file (typically using the **./spectrumscale node add** commands).
  - All installed/deployed components are upgraded.
  - Upgrades are sequential with multiple passes.

- The upgrade process using the **spectrumscale** installation toolkit comprises following passes:
  1. Pass 1 of all nodes upgrades GPFS sequentially.
  2. Pass 2 of all nodes upgrades Object sequentially.
  3. Pass 3 of all nodes upgrades NFS sequentially.
  4. Pass 4 of all nodes upgrades SMB sequentially.
  5. A post check is done to verify a healthy cluster state after the upgrade.

- | As an upgrade moves sequentially across nodes, functions such as SMB, NFS, Object, Performance Monitoring, AFM, etc. undergo failovers. This might cause outages on the nodes being upgraded.
- | Upgrading a subset of nodes is possible because the installation toolkit acts only on the nodes specified in the cluster definition file. If you want to upgrade a subset of cluster nodes, be aware of the node types and the functions being performed on these nodes. For example, all protocol nodes within a cluster must be upgraded by the installation toolkit in one batch.

### | **Understanding implications of a failed upgrade**

- | A failed upgrade might leave a cluster in a state of containing multiple code levels. It is important to analyze console output to determine which nodes or components were upgraded prior to the failure and which node or component was in the process of being upgraded when the failure occurred.
- | Once the problem has been isolated, a healthy cluster state must be achieved prior to continuing the upgrade. Use the `mmhealth` command in addition to the `mmces state show -a` command to verify that all services are up. It might be necessary to manually start services that were down when the upgrade failed. Starting the services manually helps achieve a state in which all components are healthy prior to continuing the upgrade.
- | For more information about verifying service status, see `mmhealth` command and `mmces state show` command in *IBM Spectrum Scale: Command and Programming Reference*.

---

## **Post installation and configuration problems**

This topic describes the issues that you might encounter after installing or configuring IBM Spectrum Scale.

The *IBM Spectrum Scale: Concepts, Planning, and Installation Guide* provides the step-by-step procedure for installing and migrating IBM Spectrum Scale, however, some problems might occur after installation and configuration if the procedures were not properly followed.

Some of those problems might include:

- Not being able to start GPFS after installation of the latest version. Did you reboot your IBM Spectrum Scale nodes before and after the installation/upgrade of IBM Spectrum Scale? If you did, see “GPFS daemon will not come up” on page 232. If not, reboot. For more information, see the *Initialization of the GPFS daemon* topic in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.
- E! • Not being able to access a file system. See “File system fails to mount” on page 245.
- New GPFS functions do not operate. See “GPFS commands are unsuccessful” on page 237.

---

### E! **Cluster is crashed after reinstallation**

This topic describes the steps that you need to perform when a cluster crashes after IBM Spectrum Scale reinstallation.

After reinstalling IBM Spectrum Scale code, check whether the `/var/mmfs/gen/mmsdrfs` file was lost. If it was lost, and an up-to-date version of the file is present on the primary GPFS cluster configuration server, restore the file by issuing this command from the node on which it is missing:

```
mmsdrrestore -p primaryServer
```

where `primaryServer` is the name of the primary GPFS cluster configuration server.

If the `/var/mmfs/gen/mmsdrfs` file is not present on the primary GPFS cluster configuration server, but it is present on some other node in the cluster, restore the file by issuing these commands:

```
mmsdrrestore -p remoteNode -F remoteFile
mmchcluster -p LATEST
```

where *remoteNode* is the node that has an up-to-date version of the `/var/mmfs/gen/mmsdrfs` file, and *remoteFile* is the full path name of that file on that node.

One way to ensure that the latest version of the `/var/mmfs/gen/mmsdrfs` file is always available is to use the `mmsdrbackup` user exit.

If you have made modifications to any of the users exist in `/var/mmfs/etc`, you will have to restore them before starting GPFS.

For additional information, see “Recovery from loss of GPFS cluster configuration data file” on page 230.

---

## Node cannot be added to the GPFS cluster

There is an indication leading you to the conclusion that a node cannot be added to a cluster and steps to follow to correct the problem.

That indication is:

- You issue the `mmcrcluster` or `mmaddnode` command and receive the message:

**6027-1598**

Node *nodeName* was not added to the cluster. The node appears to already belong to a GPFS cluster.

Steps to follow if a node cannot be added to a cluster:

1. Run the `mmclscluster` command to verify that the node is not in the cluster.
2. If the node is not in the cluster, issue this command on the node that could not be added:  
`mmdelnode -f`
3. Reissue the `mmaddnode` command.

---

## Problems with the `/etc/hosts` file

This topic describes the issues relating to the `/etc/hosts` file that you might come across while installing or configuring IBM Spectrum Scale.

The `/etc/hosts` file must have a unique node name for each node interface to be used by GPFS. Violation of this requirement results in the message:

**6027-1941**

Cannot handle multiple interfaces for host *hostName*.

If you receive this message, correct the `/etc/hosts` file so that each node interface to be used by GPFS appears only once in the file.

---

## Linux configuration considerations

This topic describes the Linux configuration that you need to consider while installing or configuring IBM Spectrum Scale on your cluster.

**Note:** This information applies only to Linux nodes.

Depending on your system configuration, you may need to consider:

1. **Why can only one host successfully attach to the Fibre Channel loop and see the Fibre Channel disks?**

Your host bus adapter may be configured with an enabled *Hard Loop ID* that conflicts with other host bus adapters on the same Fibre Channel loop.

To see if that is the case, reboot your machine and enter the adapter bios with <Alt-Q> when the Fibre Channel adapter bios prompt appears. Under the Configuration Settings menu, select Host Adapter Settings and either ensure that the Adapter Hard Loop ID option is disabled or assign a unique Hard Loop ID per machine on the Fibre Channel loop.

## 2. Could the GPFS daemon be terminated due to a memory shortage?

The Linux virtual memory manager (VMM) exhibits undesirable behavior for low memory situations on nodes, where the processes with the largest memory usage are killed by the kernel (using OOM killer), yet no mechanism is available for prioritizing important processes that should not be initial candidates for the OOM killer. The GPFS **mmfsd** daemon uses a large amount of pinned memory in the pagepool for caching data and metadata, and so the **mmfsd** process is a likely candidate for termination if memory must be freed up.

## 3. What are the performance tuning suggestions?

For an up-to-date list of tuning suggestions, see the IBM Spectrum Scale FAQ in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html](http://www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html)).

For Linux on z Systems, see also the Device Drivers, Features, and Commands ([www.ibm.com/support/knowledgecenter/api/content/linuxonibm/liaaf/lnz\\_r\\_dd.html](http://www.ibm.com/support/knowledgecenter/api/content/linuxonibm/liaaf/lnz_r_dd.html)) topic in the Linux on z Systems library overview.

---

## Problems with running commands on other nodes

This topic describes the problems that you might encounter relating to running remote commands during installing and configuring IBM Spectrum Scale.

Many of the GPFS administration commands perform operations on nodes other than the node on which the command was issued. This is achieved by utilizing a remote invocation shell and a remote file copy command. By default these items are **/usr/bin/ssh** and **/usr/bin/scp**. You also have the option of specifying your own remote shell and remote file copy commands to be used instead of the default **ssh** and **scp**. The remote shell and copy commands must adhere to the same syntax forms as **ssh** and **scp** but may implement an alternate authentication mechanism. For more information on the **mmcrcluster** and **mmchcluster** commands, see the *mmcrcluster command* and the *mmchcluster command* pages in the *IBM Spectrum Scale: Command and Programming Reference*. These are problems you may encounter with the use of remote commands.

## Authorization problems

This topic describes issues with running remote commands due to authorization problems in IBM Spectrum Scale.

The **ssh** and **scp** commands are used by GPFS administration commands to perform operations on other nodes. The **ssh** daemon (**sshd**) on the remote node must recognize the command being run and must obtain authorization to invoke it.

**Note:** Use the **ssh** and **scp** commands that are shipped with the OpenSSH package supported by GPFS. Refer to the IBM Spectrum Scale FAQ in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html](http://www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html)) for the latest OpenSSH information.

For more information, see “Problems due to missing prerequisites” on page 218.

For the **ssh** and **scp** commands issued by GPFS administration commands to succeed, each node in the cluster must have an **.rhosts** file in the home directory for the root user, with file permission set to 600. This **.rhosts** file must list each of the nodes and the root user. If such an **.rhosts** file does not exist on each node in the cluster, the **ssh** and **scp** commands issued by GPFS commands will fail with permission errors, causing the GPFS commands to fail in turn.

If you elected to use installation specific remote invocation shell and remote file copy commands, you must ensure:

1. Proper authorization is granted to all nodes in the GPFS cluster.
2. The nodes in the GPFS cluster can communicate without the use of a password, and without any extraneous messages.

## Connectivity problems

This topic describes the issues with running GPFS commands on remote nodes due to connectivity problems.

Another reason why **ssh** may fail is that connectivity to a needed node has been lost. Error messages from **mmdsh** may indicate that connectivity to such a node has been lost. Here is an example:

```
mmdelnode -N k145n04
Verifying GPFS is stopped on all affected nodes ...
mmdsh: 6027-1617 There are no available nodes on which to run the command.
mmdelnode: 6027-1271 Unexpected error from verifyDaemonInactive: mmcommon onall.
Return code: 1
```

If error messages indicate that connectivity to a node has been lost, use the **ping** command to verify whether the node can still be reached:

```
ping k145n04
PING k145n04: (119.114.68.69): 56 data bytes
<Ctrl- C>
----k145n04 PING Statistics----
3 packets transmitted, 0 packets received, 100% packet loss
```

If connectivity has been lost, restore it, then reissue the GPFS command.

## GPFS error messages for rsh problems

This topic describes the error messages that are displayed for rsh issues in IBM Spectrum Scale.

When **rsh** problems arise, the system may display information similar to these error messages:

```
6027-1615
      nodeName remote shell process had return code value.
```

```
6027-1617
      There are no available nodes on which to run the command.
```

---

## E! Cluster configuration data file issues

This topic describes the issues that you might encounter with respect to the cluster configuration data files while installing or configuring IBM Spectrum Scale.

## E! GPFS cluster configuration data file issues

This topic describes the issues relating to IBM Spectrum Scale cluster configuration data.

GPFS uses a file to serialize access of administration commands to the GPFS cluster configuration data files. This lock file is kept on the primary GPFS cluster configuration server in the **/var/mmfs/gen/mmLockDir** directory. If a system failure occurs before the cleanup of this lock file, the file will remain and subsequent administration commands may report that the GPFS cluster configuration data files are locked. Besides a serialization lock, certain GPFS commands may obtain an additional lock. This lock is designed to prevent GPFS from coming up, or file systems from being mounted, during critical sections of the command processing. If this happens you will see a message that shows the name of the blocking command, similar to message:

## 6027-1242

GPFS is waiting for *requiredCondition*.

To release the lock:

1. Determine the PID and the system that owns the lock by issuing:

```
mmcommon showLocks
```

The **mmcommon showLocks** command displays information about the lock server, lock name, lock holder, PID, and extended information. If a GPFS administration command is not responding, stopping the command will free the lock. If another process has this PID, another error occurred to the original GPFS command, causing it to die without freeing the lock, and this new process has the same PID. If this is the case, do not kill the process.

2. If any locks are held and you want to release them manually, from any node in the GPFS cluster issue the command:

```
mmcommon freeLocks <lockName>
```

## GPFS error messages for cluster configuration data file problems

This topic describes the error messages relating to the cluster configuration data file issues in IBM Spectrum Scale.

When GPFS commands are unable to retrieve or update the GPFS cluster configuration data files, the system may display information similar to these error messages:

### 6027-1628

Cannot determine basic environment information. Not enough nodes are available.

### 6027-1630

The GPFS cluster data on *nodeName* is back level.

### 6027-1631

The commit process failed.

### 6027-1632

The GPFS cluster configuration data on *nodeName* is different than the data on *nodeName*.

### 6027-1633

Failed to create a backup copy of the GPFS cluster data on *nodeName*.

## Recovery from loss of GPFS cluster configuration data file

This topic describes the procedure for recovering the cluster configuration data file in IBM Spectrum Scale.

A copy of the IBM Spectrum Scale cluster configuration data files is stored in the **/var/mmfs/gen/mmsdrfs** file on each node. For proper operation, this file must exist on each node in the IBM Spectrum Scale cluster. The latest level of this file is guaranteed to be on the primary, and secondary if specified, GPFS cluster configuration server nodes that were defined when the IBM Spectrum Scale cluster was first created with the **mmcrcluster** command.

If the **/var/mmfs/gen/mmsdrfs** file is removed by accident from any of the nodes, and an up-to-date version of the file is present on the primary IBM Spectrum Scale cluster configuration server, restore the file by issuing this command from the node on which it is missing:

```
mmsdrrestore -p primaryServer
```

where *primaryServer* is the name of the primary GPFS cluster configuration server.

If the `/var/mmfs/gen/mmsdrfs` file is not present on the primary GPFS cluster configuration server, but is present on some other node in the cluster, restore the file by issuing these commands:

```
mmsdrrestore -p remoteNode -F remoteFile
mmchcluster -p LATEST
```

where *remoteNode* is the node that has an up-to-date version of the `/var/mmfs/gen/mmsdrfs` file and *remoteFile* is the full path name of that file on that node.

One way to ensure that the latest version of the `/var/mmfs/gen/mmsdrfs` file is always available is to use the `mmsdrbackup` user exit.

## Automatic backup of the GPFS cluster data

This topic describes the procedure for automatically backing up the cluster data in IBM Spectrum Scale.

The IBM Spectrum Scale provides an exit, `mmsdrbackup`, that can be used to automatically back up the IBM Spectrum Scale configuration data every time it changes. To activate this facility, follow these steps:

1. Modify the IBM Spectrum Scale-provided version of `mmsdrbackup` as described in its prologue, to accomplish the backup of the `mmsdrfs` file however the user desires. This file is `/usr/lpp/mmfs/samples/mmsdrbackup.sample`.
2. Copy this modified `mmsdrbackup.sample` file to `/var/mmfs/etc/mmsdrbackup` on all of the nodes in the cluster. Make sure that the permission bits for `/var/mmfs/etc/mmsdrbackup` are set to permit execution by root.

The IBM Spectrum Scale system invokes the user-modified version of `mmsdrbackup` in `/var/mmfs/etc` every time a change is made to the `mmsdrfs` file. This will perform the backup of the `mmsdrfs` file according to the user's specifications. For more information on GPFS user exits, see the *GPFS user exits* topic in the *IBM Spectrum Scale: Command and Programming Reference*.

---

## E! GPFS application calls

### Error numbers specific to GPFS applications calls

This topic describes the error numbers specific to GPFS application calls.

When experiencing installation and configuration problems, GPFS may report these error numbers in the operating system error log facility, or return them to an application:

#### **ECONFIG = 215, Configuration invalid or inconsistent between different nodes.**

This error is returned when the levels of software on different nodes cannot coexist. For information about which levels may coexist, see the IBM Spectrum Scale FAQ in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html](http://www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html)).

#### **ENO\_QUOTA\_INST = 237, No Quota management enabled.**

To enable quotas for the file system issue the `mmchfs -Q yes` command. To disable quotas for the file system issue the `mmchfs -Q no` command.

#### **EOFFLINE = 208, Operation failed because a disk is offline**

This is most commonly returned when an open of a disk fails. Since GPFS will attempt to continue operation with failed disks, this will be returned when the disk is first needed to complete a command or application request. If this return code occurs, check your disk subsystem for stopped states and check to determine if the network path exists. In rare situations, this will be reported if disk definitions are incorrect.

#### **EALL\_UNAVAIL = 218, A replicated read or write failed because none of the replicas were available.**

Multiple disks in multiple failure groups are unavailable. Follow the procedures in Chapter 16, "Disk issues," on page 275 for unavailable disks.

#### 6027-341 [D]

Node *nodeName* is incompatible because its maximum compatible version (*number*) is less than the version of this node (*number*).

#### 6027-342 [E]

Node *nodeName* is incompatible because its minimum compatible version is greater than the version of this node (*number*).

#### 6027-343 [E]

Node *nodeName* is incompatible because its version (*number*) is less than the minimum compatible version of this node (*number*).

#### 6027-344 [E]

Node *nodeName* is incompatible because its version is greater than the maximum compatible version of this node (*number*).

---

## GPFS modules cannot be loaded on Linux

You must build the GPFS portability layer binaries based on the kernel configuration of your system. For more information, see *The GPFS open source portability layer* topic in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*. During **mmstartup** processing, GPFS loads the **mmfslinux** kernel module.

Some of the more common problems that you may encounter are:

1. If the portability layer is not built, you may see messages similar to:

```
Mon Mar 26 20:56:30 EDT 2012: runmmfs starting
Removing old /var/adm/ras/mmfs.log.* files:
Unloading modules from /lib/modules/2.6.32.12-0.6-ppc64/extra
runmmfs: The /lib/modules/2.6.32.12-0.6-ppc64/extra/mmfslinux.ko kernel extension does not exist.
runmmfs: Unable to verify kernel/module configuration.
Loading modules from /lib/modules/2.6.32.12-0.6-ppc64/extra
runmmfs: The /lib/modules/2.6.32.12-0.6-ppc64/extra/mmfslinux.ko kernel extension does not exist.
runmmfs: Unable to verify kernel/module configuration.
Mon Mar 26 20:56:30 EDT 2012 runmmfs: error in loading or unloading the mmfs kernel extension
Mon Mar 26 20:56:30 EDT 2012 runmmfs: stopping GPFS
```

2. The GPFS kernel modules, **mmfslinux** and **tracedev**, are built with a kernel version that differs from that of the currently running Linux kernel. This situation can occur if the modules are built on another node with a different kernel version and copied to this node, or if the node is rebooted using a kernel with a different version.
3. If the **mmfslinux** module is incompatible with your system, you may experience a kernel panic on GPFS startup. Ensure that the **site.mcr** has been configured properly from the **site.mcr.proto**, and GPFS has been built and installed properly.

For more information about the **mmfslinux** module, see the *Building the GPFS portability layer* topic in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.

---

## E! GPFS daemon issues

This topic describes the GPFS daemon issues that you might encounter while installing or configuring IBM Spectrum Scale.

### GPFS daemon will not come up

There are several indications that could lead you to the conclusion that the GPFS daemon (**mmfsd**) will not come up and there are some steps to follow to correct the problem.

Those indications include:

- The file system has been enabled to mount automatically, but the mount has not completed.
- You issue a GPFS command and receive the message:



## 6027-665

Failed to connect to file system daemon: Connection refused.

- The GPFS log does not contain the message:

## 6027-300 [N]

mmfsd ready

- The GPFS log file contains this error message: 'Error: daemon and kernel extension do not match.' This error indicates that the kernel extension currently loaded in memory and the daemon currently starting have mismatching versions. This situation may arise if a GPFS code update has been applied, and the node has not been rebooted prior to starting GPFS.

While GPFS scripts attempt to unload the old kernel extension during update and install operations, such attempts may fail if the operating system is still referencing GPFS code and data structures. To recover from this error, ensure that all GPFS file systems are successfully unmounted, and reboot the node. The **mmismount** command can be used to ensure that all file systems are unmounted.

## Steps to follow if the GPFS daemon does not come up

This topic describes the steps that you need to follow if the GPFS daemon does not come up after installation of IBM Spectrum Scale.

1. See “GPFS modules cannot be loaded on Linux” on page 232 if your node is running Linux, to verify that you have built the portability layer.
2. Verify that the GPFS daemon is active by issuing:

```
ps -e | grep mmfsd
```

The output of this command should list **mmfsd** as operational. For example:

```
12230 pts/8 00:00:00 mmfsd
```

If the output does not show this, the GPFS daemon needs to be started with the **mmstartup** command.

3. If you did not specify the **autoload** option on the **mmcrcluster** or the **mmchconfig** command, you need to manually start the daemon by issuing the **mmstartup** command.

If you specified the **autoload** option, someone may have issued the **mmshutdown** command. In this case, issue the **mmstartup** command. When using **autoload** for the first time, **mmstartup** must be run manually. The **autoload** takes effect on the next reboot.

4. Verify that the network upon which your GPFS cluster depends is up by issuing:

```
ping nodename
```

to each node in the cluster. A properly working network and node will correctly reply to the ping with no lost packets.

Query the network interface that GPFS is using with:

```
netstat -i
```

A properly working network will report no transmission errors.

5. Verify that the GPFS cluster configuration data is available by looking in the GPFS log. If you see the message:

## 6027-1592

Unable to retrieve GPFS cluster files from node *nodeName*.

Determine the problem with accessing node *nodeName* and correct it.

6. Verify that the GPFS environment is properly initialized by issuing these commands and ensuring that the output is as expected.
  - Issue the **mmiscluster** command to list the cluster configuration. This will also update the GPFS configuration data on the node. Correct any reported errors before continuing.

- List all file systems that were created in this cluster. For an AIX node, issue:

```
lsfs -v mmfs
```

For a Linux node, issue:

```
cat /etc/fstab | grep gpfs
```

If any of these commands produce unexpected results, this may be an indication of corrupted GPFS cluster configuration data file information. Follow the procedures in “Information to be collected before contacting the IBM Support Center” on page 367, and then contact the IBM Support Center.

7. GPFS requires a quorum of nodes to be active before any file system operations can be honored. This requirement guarantees that a valid single token management domain exists for each GPFS file system. Prior to the existence of a quorum, most requests are rejected with a message indicating that quorum does not exist.

To identify which nodes in the cluster have daemons **up** or **down**, issue:

```
mmgetstate -L -a
```

If insufficient nodes are active to achieve quorum, go to any nodes not listed as **active** and perform problem determination steps on these nodes. A quorum node indicates that it is part of a quorum by writing an `mmfsd ready` message to the GPFS log. Remember that your system may have quorum nodes and non-quorum nodes, and only quorum nodes are counted to achieve the quorum.

8. This step applies only to AIX nodes. Verify that GPFS kernel extension is not having problems with its shared segment by invoking:

```
cat /var/adm/ras/mmfs.log.latest
```

Messages such as:

**6027-319**

Could not create shared segment.

must be corrected by the following procedure:

- a. Issue the **mmshutdown** command.
  - b. Remove the shared segment in an AIX environment:
    - 1) Issue the **mmshutdown** command.
    - 2) Issue the **mmfsadm cleanup** command.
  - c. If you are still unable to resolve the problem, reboot the node.
9. If the previous GPFS daemon was brought down and you are trying to start a new daemon but are unable to, this is an indication that the original daemon did not completely go away. Go to that node and check the state of GPFS. Stopping and restarting GPFS or rebooting this node will often return GPFS to normal operation. If this fails, follow the procedures in “Additional information to collect for GPFS daemon crashes” on page 368, and then contact the IBM Support Center.

## Unable to start GPFS after the installation of a new release of GPFS

This topic describes the steps that you need to perform if you are unable to start GPFS after installing a new version of IBM Spectrum Scale.

If one or more nodes in the cluster will not start GPFS, these are the possible causes:

- If message:

**6027-2700 [E]**

A node join was rejected. This could be due to incompatible daemon versions, failure to find the node in the configuration database, or no configuration manager found.

is written to the GPFS log, incompatible versions of GPFS code exist on nodes within the same cluster.

- If messages stating that functions are not supported are written to the GPFS log, you may not have the correct kernel extensions loaded.
  1. Ensure that the latest GPFS install packages are loaded on your system.

2. If running on Linux, ensure that the latest kernel extensions have been installed and built. See the *Building the GPFS portability layer* topic in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.
  3. Reboot the GPFS node after an installation to ensure that the latest kernel extension is loaded.
- The daemon will not start because the configuration data was not migrated. See “Post installation and configuration problems” on page 226.

## GPFS error messages for shared segment and network problems

This topic describes the error messages relating to issues in shared segment and network in IBM Spectrum Scale.

For shared segment problems, follow the problem determination and repair actions specified with the following messages:

### 6027-319

Could not create shared segment.

### 6027-320

Could not map shared segment.

### 6027-321

Shared segment mapped at wrong address (is *value*, should be *value*).

### 6027-322

Could not map shared segment in kernel extension.

For network problems, follow the problem determination and repair actions specified with the following message:

### 6027-306 [E]

Could not initialize inter-node communication

## Error numbers specific to GPFS application calls when the daemon is unable to come up

This topic describes the application call error numbers when the daemon is unable to come up.

When the daemon is unable to come up, GPFS may report these error numbers in the operating system error log, or return them to an application:

### ECONFIG = 215, Configuration invalid or inconsistent between different nodes.

This error is returned when the levels of software on different nodes cannot coexist. For information about which levels may coexist, see the IBM Spectrum Scale FAQ in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html](http://www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html)).

### 6027-341 [D]

Node *nodeName* is incompatible because its maximum compatible version (*number*) is less than the version of this node (*number*).

### 6027-342 [E]

Node *nodeName* is incompatible because its minimum compatible version is greater than the version of this node (*number*).

### 6027-343 [E]

Node *nodeName* is incompatible because its version (*number*) is less than the minimum compatible version of this node (*number*).

### 6027-344 [E]

Node *nodeName* is incompatible because its version is greater than the maximum compatible version of this node (*number*).

## GPFS daemon went down

There are a number of conditions that can cause the GPFS daemon to exit.

These are all conditions where the GPFS internal checking has determined that continued operation would be dangerous to the consistency of your data. Some of these conditions are errors within GPFS processing but most represent a failure of the surrounding environment.

In most cases, the daemon will exit and restart after recovery. If it is not safe to simply force the unmounted file systems to recover, the GPFS daemon will exit.

Indications leading you to the conclusion that the daemon went down:

- Applications running at the time of the failure will see either ENODEV or ESTALE errors. The ENODEV errors are generated by the operating system until the daemon has restarted. The ESTALE error is generated by GPFS as soon as it restarts.

When quorum is lost, applications with open files receive an ESTALE error return code until the files are closed and reopened. New file open operations will fail until quorum is restored and the file system is remounted. Applications accessing these files prior to GPFS return may receive a ENODEV return code from the operating system.

- The GPFS log contains the message:

**6027-650 [X]**

The mmfs daemon is shutting down abnormally.

Most GPFS daemon down error messages are in the **mmfs.log.previous** log for the instance that failed. If the daemon restarted, it generates a new **mmfs.log.latest**. Begin problem determination for these errors by examining the operating system error log.

If an existing quorum is lost, GPFS stops all processing within the cluster to protect the integrity of your data. GPFS will attempt to rebuild a quorum of nodes and will remount the file system if automatic mounts are specified.

- Open requests are rejected with no such file or no such directory errors.

When quorum has been lost, requests are rejected until the node has rejoined a valid quorum and mounted its file systems. If messages indicate lack of quorum, follow the procedures in “GPFS daemon will not come up” on page 232.

- Removing the setuid bit from the permissions of these commands may produce errors for non-root users:

**mmdf**  
**mmgetacl**  
**mmlsdisk**  
**mmlsfs**  
**mmlsmgr**  
**mmlspolicy**  
**mmlsquota**  
**mmlssnapshot**  
**mmpuatacl**  
**mmsnapdir**  
**mmsnaplatest**

The GPFS system-level versions of these commands (prefixed by **ts**) may need to be checked for how permissions are set if non-root users see the following message:

**6027-1209**

GPFS is down on this node.

If the setuid bit is removed from the permissions on the system-level commands, the command cannot be executed and the node is perceived as being down. The system-level versions of the commands are:

**tsdf**  
**tslsdisk**

**tslfs**  
**tslsmgr**  
**tslspolicy**  
**tslsquota**  
**tslssnapshot**  
**tssnapdir**  
**tssnaplatest**

These are found in the `/usr/lpp/mmfs/bin` directory.

**Note:** The mode bits for all listed commands are 4555 or `-r-sr-xr-x`. To restore the default (shipped) permission, enter:

```
chmod 4555 tscommand
```

**Attention:** Only administration-level versions of GPFS commands (prefixed by **mm**) should be executed. Executing system-level commands (prefixed by **ts**) directly will produce unexpected results.

- For all other errors, follow the procedures in “Additional information to collect for GPFS daemon crashes” on page 368, and then contact the IBM Support Center.

---

## GPFS commands are unsuccessful

GPFS commands can be unsuccessful for various reasons.

Unsuccessful command results will be indicated by:

- Return codes indicating the GPFS daemon is no longer running.
- Command specific problems indicating you are unable to access the disks.
- A nonzero return code from the GPFS command.

Some reasons that GPFS commands can be unsuccessful include:

1. If all commands are generically unsuccessful, this may be due to a daemon failure. Verify that the GPFS daemon is active. Issue:

```
mmgetstate
```

If the daemon is not active, check `/var/adm/ras/mmfs.log.latest` and `/var/adm/ras/mmfs.log.previous` on the local node and on the file system manager node. These files enumerate the failing sequence of the GPFS daemon.

If there is a communication failure with the file system manager node, you will receive an error and the **errno** global variable may be set to EIO (I/O error).

2. Verify the GPFS cluster configuration data files are not locked and are accessible. To determine if the GPFS cluster configuration data files are locked, see “GPFS cluster configuration data file issues” on page 229.

3. The **ssh** command is not functioning correctly. See “Authorization problems” on page 228.

If **ssh** is not functioning properly on a node in the GPFS cluster, a GPFS administration command that needs to run work on that node will fail with a 'permission is denied' error. The system displays information similar to:

```
mmlscluster
sshd: 0826-813 Permission is denied.
mmdsh: 6027-1615 k145n02 remote shell process had return code 1.
mmlscluster: 6027-1591 Attention: Unable to retrieve GPFS cluster files from node k145n02
sshd: 0826-813 Permission is denied.
mmdsh: 6027-1615 k145n01 remote shell process had return code 1.
mmlscluster: 6027-1592 Unable to retrieve GPFS cluster files from node k145n01
```

These messages indicate that **ssh** is not working properly on nodes **k145n01** and **k145n02**.

If you encounter this type of failure, determine why **ssh** is not working on the identified node. Then fix the problem.

4. Most problems encountered during file system creation fall into three classes:
  - You did not create network shared disks which are required to build the file system.
  - The creation operation cannot access the disk.  
Follow the procedures for checking access to the disk. This can result from a number of factors including those described in “NSD and underlying disk subsystem failures” on page 275.
  - Unsuccessful attempt to communicate with the file system manager.  
The file system creation runs on the file system manager node. If that node goes down, the **mmcrfs** command may not succeed.
5. If the **mmdelnode** command was unsuccessful and you plan to permanently de-install GPFS from a node, you should first remove the node from the cluster. If this is not done and you run the **mmdelnode** command after the **mmfs** code is removed, the command will fail and display a message similar to this example:

```
Verifying GPFS is stopped on all affected nodes ...  
k145n05: ksh: /usr/lpp/mmfs/bin/mmremote: not found.
```

If this happens, power off the node and run the **mmdelnode** command again.

6. If you have successfully installed and are operating with the latest level of GPFS, but cannot run the new functions available, it is probable that you have not issued the **mmchfs -V full** or **mmchfs -V compat** command to change the version of the file system. This command must be issued for *each* of your file systems.

In addition to **mmchfs -V**, you may need to run the **mmmigratefs** command. See the *File system format changes between versions of GPFS* topic in the *IBM Spectrum Scale: Administration Guide*.

**Note:** Before issuing the **-V** option (with **full** or **compat**), see the *Migration, coexistence and compatibility* topic in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*. You must ensure that all nodes in the cluster have been migrated to the latest level of GPFS code and that you have successfully run the **mmchconfig release=LATEST** command.

Make sure you have operated with the new level of code for some time and are certain you want to migrate to the latest level of GPFS. Issue the **mmchfs -V full** command only after you have definitely decided to accept the latest level, as this will cause disk changes that are incompatible with previous levels of GPFS.

For more information about the **mmchfs** command, see the *IBM Spectrum Scale: Command and Programming Reference*.

## GPFS error messages for unsuccessful GPFS commands

This topic describes the error messages for unsuccessful GPFS commands.

If message **6027-538** is returned from the **mmcrfs** command, verify that the disk descriptors are specified correctly and that all named disks exist and are online. Issue the **mmlnsd** command to check the disks.

### 6027-538

Error accessing disks.

If the daemon failed while running the command, you will see message **6027-663**. Follow the procedures in “GPFS daemon went down” on page 236.

### 6027-663

Lost connection to file system daemon.

If the daemon was not running when you issued the command, you will see message **6027-665**. Follow the procedures in “GPFS daemon will not come up” on page 232.

6027-665

Failed to connect to file system daemon: *errorString*.

When GPFS commands are unsuccessful, the system may display information similar to these error messages:

6027-1627

The following nodes are not aware of the configuration server change: *nodeList*. Do not start GPFS on the preceding nodes until the problem is resolved.

---

## Quorum loss

Each GPFS cluster has a set of quorum nodes explicitly set by the cluster administrator.

These quorum nodes and the selected quorum algorithm determine the availability of file systems owned by the cluster. For more information, see *Quorum* in *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.

When quorum loss or loss of connectivity occurs, any nodes still running GPFS suspend the use of file systems owned by the cluster experiencing the problem. This may result in GPFS access within the suspended file system receiving **ESTALE** errors. Nodes continuing to function after suspending file system access will start contacting other nodes in the cluster in an attempt to rejoin or reform the quorum. If they succeed in forming a quorum, access to the file system is restarted.

Normally, quorum loss or loss of connectivity occurs if a node goes down or becomes isolated from its peers by a network failure. The expected response is to address the failing condition.

---

## E! CES configuration issues

E! The following are the issues that you might encounter while configuring cluster export services in IBM E! Spectrum Scale.

E! • Issue: The **mmces** command shows a socket-connection-error.

E! Error: Cannot connect to server(localhost), port(/var/mmfs/mmsysmon/mmsysmonitor.socket): Connection refused

E! Solution: The **mmsysmon-daemon** is not running or is malfunctioning. Submit the **mmsysmoncontrol restart** command to restore the functionality.

E! • Issue: The **mmcluster --ces** command does not show any CES IPs, bound to the CES-nodes.

E! Solution: Either all CES nodes are unhealthy or no IPs are defined as CES IPs. Try out the following E! steps to resolve this issue:

E! 1. Use the **mmces state show -ato** to find out the nodes in which the CES service is in the FAILED state.

E! Using the **ssh <nodeName> mmhealth node show** command displays the component that is creating the issue. In some cases, events are created if there are issues with the node health.

E! 2. Use the **mmces address list** command to list the IPs are defined as CES IPs. You can extend this E! list by issuing the command **mmces address add --ces-node ip <ipAddress>**.

---

## Application program errors

When receiving application program errors, there are various courses of action to take.

Follow these steps to help resolve application program errors:

1. Loss of file system access usually appears first as an error received by an application. Such errors are normally encountered when the application tries to access an unmounted file system.

The most common reason for losing access to a single file system is a failure somewhere in the path to a large enough number of disks to jeopardize your data if operation continues. These errors may be reported in the operating system error log on any node because they are logged in the first node to detect the error. Check all error logs for errors.

The **mmlsmount all -L** command can be used to determine the nodes that have successfully mounted a file system.

2. There are several cases where the state of a given disk subsystem will prevent access by GPFS. This will be seen by the application as I/O errors of various types and will be reported in the error logs as **MMFS\_SYSTEM\_UNMOUNT** or **MMFS\_DISKFAIL** records. This state can be found by issuing the **mmlsdisk** command.

3. If allocation of data blocks or files (which quota limits should allow) fails, issue the **mmlsquota** command for the user, group or fileset.

If filesets are involved, use these steps to determine which fileset was being accessed at the time of the failure:

- a. From the error messages generated, obtain the path name of the file being accessed.
- b. Go to the directory just obtained, and use this **mmlsattr -L** command to obtain the fileset name:

```
mmlsattr -L . | grep "fileset name:"
```

The system produces output similar to:

```
fileset name: myFileset
```

- c. Use the **mmlsquota -j** command to check the quota limit of the fileset. For example, using the fileset name found in the previous step, issue this command:

```
mmlsquota -j myFileset -e
```

The system produces output similar to:

Filesystem type	Block Limits				File Limits				Remarks		
	KB	quota	limit	in_doubt	grace	files	quota	limit		in_doubt	grace
fs1	FILESET	2152	0	0	0	none	250	0	250	0	none

The **mmlsquota** output is similar when checking the user and group quota. If usage is equal to or approaching the hard limit, or if the grace period has expired, make sure that no quotas are lost by checking *in doubt* values.

If quotas are exceeded in the *in doubt* category, run the **mmcheckquota** command. For more information, see “The mmcheckquota command” on page 199.

**Note:** There is no way to force GPFS nodes to relinquish all their local shares in order to check for lost quotas. This can only be determined by running the **mmcheckquota** command immediately after mounting the file system, and before any allocations are made. In this case, the value *in doubt* is the amount lost.

To display the latest quota usage information, use the **-e** option on either the **mmlsquota** or the **mmrepquota** commands. Remember that the **mmquotaon** and **mmquotaoff** commands do not enable and disable quota management. These commands merely control enforcement of quota limits. Usage continues to be counted and recorded in the quota files regardless of enforcement.

Reduce quota usage by deleting or compressing files or moving them out of the file system. Consider increasing quota limit.

## GPFS error messages for application program errors

This topic describes the error messages that IBM Spectrum Scale displays for application program errors.

Application program errors can be associated with these GPFS message numbers:

**6027-506**

*program: loadFile* is already loaded at *address*.



## E! Windows issues

The topics that follow apply to Windows Server 2008.

### Home and .ssh directory ownership and permissions

This topic describes the issues related to .ssh directory ownership and permissions.

Make sure users own their home directories, which is not normally the case on Windows. They should also own `~/ssh` and the files it contains. Here is an example of file attributes that work:

```
bash-3.00$ ls -l -d ~
drwx----- 1 demyn Domain Users 0 Dec 5 11:53 /dev/fs/D/Users/demyn
bash-3.00$ ls -l -d ~/.ssh
drwx----- 1 demyn Domain Users 0 Oct 26 13:37 /dev/fs/D/Users/demyn/.ssh
bash-3.00$ ls -l ~/.ssh
total 11
drwx----- 1 demyn Domain Users 0 Oct 26 13:37 .
drwx----- 1 demyn Domain Users 0 Dec 5 11:53 ..
-rw-r--r-- 1 demyn Domain Users 603 Oct 26 13:37 authorized_keys2
-rw----- 1 demyn Domain Users 672 Oct 26 13:33 id_dsa
-rw-r--r-- 1 demyn Domain Users 603 Oct 26 13:33 id_dsa.pub
-rw-r--r-- 1 demyn Domain Users 2230 Nov 11 07:57 known_hosts
bash-3.00$
```

### Problems running as Administrator

You might have problems using SSH when running as the domain **Administrator** user. These issues do not apply to other accounts, even if they are members of the **Administrators** group.

### GPFS Windows and SMB2 protocol (CIFS serving)

SMB2 is a version of the Server Message Block (SMB) protocol that was introduced with Windows Vista and Windows Server 2008.

Various enhancements include the following (among others):

- reduced “chattiness” of the protocol
- larger buffer sizes
- faster file transfers
- caching of metadata such as directory content and file properties
- better scalability by increasing the support for number of users, shares, and open files per server

The SMB2 protocol is negotiated between a client and the server during the establishment of the SMB connection, and it becomes active only if both the client and the server are SMB2 capable. If either side is not SMB2 capable, the default SMB (version 1) protocol gets used.

The SMB2 protocol does active metadata caching on the client redirector side, and it relies on Directory Change Notification on the server to invalidate and refresh the client cache. However, GPFS on Windows currently does not support Directory Change Notification. As a result, if SMB2 is used for serving out a IBM Spectrum Scale file system, the SMB2 redirector cache on the client will not see any cache-invalidate operations if the actual metadata is changed, either directly on the server or via another CIFS client. In such a case, the SMB2 client will continue to see its cached version of the directory contents until the redirector cache expires. Therefore, the use of SMB2 protocol for CIFS sharing of GPFS file systems can result in the CIFS clients seeing an inconsistent view of the actual GPFS namespace.

A workaround is to disable the SMB2 protocol on the CIFS server (that is, the GPFS compute node). This will ensure that the SMB2 never gets negotiated for file transfer even if any CIFS client is SMB2 capable.

To disable SMB2 on the GPFS compute node, follow the instructions under the “MORE INFORMATION” section at the Microsoft Support website ([support.microsoft.com/kb/974103](http://support.microsoft.com/kb/974103)).

E!

---

## E! Chapter 14. Network issues

E! This topic describes network issues that you might encounter while using IBM Spectrum Scale.

E!

### E! IBM Spectrum Scale failures due to a network failure

E! For proper functioning, GPFS depends both directly and indirectly on correct network operation.

E! This dependency is direct because various IBM Spectrum Scale internal messages flow on the network, E! and may be indirect if the underlying disk technology is dependent on the network. Symptoms included E! in an indirect failure would be inability to complete I/O or GPFS moving disks to the **down** state.

E! The problem can also be first detected by the GPFS network communication layer. If network E! connectivity is lost between nodes or GPFS heart beating services cannot sustain communication to a E! node, GPFS will declare the node dead and perform recovery procedures. This problem will manifest E! itself by messages appearing in the GPFS log such as:

```
E! Mon Jun 25 22:23:36.298 2007: Close connection to 192.168.10.109 c5n109. Attempting reconnect.  
E! Mon Jun 25 22:23:37.300 2007: Connecting to 192.168.10.109 c5n109  
E! Mon Jun 25 22:23:37.398 2007: Close connection to 192.168.10.109 c5n109  
E! Mon Jun 25 22:23:38.338 2007: Recovering nodes: 9.114.132.109  
E! Mon Jun 25 22:23:38.722 2007: Recovered 1 nodes.
```

E! Nodes mounting file systems owned and served by other clusters may receive error messages similar to E! this:

```
E! Mon Jun 25 16:11:16 2007: Close connection to 89.116.94.81 k155n01  
E! Mon Jun 25 16:11:21 2007: Lost membership in cluster remote.cluster. Unmounting file systems.
```

E! If a sufficient number of nodes fail, GPFS will lose the quorum of nodes, which exhibits itself by E! messages appearing in the GPFS log, similar to this:

```
E! Mon Jun 25 11:08:10 2007: Close connection to 179.32.65.4 gpfs2  
E! Mon Jun 25 11:08:10 2007: Lost membership in cluster gpfsxx.kgn.ibm.com. Unmounting file system.
```

E! When either of these cases occur, perform problem determination on your network connectivity. Failing E! components could be network hardware such as switches or host bus adapters.

E!

### E! OpenSSH connection delays

E! OpenSSH can be sensitive to network configuration issues that often do not affect other system E! components. One common symptom is a substantial delay (20 seconds or more) to establish a connection. E! When the environment is configured correctly, a command such as **ssh gandalf date** should only take one E! or two seconds to complete.

E! If you are using OpenSSH and experiencing an SSH connection delay (and if IPv6 is not supported in E! your environment), try disabling IPv6 on your Windows nodes and remove or comment out any IPv6 E! addresses from the **/etc/resolv.conf** file.



---

## Chapter 15. File system issues

Suspect a GPFS file system problem when a file system will not mount or unmount.

You can also suspect a file system problem if a file system unmounts unexpectedly, or you receive an error message indicating that file system activity can no longer continue due to an error, and the file system is being unmounted to preserve its integrity. Record all error messages and log entries that you receive relative to the problem, making sure that you look on all affected nodes for this data.

These are some of the errors encountered with GPFS file systems:

- E! • “File system fails to mount”
- E! • “File system fails to unmount” on page 249
  - “File system forced unmount” on page 250
  - “Unable to determine whether a file system is mounted” on page 259
  - “Multiple file system manager failures” on page 259
  - “Discrepancy between GPFS configuration data and the on-disk data for a file system” on page 260
  - “Errors associated with storage pools, filesets and policies” on page 261
  - “Failures using the mmbackup command” on page 273
  - “Snapshot problems” on page 267
  - “Failures using the mmpmon command” on page 271
  - “NFS issues” on page 297
  - “File access failure from an SMB client with sharing conflict” on page 305
  - “Data integrity” on page 273
  - “Messages requeueing in AFM” on page 274

---

### E! File system fails to mount

There are indications leading you to the conclusion that your file system will not mount and courses of action you can take to correct the problem.

Some of those indications include:

- On performing a manual mount of the file system, you get errors from either the operating system or GPFS.
- If the file system was created with the option of an automatic mount, you will have failure return codes in the GPFS log.
- Your application cannot access the data it needs. Check the GPFS log for messages.
- Return codes or error messages from the **mmm mount** command.
- The **mmlsmount** command indicates that the file system is not mounted on certain nodes.

If your file system will not mount, follow these steps:

1. On a quorum node in the cluster that owns the file system, verify that quorum has been achieved. Check the GPFS log to see if an **mmfsd** ready message has been logged, and that no errors were reported on this or other nodes.
2. Verify that a conflicting command is not running. This applies only to the cluster that owns the file system. However, other clusters would be prevented from mounting the file system if a conflicting command is running in the cluster that owns the file system.

For example, a **mount** command may not be issued while the **mmfsck** command is running. The **mount** command may not be issued until the conflicting command completes. Note that interrupting the **mmfsck** command is not a solution because the file system will not be mountable until the command completes. Try again after the conflicting command has completed.

3. Verify that sufficient disks are available to access the file system by issuing the **mmlsdisk** command. GPFS requires a minimum number of disks to find a current copy of the core metadata. If sufficient disks cannot be accessed, the mount will fail. The corrective action is to fix the path to the disk. See “NSD and underlying disk subsystem failures” on page 275.

Missing disks can also cause GPFS to be unable to find critical metadata structures. The output of the **mmlsdisk** command will show any unavailable disks. If you have not specified metadata replication, the failure of one disk may result in your file system being unable to mount. If you have specified metadata replication, it will require two disks in different failure groups to disable the entire file system. If there are down disks, issue the **mmchdisk start** command to restart them and retry the mount.

For a remote file system, **mmlsdisk** provides information about the disks of the file system. However **mmchdisk** must be run from the cluster that owns the file system.

If there are no disks down, you can also look locally for error log reports, and follow the problem determination and repair actions specified in your storage system vendor problem determination guide. If the disk has failed, follow the procedures in “NSD and underlying disk subsystem failures” on page 275.

4. Verify that communication paths to the other nodes are available. The lack of communication paths between all nodes in the cluster may impede contact with the file system manager.
5. Verify that the file system is not already mounted. Issue the **mount** command.
6. Verify that the GPFS daemon on the file system manager is available. Run the **mmlsmgr** command to determine which node is currently assigned as the file system manager. Run a trivial data access command such as an **ls** on the mount point directory. If the command fails, see “GPFS daemon went down” on page 236.
7. Check to see if the mount point directory exists and that there is an entry for the file system in the **/etc/fstab** file (for Linux) or **/etc/filesystems** file (for AIX). The device name for a file system mount point will be listed in column one of the **/etc/fstab** entry or as a **dev=** attribute in the **/etc/filesystems** stanza entry. A corresponding device name must also appear in the **/dev** file system.

If any of these elements are missing, an update to the configuration information may not have been propagated to this node. Issue the **mmrefresh** command to rebuild the configuration information on the node and reissue the **mmmount** command.

Do not add GPFS file system information to **/etc/filesystems** (for AIX) or **/etc/fstab** (for Linux) directly. If after running **mmrefresh -f** the file system information is still missing from **/etc/filesystems** (for AIX) or **/etc/fstab** (for Linux), follow the procedures in “Information to be collected before contacting the IBM Support Center” on page 367, and then contact the IBM Support Center.

8. Check the number of file systems that are already mounted. There is a maximum number of 256 mounted file systems for a GPFS cluster. Remote file systems are included in this number.
9. If you issue **mmchfs -V compat**, it enables backwardly-compatible format changes only. Nodes in remote clusters that were able to mount the file system before will still be able to do so.

If you issue **mmchfs -V full**, it enables all new functions that require different on-disk data structures. Nodes in remote clusters running an older GPFS version will no longer be able to mount the file system. If there are any nodes running an older GPFS version that have the file system mounted at the time this command is issued, the **mmchfs** command will fail. For more information about completing the migration to a new level of GPFS, see the *Completing the migration to a new level of IBM Spectrum Scale* section in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.

All nodes that access the file system must be upgraded to the same level of GPFS. Check for the possibility that one or more of the nodes was accidentally left out of an effort to upgrade a multi-node

system to a new GPFS release. If you need to return to the earlier level of GPFS, you must re-create the file system from the backup medium and restore the content in order to access it.

10. If DMAPI is enabled for the file system, ensure that a data management application is started and has set a disposition for the mount event. Refer to the *IBM Spectrum Scale: Command and Programming Reference* and the user's guide from your data management vendor. The data management application must be started in the cluster that owns the file system. If the application is not started, other clusters will not be able to mount the file system. Remote mounts of DMAPI managed file systems may take much longer to complete than those not managed by DMAPI.
11. Issue the **mmfsfs -A** command to check whether the automatic mount option has been specified. If automatic mount option is expected, check the GPFS log in the cluster that owns and serves the file system, for progress reports indicating:  
starting ...  
mounting ...  
mounted ....
12. If quotas are enabled, check if there was an error while reading quota files. See "MMFS\_QUOTA" on page 159.
13. Verify the **maxblocksize** configuration parameter on all clusters involved. If **maxblocksize** is less than the block size of the local or remote file system you are attempting to mount, you will not be able to mount it.
14. If the file system has encryption rules, see "Mount failure for a file system with encryption rules" on page 293.
15. To mount a file system on a remote cluster, ensure that the cluster that owns and serves the file system and the remote cluster have proper authorization in place. The authorization between clusters is set up with the **mmauth** command.

Authorization errors on AIX are similar to the following:

```
c13clapv6.gpfs.net: Failed to open remotefs.  
c13clapv6.gpfs.net: Permission denied  
c13clapv6.gpfs.net: Cannot mount /dev/remotefs on /gpfs/remotefs: Permission denied
```

Authorization errors on Linux are similar to the following:

```
mount: /dev/remotefs is write-protected, mounting read-only  
mount: cannot mount /dev/remotefs read-only  
mmount: 6027-1639 Command failed. Examine previous error messages to determine cause.
```

For more information about mounting a file system that is owned and served by another GPFS cluster, see the *Mounting a remote GPFS file system* topic in the *IBM Spectrum Scale: Administration Guide*.

## GPFS error messages for file system mount problems

There are error messages specific to file system reading, failure, mounting, and remounting.

### 6027-419

Failed to read a file system descriptor.

### 6027-482 [E]

Remount failed for device *name: errnoDescription*

### 6027-549

Failed to open *name*.

### 6027-580

Unable to access vital system metadata. Too many disks are unavailable.

### 6027-645

Attention: mmcommon getEFOptions *fileSystem* failed. Checking *fileName*.

## Error numbers specific to GPFS application calls when a file system mount is not successful

There are specific error numbers for unsuccessful file system mounting.

When a mount of a file system is not successful, GPFS may report these error numbers in the operating system error log or return them to an application:

**ENO\_QUOTA\_INST = 237, No Quota management enabled.**

To enable quotas for the file system, issue the **mmchfs -Q yes** command. To disable quotas for the file system issue the **mmchfs -Q no** command.

## Mount failure due to client nodes joining before NSD servers are online

While mounting a file system, specially during automounting, if a client node joins the GPFS cluster and attempts file system access prior to the file system's NSD servers being active, the mount fails. Use **mmchconfig** command to specify the amount of time for GPFS mount requests to wait for an NSD server to join the cluster.

If a client node joins the GPFS cluster and attempts file system access prior to the file system's NSD servers being active, the mount fails. This is especially true when automount is used. This situation can occur during cluster startup, or any time that an NSD server is brought online with client nodes already active and attempting to mount a file system served by the NSD server.

The file system mount failure produces a message similar to this:

```
Mon Jun 25 11:23:34 EST 2007: mmmount: Mounting file systems ...
No such device
Some file system data are inaccessible at this time.
Check error log for additional information.
After correcting the problem, the file system must be unmounted and then
mounted again to restore normal data access.
Failed to open fs1.
No such device
Some file system data are inaccessible at this time.
Cannot mount /dev/fs1 on /fs1: Missing file or filesystem
```

The GPFS log contains information similar to this:

```
Mon Jun 25 11:23:54 2007: Command: mount fs1 32414
Mon Jun 25 11:23:58 2007: Disk failure. Volume fs1. rc = 19. Physical volume sdcnsd.
Mon Jun 25 11:23:58 2007: Disk failure. Volume fs1. rc = 19. Physical volume sddnsd.
Mon Jun 25 11:23:58 2007: Disk failure. Volume fs1. rc = 19. Physical volume sdcnsd.
Mon Jun 25 11:23:58 2007: Disk failure. Volume fs1. rc = 19. Physical volume sdcnsd.
Mon Jun 25 11:23:58 2007: Disk failure. Volume fs1. rc = 19. Physical volume sdhnsd.
Mon Jun 25 11:23:58 2007: Disk failure. Volume fs1. rc = 19. Physical volume sdinsd.
Mon Jun 25 11:23:58 2007: File System fs1 unmounted by the system with return code 19
reason code 0
Mon Jun 25 11:23:58 2007: No such device
Mon Jun 25 11:23:58 2007: File system manager takeover failed.
Mon Jun 25 11:23:58 2007: No such device
Mon Jun 25 11:23:58 2007: Command: err 52: mount fs1 32414
Mon Jun 25 11:23:58 2007: Missing file or filesystem
```

Two **mmchconfig** command options are used to specify the amount of time for GPFS mount requests to wait for an NSD server to join the cluster:

### **nsdServerWaitTimeForMount**

Specifies the number of seconds to wait for an NSD server to come up at GPFS cluster startup time, after a quorum loss, or after an NSD server failure.



Valid values are between 0 and 1200 seconds. The default is 300. The interval for checking is 10 seconds. If `nsdServerWaitTimeForMount` is 0, `nsdServerWaitTimeWindowOnMount` has no effect.

#### **nsdServerWaitTimeWindowOnMount**

Specifies a time window to determine if quorum is to be considered *recently formed*.

Valid values are between 1 and 1200 seconds. The default is 600. If `nsdServerWaitTimeForMount` is 0, `nsdServerWaitTimeWindowOnMount` has no effect.

The GPFS daemon need not be restarted in order to change these values. The scope of these two operands is the GPFS cluster. The `-N` flag can be used to set different values on different nodes. In this case, the settings on the file system manager node take precedence over the settings of nodes trying to access the file system.

When a node rejoins the cluster (after it was expelled, experienced a communications problem, lost quorum, or other reason for which it dropped connection and rejoined), that node resets all the failure times that it knows about. Therefore, when a node rejoins it sees the NSD servers as never having failed. From the node's point of view, it has rejoined the cluster and old failure information is no longer relevant.

GPFS checks the cluster formation criteria first. If that check falls outside the window, GPFS then checks for NSD server fail times being within the window.

---

## E! **File system fails to unmount**

There are indications leading you to the conclusion that your file system will not unmount and a course of action to correct the problem.

Those indications include:

- Return codes or error messages indicate the file system will not unmount.
- The `mmismount` command indicates that the file system is still mounted on one or more nodes.
- Return codes or error messages from the `mmumount` command.

If your file system will not unmount, follow these steps:

1. If you get an error message similar to:

```
umount: /gpfs1: device is busy
```

the file system will not unmount until all processes are finished accessing it. If `mmfsd` is up, the processes accessing the file system can be determined. See “The `lsdf` command” on page 192. These processes can be killed with the command:

```
lsdf filesystem | grep -v COMMAND | awk '{print $2}' | xargs kill -9
```

If `mmfsd` is not operational, the `lsdf` command will not be able to determine which processes are still accessing the file system.

For Linux nodes it is possible to use the `/proc` pseudo file system to determine current file access. For each process currently running on the system, there is a subdirectory `/proc/pid/fd`, where `pid` is the numeric process ID number. This subdirectory is populated with symbolic links pointing to the files that this process has open. You can examine the contents of the `fd` subdirectory for all running processes, manually or with the help of a simple script, to identify the processes that have open files in GPFS file systems. Terminating all of these processes may allow the file system to unmount successfully.

2. Verify that there are no disk media failures.

Look on the NSD server node for error log entries. Identify any NSD server node that has generated an error log entry. See “Disk media failure” on page 283 for problem determination and repair actions to follow.

3. If the file system *must* be unmounted, you can force the unmount by issuing the **mmumount -f** command:

**Note:**

- a. See “File system forced unmount” for the consequences of doing this.
- b. Before forcing the unmount of the file system, issue the **lsof** command and close any files that are open.
- c. On Linux, you might encounter a situation where a GPFS file system cannot be unmounted, even if you issue the **mmumount -f** command. In this case, you must reboot the node to clear the condition. You can also try the system **umount** command before you reboot. For example:

```
umount -f /fileSystem
```

4. If a file system that is mounted by a remote cluster needs to be unmounted, you can force the unmount by issuing the command:

```
mmumount fileSystem -f -C RemoteClusterName
```

---

## Remote node expelled after remote file system successfully mounted

This problem produces 'node expelled from cluster' messages.

One cause of this condition is when the **subnets** attribute of the **mmchconfig** command has been used to specify subnets to GPFS, and there is an incorrect netmask specification on one or more nodes of the clusters involved in the remote mount. Check to be sure that all netmasks are correct for the network interfaces used for GPFS communication.

---

## File system forced unmount

There are indications that lead you to the conclusion that your file system has been forced to unmount and various courses of action that you can take to correct the problem.

Those indications are:

- Forced unmount messages in the GPFS log.
- Your application no longer has access to data.
- Your application is getting ESTALE or ENOENT return codes.
- Multiple unsuccessful attempts to appoint a file system manager may cause the cluster manager to unmount the file system everywhere.

Such situations involve the failure of paths to disk resources from many, if not all, nodes. The underlying problem may be at the disk subsystem level, or lower. The error logs for each node that unsuccessfully attempted to appoint a file system manager will contain records of a file system unmount with an error that are either coded **212**, or that occurred when attempting to assume management of the file system. Note that these errors apply to a specific file system although it is possible that shared disk communication paths will cause the unmount of multiple file systems.

- File system unmounts with an error indicating too many disks are unavailable.

The **mmlsmount -L** command can be used to determine which nodes currently have a given file system mounted.

If your file system has been forced to unmount, follow these steps:

1. With the failure of a single disk, if you have not specified multiple failure groups and replication of metadata, GPFS will not be able to continue because it cannot write logs or other critical metadata. If you have specified multiple failure groups and replication of metadata, the failure of multiple disks in different failure groups will put you in the same position. In either of these situations, GPFS will forcibly unmount the file system. This will be indicated in the error log by records indicating exactly which access failed, with an **MMFS\_SYSTEM\_UNMOUNT** record indicating the forced unmount.

The user response to this is to take the needed actions to restore the disk access and issue the **mmchdisk** command to disks that are shown as down in the information displayed by the **mmlsdisk** command.

2. Internal errors in processing data on a single file system may cause loss of file system access. These errors may clear with the invocation of the **umount** command, followed by a remount of the file system, but they should be reported as problems to the IBM Support Center.
3. If an **MMFS\_QUOTA** error log entry containing Error writing quota file... is generated, the quota manager continues operation if the next write for the user, group, or fileset is successful. If not, further allocations to the file system will fail. Check the error code in the log and make sure that the disks containing the quota file are accessible. Run the **mmcheckquota** command. For more information, see “The mmcheckquota command” on page 199.

If the file system must be repaired without quotas:

- a. Disable quota management by issuing the command:  
`mmchfs Device -Q no`
  - b. Issue the **mmmumount** command for the file system.
  - c. Make any necessary repairs and install the backup quota files.
  - d. Issue the **mmumount -a** command for the file system.
  - e. Restore quota management by issuing the **mmchfs Device -Q yes** command.
  - f. Run the **mmcheckquota** command with the **-u**, **-g**, and **-j** options. For more information, see “The mmcheckquota command” on page 199.
  - g. Issue the **mmmumount** command for the file system.
4. If errors indicate that too many disks are unavailable, see “Additional failure group considerations.”

## Additional failure group considerations

GPFS uses *file system descriptor* to be replicated on a subset of the disks as changes to the file system occur, such as adding or deleting disks. To reduce the risk of multiple failure GPFS picks disks to hold the replicas in different failure group.

There is a structure in GPFS called the *file system descriptor* that is initially written to every disk in the file system, but is replicated on a subset of the disks as changes to the file system occur, such as adding or deleting disks. Based on the number of failure groups and disks, GPFS creates between one and five replicas of the descriptor:

- If there are at least five different failure groups, five replicas are created.
- If there are at least three different disks, three replicas are created.
- If there are only one or two disks, a replica is created on each disk.

Once it is decided how many replicas to create, GPFS picks disks to hold the replicas, so that all replicas will be in different failure groups, if possible, to reduce the risk of multiple failures. In picking replica locations, the current state of the disks is taken into account. Stopped or suspended disks are avoided. Similarly, when a failed disk is brought back online, GPFS may modify the subset to rebalance the file system descriptors across the failure groups. The subset can be found by issuing the **mmlsdisk -L** command.

GPFS requires a majority of the replicas on the subset of disks to remain available to sustain file system operations:

- If there are at least five different failure groups, GPFS will be able to tolerate a loss of two of the five groups. If disks out of three different failure groups are lost, the file system descriptor may become inaccessible due to the loss of the majority of the replicas.
- If there are at least three different failure groups, GPFS will be able to tolerate a loss of one of the three groups. If disks out of two different failure groups are lost, the file system descriptor may become inaccessible due to the loss of the majority of the replicas.

- If there are fewer than three failure groups, a loss of one failure group may make the descriptor inaccessible.

If the subset consists of three disks and there are only two failure groups, one failure group must have two disks and the other failure group has one. In a scenario that causes one entire failure group to disappear all at once, if the half of the disks that are unavailable contain the single disk that is part of the subset, everything stays up. The file system descriptor is moved to a new subset by updating the remaining two copies and writing the update to a new disk added to the subset. But if the downed failure group contains a majority of the subset, the file system descriptor cannot be updated and the file system has to be force unmounted.

Introducing a third failure group consisting of a single disk that is used solely for the purpose of maintaining a copy of the file system descriptor can help prevent such a scenario. You can designate this disk by using the **descOnly** designation for disk usage on the disk descriptor. For more information on disk replication, see the *NSD creation considerations* topic in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide* and the *Data mirroring and replication* topic in the *IBM Spectrum Scale: Administration Guide*.

## GPFS error messages for file system forced unmount problems

There are many error messages for file system forced unmount problems due to unavailable disk space.

Indications there are not enough disks available:

**6027-418**

Inconsistent file system quorum. readQuorum=*value* writeQuorum=*value* quorumSize=*value*.

**6027-419**

Failed to read a file system descriptor.

Indications the file system has been forced to unmount:

**6027-473 [X]**

File System *fileSystem* unmounted by the system with return code *value* reason code *value*

**6027-474 [X]**

Recovery Log I/O failed, unmounting file system *fileSystem*

## Error numbers specific to GPFS application calls when a file system has been forced to unmount

There are error numbers to indicate that a file system is forced to unmount for GPFS application calls.

When a file system has been forced to unmount, GPFS may report these error numbers in the operating system error log or return them to an application:

**EPANIC = 666, A file system has been forcibly unmounted because of an error. Most likely due to the failure of one or more disks containing the last copy of metadata.**

E! See "Operating system error logs" on page 157 for details.

**EALL\_UNAVAIL = 218, A replicated read or write failed because none of the replicas were available.**

Multiple disks in multiple failure groups are unavailable. Follow the procedures in Chapter 16, "Disk issues," on page 275 for unavailable disks.

---

## Automount file system will not mount

The automount fails to mount the file system and the courses of action that you can take to correct the problem.

If an automount fails when you `cd` into the mount point directory, first check that the file system in question is of automount type. Use the `mmlsfs -A` command for local file systems. Use the `mmremotefs show` command for remote file systems.

## Steps to follow if automount fails to mount on Linux

There are course of actions that you can take if the automount fails to mount on Linux system.

On Linux, perform these steps:

1. Verify that the GPFS file system mount point is actually a symbolic link to a directory in the `automountdir` directory. If `automountdir=/gpfs/autmountdir` then the mount point `/gpfs/gpfs66` would be a symbolic link to `/gpfs/automountdir/gpfs66`.
  - a. First, verify that GPFS is up and running.
  - b. Use the `mmlsconfig` command to verify the `automountdir` directory. The default `automountdir` is named `/gpfs/automountdir`. If the GPFS file system mount point is not a symbolic link to the GPFS `automountdir` directory, then accessing the mount point will not cause the automounter to mount the file system.
  - c. If the command `/bin/ls -ld` of the mount point shows a directory, then run the command `mmrefresh -f`. If the directory is empty, the command `mmrefresh -f` will remove the directory and create a symbolic link. If the directory is not empty, you need to move or remove the files contained in that directory, or change the mount point of the file system. For a local file system, use the `mmchfs` command. For a remote file system, use the `mmremotefs` command.
  - d. Once the mount point directory is empty, run the `mmrefresh -f` command.
2. Verify that the `autofs` mount has been established. Issue this command:  

```
mount | grep automount
```

The output must be similar to this:

```
automount(pid20331) on /gpfs/automountdir type autofs (rw,fd=5,pgrp=20331,minproto=2,maxproto=3)
```

For Red Hat Enterprise Linux 5, verify the following line is in the default master map file (`/etc/auto.master`):

```
/gpfs/automountdir program:/usr/lpp/mmfs/bin/mmdynamicmap
```

For example, issue:

```
grep mmdynamicmap /etc/auto.master
```

Output should be similar to this:

```
/gpfs/automountdir program:/usr/lpp/mmfs/bin/mmdynamicmap
```

This is an `autofs` program map, and there will be a single mount entry for all GPFS automounted file systems. The symbolic link points to this directory, and access through the symbolic link triggers the mounting of the target GPFS file system. To create this GPFS `autofs` mount, issue the `mmcommon startAutomounter` command, or stop and restart GPFS using the `mmshutdown` and `mmstartup` commands.

3. Verify that the automount daemon is running. Issue this command:

```
ps -ef | grep automount
```

The output must be similar to this:

```
root 5116 1 0 Jun25 pts/0 00:00:00 /usr/sbin/automount /gpfs/automountdir program
      /usr/lpp/mmfs/bin/mmdynamicmap
```

For Red Hat Enterprise Linux 5, verify that the `autofs` daemon is running. Issue this command:

```
ps -ef | grep automount
```

The output must be similar to this:

```
root 22646 1 0 01:21 ? 00:00:02 automount
```

To start the automount daemon, issue the **mmcommon startAutomounter** command, or stop and restart GPFS using the **mmshutdown** and **mmstartup** commands.

**Note:** If **automountdir** is mounted (as in step 2) and the **mmcommon startAutomounter** command is not able to bring up the automount daemon, manually **umount** the **automountdir** before issuing the **mmcommon startAutomounter** again.

4. Verify that the mount command was issued to GPFS by examining the GPFS log. You should see something like this:

```
Mon Jun 25 11:33:03 2004: Command: mount gpfsx2.kgn.ibm.com:gpfs55 5182
```

5. Examine **/var/log/messages** for **autofs** error messages. The following is an example of what you might see if the remote file system name does not exist.

```
Jun 25 11:33:03 linux automount[20331]: attempting to mount entry /gpfs/automountdir/gpfs55
Jun 25 11:33:04 linux automount[28911]: >> Failed to open gpfs55.
Jun 25 11:33:04 linux automount[28911]: >> No such device
Jun 25 11:33:04 linux automount[28911]: >> mount: fs type gpfs not supported by kernel
Jun 25 11:33:04 linux automount[28911]: mount(generic): failed to mount /dev/gpfs55 (type gpfs)
on /gpfs/automountdir/gpfs55
```

6. After you have established that GPFS has received a mount request from **autofs** (Step 4) and that mount request failed (Step 5), issue a mount command for the GPFS file system and follow the

E! directions in “File system fails to mount” on page 245.

## Steps to follow if automount fails to mount on AIX

There are course of actions that you can take if the automount fails to mount on AIX server.

On AIX, perform these steps:

1. First, verify that GPFS is up and running.
2. Verify that GPFS has established **autofs** mounts for each automount file system. Issue the following command:

```
mount | grep autofs
```

The output is similar to this:

```
/var/mmfs/gen/mmDirectMap /gpfs/gpfs55 autofs Jun 25 15:03 ignore
/var/mmfs/gen/mmDirectMap /gpfs/gpfs88 autofs Jun 25 15:03 ignore
```

These are direct mount **autofs** mount entries. Each GPFS automount file system will have an **autofs** mount entry. These **autofs** direct mounts allow GPFS to mount on the GPFS mount point. To create any missing GPFS **autofs** mounts, issue the **mmcommon startAutomounter** command, or stop and restart GPFS using the **mmshutdown** and **mmstartup** commands.

3. Verify that the **autofs** daemon is running. Issue this command:

```
ps -ef | grep automount
```

Output is similar to this:

```
root 9820 4240 0 15:02:50 - 0:00 /usr/sbin/automountd
```

To start the automount daemon, issue the **mmcommon startAutomounter** command, or stop and restart GPFS using the **mmshutdown** and **mmstartup** commands.

4. Verify that the mount command was issued to GPFS by examining the GPFS log. You should see something like this:

```
Mon Jun 25 11:33:03 2007: Command: mount gpfsx2.kgn.ibm.com:gpfs55 5182
```

5. Since the **autofs** daemon logs status using **syslogd**, examine the **syslogd** log file for status information from **automountd**. Here is an example of a failed automount request:

```
Jun 25 15:55:25 gpfsa1 automountd [9820 ] :mount of /gpfs/gpfs55:status 13
```

6. After you have established that GPFS has received a mount request from **autofs** (Step 4 on page 254) and that mount request failed (Step 5 on page 254), issue a mount command for the GPFS file system and follow the directions in “File system fails to mount” on page 245.
7. If automount fails for a non-GPFS file system and you are using file **/etc/auto.master**, use file **/etc/auto\_master** instead. Add the entries from **/etc/auto.master** to **/etc/auto\_master** and restart the automount daemon.

---

## Remote file system will not mount

The remote file system mounting failure reasons and the course of action that you can take to resolve the issue.

When a remote file system does not mount, the problem might be with how the file system was defined to both the local and remote nodes, or the communication paths between them. Review the *Mounting a file system owned and served by another GPFS cluster* topic in the *IBM Spectrum Scale: Administration Guide* to ensure that your setup is correct.

These are some of the errors encountered when mounting remote file systems:

- “Remote file system I/O fails with the “Function not implemented” error message when UID mapping is enabled”
- “Remote file system will not mount due to differing GPFS cluster security configurations” on page 256
- “Cannot resolve contact node address” on page 256
- “The remote cluster name does not match the cluster name supplied by the `mmremoteclass` command” on page 256
- “Contact nodes down or GPFS down on contact nodes” on page 257
- “GPFS is not running on the local node” on page 257
- “The NSD disk does not have an NSD server specified and the mounting cluster does not have direct access to the disks” on page 258
- “The `cipherList` option has not been set properly” on page 258
- “Remote mounts fail with the “permission denied” error message” on page 259

## Remote file system I/O fails with the “Function not implemented” error message when UID mapping is enabled

There are error messages when remote file system has an I/O failure and the course of action that you can take to correct this issue.

When user ID (UID) mapping in a multi-cluster environment is enabled, certain kinds of mapping infrastructure configuration problems might result in I/O requests on a remote file system failing:

```
ls -l /fs1/testfile
ls: /fs1/testfile: Function not implemented
```

To troubleshoot this error, verify the following configuration details:

1. That **/var/mmfs/etc/mmuid2name** and **/var/mmfs/etc/mmname2uid** helper scripts are present and executable on all nodes in the local cluster and on all quorum nodes in the file system home cluster, along with any data files needed by the helper scripts.
2. That UID mapping is enabled in both local cluster and remote file system home cluster configuration by issuing the **mmfsconfig enableUIDremap** command.
3. That UID mapping helper scripts are working correctly.

For more information about configuring UID mapping, see the IBM white paper entitled *UID Mapping for GPFS in a Multi-cluster Environment* in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/SSFKN/com.ibm.cluster.gpfs.doc/gpfs\\_uid/uid\\_gpfs.html](http://www.ibm.com/support/knowledgecenter/SSFKN/com.ibm.cluster.gpfs.doc/gpfs_uid/uid_gpfs.html)).

## Remote file system will not mount due to differing GPFS cluster security configurations

There are indications leading you to the conclusion that the remote file system will not mount and courses of action you can take to correct the problem.

A mount command fails with a message similar to this:

```
Cannot mount gpfsxx2.ibm.com:gpfs66: Host is down.
```

The GPFS log on the cluster issuing the mount command should have entries similar to these:

```
There is more information in the log file /var/adm/ras/mmfs.log.latest
Mon Jun 25 16:39:27 2007: Waiting to join remote cluster gpfsxx2.ibm.com
Mon Jun 25 16:39:27 2007: Command: mount gpfsxx2.ibm.com:gpfs66 30291
Mon Jun 25 16:39:27 2007: The administrator of 199.13.68.12 gpfs1x2 requires
secure connections. Contact the administrator to obtain the target clusters
key and register the key using "mmremoteccluster update".
Mon Jun 25 16:39:27 2007: A node join was rejected. This could be due to
incompatible daemon versions, failure to find the node
in the configuration database, or no configuration manager found.
Mon Jun 25 16:39:27 2007: Failed to join remote cluster gpfsxx2.ibm.com
Mon Jun 25 16:39:27 2007: Command err 693: mount gpfsxx2.ibm.com:gpfs66 30291
```

The GPFS log file on the cluster that owns and serves the file system will have an entry indicating the problem as well, similar to this:

```
Mon Jun 25 16:32:21 2007: Kill accepted connection from 199.13.68.12 because security is required, err 74
```

To resolve this problem, contact the administrator of the cluster that owns and serves the file system to obtain the key and register the key using **mmremoteccluster** command.

The SHA digest field of the **mmauth show** and **mmremoteccluster** commands may be used to determine if there is a key mismatch, and on which cluster the key should be updated. For more information on the SHA digest, see "The SHA digest" on page 204.

## Cannot resolve contact node address

There are error messages which are displayed if the contact node address does not get resolved and the courses of action you can take to correct the problem.

The following error may occur if the contact nodes for **gpfsyy2.ibm.com** could not be resolved. You would expect to see this if your DNS server was down, or the contact address has been deleted.

```
Mon Jun 25 15:24:14 2007: Command: mount gpfsyy2.ibm.com:gpfs14 20124
Mon Jun 25 15:24:14 2007: Host 'gpfs123.ibm.com' in gpfsyy2.ibm.com is not valid.
Mon Jun 25 15:24:14 2007: Command err 2: mount gpfsyy2.ibm.com:gpfs14 20124
```

To resolve the problem, correct the contact list and try the mount again.

## The remote cluster name does not match the cluster name supplied by the mmremoteccluster command

There are error messages that gets displayed if the remote cluster name does not match with the cluster name , provided by mmremoteccluster command, and the courses of action you can take to correct the problem.

A mount command fails with a message similar to this:

```
Cannot mount gpfs1x2:gpfs66: Network is unreachable
```

and the GPFS log contains message similar to this:



```

Mon Jun 25 12:47:18 2007: Waiting to join remote cluster gpfs1x2
Mon Jun 25 12:47:18 2007: Command: mount gpfs1x2:gpfs66 27226
Mon Jun 25 12:47:18 2007: Failed to join remote cluster gpfs1x2
Mon Jun 25 12:47:18 2007: Command err 719: mount gpfs1x2:gpfs66 27226

```

Perform these steps:

1. Verify that the remote cluster name reported by the **mmremotefs show** command is the same name as reported by the **mmlscluster** command from one of the contact nodes.
2. Verify the list of contact nodes against the list of nodes as shown by the **mmlscluster** command from the remote cluster.

In this example, the correct cluster name is **gpfs1x2.ibm.com** and not **gpfs1x2**

```
mmlscluster
```

Output is similar to this:

```
GPFS cluster information
```

```
=====
```

```

GPFS cluster name:      gpfs1x2.ibm.com
GPFS cluster id:       649437685184692490
GPFS UID domain:      gpfs1x2.ibm.com
Remote shell command:  /usr/bin/ssh
Remote file copy command: /usr/bin/scp
Repository type:      server-based

```

```
GPFS cluster configuration servers:
```

```
-----
```

```

Primary server:  gpfs1x2.ibm.com
Secondary server: (none)

```

Node	Daemon node name	IP address	Admin node name	Designation
1	gpfs1x2	198.117.68.68	gpfs1x2.ibm.com	quorum

## Contact nodes down or GPFS down on contact nodes

There are error messages that gets displayed if the contact nodes are down or the GPFS is down on the contact nodes, and the courses of action you can take to correct the problem.

A mount command fails with a message similar to this:

```

GPFS: 6027-510 Cannot mount /dev/gpfs22 on /gpfs22: A remote host did not respond
within the timeout period.

```

The GPFS log will have entries similar to this:

```

Mon Jun 25 13:11:14 2007: Command: mount gpfs1x22:gpfs22 19004
Mon Jun 25 13:11:14 2007: Waiting to join remote cluster gpfs1x22
Mon Jun 25 13:11:15 2007: Connecting to 199.13.68.4 gpfs1x22
Mon Jun 25 13:16:36 2007: Failed to join remote cluster gpfs1x22
Mon Jun 25 13:16:36 2007: Command err 78: mount gpfs1x22:gpfs22 19004

```

To resolve the problem, use the **mmremotecluster show** command and verify that the cluster name matches the remote cluster and the contact nodes are valid nodes in the remote cluster. Verify that GPFS is active on the contact nodes in the remote cluster. Another way to resolve this problem is to change the contact nodes using the **mmremotecluster update** command.

## GPFS is not running on the local node

There are error messages that gets displayed if the GPFS does not run on the local nodes, and the courses of action that you can take to correct the problem.

A mount command fails with a message similar to this:

mount: fs type gpfs not supported by kernel

Follow your procedures for starting GPFS on the local node.

## The NSD disk does not have an NSD server specified and the mounting cluster does not have direct access to the disks

There are error messages that gets displayed if the file system mounting gets failed, and the courses of action that you can take to correct the problem.

A file system mount fails with a message similar to this:

```
Failed to open gpfs66.  
No such device  
mount: Stale NFS file handle  
Some file system data are inaccessible at this time.  
Check error log for additional information.  
Cannot mount gpfs1x2.ibm.com:gpfs66: Stale NFS file handle
```

The GPFS log will contain information similar to this:

```
Mon Jun 25 14:10:46 2007: Command: mount gpfs1x2.ibm.com:gpfs66 28147  
Mon Jun 25 14:10:47 2007: Waiting to join remote cluster gpfs1x2.ibm.com  
Mon Jun 25 14:10:47 2007: Connecting to 199.13.68.4 gpfs1x2  
Mon Jun 25 14:10:47 2007: Connected to 199.13.68.4 gpfs1x2  
Mon Jun 25 14:10:47 2007: Joined remote cluster gpfs1x2.ibm.com  
Mon Jun 25 14:10:48 2007: Global NSD disk, gpfs1nsd, not found.  
Mon Jun 25 14:10:48 2007: Disk failure. Volume gpfs66. rc = 19. Physical volume gpfs1nsd.  
Mon Jun 25 14:10:48 2007: File System gpfs66 unmounted by the system with return code 19 reason code 0  
Mon Jun 25 14:10:48 2007: No such device  
Mon Jun 25 14:10:48 2007: Command err 666: mount gpfs1x2.ibm.com:gpfs66 28147
```

To resolve the problem, the cluster that owns and serves the file system must define one or more NSD servers.

## The cipherList option has not been set properly

There remote mount failure,due to invalid value of cipherList, leads the error messages and the course of actions that you can take to resolve the issue.

Another reason for remote mount to fail is if **cipherList** is not set to a valid value. A mount command would fail with messages similar to this:

```
6027-510 Cannot mount /dev/dqfs1 on /dqfs1: A remote host is not available.
```

The GPFS log would contain messages similar to this:

```
Wed Jul 18 16:11:20.496 2007: Command: mount remote.cluster:fs3 655494  
Wed Jul 18 16:11:20.497 2007: Waiting to join remote cluster remote.cluster  
Wed Jul 18 16:11:20.997 2007: Remote mounts are not enabled within this cluster. \  
See the Advanced Administration Guide for instructions. In particular ensure keys have been \  
generated and a cipherlist has been set.  
Wed Jul 18 16:11:20.998 2007: A node join was rejected. This could be due to  
incompatible daemon versions, failure to find the node  
in the configuration database, or no configuration manager found.  
Wed Jul 18 16:11:20.999 2007: Failed to join remote cluster remote.cluster  
Wed Jul 18 16:11:20.998 2007: Command: err 693: mount remote.cluster:fs3 655494  
Wed Jul 18 16:11:20.999 2007: Message failed because the destination node refused the connection.
```

The **mmchconfig cipherlist=AUTHONLY** command must be run on both the cluster that owns and controls the file system, and the cluster that is attempting to mount the file system.

## Remote mounts fail with the “permission denied” error message

There are many reasons why remote mounts can fail with a “permission denied” error message.

Follow these steps to resolve permission denied problems:

1. Check with the remote cluster's administrator to make sure that the proper keys are in place. The **mmauth show** command on both clusters will help with this.
2. Check that the grant access for the remote mounts has been given on the remote cluster with the **mmauth grant** command. Use the **mmauth show** command from the remote cluster to verify this.
3. Check that the file system access permission is the same on both clusters using the **mmauth show** command and the **mmremotefs show** command. If a remote cluster is only allowed to do a read-only mount (see the **mmauth show** command), the remote nodes must specify **-o ro** on their mount requests (see the **mmremotefs show** command). If you try to do remote mounts with read/write (**rw**) access for remote mounts that have read-only (**ro**) access, you will get a “permission denied” error.

For detailed information about the **mmauth** command and the **mmremotefs** command, see the *mmauth command* and the *mmremotefs command* pages in the *IBM Spectrum Scale: Command and Programming Reference*.

---

## Unable to determine whether a file system is mounted

Certain GPFS file system commands cannot be performed when the file system in question is mounted.

In certain failure situations, GPFS cannot determine whether the file system in question is mounted or not, and so cannot perform the requested command. In such cases, message **6027-1996** (Command was unable to determine whether file system *fileSystem* is mounted) is issued.

If you encounter this message, perform problem determination, resolve the problem, and reissue the command. If you cannot determine or resolve the problem, you may be able to successfully run the command by first shutting down the GPFS daemon on all nodes of the cluster (using **mmshutdown -a**), thus ensuring that the file system is not mounted.

## GPFS error messages for file system mount status

The GPFS file system commands displays error message when they are unable to determine if the file system in question is mounted.

### 6027-1996

Command was unable to determine whether file system *fileSystem* is mounted.

---

## Multiple file system manager failures

The correct operation of GPFS requires that one node per file system function as the file system manager at all times. This instance of GPFS has additional responsibilities for coordinating usage of the file system.

When the file system manager node fails, another file system manager is appointed in a manner that is not visible to applications except for the time required to switch over.

There are situations where it may be impossible to appoint a file system manager. Such situations involve the failure of paths to disk resources from many, if not all, nodes. In this event, the cluster manager nominates several host names to successively try to become the file system manager. If none succeed, the cluster manager unmounts the file system everywhere. See “NSD and underlying disk subsystem failures” on page 275.

The required action here is to address the underlying condition that caused the forced unmounts and then remount the file system. In most cases, this means correcting the path to the disks required by GPFS.

If NSD disk servers are being used, the most common failure is the loss of access through the communications network. If SAN access is being used to all disks, the most common failure is the loss of connectivity through the SAN.

## GPFS error messages for multiple file system manager failures

Certain GPFS error messages are displayed for multiple file system manager failures.

The inability to successfully appoint a file system manager after multiple attempts can be associated with both the error messages listed in “File system forced unmount” on page 250, as well as these additional messages:

- When a forced unmount occurred on all nodes:

**6027-635 [E]**

The current file system manager failed and no new manager will be appointed.

- If message **6027-636** is displayed, it means that there may be a disk failure. See “NSD and underlying disk subsystem failures” on page 275 for NSD problem determination and repair procedures.

**6027-636 [E]**

Disk marked as stopped or offline.

- Message **6027-632** is the last message in this series of messages. See the accompanying messages:

**6027-632**

Failed to appoint new manager for *fileSystem*.

- Message **6027-631** occurs on each attempt to appoint a new manager (see the messages on the referenced node for the specific reason as to why it failed):

**6027-631**

Failed to appoint node *nodeName* as manager for *fileSystem*.

- Message **6027-638** indicates which node had the original error (probably the original file system manager node):

**6027-638 [E]**

File system *fileSystem* unmounted by node *nodeName*

## Error numbers specific to GPFS application calls when file system manager appointment fails

Certain error numbers and messages are displayed when the file system manager appointment fails .

When the appointment of a file system manager is unsuccessful after multiple attempts, GPFS may report these error numbers in error logs, or return them to an application:

**ENO\_MGR = 212, The current file system manager failed and no new manager could be appointed.**

This usually occurs when a large number of disks are unavailable or when there has been a major network failure. Run **mmlsdisk** to determine whether disks have failed and take corrective action if they have by issuing the **mmchdisk** command.

---

## Discrepancy between GPFS configuration data and the on-disk data for a file system

There is an indication leading you to the conclusion that there may be a discrepancy between the GPFS configuration data and the on-disk data for a file system.

You issue a disk command (for example, **mmaddisk**, **mmdeldisk**, or **mmrpldisk**) and receive the message:

6027-1290

GPFS configuration data for file system *fileSystem* may not be in agreement with the on-disk data for the file system. Issue the command:

```
mmcommon recoverfs fileSystem
```

Before a disk is added to or removed from a file system, a check is made that the GPFS configuration data for the file system is in agreement with the on-disk data for the file system. The preceding message is issued if this check was not successful. This may occur if an earlier GPFS disk command was unable to complete successfully for some reason. Issue the **mmcommon recoverfs** command to bring the GPFS configuration data into agreement with the on-disk data for the file system.

If running **mmcommon recoverfs** does not resolve the problem, follow the procedures in “Information to be collected before contacting the IBM Support Center” on page 367, and then contact the IBM Support Center.

---

## Errors associated with storage pools, filesets and policies

There are certain error messages associated with the storage pools, filesets and policies.

When an error is suspected while working with storage pools, policies and filesets, check the relevant section in the *IBM Spectrum Scale: Administration Guide* to ensure that your setup is correct.

When you are sure that your setup is correct, see if your problem falls into one of these categories:

- “A NO\_SPACE error occurs when a file system is known to have adequate free space”
- “Negative values occur in the 'predicted pool utilizations', when some files are 'ill-placed'” on page 263
- “Policies - usage errors” on page 263
- “Errors encountered with policies” on page 264
- “Filesets - usage errors” on page 265
- “Errors encountered with filesets” on page 265
- “Storage pools - usage errors” on page 266
- “Errors encountered with storage pools” on page 267

### A NO\_SPACE error occurs when a file system is known to have adequate free space

The GPFS commands display a NO\_SPACE error even if a file system has free space and the course of actions that you can take to correct this issue.

A ENOSPC (NO\_SPACE) message can be returned even if a file system has remaining space. The NO\_SPACE error might occur even if the **df** command shows that the file system is not full.

The user might have a policy that writes data into a specific storage pool. When the user tries to create a file in that storage pool, it returns the ENOSPC error if the storage pool is full. The user next issues the **df** command, which indicates that the file system is not full, because the problem is limited to the one storage pool in the user's policy. In order to see if a particular storage pool is full, the user must issue the **mmdf** command.

The following is a sample scenario:

1. The user has a policy rule that says files whose name contains the word 'tmp' should be put into storage pool **sp1** in the file system **fs1**. This command displays the rule:

```
mmfspolicy fs1 -L
```

The system displays an output similar to this:

```

/* This is a policy for GPFS file system fs1 */

/* File Placement Rules */
RULE SET POOL 'sp1' WHERE name like '%tmp%'
RULE 'default' SET POOL 'system'
/* End of Policy */

```

- The user moves a file from the **/tmp** directory to **fs1** that has the word 'tmp' in the file name, meaning data of **tmpfile** should be placed in storage pool **sp1**:

```
mv /tmp/tmpfile /fs1/
```

The system produces output similar to this:

```
mv: writing `~/fs1/tmpfile': No space left on device
```

This is an out-of-space error.

- This command shows storage information for the file system:

```
df |grep fs1
```

The system produces output similar to this:

```
/dev/fs1          280190976 140350976 139840000  51% /fs1
```

This output indicates that the file system is only 51% full.

- To query the storage usage for an individual storage pool, the user must issue the **mmdf** command.

```
mmdf fs1
```

The system produces output similar to this:

disk name	disk size in KB	failure group	holds metadata	holds data	free KB in full blocks	free KB in fragments
Disks in storage pool: system						
gpfs1nsd	140095488	4001	yes	yes	139840000 (100%)	19936 ( 0%)
(pool total)	140095488				139840000 (100%)	19936 ( 0%)
Disks in storage pool: sp1						
gpfs2nsd	140095488	4001	no	yes	0s ( 0%)	248 ( 0%)
(pool total)	140095488				0 ( 0%)	248 ( 0%)
(data)	280190976				139840000 ( 50%)	20184 ( 0%)
(metadata)	140095488				139840000 (100%)	19936 ( 0%)
(total)	280190976				139840000 ( 50%)	20184 ( 0%)

Inode Information

```

-----
Number of used inodes: 74
Number of free inodes: 137142
Number of allocated inodes: 137216
Maximum number of inodes: 150016

```

In this case, the user sees that storage pool **sp1** has 0% free space left and that is the reason for the **NO\_SPACE** error message.

- To resolve the problem, the user must change the placement policy file to avoid putting data in a full storage pool, delete some files in storage pool **sp1**, or add more space to the storage pool.

## Negative values occur in the 'predicted pool utilizations', when some files are 'ill-placed'

A scenario where an ill-placed files may cause GPFS to produce a 'Predicted Pool Utilization' of a negative value and the course of action that you can take to resolve this issue.

This is a hypothetical situation where ill-placed files can cause GPFS to produce a 'Predicted Pool Utilization' of a negative value.

Suppose that 2 GB of data from a 5 GB file named **abc**, that is supposed to be in the **system** storage pool, are actually located in another pool. This 2 GB of data is said to be 'ill-placed'. Also, suppose that 3 GB of this file are in the **system** storage pool, and no other file is assigned to the **system** storage pool.

If you run the **mmapplypolicy** command to schedule file **abc** to be moved from the **system** storage pool to a storage pool named **YYY**, the **mmapplypolicy** command does the following:

1. Starts with the 'Current pool utilization' for the **system** storage pool, which is 3 GB.
2. Subtracts 5 GB, the size of file **abc**.
3. Arrives at a 'Predicted Pool Utilization' of negative 2 GB.

The **mmapplypolicy** command does not know how much of an 'ill-placed' file is currently in the wrong storage pool and how much is in the correct storage pool.

When there are ill-placed files in the **system** storage pool, the 'Predicted Pool Utilization' can be any positive or negative value. The positive value can be capped by the **LIMIT** clause of the **MIGRATE** rule. The 'Current Pool Utilizations' should always be between 0% and 100%.

## Policies - usage errors

Certain misunderstandings that may be encountered while using policies and the suggestions to overcome such mistakes.

The following are common mistakes and misunderstandings encountered when dealing with policies:

- You are advised to test your policy rules using the **mmapplypolicy** command with the **-I test** option. Also consider specifying a test-subdirectory within your file system. Do not apply a policy to an entire file system of vital files until you are confident that the rules correctly express your intentions. Even then, you are advised to do a sample run with the **mmapplypolicy -I test** command using the option **-L 3** or higher, to better understand which files are selected as candidates, and which candidates are chosen. The **-L** flag of the **mmapplypolicy** command can be used to check a policy before it is applied. For examples and more information on this flag, see "The **mmapplypolicy -L** command" on page 193.
- There is a 1 MB limit on the total size of the policy file installed in GPFS.
- Ensure that all clocks on all nodes of the GPFS cluster are synchronized. Depending on the policies in effect, variations in the clock times can cause unexpected behavior.

The **mmapplypolicy** command uses the time on the node on which it is run as the current time. Policy rules may refer to a file's last access time or modification time, which is set by the node which last accessed or modified the file. If the clocks are not synchronized, files may be treated as older or younger than their actual age, and this could cause files to be migrated or deleted prematurely, or not at all. A suggested solution is to use NTP to keep the clocks synchronized on all nodes in the cluster.

- The rules of a policy file are evaluated in order. A new file is assigned to the storage pool of the first rule that it matches. If the file fails to match any rule, the file creation fails with an **EINVAL** error code. A suggested solution is to put a **DEFAULT** clause as the last entry of the policy file.
- When a policy file is installed, GPFS verifies that the named storage pools exist. However, GPFS allows an administrator to delete pools that are mentioned in the policy file. This allows more freedom for recovery from hardware errors. Consequently, the administrator must be careful when deleting storage pools referenced in the policy.

1.

## Errors encountered with policies

The analysis of those errors which may be encountered while dealing with the policies.

These are errors encountered with policies and how to analyze them:

- Policy file never finishes, appears to be looping.

The **mmapplypolicy** command runs by making two passes over the file system - one over the inodes and one over the directory structure. The policy rules are applied to each file to determine a list of candidate files. The list is sorted by the weighting specified in the rules, then applied to the file system. No file is ever moved more than once. However, due to the quantity of data involved, this operation may take a long time and appear to be hung or looping.

The time required to run **mmapplypolicy** is a function of the number of files in the file system, the current load on the file system, and on the node in which **mmapplypolicy** is run. If this function appears to not finish, you may need to reduce the load on the file system or run **mmapplypolicy** on a less loaded node in the cluster.

- Initial file placement is not correct.

The placement rules specify a single pool for initial placement. The first rule that matches the file's attributes selects the initial pool. If that pool is incorrect, then the placement rules must be updated to select a different pool. You may see current placement rules by running **mmlspolicy -L**. For existing files, the file can be moved to its desired pool using the **mmrestripefile** or **mmchattr** commands.

For examples and more information on **mmlspolicy -L**, see "The mmapplypolicy -L command" on page 193.

- Data migration, deletion or exclusion not working properly.

The **mmapplypolicy** command selects a list of candidate files to be migrated or deleted. The list is sorted by the weighting factor specified in the rules, then applied to a sufficient number of files on the candidate list to achieve the utilization thresholds specified by the pools. The actual migration and deletion are done in parallel. The following are the possibilities for an incorrect operation:

- The file was not selected as a candidate for the expected rule. Each file is selected as a candidate for only the first rule that matched its attributes. If the matched rule specifies an invalid storage pool, the file is not moved. The **-L 4** option on **mmapplypolicy** displays the details for candidate selection and file exclusion.
- The file was a candidate, but was not operated on. Only the candidates necessary to achieve the desired pool utilizations are migrated. Using the **-L 3** option displays more information on candidate selection and files chosen for migration.

For more information on **mmlspolicy -L**, see "The mmapplypolicy -L command" on page 193.

- The file was scheduled for migration but was not moved. In this case, the file will be shown as 'ill-placed' by the **mmlsattr -L** command, indicating that the migration did not succeed. This occurs if the new storage pool assigned to the file did not have sufficient free space for the file when the actual migration was attempted. Since migrations are done in parallel, it is possible that the target pool had files which were also migrating, but had not yet been moved. If the target pool now has sufficient free space, the files can be moved using the commands: **mmrestripefs**, **mmrestripefile**, **mmchattr**.

- Asserts or error messages indicating a problem.

The policy rule language can only check for some errors at runtime. For example, a rule that causes a divide by zero cannot be checked when the policy file is installed. Errors of this type generate an error message and stop the policy evaluation for that file.

**Note:** I/O errors while migrating files indicate failing storage devices and must be addressed like any other I/O error. The same is true for any file system error or panic encountered while migrating files.



## Filesets - usage errors

The misunderstandings while dealing with the filesets and the course of actions to correct them.

These are common mistakes and misunderstandings encountered when dealing with filesets:

1. Fileset junctions look very much like ordinary directories, but they cannot be deleted by the usual commands such as **rm -r** or **rmdir**. Using these commands on a fileset junction could result in a Not owner message on an AIX system, or an Operation not permitted message on a Linux system.

As a consequence these commands may fail when applied to a directory that is a fileset junction. Similarly, when **rm -r** is applied to a directory that contains a fileset junction, it will fail as well.

On the other hand, **rm -r** will delete all the files contained in the filesets linked under the specified directory. Use the **mmunlinkfileset** command to remove fileset junctions.

2. Files and directories may not be moved from one fileset to another, nor may a hard link cross fileset boundaries.

If the user is unaware of the locations of fileset junctions, **mv** and **ln** commands may fail unexpectedly. In most cases, the **mv** command will automatically compensate for this failure and use a combination of **cp** and **rm** to accomplish the desired result. Use the **mmlsfileset** command to view the locations of fileset junctions. Use the **mmlsattr -L** command to determine the fileset for any given file.

3. Because a snapshot saves the contents of a fileset, deleting a fileset included in a snapshot cannot completely remove the fileset.

The fileset is put into a 'deleted' state and continues to appear in **mmlsfileset** output. Once the last snapshot containing the fileset is deleted, the fileset will be completely removed automatically. The **mmlsfileset --deleted** command indicates deleted filesets and shows their names in parentheses.

4. Deleting a large fileset may take some time and may be interrupted by other failures, such as disk errors or system crashes.

When this occurs, the recovery action leaves the fileset in a 'being deleted' state. Such a fileset may not be linked into the namespace. The corrective action is to finish the deletion by reissuing the fileset delete command:

```
mmdeletfileset fs1 fsname1 -f
```

The **mmlsfileset** command identifies filesets in this state by displaying a status of 'Deleting'.

5. If you unlink a fileset that has other filesets linked below it, any filesets linked to it (that is, child filesets) become inaccessible. The child filesets remain linked to the parent and will become accessible again when the parent is re-linked.

6. By default, the **mmdeletfileset** command will not delete a fileset that is not empty.

To empty a fileset, first unlink all its immediate child filesets, to remove their junctions from the fileset to be deleted. Then, while the fileset itself is still linked, use **rm -rf** or a similar command, to remove the rest of the contents of the fileset. Now the fileset may be unlinked and deleted.

Alternatively, the fileset to be deleted can be unlinked first and then **mmdeletfileset** can be used with the **-f** (force) option. This will unlink its child filesets, then destroy the files and directories contained in the fileset.

7. When deleting a small dependent fileset, it may be faster to use the **rm -rf** command instead of the **mmdeletfileset** command with the **-f** option.

## Errors encountered with filesets

The analysis of those errors which may be encountered while dealing with the filesets.

These are errors encountered with filesets and how to analyze them:

1. Problems can arise when running backup and archive utilities against a file system with unlinked filesets. See the *Filesets and backup* topic in the *IBM Spectrum Scale: Administration Guide* for details.

2. In the rare case that the **mmfsck** command encounters a serious error checking the file system's fileset metadata, it may not be possible to reconstruct the fileset name and comment. These cannot be inferred from information elsewhere in the file system. If this happens, **mmfsck** will create a dummy name for the fileset, such as 'Fileset911' and the comment will be set to the empty string.
3. Sometimes **mmfsck** encounters orphaned files or directories (those without a parent directory), and traditionally these are reattached in a special directory called 'lost+found' in the file system root. When a file system contains multiple filesets, however, orphaned files and directories are reattached in the 'lost+found' directory in the root of the fileset to which they belong. For the root fileset, this directory appears in the usual place, but other filesets may each have their own 'lost+found' directory.

## Active file management fileset errors

When the **mmafmctl Device getstate** command displays a NeedsResync target/fileset state, inconsistencies exist between the home and cache. To ensure that the cached data is synchronized with the home and the fileset is returned to Active state, either the file system must be unmounted and mounted or the fileset must be unlinked and linked. Once this is done, the next update to fileset data will trigger an automatic synchronization of data from the cache to the home.

## Storage pools - usage errors

The misunderstandings while dealing with the storage pools and the course of actions to correct them.

These are common mistakes and misunderstandings encountered when dealing with storage pools:

1. Only the **system** storage pool is allowed to store metadata. All other pools must have the **dataOnly** attribute.
2. Take care to create your storage pools with sufficient numbers of failure groups to enable the desired level of replication.

When the file system is created, GPFS requires all of the initial pools to have at least as many failure groups as defined by the default replication (**-m** and **-r** flags on the **mmcrfs** command). However, once the file system has been created, the user can create a storage pool with fewer failure groups than the default replication.

The **mmadddisk** command issues a warning, but it allows the disks to be added and the storage pool defined. To use the new pool, the user must define a policy rule to create or migrate files into the new pool. This rule should be defined to set an appropriate replication level for each file assigned to the pool. If the replication level exceeds the number of failure groups in the storage pool, all files assigned to the pool incur added overhead on each write to the file, in order to mark the file as ill-replicated.

To correct the problem, add additional disks to the storage pool, defining a different failure group, or insure that all policy rules that assign files to the pool also set the replication appropriately.

3. GPFS does not permit the **mmchdisk** or **mmrpldisk** command to change a disk's storage pool assignment. Changing the pool assignment requires all data residing on the disk to be moved to another disk before the disk can be reassigned. Moving the data is a costly and time-consuming operation; therefore GPFS requires an explicit **mmdeldisk** command to move it, rather than moving it as a side effect of another command.
4. Some storage pools allow larger disks to be added than do other storage pools.

When the file system is created, GPFS defines the maximum size disk that can be supported using the on-disk data structures to represent it. Likewise, when defining a new storage pool, the newly created on-disk structures establish a limit on the maximum size disk that can be added to that pool.

To add disks that exceed the maximum size allowed by a storage pool, simply create a new pool using the larger disks.

The **mmdf** command can be used to find the maximum disk size allowed for a storage pool.

5. If you try to delete a storage pool when there are files still assigned to the pool, consider this:

A storage pool is deleted when all disks assigned to the pool are deleted. To delete the last disk, all data residing in the pool must be moved to another pool. Likewise, any files assigned to the pool, whether or not they contain data, must be reassigned to another pool. The easiest method for reassigning all files and migrating all data is to use the **mmapplypolicy** command with a single rule to move all data from one pool to another. You should also install a new placement policy that does not assign new files to the old pool. Once all files have been migrated, reissue the **mmdeletedisk** command to delete the disk and the storage pool.

If all else fails, and you have a disk that has failed and cannot be recovered, follow the procedures in “Information to be collected before contacting the IBM Support Center” on page 367, and then contact the IBM Support Center for commands to allow the disk to be deleted without migrating all data from it. Files with data left on the failed device will lose data. If the entire pool is deleted, any existing files assigned to that pool are reassigned to a “broken” pool, which prevents writes to the file until the file is reassigned to a valid pool.

#### 6. Ill-placed files - understanding and correcting them.

The **mmapplypolicy** command migrates a file between pools by first assigning it to a new pool, then moving the file's data. Until the existing data is moved, the file is marked as 'ill-placed' to indicate that some of its data resides in its previous pool. In practice, **mmapplypolicy** assigns all files to be migrated to their new pools, then it migrates all of the data in parallel. Ill-placed files indicate that the **mmapplypolicy** or **mmchattr** command did not complete its last migration or that **-I defer** was used.

To correct the placement of the ill-placed files, the file data needs to be migrated to the assigned pools. You can use the **mmrestripefs**, or **mmrestripefile** commands to move the data.

#### 7. Using the **-P PoolName** option on the **mmrestripefs**, command:

This option restricts the restripe operation to a single storage pool. For example, after adding a disk to a pool, only the data in that pool needs to be restriped. In practice, **-P PoolName** simply restricts the operation to the files assigned to the specified pool. Files assigned to other pools are not included in the operation, even if the file is ill-placed and has data in the specified pool.

## Errors encountered with storage pools

The analysis of those errors which may be encountered while dealing with the storage pools.

These are error encountered with policies and how to analyze them:

#### 1. Access time to one pool appears slower than the others.

A consequence of striping data across the disks is that the I/O throughput is limited by the slowest device. A device encountering hardware errors or recovering from hardware errors may effectively limit the throughput to all devices. However using storage pools, striping is done only across the disks assigned to the pool. Thus a slow disk impacts only its own pool; all other pools are not impeded.

To correct the problem, check the connectivity and error logs for all disks in the slow pool.

#### 2. Other storage pool problems might really be disk problems and should be pursued from the standpoint of making sure that your disks are properly configured and operational. See Chapter 16, “Disk issues,” on page 275.

---

## Snapshot problems

Use the **mmlsnapshot** command as a general hint for snapshot-related problems, to find out what snapshots exist, and what state they are in. Use the **mmsnapdir** command to find the snapshot directory name used to permit access.

The **mmlsnapshot** command displays the list of **all** snapshots of a file system. This command lists the snapshot name, some attributes of the snapshot, as well as the snapshot's status. The **mmlsnapshot** command does not require the file system to be mounted.

## Problems with locating a snapshot

Use the `mmlsnapshot` command and `mmsnapdir` command to find the snapshot detail and locate them.

The `mmlsnapshot` and `mmsnapdir` commands are provided to assist in locating the snapshots in the file system directory structure. Only valid snapshots are visible in the file system directory structure. They appear in a hidden subdirectory of the file system's root directory. By default the subdirectory is named `.snapshots`. The valid snapshots appear as entries in the snapshot directory and may be traversed like any other directory. The `mmsnapdir` command can be used to display the assigned snapshot directory name.

## Problems not directly related to snapshots

There are errors which are returned from the snapshot commands but are not linked with the snapshots directly.

Many errors returned from the snapshot commands are not specifically related to the snapshot. For example, disk failures or node failures could cause a snapshot command to fail. The response to these types of errors is to fix the underlying problem and try the snapshot command again.

### GPFS error messages for indirect snapshot errors

There are GPFS error messages which may be associated with snapshots directly but does not show a clear relation to snapshot issues.

The error messages for this type of problem do not have message numbers, but can be recognized by their message text:

- 'Unable to sync all nodes, rc=*errorCode*.'
- 'Unable to get permission to create snapshot, rc=*errorCode*.'
- 'Unable to quiesce all nodes, rc=*errorCode*.'
- 'Unable to resume all nodes, rc=*errorCode*.'
- 'Unable to delete snapshot *filesystemName* from file system *snapshotName*, rc=*errorCode*.'
- 'Error restoring inode *number*, error *errorCode*.'
- 'Error deleting snapshot *snapshotName* in file system *filesystemName*, error *errorCode*.'
- '*commandString* failed, error *errorCode*.'
- 'None of the nodes in the cluster is reachable, or GPFS is down on all of the nodes.'
- 'File system *filesystemName* is not known to the GPFS cluster.'

## Snapshot usage errors

Certain error in the GPFS error messages are related to the snapshot usage restrictions or incorrect snapshot names .

Many errors returned from the snapshot commands are related to usage restrictions or incorrect snapshot names.

An example of a snapshot restriction error is exceeding the maximum number of snapshots allowed at one time. For simple errors of these types, you can determine the source of the error by reading the error message or by reading the description of the command. You can also run the `mmlsnapshot` command to see the complete list of existing snapshots.

Examples of incorrect snapshot name errors are trying to delete a snapshot that does not exist or trying to create a snapshot using the same name as an existing snapshot. The rules for naming global and filesset snapshots are designed to minimize conflicts between the file system administrator and the filesset owners. These rules can result in errors when filesset snapshot names are duplicated across different filesets or when the snapshot command `-j` option (specifying a qualifying filesset name) is provided or

omitted incorrectly. To resolve name problems review the **mmlssnapshot** output with careful attention to the Fileset column. You can also specify the **-s** or **-j** options of the **mmlssnapshot** command to limit the output. For snapshot deletion, the **-j** option must exactly match the Fileset column.

For more information about snapshot naming conventions, see the **mmcrsnapshot** command in the *IBM Spectrum Scale: Command and Programming Reference*.

## GPFS error messages for snapshot usage errors

Certain error messages for snapshot usage errors have no error message numbers but may be recognized using the message texts.

The error messages for this type of problem do not have message numbers, but can be recognized by their message text:

- 'File system *filesystemName* does not contain a snapshot *snapshotName*, rc=*errorCode*.'
- 'Cannot create a new snapshot until an existing one is deleted. File system *filesystemName* has a limit of *number* online snapshots.'
- 'Cannot restore snapshot. *snapshotName* is mounted on *number* nodes and in use on *number* nodes.'
- 'Cannot create a snapshot in a DM enabled file system, rc=*errorCode*.'

## Snapshot status errors

There are certain snapshot commands like **mmdelsnapshot** and **mmrestorefs**, which lets snapshot go invalid if they got interrupted while running.

Some snapshot commands like **mmdelsnapshot** and **mmrestorefs** may require a substantial amount of time to complete. If the command is interrupted, say by the user or due to a failure, the snapshot may be left in an invalid state. In many cases, the command must be completed before other snapshot commands are allowed to run. The source of the error may be determined from the error message, the command description, or the snapshot status available from **mmlssnapshot**.

## GPFS error messages for snapshot status errors

Certain error messages for snapshot status error have no error message numbers and may be recognized by the message texts only.

The error messages for this type of problem do not have message numbers, but can be recognized by their message text:

- 'Cannot delete snapshot *snapshotName* which is *snapshotState*, error = *errorCode*.'
- 'Cannot restore snapshot *snapshotName* which is *snapshotState*, error = *errorCode*.'
- 'Previous snapshot *snapshotName* is invalid and must be deleted before a new snapshot may be created.'
- 'Previous snapshot *snapshotName* must be restored before a new snapshot may be created.'
- 'Previous snapshot *snapshotName* is invalid and must be deleted before another snapshot may be deleted.'
- 'Previous snapshot *snapshotName* is invalid and must be deleted before another snapshot may be restored.'
- 'More than one snapshot is marked for restore.'
- 'Offline snapshot being restored.'

## Snapshot directory name conflicts

The snapshot generated by **mmcrsnapshot** command may not be accessed due to directory conflicts and the course of action to correct the snapshot directory name conflict.

By default, all snapshots appear in a directory named **.snapshots** in the root directory of the file system. This directory is dynamically generated when the first snapshot is created and continues to exist even

after the last snapshot is deleted. If the user tries to create the first snapshot, and a normal file or directory named `.snapshots` already exists, the `mmcrsnapshot` command will be successful but the snapshot may not be accessed.

There are two ways to fix this problem:

1. Delete or rename the existing file or directory
2. Tell GPFS to use a different name for the dynamically-generated directory of snapshots by running the `mmsnapdir` command.

It is also possible to get a name conflict as a result of issuing the `mmrestorefs` command. Since `mmsnapdir` allows changing the name of the dynamically-generated snapshot directory, it is possible that an older snapshot contains a normal file or directory that conflicts with the current name of the snapshot directory. When this older snapshot is restored, the `mmrestorefs` command will recreate the old, normal file or directory in the file system root directory. The `mmrestorefs` command will not fail in this case, but the restored file or directory will hide the existing snapshots. After invoking `mmrestorefs` it may therefore appear as if the existing snapshots have disappeared. However, `mmlssnapshot` should still show all existing snapshots.

The fix is the similar to the one mentioned before. Perform one of these two steps:

1. After the `mmrestorefs` command completes, rename the conflicting file or directory that was restored in the root directory.
2. Run the `mmsnapdir` command to select a different name for the dynamically-generated snapshot directory.

Finally, the `mmsnapdir -a` option enables a dynamically-generated snapshot directory in every directory, not just the file system root. This allows each user quick access to snapshots of their own files by going into `.snapshots` in their home directory or any other of their directories.

Unlike `.snapshots` in the file system root, `.snapshots` in other directories is invisible, that is, an `ls -a` command will not list `.snapshots`. This is intentional because recursive file system utilities such as `find`, `du` or `ls -R` would otherwise either fail or produce incorrect or undesirable results. To access snapshots, the user must explicitly specify the name of the snapshot directory, for example: `ls ~/.snapshots`. If there is a name conflict (that is, a normal file or directory named `.snapshots` already exists in the user's home directory), the user must rename the existing file or directory.

The inode numbers that are used for and within these special `.snapshots` directories are constructed dynamically and do not follow the standard rules. These inode numbers are visible to applications through standard commands, such as `stat`, `readdir`, or `ls`. The inode numbers reported for these directories can also be reported differently on different operating systems. Applications should not expect consistent numbering for such inodes.

## Errors encountered when restoring a snapshot

There are errors which may be displayed while restoring a snapshot.

The following errors might be encountered when restoring from a snapshot:

- The `mmrestorefs` command fails with an `ENOSPC` message. In this case, there are not enough free blocks in the file system to restore the selected snapshot. You can add space to the file system by adding a new disk. As an alternative, you can delete a different snapshot from the file system to free some existing space. You cannot delete the snapshot that is being restored. After there is additional free space, issue the `mmrestorefs` command again.
- The `mmrestorefs` command fails with quota exceeded errors. Try adjusting the quota configuration or disabling quota, and then issue the command again.
- The `mmrestorefs` command is interrupted and some user data is not be restored completely. Try repeating the `mmrestorefs` command in this instance.

- The **mmrestorefs** command fails because of an incorrect file system, fileset, or snapshot name. To fix this error, issue the command again with the correct name.
- The **mmrestorefs -j** command fails with the following error:

6027-953

Failed to get a handle for fileset *filesetName*, snapshot *snapshotName* in file system *fileSystem*.  
*errorMessage*.

In this case, the file system that contains the snapshot to restore should be mounted, and then the fileset of the snapshot should be linked.

If you encounter additional errors that cannot be resolved, contact the IBM Support Center.

## Errors encountered when restoring a snapshot

There are errors which may be displayed while restoring a snapshot.

The following errors might be encountered when restoring from a snapshot:

- The **mmrestorefs** command fails with an **ENOSPC** message. In this case, there are not enough free blocks in the file system to restore the selected snapshot. You can add space to the file system by adding a new disk. As an alternative, you can delete a different snapshot from the file system to free some existing space. You cannot delete the snapshot that is being restored. After there is additional free space, issue the **mmrestorefs** command again.
- The **mmrestorefs** command fails with quota exceeded errors. Try adjusting the quota configuration or disabling quota, and then issue the command again.
- The **mmrestorefs** command is interrupted and some user data is not be restored completely. Try repeating the **mmrestorefs** command in this instance.
- The **mmrestorefs** command fails because of an incorrect file system, fileset, or snapshot name. To fix this error, issue the command again with the correct name.
- The **mmrestorefs -j** command fails with the following error:

6027-953

Failed to get a handle for fileset *filesetName*, snapshot *snapshotName* in file system *fileSystem*.  
*errorMessage*.

In this case, the file system that contains the snapshot to restore should be mounted, and then the fileset of the snapshot should be linked.

If you encounter additional errors that cannot be resolved, contact the IBM Support Center.

## Failures using the mmpmon command

The **mmpmon** command manages performance monitoring and displays performance information.

The **mmpmon** command is thoroughly documented in “Monitoring GPFS I/O performance with the mmpmon command” on page 3 and the *mmpmon command* page in the *IBM Spectrum Scale: Command and Programming Reference*. Before proceeding with **mmpmon** problem determination, review all of this material to ensure that you are using the **mmpmon** command correctly.

### Setup problems using mmpmon

The issues associated with the set up of **mmpmon** command and limitations of this command.

Remember these points when using the **mmpmon** command:

- You must have root authority.
- The GPFS daemon must be active.
- The input file must contain valid input requests, one per line. When an incorrect request is detected by **mmpmon**, it issues an error message and terminates.

- Input requests that appear in the input file before the first incorrect request are processed by **mmpmon**.
- Do not alter the input file while **mmpmon** is running.
  - Output from **mmpmon** is sent to standard output (STDOUT) and errors are sent to standard (STDERR).
  - Up to five instances of **mmpmon** may run on a given node concurrently. See “Monitoring GPFS I/O performance with the **mmpmon** command” on page 3. For the limitations regarding concurrent usage of **mmpmon**, see “Running **mmpmon** concurrently from multiple users on the same node” on page 4.
  - The **mmpmon** command does **not** support:
    - Monitoring read requests without monitoring writes, or the other way around.
    - Choosing which file systems to monitor.
    - Monitoring on a per-disk basis.
    - Specifying different size or latency ranges for reads and writes.
    - Specifying different latency values for a given size range.

### Incorrect output from **mmpmon**

The analysis of incorrect output of **mmpmon** command.

If the output from **mmpmon** is incorrect, such as zero counters when you know that I/O activity is taking place, consider these points:

1. Someone may have issued the `reset` or `rhist reset` requests.
2. Counters may have wrapped due to a large amount of I/O activity, or running **mmpmon** for an extended period of time. For a discussion of counter sizes and counter wrapping, see *Counter sizes and counter wrapping* section in “Monitoring GPFS I/O performance with the **mmpmon** command” on page 3.
3. See the *Other information about **mmpmon** output* section in “Monitoring GPFS I/O performance with the **mmpmon** command” on page 3. This section gives specific instances where **mmpmon** output may be different than what was expected.

### Abnormal termination or hang in **mmpmon**

The course of action which must be followed if **mmpmon** command hangs or terminates.

If **mmpmon** hangs, perform these steps:

1. Ensure that sufficient time has elapsed to cover the **mmpmon** timeout value. It is controlled using the `-t` flag on the **mmpmon** command.
2. Issue the `ps` command to find the PID for **mmpmon**.
3. Issue the `kill` command to terminate this PID.
4. Try the function again.
5. If the problem persists, issue this command:  
`mmfsadm dump eventsExporter`
6. Copy the output of **mmfsadm** to a safe location.
7. Follow the procedures in “Information to be collected before contacting the IBM Support Center” on page 367, and then contact the IBM Support Center.

If **mmpmon** terminates abnormally, perform these steps:

1. Determine if the GPFS daemon has failed, and if so restart it.
2. Review your invocation of **mmpmon**, and verify the input.
3. Try the function again.
4. If the problem persists, follow the procedures in “Information to be collected before contacting the IBM Support Center” on page 367, and then contact the IBM Support Center.



## Tracing the mmpmon command

The course of action to be followed if the **mmpmon** command does not perform as expected.

When the **mmpmon** command does not work properly, there are two trace classes used to determine the cause of the problem. Use these only when requested by the IBM Support Center.

### eventsExporter

Reports attempts to connect and whether or not they were successful.

### mmpmon

Shows the command string that came in to the **mmpmon** command, and whether it was successful or not.

**Note:** Do not use the **perfmon** trace class of the GPFS trace to diagnose **mmpmon** problems. This trace event does not provide the necessary data.

---

## Failures using the mmbackup command

Use the **mmbackup** command to back up the files in a GPFS file system to storage on a IBM Spectrum Protect™ server. A number of factors can cause **mmbackup** to fail.

The most common of these are:

- The file system is not mounted on the node issuing the **mmbackup** command.
- The file system is not mounted on the IBM Spectrum Protect client nodes.
- The **mmbackup** command was issued to back up a file system owned by a remote cluster.
- The IBM Spectrum Protect clients are not able to communicate with the IBM Spectrum Protect server due to authorization problems.
- The IBM Spectrum Protect server is down or out of storage space.
- When the target of the backup is tape, the IBM Spectrum Protect server may be unable to handle all of the backup client processes because the value of the IBM Spectrum Protect server's MAXNUMMP parameter is set lower than the number of client processes. This failure is indicated by message ANS1312E from IBM Spectrum Protect.

The errors from **mmbackup** normally indicate the underlying problem.

## GPFS error messages for mmbackup errors

Error messages that are displayed for mmbackup errors

### 6027-1995

Device *deviceName* is not mounted on node *nodeName*.

## IBM Spectrum Protect error messages

Error message displayed for server media mount.

### ANS1312E

Server media mount not possible.

---

## Data integrity

GPFS takes extraordinary care to maintain the integrity of customer data. However, certain hardware failures, or in extremely unusual circumstances, the occurrence of a programming error can cause the loss of data in a file system.

GPFS performs extensive checking to validate metadata and ceases using the file system if metadata becomes inconsistent. This can appear in two ways:

1. The file system will be unmounted and applications will begin seeing ESTALE return codes to file operations.
2. Error log entries indicating an **MMFS\_SYSTEM\_UNMOUNT** and a corruption error are generated.

If actual disk data corruption occurs, this error will appear on each node in succession. Before proceeding with the following steps, follow the procedures in “Information to be collected before contacting the IBM Support Center” on page 367, and then contact the IBM Support Center.

1. Examine the error logs on the NSD servers for any indication of a disk error that has been reported.
2. Take appropriate disk problem determination and repair actions prior to continuing.
3. After completing any required disk repair actions, run the offline version of the **mmfsck** command on the file system.
4. If your error log or disk analysis tool indicates that specific disk blocks are in error, use the **mmfileid** command to determine which files are located on damaged areas of the disk, and then restore these files. See “The mmfileid command” on page 201 for more information.
5. If data corruption errors occur in only one node, it is probable that memory structures within the node have been corrupted. In this case, the file system is probably good but a program error exists in GPFS or another authorized program with access to GPFS data structures.

Follow the directions in “Data integrity” on page 273 and then reboot the node. This should clear the problem. If the problem repeats on one node without affecting other nodes check the programming specifications code levels to determine that they are current and compatible and that no hardware errors were reported. Refer to the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide* for correct software levels.

## Error numbers specific to GPFS application calls when data integrity may be corrupted

If there is a possibility of the corruption of data integrity, GPFS displays specific error messages or returns them to the application.

When there is the possibility of data corruption, GPFS may report these error numbers in the operating system error log, or return them to an application:

### **EVALIDATE=214, Invalid checksum or other consistency check failure on disk data structure.**

This indicates that internal checking has found an error in a metadata structure. The severity of the error depends on which data structure is involved. The cause of this is usually GPFS software, disk hardware or other software between GPFS and the disk. Running **mmfsck** should repair the error. The urgency of this depends on whether the error prevents access to some file or whether basic metadata structures are involved.

---

## Messages requeuing in AFM

The course of actions to be followed for resolving the requeued messages in the Gateway node

Sometimes requests in the AFM messages queue on the gateway node get requested because of errors at home. For example, if there is no space at home to perform a new write, a write message that is queued is not successful and gets requeued. The administrator would see the failed message getting requeued in the queue on the gateway node. The administrator has to resolve the issue by adding more space at home and running the **mmafmctl resumeRequeued** command, so that the requeued messages are executed at home again. If **mmafmctl resumeRequeued** is not run by an administrator, AFM would still execute the message in the regular order of message executions from cache to home.

Running the **mmfsadm dump afm all** command on the gateway node shows the queued messages. Requeued messages show in the dumps similar to the following example:

```
c12c4apv13.gpfs.net: Normal Queue: (listed by execution order) (state: Active)
c12c4apv13.gpfs.net: Write [612457.552962] requeued file3 (43 @ 293) chunks 0 bytes 0 0
```

---

## Chapter 16. Disk issues

GPFS uses only disk devices prepared as Network Shared Disks (NSDs). However NSDs might exist on top of a number of underlying disk technologies.

NSDs, for example, might be defined on top of Fibre Channel SAN connected disks. This information provides detail on the creation, use, and failure of NSDs and their underlying disk technologies.

These are some of the errors encountered with GPFS disks and NSDs:

- “NSD and underlying disk subsystem failures”
- “GPFS has declared NSDs built on top of AIX logical volumes as down” on page 282
- “Disk accessing commands fail to complete due to problems with some non-IBM disks” on page 283
- “Persistent Reserve errors” on page 287
- “GPFS is not using the underlying multipath device” on page 290

---

### NSD and underlying disk subsystem failures

There are indications that will lead you to the conclusion that your file system has disk failures.

Some of those indications include:

- Your file system has been forced to unmount. For more information about forced file system unmount, see “File system forced unmount” on page 250.
- The **mmismount** command indicates that the file system is not mounted on certain nodes.
- Your application is getting EIO errors.
- Operating system error logs indicate you have stopped using a disk in a replicated system, but your replication continues to operate.
- The **mmisdisk** command shows that disks are down.

**Note:** If you are reinstalling the operating system on one node and erasing all partitions from the system, GPFS descriptors will be removed from any NSD this node can access locally. The results of this action might require recreating the file system and restoring from backup. If you experience this problem, do not unmount the file system on any node that is currently mounting the file system. Contact the IBM Support Center immediately to see if the problem can be corrected.

### Error encountered while creating and using NSD disks

Use **mmcrnsd** command to prepare NSD disks. While preparing the NSD disks, there are several errors conditions encountered.

GPFS requires that disk devices be prepared as NSDs. This is done using the **mmcrnsd** command. The input to the **mmcrnsd** command is given in the form of disk stanzas. For a complete explanation of disk stanzas, see the *Stanza files* section in the *IBM Spectrum Scale: Administration Guide*, and the following topics from the *IBM Spectrum Scale: Command and Programming Reference*:

- **mmcdisk** command
- **mmchnsd** command
- **mmcrfs** command
- **mmcrnsd** command

For disks that are SAN-attached to all nodes in the cluster, **device=DiskName** should refer to the disk device name in **/dev** on the node where the **mmcrnsd** command is issued. If a server list is specified, **device=DiskName** must refer to the name of the disk on the first server node. The same disk can have different local names on different nodes.

When you specify an NSD server node, that node performs all disk I/O operations on behalf of nodes in the cluster that do not have connectivity to the disk. You can also specify up to eight additional NSD server nodes. These additional NSD servers will become active if the first NSD server node fails or is unavailable.

When the **mmcrnsd** command encounters an error condition, one of these messages is displayed:

**6027-2108**

Error found while processing stanza

or

**6027-1636**

Error found while checking disk descriptor *descriptor*

Usually, this message is preceded by one or more messages describing the error more specifically.

Another possible error from **mmcrnsd** is:

**6027-2109**

Failed while processing disk stanza on node *nodeName*.

or

**6027-1661**

Failed while processing disk descriptor *descriptor* on node *nodeName*.

One of these errors can occur if an NSD server node does not have read and write access to the disk. The NSD server node needs to write an NSD volume ID to the raw disk. If an additional NSD server node is specified, that NSD server node will scan its disks to find this NSD volume ID string. If the disk is SAN-attached to all nodes in the cluster, the NSD volume ID is written to the disk by the node on which the **mmcrnsd** command is running.

## Displaying NSD information

Use **mmlnsd** command to display the NSD information and analyze the cluster details pertaining to NSDs.

Use the **mmlnsd** command to display information about the currently defined NSDs in the cluster. For example, if you issue **mmlnsd**, your output may be similar to this:

File system	Disk name	NSD servers
fs1	t65nsd4b	(directly attached)
fs5	t65nsd12b	c26f4gp01.ppd.pok.ibm.com,c26f4gp02.ppd.pok.ibm.com
fs6	t65nsd13b	c26f4gp01.ppd.pok.ibm.com,c26f4gp02.ppd.pok.ibm.com,c26f4gp03.ppd.pok.ibm.com

This output shows that:

- There are three NSDs in this cluster: **t65nsd4b**, **t65nsd12b**, and **t65nsd13b**.
- NSD disk **t65nsd4b** of file system **fs1** is SAN-attached to all nodes in the cluster.
- NSD disk **t65nsd12b** of file system **fs5** has 2 NSD server nodes.
- NSD disk **t65nsd13b** of file system **fs6** has 3 NSD server nodes.

If you need to find out the local device names for these disks, you could use the **-m** option on the **mmlsnsd** command. For example, issuing:

```
mmlsnsd -m
```

produces output similar to this example:

Disk name	NSD volume ID	Device	Node name	Remarks
t65nsd12b	0972364D45EF7B78	/dev/hdisk34	c26f4gp01.ppd.pok.ibm.com	server node
t65nsd12b	0972364D45EF7B78	/dev/hdisk34	c26f4gp02.ppd.pok.ibm.com	server node
t65nsd12b	0972364D45EF7B78	/dev/hdisk34	c26f4gp04.ppd.pok.ibm.com	
t65nsd13b	0972364D000000001	/dev/hdisk35	c26f4gp01.ppd.pok.ibm.com	server node
t65nsd13b	0972364D000000001	/dev/hdisk35	c26f4gp02.ppd.pok.ibm.com	server node
t65nsd13b	0972364D000000001	-	c26f4gp03.ppd.pok.ibm.com	(not found) server node
t65nsd4b	0972364D45EF7614	/dev/hdisk26	c26f4gp04.ppd.pok.ibm.com	

From this output we can tell that:

- The local disk name for **t65nsd12b** on NSD server **c26f4gp01** is **hdisk34**.
- NSD disk **t65nsd13b** is not attached to node on which the **mmlsnsd** command was issued, node **c26f4gp04**.
- The **mmlsnsd** command was not able to determine the local device for NSD disk **t65nsd13b** on **c26f4gp03** server.

To find the nodes to which disk **t65nsd4b** is attached and the corresponding local devices for that disk, issue:

```
mmlsnsd -d t65nsd4b -M
```

Output is similar to this example:

Disk name	NSD volume ID	Device	Node name	Remarks
t65nsd4b	0972364D45EF7614	/dev/hdisk92	c26f4gp01.ppd.pok.ibm.com	
t65nsd4b	0972364D45EF7614	/dev/hdisk92	c26f4gp02.ppd.pok.ibm.com	
t65nsd4b	0972364D45EF7614	-	c26f4gp03.ppd.pok.ibm.com	(not found) directly attached
t65nsd4b	0972364D45EF7614	/dev/hdisk26	c26f4gp04.ppd.pok.ibm.com	

From this output we can tell that NSD **t65nsd4b** is:

- Known as **hdisk92** on node **c26f4gp01** and **c26f4gp02**.
- Known as **hdisk26** on node **c26f4gp04**
- Is not attached to node **c26f4gp03**

To display extended information about a node's view of its NSDs, the **mmlsnsd -X** command can be used:

```
mmlsnsd -X -d "hd3n97;sdfnsd;hd5n98"
```

The system displays information similar to:

Disk name	NSD volume ID	Device	Devtype	Node name	Remarks
hd3n97	0972846145C8E927	/dev/hdisk3	hdisk	c5n97g.ppd.pok.ibm.com	server node,pr=no
hd3n97	0972846145C8E927	/dev/hdisk3	hdisk	c5n98g.ppd.pok.ibm.com	server node,pr=no
hd5n98	0972846245EB501C	/dev/hdisk5	hdisk	c5n97g.ppd.pok.ibm.com	server node,pr=no
hd5n98	0972846245EB501C	/dev/hdisk5	hdisk	c5n98g.ppd.pok.ibm.com	server node,pr=no
sdfnsd	0972845E45F02E81	/dev/sdf	generic	c5n94g.ppd.pok.ibm.com	server node
sdfnsd	0972845E45F02E81	/dev/sdm	generic	c5n96g.ppd.pok.ibm.com	server node

From this output we can tell that:

- Disk **hd3n97** is an hdisk known as **/dev/hdisk3** on NSD server node **c5n97** and **c5n98**.

- Disk **sdfnsd** is a generic disk known as **/dev/sdf** and **/dev/sdm** on NSD server node **c5n94g** and **c5n96g**, respectively.
- In addition to the preceding information, the NSD volume ID is displayed for each disk.

**Note:** The **-m**, **-M** and **-X** options of the **mmlnsd** command can be very time consuming, especially on large clusters. Use these options judiciously.

## NSD creation fails with a message referring to an existing NSD

The **mmdelnsd** command failure, error messages and the recovery steps.

NSDs are deleted with the **mmdelnsd** command. Internally, this is a two-step process:

1. Remove the NSD definitions from the GPFS control information.
2. Zero-out GPFS-specific data structures on the disk.

If for some reason the second step fails, for example because the disk is damaged and cannot be written to, the **mmdelnsd** command issues a message describing the error and then another message stating the exact command to issue to complete the deletion of the NSD. If these instructions are not successfully completed, a subsequent **mmcrnsd** command can fail with

**6027-1662**

Disk device *deviceName* refers to an existing NSD *name*.

This error message indicates that the disk is either an existing NSD, or that the disk was previously an NSD that had been removed from the GPFS cluster using the **mmdelnsd -p** command, and had not been marked as available.

If the GPFS data structures are not removed from the disk, it might be unusable for other purposes. For example, if you are trying to create an AIX volume group on the disk, the **mkvg** command might fail with messages similar to:

```
0516-1339 /usr/sbin/mkvg: Physical volume contains some 3rd party volume group.
0516-1397 /usr/sbin/mkvg: The physical volume hdisk5, will not be added to the volume group.
0516-862 /usr/sbin/mkvg: Unable to create volume group.
```

The easiest way to recover such a disk is to temporarily define it as an NSD again (using the **-v no** option) and then delete the just-created NSD. For example:

```
mmcrnsd -F filename -v no
mmdelnsd -F filename
```

## GPFS has declared NSDs as down

GPFS reactions to NSD failures and the recovery procedure.

There are several situations in which disks can appear to fail to GPFS. Almost all of these situations involve a failure of the underlying disk subsystem. The following information describes how GPFS reacts to these failures and how to find the cause.

GPFS will stop using a disk that is determined to have failed. This event is marked as **MMFS\_DISKFAIL E!** in an error log entry (see “Operating system error logs” on page 157). The state of a disk can be checked by issuing the **mmlsdisk** command.

The consequences of stopping disk usage depend on what is stored on the disk:

- Certain data blocks may be unavailable because the data residing on a stopped disk is not replicated.
- Certain data blocks may be unavailable because the controlling metadata resides on a stopped disk.
- In conjunction with other disks that have failed, all copies of critical data structures may be unavailable resulting in the unavailability of the entire file system.

The disk will remain unavailable until its status is explicitly changed through the **mmchdisk** command. After that command is issued, any replicas that exist on the failed disk are updated before the disk is used.

GPFS can declare disks **down** for a number of reasons:

- If the first NSD server goes down and additional NSD servers were not assigned, or all of the additional NSD servers are also down and no local device access is available on the node, the disks are marked as stopped.
- A failure of an underlying disk subsystem may result in a similar marking of disks as stopped.
  1. Issue the **mmlsdisk** command to verify the status of the disks in the file system.
  2. Issue the **mmchdisk** command with the **-a** option to start all stopped disks.
- Disk failures should be accompanied by error log entries (see The operating system error log facility) for the failing disk. GPFS error log entries labelled **MMFS\_DISKFAIL** will occur on the node detecting the error. This error log entry will contain the identifier of the failed disk. Follow the problem determination and repair actions specified in your disk vendor problem determination guide. After performing problem determination and repair issue the **mmchdisk** command to bring the disk back up.

## Unable to access disks

Accessing disk may be restricted due to incorrect disk specification or configuration failure during disk subsystem initialization.

If you cannot open a disk, the specification of the disk may be incorrect. It is also possible that a configuration failure may have occurred during disk subsystem initialization. For example, on Linux you should consult **/var/log/messages** to determine if disk device configuration errors have occurred.

```
Feb 16 13:11:18 host123 kernel: SCSI device sdu: 35466240 512-byte hdwr sectors (18159 MB)
Feb 16 13:11:18 host123 kernel: sdu: I/O error: dev 41:40, sector 0
Feb 16 13:11:18 host123 kernel: unable to read partition table
```

E! On AIX, consult “Operating system error logs” on page 157 for hardware configuration error log entries.

Accessible disk devices will generate error log entries similar to this example for a SSA device:

```
-----
LABEL:          SSA_DEVICE_ERROR
IDENTIFIER:     FE9E9357

Date/Time:      Wed Sep  8 10:28:13 edt
Sequence Number: 54638
Machine Id:     000203334C00
Node Id:       c154n09
Class:         H
Type:         PERM
Resource Name: pdisk23
Resource Class: pdisk
Resource Type: scsd
Location:      USSA4B33-D3
VPD:
  Manufacturer.....IBM
  Machine Type and Model.....DRVC18B
  Part Number.....09L1813
  ROS Level and ID.....0022
  Serial Number.....6800D2A6HK
  EC Level.....E32032
  Device Specific.(Z2).....CUSHA022
  Device Specific.(Z3).....09L1813
  Device Specific.(Z4).....99168
```

```
Description
DISK OPERATION ERROR
```

Probable Causes  
DASD DEVICE

Failure Causes  
DISK DRIVE

Recommended Actions  
PERFORM PROBLEM DETERMINATION PROCEDURES

Detail Data  
ERROR CODE  
2310 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000

-----  
or this one from GPFS:

-----  
LABEL: MMFS\_DISKFAIL  
IDENTIFIER: 9C6C05FA  
  
Date/Time: Tue Aug 3 11:26:34 edt  
Sequence Number: 55062  
Machine Id: 000196364C00  
Node Id: c154n01  
Class: H  
Type: PERM  
Resource Name: mmfs  
Resource Class: NONE  
Resource Type: NONE  
Location:

Description  
DISK FAILURE

Probable Causes  
STORAGE SUBSYSTEM  
DISK

Failure Causes  
STORAGE SUBSYSTEM  
DISK

Recommended Actions  
CHECK POWER  
RUN DIAGNOSTICS AGAINST THE FAILING DEVICE

Detail Data  
EVENT CODE  
1027755  
VOLUME  
fs3  
RETURN CODE  
19  
PHYSICAL VOLUME  
vp31n05  
-----

## Guarding against disk failures

Protection methods to guard against data loss due to disk media failure.

There are various ways to guard against the loss of data due to disk media failures. For example, the use of a RAID controller, which masks disk failures with parity disks, or a twin-tailed disk, could prevent the need for using these recovery steps.



GPFS offers a method of protection called *replication*, which overcomes disk failure at the expense of additional disk space. GPFS allows replication of data and metadata. This means that three instances of data, metadata, or both can be automatically created and maintained for any file in a GPFS file system. If one instance becomes unavailable due to disk failure, another instance is used instead. You can set different replication specifications for each file, or apply default settings specified at file system creation. Refer to the *File system replication parameters* topic in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.

## Disk connectivity failure and recovery

The GPFS has certain error messages defined for local connection failure from NSD servers.

If a disk is defined to have a local connection and to be connected to defined NSD servers, and the local connection fails, GPFS bypasses the broken local connection and uses the NSD servers to maintain disk access. The following error message appears in the GPFS log:

### 6027-361 [E]

Local access to *disk* failed with EIO, switching to access the disk remotely.

This is the default behavior, and can be changed with the `useNSDserver` file system mount option. See the *NSD server considerations* topic in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.

For a file system using the default mount option `useNSDserver=asneeded`, disk access fails over from local access to remote NSD access. Once local access is restored, GPFS detects this fact and switches back to local access. The detection and switch over are not instantaneous, but occur at approximately five minute intervals.

**Note:** In general, after fixing the path to a disk, you must run the `mmnsddiscover` command on the server that lost the path to the NSD. (Until the `mmnsddiscover` command is run, the reconnected node will see its local disks and start using them by itself, but it will not act as the NSD server.)

After that, you must run the command on all client nodes that need to access the NSD on that server; or you can achieve the same effect with a single `mmnsddiscover` invocation if you utilize the `-N` option to specify a node list that contains all the NSD servers and clients that need to rediscover paths.

## Partial disk failure

Partial disk failures when you have chosen not to implement hardware protection against media failures and the course of action to correct this problem.

If the disk has only partially failed and you have chosen not to implement hardware protection against media failures, the steps to restore your data depends on whether you have used replication. If you have replicated neither your data nor metadata, you will need to issue the offline version of the `mmfsck` command, and then restore the lost information from the backup media. If it is just the data which was not replicated, you will need to restore the data from the backup media. There is no need to run the `mmfsck` command if the metadata is intact.

If both your data and metadata have been replicated, implement these recovery actions:

1. Unmount the file system:  
`mmumount fs1 -a`
2. Delete the disk from the file system:  
`mmdisk fs1 gpfs10nsd -c`
3. If you are replacing the disk, add the new disk to the file system:  
`mmadddisk fs1 gpfs11nsd`
4. Then restripe the file system:  
`mmrestripefs fs1 -b`

**Note:** Ensure there is sufficient space elsewhere in your file system for the data to be stored by using the **mmdf** command.

---

## GPFS has declared NSDs built on top of AIX logical volumes as down

Earlier releases of GPFS allowed AIX logical volumes to be used in GPFS file systems. Using AIX logical volumes in GPFS file systems is now discouraged as they are limited with regard to their clustering ability and cross platform support.

Existing file systems using AIX logical volumes are however still supported, and this information might be of use when working with those configurations.

## Verify logical volumes are properly defined for GPFS use

The logical volume are configured properly to map between the GPFS NSD and the underlying disk device.

To verify your logical volume configuration, you must first determine the mapping between the GPFS NSD and the underlying disk device. Issue the command:

```
mm|nsd -m
```

which will display any underlying physical device present on this node which is backing the NSD. If the underlying device is a logical volume, perform a mapping from the logical volume to the volume group.

Issue the commands:

```
lsvg -o | lsvg -i -l
```

The output will be a list of logical volumes and corresponding volume groups. Now issue the **lsvg** command for the volume group containing the logical volume. For example:

```
lsvg gpfs1vg
```

The system displays information similar to:

VOLUME GROUP:	gpfs1vg	VG IDENTIFIER:	000195600004c00000000ee60c66352
VG STATE:	active	PP SIZE:	16 megabyte(s)
VG PERMISSION:	read/write	TOTAL PPs:	542 (8672 megabytes)
MAX LVs:	256	FREE PPs:	0 (0 megabytes)
LVs:	1	USED PPs:	542 (8672 megabytes)
OPEN LVs:	1	QUORUM:	2
TOTAL PVs:	1	VG DESCRIPTORS:	2
STALE PVs:	0	STALE PPs:	0
ACTIVE PVs:	1	AUTO ON:	no
MAX PPs per PV:	1016	MAX PVs:	32
LTG size:	128 kilobyte(s)	AUTO SYNC:	no
HOT SPARE:	no		

## Check the volume group on each node

All the disks in GPFS cluster has to be properly defined to all the nodes.

Make sure that all disks are properly defined to all nodes in the GPFS cluster:

1. Issue the AIX **lspv** command on all nodes in the GPFS cluster and save the output.
2. Compare the **pvid** and volume group fields for all GPFS volume groups.

Each volume group must have the same **pvid** and volume group name on each node. The **hdisk** name for these disks may vary.

For example, to verify the volume group **gpfs1vg** on the five nodes in the GPFS cluster, for each node in the cluster issue:

```
lspv | grep gpfs1vg
```

The system displays information similar to:

k145n01:	hdisk3	00001351566acb07	gpfs1vg	active
k145n02:	hdisk3	00001351566acb07	gpfs1vg	active
k145n03:	hdisk5	00001351566acb07	gpfs1vg	active
k145n04:	hdisk5	00001351566acb07	gpfs1vg	active
k145n05:	hdisk7	00001351566acb07	gpfs1vg	active

Here the output shows that on each of the five nodes the volume group **gpfs1vg** is the same physical disk (has the same **pvid**). The **hdisk** numbers vary, but the fact that they may be called different **hdisk** names on different nodes has been accounted for in the GPFS product. This is an example of a properly defined volume group.

If any of the **pvids** were different for the same volume group, this would indicate that the same volume group name has been used when creating volume groups on different physical volumes. This will not work for GPFS. A volume group name can be used only for the same physical volume shared among nodes in a cluster. For more information, refer to AIX in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter/ssw\\_aix/welcome](http://www.ibm.com/support/knowledgecenter/ssw_aix/welcome)) and search for *operating system and device management*.

## Volume group varyon problems

Use **varyoffvg** command for the volume group at all nodes to correct **varyonvg** issues at the volume group layer.

If an NSD backed by an underlying logical volume will not come online to a node, it may be due to **varyonvg** problems at the volume group layer. Issue the **varyoffvg** command for the volume group at all nodes and restart GPFS. On startup, GPFS will **varyon** any underlying volume groups in proper sequence.

---

## Disk accessing commands fail to complete due to problems with some non-IBM disks

Certain disk commands, such as **mmcrfs**, **mmadddisk**, **mmrpldisk**, **mmmout** and the operating system's **mount**, might issue the **varyonvg -u** command if the NSD is backed by an AIX logical volume.

For some non-IBM disks, when many **varyonvg -u** commands are issued in parallel, some of the AIX **varyonvg -u** invocations do not complete, causing the disk command to hang.

This situation is recognized by the GPFS disk command not completing after a long period of time, and the persistence of the **varyonvg** processes as shown by the output of the **ps -ef** command on some of the nodes of the cluster. In these cases, **kill** the **varyonvg** processes that were issued by the GPFS disk command on the nodes of the cluster. This allows the GPFS disk command to complete. Before mounting the affected file system on any node where a **varyonvg** process was killed, issue the **varyonvg -u** command (**varyonvg -u vgname**) on the node to make the disk available to GPFS. Do this on each of the nodes in question, one by one, until all of the GPFS volume groups are varied online.

---

## Disk media failure

Recovery procedures to recover lost data in case of disk media failure.

Regardless of whether you have chosen additional hardware or replication to protect your data against media failures, you first need to determine that the disk has completely failed. If the disk has completely failed and it is not the path to the disk which has failed, follow the procedures defined by your disk vendor. Otherwise:

1. Check on the states of the disks for the file system:

```
mmlsdisk fs1 -e
```

GPFS will mark disks **down** if there have been problems accessing the disk.

2. To prevent any I/O from going to the down disk, issue these commands *immediately*:

```
mmchdisk fs1 suspend -d gpfs1nsd  
mmchdisk fs1 stop -d gpfs1nsd
```

**Note:** If there are any GPFS file systems with pending I/O to the down disk, the I/O will timeout if the system administrator does not stop it.

To see if there are any threads that have been waiting a long time for I/O to complete, on all nodes issue:

```
mmfsadm dump waiters 10 | grep "I/O completion"
```

3. The next step is *irreversible*! Do not run this command unless data and metadata have been replicated. This command scans file system metadata for disk addresses belonging to the disk in question, then replaces them with a special “broken disk address” value, which may take a while.

**CAUTION:**

**Be extremely careful with using the -p option of mmdeldisk, because by design it destroys references to data blocks, making affected blocks unavailable. This is a last-resort tool, to be used when data loss may have already occurred, to salvage the remaining data—which means it cannot take any precautions. If you are not absolutely certain about the state of the file system and the impact of running this command, do not attempt to run it without first contacting the IBM Support Center.**

```
mmdeldisk fs1 gpfs1n12 -p
```

4. Invoke the **mmfileid** command with the operand **:BROKEN**:

```
mmfileid :BROKEN
```

For more information, see “The mmfileid command” on page 201.

5. After the disk is properly repaired and available for use, you can add it back to the file system.

## Replicated metadata and data

The course of actions to be followed to recover the lost files if you have replicated metadata and data and only disks in a single failure group has failed.

If you have replicated metadata and data and only disks in a single failure group have failed, everything should still be running normally but with slightly degraded performance. You can determine the replication values set for the file system by issuing the **mmlsfs** command. Proceed with the appropriate course of action:

1. After the failed disk has been repaired, issue an **mmadddisk** command to add the disk to the file system:

```
mmadddisk fs1 gpfs12nsd
```

You can rebalance the file system at the same time by issuing:

```
mmadddisk fs1 gpfs12nsd -r
```

**Note:** Rebalancing of files is an I/O intensive and time consuming operation, and is important only for file systems with large files that are mostly invariant. In many cases, normal file update and creation will rebalance your file system over time, without the cost of the rebalancing.

2. To re-replicate data that only has single copy, issue:

```
mmrestripefs fs1 -r
```

Optionally, use the **-b** flag instead of the **-r** flag to rebalance across all disks.

**Note:** Rebalancing of files is an I/O intensive and time consuming operation, and is important only for file systems with large files that are mostly invariant. In many cases, normal file update and creation will rebalance your file system over time, without the cost of the rebalancing.

3. Optionally, check the file system for metadata inconsistencies by issuing the offline version of **mmfsck**:

```
mmfsck fs1
```

If **mmfsck** succeeds, you may still have errors that occurred. Check to verify no files were lost. If files containing user data were lost, you will have to restore the files from the backup media.

If **mmfsck** fails, sufficient metadata was lost and you need to recreate your file system and restore the data from backup media.

## Replicated metadata only

Using replicated metadata for lost data recovery.

If you have only replicated metadata, you should be able to recover some, but not all, of the user data. Recover any data to be kept using normal file operations or erase the file. If you read a file in block-size chunks and get a failure return code and an **EIO** errno, that block of the file has been lost. The rest of the file may have useful data to recover, or it can be erased.

## Strict replication

Use **mmchfs -K no** command to perform disk action for strict replication.

If data or metadata replication is enabled, and the status of an existing disk changes so that the disk is no longer available for block allocation (if strict replication is enforced), you may receive an **errno** of **ENOSPC** when you create or append data to an existing file. A disk becomes unavailable for new block allocation if it is being deleted, replaced, or it has been suspended. If you need to delete, replace, or suspend a disk, and you need to write new data while the disk is offline, you can disable strict replication by issuing the **mmchfs -K no** command before you perform the disk action. However, data written while replication is disabled will not be replicated properly. Therefore, after you perform the disk action, you must re-enable strict replication by issuing the **mmchfs -K** command with the original value of the **-K** option (**always** or **whenpossible**) and then run the **mmrestripefs -r** command. To determine if a disk has strict replication enforced, issue the **mmlsfs -K** command.

**Note:** A disk in a **down** state that has not been explicitly suspended is still available for block allocation, and thus a spontaneous disk failure will not result in application I/O requests failing with **ENOSPC**. While new blocks will be allocated on such a disk, nothing will actually be written to the disk until its availability changes to **up** following an **mmchdisk start** command. Missing replica updates that took place while the disk was down will be performed when **mmchdisk start** runs.

## No replication

Perform unmounting yourself if no replication has been done and the system metadata has been lost. You can follow the course of actions for manual unmounting.

When there is no replication, the system metadata has been lost and the file system is basically irrecoverable. You may be able to salvage some of the user data, but it will take work and time. A forced unmount of the file system will probably already have occurred. If not, it probably will very soon if you try to do any recovery work. You can manually force the unmount yourself:

1. Mount the file system in **read-only** mode (see “Read-only mode mount” on page 192). This will bypass recovery errors and let you read whatever you can find. Directories may be lost and give errors, and parts of files will be missing. Get what you can now, for all will soon be gone. On a single node, issue:

```
mount -o ro /dev/fs1
```

2. If you read a file in block-size chunks and get an **EIO** return code that block of the file has been lost. The rest of the file may have useful data to recover or it can be erased. To save the file system parameters for recreation of the file system, issue:

```
mmlsfs fs1 > fs1.saveparms
```

**Note:** This next step is *irreversible!*

To delete the file system, issue:

```
mmdelfs fs1
```

3. To repair the disks, see your disk vendor problem determination guide. Follow the problem determination and repair actions specified.
4. Delete the affected NSDs. Issue:

```
mmdelnsd nsdname
```

The system displays output similar to this:

```
mmdelnsd: Processing disk nsdname  
mmdelnsd: 6027-1371 Propagating the cluster configuration data to all  
affected nodes. This is an asynchronous process.
```

5. Create a disk descriptor file for the disks to be used. This will include recreating NSDs for the new file system.
6. Recreate the file system with either different parameters or the same as you used before. Use the disk descriptor file.
7. Restore lost data from backups.

## GPFS error messages for disk media failures

There are some GPFS error messages associated with disk media failures.

Disk media failures can be associated with these GPFS message numbers:

### 6027-418

Inconsistent file system quorum. readQuorum=*value* writeQuorum=*value* quorumSize=*value*

### 6027-482 [E]

Remount failed for device *name*: *errnoDescription*

### 6027-485

Perform **mmchdisk** for any disk failures and re-mount.

### 6027-636 [E]

Disk marked as stopped or offline.

## Error numbers specific to GPFS application calls when disk failure occurs

There are certain error numbers associated with GPFS application calls when disk failure occurs.

When a disk failure has occurred, GPFS may report these error numbers in the operating system error log, or return them to an application:

### **EOffline = 208, Operation failed because a disk is offline**

This error is most commonly returned when an attempt to open a disk fails. Since GPFS will attempt to continue operation with failed disks, this will be returned when the disk is first needed to complete a command or application request. If this return code occurs, check your disk for stopped states, and check to determine if the network path exists.

To repair the disks, see your disk vendor problem determination guide. Follow the problem determination and repair actions specified.

**ENO\_MGR = 212, The current file system manager failed and no new manager could be appointed.**

This error usually occurs when a large number of disks are unavailable or when there has been a major network failure. Run the **mmlsdisk** command to determine whether disks have failed. If disks have failed, check the operating system error log on all nodes for indications of errors. Take corrective action by issuing the **mmchdisk** command.

To repair the disks, see your disk vendor problem determination guide. Follow the problem determination and repair actions specified.

---

## Persistent Reserve errors

You can use Persistent Reserve (PR) to provide faster failover times between disks that support this feature. PR allows the stripe group manager to "fence" disks during node failover by removing the reservation keys for that node. In contrast, non-PR disk failovers cause the system to wait until the disk lease expires.

GPFS allows file systems to have a mix of PR and non-PR disks. In this configuration, GPFS will fence PR disks for node failures and recovery and non-PR disk will use disk leasing. If all of the disks are PR disks, disk leasing is not used, so recovery times improve.

GPFS uses the **mmchconfig** command to enable PR. Issuing this command with the appropriate **usePersistentReserve** option configures disks automatically. If this command fails, the most likely cause is either a hardware or device driver problem. Other PR-related errors will probably be seen as file system unmounts that are related to disk reservation problems. This type of problem should be debugged with existing trace tools.

## Understanding Persistent Reserve

The AIX server displays the value of *reserve\_policy* and *PR\_key\_value* for Persistent Reserve. Use **chdev** command to set the values for *reserve\_policy* and *PR\_key\_value*.

**Note:** While Persistent Reserve (PR) is supported on both AIX and Linux, *reserve\_policy* is applicable only to AIX.

Persistent Reserve refers to a set of Small Computer Systems Interface-3 (SCSI-3) standard commands and command options. These PR commands and command options give SCSI initiators the ability to establish, preempt, query, and reset a reservation policy with a specified target disk. The functions provided by PR commands are a superset of current reserve and release mechanisms. These functions are not compatible with legacy reserve and release mechanisms. Target disks can only support reservations from either the legacy mechanisms or the current mechanisms.

**Note:** Attempting to mix Persistent Reserve commands with legacy reserve and release commands will result in the target disk returning a reservation conflict error.

Persistent Reserve establishes an interface through a *reserve\_policy* attribute for SCSI disks. You can optionally use this attribute to specify the type of reservation that the device driver will establish before accessing data on the disk. For devices that do not support the *reserve\_policy* attribute, the drivers will use the value of the *reserve\_lock* attribute to determine the type of reservation to use for the disk. GPFS supports four values for the *reserve\_policy* attribute:

**no\_reserve::**

Specifies that no reservations are used on the disk.

**single\_path::**

Specifies that legacy reserve/release commands are used on the disk.

**PR\_exclusive::**

Specifies that Persistent Reserve is used to establish exclusive host access to the disk.

## PR\_shared::

Specifies that Persistent Reserve is used to establish shared host access to the disk.

Persistent Reserve support affects both the parallel (scdisk) and SCSI-3 (scsidisk) disk device drivers and configuration methods. When a device is opened (for example, when the **varyonvg** command opens the underlying **hdiskx**), the device driver checks the ODM for *reserve\_policy* and *PR\_key\_value* and then opens the device appropriately. For PR, each host attached to the shared disk must use unique registration key values for *reserve\_policy* and *PR\_key\_value*. On AIX, you can display the values assigned to *reserve\_policy* and *PR\_key\_value* by issuing:

```
lsattr -E1 hdiskx -a reserve_policy,PR_key_value
```

If needed, use the AIX **chdev** command to set *reserve\_policy* and *PR\_key\_value*.

**Note:** GPFS manages *reserve\_policy* and *PR\_key\_value* using *reserve\_policy=PR\_shared* when Persistent Reserve support is enabled and *reserve\_policy=no\_reserve* when Persistent Reserve is disabled.

## Checking Persistent Reserve

For Persistent Reserve to function properly, follow the course of actions to determine the PR status.

For Persistent Reserve to function properly, you must have PR enabled on all of the disks that are PR-capable. To determine the PR status in the cluster:

1. Determine if PR is enabled on the cluster
  - a. Issue **mmlsconfig**
  - b. Check for `usePersistentReserve=yes`
2. s
3. Determine if PR is enabled for all disks on all nodes
  - a. Make sure that GPFS has been started and mounted on all of the nodes
  - b. Enable PR by issuing **mmchconfig**
  - c. Issue the command **mmlsnsd -X** and look for `pr=yes` on all the `hdisk` lines

### Notes:

1. To view the keys that are currently registered on a disk, issue the following command from a node that has access to the disk:

```
/usr/lpp/mmfs/bin/tsprrreadkeys hdiskx
```
2. To check the AIX ODM status of a single disk on a node, issue the following command from a node that has access to the disk:

```
lsattr -E1 hdiskx -a reserve_policy,PR_key_value
```

## Clearing a leftover Persistent Reserve reservation

The following tables list the RAS events that are applicable to various components of the IBM Spectrum Scale system./////You can collect debugging information in Cluster Export Services. /////

Message number **6027-2202** indicates that a specified disk has a SCSI-3 PR reservation, which prevents the **mmcrnsd** command from formatting it. The following example is specific to a Linux environment. Output on AIX is similar but not identical.

Before trying to clear the PR reservation, use the following instructions to verify that the disk is really intended for GPFS use. Note that in this example, the device name is specified without a prefix (**/dev/sdp** is specified as **sdp**).

1. Display all the registration key values on the disk:

```
/usr/lpp/mmfs/bin/tsprrreadkeys sdp
```



The system displays information similar to:

```
Registration keys for sdp
1. 00006d0000000001
```

If the registered key values all start with 0x00006d, which indicates that the PR registration was issued by GPFS, proceed to the next step to verify the SCSI-3 PR reservation type. Otherwise, contact your system administrator for information about clearing the disk state.

2. Display the reservation type on the disk:

```
/usr/lpp/mmfs/bin/tsprradres sdp
```

The system displays information similar to:

```
yes:LU_SCOPE:WriteExclusive-AllRegistrants:0000000000000000
```

If the output indicates a PR reservation with type **WriteExclusive-AllRegistrants**, proceed to the following instructions for clearing the SCSI-3 PR reservation on the disk.

If the output does not indicate a PR reservation with this type, contact your system administrator for information about clearing the disk state.

To clear the SCSI-3 PR reservation on the disk, follow these steps:

1. Choose a hex value (*HexValue*); for example, **0x111abc** that is not in the output of the **tsprradkeys** command run previously. Register the local node to the disk by entering the following command with the chosen *HexValue*:

```
/usr/lpp/mmfs/bin/tsprrregister sdp 0x111abc
```

2. Verify that the specified *HexValue* has been registered to the disk:

```
/usr/lpp/mmfs/bin/tsprradkeys sdp
```

The system displays information similar to:

```
Registration keys for sdp
1. 00006d0000000001
2. 0000000000111abc
```

3. Clear the SCSI-3 PR reservation on the disk:

```
/usr/lpp/mmfs/bin/tsprrclear sdp 0x111abc
```

4. Verify that the PR registration has been cleared:

```
/usr/lpp/mmfs/bin/tsprradkeys sdp
```

The system displays information similar to:

```
Registration keys for sdp
```

5. Verify that the reservation has been cleared:

```
/usr/lpp/mmfs/bin/tsprradres sdp
```

The system displays information similar to:

```
no:::
```

The disk is now ready to use for creating an NSD.

## Manually enabling or disabling Persistent Reserve

The PR status can be set manually with the help of IBM Support Center.

**Attention:** Manually enabling or disabling Persistent Reserve should only be done under the supervision of the IBM Support Center with GPFS stopped on the node.

The IBM Support Center will help you determine if the PR state is incorrect for a disk. If the PR state is incorrect, you may be directed to correct the situation by manually enabling or disabling PR on that disk.

---

## GPFS is not using the underlying multipath device

You can view the underlying disk device where I/O is performed on an NSD disk by using the **mmlsdisk** command with the **-M** option.

The **mmlsdisk** command output might show unexpected results for multipath I/O devices. For example if you issue this command:

```
mmlsdisk dmfs2 -M
```

The system displays information similar to:

Disk name	I/O performed on node	Device	Availability
m0001	localhost	/dev/sdb	up

The following command is available on Linux only.

```
# multipath -ll
mpathae (36005076304ffc0e50000000000000001) dm-30 IBM,2107900
[size=10G][features=1 queue_if_no_path][hwhandler=0]
\_ round-robin 0 [prio=8][active]
  \_ 1:0:5:1 sdhr 134:16 [active][ready]
  \_ 1:0:4:1 sdgl 132:16 [active][ready]
  \_ 1:0:1:1 sdff 130:16 [active][ready]
  \_ 1:0:0:1 sddz 128:16 [active][ready]
  \_ 0:0:7:1 sdct 70:16 [active][ready]
  \_ 0:0:6:1 sdbn 68:16 [active][ready]
  \_ 0:0:5:1 sdah 66:16 [active][ready]
  \_ 0:0:4:1 sdb 8:16 [active][ready]
```

The **mmlsdisk** output shows that I/O for NSD **m0001** is being performed on disk **/dev/sdb**, but it should show that I/O is being performed on the device-mapper multipath (DMM) **/dev/dm-30**. Disk **/dev/sdb** is one of eight paths of the DMM **/dev/dm-30** as shown from the **multipath** command.

This problem could occur for the following reasons:

- The previously installed user exit **/var/mmfs/etc/nsddevices** is missing. To correct this, restore user exit **/var/mmfs/etc/nsddevices** and restart GPFS.
- The multipath device type does not match the GPFS known device type. For a list of known device types, see **/usr/lpp/mmfs/bin/mmdevdiscover**. After you have determined the device type for your multipath device, use the **mmchconfig** command to change the NSD disk to a known device type and then restart GPFS.

The following output shows that device type **dm-30** is **dmm**:

```
/usr/lpp/mmfs/bin/mmdevdiscover | grep dm-30
dm-30 dmm
```

To change the NSD device type to a known device type, create a file that contains the NSD name and device type pair (one per line) and issue this command:

```
mmchconfig updateNsdType=/tmp/filename
```

where the contents of **/tmp/filename** are:

```
m0001 dmm
```

The system displays information similar to:

mmchconfig: Command successfully completed  
mmchconfig: Propagating the cluster configuration data to all  
affected nodes. This is an asynchronous process.

---

## Kernel panics with a 'GPFS dead man switch timer has expired, and there's still outstanding I/O requests' message

This problem can be detected by an error log with a label of `KERNEL_PANIC`, and the `PANIC MESSAGES` or a `PANIC STRING`.

For example:

```
GPFS Deadman Switch timer has expired, and there's still outstanding I/O requests
```

GPFS is designed to tolerate node failures through per-node metadata logging (journaling). The log file is called the *recovery log*. In the event of a node failure, GPFS performs recovery by replaying the recovery log for the failed node, thus restoring the file system to a consistent state and allowing other nodes to continue working. Prior to replaying the recovery log, it is critical to ensure that the failed node has indeed failed, as opposed to being active but unable to communicate with the rest of the cluster.

In the latter case, if the failed node has direct access (as opposed to accessing the disk with an NSD server) to any disks that are a part of the GPFS file system, it is necessary to ensure that no I/O requests submitted from this node complete once the recovery log replay has started. To accomplish this, GPFS uses the disk lease mechanism. The disk leasing mechanism guarantees that a node does not submit any more I/O requests once its disk lease has expired, and the surviving nodes use disk lease time out as a guideline for starting recovery.

This situation is complicated by the possibility of 'hung I/O'. If an I/O request is submitted prior to the disk lease expiration, but for some reason (for example, device driver malfunction) the I/O takes a long time to complete, it is possible that it may complete after the start of the recovery log replay during recovery. This situation would present a risk of file system corruption. In order to guard against such a contingency, when I/O requests are being issued directly to the underlying disk device, GPFS initiates a kernel timer, referred to as **dead man switch**. The **dead man switch** timer goes off in the event of disk lease expiration, and checks whether there is any outstanding I/O requests. If there is any I/O pending, a kernel panic is initiated to prevent possible file system corruption.

Such a kernel panic is not an indication of a software defect in GPFS or the operating system kernel, but rather it is a sign of

1. Network problems (the node is unable to renew its disk lease).
2. Problems accessing the disk device (I/O requests take an abnormally long time to complete). See "MMFS\_LONGDISKIO" on page 158.



---

## Chapter 17. Security issues

This topic describes some security issues that you might encounter while using IBM Spectrum Scale.

---

### Encryption issues

The topics that follow provide solutions for problems that may be encountered while setting up or using encryption.

#### Unable to add encryption policies

If the `mmchpolicy` command fails when you are trying to add encryption policies, perform the following diagnostic steps:

1. Confirm that the `gpfs.crypto` and `gpfs.gskit` packages are installed.
2. Confirm that the file system is at GPFS 4.1 or later and the fast external attributes (`--fastea`) option is enabled.
3. Examine the error messages that are logged in the `mmfs.log.latest` file, which is located at: `/var/adm/ras/mmfs.log.latest`.

#### Receiving “Permission denied” message

If you experience a “Permission denied” failure while creating, opening, reading, or writing to a file, perform the following diagnostic steps:

1. Confirm that the key server is operational and correctly set up and can be accessed through the network.
2. Confirm that the `/var/mmfs/etc/RKM.conf` file is present on all nodes from which the file is supposed to be accessed. The `/var/mmfs/etc/RKM.conf` file must contain entries for all the RKMs needed to access the file.
3. Verify that the master keys needed by the file and the keys that are specified in the encryption policies are present on the key server.
4. Examine the error messages in the `/var/adm/ras/mmfs.log.latest` file.

#### “Value too large” failure when creating a file

If you experience a “Value too large to be stored in data type” failure when creating a file, follow these diagnostic steps.

1. Examine error messages in `/var/adm/ras/mmfs.log.latest` to confirm that the problem is related to the extended attributes being too large for the inode. The size of the encryption extended attribute is a function of the number of keys used to encrypt a file. If you encounter this issue, update the encryption policy to reduce the number of keys needed to access any given file.
2. If the previous step does not solve the problem, create a new file system with a larger inode size.

#### Mount failure for a file system with encryption rules

If you experience a mount failure for a file system with encryption rules, follow these diagnostic steps.

1. Confirm that the `gpfs.crypto` and `gpfs.gskit` packages are installed.
2. Confirm that the `/var/mmfs/etc/RKM.conf` file is present on the node and that the content in `/var/mmfs/etc/RKM.conf` is correct.
3. Examine the error messages in `/var/adm/ras/mmfs.log.latest`.

## “Permission denied” failure of key rewrap

If you experience a “Permission denied” failure of a key rewrap, follow these diagnostic steps.

When `mmapplypolicy` is invoked to perform a key rewrap, the command may issue messages like the following:

```
[E] Error on gpfs_enc_file_rewrap_key(/fs1m/s1s/test4,KEY-d7bd45d8-9d8d-4b85-a803-e9b794ec0af2:hs21n56_new,KEY-40a0b68b-c86d-4519-9e48-3714d3b71e20:js21n92)
Permission denied(13)
```

If you receive a message similar to this, follow these steps:

1. Check for syntax errors in the migration policy syntax.
2. Ensure that the new key is not already being used for the file.
3. Ensure that both the original and the new keys are retrievable.
4. Examine the error messages in `/var/adm/ras/mmfs.log.latest` for additional details.

---

## E! Authentication issues

E! This topic describes the authentication issues that you might experience while using file and object E! protocols.

### E! File protocol authentication setup issues

E! When trying to enable Active Directory Authentication for file (SMB, NFS), the operation might fail due E! to a timeout. In some cases, the AD server can return multiple IPs that cannot be queried within the E! allotted timeout period and/or IPs that belong to networks inaccessible by the IBM Spectrum Scale E! nodes.

E! **You can try the following workarounds to resolve this issue:**

E! • Remove any invalid/unreachable IPs from the AD DNS.

E! If you removed any invalid/unreachable IPs, retry the `mmuserauth service create` command that E! previously failed.

E! • You can also try to disable any adapters that might not be in use.

E! For example, on Windows 2008: **Start -> Control Panel -> Network and Sharing Center -> Change E! adapter settings** -> Right-click the adapter that you are trying to disable and click **Disable**

E! If you disabled any adapters, retry the `mmuserauth service create` command that previously failed.

### E! Protocol authentication issues

E! You can use a set of GPFS commands to identify and rectify issues that are related to authentication E! configurations.

E! To do basic authentication problem determination, perform the following steps:

E! 1. Issue the `mmces state show auth` command to view the current state of authentication.

E! 2. Issue the `mmces events active auth` command to see whether events are currently contributing to E! make the state of the authentication component unhealthy.

E! 3. Issue the `mmuserauth service list` command to view the details of the current authentication E! configuration.

E! 4. Issue the `mmuserauth service check -N cesNodes --server-reachability` command to verify the state E! of the authentication configuration across the cluster.

E! 5. Issue the `mmuserauth service check -N cesNodes --rectify` command to rectify the authentication E! configuration.

E! **Note:** Server reachability cannot be rectified by using the `--rectify` parameter.

## E! **Authentication error events**

E! This topic describes how to verify and resolve Authentication errors.

E! Following is a list of possible events that may cause a node to go into a failed state and possible solutions E! for each of the issues. To determine what state a component is in, issue the **mmces command**.

### E! **SSD/YPBIND process not running (sssd\_down)**

#### E! **Cause**

E! The SSSD or the YPBIND process is not running.

#### E! **Determination**

E! To learn the authentication current state, run the following command:

```
E! mmces state show auth
```

E! To check the active events for authentication, run the following command:

```
E! mmces events active auth
```

E! To check the current authentication state, run the following command:

```
E! mmces state show auth
```

E! To check the current authentication configuration, run the following command:

```
E! mmuserauth service list
```

E! To check the current authentication configuration across the cluster, run the following command:

```
E! mmuserauth service check -N cesNodes --server-reachability
```

#### E! **Solution**

E! Rectify the configuration by running the following command:

```
E! mmuserauth service check -N cesNodes --rectify
```

E! **Note:** Server reachability cannot be rectified by using the `--rectify` flag.

### E! **Winbind process not running (wnbd\_down)**

#### E! **Cause**

E! The Winbind process is not running.

#### E! **Determination**

E! Run the same command as recommended in the section above, SSD/YPBIND process not running E! (sssd\_down).

#### E! **Solution**

E! Follow the steps in the previous section, SSD/YPBIND process not running (sssd\_down). Then, run the E! following command:

```
E! mmces service stop smb -N <Node on which the problem exists>
```

```
E! mmces service start smb -N <Node on which the problem existed>
```





E!

---

## E! Chapter 18. Protocol issues

E! This topic describes the protocol-related issues (NFS, SMB, and Object) that you might come across while E! using IBM Spectrum Scale.

E!

### E! NFS issues

E! This topic describes some of the possible problems that can be encountered when GPFS interacts with E! NFS.

E! For details on how GPFS and NFS interact, see the *NFS and GPFS* topic in the *IBM Spectrum Scale: Administration Guide*.

E! These are some of the problems encountered when GPFS interacts with NFS:

E! • “NFS client with stale inode data”

E! • “NFSV4 problems”

#### | CES NFS failure due to network failure

| This topic provides information on how to resolve a CES NFS failure caused by a network failure.

| When a network failure occurs because a cable is disconnected, a switch fails, or an adapter fails, CES  
| NFS I/O operations will not complete. To resolve the failure, run the **systemctl restart network**  
| command on the CES node to which the IP is failing back (where the failure occurred). This clears the  
| client suspension and refreshes the network.

#### E! NFS client with stale inode data

E! The NFS client may have to a stale inode data due to caching and the course of action to be followed to E! correct this issue.

E! For performance reasons, some NFS implementations cache file information on the client. Some of the E! information (for example, file state information such as file size and timestamps) is not kept up-to-date in E! this cache. The client may view stale inode data (on **ls -l**, for example) if exporting a GPFS file system E! with NFS. If this is not acceptable for a given installation, caching can be turned off by mounting the file E! system on the client using the appropriate operating system **mount** command option (for example, **-o** E! **noac** on Linux NFS clients).

E! Turning off NFS caching will result in extra file systems operations to GPFS, and negatively affect its E! performance.

E! The clocks of all nodes in the GPFS cluster must be synchronized. If this is not done, NFS access to the E! data, as well as other GPFS file system operations, may be disrupted. NFS relies on metadata timestamps E! to validate the local operating system cache. If the same directory is either NFS-exported from more than E! one node, or is accessed with both the NFS and GPFS mount point, it is critical that clocks on all nodes E! that access the file system (GPFS nodes and NFS clients) are constantly synchronized using appropriate E! software (for example, NTP). Failure to do so may result in stale information seen on the NFS clients.

#### E! NFSV4 problems

E! The analysis of NFS V4 issues and suggestions to resolve these issues.

E! Before analyzing an NFS V4 problem, review this documentation to determine if you are using NFS V4 E! ACLs and GPFS correctly:

- E! 1. The *NFS Version 4 Protocol* paper and other related information that are available in the Network File System Version 4 (nfsv4) section of the IETF Datatracker website ([datatracker.ietf.org/wg/nfsv4/documents](http://datatracker.ietf.org/wg/nfsv4/documents)).
- E! 2. The *Managing GPFS access control lists and NFS export* topic in the *IBM Spectrum Scale: Administration Guide*.
- E! 3. The *GPFS exceptions and limitations to NFS V4 ACLs* topic in the *IBM Spectrum Scale: Administration Guide*.

E! The commands **mmdeacl** and **mmputacl** can be used to revert an NFS V4 ACL to a traditional ACL. Use the **mmdeacl** command to remove the ACL, leaving access controlled entirely by the permission bits in the mode. Then use the **chmod** command to modify the permissions, or the **mmputacl** and **mmeditACL** commands to assign a new ACL.

E! For files, the **mmputacl** and **mmeditACL** commands can be used at any time (without first issuing the **mmdeacl** command) to assign any type of ACL. The command **mmeditACL -k posix** provides a translation of the current ACL into traditional POSIX form and can be used to more easily create an ACL to edit, instead of having to create one from scratch.

## | **NFS mount issues**

| This topic provides information on how to verify and resolve NFS mount errors.

| There are several possible NFS mount error conditions, including

- | • Mount times out
- | • NFS mount fails with a “No such file or directory” error
- | • NFS client cannot mount NFS exports.

### | **Mount times out**

#### | **Description**

| The user is trying to do an NFS mount and receives a timeout error.

#### | **Verification**

| When a timeout error occurs, check the following.

- | 1. Check to see whether the server is reachable by issuing either or both of the following commands:

```
| ping <server-ip>
| ping <server-name>
```

| The expected result is that the server responds.

- | 2. Check to see whether portmapper, NFS, and mount daemons are running on the server.

- | a. On a IBM Spectrum Scale CES node, issue the following command:

```
| mmces service list
```

| The expected results are that the output indicates that the NFS service is running as in this example:

```
| Enabled services: SMB NFS
| SMB is running, NFS is running
```

- | b. On the NFS server node, issue the following command:

```
| rpcinfo -p
```

| The expected result is that portmapper, mountd, and NFS are running as shown in the following sample output.

program	vers	proto	port	service
100000	4	tcp	111	portmapper
100000	4	tcp	111	portmapper
100000	3	tcp	111	portmapper
100000	2	tcp	111	portmapper
100000	4	udp	111	portmapper
100000	3	udp	111	portmapper
100000	2	udp	111	portmapper
100024	1	udp	53111	status
100024	1	tcp	58711	status
100003	3	udp	2049	nfs
100003	3	tcp	2049	nfs
100003	4	udp	2049	nfs
100003	4	tcp	2049	nfs
100005	1	udp	59149	mountd
100005	1	tcp	54013	mountd
100005	3	udp	59149	mountd
100005	3	tcp	54013	mountd
100021	4	udp	32823	nlockmgr
100021	4	tcp	33397	nlockmgr
100011	1	udp	36650	rquotad
100011	1	tcp	36673	rquotad
100011	2	udp	36650	rquotad
100011	2	tcp	36673	rquotad

3. Check to see whether the firewall is blocking NFS traffic on Linux systems by issuing the following command on the NFS client and the NFS server:

```
iptables -L
```

Then check whether any hosts or ports that are involved with the NFS connection are blocked (denied).

If the client and the server are running in different subnets, then a firewall could be running on the router also.

4. Check to see whether the firewall is blocking NFS traffic on the client or router, using the appropriate commands.

## NFS mount fails with a “No such file or directory” error

### Description

The user is trying to do an NFS mount on Linux and receives this message:

```
No such file or directory
```

Following are the root causes of this error.

#### Root cause #1 - Access type is none

An NFS export was created on the server without a specified access type. Therefore, for security reasons, the default access is none, mounting does not work.

### Solution

On the NFS server, specify an access type (for example, RW for Read and Write) for the export. If the export has been created already, you can achieve this by issuing the **mmnfs export change** command. See the following example. The backslash (\) is a line continuation character:

```
mmnfs export change /mnt/gpfs0/nfs_share1 \  
--nfschange "*(Access_Type=RW,Squash=NO_ROOT_SQUASH)"
```

### Verification

| To verify the access type that is specified for the export, issue the **mmnfs export list** on the NFS server.  
| For example:

| `mmnfs export list --nfsdefs /mnt/gpfs0/nfs_share1`

| The system displays output similar to this:

Path	Delegations	Clients	Access_Type	Protocols	Transports	Squash	Anonymous_uid	Anonymous_gid	SecType	PrivilegedPort	Export_id	DefaultDelegation	Manage_Gids	NFS_Commit
/mnt/gpfs0/nfs_share1	none	*	RW	3,4	TCP	NO_ROOT_SQUASH	-2	-2	KRB5	FALSE	2	none	FALSE	FALSE

| "NONE" indicates the root cause; the access type is none .

| "RO" or "RW" indicates that the solution was successful.

### | **Root cause # 2 - Protocol version that is not supported by the server**

#### | **Solution**

| On the NFS server, specify the protocol version needed by the client for export (for example, 3:4). If the export already exists, you can achieve this by issuing the **mmnfs export change** command. For example:

| `mmnfs export change /mnt/gpfs0/nfs_share1 --nfschange "*" (Protocols=3:4)`

#### | **Verification**

| To verify the protocols that are specified for the export, issue the **mmnfs export change** command. For example:

| `mmnfs export list --nfsdefs /mnt/gpfs0/nfs_share1`

| The system displays output similar to this:

Path	Delegations	Clients	Access_Type	Protocols	Transports	Squash	Anonymous_uid	Anonymous_gid	SecType	PrivilegedPort	DefaultDelegations	Manage_Gids	NFS_Commit
/mnt/gpfs0/nfs_share1	none	*	RW	3,4	TCP	NO_ROOT_SQUASH	-2	-2	SYS	FALSE	none	FALSE	FALSE

### | **NFS client cannot mount NFS exports**

#### | **Problem**

| The NFS client cannot mount NFS exports. The **mount** command on the client either returns an error or times out.

#### | **Determination**

| The error itself occurs on the NFS client side. Additionally, and based on the nature of the problem, the server-side NFS logs can provide more details about the origin of the error.

#### | **Solution**

| These are the reasons for client-side mount errors:

- | • The NFS server is not running
- | • The firewall is blocking NFS traffic
- | • The client does not have permissions to mount the export.

| 1. Ensure that the NFS server is running correctly on all of the CES nodes and that the CES IP address used to mount is active in the CES cluster. To check the CES IP address and the NFS server status run:

| `mmlscluster --ces`  
| `mmces service list -a`

- | 2. Ensure that the firewall allows NFS traffic to pass through. In order for this, the CES NFS service must be configured with explicit NFS ports so that discrete firewall rules can be established. On the client, run:  
| `rpcinfo -t <CES_IP_ADDRESS> nfs`
- | 3. Verify that the NFS client is allowed to mount the export. In NFS terms, a definition exists for this client for the export to be mounted. To check NFS export details, enter the following command:  
| `mmnfs export list --nfsdefs <NFS_EXPORT_PATH>`

| The system displays output similar to this:

Path	Delegations	Clients	Access_Type	Protocols	Transports	Squash	Anonymous_uid	Anonymous_gid	SecType	PrivilegedPort	DefaultDelegations	Manage_Gids	NFS_Commit
/mnt/gpfs0/nfs_share1	none	*	RW	3,4	TCP	NO_ROOT_SQUASH	-2	-2	SYS	FALSE	none	FALSE	FALSE

| On the client, run:

| `showmount -e <CES_IP_ADDRESS>`

## | NFS error events

E! This topic provides information on how to verify and resolve NFS errors.

E! Following is a list of possible events that might cause a node to go into a failed state and possible solutions for each of the issues. To determine what state a component is in, run the `mmces events active nfs` command.

### E! NFS is not active (nfs\_not\_active)

#### E! Cause

E! Statistics query indicates that CES NFS is not responding.

#### E! Determination

E! Call the CES NFS statistics command with some delay and compare the NFS server time stamp, then determine if the NFS operation counts are increasing. Run this command:

```
E! /usr/bin/ganesha_stats ; sleep 5 ; /usr/bin/ganesha_stats
E! Timestamp: Wed Apr 27 19:27:22 201634711407 nsecs
E! Total NFSv3 ops: 0
E! Total NFSv4.0 ops: 86449
E! Total NFSv4.1 ops: 0
E! Total NFSv4.2 ops: 0
E! Timestamp: Wed Apr 27 19:27:27 201687146242 nsecs
E! Total NFSv3 ops: 0
E! Total NFSv4.0 ops: 105271
E! Total NFSv4.1 ops: 0
E! Total NFSv4.2 ops: 0
```

#### E! Solution

E! Restart CES NFS on the local CES node using commands `mmces service stop nfs` and `mmces service start nfs`.

### E! CES NFSD process not running (nfsd\_down)

#### E! Cause

E! CES NFS server protocol is no longer running.

#### E! Determination

E! 1. Check to see whether the CES NFS daemon is running:

```
E! ps -C ganesha.nfsd
```

E! 2. Check whether d-bus is alive. Run:

```
E! /usr/bin/ganesha_stats
```

E! If either CES NFS or d-bus is down, you will receive an error:

```
E! ERROR: Can't talk to ganesha service on d-bus. Looks like Ganesh is down.
```

#### E! **Solution**

E! Restart CES NFS on the local CES node by using commands `mmces service stop nfs` and `mmces service start nfs`.

#### E! **RPC statd process is not running (statd\_down)**

E! This applies only if NFS version 3 is enabled in the CES NFS configuration

#### E! **Cause**

E! The `rpc.statd` process is no longer running.

#### E! **Determination**

E! Check `rpc.statd` by running:

```
E! ps -C rpc.statd
```

#### E! **Solution**

E! Restart CES NFS on the local CES node by using commands `mmces service stop nfs` and `mmces service start nfs`.

#### E! **Portmapper port 111 is not active (portmapper\_down)**

#### E! **Cause**

E! RPC call to port 111 failed or timed out.

#### E! **Determination**

E! Check portmapper output by running:

```
E! rpcinfo -n 111 -t localhost portmap
```

```
E! rpcinfo -t localhost nfs 3
```

```
E! rpcinfo -t localhost nfs 4
```

#### E! **Solution**

E! Check to see whether portmapper is running and if portmapper (`rpcbind`) is configured to automatically start on system startup.

E! For more information on NFS events, see “Events” on page 371.

## | **NFS error scenarios**

| This topic provides information on how to verify and resolve NFS errors.

## | **NFS client cannot access exported NFS data**

### | **Problem**

| The NFS client cannot access the exported data even though the export is mounted. This often results in errors to occur while writing data, creating files, or traversing the directory hierarchy (permission denied).

### | **Determination**

| The error itself occurs on the NFS client side. Additionally, and based on the nature of the problem, the server-side NFS logs can provide more details about the origin of the error.

### | **Solution**

| There are multiple reasons for this problem:

| The ACL definition in the file system does not allow the requested operation

| The export and/or client definition of the export do not allow that operation (such as a "read only" definition)

| 1. Verify the ACL definition of the export path in the file system. To check ACL definitions, run:

| `mmgetacl Path`

| 2. Verify the definition of the export and the client (especially the access type). To check the NFS export details, run:

| `mmnfs export list -n Path`

## | **NFS client I/O temporarily stalled**

### | **Problem**

| The NFS client temporarily encounters stalled I/O or access requests to the export. The problem goes away after a short time (about 1 minute.)

### | **Determination**

| The error itself occurs on the NFS client side, but due to an action on the NFS server side. The server-side NFS logs can provide more details about the origin of the error (such as a restart of the NFS server) along with the CES logs (such as manual move of a CES IP or a failover condition).

### | **Origin**

| A restart of the NFS server might temporarily suspend further access to the export from the NFS client (depending on the type of request). The suspension occurs because a restart of the NFS server causes the grace period to start. During the grace period, certain NFS operations are not allowed:

| 1. An explicit restart triggered manually through the CLI by running: **mmces service stop / start ...**

| 2. An explicit move of CES IPs manually through the CLI by running: **mmces address move ...**

| 3. **A change in the definition of an existing export.**

| **Note:** Adding or removing NFS exports does not initiate a restart.

| 4. The creation of the first export.

| 5. A critical error condition that triggers CES failover, which in turn causes IP addresses to move.

| 6. A failback of CES IPs (depending on the setting of the address distribution policy).

---

## E! **SMB issues**

E! This topic describes SMB-related issues that you might come across while using the IBM Spectrum Scale E! system.

### E! **Determining the health of integrated SMB server**

E! There are some GPFS commands to determine the health of the SMB server.

E! The following commands can be used to determine the health of SMB services:

E! • To check the overall CES cluster state, issue the following command:

```
E! mmisccluster --ces
```

E! The system displays output similar to this:

```
E! GPFS cluster information
E! =====
E!   GPFS cluster name:      boris.nsd001st001
E!   GPFS cluster id:       3992680047366063927
E!
E! Cluster Export Services global parameters
E! -----
E!   Shared root directory:  /gpfs/fs0
E!   Enabled Services:      NFS SMB
E!   Log level:             2
E!   Address distribution policy: even-coverage
E!
E! Node  Daemon node name      IP address      CES IP address list
E! -----
E!   4    prt001st001             172.31.132.1   10.18.24.25 10.18.24.32 10.18.24.34 10.18.24.36 9.11.102.89
E!   5    prt002st001             172.31.132.2   9.11.102.90 10.18.24.19 10.18.24.21 10.18.24.23 10.18.24.30
E!   6    prt003st001             172.31.132.3   10.18.24.38 10.18.24.39 10.18.24.41 10.18.24.42 9.11.102.43
E!   7    prt004st001             172.31.132.4   9.11.102.37 10.18.24.26 10.18.24.28 10.18.24.18 10.18.24.44
E!   8    prt005st001             172.31.132.5   9.11.102.36 10.18.24.17 10.18.24.33 10.18.24.35 10.18.24.37
E!   9    prt006st001             172.31.132.6   9.11.102.41 10.18.24.24 10.18.24.20 10.18.24.22 10.18.24.40
E!  10    prt007st001             172.31.132.7   9.11.102.42 10.18.24.31 10.18.24.27 10.18.24.29 10.18.24.43
```

E! This shows at a glance whether nodes are failed or whether they host public IP addresses. For E! successful SMB operation at least one CES node must be HEALTHY and hosting at least one IP E! address.

E! • To show which services are enabled, issue the following command:

```
E! mmces service list
```

E! The system displays output similar to this:

```
E! Enabled services: NFS SMB
E! NFS is running, SMB is running
```

E! For successful SMB operation, SMB needs to be enabled and running.

E! • To determine the overall health state of SMB on all CES nodes, issue the following command:

```
E! mmces state show smb -a
```

E! The system displays output similar to this:

```
E! NODE      SMB
E! prt001st001 HEALTHY
E! prt002st001 HEALTHY
E! prt003st001 HEALTHY
E! prt004st001 HEALTHY
E! prt005st001 HEALTHY
E! prt006st001 HEALTHY
E! prt007st001 HEALTHY
```

E! • To show the reason for a currently active (failed) state on all nodes, issue the following command:

```
E! mmces events active SMB -a
```



E! The system displays output similar to this:

```
E! NODE COMPONENT EVENT NAME SEVERITY DETAILS
```

E! In this case nothing is listed because all nodes are healthy and so there are no active events. If a node was unhealthy it would look similar to this:

```
E! NODE COMPONENT EVENT NAME SEVERITY DETAILS
E! prt001st001 SMB ctdb_down ERROR CTDB process not running
E! prt001st001 SMB smbd_down ERROR SMBD process not running
```

E! • To show the history of events generated by the monitoring framework, issue the following command

```
E! mmces events list SMB
```

E! The system displays output similar to this:

```
E! NODE          TIMESTAMP          EVENT NAME          SEVERITY  DETAILS
E! prt001st001  2015-05-27  14:15:48.540577+07:07MST  smbd_up      INFO      SMBD process now running
E! prt001st001  2015-05-27  14:16:03.572012+07:07MST  smbport_up   INFO      SMB port 445 is now active
E! prt001st001  2015-05-27  14:28:19.306654+07:07MST  ctdb_recovery WARNING   CTDB Recovery detected
E! prt001st001  2015-05-27  14:28:34.329090+07:07MST  ctdb_recovered INFO      CTDB Recovery finished
E! prt001st001  2015-05-27  14:33:06.002599+07:07MST  ctdb_recovery WARNING   CTDB Recovery detected
E! prt001st001  2015-05-27  14:33:19.619583+07:07MST  ctdb_recovered INFO      CTDB Recovery finished
E! prt001st001  2015-05-27  14:43:50.331985+07:07MST  ctdb_recovery WARNING   CTDB Recovery detected
E! prt001st001  2015-05-27  14:44:20.285768+07:07MST  ctdb_recovered INFO      CTDB Recovery finished
E! prt001st001  2015-05-27  15:06:07.302641+07:07MST  ctdb_recovery WARNING   CTDB Recovery detected
E! prt001st001  2015-05-27  15:06:21.609064+07:07MST  ctdb_recovered INFO      CTDB Recovery finished
E! prt001st001  2015-05-27  22:19:31.773404+07:07MST  ctdb_recovery WARNING   CTDB Recovery detected
E! prt001st001  2015-05-27  22:19:46.839876+07:07MST  ctdb_recovered INFO      CTDB Recovery finished
E! prt001st001  2015-05-27  22:22:47.346001+07:07MST  ctdb_recovery WARNING   CTDB Recovery detected
E! prt001st001  2015-05-27  22:23:02.050512+07:07MST  ctdb_recovered INFO      CTDB Recovery finished
```

E! • To retrieve monitoring state from health monitoring component, issue the following command:

```
E! mmces state show
```

E! The system displays output similar to this:

```
E! NODE AUTH NETWORK NFS OBJECT SMB CES
E! prt001st001 DISABLED HEALTHY HEALTHY DISABLED DISABLED HEALTHY
```

E! • To check the monitor log, issue the following command:

```
| grep smb /var/adm/ras/mmsysmonitor.log | head -n 10
```

E! The system displays output similar to this:

```
E! 2016-04-27T03:37:12.2 prt2st1 I Monitor smb service LocalState:HEALTHY Events:0 Entities:0 - Service.monitor:596
E! 2016-04-27T03:37:27.2 prt2st1 I Monitor smb service LocalState:HEALTHY Events:0 Entities:0 - Service.monitor:596
E! 2016-04-27T03:37:42.3 prt2st1 I Monitor smb service LocalState:HEALTHY Events:0 Entities:0 - Service.monitor:596
E! 2016-04-27T03:37:57.2 prt2st1 I Monitor smb service LocalState:HEALTHY Events:0 Entities:0 - Service.monitor:596
E! 2016-04-27T03:38:12.4 prt2st1 I Monitor smb service LocalState:HEALTHY Events:0 Entities:0 - Service.monitor:596
E! 2016-04-27T03:38:27.2 prt2st1 I Monitor smb service LocalState:HEALTHY Events:0 Entities:0 - Service.monitor:596
E! 2016-04-27T03:38:42.5 prt2st1 I Monitor smb service LocalState:HEALTHY Events:0 Entities:0 - Service.monitor:596
E! 2016-04-27T03:38:57.2 prt2st1 I Monitor smb service LocalState:HEALTHY Events:0 Entities:0 - Service.monitor:596
E! 2016-04-27T03:39:12.2 prt2st1 I Monitor smb service LocalState:HEALTHY Events:0 Entities:0 - Service.monitor:596
E! 2016-04-27T03:39:27.6 prt2st1 I Monitor smb service LocalState:HEALTHY Events:0 Entities:0 - Service.monitor:596
```

E! • The following logs can also be checked:

```
E! /var/adm/ras/*
E! /var/log/messages
```

## E! File access failure from an SMB client with sharing conflict

E! If SMB clients fail to access files with file sharing conflict messages, and no such conflict exists, there can be a mismatch with file locking rules.

E! File systems that are exported with the CES SMB service, or a customized deployment version of Samba, require the **-D nfs4** flag on the **mmchfs** or **mmcrfs** command. This setting enables NFSv4 and SMB sharing rules.

## E! **SMB client on Linux fails with an “NT status logon failure”**

E! This topic describes how to verify and resolve an “NT status logon failure” on the SMB client on Linux.

### E! **Description**

E! The user is trying to log on to the SMB client using AD authentication on Linux and receives this E! message:

E! NT STATUS LOGON FAILURE

E! Following are the root causes of this error.

### E! **Description of Root cause #1**

E! The user is trying to log on to the SMB client using AD authentication on Linux and receives this E! message:

E! Password Invalid

### E! **Cause**

E! The system did not recognize the specified password.

### E! **Verification**

E! Verify the password by running the following command on an IBM Spectrum Scale protocol node:

```
E! /usr/lpp/mmfs/bin wbinfo -a '<domain>\<user>'
```

E! The expected result is that the following messages display:

E! plaintext password authentication succeeded.

E! challenge/response password authentication succeeded.

E! If this message displays:

E! plaintext password authentication failed.

E! Could not authenticate user USER with plain text password

E! the domain for that user was not specified correctly.

### E! **Resolution**

E! To resolve the error, enter the correct password.

E! If you do not know the correct password, follow your IT procedures to request a new password.

### E! **Description of root cause # 2**

E! The user is trying to log on to the SMB client using AD authentication on Linux and receives this E! message:

E! The Userid is not recognized

### E! **Cause**

E! The system did not recognize the specified password.

### E! **Verification**

E! Verify the password by running the following command on an IBM Spectrum Scale protocol node:

```
E! /usr/lpp/mmfs/bin wbinfo -a '<domain>\<user>'
```

E! The expected result is that the following messages display:

```
E! plaintext password authentication succeeded.
```

```
E! challenge/response password authentication succeeded
```

E! If this message displays:

```
E! Could not authenticate user USER with challenge/response password
```

E! the specified user is not known by the system.

### E! **Resolution**

E! To resolve the error, enter the correct userid.

E! If you think the correct user was specified, contact your IT System or AD Server administrator to get

E! your userid verified.

## | **SMB client on Linux fails with the NT status password must change error message**

| This topic describes how to verify and resolve an NT status password must change error on the SMB client on Linux.

### | **Description**

| The user is trying to access the SMB client on Linux and receives this error message:

```
| NT_STATUS_PASSWORD_MUST_CHANGE
```

### | **Cause**

| The specified password expired.

### | **Verification**

| Verify the password by running the following command on an IBM Spectrum Scale protocol node:

```
| /usr/lpp/mmfs/bin wbinfo -a '<domain>\<user>'
```

| The expected result is that the following messages display:

```
| plaintext password authentication succeeded.
```

```
| challenge/response password authentication succeeded.
```

| If this message displays:

```
| Could not authenticate user mzdom\aduser1 with challeng/response
```

| the specified password probably expired.

### | **Resolution**

| Log on to a Windows client, and when prompted, enter a new password. If the problem persists, ask the AD administrator to unlock the account.

## | **SMB mount issues**

| This topic describes how to verify and resolve SMB mount errors.

| Possible SMB mount error conditions include:

- | • Mount.CIFS on Linux fails with mount error (13) “Permission denied”
- | • Mount.CIFS on Linux fails with mount error (127) “Key expired”
- | • Mount on Mac fails with an authentication error.

E! If you receive any of these errors, verify your authentication settings. For more information, see “Protocol E! authentication issues” on page 294

### E! **Mount.Cifs on Linux fails with mount error (13) “Permission denied”**

#### E! **Description**

E! The user is trying to mount CIFS on Linux and receives the following error message:

E! Permission Denied

E! The root causes for this error are the same as for “SMB client on Linux fails with an “NT status logon E! failure”” on page 306.

### E! **Mount.Cifs on Linux fails with mount error (127) “Key has expired”**

#### E! **Description**

E! The user is trying to access a CIFS share and receives the following error message:

E! key has expired

E! The root causes for this error are the same as for “SMB client on Linux fails with an “NT status logon E! failure”” on page 306.

### E! **Mount on Mac fails with an authentication error**

#### E! **Description**

E! The user is attempting a mount on a Mac and receives this error message:

E! mount\_smbfs: server rejected the connection: Authentication error

E! The root causes for this error are the same as for “SMB client on Linux fails with an “NT status logon E! failure”” on page 306.

## | **Net use on Windows fails with “System error 86”**

| This topic describes how to verify and solve a “System error 86” when the user is attempting to access net use on Windows.

### | **Description**

| The user is attempting to access the network and receives this error:

- | System error 86 has occurred.
- | The specified password is not correct.

| The root causes for this error are the same as that for the failure of SMB client on Linux. For more information on the root cause, see “SMB client on Linux fails with an “NT status logon failure”” on page 306.

## | **SMB error events**

| This topic describes how to verify and resolve SMB errors.

### | **CTDB process is not running (ctdb\_down)**

#### | **Cause**

| CTDB process is not running.

#### | **Determination**

| Check /var/log/messages for CTDB error messages or crashes.

#### | **Solution**

| Fix any obvious issues and run this command:

```
| mmces service stop SMB  
| mmces service start SMB
```

### | **CTDB recovery detected (ctdb\_recovery)**

#### | **Cause**

| CTDB status is stuck in Recovery mode for an extended amount of time.

#### | **Determination**

| If the service status is Degraded for a while, there is an issue. The service status should be Transient.  
| Check the logs for a possible issue.

#### | **Solution**

| Run:

```
| mmces service stop SMB && mmces service start SMB
```

| If still not fixed, run:

```
| gpfs.snap
```

| and contact IBM support.

### | **CTDB state is not healthy (ctdb\_state\_down)**

#### | **Determination**

- | 1. Check /var/log/messages for errors and correct any that you find.
- | 2. Check CTDB status by running the **ctdb status** command.
- | 3. Check the network connectivity.
- | 4. Check the banned reason by running '**ctdb showban**' and attempt to correct the error.

#### | **Solution**

| After the error is resolved, the CTDB node should recover. If you have not resolved the error, restart SMB by running this command:

```
| mmces service stop SMB && mmces service start SMB
```

## | **SMDB process not running**

### | **Determination**

- | 1. Check `/var/log/messages` and `/var/adm/ras/log.smbd` for errors and correct if found.
- | 2. Restart by running this command:  
| `mmces service stop SMB && mmces services start SMB`

## | **SMB port (?) is not active (smbport\_down\_)**

### | **Cause**

| The SMB port (?) is not listening for connections.

### | **Determination**

| Check the network connectivity.

### | **Solution**

- | Restart by running:  
| `mmces service stop SMB && mmces services start SMB`

## | **SMB access issues**

E! This topic describes how to analyze and resolve SMB access issues.

E! The most common issue with ACLs is getting an unexpected Access denied message. Check the following:

- E! 1. Export ACLs: Use the MMC tool or `mmsmb exportacl` to see that the share allows access for the logged in user.
- E! 2. File system object ACLs: Use the Windows Explorer ACL dialog and/or `mmgetacl` to make sure the correct ACLs are in place on all components in the path.
- E! 3. Make sure that the `READ_ATTR` right is set on folders to be traversed.
- E! 4. Keep in mind that even if `READ_NAMED` and `WRITE_NAMED` are not enforced by the file system, the SMB server enforces them.
- E! 5. Export settings: Check the export settings by running `mmsmb export list --all` so that export options like `read only = no` or `available = no` do not restrict access.
- E! 6. Make sure your clients try to negotiate a supported protocol level.
- E! 7. For `smbclient`: make sure the option `-m SMB2` is used and supported by your version of `smbclient` (`smbclient -L localhost -U<user>%<password> -m SMB2`)
- E! 8. Windows XP, Windows Server 2003 and older Windows versions are not supported, because they only support SMB1.
- E! 9. For the Linux kernel client, make sure you check the version option to use `smb2`.

E! **Note:** For known issues in the Linux kernel client, see the documentation for your Linux distribution.

E! If the root cause cannot be narrowed down, perform these steps the results of which will help make a more detailed analysis.

- E! 1. Provide exact information about what happened.
- E! 2. Provide screen captures of Windows ACL dialogs with the problem before and after the issue.
- E! 3. Provide the output of `mmgetacl` for all files and folders related to the ACL/permission problem before and after the problematic event.

- E! 4. Trace how the client has mounted the share.
- E! 5. You can force a re-connect by stopping the `smbd` process that serves that connection.
- E! 6. Describe how the user has mounted the export.
- E! 7. List all users and groups that are in the test along with their memberships.
- E! 8. Collect export information by running: `mmsmb export list --all`.
- E! 9. Provide the version of Windows used for each client.
- E! 10. Provide a Samba level 10 trace for the test by running the `mmprotocoltrace` tool.
- E! 11. Provide IBM Spectrum Scale traces for the test by running `mmtracectl --start and --stop`.
- E! 12. Collect the network trace of the re-create by running `mmprotocoltrace`.

---

## E! Object issues

E! This topic describes some of the Object-related issues that you might come across while using IBM Spectrum Scale.

### E! Getting started with troubleshooting object issues

E! Use the following checklists to troubleshoot object issues.

#### | Checklist 1

| This checklist must be referred to before using an object service.

- | 1. Check the cluster state by running the `mmgetstate -a` command.  
| The cluster state must be Active.
- | 2. If the server is not functioning properly, restart the server by running the `mmstartup -a` command.
- | 3. Check the status of the CES IP by running the `mmiscluster -ces` command.  
| The system displays the all the CES nodes along with their assigned IP addresses.
- | 4. Check the service states of the CES by running the `mmces state show -a` or `mmhealth node show ces -N cesnodes` command.  
| The overall CES state and object service states must be Healthy.
- | 5. Check the service listing of all the service states by running the `mmces service list -verbose` command.
- | 6. Check the authentication status by running the `mmuserauth service check` command.
- | 7. Check the object auth listing by running the `source openrc ; openstack user list` command.  
| The system lists all the users IDs.

#### | Checklist 2

| This checklist must be referred to before using the keystone service.

- | 1. Check if object authentication has been configured by running the `mmuserauth service list --data-access-method object` command.
- | 2. Check the state of object authentication by running the `mmces state show AUTH_OBJ -a` command.
- | 3. Check if the protocol node is serving the CES IP by running the `mmiscluster --ces` command.
- | 4. Check if the `object_database_node` tag is present in one of the CES IP by running the `mmces address list` command.
- | 5. Check if `httpd` is running on all the CES nodes and `postgres` is running on the node that has CES IP with the `object_database_node` tag by running the `mmces service list -v a` command.
- | 6. Check if authentication configuration is correct on all nodes by running the `mmuserauth service check --data-access-method object -N cesNodes` command.
- | 7. If the `mmuserauth service check` reports an error, run the `mmuserauth service check --data-access-method object --rectify -N <node>` command where `node` is the number of the node on which the error is reported.

## | **Authenticating the object service**

| This topic provides troubleshooting references and steps for resolving system errors when you are authenticating the object service.

### | **Description**

| When the user authenticates or runs any create, update, or delete operation, the system displays one of the following errors:

```
| {"error": {"message": "An unexpected error prevented the server from fulfilling your request.",  
| "code": 500, "title": "Internal Server Error"}}
```

```
| ERROR: openstack An unexpected error prevented the server from fulfilling your request.  
| (HTTP 500)(Request-ID: req-11399fd1-a601-4615-8f70-6ba275ec3cd6)
```

### | **Cause**

| The system displays this error under one or all of the following conditions:

- | • The authentication service is not running.
- | • The system is unable to reach the authentication server.
- | • The user credentials for keystone have been changed or have expired.

### | **Proposed workaround**

- | • Perform all the steps in Checklist 1.
- | • Check if the IP addresses of the keystone endpoints are correct and reachable. If you are using a local keystone, check if the postgresql-obj service is running.

## | **Authenticating or using the object service**

E! This topic provides troubleshooting references and steps for resolving system errors when you are authenticating or using the object service.

### E! **Description**

E! When the user is authenticating the object service or running the create, update, retrieve, and delete operations, the system displays the following error:

```
E! Error: {"error": {"message": "The request you have made requires authentication.",  
E! "code": 401, "title": "Unauthorized"}}
```

### E! **Cause**

E! The system displays this error under one or all of the following conditions:

- E! • The password, user ID, or service ID that you have entered is incorrect.
- E! • The token that you are using has expired.

### E! **Proposed workaround**

- E! • Check your user ID and password. All user IDs in the system can be viewed in the OpenStack user list.
- E! • While using the curl command, To check if your token has expired, user must get a new token from keystone for running the swift command.
- E! • Check if a valid service ID is provided in the /etc/swift/proxy-server.conf file, in the filter:authtoken section. Also, check if the password for the service ID is still valid. The service ID can be viewed in the OpenStack service, project, and endpoint lists.



## E! **Accessing resources**

E! This topic provides troubleshooting references and steps for resolving system errors when you are accessing resources.

### E! **Description**

E! When an unauthorized user is accessing an object resource, the system displays the following error:

```
E! Error: Error: HTTP/1.1 403 Forbidden
E! Content-Length: 73 Content-Type: text/html; charset=UTF-8 X-Trans-Id: tx90ad4ac8da9242068d111-
E! 0056a88ff0 Date: Wed, 27 Jan 2016 09:37:52 GMT <html><h1>Forbidden</h1><p>Access was denied to
E! this resource.</p>
```

### E! **Cause**

E! The system displays this error under one or all of the following conditions:

- E! • The user is not authorized by the system to access the resources for a certain operation.
- E! • The endpoint, auth URL, service ID, keystone version, or API version is incorrect.

### E! **Proposed workaround**

- E! • Contact your system administrator for gaining authorization to access the resources.
- E! • Check your service ID. The service ID can be viewed in the OpenStack service, project, and endpoint lists.

## E! **Connecting to the object services**

E! This topic provides troubleshooting references and steps for resolving system errors when you are connecting to the object services.

### E! **Description**

E! When the user is unable to connect to the object services, the system displays the following error:

```
E! curl: (7) Failed connect
E! to sptscl2.in.ibm.com:8080; No route to host
```

### E! **Cause**

E! The system displays this error under one or all of the following conditions:

- E! • The firewall is running.
- E! • The firewall rules have been configured incorrectly.

### E! **Proposed workaround**

E! Set up the firewall rules correctly in your system.

E! For more information about the firewall rules, see *Installation prerequisites* in *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.

## | **Creating a path**

| This topic provides troubleshooting references and steps for resolving system errors when you are creating a path.

## | **Description**

| When the user is trying to create an inexistent path while performing the create, update, retrieve, or delete task, the system displays the following error:

```
| Error: HTTP/1.1 404 Not Found  
| Content-Length: 70 Content-Type: text/html; charset=UTF-8 X-Trans-Id:  
| tx88ec3b783bc04b78b5608-0056a89b52 Date: Wed, 27 Jan 2016 10:26:26  
| GMT <html><h1>Not Found</h1><p>The resource could not be found.</p></html>
```

## | **Cause**

| The system displays this error because the path you are creating does not exist.

## | **Proposed workaround**

| Recreate the object or the container before performing the GET operation.

## | **Constraints for creating objects and containers**

| This topic provides the constraints that must be kept in mind while creating objects and containers.

## | **Description**

| When the user is trying to create objects and containers for unified file and object access, the system displays the 400 Bad request error.

## | **Cause**

| The system displays this error under one or all of the following conditions:

- | • The name of the container is longer than 255 characters.
- | • The name of the object is longer than 214 characters.
- | • The name of any container in the object hierarchy is longer than 214 characters.
- | • The path name of the object includes successive forward slashes.
- | • The name of the container and the object is a single period (.) or a double period (..).

## | **Proposed workaround**

| Keep in mind the following constraints while creating objects and containers for unified file and object access:

- | • The name of the container must not exceed 255 characters.
- | • The name of the object must not exceed 214 characters.
- | • The name of any container in the object hierarchy must not exceed 214 characters.
- | • The path name of the object must not include successive forward slashes.
- | • The name of the container and the object must not be a single period (.) or a double period (..).  
| However, a single period or a double period can be part of the name of the container and the object.

## | **The Bind password is used when the object authentication E! configuration has expired**

E! This topic provides troubleshooting references and steps for resolving system errors when you are using E! the Bind password and the object authentication configuration has expired.

## E! **Description**

E! When object is configured with the AD/LDAP authentication and the bind password is being used for E! LDAP communication, the system displays the following error:

```
E! [root@SSClusterNode3 ~]# openstack user list ERROR: openstack An unexpected error prevented the
E! server from fulfilling your request. (HTTP 500) (Request-ID: req-d2ca694a-31e3-46cc-98b2-
E! 93556571aa7d) Authorization Failure. Authorization failed: An unexpected error prevented the
E! server from fulfilling your request. (HTTP 500) (Request-ID: req-d6ccba54-baea-4a42-930e-
E! e9576466de3c)
```

## E! **Cause**

E! The system displays this error when the Bind password has been changed on the AD/LDAP server.

## E! **Proposed workaround**

E! 1. Obtain the new password from the AD/LDAP server.

E! 2. Run the following command to restart keystone on all protocol nodes: `mmobj config change`

E! `--ccrfile keystone.conf --section ldap --property password --value <password>` where password

E! is the new password obtained in Step 1.

E! **Note:** This command restarts Keystone on all protocol nodes.

## E! **The password used for running the keystone command has expired or is incorrect**

E! This topic provides troubleshooting references and steps for resolving system errors when you are using E! an expired or incorrect password for running the keystone command.

## E! **Description**

E! When the user is trying to run the keystone command by using a password that has expired or is E! incorrect, the system displays the following error:[root@SSClusterNode5 ~]# openstack user listERROR: E! openstack The request you have made requires authentication. (HTTP 401) (Request-ID: E! req-9e8d91b6-0ad4-42a8-b0d4-797a08150cea)

## E! **Cause**

E! The system displays this error when the user has changed the password but is still using the expired E! password to access keystone.

## E! **Proposed workaround**

E! Use the correct password to access keystone.

## | **The LDAP server is not reachable**

| This topic provides troubleshooting references and steps for resolving system errors when you are trying | to reach an LDAP server.

## | **Description**

| When object authentication is configured with AD/LDAP and the user is trying to run the keystone | commands, the system displays the following error:[root@SSClusterNode3 ~]# openstack user | listERROR: openstack An unexpected error prevented the server from fulfilling your request. | (HTTP 500) (Request-ID: req-d3fe863e-da1f-4792-86cf-bd2f4b526023)

## | **Cause**

| The system displays this error under one or all of the following conditions:

- | • The LDAP server is not reachable due to network issues.
- | • The LDAP server is not reachable because the system firewall is running.
- | • The LDAP server has been shut down.

## | **Note:**

| When the LDAP server is not reachable, the keystone logs can be viewed in the /var/log/keystone directory.

| The following example is an LDAP error found in /var/log/keystone/keystone.log:

```
| /var/log/keystone/keystone.log:2016-01-28 14:21:00.663 25720 TRACE keystone.common.wsgi result =  
| func(*args,**kwargs)2016-01-28 14:21:00.663 25720 TRACE keystone.common.wsgi SERVER_DOWN:  
| {'desc': "Can't contact LDAP server"}.
```

## | **Proposed workaround**

- | • Check your network settings.
- | • Configure your firewall correctly.
- | • Repair the LDAP server.

## | **The TLS certificate has expired**

E! This topic provides troubleshooting references and steps for resolving system errors when the TLS certificate has expired.

### E! **Description**

E! When the user is trying to configure object authentication with AD/LDAP by using the TLS certificate for E! configuration, the system displays the following error:

```
E! [E] Failed to execute command  
E! ldapsearchldap_start_tls: Connect error (-11)additional info: TLS error -8174:security library  
E! : bad database.mmuserauth service create: Command failed.  
E! Examine previous error messages to determine cause.
```

### E! **Cause**

E! The system displays this error because the TLS certificate has expired.

### E! **Proposed workaround**

- E! 1. Update the TLS certificate on the AD/LDAP server.
- E! 2. Rerun the command.

## E! **The TLS CACERT certificate has expired**

E! This topic provides troubleshooting references and steps for resolving system errors when the TLS certificate has expired.

### E! **Description**

E! When the system is configured with AD/LDAP and TLS, the TLS CACERT has expired after E! configuration, and the user is trying to run the keystone command, the system displays the following E! error:

```
E! [root@SSClusterNode3 ~]# openstack user list
E! ERROR: openstack An unexpected error prevented the server from fulfilling your request.
E! (HTTP 500) (Request-ID: req-dfd63d79-39e5-4c4a-951d-44b72e8fd9ef)
E! Logfile /var/log/keystone/keystone.log2045-01-14 10:50:40.809 30518
E! TRACE keystone.common.wsgi CONNECT_ERROR:
E! {'info': "TLS error -8162:The certificate issuer's certificate has expired."}
E! Check your system date and time.", 'desc': 'Connect error'}
```

E! **Note:**

E! The log files for this error can be viewed in /var/log/keystone/keystone.log.

### E! **Cause**

E! The system displays this error because the TLS CACERT certificate has expired.

### E! **Proposed workaround**

- E! 1. Obtain the updated TLS CACERT certificate on the system.
- E! 2. Rerun the object authentication command.

E! **Note:**

E! Do not run the `-idmdelete` command before you run the object authentication command.

## E! **The TLS certificate on the LDAP server has expired**

E! This topic provides troubleshooting references and steps for resolving system errors when the TLS certificate on the LDAP server has expired.

### E! **Description**

E! When the system is configured with AD/LDAP using TLS, and the certificate on AD/LDAP has expired, E! the system displays the following error when the user is trying to run the keystone commands:

```
E! [root@SSClusterNode3 ~]# openstack user list
E! ERROR: openstack An unexpected error prevented the server from fulfilling your request.
E! (HTTP 500) (Request-ID: req-5b3422a1-fc43-4210-b092-1201e38b8cd5)2017-05-08 22:08:35.443 30518
E! TRACE keystone.common.wsgi CONNECT_ERROR: {'info': 'TLS error -8157:Certificate extension not found.',
E! 'desc': 'Connect error'}
E! 2017-05-08 22:08:35.443 30518 TRACE keystone.common.wsgi
```

### E! **Cause**

E! The system displays this error because the TLS certificate on the LDAP server has expired.

### E! **Proposed workaround**

E! Update the TLS certificate on the LDAP server.

## | **The SSL certificate has expired**

| This topic provides troubleshooting references and steps for resolving system errors when the SSL certificate has expired.

### | **Description**

| When object authentication is configured with SSL and the user is trying to run the authentication commands, the system displays the following error:

```
| [root@SSClusterNode3 ~]# openstack user list
| ERROR: openstack SSL exception connecting to https://SSCluster:35357/v3/auth/tokens:
| [Errno 1] _ssl.c:504: error:14090086:SSL routines:SSL3_GET_SERVER_CERTIFICATE:certificate verify failed
```

## | Cause

| The system displays this error because the SSL certificate has expired. The user may have used the same certificate earlier for keystone configuration, but now the certificate has expired.

## | Proposed workaround

- | 1. Remove the object authentication.
- | 2. Reconfigure the authentication with the new SSL certificate.

## | Note:

| Do not run the `mmuserauth service remove --data-access-method object --idmapdelete` command during removing and reconfiguring the authentication.

## | The signing certificate has expired

E! This topic provides troubleshooting references and steps for resolving system errors when the signing E! certificate has expired.

## E! Description

E! When object authentication uses PKI as a token driver, it requires signing certificates. If the system is E! configured in such a manner and the user tries to run the keystone commands, the system displays the E! following error:

```
E! [root@SSClusterNode3 tmp]# openstack user listERROR: openstack SSL exception connecting to
E! https://SSCluster:35357/v3/auth/tokens: [Errno 1] _ssl.c:504: error:140770FC:SSL
E! routines:SSL23_GET_SERVER_HELLO:unknown protocol[root@SSClusterNode3 ~]#
E! swift listAuthorization Failure. Authorization failed: SSL exception connecting to
E! https://SSCluster:35357/v3/auth/tokens: [Errno 1] _ssl.c:504:
E! error:140770FC:SSL routines:SSL23_GET_SERVER_HELLO:unknown protocol
```

## E! Cause

E! The system displays this error because the signing certificate has expired after configuration.

## E! Proposed workaround

- E! 1. Remove the authentication.
- E! 2. Reconfigure the authentication with the new signing certificate.

## E! Note:

E! Do not run the `mmuserauth service remove --data-access-method object --idmapdelete` command E! during removing and reconfiguring the authentication.

## E! Users are not listed in the OpenStack user list

E! This topic provides troubleshooting references and steps for resolving system errors when the user is not E! listed in the OpenStack user list.

## E! Description

E! When the authentication type is AD/LDAP, the users are not listed in the OpenStack user list.

## E! Cause

E! The system displays this error under one or all of the following conditions:

- E! • Only the users under the specified user DN are visible to keystone.
- E! • The users do not have the specified object class.

## E! Proposed workaround

E! Change the object authentication or modify the AD/LDAP for the users who do not have the specified E! object class.

## | IBM Security Lifecycle Manager cannot be installed

| This topic provides troubleshooting references and steps for resolving system errors when IBM Security Lifecycle Manager cannot be installed.

### | Description

| When the user tries to install IBM Security Lifecycle Manager, the system displays the following errors:

```
| eclipse.buildId=unknownjava.fullversion=JRE 1.6.0 IBM J9 2.4 Linux x86-32
| jvmxi3260sr9-20110203_74623 (JIT enabled, AOT enabled)J9VM -
| 20110203_074623JIT - r9_20101028_17488ifx3GC - 20101027_AABootLoader
| constants: OS=linux, ARCH=x86, WS=gtk, NL=enFramework arguments: -toolId
| install -accessRights admin input @osgi.install.area/install.xmlCommand-
| line arguments: -os linux -ws gtk -arch x86 -toolId install -accessRights
| admin input @osgi.install.area/install.xml!ENTRY com.ibm.cic.agent.ui 4 0
| 2013-07-09 14:11:47.692!MESSAGE Could not load SWT library.
| Reasons:/home/tklm-v3/disk1/im/configuration/org.eclipse.osgi/bundles/207/1/
| .cp/libswt-pi-gtk-4234.so (libgtk-x11-2.0.so.0: cannot open shared object
| file: No such file or directory)
| swt-pi-gtk (Not found in java.library.path)/root/.swt/lib/linux/x86/libswt-
| pi-gtk-4234.so (libgtk-x11-2.0.so.0: cannot open shared object file: No
| such file or directory)
| /root/.swt/lib/linux/x86/libswt-pi-gtk.so (/root/.swt/lib/linux/x86/liblib
| swt-pi-gtk.so.so:cannot open shared object file: No such file or directory)"
```

### | Cause

| The system displays this error when the system packages are not upgraded.

### | Proposed workaround

- | • All system packages must be upgraded, except the kernel that should be 6.3 in order for encryption to work correctly.
- | • Run through the following checklist before installing IBM Security Lifecycle Manager:

| *Table 52.*

System components	Minimum values	Header
System memory (RAM)	4 GB	4 GB
Processor speed	Linux and Windows systems	Linux and Windows systems 3.0 GHz dual processors AIX and Sun Solaris systems 1.5 GHz (4-way)
Disk space free for IBM Security Key Lifecycle Manager and prerequisite products such as DB2®	3.0 GHz single processor AIX and Sun Solaris systems 1.5 GHz (2-way)	5 GB
Disk space free in /tmp or C:\temp	5 GB	2 GB

Table 52. (continued)

System components	Minimum values	Header
Disk space free in /home directory for DB2	2 GB	6 GB
Disk space free in /var directory for DB2	5 GB 512 MB on Linux and UNIX operating systems	512 MB on Linux and UNIX operating systems

## The IBM Security Lifecycle Manager prerequisites cannot be installed

This topic provides troubleshooting references and steps for resolving system errors when the IBM Security Lifecycle Manager prerequisites cannot be installed.

### Description

When the user tries to install the IBM Security Lifecycle Manager prerequisites, the system displays the following error:

```
JVMJ9VM011W Unable to load j9dmp24: libstdc++.so.5: cannot open shared
object file: No such file or directory
JVMJ9VM011W Unable to load j9jit24: libstdc++.so.5: cannot open shared
object file: No such file or directory
JVMJ9VM011W Unable to load j9gc24: libstdc++.so.5: cannot open shared
object file: No such file or directory
JVMJ9VM011W Unable to load j9vrb24: libstdc++.so.5: cannot open shared
object file: No such file or directory
```

### Cause

The system displays this error when the system packages are not upgraded.

### Proposed workaround

- All system packages must be upgraded, except the kernel that should be 6.3 in order for encryption to work correctly.
- Update all packages excluding kernel:
 

```
yum update --exclude=kernel*
```
- Modify: /etc/yum.conf
 

```
[main]
...
exclude=kernel* redhat-release*
```

## The error code signature does not match

This topic provides troubleshooting references and steps for resolving system errors when the error code signature does not match.

### Description

When there is an error code signature mismatch, the system displays the following error:

```
<?xml version="1.0" encoding="UTF-8"?><Error> <Code>SignatureDoesNotMatch</Code> <Message>The
request signature wecalculated does not match the signature you provided. Check your keyand
signing method.</Message> <RequestId>tx48ae6acd398044b5b1ebd-005637c767</RequestId></Error>
```



## E! Cause

E! The system displays this error when the specified user ID does not exist and the user ID does not have E! the defined credentials or has not assigned a role to the account.

## E! Proposed workaround

- E! • For role assignments, review the output of these commands to identify the role assignment for the E! affected user:
  - E! – OpenStack user list
  - E! – OpenStack role assignment list
  - E! – OpenStack role list
  - E! – OpenStack project list
- E! • For credential issues, review the credentials assigned to that user id:
  - E! – OpenStack credential list
  - E! – OpenStack credential show <ID>

## E! Swift PUT returns the 202 error and S3 PUT returns the 500 error due E! to the missing time synchronization

E! This topic provides troubleshooting references and steps for resolving system errors when Swift PUT E! returns the 202 error and S3 PUT returns the 500 error due to the missing time synchronization.

## E! Description

E! The swift object servers require monotonically-increasing timestamps on the PUT requests. If the time E! between all the nodes is not synchronized, the PUT request can be rejected, resulting in the object server E! returning a 409 status code that is turned into 202 in the proxy-server. When the swift3 middleware E! receives the 202 code, it returns a 500 to the client. When enabling DEBUG logging, the system displays E! the following message:

E! From the object server:

```
E! Feb 9 14:41:09 prt001st001 object-server: 10.0.5.6 - - [09/Feb/2016:21:41:09 +0000] "PUT
E! /z1device119/14886/AUTH_bfd953e691c4481d8fa0249173870a56/mycontainers12/myobjects407"
```

E! From the proxy server:

```
E! Feb 9 14:14:10 prt003st001 proxy-server: Object PUT returning 202 for 409: 1455052450.83619 <=
E! '409 (1455052458.12105)' (txn: txf7611c330872416aabcc1-0056ba56a2) (client_ip:
```

E! If S3 is used, the following error is displayed from Swift3:

```
E! Feb 9 14:25:52 prt005st001 proxy-server: 500 Internal Server Error: #012Traceback (most recent
E! call last):#012 File "/usr/lib/python2.7/site-packages/swift3/middleware.py", line 81, in
E! __call__#012 resp = self.handle_request(req)#012 File "/usr/lib/python2.7/site-packages/swift3/
E! middleware.py", line 104, in handle_request#012 res = getattr(controller, req.method)(req)#012
E! File "/usr/lib/python2.7/site-packages/swift3/controllers/obj.py", line 97, in PUT#012 resp =
E! req.get_response(self.app)#012 File "/usr/lib/python2.7/site-packages/swift3/request.py", line
E! 825, in get_response#012 headers, body, query)#012 File "/usr/lib/python2.7/site-packages/
E! swift3/request.py", line 805, in get_acl_response#012 app, method, container, obj, headers,
E! body, query)#012 File "/usr/lib/python2.7/site-packages/swift3/request.py", line 669, in
E! _get_response#012 raise InternalError('unexpected status code %d' % status)#012InternalError:
E! 500 Internal Server Error (txn: tx40d4ff7ca5b94b1bb6881-0056ba5960) (client_ip: 10.0.5.1) Feb 9
E! 14:25:52 prt005st001 proxy-server: 500 Internal Server Error: #012Traceback (most recent call
```

```
E! last):#012 File "/usr/lib/python2.7/site-packages/swift3/middleware.py", line 81, in
E! __call__#012 resp = self.handle_request(req)#012 File "/usr/lib/python2.7/site-packages/swift3/
E! middleware.py", line 104, in handle_request#012 res = getattr(controller, req.method)(req)#012
E! File "/usr/lib/python2.7/site-packages/swift3/controllers/obj.py", line 97, in PUT#012 resp =
E! req.get_response(self.app)#012 File "/usr/lib/python2.7/site-packages/swift3/request.py", line
E! 825, in get_response#012 headers, body, query)#012 File "/usr/lib/python2.7/site-packages/
E! swift3/request.py", line 805, in get_acl_response#012 app, method, container, obj, headers,
E! body, query)#012 File "/usr/lib/python2.7/site-packages/swift3/request.py", line 669, in
E! _get_response#012 raise InternalError('unexpected status code %d' % status)#012InternalError:
E! 500 Internal Server Error (txn: tx40d4ff7ca5b94b1bb6881-0056ba5960) (client_ip: 10.0.5.1)
```

## E! **Cause**

E! The system displays these errors when the time is not in sync.

## E! **Proposed workaround**

- E! • To check for this problem, run the **mmdsh date** command.
- E! • Enable the NTPD service on all protocol nodes and have the time synchronized from an NTP server.

## E! **Unable to generate the accurate container listing by performing the GET operation for unified file and object access container**

E! This topic provides troubleshooting references and steps for resolving system errors when the system is unable to generate the accurate container listing by performing the GET operation for unified file and object access container.

## E! **Description**

E! The system does not display the accurate container listing for a unified file and object access enabled container.

## E! **Cause**

E! This error occurs under one or all of the following conditions:

- E! • The `ibmObjectizer` list too long. Therefore, a longer time is taken to update and display the listing.
- E! • The files created on the file system are not supported for objectization.

## E! **Proposed workaround**

E! Tune the `ibmobjectizer` interval configuration by running the **mmobj config change** command.

E! The following is an example of setting up the objectization interval by using the **mmobj config change**:

```
E! mmobj config change --ccrfile spectrum-scale-objectizer.conf \  
E! --section DEFAULT --property objectization_interval --value 2400
```

E! This command sets an interval of 40 minutes between the completion of an objectization cycle and the start of the next cycle.

## | **Fatal error of object configuration during deployment**

| This topic provides troubleshooting references and steps for resolving fatal system errors in object configuration during deployment.

## | **Description**

| When the user enables object by using **spectrumscale** installation toolkit, the system displays the following error:

```
| [ FATAL ] Required option 'endpoint_hostname' missing in section:  
| 'object'. To set this, use: ./spectrumscale config object -endpoint  
|  
| [ FATAL ] Invalid configuration for setting up Object Store.
```

## | **Cause**

| The system displays this error when the object authentication not completed with the required parameters.

## | **Proposed workaround**

| Run the **spectrumscale config obj** command with the mandatory arguments.

## | **Object authentication configuration fatal error during deployment**

| This topic provides troubleshooting references and steps for resolving fatal system errors in object authentication configuration during deployment.

## | **Description**

| When the user configures the authentication by using the **spectrumscale** installation toolkit, the system displays the following error:

```
| 2016-02-16 13:48:07,799 [ FATAL ] <nodename> failure whilst: Configuring object authentication  
| (SS98)
```

## | **Cause**

| The system displays this error under one or all of the following conditions:

- | • Only the users under the specified user DN are visible to Keystone.
- | • The users do not have the specified object class.

## | **Proposed workaround**

| Change the object authentication or modify the AD/LDAP for the users who do not have the specified object class.

## | **Fatal error of object authentication during deployment**

| This topic provides troubleshooting references and steps for resolving fatal errors in object authentication during deployment.

## | **Description**

| When the user configures authentication by using **spectrumscale** installation toolkit, the system displays the following error:

```
| 2016-02-16 13:48:07,799 [ FATAL ] <nodename> failure whilst: Configuring object authentication  
| (SS98)
```

## | **Cause**

| The system displays this error under one or all of the following conditions:

- | • IBM Spectrum Scale for the object storage program is currently running.
- | • Parameters provided in the configuration.txt and authconfig.txt files are incorrect.
- | • The system is unable to connect to the authentication server with the given credentials or network issues.

## | **Proposed workaround**

- | • Shut down IBM Spectrum Scale for the object storage program before continuing.
- | • Check the connectivity of protocol nodes with the authentication server with valid credentials.
- | • Stop the service manually with the **mmces service stop obj -a** command. Manually run the **mmuserauth service create** command to complete the authentication configuration for object.
- | • Fix the configuration.txt and authconfig.txt files and rerun the IBM Spectrum Scale deployment with the **spectrumscale deploy** command.

---

## Chapter 19. Disaster recovery issues

As with any type of problem or failure, obtain the GPFS log files (**mmfs.log.\***) from all nodes in the cluster and, if available, the content of the internal dumps.

For more information, see:

- The *Data mirroring and replication* topic in the *IBM Spectrum Scale: Administration Guide* for detailed information about GPFS disaster recovery
- “Creating a master GPFS log file” on page 141
- “Information to be collected before contacting the IBM Support Center” on page 367

The following two messages might appear in the GPFS log for active/active disaster recovery scenarios with GPFS replication. The purpose of these messages is to record quorum override decisions that are made after the loss of most of the disks:

6027-435 [N]

The file system descriptor quorum has been overridden.

6027-490 [N]

The descriptor replica on disk *diskName* has been excluded.

A message similar to these appear in the log on the file system manager, node every time it reads the file system descriptor with an overridden quorum:

```
...
6027-435 [N] The file system descriptor quorum has been overridden.
6027-490 [N] The descriptor replica on disk gpfs23nsd has been excluded.
6027-490 [N] The descriptor replica on disk gpfs24nsd has been excluded.
...
```

For more information on node override, see the section on *Quorum*, in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*

For PPRC and FlashCopy®-based configurations, more problem determination information can be collected from the ESS log file. This information and the appropriate ESS documentation must be referred while working with various types disk subsystem-related failures. For instance, if users are unable to perform a PPRC failover (or failback) task successfully or unable to generate a FlashCopy of a disk volume, they should consult the subsystem log and the appropriate ESS documentation. For more information, see the following topics:

- *IBM Enterprise Storage Server*® ([www.redbooks.ibm.com/redbooks/pdfs/sg245465.pdf](http://www.redbooks.ibm.com/redbooks/pdfs/sg245465.pdf))
- *IBM TotalStorage Enterprise Storage Server Web Interface User's Guide*([publibfp.boulder.ibm.com/epubs/pdf/f2bui05.pdf](http://publibfp.boulder.ibm.com/epubs/pdf/f2bui05.pdf)).

---

### Disaster recovery setup problems

The following setup problems might impact disaster recovery implementation:

1. Considerations of data integrity require proper setup of PPRC consistency groups in PPRC environments. Additionally, when using the FlashCopy facility, make sure to suspend all I/O activity before generating the FlashCopy image. See “Data integrity” on page 273.
2. In certain cases, it might not be possible to restore access to the file system even after relaxing the node and disk quorums. For example, in a three failure group configuration, GPFS tolerates and recovers from a complete loss of a single failure group (and the tiebreaker with a quorum override). However, all disks in the remaining failure group must remain active and usable in order for the file

system to continue its operation. A subsequent loss of at least one of the disks in the remaining failure group would render the file system unusable and trigger a forced unmount. In such situations, users might still be able to perform a restricted mount and attempt to recover parts of their data from the damaged file system. For more information on restricted mounts, see “Restricted mode mount” on page 192.

3. When you issue **mmfsctl syncFSconfig**, you might get an error similar to the following:

```
mmfsctl: None of the nodes in the peer cluster can be reached
```

In such scenarios, check the network connectivity between the peer GPFS clusters and verify their remote shell setup. This command requires full TCP/IP connectivity between the two sites, and all nodes must be able to communicate by using ssh or rsh without the use of a password.

---

## Protocols cluster disaster recovery issues

| Sometimes issuing an **mmcesdr** command can cause problems with protocols disaster recovery in IBM Spectrum Scale.

| Whenever such an error or problem is encountered, view the Protocols DR log file for more information on the issue. This log file is at `/var/adm/ras/mmcesdr.log` on the node where the command was run.

---

## Other problems with disaster recovery

You might encounter the following issues that are related to disaster recovery in IBM Spectrum Scale:

1. Currently, users are advised to always specify the **all** option when you issue the **mmfsctl syncFSconfig** command, rather than the device name of one specific file system. Issuing this command enables GPFS to detect and correctly resolve the configuration discrepancies that might occur as a result of the manual administrative action in the target GPFS cluster to which the configuration is imported.
2. The optional **SpecFile** parameter to the **mmfsctl syncFSconfig** that is specified with the **-S** flag must be a fully qualified path name that defines the location of the spec data file on nodes in the target cluster. It is not the local path name to the file on the node from which the **mmfsctl** command is issued. A copy of this file must be available at the provided path name on all peer contact nodes that are defined in the **RemoteNodesFile**.

---

## Chapter 20. Performance issues

The performance issues might occur because of the system components or configuration or maintenance issues.

---

### Issues caused by the low-level system components

This section discusses the issues caused by the low-level system components used in the IBM Spectrum Scale cluster.

### Suboptimal performance due to high utilization of the system level components

In some cases, the CPU or memory utilization on an IBM Spectrum Scale node is higher than 90%. Such heavy utilization can adversely impact the system performance as it affects the cycles allocated to the IBM Spectrum Scale daemon service.

#### Problem identification

On the node, issue an Operating System command such as **top** or **dstat** to verify whether the system level resource utilization is higher than 90%. The following example shows the sample output for the **dstat** command:

```
# dstat 1 10
```

```
----total-cpu-usage---- -dsk/total- -net/total- ---paging-- ---system--
usr sys idl wai hiq siq | read writ | recv send | in out | int csw
| 0 0 100 0 0 0 | 7308k 9236k | 0 0 | 0 0 | 812 3691
| 0 0 100 0 0 0 | 0 0 | 3977B 1038B | 0 0 | 183 317
| 1 2 98 0 0 0 | 0 0 | 2541B 446B | 0 0 | 809 586
| 0 1 99 0 0 0 | 0 0 | 4252B 346B | 0 0 | 427 405
| 0 0 100 0 0 0 | 0 0 | 3880B 346B | 0 0 | 196 349
| 0 0 100 0 0 0 | 0 0 | 3594B 446B | 0 0 | 173 320
| 1 1 98 0 0 0 | 0 0 | 3969B 446B | 0 0 | 692 662
| 0 0 100 0 0 0 | 0 116k | 3120B 346B | 0 0 | 189 312
| 0 0 100 0 0 0 | 0 0 | 3050B 346B | 0 0 | 209 342
| 0 0 100 0 0 0 | 0 4096B | 4555B 346B | 0 0 | 256 376
| 0 0 100 0 0 0 | 0 0 | 3232B 346B | 0 0 | 187 340
```

#### Problem resolution and verification

If the system level resource utilization is high, determine the process or application that contributes to the performance issue and take appropriate action to minimize the utilization to an acceptable level.

### Suboptimal performance due to long IBM Spectrum Scale waiters

Low-level system issues, like slow disks, or slow network, might cause long GPFS waiters. These long waiters cause performance degradation. You can use the **mmdiag --waiters** command to display the **mmfsd** threads waiting for events. This information can help resolve deadlocks and improve the system performance.

#### Problem identification

On the node, issue the **mmdiag --waiters** command to check whether any long waiters are present. The following example shows long waiters that are contributed by the slow disk, **dm-14**:







## Suboptimal performance due to low values assigned to IBM Spectrum Scale configuration parameters

Most GPFS configuration parameters have default values. For example, in IBM Spectrum Scale version 4.2 and above, pagepool parameter defaults to either one-third of the physical memory on the node or 1 GiB (whichever is smaller), maxMBpS defaults to 2048 and maxFilesToCache defaults to 4000. However, if the IBM Spectrum Scale configuration parameters are explicitly set to values lower than their default values by the user, it can impact the I/O performance.

### Problem identification

On the GPFS node, issue the `mmdiag --config` command to display and verify the values of the GPFS configuration parameters. Check whether these values match the optimal values set for IBM Spectrum Scale system configuration. For more information on optimal values for configuration parameter see Tuning Parameters.

### Problem resolution and verification

Issue the `mmchconfig Attribute=value -i` command to set the configuration parameters to their optimal values based on the best practices followed for an IBM Spectrum Scale system configuration.

You might need to restart GPFS for certain configuration parameter values to take effect. Issue the `mmshutdown` command, followed by the `mmstartup` command to restart GPFS. Issue the `mmdiag --config` command to verify the configuration changes and updates.

## Suboptimal performance due to new nodes with default parameter values added to the cluster

When new nodes are added to the IBM Spectrum Scale cluster, ensure that the GPFS configuration parameter values on the new nodes are not set to default values, unless explicitly set so by the user based on the GPFS node class. Instead, the GPFS configuration parameter values on the new nodes must be similar to the values of the existing nodes of similar type for optimal performance. The necessary system level component settings, like BIOS, network and others on the new nodes, also need to match the system level component settings of the existing nodes.

### Problem identification

The `mmisconfig` command can be used to display and verify the configuration values for a IBM Spectrum Scale cluster.

Issue the `mmdiag --config` command on the newly added GPFS nodes to verify whether the configuration parameter values for the new nodes are same as values for the existing nodes. If the newly added nodes have special roles or higher capability, then the configuration values must be adjusted accordingly.

Certain applications like SAS benefit from a larger GPFS pagepool. The GPFS pagepool is used to cache user file data and file system metadata. The default value of the GPFS pagepool parameter is 1 GiB in GPFS version 3.5 and higher. For SAS application, a minimum of 4 GiB pagepool is recommended. When new SAS application nodes are added to the IBM Spectrum Scale cluster, ensure that the pagepool is set to at least 4 GiB. If left to its default value, the pagepool value is set to 1 GiB. This negatively impacts the application performance.

### Problem resolution and verification

Issue the `mmchconfig Attribute=value -N <new_nodes> -i` command to set the configuration parameters to either their optimal values based on the best practices, or values similar to the existing nodes. It might be necessary to restart the GPFS daemon for the values to take effect. Issue the `mmshutdown` command,

| followed by the **mmstartup** command to restart the GPFS daemon. Verify the changes by running the  
| **mm1sconfig** on a node that is part of the GPFS cluster, and the **mmdiag --config** command on the new  
| nodes.

| The following sample output shows that the value for the pagepool on the existing application nodes  
| **c25m3n03-ib** and **c25m3n04-ib** is set to 2G.

| **Note:** Here **Application node** refers to NSD or SAN GPFS client nodes where applications are executed.  
| These nodes have GPFS RPM installed for good performance.

```
| #mm1sconfig  
| [c25m3n03-ib,c25m3n04-ib]  
| pagepool 2G
```

| If you add new application nodes **c25m3n05-ib** and **c25m3n06-ib** to the cluster, the pagepool and other  
| GPFS parameter values for the new node must be set according to the corresponding parameter values  
| for the existing nodes **c25m3n03-ib** and **c25m3n04-ib**. Therefore, the pagepool on these new nodes must  
| also be set to 2G by using the **mmchconfig** command.

```
| mmchconfig pagepool=2G -N c25m3n05-ib,c25m3n06-ib -i
```

| **Note:** The **-i** option specifies that the changes take effect immediately and are permanent. This option is  
| valid only for certain attributes. For more information on block allocation, see the *mmchconfig* command in  
| the *IBM Spectrum Scale: Command and Programming Reference*.

| Issue the **mm1sconfig** command to verify whether all the nodes have similar values. The following sample  
| output shows that all the nodes have pagepool value set to 2G:

```
| [c25m3n03-ib,c25m3n04-ib,c25m3n05-ib,c25m3n06-ib]  
| pagepool 2G
```

| **Note:** If the pagepool is set to a custom value (2G for this example), then the pagepool value is listed  
| when you issue the **mm1sconfig** command. If the pagepool is set to a default value (1G) then this will be  
| listed when you issue the **mm1sconfig pagepool** command.

| On the new node, issue the **mmdiag --config** command to verify that the new values are in effect. The  
| sample output displays that the pagepool value has been effectively set to 2G for the nodes **c25m3n03-ib**,  
| **c25m3n04-ib**,**c25m3n05-ib**, **c25m3n06-ib**:

```
| ! pagepool 2147483648
```

| **Note:** The exclamation (!) in the front of the parameter denotes that the value of this parameter was set  
| by the user, and is not the default value for the parameter.

## | **Suboptimal performance due to low value assigned to QoSIO | operation classes**

| If Quality of Service for I/O (QoSIO) feature is enabled on the file system, verify whether any of the  
| storage pools are assigned low values for **other** and **maintenance** class. Assigning low values for **other**  
| and **maintenance** class can impact the performance when I/O is performed on that specific storage pool.

### | **Problem identification**

| On the GPFS node, issue the **mm1sqos <fs>** command and check the **other** and **maintenance** class settings.  
| In the sample output below, the **maintenance** class IOPS for **datapool1** storage pool is set to 200 IOPS,  
| and the **other** class IOPS for **datapool2** storage pool is set to 400 IOPS. This IOPS value might be low for  
| an environment with high performing storage subsystem.

```
| # mm1sqos gpfs1b
```

```

QoS config::      enabled -- pool=*,other=inf,maintenance=inf:pool=datapool1,other=inf,
maintenance=200Iops:pool=datapool2,other=400Iops,maintenance=inf
QoS values::     pool=system,other=inf,maintenance=inf:pool=datapool1,other=inf,
maintenance=200Iops:pool=datapool2,other=400Iops,maintenance=inf
QoS status::     throttling active, monitoring active

```

## Problem resolution and verification

On the GPFS node, issue the **mmchqos** command to change the QoS values for a storage pool in the file system. Issue the **mmqsos** command to verify whether the changes are reflected in the QoS settings.

For example, if the IOPS corresponding to datapool2 **other** class must be set to unlimited then issue the following command.

```
mmchqos gpfs1b --enable pool=datapool2,other=unlimited
```

Issue the **# mmqsos gpfs1b** command to verify whether the change is reflected.

```
# mmqsos gpfs1b
```

```

QoS config::      enabled -- pool=*,other=inf,maintenance=inf:pool=datapool1,
other=inf,maintenance=200Iops:pool=datapool2,
other=inf,maintenance=inf
QoS values::     pool=system,other=inf,maintenance=inf:pool=datapool1,
other=inf,maintenance=200Iops:pool=datapool2,
other=inf,maintenance=inf
QoS status::     throttling active, monitoring active

```

## Suboptimal performance due to improper mapping of the file system NSDs to the NSD servers

The NSDs in a file system need to be optimally assigned to the NSD servers so that the client I/O is equally distributed across all the NSD servers. For example, consider a file system with 10 NSDs and 2 NSD servers. The NSD-to-server mapping must be done in such a way that each server acts as the **primary** server for 5 of the NSD in the file system. If the NSD-to-server mapping is unbalanced, it can result in hot spots in one or more of the NSD servers. Presence of hot spots within a system can cause performance degradation.

### Problem identification

Issue the **mmnsd** command, and verify that the primary NSD server allocated to a file system is evenly distributed.

**Note:** The primary server is the first server listed under the **NSD server** column for a particular file system.

On the NSD client, issue the **mmldisk <fs> -m** command to ensure that the NSD client I/O is distributed evenly across all the NSD servers.

In the following sample output, all the NSDs are assigned to the same primary server **c80f1m5n03ib0**.

```
# mmnsd
```

File system	Disk name	NSD servers
gpfs2	Perf2a_NSD01	c80f1m5n03ib0,c80f1m5n02ib0
gpfs2	Perf2a_NSD02	c80f1m5n03ib0,c80f1m5n02ib0
gpfs2	Perf2a_NSD03	c80f1m5n03ib0,c80f1m5n02ib0
gpfs2	Perf2a_NSD04	c80f1m5n03ib0,c80f1m5n02ib0
gpfs2	Perf2a_NSD05	c80f1m5n03ib0,c80f1m5n02ib0
gpfs2	Perf2a_NSD06	c80f1m5n03ib0,c80f1m5n02ib0
gpfs2	Perf2a_NSD07	c80f1m5n03ib0,c80f1m5n02ib0

```

| gpfs2      Perf2a_NSD08 c80f1m5n03ib0,c80f1m5n02ib0
| gpfs2      Perf2a_NSD09 c80f1m5n03ib0,c80f1m5n02ib0
| gpfs2      Perf2a_NSD10 c80f1m5n03ib0,c80f1m5n02ib0

```

In this case, all the NSD client I/O for the **gpfs2** file system are processed by the single NSD server **c80f1m5n03ib0**, instead of being equally distributed across both the NSD servers **c80f1m5n02ib0** and **c80f1m5n03ib0**. This can be verified by issuing the **mm1sdisk <fs> -m** command on the NSD client, as shown in the following sample output:

```

| # mm1sdisk gpfs2 -m
| Disk name      IO performed on node      Device                      Availability
| -----
| Perf2a_NSD01   c80f1m5n03ib0            -                            up
| Perf2a_NSD02   c80f1m5n03ib0            -                            up
| Perf2a_NSD03   c80f1m5n03ib0            -                            up
| Perf2a_NSD04   c80f1m5n03ib0            -                            up
| Perf2a_NSD05   c80f1m5n03ib0            -                            up
| Perf2a_NSD06   c80f1m5n03ib0            -                            up
| Perf2a_NSD07   c80f1m5n03ib0            -                            up
| Perf2a_NSD08   c80f1m5n03ib0            -                            up
| Perf2a_NSD09   c80f1m5n03ib0            -                            up
| Perf2a_NSD10   c80f1m5n03ib0            -                            up

```

## Problem resolution and verification

If the NSD-to-primary mapping is unbalanced, issue the **mmchnsd** command to balance the NSD distribution across the NSD servers. Issue the **mm1snsd** command or the **mm1sdisk <fs> -m** command on the NSD client to ensure that the NSD distribution across the servers is balanced.

In the following sample output, there are 10 NSDs in the **gpfs2** file system. The NSDs are evenly distributed between the two servers, such that both servers, **c80f1m5n03ib0** and **c80f1m5n02ib0** act as primary servers for 5NSDs each.

```

| # mm1snsd
| File system    Disk name      NSD servers
| -----
| gpfs2          Perf2a_NSD01   c80f1m5n03ib0,c80f1m5n02ib0
| gpfs2          Perf2a_NSD02   c80f1m5n02ib0,c80f1m5n03ib0
| gpfs2          Perf2a_NSD03   c80f1m5n03ib0,c80f1m5n02ib0
| gpfs2          Perf2a_NSD04   c80f1m5n02ib0,c80f1m5n03ib0
| gpfs2          Perf2a_NSD05   c80f1m5n03ib0,c80f1m5n02ib0
| gpfs2          Perf2a_NSD06   c80f1m5n02ib0,c80f1m5n03ib0
| gpfs2          Perf2a_NSD07   c80f1m5n03ib0,c80f1m5n02ib0
| gpfs2          Perf2a_NSD08   c80f1m5n02ib0,c80f1m5n03ib0
| gpfs2          Perf2a_NSD09   c80f1m5n03ib0,c80f1m5n02ib0
| gpfs2          Perf2a_NSD10   c80f1m5n02ib0,c80f1m5n03ib0

```

The NSD client I/O is also evenly distributed across the two NSD servers, as seen in the following sample output:

```

| # mm1sdisk gpfs2 -m
| Disk name      IO performed on node      Device                      Availability
| -----
| Perf2a_NSD01   c80f1m5n03ib0            -                            up
| Perf2a_NSD02   c80f1m5n02ib0            -                            up
| Perf2a_NSD03   c80f1m5n03ib0            -                            up
| Perf2a_NSD04   c80f1m5n02ib0            -                            up
| Perf2a_NSD05   c80f1m5n03ib0            -                            up
| Perf2a_NSD06   c80f1m5n02ib0            -                            up
| Perf2a_NSD07   c80f1m5n03ib0            -                            up

```

```

| Perf2a_NSD08 c80f1m5n02ib0 - up
| Perf2a_NSD09 c80f1m5n03ib0 - up
| Perf2a_NSD10 c80f1m5n02ib0 - up

```

## | Suboptimal performance due to incompatible file system block allocation type

| In some cases, proof-of-concept (POC) is done on a smaller setup that consists of clusters with eight or fewer nodes and file system with eight or fewer disks. When the necessary performance requirements are met, the production file system is deployed on a larger cluster and storage setup. It is possible that on a larger cluster, the file performance per NSD is less compared to the smaller POC setup, even if all the cluster and storage component are healthy and performing optimally. In such cases, it is likely that the file system is configured with the default **cluster** block allocation type during the smaller POC setup and the larger file system setup are configured with **scatter** block allocation type.

### | Problem identification

| Issue the **mm1sfs** command to verify the block allocation type that is in effect on the smaller and larger setup file system.

| In the sample output below, the **Block allocation type** for the **gpfs2** file system is set to **scatter**.

```

| # mm1sfs gpfs2 | grep 'Block allocation type'
| -j scatter Block allocation type

```

### | Problem resolution and verification

| `layoutMap={scatter|cluster}` specifies the block allocation map type. When allocating blocks for a file, GPFS first uses a round-robin algorithm to spread the data across all disks in the storage pool. After a disk is selected, the location of the data block on the disk is determined by the block allocation map type.

| For cluster block allocation map type, GPFS attempts to allocate blocks in clusters. Blocks that belong to a particular file are kept adjacent to each other within each cluster. For scatter block allocation map type, the location of the block is chosen randomly. For production setup, where performance consistency throughout the life time of the file system is paramount, scatter block allocation type is recommended. The IBM Spectrum Scale storage I/O performance sizing also needs to be performed by using the scatter block allocation.

| The cluster allocation method might provide better disk performance for some disk subsystems in relatively small installations. However, the benefits of clustered block allocation diminish when the number of nodes in the cluster or the number of disks in a file system increases, or when the file system's free space becomes fragmented. The cluster allocation is the default allocation method for GPFS clusters with eight or fewer nodes and for file systems with eight or fewer disks.

| The scatter allocation method provides more consistent file system performance by averaging out performance variations. This is so because for many disk subsystems, the location of the data relative to the disk edge has a substantial effect on the performance. This allocation method is appropriate in most cases and is the default allocation type for GPFS clusters with more than eight nodes or file systems with more than eight disks.

| The block allocation map type cannot be change after the storage pool is created. For more information on block allocation, see the *mmcrfs* command in the *IBM Spectrum Scale: Command and Programming Reference*.

| **Attention:** Scatter block allocation is recommended for a production setup where performance consistency is paramount throughout the lifetime of the file system. However, in an FPO environments (Hadoop or Big Data), cluster block allocation is recommended.

---

## | Issues caused by the unhealthy state of the components used

| This section discusses the issues caused due to the unhealthy state of the components used in the IBM Spectrum Scale stack

### | Suboptimal performance due to failover of NSDs to secondary server - NSD server failure

| In a shared storage configuration, failure of an NSD server might result in the failover of its NSDs to the secondary server, if the secondary server is active. This can reduce the total number of NSD servers actively serving the file system, which in turn impacts the file system's performance.

#### | Problem identification

| In IBM Spectrum Scale, the system-defined node class "nsdnodes" contains all the NSD server nodes in the IBM Spectrum Scale cluster. Issue the `mmgetstate -N nsdnodes` command to verify the state of the GPFS daemon. The GPFS file system performance might degrade if one or more NSD servers are in the **down** or **arbitrating** or **unknown** state.

| The following example displays two nodes: one in **active** state and the other in **down** state

```
| # mmgetstate -N nsdnodes
| Node number  Node name      GPFS state
| -----
|              1      c25m3n07-ib   active
|              2      c25m3n08-ib   down
```

#### | Problem resolution and verification

| Resolve any system-level or software issues that exist. For example, confirm that NSD server have no network connectivity problems, or that the GPFS portability modules are correctly built for the kernel that is running. Also, perform necessary low-level tests to ensure that both the NSD server and the communication to the node are healthy and stable.

| Verify that no system or software issues exist, and start GPFS on the NSD server by using the `mmstartup -N <NSD_server_to_revive>` command. Use the `mmgetstate -N nsdnodes` command to verify that the GPFS daemon is in active state as shown:

```
| # mmgetstate -N nsdnodes
| Node number  Node name      GPFS state
| -----
|              1      c25m3n07-ib   active
|              2      c25m3n08-ib   active
```

### | Suboptimal performance due to failover of NSDs to secondary server - Disk connectivity failure

| In a shared storage configuration, disk connectivity failure on an NSD server might result in failover of its NSDs to the secondary server, if the secondary server is active. This can reduce the total number of NSD servers actively serving the file system, which in turn impacts the overall performance of the file system.

## Problem identification

The `mm1snsd` command displays information about the currently defined disks in a cluster. In the following sample output, the NSD client is configured to perform file system I/O on the primary NSD server `c25m3n07-ib` for odd-numbered NSDs like `DMD_NSD01`, `DMD_NSD03`. In this case, `c25m3n08-ib` acts as a secondary server.

The NSD client is configured to perform file system I/O on the NSD server `c25m3n08-ib` for even-numbered NSDs like `DMD_NSD02`, `DMD_NSD04`. In this case, `c25m3n08-ib` is the primary server, while `c25m3n07-ib` acts as the secondary server.

Issue the `#mm1snsd` command to display the NSD server information for the disks in a file system. The following sample output shows the various disks in the `gpfs1b` file system and the NSD servers that are supposed to act as primary and secondary servers for these disks.

```
# mm1snsd
File system  Disk name  NSD servers
-----
gpfs1b      DMD_NSD01  c25m3n07-ib,c25m3n08-ib
gpfs1b      DMD_NSD02  c25m3n08-ib,c25m3n07-ib
gpfs1b      DMD_NSD03  c25m3n07-ib,c25m3n08-ib
gpfs1b      DMD_NSD04  c25m3n08-ib,c25m3n07-ib
gpfs1b      DMD_NSD05  c25m3n07-ib,c25m3n08-ib
gpfs1b      DMD_NSD06  c25m3n08-ib,c25m3n07-ib
gpfs1b      DMD_NSD07  c25m3n07-ib,c25m3n08-ib
gpfs1b      DMD_NSD08  c25m3n08-ib,c25m3n07-ib
gpfs1b      DMD_NSD09  c25m3n07-ib,c25m3n08-ib
gpfs1b      DMD_NSD10  c25m3n08-ib,c25m3n07-ib
```

However, the `mm1sdisk <fsdevice> -m` command that is issued on the NSD client indicates that the NSD client is currently performing all the file system I/O on a single NSD server, `c25m3n07-ib`.

```
# mm1sdisk <fsdevice> -m
Disk name    IO performed on node  Device  Availability
-----
DMD_NSD01   c25m3n07-ib          -        up
DMD_NSD02   c25m3n07-ib          -        up
DMD_NSD03   c25m3n07-ib          -        up
DMD_NSD04   c25m3n07-ib          -        up
DMD_NSD05   c25m3n07-ib          -        up
DMD_NSD06   c25m3n07-ib          -        up
DMD_NSD07   c25m3n07-ib          -        up
DMD_NSD08   c25m3n07-ib          -        up
DMD_NSD09   c25m3n07-ib          -        up
DMD_NSD10   c25m3n07-ib          -        up
```

## Problem resolution and verification

Resolve any system-level or disk-level software issues that exist. For example, storage connectivity issues on the NSD server, or driver issues. Rediscover the NSD disk paths by using the `mmnsddiscover -a -N all` command. On the NSD client, first issue the `mm1snsd` command to obtain the primary NSD server configured for the NSD pertaining to a file system. The `echo "NSD-Name Primary-NSD-Server"; mm1snsd | grep <fsdevice> | awk` command parses the output that is generated by the `mm1snsd` command and displays the primary NSD server for each of the NSDs. Perform file I/O on the NSD client and issue the `mm1sdisk <fs> -m` command to verify that the NSD client is performing file system I/O by using all the configured NSD servers. On the NSD client, first issue the `mm1snsd` command to obtain the configured primary NSD server for the NSD pertaining to a file system. The `# echo "NSD-Name Primary-NSD-Server"; mm1snsd | grep <fsdevice> | awk` command parses the output that is generated by the `mm1snsd` command and displays the primary NSD server for each of the NSDs.



```
| # echo "NSD-Name Primary-NSD-Server"; mmlsnsd | grep <gpfs1b> | awk -F ',' '{print $1}' | awk
| '{print $2 " " $3}'
| NSD-Name Primary-NSD-Server
| DMD_NSD01 c25m3n07-ib
| DMD_NSD02 c25m3n08-ib
| DMD_NSD03 c25m3n07-ib
| DMD_NSD04 c25m3n08-ib
| DMD_NSD05 c25m3n07-ib
| DMD_NSD06 c25m3n08-ib
| DMD_NSD07 c25m3n07-ib
| DMD_NSD08 c25m3n08-ib
| DMD_NSD09 c25m3n07-ib
| DMD_NSD10 c25m3n08-ib
```

## Suboptimal performance due to file system being fully utilized

As a file system nears full utilization, it becomes difficult to find free space for new blocks. This impacts the performance of the write, append, and create operations.

### Problem identification

On the GPFS node, issue the `mmdf <fs>` command to determine the available space.

```
| # mmdf gpfs1b
| disk          disk size  failure holds   holds          free KB          free KB
| name          in KB      group metadata data           in full blocks  in fragments
| -----
| Disks in storage pool: system (Maximum disk size allowed is 18 TB)
| DMD_NSD01    1756094464    101 Yes    Yes    1732298752 ( 99%)    18688 ( 0%)
| DMD_NSD09    1756094464    101 Yes    Yes    1732296704 ( 99%)    13440 ( 0%)
| DMD_NSD03    1756094464    101 Yes    Yes    1732304896 ( 99%)    17728 ( 0%)
| DMD_NSD07    1756094464    101 Yes    Yes    1732300800 ( 99%)    14272 ( 0%)
| DMD_NSD05    1756094464    101 Yes    Yes    1732298752 ( 99%)    13632 ( 0%)
| DMD_NSD06    1756094464    102 Yes    Yes    1732300800 ( 99%)    13632 ( 0%)
| DMD_NSD04    1756094464    102 Yes    Yes    1732300800 ( 99%)    15360 ( 0%)
| DMD_NSD08    1756094464    102 Yes    Yes    1732294656 ( 99%)    13504 ( 0%)
| DMD_NSD02    1756094464    102 Yes    Yes    1732302848 ( 99%)    18688 ( 0%)
| DMD_NSD10    1756094464    102 Yes    Yes    1732304896 ( 99%)    18560 ( 0%)
| -----
| (pool total)    17560944640                                17323003904 ( 99%)    157504 ( 0%)
| -----
| (total)          17560944640                                17323003904 ( 99%)    157504 ( 0%)
| -----
| Inode Information
| -----
| Number of used inodes:          4048
| Number of free inodes:         497712
| Number of allocated inodes:    501760
| Maximum number of inodes:     17149440
```

The UNIX command `df` also can be used to determine the use percentage (Use%) of a file system. The following sample output displays a file system with 2% capacity used.

```
| # df -h
| Filesystem      Size  Used Avail Use% Mounted on
| /dev/gpfs1b    17T  227G  17T   2% /mnt/gpfs1b
```

## Problem resolution and verification

Use the **mmaddisk** command to add new disks or NSDs to increase the GPFS file system capacity. You can also delete unnecessary files from the file system by using the **rm** command in UNIX environments to free up space.

In the sample output below, the **df -h** and **mmdf** commands show the file system use percentage to be around 2%. This indicates that the file system has sufficient capacity available.

```
# df -h
```

```
Filesystem      Size  Used Avail Use% Mounted on
/dev/gpfs1b    17T  211G  17T   2% /mnt/gpfs1b
```

```
# mmdf gpfs1b
```

```
disk          disk size  failure holds   holds          free KB          free KB
            name              in KB    group metadata data      in full blocks  in fragments
-----
Disks in storage pool: system (Maximum disk size allowed is 18 TB)
DMD_NSD01    1756094464  101 Yes   Yes  1734092800 ( 99%)  12992 ( 0%)
DMD_NSD09    1756094464  101 Yes   Yes  1734094848 ( 99%)  14592 ( 0%)
DMD_NSD03    1756094464  101 Yes   Yes  1734045696 ( 99%)  15360 ( 0%)
DMD_NSD07    1756094464  101 Yes   Yes  1734043648 ( 99%)  10944 ( 0%)
DMD_NSD05    1756094464  101 Yes   Yes  1734053888 ( 99%)  11584 ( 0%)
DMD_NSD06    1756094464  102 Yes   Yes  1734103040 ( 99%)  11584 ( 0%)
DMD_NSD04    1756094464  102 Yes   Yes  1734096896 ( 99%)  10048 ( 0%)
DMD_NSD08    1756094464  102 Yes   Yes  1734053888 ( 99%)  14592 ( 0%)
DMD_NSD02    1756094464  102 Yes   Yes  1734092800 ( 99%)  13504 ( 0%)
DMD_NSD10    1756094464  102 Yes   Yes  1734062080 ( 99%)  13632 ( 0%)
-----
(pool total) 17560944640                                17340739584 ( 99%)  128832 ( 0%)
-----
(total)      17560944640                                17340739584 ( 99%)  128832 ( 0%)
```

```
Inode Information
```

```
-----
Number of used inodes:      4075
Number of free inodes:     497685
Number of allocated inodes: 501760
Maximum number of inodes:  17149440
```

### CAUTION:

Exercise extreme caution when you delete files. Ensure that the files are no longer required for any purpose or are backed up before you delete them.

## Suboptimal performance due to VERBS RDMA being inactive

IBM Spectrum Scale for Linux supports InfiniBand Remote Direct Memory Access (RDMA) using the Verbs API for data transfer between an NSD client and the NSD server. If InfiniBand (IB) VERBS RDMA is enabled on the IBM Spectrum Scale cluster, and if there is drop in the file system performance, verify whether the NSD client nodes are using VERBS RDMA for communication to the NSD server nodes. If the nodes are not using RDMA, then the communication switches to using the GPFS node's TCP/IP interface, which can cause performance degradation.

## Problem identification

Issue the **mmisconfig | grep verbsRdma** command to verify whether VERBS RDMA is enabled on the IBM Spectrum Scale cluster.

```
# mmisconfig | grep verbsRdma
```

```
verbsRdma enable
```

| If VERBS RDMA is enabled, check whether the status of VERBS RDMA on a node is **Started** by running the **mmfsadm test verbs status** command.

| # mmfsadm test verbs status

| VERBS RDMA status: started

| The following sample output shows the various disks in the **gpfs1b** file system and the NSD servers that are supposed to act as primary and secondary servers for these disks.

| # mmlnsd

```
| File system   Disk name     NSD servers
|-----|
| gpfs1b       DMD_NSD01    c25m3n07-ib,c25m3n08-ib
| gpfs1b       DMD_NSD02    c25m3n08-ib,c25m3n07-ib
| gpfs1b       DMD_NSD03    c25m3n07-ib,c25m3n08-ib
| gpfs1b       DMD_NSD04    c25m3n08-ib,c25m3n07-ib
| gpfs1b       DMD_NSD05    c25m3n07-ib,c25m3n08-ib
| gpfs1b       DMD_NSD06    c25m3n08-ib,c25m3n07-ib
| gpfs1b       DMD_NSD07    c25m3n07-ib,c25m3n08-ib
| gpfs1b       DMD_NSD08    c25m3n08-ib,c25m3n07-ib
| gpfs1b       DMD_NSD09    c25m3n07-ib,c25m3n08-ib
| gpfs1b       DMD_NSD10    c25m3n08-ib,c25m3n07-ib
```

| Issue the **mmfsadm test verbs conn** command to verify whether the NSD client node is communicating with all the NSD servers that use VERBS RDMA. In the following sample output, the NSD client node has VERBS RDMA communication active on only one of the two NSD servers.

| # mmfsadm test verbs conn

```
| RDMA Connections between nodes:
| destination idx cook sta cli peak cli RD cli WR cli RD KBcli WR KB srv wait serv RD serv WR serv RD KB serv WR KB vrecv vsend vrecv KB vsend KB
|-----|
| c25m3n07-ib 1 2 RTS 0 24 198 16395 12369 34360606 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

## | Problem resolution

| Resolve any low-level IB RDMA issue like loose IB cables or IB fabric issues. When the low-level RDMA issues are resolved, issue system commands like **ibstat** or **ibv\_devinfo** to verify whether the IB **port state** is **active**. The following system output displays the output for a **ibstat** command issued. In the sample output, the port state for **Port 1** is **Active**, while that for **Port 2** is **Down**.

| # ibstat

```
| CA 'm1x5_0'
|   CA type: MT4113
|   Number of ports: 2
|   Firmware version: 10.100.6440
|   Hardware version: 0
|   Node GUID: 0xe41d2d03001fa210
|   System image GUID: 0xe41d2d03001fa210
|   Port 1:
|     State: Active
|     Physical state: LinkUp
|     Rate: 56
|     Base lid: 29
|     LMC: 0
|     SM lid: 1
|     Capability mask: 0x26516848vverify
|     Port GUID: 0xe41d2d03001fa210
|     Link layer: InfiniBand
|   Port 2:
|     State: Down
|     Physical state: Disabled
|     Rate: 10
```

```

|         Base lid: 65535
|         LMC: 0
|         SM lid: 0
|         Capability mask: 0x26516848
|         Port GUID: 0xe41d2d03001fa218
|         Link layer: InfiniBand

```

Restart GPFS on the node and check whether the status of VERBS RDMA on a node is **Started** by running the **mmfsadm test verbs status** command.

In the following sample output, the NSD client (c25m3n03-ib) and the two NSD servers all show VERBS RDMA status as **started**.

```

| # mmdsh -N nsdnodes,c25m3n03-ib '/usr/lpp/mmfs/bin/mmfsadm test verbs status'
| c25m3n03-ib: VERBS RDMA status: started
| c25m3n07-ib: VERBS RDMA status: started
| c25m3n08-ib: VERBS RDMA status: started

```

Perform a large I/O activity on the NSD client, and issue the **mmfsadm test verbs conn** command to verify whether the NSD client node is communicating with all the NSD servers that use VERBS RDMA.

In the sample output below, the NSD client node has VERBS RDMA communication active on all the active NSD servers.

```

| # mmfsadm test verbs conn
| RDMA Connections between nodes:
| destination  idx cook sta cli peak cli RD  cli WR cli RD KB cli WR KB  srv wait serv RD serv WR  serv RD KB  serv WR KB  vrecv vsend  vrecv KB  vsend KB
|-----
| c25m3n08-ib  0  3  RTS  0 13  8193  8205  17179930 17181212  0  0  0  0  0  0  0  0  0  0  0  0  0
| c25m3n07-ib  1  2  RTS  0 14  8192  8206  17179869 17182162  0  0  0  0  0  0  0  0  0  0  0  0

```

## Issues caused by the use of configurations or commands related to maintenance and operation

This section discusses the issues caused due to the unhealthy state of the components used in the IBM Spectrum Scale stack

### Suboptimal performance due to maintenance commands in progress

When in progress, long-running GPFS maintenance operations like **mmrestripefs**, **mmapplypolicy**, **mmadddisk**, and **mmdeiddisk**, consume some percentage of the system resources. Significant consumption of the system resources can impact the I/O performance of the application.

#### Problem identification

Check the GPFS log file `/var/adm/ras/mmfs.log.latest` on the File System Manager node **mmismgr** to verify whether any GPFS maintenance operations are in progress.

The following sample output shows that the **mmrestripefs** operation was initiated on Jan 19 at 14:32:41, and the operation was successfully completed at 14:45:42. The I/O performance of the application is impacted during this time frame due to the execution of the **mmrestripefs** command.

```

| Tue Jan 19 14:32:41.625 2016: [I] Command: mmrestripefs /dev/gpfs2 -r -N all
| Tue Jan 19 14:45:42.975 2016: [I] Command: successful mmrestripefs /dev/gpfs2 -r -N all

```

#### Problem resolution and verification

The Quality of Service (QoS) feature for I/O operations in IBM Spectrum Scale 4.2 and higher versions is used to allocate appropriate maintenance IOPS to reduce the impact of the maintenance operation on the

| application. In the following sample output, the file system consists of a single storage pool – the default  
| ‘system’ pool. The QoS feature is disabled and inactive.

```
| # mm1sqos gpfs1a
| QoS config::      disabled
| QoS status::      throttling inactive, monitoring inactive
```

| You can use the **mmchqos** command to allocate appropriate maintenance IOPS to the IBM Spectrum Scale  
| system. For example, consider that the storage system has 100 K IOPS. If you want to allocate 1000 IOPS  
| to the long running GPFS maintenance operations for the system storage pool, use the **mmchqos** command  
| to enable the QoS feature, and allocate the IOPS as shown:

```
| # mmchqos gpfs1a --enable pool=system,maintenance=1000IOPS
| Adjusted QoS Class specification: pool=system,other=inf,maintenance=1000Iops
| QoS configuration has been installed and broadcast to all nodes.
```

| Verify the QoS setting and values on a file system by using the **mm1sqos** command.

```
| # mm1sqos gpfs1aQoS config:: enabled --
|           pool=system,other=inf,maintenance=1000IopsQoS status::      throttling active,
|           monitoring active
```

| **Note:** Allocating a small share of IOPS, for example 1000 IOPS, to the long running GPFS maintenance  
| operations can increase the maintenance command execution times. So depending on the operation's  
| needs, the IOPS assigned to the ‘other’ and ‘maintenance’ class must be adjusted by using the **mmchqos**  
| command. This balances the application as well as the I/O requirements for the GPFS maintenance  
| operation.

| For more information on setting the QoS for I/O operations, see the *mm1sqos command* section in the *IBM  
| Spectrum Scale: Command and Programming Reference* and *Setting the Quality of Service for I/O operations  
| (QoS)* section in the *IBM Spectrum Scale: Administration Guide*.

## | **Suboptimal performance due to frequent invocation or execution of | maintenance commands**

| When the GPFS maintenance operations like **mmbackup**, **mmapplypolicy**, **mmdf**, **mmcrsnapshot**,  
| **mmde1snapshot**, and others are in progress, they can consume some percentage of system resources. This  
| can impact the I/O performance of applications. If these maintenance operations are scheduled  
| frequently, for example within every few seconds or minutes, the performance impact can be significant,  
| unless the I/O subsystem is sized adequately to handle both the application and the maintenance  
| operation I/O load.

## | **Problem identification**

| Check the GPFS log file `/var/adm/ras/mmfs.log.1latest` on the file system manager node **mm1smgr** to  
| verify whether any GPFS maintenance operations are being invoked frequently by a cron job or other  
| cluster management software like **Nagios**.

| In the sample output below, the **mmdf** command is being invoked periodically every 3-4 seconds.

```
| Tue Jan 19 15:13:47.389 2016: [I] Command: mmdf /dev/gpfs2
| Tue Jan 19 15:13:47.518 2016: [I] Command: successful mmdf /dev/gpfs2
| Tue Jan 19 15:13:51.109 2016: [I] Command: mmdf /dev/gpfs2
| Tue Jan 19 15:13:51.211 2016: [I] Command: successful mmdf /dev/gpfs2
| Tue Jan 19 15:13:54.816 2016: [I] Command: mmdf /dev/gpfs2
| Tue Jan 19 15:13:54.905 2016: [I] Command: successful mmdf /dev/gpfs2
| Tue Jan 19 15:13:58.481 2016: [I] Command: mmdf /dev/gpfs2
| Tue Jan 19 15:13:58.576 2016: [I] Command: successful mmdf /dev/gpfs2
| Tue Jan 19 15:14:02.164 2016: [I] Command: mmdf /dev/gpfs2
| Tue Jan 19 15:14:02.253 2016: [I] Command: successful mmdf /dev/gpfs2
```

```

| Tue Jan 19 15:14:05.850 2016: [I] Command: mmdf /dev/gpfs2
| Tue Jan 19 15:14:05.945 2016: [I] Command: successful mmdf /dev/gpfs2
| Tue Jan 19 15:14:09.536 2016: [I] Command: mmdf /dev/gpfs2
| Tue Jan 19 15:14:09.636 2016: [I] Command: successful mmdf /dev/gpfs2
| Tue Jan 19 15:14:13.210 2016: [I] Command: mmdf /dev/gpfs2
| Tue Jan 19 15:14:13.299 2016: [I] Command: successful mmdf /dev/gpfs2
| Tue Jan 19 15:14:16.886 2016: [I] Command: mmdf /dev/gpfs2
| Tue Jan 19 15:14:16.976 2016: [I] Command: successful mmdf /dev/gpfs2
| Tue Jan 19 15:14:20.557 2016: [I] Command: mmdf /dev/gpfs2
| Tue Jan 19 15:14:20.645 2016: [I] Command: successful mmdf /dev/gpfs2

```

## | Problem resolution and verification

| Adjust the frequency of the GPFS maintenance operations so that it does not impact the applications performance. The I/O subsystem must be designed in such a way that it is able to handle both the application and the maintenance operation I/O load.

| You can also use the `mmchqos` command to allocate appropriate maintenance IOPS, which can reduce the impact of the maintenance operations on the application.

## | Suboptimal performance when a tracing is active on a cluster

| Tracing is usually enabled on the IBM Spectrum Scale cluster for troubleshooting purposes. However, running a trace on a node might cause performance degradation.

## | Problem identification

| Issue the `mmfsconfig` command and verify whether GPFS tracing is configured. The following sample output displays a cluster in which tracing is configured:

```

| # mmfsconfig | grep trace
| trace all 4 tm 2 thread 1 mutex 1 vnode 2 ksvfs 3 klockl 2 io 3 pgallo 1 mb 1 lock 2 fsck 3
| tracedevOverwriteBufferSize 1073741824
| tracedevWriteMode overwrite 268435456

```

| Issue the `# ps -aux | grep lxtrace | grep mmfs` command to determine whether GPFS tracing process is running on a node. The following sample output shows that GPFS tracing process is running on the node:

```

| # ps -aux | grep lxtrace | grep mmfs
| root      19178  0.0  0.0  20536  128 ?        Ss   14:06   0:00
| /usr/lpp/mmfs/bin/lxtrace-3.10.0-229.e17.x86_64 on
| /tmp/mmfs/lxtrace.trc.c80f1m5n08ib0 --overwrite-mode --buffer-size
| 268435456

```

## | Problem resolution and verification

| When the traces have met their purpose and are no longer needed, use one of the following commands to stop the tracing on all nodes:

- Use this command to stop tracing:

```
| mmtracectl --stop -N all
```

- Use this command to clear all the trace setting variables and stop the tracing:

```
| mmtracectl --off -N all
```

## | Suboptimal performance due to replication settings being set to 2 or 3

| The file system write performance depends on the write performance of the storage volumes and its RAID configuration. However, in case the backend storage write performance is on par with its read

| performance, but the file system write performance is just 50% (half) or 33% (one-third) of the read performance, check if the file system replication is enabled.

## | **Problem identification**

| When file system replication is enabled and set to 2, effective write performance becomes 50% of the raw write performance, since for every write operation, there are two internal write operation due to replication. Similarly, when file system replication is enabled and set to 3, effective write performance becomes approximately 33% of the raw write performance, since for every write operation, there are three internal write operation.

| Issue the `mm1sfs` command, and verify the default number of metadata and data replicas enabled on the file system. In the following sample output the metadata and data replication on the file system is set to 2:

```
| # mm1sfs <fs> | grep replica | grep -i default
| -m                2                Default number of metadata replicas
| -r                2                Default number of data replicas
```

| Issue the `mm1sattr` command to check whether replication is enabled at file level

```
| # mm1sattr -L largefile.foo | grep replication
| metadata replication: 2 max 2
| data replication:    2 max 2
```

## | **Problem resolution and verification**

| The GPFS placement policy can be enforced to set the replication factor of temporary files for non-critical datasets to one. For example, temporary files like log files that can be recreated if necessary.

| Follow these steps to set the replication value for log files to 1:

| 1. Create a `placement_policy.txt` file by using the following rule:

```
| rule 'non-replicate-log-files' SET POOL 'SNCdata' REPLICATE (1) where lower(NAME) like '%.log'
| rule 'default' SET POOL 'SNCdata'
```

| 2. Install the placement policy on the file system by using the following command:

```
| mmchpolicy <fs> placement_policy.txt
```

| **Note:** You can test the placement policy before installing it by using the following command:

```
| mmchpolicy <fs> placement_policy.txt -I test
```

| 3. Issue one of the following commands to remount the file system for the policy to take effect:

| Remount the file system on all the nodes by using one of the following commands:

- | • `mmumount <fs> -N all`
- | • `mmmount <fs> -N all`

| 4. Issue the `mm1spolicy <fs> -L` command to verify whether the output is as shown:

```
| rule 'non-replicate-log-files' SET POOL 'SNCdata' REPLICATE (1) where lower(NAME) like '%.log'
| rule 'default' SET POOL 'SNCdata'
```

## | **Suboptimal performance due to updates made on a file system or fileset with snapshot**

| If a file is modified after its snapshot creation, the system can face performance degradation due to the copy-on-write property enforced on updated data files.

## | Problem identification

| Updating a file that has a snapshot might create unnecessary load on a system because each application update or write operation goes through the following steps:

- | 1. Read the original data block pertaining to the file region that must be updated.
- | 2. Write the data block read in the step 1 above to the corresponding snapshot location.
- | 3. Perform the application write or update operation on the desired file region.

| Issue the **mm1ssnapshot** to verify whether the snapshot was created before the file data update operation.

| In the following sample output, the **gpfs2** file system contains a snapshot.

```
| # mm1ssnapshot gpfs2
| Snapshots in file system gpfs2:
| Directory          SnapId   Status  Created
| snap1              2       Valid   Mon Jan 25 12:42:30 2016
```

## | Problem resolution and verification

| Use the **mmdel snapshot** command to delete the file system snapshot, if it is no longer necessary. For more information on the **mmdel snapshot** command, see the *mmdel snapshot command* in the *IBM Spectrum Scale: Command and Programming Reference*.

---

## Delays and deadlocks

The first item to check when a file system appears hung is the condition of the networks including the network used to access the disks.

Look for increasing numbers of dropped packets on all nodes by issuing:

- The **netstat -D** command on an AIX node.
- The **ifconfig interfacename** command, where *interfacename* is the name of the interface being used by GPFS for communication.

When using subnets ( see the *Using remote access with public and private IP addresses* topic in the *IBM Spectrum Scale: Administration Guide* .), different interfaces may be in use for intra-cluster and intercluster communication. The presence of a hang or dropped packed condition indicates a network support issue that should be pursued first. Contact your local network administrator for problem determination for your specific network configuration.

If file system processes appear to stop making progress, there may be a system resource problem or an internal deadlock within GPFS.

**Note:** A deadlock can occur if user exit scripts that will be called by the **mmaddcallback** facility are placed in a GPFS file system. The scripts should be placed in a local file system so they are accessible even when the networks fail.

To debug a deadlock, do the following:

1. Check how full your file system is by issuing the **mmdf** command. If the **mmdf** command does not respond, contact the IBM Support Center. Otherwise, the system displays information similar to:

disk name	disk size in KB	failure group	holds metadata	holds data	free KB in full blocks	free KB in fragments
-----						
Disks in storage pool: system (Maximum disk size allowed is 1.1 TB)						
dm2	140095488	1	yes	yes	136434304 ( 97%)	278232 ( 0%)
dm4	140095488	1	yes	yes	136318016 ( 97%)	287442 ( 0%)
dm5	140095488	4000	yes	yes	133382400 ( 95%)	386018 ( 0%)



dm0nsd	140095488	4005	yes	yes	134701696 ( 96%)	456188 ( 0%)
dm1nsd	140095488	4006	yes	yes	133650560 ( 95%)	492698 ( 0%)
dm15	140095488	4006	yes	yes	140093376 (100%)	62 ( 0%)
-----						
(pool total)	840572928				814580352 ( 97%)	1900640 ( 0%)
=====						
(total)	840572928				814580352 ( 97%)	1900640 ( 0%)

#### Inode Information

```
-----
Number of used inodes:      4244
Number of free inodes:     157036
Number of allocated inodes: 161280
Maximum number of inodes:  512000
```

GPFS operations that involve allocation of data and metadata blocks (that is, file creation and writes) will slow down significantly if the number of free blocks drops below 5% of the total number. Free up some space by deleting some files or snapshots (keeping in mind that deleting a file will not necessarily result in any disk space being freed up when snapshots are present). Another possible cause of a performance loss is the lack of free inodes. Issue the **mmchfs** command to increase the number of inodes for the file system so there is at least a minimum of 5% free. If the file system is approaching these limits, you may notice the following error messages:

#### 6027-533 [W]

Inode space *inodeSpace* in file system *fileSystem* is approaching the limit for the maximum number of inodes.

#### operating system error log entry

```
Jul 19 12:51:49 node1 mmfs: Error=MMFS_SYSTEM_WARNING, ID=0x4DC797C6,
Tag=3690419: File system warning. Volume fs1. Reason: File system fs1 is approaching the
limit for the maximum number of inodes/files.
```

2. If automated deadlock detection and deadlock data collection are enabled, look in the latest GPFS log file to determine if the system detected the deadlock and collected the appropriate debug data. Look in **/var/adm/ras/mmfs.log.latest** for messages similar to the following:

```
Thu Feb 13 14:58:09.524 2014: [A] Deadlock detected: 2014-02-13 14:52:59: waiting 309.888 seconds on node
p7fbn12: SyncHandlerThread 65327: on LkObjCondvar, reason 'waiting for R0 lock'
Thu Feb 13 14:58:09.525 2014: [I] Forwarding debug data collection request to cluster manager p7fbn11 of
cluster cluster1.gpfs.net
Thu Feb 13 14:58:09.524 2014: [I] Calling User Exit Script gpfsDebugDataCollection: event deadlockDebugData,
Async command /usr/lpp/mmfs/bin/mmcommon.
Thu Feb 13 14:58:10.625 2014: [N] sdrServ: Received deadlock notification from 192.168.117.21
Thu Feb 13 14:58:10.626 2014: [N] GPFS will attempt to collect debug data on this node.
mmtrace: move /tmp/mmfs/lxtrace.trc.p7fbn12.recycle.cpu0
/tmp/mmfs/trcfile.140213.14.58.10.deadlock.p7fbn12.recycle.cpu0
mmtrace: formatting /tmp/mmfs/trcfile.140213.14.58.10.deadlock.p7fbn12.recycle to
/tmp/mmfs/trcrpt.140213.14.58.10.deadlock.p7fbn12.gz
```

This example shows that deadlock debug data was automatically collected in **/tmp/mmfs**. If deadlock debug data was not automatically collected, it would need to be manually collected.

To determine which nodes have the longest waiting threads, issue this command on each node:

```
/usr/lpp/mmfs/bin/mmdiag --waiters waitTimeInSeconds
```

For all nodes that have threads waiting longer than *waitTimeInSeconds* seconds, issue:

```
mmfsadm dump all
```

#### Notes:

- a. Each node can potentially dump more than 200 MB of data.
- b. Run the **mmfsadm dump all** command only on nodes that you are sure the threads are really hung. An **mmfsadm dump all** command can follow pointers that are changing and cause the node to crash.

3. If the deadlock situation cannot be corrected, follow the instructions in “Additional information to collect for delays and deadlocks” on page 368, then contact the IBM Support Center.

---

## E! Chapter 21. GUI issues

E! The topics listed in this section provide the list of most frequent and important issues reported with the E! IBM Spectrum Scale GUI.

### E! Related concepts:

E! Chapter 2, “Monitoring system health through the IBM Spectrum Scale GUI,” on page 79

E! Use **Monitoring > Events** page to monitor the events that are reported in the system. The Events page E! displays events and you can monitor and troubleshoot errors on your system.

E! “Collecting diagnostic data through GUI” on page 174

IBM Support might ask you to collect logs, trace files, and dump files from the system to help them resolve a problem. You can perform this task from the management GUI or by using the **gpfs.snap** command. Use the **Settings > Diagnostic Data** page in the IBM Spectrum Scale GUI to collect details of the issues reported in the system.

---

## Understanding GUI support matrix and limitations

It is important to understand the supported versions and limitations to analyze whether you are facing a real issue in the system.

The IBM Spectrum Scale FAQ in IB Knowledge Center contains the GUI support matrix. The IBM Spectrum Scale FAQ is available at <http://www.ibm.com/support/knowledgecenter/STXKQY/gpfsclustersfaq.html>.

To know more about GUI limitations, see *GUI limitations* in *IBM Spectrum Scale: Administration Guide*.

---

## Examples for GUI issues and their resolutions

This topic covers some of the examples for the most frequent GUI issues and its resolutions.

### GUI fails to start

This issue is primarily because of the database issue. In ideal scenarios, the service script automatically initializes and starts PostgreSQL. However, in rare cases, the database might be either inconsistent or corrupted.

If the PostgreSQL database is corrupted, it might be because of the following reasons:

- The additional (non-distro) PostgreSQL package is installed and it occupies the port 5432.
- Details that are stored in the `/etc/hosts` file are corrupted so the "localhost" is not listed as the first item for the IP127.0.0.1.
- An incompatible schema exists in the database from a previous release.

If the GUI logs show any of the database errors, try the following steps:

1. Issue **systemctl stop gpfsGUI** to stop GUI services.
2. Issue `'su postgres -c 'psql -d postgres -c "DROP SCHEMA FSCC CASCADE"'`.
3. If the previous step does not help, issue `'rm -rf /var/lib/pgsql/data'`.
4. Issue **systemctl start gpfsGUI** to start GUI.

If the problem still persists, it might be because of a corrupted GUI installation, missing GUI dependency, or some other unknown issue. In this scenario, you can remove and reinstall the GUI rpm. For more information on how to install and uninstall GUI rpms, see *Manually installing IBM Spectrum Scale management GUI* in the *IBM Spectrum Scale: Concepts, Planning, and Installation Guide*.

You can collect the logs that are available in the `/var/log/cnlog/mgtsrv` folder to investigate further. You can also use the **gpfs.snap** command as shown in the following example to collect logs and dumps in case of a GUI issue:

```
gpfs.snap -N GUI_MGMT_SERVERS
```

Collecting logs and dumps through the **gps.snap** command also collects the GPFS logs. So, manually getting the logs from the folder `/var/log/cnlog/mgtsrv` is quicker and provides only the required data that is required to search for the details of the GUI issue.

## GUI login page does not open

The management GUI is accessible through the following URL after the installation: `https://<ip or host name>`.

If the GUI login page does not open, try out the following:

1. Issue the following command to verify the status:  
`systemctl status gpfsGUI`
2. Check the status of java components by issuing the following command:

```
netstat -ltn | grep java
```

The system must display the following output in the ideal scenarios:

There can be more lines in the output as given in the following example but the port 443 is the most important for the GUI service:

```
tcp6  0  0  :::49679          :::*      LISTEN    67869/java
tcp6  0  0  127.0.0.1:38743   :::*      LISTEN    67869/java
tcp6  0  0  :::443            :::*      LISTEN    67869/java
tcp6  0  0  :::4444           :::*      LISTEN    67869/java
```

**Note:** The port 4444 is used by the GUI CLI to interact with the GUI back-end service. Other ports that are listed here are used by Java internally.

If you find that the port 443 is not opened by WebSphere® Liberty, restart the GUI service by issuing the **systemctl restart gpfsGUI** command. The GUI uses the default https port 443. If some other application or process listens to this port, it causes a port conflict and the GUI does not work.

## GUI performance monitoring issues

The sensor gets the performance data for the collector. The collector application that is called *pmcollector* runs on every GUI node to display the performance details in the GUI. A sensor application is running on every node of the system.

If GUI is not displaying the performance data, the following might be the reasons:

1. Collectors are not enabled
2. Sensors are not enabled
3. NTP failure

### Collectors are not enabled

Do the following to verify whether collectors are working properly:

1. Issue **systemctl status pmcollector** on the GUI node to confirm that the collector is running.
2. If collector service is not started already, start the collector on the GUI nodes by issuing the **systemctl restart pmcollector** command. Depending on the system requirement, the *pmcollector* service can be configured to be run on the nodes other than GUI nodes. You need to verify the status of *pmcollector* service on all nodes where collector is configured.
3. If you cannot start the service, verify its log file that is located at `/var/log/zimon/ZIMonCollector.log` to see whether it logs any other details of the issues related to the collector service status.

4. Use a sample CLI query to test if data collection works properly. For example:

```
mmperfmon query cpu_user
```

**Note:** After migrating from release 4.2.0.x to 4.2.1, you might see the *pmcollector* service critical error on GUI nodes. In this case, restart the *pmcollector* service by running the **systemctl restart pmcollector** command on all GUI nodes.

### Sensors are not enabled

The following table lists sensors that are used to get the performance data for each resource type:

Table 53. Sensors available for each resource type

Resource type	Sensor name	Candidate nodes
Network	Network	All
System Resources	CPU	All
	Load	
	Memory	
NSD Server	GPFSNSDDisk	NSD Server nodes
IBM Spectrum Scale Client	GPFSFilesystem	IBM Spectrum Scale Client nodes
	GPFSVFS	
	GPFSFilesystemAPI	
NFS	NFSIO	Protocol nodes running NFS service
SMB	SMBStats	Protocol nodes running SMB service
	SMBGlobalStats	
CTDB	CTDBStats	Protocol nodes running SMB service
Object	SwiftAccount	Protocol nodes running Object service
	SwiftContainer	
	SwiftObject	
	SwiftProxy	
Transparent Cloud Tiering	MCStoreGPFSStats	Cloud gateway nodes
	MCStoreIcstoreStats	
	MCStoreLWESStats	
Capacity	DiskFree	All nodes
	GPFSFilesetQuota	Only a single node
	GPFSDiskCap	Only a single node

Do the following to verify whether sensors are working properly:

1. Confirm that the sensor is configured correctly by issuing the **mmperfmon config show** command. This command lists the content of the sensor configuration that is located at `/opt/IBM/zimon/ZIMonSensors.cfg`.
2. The configuration must point to the node where the collector is running and all the expected sensors must be enabled. An enabled sensor has a period greater than 0 in the same configuration file. After the configuration file is updated, the *pmsensor* service needs to be restarted.
3. Issue **systemctl start pmsensors** to start the service if it is stopped.

If sensors and collectors are properly configured and enabled, you can issue the **mmperfmon** and **mmpmmon** commands to see whether performance data is really generated.

You can query the data displayed in the performance charts through CLI as well. For more information on how to query performance data displayed in GUI, see “Querying performance data shown in the GUI through CLI” on page 76.

## NTP failure

The performance monitoring fails if the clock is not properly synchronized in the cluster. Issue the `ntpq -c peers` command to verify the NTP state.

### Related concepts:

“Performance monitoring through IBM Spectrum Scale GUI” on page 69

The IBM Spectrum Scale GUI provides a graphical representation of the status and historical trends of the key performance indicators. This helps the users to make decisions easily without wasting time. Use **Monitoring > Dashboard** and **Monitoring > Statistics** pages in the GUI to monitor the performance of the system based on various aspects.

Chapter 20, “Performance issues,” on page 327

The performance issues might occur because of the system components or configuration or maintenance issues.

## GUI is showing “Server was unable to process the request” error

The GUI might not respond on user actions or it might show “Server was unable to process the request” error. This is primarily because of JavaScript Console issue. JavaScript errors are not collected in the diagnostic data. The IBM Support might need the JavaScript error details to troubleshoot this issue.

The location where the JavaScript console is stored depends on the web browser.

- **For Google Chrome:** Select menu item **Tools > Javascript Console**.
- **For Mozilla Firefox:** Install and run the firebug plug-in to get the JavaScript console.

## GUI is displaying outdated information

The IBM Spectrum Scale GUI caches configuration data in an SQL database. Refresh tasks update the cached information. Many refresh tasks are invoked by events when the configuration is changed in the cluster. In those cases, the GUI pages reflect changes in a minute. For certain types of data, events are not raised by itself to invoke the refresh tasks. In such cases, the system must poll the data on a regular interval to reflect up-to-date information in the GUI pages. All the refresh tasks run on a schedule. The system also polls the data frequently even for those tasks that are triggered by events.

If the GUI shows stale data and the user does not want to wait until the next issue of refresh task, you can run those refresh tasks manually as shown in the following example:

```
/usr/lpp/mmfs/gui/cli/runtask <task_name>
```

**Note:** Many file system-related tasks require the corresponding file system to be mounted on the GUI to collect data.

The following table lists the details of the available GUI refresh tasks.

*Table 54. GUI refresh tasks*

Refresh task	Frequency	Collected information	Prerequisite - File system must be mounted	Invoked by event	CLI commands used
ALTER_HOST_NAME	12 h	Host names and IP addresses in Monitor > Nodes page			mmremote networkinfo
CALLBACK	6 h	Checks and registers callbacks used by GUI		Yes	mm1scallback and mmaddcallback
CES_ADDRESS	1 h	CES IP addresses in Monitor > Nodes page		Yes	mmces node list

Table 54. GUI refresh tasks (continued)

Refresh task	Frequency	Collected information	Prerequisite - File system must be mounted	Invoked by event	CLI commands used
CES_STATE	10 min	CES state in Monitor > Nodes		Yes	mmces state show -N cesNodes  mmces events active -N cesNodes (used for the information field)
CES_SERVICE_STATE	1 h	CES service state in Monitor > Nodes page		Yes	mmces service list -N cesNodes -Y
CES_USER_AUTH_SERVICE	1 h	Not displayed		Yes	mmuserauth service list -Y
CHECK_FIRMWARE	6 h	Monitor > Events page			Checks whether the reported firmware is up to date
CLUSTER_CONFIG	1 h	List of nodes and node classes in Monitoring > Nodes		Yes	mmsdrquery and mmlsnodeclass
CONNECTION_STATUS	10 min	Connections status in Monitoring > Nodes page			Nodes reachable through SSH
DAEMON_CONFIGURATION	1 h	Not displayed		Yes	mmlsconfig
DF	1 h	Not directly displayed; used to generate low space events	Yes	Yes	df, df -i, mmlspool
DISK_USAGE	3:00 AM	Not directly displayed; used to generate low space events	Yes		mmdf, mmsdrquery (mmlsnd and mmremote getdisksize for non-GNR-NSDs that is not assigned to the file system)
DISKS	1 h	NSD list in Monitoring > NSDs		Yes	mmsrquery, mmlsnd, and mmlsdisk
FILESETS_NOUSAGE	1 h	List of filesets in Files > Filesets	Yes	Yes	mmlsfileset
FILESYSTEM_MOUNT	1 h	Mount state in Files > File Systems		Yes	mmlsmount
FILESYSTEMS	1 h	List of file systems in Files > File Systems	Yes	Yes	mmsdrquery, mmlsfs, , and mmlsnadir
GUI_CONFIG_CHECK	12 h	Checks that cluster configuration is compatible with GUI requirements		Yes	mmsdrquery, mmgetstate, and getent
HEALTH_STATES	10 min	Health events in Monitoring > Events		Yes	mmhealth node show {component} -v -N {nodes} -Y  mmhealth node eventlog -Y
HOST_STATES	1 h	GPFS state in Monitoring > Nodes		Yes	mmgetstate
NFS_EXPORTS	1 h	Exports in Protocols > NFS Exports		Yes	mmceservice list and mncesnfslexport
NFS_EXPORTS_DEFAULTS	1 h	Not displayed		Yes	mmceservice list and mncesnfs1scfg
NFS_SERVICE	1 h	NFS settings in Settings > NFS Service		Yes	mmceservice list and mncesnfs1scfg
NODECLASS	6 h	Node classes in Monitor>Nodes		Yes	mmlsnodeclass
OBJECT_STORAGE_POLICY	6 h	Storage policies of containers in Object > Accounts		Yes	mmobj policy list
OS_DETECT	6 h	Information about operating system, cpu architecture, hardware vendor, type, serial in Monitoring > Nodes		Yes	mmremote nodeinfo
PM_MONITOR	10 min	Checks if the performance collector is up and running and also checks the CPU data for each node			systemctl status pmcollector and zimon query
POLICIES	1 h	Policies in Files > Information Lifecycle	Yes	Yes	mmlspolicy
QUOTA	2:15 AM	Quotas in Files > Quota  Fileset capacity in Monitoring > Capacity	Yes	Yes	mmrepquota and mmlsdefaultquota

Table 54. GUI refresh tasks (continued)

Refresh task	Frequency	Collected information	Prerequisite - File system must be mounted	Invoked by event	CLI commands used
REMOTE_CONFIG	1 h	Not displayed		Yes	mmauth, gets and parses sdr file
SMB_GLOBALS	1 h	SMB settings in Settings > SMB Service		Yes	mmcesssmbconfig
SMB_SHARES	1 h	Shares in Protocols > SMB Shares		Yes	mmcesssmbsexport
SNAPSHOTS	1 h	Snapshots in Files > Snapshots	Yes	Yes	mm1ssnapshot
SNAPSHOTS_USAGE	Never	Snapshot usage in Files > Snapshots <b>Important:</b> This is an expensive operation that might take down the snapshot subsystem for hours in larger file systems.	Yes		mm1ssnapshot -d
SYSTEMUTIL_DF	1 h	Used to generate warnings if nodes run out of local disk space			Checks local disk space of node
STORAGE_POOL	1 h	Pool properties in Files > File Systems		Yes	mm1spool <device> all -L -Y
TIMESYNC_MONITOR	6 h	Generates warnings if the clocks on nodes are out of sync			date

## Capacity information is not available in GUI pages

The IBM Spectrum Scale management GUI does not display the capacity information on various GUI pages if *GPFSDiskCap* and *GPFSEntQuota* sensors are disabled and quota is disabled on the file system.

The following table provides the solution for the capacity data display issues in the corresponding GUI pages.

Table 55. Troubleshooting details for capacity data display issues in GUI

GUI page	Solution
<b>Files &gt; File Systems, Monitoring &gt; Capacity &gt; File Systems, and Monitoring &gt; Capacity &gt; Pools</b>	Verify whether <i>GPFSDiskCap</i> sensor is enabled on at least one node and ensure that the file system is mounted on this node. It is recommended to enable sensor and mount file system on the GUI node itself to get most updated capacity data.  As the period is usually set to one day (86400 seconds), it can take up to one day until the capacity information is displayed or updated.
<b>Files &gt; Filesets, Monitoring &gt; Capacity &gt; Filesets, Monitoring &gt; Capacity &gt; Users, and Monitoring &gt; Capacity &gt; User Groups</b> do not display capacity information for the fileset.	In this case, the quota is not enabled for the file system that hosts this fileset. Go to <b>Files &gt; Quotas</b> page and enable quotas for the corresponding file system. By default, the quotas are disabled for all file systems.
<b>Monitoring &gt; Statistics</b>	Verify whether the <i>GPFSDiskCap</i> and <i>GPFSEntQuota</i> sensors are enabled and quota is enabled for the file systems. For more information on how to enable performance monitoring sensors, see "Configuring performance monitoring options in GUI" on page 70.

E!



## Chapter 22. AFM issues

The following table lists the common questions in AFM.

Table 56. Common questions in AFM with their resolution

Question	Answer / Resolution
How do I flush requeued messages?	<p>Sometimes, requests in the AFM messages queue on the gateway node get requeued because of errors at the home cluster. For example, if space is not available at the home cluster to perform a new write, a write message that is queued is not successful and gets requeued. The administrator views the failed message being requeued on the Primary gateway. Add more space to the home cluster and run <code>mmafmctl resumeRequeued</code> so that the requeued messages are executed at home again. If <code>mmafmctl resumeRequeued</code> is not run by an administrator, AFM executes the message in the regular order of message executions from the cache cluster to the home cluster.</p> <p>Running the <code>mmfsadm saferdump afm all</code> command on the gateway node displays the queued messages. The requeued messages are displayed in the dumps. An example:</p> <pre>c12c4apv13.gpfs.net: Normal Queue: (listed by execution order) (state: Active) c12c4apv13.gpfs.net: Write [612457.552962] requeued file3 (43 @ 293) chunks 0 bytes 0 0</pre>
Why is a fileset in the Unmounted or Disconnected state when parallel I/O is set up?	Filesets that are using a mapping target go to the Disconnected mode if the NFS server of the Primary gateway is unreachable, even if NFS servers of all participating gateways are reachable. The NFS server of the Primary gateway must be checked to fix this problem.
How do I activate an inactive fileset?	The <code>mmafmctl prefetch</code> command without options, where prefetch statistics are procured, activates an inactive fileset.
How do I reactivate a fileset in the Dropped state?	The <code>mmafmctl prefetch</code> command without options, where prefetch statistics are procured, activates a fileset in a dropped state.
How to clean unmount the home filesystem if there are caches using GPFS protocol as backend?	<p>To have a clean unmount of the home filesystem, the filesystem must first be unmounted on the cache cluster where it is remotely mounted and the home filesystem must be unmounted. Unmounting the remote file system from all nodes in the cluster might not be possible until the relevant cache cluster is unlinked or the local file system is unmounted.</p> <p>Force unmount, shutdown, or crash of the remote cluster results in panic of the remote filesystem at the cache cluster and the queue is dropped. The next access to the fileset runs the recovery. However, this should not affect the cache cluster.</p>
What should be done if the <code>df</code> command hangs on the cache cluster?	<p>On RHEL 7.0 or later, <code>df</code> does not support hidden NFS mounts. As AFM uses regular NFS mounts on the gateway nodes, this change causes commands like <code>df</code> to hang if the secondary gets disconnected.</p> <p>The following workaround can be used that allows NFS mounts to continue to be hidden:</p> <p>Remove <code>/etc/mtab</code> symlink, and create a new file <code>/etc/mtab</code> and copy <code>/proc/mounts</code> to <code>/etc/mtab</code> file during the startup. In this solution, the <code>mtab</code> file might go out of synchronization with <code>/proc/mounts</code>.</p>
What happens when the hard quota is reached in an AFM cache?	Like any filesystem that reaches the hard quota limit, requests fail with <code>E_NO_SPACE</code> .

Table 56. Common questions in AFM with their resolution (continued)

Question	Answer / Resolution
When are inodes deleted from the cache?	After an inode is allocated, it is never deleted. The space remains allocated and they are re-used.
If inode quotas are set on the cache, what happens when the inode quotas are reached?	Attempts to create new files fail, but cache eviction is not triggered. Cache eviction is triggered only when block quota is reached, not the inode quotas.
How can the cache use more inodes than the home?	One way is for file deletions. If a file is renamed at the home site, the file in cache is deleted and created again in cache. This results in the file being assigned a different inode number at the cache site. Also, if a cache fileset is LU mode or SW mode, then there can be changes made at the cache that cause it to be bigger than the home.
Why does fileset go to Unmounted state even if home is accessible on the cache cluster?	Sometimes, it is possible that the same home is used by multiple clusters, one set of filesets doing a quiesce turn the home unresponsive to the second cluster's filesets, which show home as unmounted
What could be impact of not running <b>mmafmconfig</b> command despite having a GPFS home?	Sparse file support is not present even if home is GPFS. Recovery and many AFM functions do not work. Crashes can happen for <b>readdir</b> or <b>lookup</b> , if the backend is using NSD protocol and remote mount is not available at the gateway node.
What should be done if there are cluster wide waiters but everything looks normal, such as home is accessible from gateway nodes, applications are in progress on the cache fileset?	This can happen when the application is producing requests at a faster pace. Check <b>iohist</b> to check disk rates.
Read seems to be stuck/inflight for a long time. What should be done?	Restart <b>nfs</b> at home to see if error resolves. Check the status of the fileset using <b>mmafmctl getstate</b> command to see if you fileset is in unmounted state.
The <b>mmfs.log</b> show errors during read such as error 233 :	These are temporary issues during read: Tue Feb 16 03:32:40.300 2016: [E] AFM: Read file system fs1 fileset newSanity-160216-020201-KNFS-TC8-SW file IDs [58195972.58251658.-1.-1,R] name file-3G remote error 233 These go away automatically and read should be successful.
Can the home have different sub-directories exported using unique FSIDs, while parent directory is also exported using an FSID?	This is not a recommended configuration.
I have a non-GPFS home, I have applications running in cache and some requests are requeued with the following error: SetXAttr file system fs1 fileset sw_gpfs file IDs [-1.1067121.-1.-1,N] name local error 124	<b>mmafmconfig</b> is not setup at home. Running <b>mmafmconfig</b> command at home and relinking cache should resolve this issue.
During failover process, some gateway nodes might show error 233 in <b>mmfs.log</b> .	This error is harmless. The failover completes successfully.
Resync fails with No buffer space available error, but <b>mmdiag --memory</b> shows that memory is available.	Increase <b>afmHardMemThreshold</b> .

Table 56. Common questions in AFM with their resolution (continued)

Question	Answer / Resolution
How can I change the mode of a fileset?	<p>The mode of an AFM client cache fileset cannot be changed from local-update mode to any other mode; however, it can be changed from read-only to single-writer (and vice versa), and from either read-only or single-writer to local-update. Complete the following steps to change the mode:</p> <ol style="list-style-type: none"> <li>1. Ensure that fileset status is active and that the gateway is available.</li> <li>2. Unmount the file system</li> <li>3. Unlink the fileset.</li> <li>4. Run the <code>mmchfileset</code> command to change the mode.</li> <li>5. Mount the file system again.</li> <li>6. Link the fileset again.</li> </ol>
Why are <code>setuid</code> or <code>setgid</code> bits in a single-writer cache reset at home after data is appended?	<p>The <code>setuid</code> or <code>setgid</code> bits in a single-writer cache are reset at home after data is appended to files on which those bits were previously set and synced. This is because over NFS, a write operation to a <code>setuid</code> file resets the <code>setuid</code> bit.</p>
How can I traverse a directory that is not cached?	<p>On a fileset whose metadata in all subdirectories is not cached, any application that optimizes by assuming that directories contain two fewer subdirectories than their hard link count do not traverse the last subdirectory. One such example is <code>find</code>; on Linux, a workaround for this is to use <code>find -noleaf</code> to correctly traverse a directory that has not been cached</p>
What extended attribute size is supported?	<p>For an operating system in the gateway whose Linux kernel version is below 2.6.32, the NFS max <code>rsiz</code> is 32K, so AFM does not support an extended attribute size of more than 32K on that gateway.</p>
What should I do when my file system or fileset is getting full?	<p>The <code>.ptrash</code> directory is present in cache and home. In some cases, where there is a conflict that AFM cannot resolve automatically, the file is moved to <code>.ptrash</code> at cache or home. In cache the <code>.ptrash</code> gets cleaned up when eviction is triggered. At home, it is not cleared automatically. When the administrator is looking to clear some space, the <code>.ptrash</code> must be cleaned up first.</p>



## Chapter 23. AFM DR issues

This topic lists the answers to the common AFM DR questions.

Table 57. Common questions in AFM DR with their resolution

Issue	Resolution
How do I flush requeued messages?	<p>Sometimes, requests in the AFM messages queue on the gateway node get requeued due to errors at the home cluster. For example, if space is not available at the home cluster to perform a new write, a write message that is queued is not successful and gets requeued. The administrator views the failed message being requeued on the MDS. Add more space to the home cluster and run <code>mmafmctl resumeRequeued</code> so that the requeued messages are executed at home again. If <code>mmafmctl resumeRequeued</code> is not run by an administrator, AFM executes the message in the regular order of message executions from the cache cluster to the home cluster. Running <code>mmfsadm saferdump afm all</code> on the gateway node displays the queued messages. The requeued messages are displayed in the dumps. An example:</p> <pre>c12c4apv13.gpfs.net: Normal Queue: (listed by execution order) (state: Active)c12c4apv13.gpfs.net: Write [612457.552962] requeued file3 (43 @ 293) chunks 0 bytes 0 0</pre>
Why is a fileset in the Unmounted or Disconnected state when parallel I/O is set up?	<p>Filesets that are using a mapping target go to the Disconnected mode if the NFS server of the MDS is unreachable, even if NFS servers of all participating gateways are reachable. The NFS server of the MDS must be checked to fix this problem.</p>
How to clean unmount of the secondary filesystem fails if there are caches using GPFS protocol as backend?	<p>To have a clean unmount of secondary filesystem, the filesystem should first be unmounted on the primary cluster where it has been remotely mounted and then the secondary filesystem should be unmounted. It might not be possible to unmount the remote file system from all nodes in the cluster until the relevant primary is unlinked or the local file system is unmounted.</p> <p>Force unmount/shutdown/crash of remote cluster results panic of the remote filesystem at primary cluster and queue gets dropped, next access to fileset runs recovery. However this should not affect primary cluster.</p>

Table 57. Common questions in AFM DR with their resolution (continued)

Issue	Resolution
<p>'DF' command hangs on the primary cluster</p>	<p>On RHEL 7.0 or later, <b>df</b> does not support hidden NFS mounts. As AFM uses regular NFS mounts on the gateway nodes, this change causes commands like <b>df</b> to hang if the secondary gets disconnected. The following workaround can be used that allows NFS mounts to continue to be hidden:</p> <p>Remove /etc/mtab symlink, and create new file /etc/mtab and copy /proc/mounts to /etc/mtab file during startup. In this solution, mtab file might go out of sync with /proc/mounts</p>
<p>What does NeedsResync state imply ?</p>	<p>NeedsResync state does not necessarily mean a problem. If this state is during a conversion or recovery, the problem gets automatically fixed in the subsequent recovery. You can monitor the <b>mmafmctl \$fsname getstate</b> to check if its queue number is changing. And also can check the gpfs logs and for any errors, such as unmounted.</p>
<p>Is there a single command to delete all RPO snapshots from a primary fileset?</p>	<p>No. All RPOs need to be manually deleted.</p>
<p>Suppose there are more than 2 RPO snapshots on the primary. Where did these snapshots come from?</p>	<p>Check the queue. Check if recovery happened in the recent past. The extra snapshots will get deleted during subsequent RPO cycles.</p>

---

## Chapter 24. Transparent Cloud Tiering issues

This topic describes the common issues (along with workarounds) that you might encounter while using Transparent Cloud Tiering.

### **Migration of files fails with GpfsConnectorException:<init> - Error converting path to DMAPI handle**

This is caused if the file names exceed the GPFS limit of 4 K bytes.

Ensure that file names are within IBM Spectrum Scale limits.

### **Migration/Recall failures**

If a migration or recall fails, simply retry the policy or CLI command that failed two times after clearing the condition causing the failure. This works because the Transparent Cloud Tiering service is idempotent.

### **Starting or stopping Transparent Cloud Tiering service fails with the The Multi Cloud Storage Gateway seems to be in startup phase message**

This is typically caused if the Gateway service is killed manually by using the `kill` command, without the graceful shutdown by using the `mmcloudgateway service stop` command.

Remove the Gateway-related lock files present under `/var/lock/mcstore/` to allow the Gateway service to start again. If there are exceptions related to SQL in the logs, remove the `CloudDirectoryDb` file under `/<gpfs-filesystem- associated-to-cloud>/mcstore/`.

### **Adding a cloud account to configure IBM Cloud Object Storage fails with this error, 56: Cloud Account Validation failed. Invalid credential for Cloud Storage Provider. Details: Endpoint URL Validation Failed, invalid username or password.**

Ensure that the appropriate user role is set through IBM Cloud Object StoragedsNet Manager GUI.

### **HTTP Error 401 Unauthorized exception while you configure a cloud account**

This issue happens when the time between the object storage server and the Gateway node is not synced up.

Sync up the time with an NTP server and retry the operation.

### **Account creation command fails after a long wait and IBM Cloud Object Storage displays an error message saying that the vault cannot be created; but the vault is created**

When you look at the IBM Cloud Object Storage manager UI, you see that the vault exists. This problem can occur if Transparent Cloud Tiering does not receive a successful return code from IBM Cloud Object Storage for the vault creation request.

The most common reason for this problem is that the threshold setting on the vault template is incorrect. If you have 6 IBM Cloud Object Storage slicestors and the write threshold is 6, then IBM Cloud Object Storage expects that all the slicestors are healthy. Check the IBM Cloud Object Storage manager UI. If any slicestors are in a warning or error state, update the threshold of the vault template.

## Account creation command fails with error MCSTG00065E, but the data vault and the metadata vault exist

The full error message for this error is as follows:

```
MCSTG00065E: Command Failed with following reason: Error checking existence of, or creating, cloud container container_name or cloud metadata container container_name.meta.
```

But the data vault and the metadata vault are visible on the IBM Cloud Object Storage UI.

This error can occur if the metadata vault was created but its name index is disabled. To resolve this problem, do one of the follow actions:

- Enter the command again with a new vault name and vault template.
- Delete the vault on the IBM Cloud Object Storage UI and run the command again with the correct *--metadata-location*.

**Note:** It is a good practice to disable the name index of the data vault. The name index of the metadata vault must be enabled.

## File or metadata transfer fails with `koffLimitedRetryHandler:logError - Cannot retry after server error, command has exceeded retry limit, followed by RejectingHandler:exceptionCaught - Caught an exception com.ibm.gpfsconnector.messages.GpfsConnectorException: Unable to migrate`

This is most likely caused by a network connectivity and/or bandwidth issue.

Make sure that the network is functioning properly and retry the operation.

For policy-initiated migrations, IBM Spectrum Scale policy scan might automatically retry the migration of the affected files on a subsequent run.

## **gpfs.snap: An Error was detected on node XYZ while invoking a request to collect the snap file for Transparent Cloud Tiering: (return code: 137).**

If the `gpfs.snap` command fails with this error, increase the value of the *timeout* parameter by using the `gpfs.snap --timeout Seconds` option.

## Statistics are missing when the Transparent Cloud Tiering service on some nodes are manually stopped

In the event of manually stopping the Transparent Cloud Tiering service on some nodes, ensure that the `pmsensors` service is restarted in order to display the statistics values of the nodes under `GUI>Statistics>Resource>Transparent Cloud Tiering`. Restart the sensors by using this command: `service pmsensors restart`.



E!

---

## E! Chapter 25. Recovery procedures

E! You need to perform certain procedures to recover the system to minimize the impact of the issue E! reported in the system and to bring the system back to the normal operating state. The procedure E! re-creates the system by using saved configuration data or by restarting the affected services.

E!

---

### E! Restoring data and system configuration

E! You can back up and restore the configuration data for the system after preliminary recovery tasks are E! completed.

E! You can maintain your configuration data for the system by completing the following tasks:

- E! 1. Backing up the configuration data
- E! 2. Restoring the configuration data
- E! 3. Deleting unwanted backup configuration data files

E! The following topics describes how to perform backup and restore data and configuration in the IBM E! Spectrum Scale system:

- E! • *Protocols cluster disaster recovery in IBM Spectrum Scale: Administration Guide*
- E! • *Restore procedure with SOBAR in IBM Spectrum Scale: Administration Guide*
- E! • *Encryption and backup/restore in IBM Spectrum Scale: Administration Guide*
- E! • *Backup and restore with storage pools in IBM Spectrum Scale: Administration Guide*
- E! • *Restoring quota files in IBM Spectrum Scale: Administration Guide*
- E! • *Backing up and restoring protocols and CES configuration information in IBM Spectrum Scale: Administration Guide*
- E! • *Failback or restore steps for object configuration in IBM Spectrum Scale: Administration Guide*

E!

---

### E! Automatic recovery

E! The IBM Spectrum Scale recovers itself from certain issues without manual intervention.

E! The following automatic recovery options are available in the system:

- E! • Failover of CES IP addresses to recover from node failures. That is, if any important service or protocol E! service is broken on a node, the system changes the status of that node as *Failed* and moves the public E! IPs to healthy nodes in the cluster. This change is transparent to the users of protocol services and they E! do not experience any service outage.
- E! • In case of any errors with the protocol services such as SMB and Object, the system restarts the E! corresponding daemons. If restarting the protocol service daemons does not resolve the issue and the E! maximum retry count is reached, the system changes the status of the node as *Failed*. The protocol E! service restarts are logged in the event log. Issue either `mmhealth node eventlog` or `mmces events list` E! commands to view the details of such events.

E!

---

### E! Upgrade recovery

E! Use this information to recover from a failed upgrade.

E! A failed upgrade might leave a cluster in a state of containing multiple code levels. It is important to E! analyze console output to determine which nodes or components were upgraded prior to the failure and E! which node or component was in the process of being upgraded when the failure occurred.

E! Once the problem has been isolated, a healthy cluster state must be achieved prior to continuing the upgrade. Use the **mmhealth** command in addition to the **mmces state show -a** command to verify that all services are up. It might be necessary to manually start services that were down when the upgrade failed. Starting the services manually helps achieve a state in which all components are healthy prior to continuing the upgrade.

E! For more information about verifying service status, see **mmhealth command** and **mmces state show command** in *IBM Spectrum Scale: Command and Programming Reference*.

---

## E! Directed maintenance procedures

E! The directed maintenance procedures (DMPs) assist you to repair a problem when you select the action **Run fix procedure** on a selected event from the **Monitoring > Events** page. DMPs are present for only a few events reported in the system.

E! The following table provides details of the available DMPs and the corresponding events.

E! *Table 58. DMPs*

E! DMP	E! Event ID
E! Start NSD	MS0043
E! Start cluster	MS0063
E! Start GPFS daemon	MS0054
E! Increase fileset space	MS0147
E! Synchronize Node Clocks	MS0242
E! Start performance monitoring collector service	MS0364
E! Start performance monitoring sensor service	MS0372

### E! Start NSD

E! The Start NSD DMP assists to start NSDs that are not working.

E! The following are the corresponding event details and the proposed solution:

- E! • **Event ID:** MS0043
- E! • **Problem:** The availability of an NSD is changed to “down”.
- E! • **Solution:** Recover the NSD

E! The DMP provides the option to start the NSDs that are not functioning. If multiple NSDs are down, you can select whether to recover only one NSD or all of them.

E! The system issues the **mmchdisk** command to recover NSDs as given in the following format:

```
E! /usr/lpp/mmfs/bin/mmchdisk <device> start -d <disk description>
```

E! For example: `/usr/lpp/mmfs/bin/mmchdisk r1_FS start -d G1_r1_FS_data_0`

### E! Start cluster

E! The start cluster DMP assists you to start the IBM Spectrum Scale cluster.

E! The following are the corresponding event details and the proposed solution:

- E! • **Event ID:** MS0063
- E! • **Problem:** The cluster is down. Either quorum was lost or it was administratively shut down.
- E! • **Solution:** Restart individual nodes at least until the quorum is reached.

E! The procedure issues the **mmstartup -a** command as shown in the following example:

```
E! /usr/lpp/mmfs/bin/mmstartup -a
```

E! When the cluster is started, the procedure runs the following refresh task to remove the event from the event list:

```
E! runtask HOST_STATES -c << Cluster_Id>>
```

## E! **Start GPFS daemon**

E! When the GPFS daemon is down, GPFS functions do not work properly on the node.

E! The following are the corresponding event details and the proposed solution:

E! • **Event ID:** MS0054

E! • **Problem:** The GPFS daemon is down. GPFS is not operational on node.

E! • **Solution:** Start GPFS daemon.

E! The system issues the **mmstartup -N** command to restart GPFS daemon as given in the following format:

```
E! /usr/lpp/mmfs/bin/mmstartup -N <Node>
```

E! For example: `usr/lpp/mmfs/bin/mmstartup -N gss-05.localnet.com`

E! When the GPFS daemon is started, the procedure runs the following refresh task to remove the event from the event list:

```
E! runtask HOST_STATES -c <ClusterID>
```

## E! **Increase fileset space**

E! The system needs inodes to allow I/O on a fileset. If the inodes allocated to the fileset are exhausted, you need to either increase the number of maximum inodes or delete the existing data to free up space.

E! The procedure helps to increase the maximum number of inodes by a percentage of the already allocated inodes. The following are the corresponding event details and the proposed solution:

E! • **Event ID:** MS0147

E! • **Problem:** The inode usage in the fileset reached an exhausted level

E! • **Solution:** increase the maximum number of inodes

E! The system issues the **mmchfileset** command to recover NSDs as given in the following format:

```
E! /usr/lpp/mmfs/bin/mmchfileset <Device> <Fileset> --inode-limit <inodesMaxNumber>
```

E! For example: `/usr/lpp/mmfs/bin/mmchfileset r1_FS testFileset --inode-limit 2048`

E! After increasing the number of inodes, the procedure runs the following refresh task to remove the event from the event list:

```
E! runtask FILESETS_NOUSAGE -c <ClusterID>
```

## E! **Synchronize node clocks**

E! The time must be in sync with the time set on the GUI node. If the time is not in sync, the data that is displayed in the GUI might be wrong or it does not even display the details. For example, the GUI does not display the performance data if time is not in sync.

E! The procedure assists to fix timing issue on a single node or on all nodes that are out of sync. The following are the corresponding event details and the proposed solution:

E! • **Event ID:** MS0242

E! • **Limitation:** This DMP is not available in sudo wrapper clusters. In a sudo wrapper cluster, the user name is different from 'root'. The system detects the user name by finding the parameter `GPFSS_USER=<user name>`, which is available in the file `/usr/lpp/mmfs/gui/conf/gpfsgui.properties`.

E! • **Problem:** The time on the node is not synchronous with the time on the GUI node. It differs more than 1 minute.

E! • **Solution:** Synchronize the time with the time on the GUI node.

E! The system issues the `sync_node_time` command as given in the following format to synchronize the time in the nodes:

```
E! /usr/lpp/mmfs/gui/bin/sync_node_time <nodeName>
```

E! For example: `/usr/lpp/mmfs/gui/bin/sync_node_time c55f06n04.gpfs.net`

E! After the clocks are restarted, the procedure runs the following refresh task to remove the event from the event list:

```
E! runtask TIMESYNC_MONITOR -c <ClusterID>
```

## E! **Start performance monitoring collector service**

E! The collector services on the GUI node must be functioning properly to display the performance data in the IBM Spectrum Scale management GUI.

E! The following are the corresponding event details and the proposed solution:

E! • **Event ID:** MS0364

E! • **Limitation:** This DMP is not available in sudo wrapper clusters when a remote *pmcollector* service is used by the GUI. A remote *pmcollector* service is detected in case a different value than localhost is specified in the `ZIMonAddress` in file, which is located at: `/usr/lpp/mmfs/gui/conf/gpfsgui.properties`. In a sudo wrapper cluster, the user name is different from 'root'. The system detects the user name by finding the parameter `GPFSS_USER=<user name>`, which is available in the file `/usr/lpp/mmfs/gui/conf/gpfsgui.properties`.

E! • **Problem:** The performance monitoring collector service *pmcollector* is in inactive state.

E! • **Solution:** Issue the `systemctl status pmcollector` to check the status of the collector. If *pmcollector* service is inactive, issue `systemctl start pmcollector`.

E! The system restarts the performance monitoring services by issuing the `systemctl restart pmcollector` command.

E! The performance monitoring collector service might be on some other node of the current cluster. In this case, the DMP first connects to that node, then restarts the performance monitoring collector service.

```
E! ssh <nodeAddress> systemctl restart pmcollector
```

E! For example: `ssh 10.0.100.21 systemctl restart pmcollector`

E! When the cluster is started, the procedure runs the following refresh task to remove the event from the event list:

```
E! runtask PM_MONITOR -c <ClusterID>
```

E! In a sudo wrapper cluster, when collector on remote node is down, the DMP does not restart the collector services by itself. You need to do it manually.

## E! **Start performance monitoring sensor service**

E! You need to start the sensor service to get the performance details in the collectors. If sensors and collectors are not started, the GUI and CLI do not display the performance data in the IBM Spectrum Scale management GUI.

E! The following are the corresponding event details and the proposed solution:

E! • **Event ID:** MS0372

E! • **Limitation:** This DMP is not available in sudo wrapper clusters. In a sudo wrapper cluster, the user name is different from 'root'. The system detects the user name by finding the parameter  
E! GPFS\_USER=<user name>, which is available in the file /usr/lpp/mmfs/gui/conf/gpfsgui.properties.

E! • **Problem:** The performance monitoring sensor service *pmsensor* is not sending any data. The service might be down or the difference between the time of the node and the node hosting the performance monitoring collector service *pmcollector* is more than 15 minutes.

E! • **Solution:** Issue **systemctl status pmsensor** to verify the status of the sensor service. If *pmsensor* service is inactive, issue **systemctl start pmsensor**.

E! The system restarts the sensors by issuing **systemctl restart pmsensors** command.

E! For example: `ssh gss-15.localnet.com systemctl restart pmsensors`

E! When the cluster is started, the procedure runs the following refresh task to remove the event from the event list:

E! `runtask PM_MONITOR -c <ClusterID>`

E!



E!

---

## E! Chapter 26. Support for troubleshooting

E! This topic describes the support that is available for troubleshooting any issues that you might encounter E! while using IBM Spectrum Scale .

E!

### E! Contacting IBM support center

E! Specific information about a problem such as: symptoms, traces, error logs, GPFS logs, and file system E! status is vital to IBM in order to resolve a GPFS problem.

E! Obtain this information as quickly as you can after a problem is detected, so that error logs will not wrap E! and system parameters that are always changing, will be captured as close to the point of failure as E! possible. When a serious problem is detected, collect this information and then call IBM. For more E! information, see:

E! • “Information to be collected before contacting the IBM Support Center”

E! • “How to contact the IBM Support Center” on page 369.

### E! Information to be collected before contacting the IBM Support Center

E! For effective communication with the IBM Support Center to help with problem diagnosis, you need to E! collect certain information.

### E! Information to be collected for all problems related to GPFS

E! Regardless of the problem encountered with GPFS, the following data should be available when you E! contact the IBM Support Center:

E! 1. A description of the problem.

E! 2. Output of the failing application, command, and so forth.

E! 3. A tar file generated by the **gpfs.snap** command that contains data from the nodes in the cluster. In E! large clusters, the **gpfs.snap** command can collect data from certain nodes (for example, the affected E! nodes, NSD servers, or manager nodes) using the **-N** option.

E! If the **gpfs.snap** command cannot be run, collect these items:

E! a. Any error log entries relating to the event:

E! • On an AIX node, issue this command:

E! `errpt -a`

E! • On a Linux node, create a tar file of all the entries in the **/var/log/messages** file from all nodes in E! the cluster or the nodes that experienced the failure. For example, issue the following command E! to create a tar file that includes all nodes in the cluster:

E! `mmdsh -v -N all "cat /var/log/messages" > all.messages`

E! • On a Windows node, use the **Export List...** dialog in the Event Viewer to save the event log to a E! file.

E! b. A master GPFS log file that is merged and chronologically sorted for the date of the failure (see E! “Creating a master GPFS log file” on page 141).

E! c. If the cluster was configured to store dumps, collect any internal GPFS dumps written to that E! directory relating to the time of the failure. The default directory is **/tmp/mmfs**.

E! d. On a failing Linux node, gather the installed software packages and the versions of each package E! by issuing this command:

E! `rpm -qa`

- E! e. On a failing AIX node, gather the name, most recent level, state, and description of all installed software packages by issuing this command:  
E! `lspp -l`
- E! f. File system attributes for all of the failing file systems, issue:  
E! `mmlsfs Device`
- E! g. The current configuration and state of the disks for all of the failing file systems, issue:  
E! `mmlsdisk Device`
- E! h. A copy of file `/var/mmfs/gen/mmsdrfs` from the primary cluster configuration server.
- E! 4. For Linux on z Systems, collect the data of the operating system as described in the *Linux on z Systems Troubleshooting Guide* ([www.ibm.com/support/knowledgecenter/linuxonibm/liaaf/lnz\\_r\\_sv.html](http://www.ibm.com/support/knowledgecenter/linuxonibm/liaaf/lnz_r_sv.html)).
- E! 5. If you are experiencing one of the following problems, see the appropriate section before contacting the IBM Support Center:
  - E! • For delay and deadlock issues, see “Additional information to collect for delays and deadlocks.”
  - E! • For file system corruption or MMFS\_FSSTRUCT errors, see “Additional information to collect for file system corruption or MMFS\_FSSTRUCT errors.”
  - E! • For GPFS daemon crashes, see “Additional information to collect for GPFS daemon crashes.”

### E! **Additional information to collect for delays and deadlocks**

E! When a delay or deadlock situation is suspected, the IBM Support Center will need additional information to assist with problem diagnosis. If you have not done so already, ensure you have the following information available before contacting the IBM Support Center:

- E! 1. Everything that is listed in “Information to be collected for all problems related to GPFS” on page 367.
- E! 2. The deadlock debug data collected automatically.
- E! 3. If the cluster size is relatively small and the `maxFilesToCache` setting is not high (less than 10,000), issue the following command:  
E! `gpfs.snap --deadlock`
- E! If the cluster size is large or the `maxFilesToCache` setting is high (greater than 1M), issue the following command:  
E! `gpfs.snap --deadlock --quick`

### E! **Additional information to collect for file system corruption or MMFS\_FSSTRUCT errors**

E! When file system corruption or MMFS\_FSSTRUCT errors are encountered, the IBM Support Center will need additional information to assist with problem diagnosis. If you have not done so already, ensure you have the following information available before contacting the IBM Support Center:

- E! 1. Everything that is listed in “Information to be collected for all problems related to GPFS” on page 367.
- E! 2. Unmount the file system everywhere, then run `mmfsck -n` in offline mode and redirect it to an output file.

E! The IBM Support Center will determine when and if you should run the `mmfsck -y` command.

### E! **Additional information to collect for GPFS daemon crashes**

E! When the GPFS daemon is repeatedly crashing, the IBM Support Center will need additional information to assist with problem diagnosis. If you have not done so already, ensure you have the following information available before contacting the IBM Support Center:

- E! 1. Everything that is listed in “Information to be collected for all problems related to GPFS” on page 367.



- E! 2. Ensure the `/tmp/mmfs` directory exists on all nodes. If this directory does not exist, the GPFS daemon will not generate internal dumps.
- E! 3. Set the traces on this cluster and *all* clusters that mount any file system from this cluster:  
E! `mmtracectl --set --trace=def --trace-recycle=global`
- E! 4. Start the trace facility by issuing:  
E! `mmtracectl --start`
- E! 5. Recreate the problem if possible or wait for the assert to be triggered again.
- E! 6. Once the assert is encountered on the node, turn off the trace facility by issuing:  
E! `mmtracectl --off`  
E! If traces were started on multiple clusters, `mmtracectl --off` should be issued immediately on all clusters.
- E! 7. Collect `gpfs.snap` output:  
E! `gpfs.snap`

## E! How to contact the IBM Support Center

E! The IBM Support Center is available for various types of IBM hardware and software problems that E! GPFS customers may encounter.

E! These problems include the following:

- E! • IBM hardware failure
- E! • Node halt or crash not related to a hardware failure
- E! • Node hang or response problems
- E! • Failure in other software supplied by IBM

### E! If you have an IBM Software Maintenance service contract

E! If you have an IBM Software Maintenance service contract, contact the IBM Support Center, as E! follows:

E! Your location	E! Method of contacting the IBM Support Center
E! In the United States	E! Call <b>1-800-IBM-SERV</b> for support.
E! Outside the United States	E! Contact your local IBM Support Center or see the Directory of worldwide contacts ( <a href="http://www.ibm.com/planetwide">www.ibm.com/planetwide</a> ).

E! When you contact the IBM Support Center, the following will occur:

- E! 1. You will be asked for the information you collected in “Information to be collected before contacting the IBM Support Center” on page 367.
- E! 2. You will be given a time period during which an IBM representative will return your call. Be sure that the person you identified as your contact can be reached at the phone number you provided in the PMR.
- E! 3. An online Problem Management Record (PMR) will be created to track the problem you are reporting, and you will be advised to record the PMR number for future reference.
- E! 4. You may be requested to send data related to the problem you are reporting, using the PMR number to identify it.
- E! 5. Should you need to make subsequent calls to discuss the problem, you will also use the PMR number to identify the problem.

### E! If you do not have an IBM Software Maintenance service contract

E! If you do not have an IBM Software Maintenance service contract, contact your IBM sales E! representative to find out how to proceed. Be prepared to provide the information you collected E! in “Information to be collected before contacting the IBM Support Center” on page 367.

E! For failures in non-IBM software, follow the problem-reporting procedures provided with that product.

E!

---

## E! **Call home**

E! The call home feature automatically notifies IBM Support if certain types of events occur in the system.

E! Using this information, IBM Support can contact the system administrator in case of any issues.

E! Configuring call home reduces the response time for IBM Support to address the issues.

E! The details are collected from individual nodes that are marked as call home child nodes in the cluster.

E! The details from each child node are collected by the call home node. You need to create a call home

E! group by grouping call home child nodes. One of the nodes in the group is configured as the call home

E! node and it performs data collection and upload.

E! The data gathering and upload can be configured individually on each group. Use the groups to reflect

E! logical units in the cluster. For example, it is easier to manage when you create a group for all CES nodes

E! and another group for all non-CES nodes.

E! For more information on how to configure and manage the call home feature, see Chapter 7, “Monitoring

E! the IBM Spectrum Scale system by using call home,” on page 117.

## Chapter 27. References

The IBM Spectrum Scale system displays messages if it encounters any issues when you configure the system. The message severity tags helps to assess the severity of the issue.

### Events

The following tables list the RAS events that are applicable to various components of the IBM Spectrum Scale system.

**Note:** The recorded events are stored in local database on each node. The user can get a list of recorded events using `mmces events list` command. The recorded events can also be displayed through GUI.

Table 59. Events for the AUTH component

Event	Event Type	Severity	Message	Description	Cause	User Action
ads_down	STATE_CHANGE	ERROR	External ADS server is unresponsive	External ADS server is unresponsive	The local node is unable to connect to any ADS server.	Local node is unable to connect to any Active Directory Service server. Verify network connection and check that Active Directory Service server(s) are operational.
ads_failed	STATE_CHANGE	ERROR	local winbindd is unresponsive	local winbindd is unresponsive.	The local winbindd does not respond to ping requests. It is needed for Active Directory Service.	Local winbindd does not respond to ping requests. Try to restart winbindd, and if not successful, perform winbindd troubleshooting.
ads_up	STATE_CHANGE	INFO	external ADS server is up	External ADS server is up.		External Active Directory Service server is operational, no user action required.
ads_warn	INFO	WARNING	external ADS server monitoring returned unknown result	External ADS server monitoring returned unknown result.	An internal error occurred while monitoring the external ADS server.	An internal error occurred while monitoring external Active Directory Service server. Perform trouble check.
ldap_down	STATE_CHANGE	ERROR	external LDAP server {0} is unresponsive	External LDAP server <LDAP server> is unresponsive.	The local node is unable to connect to the LDAP server.	Local node is unable to connect to LDAP server. Verify network connection and check that LDAP server is operational.
ldap_up	STATE_CHANGE	INFO	external LDAP server {0} is up	The external LDAP server is operational.		NA
nis_down	STATE_CHANGE	ERROR	external NIS server {0} is unresponsive	External NIS server <NIS server> is unresponsive.	The local node is unable to connect to any NIS server.	Local node is unable to connect to any Network Information Server. Verify network connection and check that Network Information Service server(s) are operational.

Table 59. Events for the AUTH component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
nis_failed	STATE_CHANGE	ERROR	ypbind is unresponsive	ypbind is unresponsive.	The local ypbind daemon does not respond.	Local ypbind daemon does not respond. Try to restart ypbind, and if not successful, perform ypbind troubleshooting.
nis_up	STATE_CHANGE	INFO	external NIS server {0} is up	External Network Information Service (NIS) server is operational.		NA
nis_warn	INFO	WARNING	external NIS monitoring returned unknown result	The external NIS server monitoring returned an unknown result.	An internal error occurred while monitoring external Network Information Service server.	Perform trouble check.
sssd_down	STATE_CHANGE	ERROR	SSSD process not running	The SSSD process not running.	The SSSD authentication service is not running.	Perform trouble check.
sssd_restart	INFO	INFO	SSSD process was not running. Trying to start it	Attempt to start the SSSD authentication process.	The SSSD process was not running.	NA
sssd_up	STATE_CHANGE	INFO	SSSD process is now running	SSSD process is now running.	The SSSD authentication process is running.	NA
sssd_warn	INFO	WARNING	SSSD process monitoring returned unknown result	SSSD authentication process monitoring returned an unknown result.	An internal error occurred while monitoring the SSSD.	Perform trouble check.
wnbd_down	STATE_CHANGE	ERROR	WINBINDD process not running	The WINBINDD authentication process not running.	The WINBINDD authentication service is not running.	Verify the configuration and Active Directory connection.
wnbd_restart	INFO	INFO	WINBINDD process was not running. Trying to start it	Attempt to start the WINBINDD authentication process.	The WINBINDD process was not running.	NA
wnbd_up	STATE_CHANGE	INFO	WINBINDD process is now running	The WINBINDD authentication service is operational.		NA
wnbd_warn	INFO	WARNING	WINBINDD process monitoring returned unknown result	WINBINDD process monitoring returned an unknown result.	An internal error occurred during the monitoring of WINBINDD.	Perform trouble check.
yp_down	STATE_CHANGE	ERROR	YPBIND process not running	The YPBIND process not running.	The YPBIND authentication service is not running.	Perform trouble check.
yp_restart	INFO	INFO	YPBIND process was not running. Trying to start it	Attempt to start the YPBIND process.	The YPBIND process is not running.	NA
yp_up	STATE_CHANGE	INFO	YPBIND process is now running	The YPBIND service is operational.		NA
yp_warn	INFO	WARNING	YPBIND process monitoring returned unknown result	The YPBIND process monitoring returned an unknown result.	An internal error occurred while monitoring YPBIND.	Perform trouble determination.

Table 60. Events for the CESNetwork component

Event	Event Type	Severity	Message	Description	Cause	User Action
ces_bond_down	STATE_CHANGE	ERROR	All slaves of the CES-bond {0} are down	All slaves of a CES-network bond are down.	All slaves of this network bond went down.	Check the bonding configuration, the network configuration and cabling of all slaves of the bond.
ces_bond_degraded	STATE_CHANGE	INFO	Some slaves of the CES-bond {0} went down	Some of the CES-bond parts are malfunctioning.	Some slaves of the bond are not functioning properly.	Check the bonding configuration, the network configuration and cabling of the malfunctioning slaves of the bond.
ces_bond_up	STATE_CHANGE	INFO	All slaves of the CES bond {0} are working as expected	This CES bond is functioning properly.	All slaves of this network bond are functioning properly.	NA
ces_disable_node_network	INFO	INFO	Network was disabled	Clean up after a 'mmchnode --ces-disable' command. The network configuration is modified accordingly.	Informational message. Clean up after a 'mmchnode --ces-disable' command.	NA
ces_enable_node_network	INFO	INFO	Network was enabled	Called to handle any network specific issues involved after a 'mmchnode --ces-enable' command.	Informational message. Called after a 'mmchnode --ces-enable' command.	NA
ces_many_tx_errors	STATE_CHANGE	ERROR	CES NIC {0} had many TX errors since the last monitoring cycle	This CES-related NIC had many TX errors since the last monitoring cycle.	/proc/net/dev lists much more TX errors for this adapter since the last monitoring cycle.	Check cable contacts or try a different cable.
ces_network_connectivity_down	STATE_CHANGE	ERROR	CES NIC {0} can not connect to the gateway	This CES-related network adapter can connect to the gateway.	The gateway responds to the sent connections-checking packets.	Check the network configuration of the network adapter, path to the gateway and gateway itself.
ces_network_connectivity_up	STATE_CHANGE	INFO	CES NIC {0} can connect to the gateway	A CES-related NIC can connect to the gateway.		NA
ces_network_down	STATE_CHANGE	ERROR	CES NIC {0} is down	This CES-related network adapter is down.	This network adapter was disabled.	Enable this network adapter or check for problems in system logs.
ces_network_found	INFO	INFO	CES NIC {0} was found	A new CES-related network adapter was found.	A new NIC, which is relevant for the CES, is listed by 'ip a'.	NA

Table 60. Events for the CESNetwork component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
ces_network_ips_down	STATE_CHANGE	WARNING	No CES relevant NICs detected	No CES-relevant network adapters or CES IPs detected, which makes the node unavailable for the CES clients.	No network adapters have the CES-relevant IPs, which makes the node unavailable for the CES clients.	Check the network cabling and network infrastructure.
ces_network_ips_up	STATE_CHANGE	INFO	CES-relevant IPs are served by found NICs	CES-relevant IPs are served by network adapters	At least one CES-relevant IP is assigned to a network adapter.	NA
ces_network_link_down	STATE_CHANGE	ERROR	Physical link of the CES NIC {0} is down	The physical link of this CES-related adapter is down.	The flag LOWER_UP is not set for this NIC in the output of 'ip a'.	Check the cabling of this network adapter.
ces_network_link_up	STATE_CHANGE	INFO	Physical link of the CES NIC {0} is up	The physical link of this CES-related adapter is up.	The flag LOWER_UP is set for this NIC in the output of 'ip a'.	NA
ces_network_up	STATE_CHANGE	INFO	CES NIC {0} is up	This CES-related network adapter is up.	This network adapter is enabled.	NA
ces_network_vanished	INFO	INFO	CES NIC {0} has vanished	One of CES-related network adapters can not be detected anymore.	One of the previously monitored NICs is not listed by 'ip a' anymore.	NA
ces_no_tx_errors	STATE_CHANGE	INFO	CES NIC {0} had no or a tiny number of TX errors.	A CES-related NIC had no or an insignificant number of TX errors.	/proc/net/dev lists no or an insignificant number of TX errors for this adapter since the last monitoring cycle.	Check the network cabling and network infrastructure.
ces_startup_network	INFO	INFO	CES network service was started	Information that the CES network has started	CES network IPs are started	NA
handle_network_problem_info	INFO	INFO	Handle network problem - Problem: {0},Argument: {1}	Information about network related reconfigurations. This can be enable or disable IPs, assign or unassign IPs for example.	A change in the network configuration. Details are part of the information message	NA
move_cesip_from	INFO	INFO	Address {0} was moved from this node to node {1}	Information that a CES IP address was moved from the current node to an other node.	Rebalancing of CES IP addresses.	NA

Table 60. Events for the CESNetwork component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
move_cesips_info	INFO	INFO	A move request for ip addresses was executed	CES IP addresses can be moved in case of node failovers from one node to one or more other nodes.", "This message is logged on a node observing this, not necessarily on any affected node.	A CES IP movement was detected.	NA
move_cesip_to	INFO	INFO	Address {0} was moved from node {1} to this node	Information that a CES IP address was moved from an other node to the current node.	Rebalancing of CES IP addresses.	NA

Table 61. Events for the CloudGateway component

Event	Event Type	Severity	Message	Description	Cause	User Action
cloudgateway_down	STATE_CHANGE	ERROR	Cloud gateway service is down	The cloud gateway service is down		
cloudgateway_starting	STATE_CHANGE	WARNING	Cloud gateway service is starting up	The cloud gateway service is starting.		
cloudgateway_up	STATE_CHANGE	INFO	Cloud gateway service is up	The cloud gateway service is running.		NA
cloudgateway_warn	INFO	WARNING	cloud gateway monitoring returned unknown result	The cloud gateway check returned unknown result.		Perform trouble determination.
cloudgw_restart	INFO	WARNING	The cloud gateway service failed. Trying to recover	Attempt to start the cloud gateway process.	The cloud gateway process was not running	NA

Table 62. Events for the DISK component

Event	Event Type	Severity	Message	Description	Cause	User Action
disk_down	STATE_CHANGE	WARNING	Disk {0} is reported as not up	A disk is reported as down.	This can indicate a hardware issue.	If the down state is unexpected, then refer to the section Disk issues in the Troubleshooting guide.
disk_up	STATE_CHANGE	INFO	Disk {0} is up	Disk is up.	A disk was detected in up state	N/A
disk_found	INFO	INFO	The disk {0} was found	A disk was detected	A disk was detected.	N/A
disk_vanished	INFO	INFO	The disk {0} has vanished	A declared disk was not detected.	A disk is not in use for a file system. This could be a valid situation.	N/A

Table 63. Events for the file system component

Event	Event Type	Severity	Message	Description	Cause	User Action
filesystem_found	INFO	INFO	file system {0} was found	A file system listed in the IBM Spectrum Scale configuration was detected.	N/A	N/A

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
filesystem_vanished	INFO	INFO	file system {0} has vanished	A file system listed in the IBM Spectrum Scale configuration was not detected.	A file system, listed in the IBM Spectrum Scale configuration as mounted before, is not found. This could be a valid situation.	Run <code>mmismount all_local</code> to verify that all expected file system are mounted.
fs_forced_unmount	STATE_CHANGE	ERROR	The file system {0} was {1} (forced) to unmount	A file system was forced to unmount by IBM Spectrum Scale.	A situation like a kernel panic might initiate the unmount	Check error messages and the error log for further details. Also see the topic File system forced unmount in the IBM Spectrum Scale documentation: <a href="#">Troubleshooting. File system issues</a> .
fserrallocblock	STATE_CHANGE	ERROR	FS={0},ErrNo={1},Msg={2}	Corrupted alloc segment detected while attempting to alloc disk block	A file system corruption detected.	Check error message for details and the <code>mmfs.log.latest</code> log for further details. See the topic <a href="#">Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering. Managing file systems</a> . If the file system is severely damaged, the best course of action is to follow the procedures in section: <a href="#">Additional information to collect for file system corruption or MMFS_FSSTRUCT errors</a> .
fserrbadacref	STATE_CHANGE	ERROR	FS={0},ErrNo={1},Msg={2}	File references invalid ACL.	A file system corruption detected.	Check error message for details and the <code>mmfs.log.latest</code> log for further details. See the topic <a href="#">Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering. Managing file systems</a> . If the file system is severely damaged, the best course of action is to follow the procedures in section: <a href="#">Additional information to collect for file system corruption or MMFS_FSSTRUCT errors</a> .



Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrbadirblock	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Invalid directory block	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_ FSSTRUCT errors.
fserrbadiskaddrindex	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Bad disk index in disk address	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrbaddiskaddrsector	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Bad sector number in disk address, or start sector plus length would exceed the size of the disk	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering. Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.
fserrbaddittoaddr	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Invalid ditto address	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering. Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrbadinodeorgen	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Deleted inode has dir entry or the generation number do not match to the directory	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering. Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.
fserrbadinodestatus	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Inode status bad, expected deleted status	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering. Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrbadptrreplications	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Invalid computed pointer replication factors	Invalid computed pointer replication factors.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.
fserrbadreplicationcounts	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Invalid current or maximum data or metadata replication counts	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrbadxattrblock	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Invalid extended attribute block	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.
fserrcheckheaderfailed	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	CheckHeader returned an error	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrclonetree	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Invalid cloned file tree structure	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.
fserrdeallocblock	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Corrupted alloc segment detected while attempting to dealloc disk block	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrdotdotnotfound	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Unable to locate an entry ...	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering. Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.
fserrgennummismatch	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Found generation number in '.' entry does not match actual generation number of parent directory	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering. Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrinconsistentfilesetrootdir	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Inconsistent fileset / root directory, i.e. Fileset is in use, root dir '.' points to itself.	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering. Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.
fserrinconsistentfilesetsnapshot	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Inconsistent fileset / snapshot records, i.e. Fileset snapList points to a SnapItem that doesn't exist.	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering. Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.



Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrinconsistentinode	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Size data in inode is not consistent	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.
fserrindirectblock	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Indirect block header has does not match info in inode	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrindirectionlevel	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Invalid indirection level in inode	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_ FSSTRUCT errors.
fserrinodecorrupted	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Code to catch infinite loop in the lfs layer in case of a corrupted inode or directory entry.	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_ FSSTRUCT errors.

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrinodenumismatch	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Msg={2}	Found inode number in '..' entry does not match actual inode number of parent directory	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.
fserrinvalid	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Unknown error={2}	Unrecognized FSSTRUCT error received. Check documentation.	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrinvalidfilesetmetadatarecord	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Unknown error={2}	Invalid fileset metadata record	A file system corruption detected	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_ FSSTRUCT errors.
fserrinvalidsnapshotstates	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Unknown error={2}	Invalid snapshot states, i.e. more than one in an inode space being emptied (SnapBeingDeletedOne).	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_ FSSTRUCT errors.

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fserrsnapinodemodified	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Unknown error={2}	Inode was modified without saving old content to shadow inode file	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.
fserrvalidate	STATE_CHANGE	ERROR	FS={0},ErrNo={1}, Unknown error={2}	A file system corruption detected. Validation routine failed on a disk read.	A file system corruption detected.	Check error message for details and the mmfs.log.latest log for further details. See the topic Checking and repairing a file system in the IBM Spectrum Scale documentation: Administering, Managing file systems. If the file system is severely damaged, the best course of action is to follow the procedures in section: Additional information to collect for file system corruption or MMFS_FSSTRUCT errors.
fsstruct_error	STATE_CHANGE	WARNING	A file system {0} structure error was detected Err={1} msg={2}	A file system structure error was detected. The same underlying issue might cause different events.	A file system issue was detected.	Repair actions depend on the kind of error. See documentation.
fsstruct_fixed	STATE_CHANGE	INFO	A file system {0} structure error has been marked as fixed	A file system structure error was declared or detected as fixed.	A file system issue was resolved.	N/A
fs_unmount_info	INFO	INFO	The file system {0} was unmounted {1}	A file system was unmounted.	A file system was unmounted.	N/A
fs_remount_mount	STATE_CHANGE_EXTERNAL	INFO	The file system {0} was mounted {1}	A file system was mounted.	A new or previously unmounted file system was mounted.	N/A
mounted_fs_check	STATE_CHANGE	INFO	The file system {0} is mounted	A mounted file system is reported correctly.	No mount state mismatch information detected.	N/A

Table 63. Events for the file system component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
stale_mount	STATE_CHANGE	WARNING	Found stale mounts for {0}	A mount state information mismatch was detected between mmlsmount and /proc/mounts.	A file system might not be fully mounted or unmounted.	Run mmlsmount all_local to verify that all expected file systems are mounted.
unmounted_fs_ok	STATE_CHANGE	INFO	The file system {0} is probably needed, but not declared as automount	An internally mounted or a declared but not mounted file system was detected.	A declared file system is not mounted	N/A
unmounted_fs_check	STATE_CHANGE	WARNING	The file system {0} is probably needed, but not mounted	An internally mounted or a declared but not mounted file system was detected.	A file system might not be fully mounted or unmounted.	Run mmlsmount all_local to verify that all expected file systems are mounted.

Table 64. Events for the GPFS component

Event	Event Type	Severity	Message	Description	Cause	User Action
ccr_client_init_ok	STATE_CHANGE	INFO	GPFS CCR client initialization is ok {0}.	GPFS CCR client initialization is ok.	N/A	N/A
ccr_client_init_fail	STATE_CHANGE	ERROR	GPFS CCR client initialization failed Item={0},ErrMsg={1}, Failed={2}.	GPFS CCR client initialization failed. See message for details.	The item specified in the message is either not available or corrupt.	Recover this degraded node from a still intact node using the mmsdrrestore -p <NODE> command with <NODE> specifying intact node. See the mmsdrrestore man page for more details.
ccr_client_init_warn	STATE_CHANGE	WARNING	GPFS CCR client initialization failed Item={0},ErrMsg={1}, Failed={2}.	GPFS CCR client initialization failed. See message for details.	The item specified in the message is either not available or corrupt.	Recover this degraded node from a still intact node using the mmsdrrestore -p <NODE> command with <NODE> specifying intact node. See the mmsdrrestore man page for more details.
ccr_auth_keys_ok	STATE_CHANGE	INFO	The security file used by GPFS CCR is ok {0}.	The security file used by GPFS CCR is ok.	N/A	N/A
ccr_auth_keys_fail	STATE_CHANGE	ERROR	The security file used by GPFS CCR is corrupt Item={0},ErrMsg={1}, Failed={2}	The security file used by GPFS CCR is corrupt. See message for details.	Either the security file is missing or corrupt.	Recover this degraded node from a still intact node using the mmsdrrestore -p <NODE> command with <NODE> specifying intact node. See the mmsdrrestore man page for more details.

Table 64. Events for the GPFS component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
ccr_paxos_cached_ok	STATE_CHANGE	INFO	The stored GPFS CCR state is ok {0}	The stored GPFS CCR state is ok.		N/A
ccr_paxos_cached_fail	STATE_CHANGE	ERROR	The stored GPFS CCR state is corrupt Item={0},ErrMsg={1}, Failed={2}	The stored GPFS CCR state is corrupt. See message for details.	Either the stored GPFS CCR state file is corrupt or empty.	Recover this degraded node from a still intact node using the mmsdrrestore -p <NODE> command with <NODE> specifying intact node. See the mmsdrrestore man page for more details.
ccr_paxos_12_fail	STATE_CHANGE	ERROR	The stored GPFS CCR state is corrupt Item={0},ErrMsg={1}, Failed={2}	The stored GPFS CCR state is corrupt. See message for details.	tbd	Recover this degraded node from a still intact node using the mmsdrrestore -p <NODE> command with <NODE> specifying intact node. See the mmsdrrestore man page for more details.
ccr_paxos_12_ok	STATE_CHANGE	INFO	The stored GPFS CCR state is ok {0}	The stored GPFS CCR state is ok..	N/A	N/A
ccr_paxos_12_warn	STATE_CHANGE	WARNING	The stored GPFS CCR state is corrupt Item={0},ErrMsg={1}, Failed={2}	The stored GPFS CCR state is corrupt. See message for details.	One stored GPFS state file is missing or corrupt.	No user action necessary, GPFS will repair this automatically.
ccr_local_server_ok	STATE_CHANGE	INFO	The local GPFS CCR server is reachable {0}	The local GPFS CCR server is reachable.	N/A	N/A
ccr_local_server_warn	STATE_CHANGE	WARNING	The local GPFS CCR server is not reachable Item={0},ErrMsg={1}, Failed={2}	The local GPFS CCR server is not reachable. See message for details.	Either the local network and/or firewall is configured wrong or the local GPFS daemon does not respond.	Check the network and firewall configuration with regards to the used GPFS communication port (default: 1191). Restart GPFS on this node.
ccr_ip_lookup_ok	STATE_CHANGE	INFO	The IP address lookup for the GPFS CCR component is ok {0}	The IP address lookup for the GPFS CCR component is ok.	N/A	N/A
ccr_ip_lookup_warn	STATE_CHANGE	WARNING	The IP address lookup for the GPFS CCR component takes to long Item={0},ErrMsg={1}, Failed={2}	The IP address lookup for the GPFS CCR component takes to long, resulting in slow administration commands. See message for details.	Either the local network or the DNS is misconfigured.	Check the local network and DNS configuration.

Table 64. Events for the GPFS component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
ccr_quorum_nodes_fail	STATE_CHANGE	ERROR	A majority of the quorum nodes are not reachable over the management network Item={0},ErrMsg={1}, Failed={2}	A majority of the quorum nodes are not reachable over the management network. GPFS declares quorum loss. See message for details.	The quorum nodes cannot communicate with each other caused by a network and/or firewall misconfiguration.	Check the network and/or firmware (default port 1191 must not be blocked) configuration of the not reachable quorum nodes.
ccr_quorum_nodes_ok	STATE_CHANGE	INFO	All quorum nodes are reachable {0}	All quorum nodes are reachable.	N/A	N/A
ccr_quorum_nodes_warn	STATE_CHANGE	WARNING	Clustered Configuration Repository issue with Item={0},ErrMsg={1}, Failed={2}	At least on quorum node is not reachable. See message for details.	The quorum node is not reachable caused by a network and/or firewall misconfiguration.	Check the network and/or firmware (default port 1191 must not be blocked) configuration of the not reachable quorum node.
ccr_comm_dir_fail	STATE_CHANGE	ERROR	The files committed to the GPFS CCR are not complete or corrupt Item={0},ErrMsg={1}, Failed={2}	The files committed to the GPFS CCR are not complete or corrupt. See message for details.	The local disk might be full.	Check the local disk space and remove not necessary files. Recover this degraded node from a still intact node using the mmsdrrestore -p <NODE> command with <NODE> specifying intact node. See the mmsdrrestore man page for more details.
ccr_comm_dir_ok	STATE_CHANGE	INFO	The files committed to the GPFS CCR are complete and intact {0}	The files committed to the GPFS CCR are complete and intact..	N/A	N/A
ccr_comm_dir_warn	STATE_CHANGE	WARNING	The files committed to the GPFS CCR are not complete or corrupt Item={0},ErrMsg={1}, Failed={2}	The files committed to the GPFS CCR are not complete or corrupt. See message for details.	The local disk might be full.	Check the local disk space and remove not necessary files. Recover this degraded node from a still intact node using the mmsdrrestore -p <NODE> command with <NODE> specifying intact node. See the mmsdrrestore man page for more details.
ccr_tiebreaker_dsk_fail	STATE_CHANGE	ERROR	Access to tiebreaker disks failed Item={0},ErrMsg={1}, Failed={2}	Access to all tiebreaker disks failed. See message for details.	Corrupt disk.	Check, that the tiebreaker disks are available.



Table 64. Events for the GPFS component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
ccr_tiebreaker_dsk_ok	STATE_CHANGE	INFO	All tiebreaker disks used by the GPFS CCR are accessible {0}	All tiebreaker disks used by the GPFS CCR are accessible.	N/A	N/A
ccr_tiebreaker_dsk_warn	STATE_CHANGE	WARNING	At least one tiebreaker disk is not accessible Item={0},ErrMsg={1}, Failed={2}	At least one tiebreaker disk is not accessible. See message for details.	Corrupt disk.	Check, that the tiebreaker disks is accessible.
cluster_state_manager_reset	INFO	INFO	Clear memory of cluster state manager for this node	A reset request for the monitor state manager was received	A reset request for the monitor state manager was received.	N/A
nodeleave_info	INFO	INFO	A CES node left the cluster: Node {0}	Informational. Shows the node name of the cluster leaving node. ", + "This event might be logged on a different node, not necessarily on the leaving node	A CES node left the cluster. The name of the leaving node is provided.	NA
nodestatechange_info	INFO	INFO	Message: A CES node state change: Node {0} {1} {2} flag	Informational. Shows the modified node state, like the node turned to suspended mode, network down, or others.	"A node state change was detected. Details are shown in the message.	NA
quorumloss	INFO	WARNING	The cluster detected a quorum loss	The cluster may got in an inconsistent or split-brain state. ", + "Reasons could be network or hardware issues, or if quorum node(s) were removed from the cluster. ", + "The event may not be logged on the node which causes the quorum loss	The number of required quorum nodes does not match the minimum requirements. This can be an expected situation", + "user_action": "Make sure the required cluster quorum nodes are up and running.	Recover from the underlying issue. Make sure the cluster nodes are up and running.
gpfs_down	STATE_CHANGE	ERROR	The IBM Spectrum Scale service process not running on this node. Normal operation cannot be done	The IBM Spectrum Scale service is not running. This can be an expected state when the IBM Spectrum Scale service is shutdown.	The IBM Spectrum Scale service is not running.	Check the state of the IBM Spectrum Scale file system daemon, and check for the root cause in the /var/adm/ras/mmfs.log.latest log.

Table 64. Events for the GPFS component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
gpfs_up	STATE_CHANGE	INFO	The IBM Spectrum Scale service process is running	The IBM Spectrum Scale service is running	The IBM Spectrum Scale service is running	NA
gpfs_warn	INFO	WARNING	IBM Spectrum Scale process monitoring returned unknown result. This could be a temporary issue	Check of the IBM Spectrum Scale file system daemon returned an unknown result. This could be a temporary issue, like a timeout during the check procedure.	The IBM Spectrum Scale file system daemon state could not be determined due to a problem.	Find potential issues for this kind of failure in the <code>/var/adm/ras/mmsysmonitor.log</code> file.
info_on_duplicate_events	INFO	INFO	The event {0}{id} was repeated {1} times	Multiple messages of the same type were deduplicate to avoid log flooding.	Multiple events of the same type processed.	N/A
shared_root_bad	STATE_CHANGE	ERROR	Shared root is unavailable	The CES shared root file system is bad or not available. This file system is required to run the cluster because it stores the cluster-wide information. This problem will trigger a failover.	The CES framework detects the CES shared root file system to be unavailable on the node.	Check if the CES shared root file system and other expected IBM Spectrum Scale file systems are mounted properly.
shared_root_ok	STATE_CHANGE	INFO	Shared root is available	The CES shared root file system is available. This file system is required to run the cluster, because it stores clusterwide information.	The CES framework detects the CES shared root file system to be ok.	NA

Table 64. Events for the GPFS component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
quorum_down	STATE_CHANGE	ERROR	A quorum loss was detected	The monitor has detected a quorum loss. "Reasons could be network or hardware issues, or if quorum node(s) were removed from the cluster. ", "The event may not necessarily be logged on the node which causes the quorum loss"	The local node does not have quorum. It might be in an inconsistent or split-brain state	Check if the cluster quorum nodes are running and can be reached over the network. Check local firewall settings
quorum_up	STATE_CHANGE	INFO	quorum detected	The monitor detected a valid quorum.		N/A
quorum_warn	INFO	WARNING	The IBM Spectrum Scale quorum monitor could not be executed. This could be a timeout issue	Check of the quorum state returned an unknown result. This might be a temporary issue, like a timeout during the check procedure.	The quorum state could not be determined due to a problem.	Find potential issues for this kind of failure in the /var/adm/ras/mmsysmonitor.log file.
deadlock_detected	INFO	WARNING	The cluster detected a IBM Spectrum Scale file system deadlock	The cluster detected a deadlock in the IBM Spectrum Scale file system.	High file system activity might cause this issue.	The problem might be temporary or remain. Check the /var/adm/ras/mmfs.log.latest log files for more detailed information.
gpfsport_access_up	STATE_CHANGE	INFO	Access to IBM Spectrum Scale ip {0} port {1} ok	The TCP access check of the local IBM Spectrum Scale file system daemon port was successful.	IBM Spectrum Scale file system service access check successful.	N/A
gpfsport_down	STATE_CHANGE	ERROR	IBM Spectrum Scale port {0} is not active	The expected local IBM Spectrum Scale file system daemon port was not detected.	The IBM Spectrum Scale file system daemon is not running.	Check if the IBM Spectrum Scale service is running.
gpfsport_access_down	STATE_CHANGE	ERROR	No access to IBM Spectrum Scale ip {0} port {1}. Check firewall settings	The access check of the local IBM Spectrum Scale file system daemon port failed.	The port is probably blocked by a firewall rule	Check if the IBM Spectrum Scale file system daemon is running and check the firewall for blocking rules on this port.

Table 64. Events for the GPFS component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
gpfsport_up	STATE_CHANGE	INFO	IBM Spectrum Scale port {0} is active	The expected local IBM Spectrum Scale file system daemon port was detected.	The expected local IBM Spectrum Scale file system daemon port was detected.	N/A
gpfsport_warn	INFO	WARNING	IBM Spectrum Scale monitoring ip {0} port {1} returned unknown result	The check of the IBM Spectrum Scale file system daemon port returned an unknown result.	The IBM Spectrum Scale file system daemon port could not be determined due to a problem.	Find potential issues for this kind of failure in the logs.
gpfsport_access_warn	INFO	WARNING	IBM Spectrum Scale access check ip {0} port {1} failed. Check for valid IBM Spectrum Scale-IP	The access check of the IBM Spectrum Scale file system daemon port returned an unknown result.	The IBM Spectrum Scale file system daemon port access could not be determined due to a problem.	Find potential issues for this kind of failure in the logs.
longwaiters_found	STATE_CHANGE	ERROR	Detected IBM Spectrum Scale long-waiters.	Longwaiter threads found in the IBM Spectrum Scale file system.	High load might cause this issue.	Check log files. This could be also a temporary issue.
no_longwaiters_found	STATE_CHANGE	INFO	No IBM Spectrum Scale long-waiters	No longwaiter threads found in the IBM Spectrum Scale file system.	No longwaiter threads found in the IBM Spectrum Scale file system.	N/A
longwaiters_warn	INFO	WARNING	IBM Spectrum Scale long-waiters monitoring returned unknown result	The long waiters check returned an unknown result.	The IBM Spectrum Scale file system long waiters check could not be determined due to a problem.	Find potential issues for this kind of failure in the logs.
quorumreached_detected	INFO	INFO	Quorum reached event.	The cluster has reached quorum.	The cluster has reached quorum.	N/A

Table 65. Events for the GUI component

Event	Event Type	Severity	Message	Description	Cause	User Action
gui_down	STATE_CHANGE	ERROR	GUI service should be {0} and is {1}.	The GUI service is down.	The GUI service is not running on this node, although it has the node class GUI_MGMT_SERVER_NODE.	Restart the GUI service or change the node class for this node.
gui_up	STATE_CHANGE	INFO	GUI service as expected, state is {0}	The GUI service is running	The GUI service is running as expected.	NA

Table 65. Events for the GUI component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
gui_warn	INFO	INFO	GUI service returned unknow result.	The GUI service returned an unknown result.	The 'service' or 'systemctl' command returned unknown results about the gpfsogui service.	Check with the 'service' or 'systemctl' command if the gpfsogui is in the expected status. Is there no gpfsogui service although the node has the node class GUI_MGMT_SERVER_NODE check with the GUI documentation. Otherwise monitor, if this warning appears more often.

Table 66. Events for the HadoopConnector component

Event	Event type	Severity	Message	Description	Cause	User Action
hadoop_datanode_down	STATE_CHANGE	ERROR	Hadoop DataNode service is down	The Hadoop DataNode service is down	The Hadoop DataNode process is not running.	Start the Hadoop DataNode service.
hadoop_datanode_up	STATE_CHANGE	INFO	Hadoop DataNode service is up	The Hadoop ndatanode service is running	The Hadoop DataNode process is running.	NA
hadoop_datanode_warn	INFO	WARNING	Hadoop DataNode monitoring returned unknown results	The Hadoop DataNode service check returned unknown results.	The Hadoop DataNode service status check returned unknown results.	If this status persists after a few minutes, restart the DataNode service.
hadoop_namenode_down	STATE_CHANGE	ERROR	Hadoop NameNode service is down	The Hadoop NameNode service is down	The Hadoop NameNode process is not running.	Start the Hadoop NameNode service.
hadoop_namenode_up	STATE_CHANGE	INFO	Hadoop NameNode service is up	The Hadoop NameNode service is running.	The Hadoop NameNode process is running.	
hadoop_namenode_warn	INFO	WARNING	Hadoop NameNode monitoring returned unknown results	The Hadoop NameNode service status check returned unknown results.		If this status persists after a few minutes, restart the NameNode service.

Table 67. Events for the KEYSTONE component

Event	EventType	Severity	Message	Description	Cause	User action
ks_failed	STATE_CHANGE	ERROR	keystone (httpd) process should be {0} but is {1}	The keystone (httpd) process is in an unexpected mode.	If the object authentication is local, AD or LDAP the process failed unexpectedly. If the object authentication is none or user defined the process is expected to be stopped but it was not.	Make sure that the process is in the expected state.
ks_ok	STATE_CHANGE	INFO	keystone(httpd) process as expected, state is {0}	The keystone (httpd) process is in the expected state.	If the object authentication is local, AD or LDAP the process is running. If the object authentication is none or user defined the process stopped as expected.	NA
ks_restart	INFO	WARNING	The {0} service failed. Trying to recover			
ks_url_exfail	STATE_CHANGE	WARNING	Keystone request failed using {0}			
ks_url_failed	STATE_CHANGE	ERROR	Keystone request failed using {0}	A keystone URL request failed.	An HTTP request to keystone failed.	Check that httpd / keystone is running on the expected server and is accessible with the defined ports.
ks_url_ok	STATE_CHANGE	INFO	Keystone request successfully using {0}	A keystone URL request was successful.	A HTTP request to keystone returned successfully.	NA
ks_url_warn	INFO	WARNING	Keystone request on {0} returned unknown result	A keystone URL request returned an unknown result.	A simple HTTP request to keystone returned with an unexpected error.	Check that httpd / keystone is running on the expected server and is accessible with the defined ports.
ks_warn	INFO	WARNING	keystone(httpd) process monitoring returned unknown result	The keystone (httpd) monitoring returned an unknown result.	A status query for httpd returned an unexpected error.	Check service script and settings of httpd.
postgresql_failed	STATE_CHANGE	ERROR	postgresql-obj process should be {0} but is {1}	The postgresql-obj process is in an unexpected mode.	The database back-end for object authentication is supposed to run on a single node. Either the DB is not running on the designated node or it is running on a different node.	Check that postgresql-obj is running on the expected server.

Table 67. Events for the KEYSTONE component (continued)

Event	EventType	Severity	Message	Description	Cause	User action
postgresql_ok	STATE_CHANGE	INFO	postgresql-obj process as expected, state is {0}	The postgresql-obj process is in the expected mode.	The database back-end for object authentication is supposed to running on the right node while being stopped on others.	NA
postgresql_warn	INFO	WARNING	postgresql-obj process monitoring returned unknown result	The postgresql-obj process monitoring returned an unknown result.	A status query for postgresql-obj returned with an unexpected error.	Check postgres database engine.

Table 68. Events for the NFS component

Event	EventType	Severity	Message	Description	Cause	User Action
dbus_error	STATE_CHANGE	WARNING	DBus availability check failed	DBus availability check failed.	The DBus was detected as down. This might cause several issues on the local node.	Stop the NFS service, restart the DBus, and start the NFS service again.
disable_nfs_service	INFO	INFO	Ganesha NFS service was disabled	The NFS service was disabled on this node. Disabling a Service means, that also all configuration files are removed. This is different from stopping a running service.	The user has executed 'mmces service disable nfs'.	NA
enable_nfs_service	INFO	INFO	Ganesha NFS service was enabled	The NFS service was enabled on this node. Enabling a protocol service means, that also all required configuration files are automatically installed with the current valid configuration settings.	The user has executed 'mmces service enable nfs'	NA
ganeshaexit	INFO	INFO	Ganesha NFS was stopped	An NFS server instance has terminated.	A NFS instance was terminated or was killed somehow.	Restart the NFS service when the root cause for this issue is solved.
ganeshaGrace	INFO	INFO	Ganesha NFS is set to grace	The NFS server is sent to grace mode for a limited time. This gives previously connected clients time to recover their file locks.	The grace period is always cluster wide. NFS export configurations might have changed, and one or more NFS servers were restarted.	NA

Table 68. Events for the NFS component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
nfs3_down	INFO	WARNING	NFS v3 check returned down	The NFS v3 NULL check failed when expected. The NFS v3 NULL check is done to see if the NFS server reacts on NFS v3 requests. The v3 protocol must be enabled for this check. If this down state is detected, further checks are done to figure out if the NFS server is working. If the NFS server seems to be not working, then a failover is triggered. If NFS v3 and NFS v4 protocols are configured, then only the v3 NULL test is executed.	The NFS server might hang or is under high load so that the request might not be processed.	Check the health state of the NFS server and restart, if necessarily.
nfs3_up	INFO	INFO	NFS v4 check returned up	The NFS v4 NULL check was successful.	The NFS v4 NULL check works as expected.	NA
nfs4_down	INFO	WARNING	NFS v4 check returned down	The NFS v4 NULL check failed when expected. The NFS v4 NULL check is done to see whether the NFS server reacts on NFS v4 requests. The v4 protocol must be enabled for this check. If this down state is detected, further checks are done to figure out if the NFS server is working. If the NFS server seems to be not working, then a failover is triggered.	The NFS server might hang or is under high load so that the request might not be processed.	Check the health state of the NFS server and restart, if necessary.
nfs4_up	INFO	INFO	NFS v4 check returned up	The NFS v4 NULL check was successful.	The NFS v4 NULL check works as expected.	NA
nfs_active	STATE_CHANGE	INFO	NFS is now active	The NFS service must be up and running, and in a healthy state to provide the configured file exports.	The NFS server was detected as alive (again).	NA



Table 68. Events for the NFS component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
nfs_dbus_error	STATE_CHANGE	WARNING	NFS check via Dbus failed	The NFS service must be registered on Dbus to be fully working.	The NFS service is registered on Dbus, but there was a problem accessing it.	Check the health state of the NFS service, restart the NFS service. Check the log files for reported issues.
nfs_dbus_failed	STATE_CHANGE	WARNING	NFS check via Dbus did not return expected message	NFS service configuration settings (log configuration settings) are queried via Dbus. The result is checked for expected keywords.	The NFS service is registered on Dbus, but the check via Dbus did not return the expected result.	Stop the NFS service and start it again. Check the log configuration of the NFS service.
nfs_dbus_ok	STATE_CHANGE	INFO	NFS check via Dbus successful	Check that the NFS service is registered on Dbus and working.	The NFS service is registered on Dbus and working.	NA
nfs_in_grace	STATE_CHANGE	WARNING	NFS in grace mode	The monitor detected that Ganesha is in grace mode. During this time, the ganesha state is shown as degraded.	The NFS service was started or restarted.	NA
nfs_not_active	STATE_CHANGE	ERROR	NFS is not active	A check showed that a running Ganesha instance shows no activity at all.	Process might have hung.	Restart Ganesha.
nfs_not_dbus	STATE_CHANGE	WARNING	NFS service not available as Dbus service. Consider restart of NFS server.	The NFS service is currently not registered on Dbus. In this mode, the NFS service is not fully working. Exports cannot be added or removed, and not set in grace mode, which is important for data consistency.	The NFS service might have been started while the Dbus was down.	Stop the NFS service, restart the Dbus, and start the NFS service again.
nfsd_down	STATE_CHANGE	ERROR	NFSD process not running	Checks for an NFS service process.	The NFS server process was not detected.	Check the health state of the NFS server and restart, if necessary. The process might hang or is in failed state.
nfsd_up	STATE_CHANGE	INFO	NFSD process now running	Checks for a NFS service process.	The NFS server process was detected. Some further checks are done then.	NA
nfsd_warn	INFO	WARNING	NFSD process monitoring returned unknown result	Checks for a NFS service process.	The NFS server process state might not be determined due to a problem.	Check the health state of the NFS server and restart, if necessary.

Table 68. Events for the NFS component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
portmapper_down	STATE_CHANGE	ERROR	Portmapper port 111 is not active	The portmapper is needed to provide the NFS services to clients.	The portmapper is not running on port 111.	NA
portmapper_up	STATE_CHANGE	INFO	Portmapper port is now active	The portmapper is needed to provide the NFS services to clients.	The portmapper is running on port 111.	NA
portmapper_warn	INFO	WARNING	Portmapper port monitoring (111) returned unknown result	The portmapper is needed to provide the NFS services to clients.	The portmapper status might not be determined due to a problem.	Restart the portmapper, if necessary.
postIpChange_info	INFO	INFO	IP addresses modified (post change)	Information that IP addresses are moved around the cluster nodes.	CES IP addresses were moved or added to the node, and activated.	NA
rquotad_down	INFO	INFO	rpc.rquotad not running	Currently not in use. Future.	NA	NA
rquotad_up	INFO	INFO	rpc.rquotad is running	Currently not in use. Future.		
start_nfs_service	INFO	INFO	Ganesha NFS service was started	Information about a NFS service start.	The NFS service was started (like 'mmces service start nfs').	NA
statd_down	STATE_CHANGE	ERROR	rpc.statd is not running	The statd process is used by NFS v3 to handle file locks.	The statd process is not running.	Stop and start the NFS service. This attempts to start the statd process also.
statd_up	STATE_CHANGE	INFO	rpc.statd is running	The statd process is used by NFS v3 to handle file locks.	The statd process is running.	NA
stop_nfs_service	INFO	INFO	Ganesha NFS service was stopped	Information about a NFS service stop.	The NFS service was stopped (like 'mmces service stop nfs').	NA

Table 69. Events for the Network component

Event	EventType	Severity	Message	Description	Cause	User Action
bond_degraded	STATE_CHANGE	INFO	Some slaves of the bond {0} went down	Some of the bond parts are malfunctioning	Some slaves of the bond are not functioning properly.	Check the bonding configuration, the network configuration and cabling of the malfunctioning slaves of the bond.
bond_down	STATE_CHANGE	ERROR	All slaves of the bond {0} are down	All slaves of a network bond are down.	All slaves of this network bond went down.	Check the bonding configuration, the network configuration and cabling of all slaves of the bond.

Table 69. Events for the Network component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
bond_up	STATE_CHANGE	INFO	All slaves of the bond {0} are working as expected	This bond is functioning properly.	All slaves of this network bond are functioning properly.	NA
ces_disable_node network	INFO	INFO	Network was disabled	Clean up after a 'mmchnode --ces-disable' command. The network configuration is modified accordingly.	Clean up after a 'mmchnode --ces-disable' command.	NA
ces_enable_node network	INFO	INFO	Network was disabled	Called to handle any network- sepcific issues involved after a 'mmchnode --ces-enable' command. The network configuration is modified accordingly.	Called after a 'mmchnode --ces-enable' command.	NA
ces_startup_network	INFO	INFO	CES network service was started	Information that the CES network has started.	CES network IPs are started.	NA
handle_network _problem_info	INFO	INFO	Handle network problem - Problem: {0}, Argument: {1}	Information about network- related reconfigurations. This can be enable or disable IPs, assign or unassign IPs for example.	A change in the network configuration. Details are part of the information message.	NA
many_tx_errors	STATE_CHANGE	ERROR	NIC {0} had many TX errors since the last monitoring cycle	The network adapter had many TX errors since the last monitoring cycle.	/proc/net/dev lists much more TX errors for this adapter since the last monitoring cycle.	Check the network cabling and network infrastructure
move_cesip_from	INFO	INFO	Address {0} was moved from this node to node {1}	Information that a CES IP address was moved from the current node to another node.	Rebalancing of CES IP addresses.	NA
move_cesip_to	INFO	INFO	Address {0} was moved from node {1} to this node	Information that a CES IP address was moved from another node to the current node.	Rebalancing of CES IP addresses.	NA
move_cesips_infos	INFO	INFO	A move request for ip addresses was executed	CES IP addresses can be moved in case of node failovers from one node to one or more other nodes. This message is logged on a node observing this, not necessarily on any affected node.	A CES IP movement was detected.	NA
network_connectivity_down	STATE_CHANGE	ERROR	NIC {0} can not connect to the gateway	This network adapter can not connect to the gateway.	The gateway does not respond to the sent connections-checking packets.	Check the network configuration of the network adapter, path to the gateway and gateway itself.

Table 69. Events for the Network component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
network_connectivity_up	STATE_CHANGE	INFO	NIC {0} can connect to the gateway	This network adapter can connect to the gateway.	The gateway responds to the sent connections-checking packets.	NA
network_down	STATE_CHANGE	ERROR	Network is down	This network adapter is down.	This network adapter was disabled.	Enable this network adapter.
network_found	INFO	INFO	NIC {0} was found	A new network adapter was found.	A new NIC, which is relevant for the IBM Spectrum Scale monitoring, is listed by 'ip a'.	NA
network_ips_down	STATE_CHANGE	ERROR	No relevant NICs detected	No relevant network adapters detected	No network adapters have the IBM Spectrum Scale-relevant IPs.	Find out, why the IBM Spectrum Scale-relevant IPs were not assigned to any NICs. "
network_ips_up	STATE_CHANGE	INFO	Relevant IPs are served by found NICs	Relevant IPs are served by network adapters	At least one IBM Spectrum Scale-relevant IP is assigned to a network adapter.	NA
network_link_down	STATE_CHANGE	ERROR	Physical link of the NIC {0} is down	The physical link of this adapter is down.	The flag LOWER_UP is not set for this NIC in the output of 'ip a'.	Check the cabling of this network adapter.
network_link_up	STATE_CHANGE	INFO	Physical link of the NIC {0} is up	The physical link of this adapter is up.	The flag LOWER_UP is set for this NIC in the output of 'ip a'	NA
network_up	STATE_CHANGE	INFO	Network is up	This network adapter is up.	This network adapter is enabled.	NA
network_vanished	INFO	INFO	NIC {0} has vanished	One of network adapters can not be detected anymore.	One of the previously monitored NICs is not listed by 'ip a' anymore.	NA
no_tx_errors	STATE_CHANGE	INFO	NIC {0} had no or a tiny number of TX errors.	The NIC had no or an insignificant number of TX errors.	/proc/net/dev lists no or an insignificant number of TX errors for this adapter since the last monitoring cycle.	Check the network cabling and network infrastructure.

Table 70. Events for the object component

Event	EventType	Severity	Message	Description	Cause	User Action
account-auditor_failed	STATE_CHANGE	ERROR	account-auditor process should be {0} but is {1}	The account-auditor process is not in the expected state.	The account-auditor process is expected to be running on the singleton node only.	Check the status of openstack-swift-account-auditor process and object singleton flag.

Table 70. Events for the object component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
account-auditor_ok	STATE_CHANGE	INFO	account-auditor process as expected, state is {0}	The account-auditor process is in the expected state.	The account-auditor process is expected to be running on the singleton node only.	NA
account-auditor_warn	INFO	WARNING	account-auditor process monitoring returned unknown result	The account-auditor check returned an unknown result.	A status query for openstack-swift-account-auditor returned with an unexpected error.	Check service script and settings.
account-reaper_failed	STATE_CHANGE	ERROR	account-reaper process should be {0} but is {1}	The account-reaper process is not running.	The account-reaper process is not running.	Check the status of openstack-swift-account-reaper process.
account-reaper_ok	STATE_CHANGE	INFO	account-reaper process as expected, state is {0}	The account-reaper process is running.	The account-reaper process is running.	NA
account-reaper_warn	INFO	WARNING	account-reaper process monitoring returned unknown result	The account-reaper check returned an unknown result.	A status query for openstack-swift-account-reaper returned with an unexpected error.	Check service script and settings.
account-replicator_failed	STATE_CHANGE	ERROR	account-replicator process should be {0} but is {1}	The account-replicator process is not running.	The account-replicator process is not running.	Check the status of openstack-swift-account-replicator process.
account-replicator_ok	STATE_CHANGE	INFO	account-replicator process as expected, state is {0}	The account-replicator process is running.	The account-replicator process is running.	NA
account-replicator_warn	INFO	WARNING	account-replicator process monitoring returned unknown result	The account-replicator check returned an unknown result.	A status query for openstack-swift-account-replicator returned with an unexpected error.	Check the service script and settings.
account-server_failed	STATE_CHANGE	ERROR	account process should be {0} but is {1}	The account-server process is not running.	The account-server process is not running.	Check the status of openstack-swift-account process.
account-server_ok	STATE_CHANGE	INFO	account process as expected, state is {0}	The account-server process is running.	The account-server process is running.	NA
account-server_warn	INFO	WARNING	account process monitoring returned unknown result	The account-server check returned an unknown result.	A status query for openstack-swift-account returned with an unexpected error.	Check the service script and settings.
container-auditor_failed	STATE_CHANGE	ERROR	container-auditor process should be {0} but is {1}	The container-auditor process is not in the expected state.	The container-auditor process is expected to be running on the singleton node only.	Check the status of openstack-swift-container-auditor process and object singleton flag.
container-auditor_ok	STATE_CHANGE	INFO	container-auditor process as expected, state is {0}	The container-auditor process is in the expected state.	The container-auditor process is expected to be running on the singleton node only.	NA
container-auditor_warn	INFO	WARNING	container-auditor process monitoring returned unknown result	The container-auditor check returned an unknown result.	A status query for openstack-swift-container-auditor returned with an unexpected error.	Check service script and settings.

Table 70. Events for the object component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
container-replicator_failed	STATE_CHANGE	ERROR	container-replicator process should be {0} but is {1}	The container-replicator process is not running.	The container-replicator process is not running.	Check the status of openstack-swift-container-replicator process.
container-replicator_ok	STATE_CHANGE	INFO	container-replicator process as expected, state is {0}	The container-replicator process is running.	The container-replicator process is running.	NA
container-replicator_warn	INFO	WARNING	container-replicator process monitoring returned unknown result	The container-replicator check returned an unknown result.	A status query for openstack-swift-container-replicator returned with an unexpected error.	Check service script and settings.
container-server_failed	STATE_CHANGE	ERROR	container process should be {0} but is {1}	The container-server process is not running.	The container-server process is not running.	Check the status of openstack-swift-container process.
container-server_ok	STATE_CHANGE	INFO	container process as expected, state is {0}	The container-server process is running.	The container-server process is running.	NA
container-server_warn	INFO	WARNING	container process monitoring returned unknown result	The container-server check returned an unknown result.	A status query for openstack-swift-container returned with an unexpected error.	Check the service script and settings.
container-updater_failed	STATE_CHANGE	ERROR	container-updater process should be {0} but is {1}	The container-updater process is not in the expected state.	The container-updater process is expected to be running on the singleton node only.	Check the status of openstack-swift-container-updater process and object singleton flag.
container-updater_ok	STATE_CHANGE	INFO	container-updater process as expected, state is {0}	The container-updater process is in the expected state.	The container-updater process is expected to be running on the singleton node only.	NA
container-updater_warn	INFO	WARNING	container-updater process monitoring returned unknown result	The container-updater check returned an unknown result.	A status query for openstack-swift-container-updater returned with an unexpected error.	Check the service script and settings.
disable_Address_database_node	INFO	INFO	Disable Address Database Node	Event to signal that the database flag was removed from this node.	A CES IP with a singleton/database flag linked to it was removed/moved from/to this node.	NA
disable_Address_singleton_node	INFO	INFO	Disable Address Singleton Node	Event to signal that the singleton flag was removed from this.	A CES IP with a singleton/database flag linked to it was removed/moved from/to this node.	NA
enable_Address_database_node	INFO	INFO	Enable Address Database Node	Event to signal that the database flag was moved to this node.	A CES IP with a singleton/database flag linked to it was removed/moved from/to this node.	NA
enable_Address_singleton_node	INFO	INFO	Enable Address Singleton Node	Event to signal that the singleton flag was moved to this node.	A CES IP with a singleton/database flag linked to it was removed/moved from/to this node.	NA

Table 70. Events for the object component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
ibmobjectizer_failed	STATE_CHANGE	ERROR	ibmobjectizer process should be {0} but is {1}	The ibmobjectizer process is in the expected state.	The ibmobjectizer process is expected to be running on the singleton node only.	Check the status of the ibmobjectizer process and object singleton flag.
ibmobjectizer_ok	STATE_CHANGE	INFO	ibmobjectizer process as expected, state is {0}	The ibmobjectizer process is not in the expected state.	The ibmobjectizer process is expected to be running on the singleton node only.	NA
ibmobjectizer_warn	INFO	WARNING	ibmobjectizer process monitoring returned unknown result	The ibmobjectizer check returned an unknown result.	A status query for ibmobjectizer returned with an unexpected error.	Check the service script and settings.
memcached_failed	STATE_CHANGE	ERROR	memcached process should be {0} but is {1}	The memcached process is not running.	The memcached process is not running.	Check the status of memcached process.
memcached_ok	STATE_CHANGE	INFO	memcached process as expected, state is {0}	The memcached process is running.	The memcached process is running.	NA
memcached_warn	INFO	WARNING	memcached process monitoring returned unknown result	The memcached check returned an unknown result.	A status query for memcached returned with an unexpected error.	Check the service script and settings.
obj_restart	INFO	WARNING	The {0} service failed. Trying to recover			
object-expirer_failed	STATE_CHANGE	ERROR	object-expirer process should be {0} but is {1}	The object-expirer process is not in the expected state.	The object-expirer process is expected to be running on the singleton node only.	Check the status of openstack -swift-object-expirer process and object singleton flag.
object-expirer_ok	STATE_CHANGE	INFO	object-expirer process as expected, state is {0}	The object-expirer process is in the expected state.	The object-expirer process is expected to be running on the singleton node only.	NA
object-expirer_warn	INFO	WARNING	object-expirer process monitoring returned unknown result	The object-expirer check returned an unknown result.	A status query for openstack -swift-object-expirer returned with an unexpected error.	Check the service script and settings.
object-replicator_failed	STATE_CHANGE	ERROR	object-replicator process should be {0} but is {1}	The object-replicator process is not running.	The object-replicator process is not running.	Check the status of openstack -swift-object-replicator process.
object-replicator_ok	STATE_CHANGE	INFO	object-replicator process as expected, state is {0}	The object-replicator process is running.	The object-replicator process is running.	NA
object-replicator_warn	INFO	WARNING	object-replicator process monitoring returned unknown result	The object-replicator check returned an unknown result.	A status query for openstack -swift-object-replicator returned with an unexpected error.	Check the service script and settings.
object-server_failed	STATE_CHANGE	ERROR	object process should be {0} but is {1}	The object-server process is not running.	The object-server process is not running.	Check the status of the openstack -swift-object process.
object-server_ok	STATE_CHANGE	INFO	object process as expected, state is {0}	The object-server process is running.	The object-server process is running.	NA

Table 70. Events for the object component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
object-server_warn	INFO	WARNING	object process monitoring returned unknown result	The object-server check returned an unknown result.	A status query for openstack -swift-object-server returned with an unexpected error.	Check the service script and settings.
object-updater_failed	STATE_CHANGE	ERROR	object-updater process should be {0} but is {1}	The object-updater process is not in the expected state.	The object-updater process is expected to be running on the singleton node only.	Check the status of the openstack -swift-object-updater process and object singleton flag.
object-updater_ok	STATE_CHANGE	INFO	object-updater process as expected, state is {0}	The object-updater process is in the expected state.	The object-updater process is expected to be running on the singleton node only.	NA
object-updater_warn	INFO	WARNING	object-updater process monitoring returned unknown result	The object-updater check returned an unknown result	A status query for openstack -swift-object-updater returned with an unexpected error.	Check the service script and settings.
openstack-object-sof_failed	STATE_CHANGE	ERROR	object-sof process should be {0} but is {1}	The swift-on-file process is not in the expected state.	The swift-on-file process is expected to be running then the capability is enabled and stopped when disabled.	Check the status of the openstack -swift-object-sof process and capabilities flag in spectrum-scale-object.conf.
openstack-object-sof_ok	STATE_CHANGE	INFO	object-sof process as expected, state is {0}	The swift-on-file process is in the expected state.	The swift-on-file process is expected to be running then the capability is enabled and stopped when disabled.	NA
openstack-object-sof_warn	INFO	INFO	object-sof process monitoring returned unknown result	The openstack -swift-object-sof check returned an unknown result.	A status query for openstack -swift-object-sof returned with an unexpected error.	Check the service script and settings.
postIpChange_info	INFO	INFO	IP addresses modified {0}	CES IP addresses have been moved and activated.		NA
proxy-server_failed	STATE_CHANGE	ERROR	proxy process should be {0} but is {1}	The proxy-server process is not running.	The proxy-server process is not running.	Check the status of the openstack -swift-proxy process.
proxy-server_ok	STATE_CHANGE	INFO	proxy process as expected, state is {0}	The proxy-server process is running.	The proxy-server process is running.	NA
proxy-server_warn	INFO	WARNING	proxy process monitoring returned unknown result	The proxy-server process monitoring returned an unknown result.	A status query for openstack-swift-proxy-server returned with an unexpected error.	Check the service script and settings.
ring_checksum_failed	STATE_CHANGE	ERROR	Checksum of ring file {0} does not match the one in CCR	Files for object rings have been modified unexpectedly.	Checksum of file did not match the stored value.	Check the ring files.
ring_checksum_ok	STATE_CHANGE	INFO	Checksum of ring file {0} is OK	Files for object rings were successfully checked.	Checksum of file found unchanged.	NA



Table 70. Events for the object component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
ring_checksum_warn	INFO	WARNING	Issue while checking checksum of ring file {0}	Checksum generation process failed.	The ring_checksum check returned an unknown result.	Check the ring files and the md5sum executable.

Table 71. Events for the Performance component

Event	EventType	Severity	Message	Description	Cause	User Action
pmcollector_down	STATE_CHANGE	ERROR	pmcollector process should be {0} and is {1}	The performance monitor collector is down.	This node is in the global config of perfmon (mmpfmon config show) but the pmcollector service is not running.	Start the service with the 'service' or 'systemctl' command or remove the node from the global perfmon config with the mmpfmon command.
pmsensors_down	STATE_CHANGE	ERROR	pmsensors process should be {0} and is {1}	The performance monitor sensors are down	The node has the perfmon designation in mmlscluster but the pmsensors service is not running.	Start the service with the 'service' or 'systemctl' command or remove the perfmon designation with mmchnode.
pmsensors_up	STATE_CHANGE	INFO	pmsensors process as expected, state is {0}	The performance monitor sensors are running.		NA
pmcollector_up	STATE_CHANGE	INFO	pmcollector process as expected, state is {0}	The performance monitor collector is running.	The perfmon collector service is running as expected.	NA
pmcollector_warn	INFO	INFO	pmcollector process returned unknown result	The monitoring service for performance monitor collector returned an unknown result.	The 'service' or 'systemctl' command returned unknown results about the pmcollector service.	Check with the 'service' or 'systemctl' command if the pmcollector is in the expected status. If there is no pmcollector service, although the node is in the global mmpfmon config, check with the perfmon documentation.

Table 71. Events for the Performance component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
pmsensors_warn	INFO	INFO	pmsensors process returned unknown result	The monitoring service for performance monitor sensors returned an unknown result.	The 'service' or 'systemctl' command returned unknown results about the pmsensors service.	Check with the 'service' or 'systemctl' command if the pmsensors is in the expected status. If there is no pmsensors service although the node has the perfmom designation in mmlscluster, check with the perfmom documentation.

Table 72. Events for the SMB component

Event	EventType	Severity	Message	Description	Cause	User Action
ctdb_down	STATE_CHANGE	ERROR	CTDB process not running	The CTDB process is not running.		Perform trouble check.
ctdb_recovered	STATE_CHANGE	INFO	STATE_CHANGE	CTDB completed database recovery.		NA
ctdb_recovery	STATE_CHANGE	WARNING	CTDB Recovery detected	CTDB is performing a database recovery.		NA
ctdb_state_down	STATE_CHANGE	ERROR	CTDB state is {0}	The CTDB state is unhealthy.		Perform trouble check.
ctdb_state_up	STATE_CHANGE	INFO	CTDB state is healthy	The CTDB state is healthy.		NA
ctdb_up	STATE_CHANGE	INFO	CTDB process now running	The CTDB process is running.		NA
ctdb_warn	INFO	WARNING	CTDB monitoring returned unknown result	The CTDB check returned unknown result.		Perform trouble check.
smb_restart	INFO	WARNING	The SMB service failed. Trying to recover	Attempt to start the SMBD process.	The SMBD process was not running.	NA
smbd_down	STATE_CHANGE	ERROR	SMBD process not running	The SMBD process is not running.		Perform trouble check.
smbd_up	STATE_CHANGE	INFO	SMBD process now running	The SMBD process is running.		NA
smbd_warn	INFO	WARNING	SMBD process monitoring returned unknown result	The SMBD process monitoring returned an unknown result.		Perform trouble check.
smbport_down	STATE_CHANGE	ERROR	SMB port {0} is not active	SMBD is not listening on a TCP protocol port.		Perform trouble check.

Table 72. Events for the SMB component (continued)

Event	EventType	Severity	Message	Description	Cause	User Action
smbport_up	STATE_CHANGE	INFO	SMB port {0} is now active	An SMB port was activated.		NA
smbport_warn	INFO	WARNING	SMB port monitoring {0} returned unknown result	An internal error occurred while monitoring SMB TCP protocol.		Perform trouble check.

Table 73. GUI events

Event	Event type	Severity	Message	Description	System action	User action
MM0001	INFO	WARNING	The rebuild of declustered array "{0}" in recovery group "{1}" failed.	After a disk failure the declustered array attempted a rebuild which failed. Most probably the declustered array run out of spare space.	Operation continues with reduced redundancy.	Replace all physical disks in "replaceable" state to provide more spare space.
MM0008	INFO	INFO	The file system "{0}" was mounted on node "{1}"	A file system was mounted on a node.		
MM0009	INFO	INFO	The file system "{0}" was unmounted on node "{1}"	A file system was unmounted on a node.		
MM0012	INFO	WARNING	Checksum mismatch (client: "{0}", role: "{1}", server "{2}", NSD: "{3}", reason: "{4}", start sector: {5}, length: {6}, error count client/server/NSD: {7}/{8}/{9})	There was a mismatch between the calculated checksum of a data block and the expected one. This might be caused by a transmission error between the two involved nodes, a CPU or memory error on one of the nodes or a read error on the involved disk. The block will be reread so no data is corrupted. Frequent occurrence of checksum errors hint at hidden issues somewhere in the cluster and should be analyzed.	Retransmit bad block	
MM0013	INFO	ERROR	Failure opening recovery group "{0}" on node "{1}" because of {2} / {3}	A node could not open a recovery group. Either the connection to the corresponding enclosure is broken or the enclosure is out of service.		Fix the connection issue or bring enclosure back online.
MM0014	INFO	ERROR	Recovery group "{0}" panicked on node "{1}" because of {2} / {3}	A recovery group panicked.		
MM0015	INFO	INFO	Recovery group "{0}" relinquished by node "{1}" because of {2} / {3}	A recovery group was relinquished by a node. Another node will takeover this recovery group.		
MM0016	INFO	INFO	Recovery group "{0}" taken over by node "{1}" because of {2} / {3}	A recovery group was taken over by a node.		

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MM0020	INFO	WARNING	Path "{2}" to physical disk "{0}" of recovery group "{3}" failed on node "{1}".	A path to a physical disk failed. Since each disk is connected through multiple paths, the disk is still accessible after a single path failure.	Fall back to another path	Check cabling between node and disk enclosure
MM0025	INFO	INFO	This is a SNMP test message.	Test message.	None	NA
MM1001	INFO	WARNING	Operations for rule {0} are still running at the start of the next management of rule {0}.	This is a snapshot notification event	Processing continues	No action is required
MM1002	INFO	WARNING	The number of pending operations exceeds {0} operations for rule {1}.	This is a snapshot notification event	Processing continues	No action is required
MM1003	INFO	WARNING	The total number of pending operations exceeds {0} operations	This is a snapshot notification	Processing continues	n/a
MM1004	INFO	WARNING	A snapshot operation exceeds {0} minutes for rule {1} on file system {2}, file set {3}.	This is a snapshot notification event	Processing continues	No action is required
MM1005	INFO	WARNING	A snapshot creation invoked by rule {0} failed on file system {1}, file set {2}.	This is a snapshot notification event	Processing continues	Retry the snapshot creation
MM1006	INFO	WARNING	A snapshot deletion invoked by rule {0} failed on file system {1}, file set {2}.	This is a snapshot notification event	Processing continues	Retry the snapshot delete action
MM1007	INFO	WARNING	A snapshot operation exceeds {0} minutes for rule {1} on file system {2}.	The snapshot operation resulting from the rule is exceeding the established time limit.	Processing continues	Reduce the number of snapshot operations
MM1008	INFO	ERROR	A snapshot creation invoked by rule {0} failed on file system {1}.	The snapshot was not created according to the specified rule.	Processing continues	Try to create the snapshot again manually
MM1009	INFO	WARNING	A snapshot deletion invoked by rule {0} failed on file system {1}.	The snapshot was not deleted as according to the specified rule.	Processing continues	Try to manually delete the snapshot
MM1010	INFO	INFO	The {0} node is elected as new cluster manager.	A new cluster manager node has been elected. This happens when a cluster first starts up or when the current cluster manager fails or resigns and a new node takes over as cluster manager.		NA
MM1011	INFO	WARNING	The fileset quota has been exceeded.	The event is triggered when the file system manager detects that a fileset quota has been exceeded.		n/a
MS0001	STATE_CHANGE	INFO	The state of recovery group "{0}" is unknown.	The state of a recovery group could not be determined.		
MS0003	STATE_CHANGE	INFO	Dummy message for Storage Server	Dummy message for Storage Server		

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0011	STATE_CHANGE	INFO	The state of declustered array "{0}" in recovery group "{2}" is unknown.	The state of a declustered array could not be determined.		
MS0012	STATE_CHANGE	INFO	Declustered array "{0}" in recovery group "{2}" is ok.	The state of the declustered array changed to "ok" which means it works well.		
MS0013	STATE_CHANGE	ERROR	Declustered array "{0}" in recovery group "{2}" needs service.	The state of the declustered array changed to "needs service", which means that some repair action needs to be performed.		
MS0021	STATE_CHANGE	INFO	The state of virtual disk "{0}" is unknown.	The state of a virtual disk could not be determined.		
MS0022	STATE_CHANGE	INFO	The virtual disk "{0}" is ok.	The state of a virtual disk changed to "ok", which means it works well.		
MS0023	STATE_CHANGE	INFO	The virtual disk "{0}" is degraded ({1}).	The state of a virtual disk changed to "degraded". This means that (probably caused by a PDisk failure) the data redundancy is reduced.	The expected redundancy level is rebuild with low priority.	
MS0024	STATE_CHANGE	WARNING	The virtual disk "{0}" is critical degraded ({1}).	The state of a virtual disk changed to "critical degraded". This could be because of multiple pdisk failures. This error means that there is no data redundancy left. The system rebuilds redundancy with highest priority which might cause performance degradation.	The redundancy is rebuild with highest priority.	
MS0025	STATE_CHANGE	ERROR	The virtual disk "{0}" is offline.	The state of a virtual disk changed to "offline". This means that access to the data is lost probably caused by a simultaneous failure of multiple physical disks. For example, whole enclosure went offline.		Revitalize some of the physical disks.
MS0031	STATE_CHANGE	INFO	The state of physical disk "{0}" in recovery group "{2}" is unknown.	The state of a physical disk could not be determined.		
MS0032	STATE_CHANGE	INFO	Physical disk "{0}" in recovery group "{2}" is draining.	The state of a physical disk changed to "draining". This means all data and metadata is migrated away from this disk. This is caused either by administrative action or because GPFS experienced problems with this disk.	Data and metadata gets drained from the disk.	

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0033	STATE_CHANGE	ERROR	Physical disk "{0}" in recovery group "{2}" should be replaced.	The state of a physical disk changed to "replaceable". This means the disk should be replaced soon. Disks are marked as replaceable when the replace threshold of the declustered array is reached.		Replace the disk.
MS0034	STATE_CHANGE	INFO	Physical disk "{0}" in recovery group "{2}" operates normal.	The state of a physical disk changed to "normal". This means the disk operates inside normal parameters.		
MS0041	STATE_CHANGE	INFO	The state of NSD "{0}" is unknown.	The state of an NSD could not be determined.		
MS0042	STATE_CHANGE	INFO	Nsd "{0}" is up.	The availability of an NSD changed to "up".		
MS0043	STATE_CHANGE	ERROR	Nsd "{0}" is down.	The availability of an NSD changed to "down".		Recover the NSD.
MS0044	STATE_CHANGE	INFO	Nsd "{0}" is recovering.	The availability of an NSD changed to "recovering".	Recovering of NSD.	
MS0045	STATE_CHANGE	ERROR	Nsd "{0}" is unrecovered.	The availability of an NSD changed to "unrecovered" because the attempt to recover the NSD failed.		Fix physical disk issues and retry recovery.
MS0051	STATE_CHANGE	INFO	The state of node "{0}" is unknown.	The state of an NSD could not be determined.		
MS0052	STATE_CHANGE	INFO	The GPFS daemon node "{0}" is active.	The GPFS daemon is up.		
MS0053	STATE_CHANGE	INFO	The GPFS daemon node "{0}" is arbitrating.	The GPFS daemon is starting up or waiting for the quorum to establish.		
MS0054	STATE_CHANGE	ERROR	The GPFS daemon node "{0}" is down.	The GPFS daemon is down. GPFS is not operational on this node.		Restart GPFS daemon
MS0061	STATE_CHANGE	INFO	The state of cluster "{0}" is unknown.	The state of the cluster could not be determined.		
MS0062	STATE_CHANGE	INFO	The cluster "{0}" is up. Quorum {1}.	The cluster is up and running.		
MS0063	STATE_CHANGE	ERROR	The cluster "{0}" is down. Quorum {1}.	The cluster is down. Either quorum was lost or it was administratively shut down.		Restart individual nodes at least until quorum is reached.
MS0071	STATE_CHANGE	INFO	The state of disk enclosure "{0}" is unknown.	The state of the disk enclosure could not be determined.		
MS0072	STATE_CHANGE	INFO	The disk enclosure "{0}" is ok.	The disk enclosure reports no issues.		
MS0073	STATE_CHANGE	ERROR	The disk enclosure "{0}" needs service.	The disk enclosure reports that it needs service.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0074	STATE_CHANGE	ERROR	The disk enclosure "{0}" is missing.	The disk enclosure is not visible by any node in the cluster.		Follow the maintenance procedure for checking the attachment of the enclosure.
MS0081	STATE_CHANGE	INFO	The state of environmental service module "{0}" is unknown.	The state of the environmental service module could not be determined.		
MS0082	STATE_CHANGE	INFO	The environmental service module "{0}" is ok.	The environmental service module reports no issues.		
MS0083	STATE_CHANGE	ERROR	The environmental service module "{0}" is in error.	The environmental service module reports an error.		Follow the maintenance procedure for the environmental service module.
MS0084	STATE_CHANGE	ERROR	The environmental service module "{0}" is missing.	The environmental service module is not visible by any node in the cluster.		Follow the maintenance procedure for checking the attachment of the environmental service module.
MS0091	STATE_CHANGE	INFO	The temperature of environmental service module "{0}" is unknown.	The temperature of the environmental service module could not be determined.		
MS0092	STATE_CHANGE	INFO	The temperature of environmental service module "{0}" is ok ({1}).	The environmental service module reports no temperature issue.		
MS0093	STATE_CHANGE	ERROR	The temperature of environmental service module "{0}" is too high ({1}).	The environmental service module reports too high temperature.		Follow the maintenance procedure for the environmental service module.
MS0094	STATE_CHANGE	WARNING	Communication to temperature sensor of environmental service module "{0}" lost.	The I2C bus used for reading the temperature sensor failed. Thus the temperature cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0095	STATE_CHANGE	ERROR	Temperature of environmental service module "{0}" exceeded critical threshold at {1}.	If the critical threshold is exceeded, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0096	STATE_CHANGE	WARNING	Temperature of environmental service module "{0}" exceeded warning threshold at {1}.	If the warning threshold is exceeded, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0097	STATE_CHANGE	ERROR	Temperature of environmental service module "{0}" undercut the lower critical threshold at {1}.	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0098	STATE_CHANGE	WARNING	Temperature of environmental service module "{0}" undercut the lower warning threshold at {1}.	If the lower warning threshold is undercut, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0101	STATE_CHANGE	INFO	The state of drawer control module "{0}" is unknown.	The state of the drawer control module could not be determined.		
MS0102	STATE_CHANGE	INFO	The drawer control module "{0}" is ok.	The drawer control module reports no issues.		
MS0103	STATE_CHANGE	ERROR	The drawer control module "{0}" is in error.	The drawer control module reports an error.		Follow the maintenance procedure for the drawer control module.
MS0104	STATE_CHANGE	ERROR	The drawer control module "{0}" is missing.	The drawer control module is not visible by any node in the cluster.		Follow the maintenance procedure for checking the attachment of the drawer control module.
MS0105	STATE_CHANGE	WARNING	A drawer of drawer control module "{0}" is open.	A drawer of the drawer control module is open. Drawers should be closed immediately after replacing disks.		Close the drawer.
MS0111	STATE_CHANGE	INFO	The state of fan "{0}" is unknown.	The state of the fan could not be determined.		
MS0112	STATE_CHANGE	INFO	The fan "{0}" is ok ({1}).	The fan reports no issues.		
MS0113	STATE_CHANGE	ERROR	The fan "{0}" is in error ({1}).	The fan reports an error.		Follow the maintenance procedure for the fan.
MS0114	STATE_CHANGE	ERROR	The fan "{0}" is missing.	The fan is not visible by any node in the cluster.		Follow the maintenance procedure for checking the attachment of the fan.
MS0120	STATE_CHANGE	INFO	The state of power supply "{0}" is unknown.	The state of the power supply could not be determined.		
MS0121	STATE_CHANGE	INFO	The state of power supply "{0}" is unknown.	The state of the power supply could not be determined.		
MS0122	STATE_CHANGE	INFO	The power supply "{0}" is ok.	The power supply reports no issues.		
MS0123	STATE_CHANGE	ERROR	The power supply "{0}" is in error.	The power supply reports an error.		Follow the maintenance procedure for the power supply.
MS0124	STATE_CHANGE	ERROR	The power supply "{0}" is missing.	The power supply is not visible by any node in the cluster.		Follow the maintenance procedure for checking the attachment of the power supply.
MS0126	STATE_CHANGE	INFO	The temperature of power supply "{0}" is unknown.	The temperature of the power supply could not be determined.		
MS0127	STATE_CHANGE	INFO	The temperature of power supply "{0}" is ok ({1}).	The power supply reports no temperature issue.		
MS0128	STATE_CHANGE	ERROR	The temperature of power supply "{0}" is too high ({1}).	The power supply reports too high temperature.		Follow the maintenance procedure for the power supply.



Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0130	STATE_CHANGE	INFO	The pool {0} of file system {2} reached a normal level.	The pool reached a normal level.		None.
MS0131	STATE_CHANGE	WARNING	The pool {0} of file system {2} reached a warning level. {1}	The pool reached a warning level.		Add more capacity to pool or move data to different pool or delete data and/or snapshots.
MS0132	STATE_CHANGE	ERROR	The pool {0} of file system {2} reached a nearly exhausted level. {1}	The pool reached a nearly exhausted level.		Add more capacity to pool or move data to different pool or delete data and/or snapshots.
MS0133	STATE_CHANGE	INFO	The state of pool {0} in file system {2} is unknown.	Could not determine fill state of the pool.		None.
MS0135	STATE_CHANGE	INFO	The pool {0} of file system {2} reached a normal metadata level.	The pool reached a normal level.		None.
MS0136	STATE_CHANGE	WARNING	The pool {0} of file system {2} reached a warning level for metadata. {1}	The pool reached a warning level.		Add more capacity to pool or move data to different pool or delete data and/or snapshots.
MS0137	STATE_CHANGE	ERROR	The pool {0} of file system {2} reached a nearly exhausted metadata level. {1}	The pool reached a nearly exhausted level.		Add more capacity to pool or move data to different pool or delete data and/or snapshots.
MS0138	STATE_CHANGE	INFO	The state of pool {0} in file system {2} is unknown.	Could not determine fill state of metadata of the pool.		None.
MS0140	STATE_CHANGE	INFO	The local file systems on node {0} reached a normal level.	The local file systems reached a normal level.		None.
MS0141	STATE_CHANGE	WARNING	A local file system on node {0} reached a warning level. {1}	The local file systems reached a warning level.		Delete data on the local disk.
MS0142	STATE_CHANGE	ERROR	A local file system on node {0} reached a nearly exhausted level. {1}	The local file systems reached a nearly exhausted level.		Delete data on the local disk.
MS0143	STATE_CHANGE	INFO	The fill level of local file systems on node {0} is unknown.	Could not determine fill state of the local file systems.		None.
MS0145	STATE_CHANGE	INFO	The inode usage of fileset {0} in file system {2} reached a normal level.	The inode usage in the fileset reached a normal level.		None.
MS0146	STATE_CHANGE	WARNING	The inode usage of fileset {0} in file system {2} reached a warning level. {1}	The inode usage in the fileset reached warning level.		Delete data.
MS0147	STATE_CHANGE	ERROR	The inode usage of fileset {0} in file system {2} reached a nearly exhausted level. {1}	The inode usage in the fileset reached a nearly exhausted level.		Click on Action 'Run fix procedure'.
MS0148	STATE_CHANGE	INFO	The inode usage of fileset {0} could not to determined.	Could not determine inode usage of fileset.		None.
MS0151	STATE_CHANGE	INFO	The firmware level of ESM "{0}" is unknown.	The firmware level of a environmental service module could not be determined.		

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0152	STATE_CHANGE	INFO	The environmental service module "{0}" runs the recommended firmware ({1}).	The reported firmware level of the environmental service module is compliant with the recommendation.		
MS0153	STATE_CHANGE	WARNING	The environmental service module "{0}" does not run the recommended firmware ({1} actual/ recommended).	The reported firmware level of the environmental service module is not compliant with the recommendation.		Run mmchfirmware for the enclosure to update the firmware
MS0154	STATE_CHANGE	WARNING	No supported firmware available for the environmental service module "{0}" (part {1}).	The GPFS packages do not contain a firmware for this device. Most probably this device is not supported.		Check support status of device with IBM
MS0161	STATE_CHANGE	INFO	The firmware level of physical disk "{0}" is unknown.	The firmware level of a physical disk could not be determined.		
MS0162	STATE_CHANGE	INFO	The physical disk "{0}" runs the recommended firmware ({1}).	The reported firmware level of the physical disk is compliant with the recommendation.		
MS0163	STATE_CHANGE	WARNING	The physical disk "{0}" does not run the recommended firmware ({1} actual/ recommended).	The reported firmware level of the physical disk is not compliant with the recommendation.		Run mmchfirmware for the disk to update the firmware
MS0164	STATE_CHANGE	WARNING	No supported firmware available for the physical disk "{0}" (part {1}).	The GPFS packages do not contain a firmware for this device. Most probably this device is not supported.		Check support status of device with IBM
MS0170	STATE_CHANGE	INFO	The state of the 12V power line of disk enclosure "{0}" is unknown.	The state of the disk enclosure's power line could not be determined.		
MS0171	STATE_CHANGE	INFO	The 12V power line of disk enclosure "{0}" is ok (voltage={1}).	The disk enclosure reports no issues.		
MS0172	STATE_CHANGE	ERROR	The 12V power line of disk enclosure "{0}" failed (voltage={1}).	The disk enclosure reports a failure of the 12V power line.		Follow the maintenance procedure for the enclosure.
MS0173	STATE_CHANGE	WARNING	Communication to sensor of 12V power line of disk enclosure "{0}" lost.	The I2C bus used for reading the voltage sensor failed. Thus the voltage line cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0174	STATE_CHANGE	ERROR	Voltage of 12V power line of disk enclosure "{0}" exceeded critical threshold ({1}).	If the critical threshold is exceeded, the device functions might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0175	STATE_CHANGE	WARNING	Voltage of 12V power line of disk enclosure "{0}" exceeded warning threshold ({1}).	If the warning threshold is exceeded, the device function might be affected. However, this might indicate a hardware defect.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0176	STATE_CHANGE	ERROR	Voltage of 12V power line of disk enclosure "{0}" undercut the lower critical threshold ({1}).	If the lower critical threshold is undercut, the device functions and part might get affected.		Follow the maintenance procedure for the enclosure.
MS0177	STATE_CHANGE	WARNING	Voltage of 12V power line of disk enclosure "{0}" undercut the lower warning threshold ({1}).	If the lower warning threshold is undercut, the device functions might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0180	STATE_CHANGE	INFO	The temperature of power supply "{0}" is unknown.	The temperature of the power supply could not be determined.		
MS0181	STATE_CHANGE	INFO	The temperature of power supply "{0}" is ok at {1}.	The power supply reports no temperature issues.		
MS0182	STATE_CHANGE	ERROR	The temperature of power supply "{0}" failed ({1}).	The power supply reports a temperature problem.		Follow the maintenance procedure for the enclosure.
MS0183	STATE_CHANGE	WARNING	Communication to temperature sensor power supply "{0}" lost.	The I2C bus used for reading the temperature sensor failed. Thus the temperature cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0184	STATE_CHANGE	ERROR	Temperature of power supply "{0}" exceeded critical threshold at {1}.	If the critical threshold is exceeded, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0185	STATE_CHANGE	WARNING	Temperature of power supply "{0}" exceeded warning threshold at {1}.	If the warning threshold is exceeded, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0186	STATE_CHANGE	ERROR	Temperature of power supply "{0}" undercut the lower critical threshold at {1}.	If the lower critical threshold is undercut, the device functions might get affected and the parts could get damaged.		Follow the maintenance procedure for the enclosure.
MS0187	STATE_CHANGE	WARNING	Temperature of power supply "{0}" undercut the lower warning threshold at {1}.	If the lower warning threshold is undercut, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0190	STATE_CHANGE	INFO	The state of power supply "{0}" is unknown.	The state of the power supply could not be determined.		
MS0191	STATE_CHANGE	INFO	The power supply "{0}" is ok.	The power supply reports no issues.		
MS0192	STATE_CHANGE	ERROR	The power supply "{0}" failed.	The power supply reports a problem.		Follow the maintenance procedure for the enclosure.
MS0193	STATE_CHANGE	WARNING	Communication to power supply "{0}" lost.	Communication to power supply through the I2C bus failed. Thus the power supply cannot be monitored.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0194	STATE_CHANGE	WARNING	Power supply "{0}" is absent.	The power supply was removed. The enclosure is still operating but lost power redundancy.		Reinsert the power supply.
MS0195	STATE_CHANGE	WARNING	The power supply "{0}" is switched off.	The power supply was manually switched off. The enclosure is operational but lost power redundancy.		Switch on the power supply.
MS0196	STATE_CHANGE	ERROR	The power supply "{0}" is not providing power.	The power supply is off. The enclosure is still operational but lost power redundancy.		Follow the maintenance procedure for the enclosure.
MS0197	STATE_CHANGE	ERROR	The output voltage of power supply "{0}" is critically high.	The power supply has a critically high output voltage. This might damage electronic components inside the enclosure.		Follow the maintenance procedure for the enclosure.
MS0198	STATE_CHANGE	ERROR	The output current of power supply "{0}" is critically high.	The power supply has a critically high output current. This might damage the power supply.		Follow the maintenance procedure for the enclosure.
MS0199	STATE_CHANGE	ERROR	The power supply "{0}" has no input power.	The power supply has no input power. The enclosure is still operational but has lost power redundancy.		Check the power cable of the power supply.
MS0200	STATE_CHANGE	INFO	The state of the 1.0V power line of environmental service module "{0}" is unknown.	The state of the environmental service module's power line could not be determined.		
MS0201	STATE_CHANGE	INFO	The 1.0V power line of environmental service module "{0}" is ok ({1}).	The environmental service module reports no issues.		
MS0202	STATE_CHANGE	ERROR	The 1.0V power line of environmental service module "{0}" failed ({1}).	The environmental service module reports a failure of the 1.0V power line.		Follow the maintenance procedure for the enclosure.
MS0203	STATE_CHANGE	WARNING	Communication to sensor of 1.0V power line of environmental service module "{0}" lost.	The I2C bus used for reading the voltage sensor failed. Thus the voltage line cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0204	STATE_CHANGE	ERROR	Voltage of 1.0V power line of environmental service module "{0}" exceeded critical threshold ({1}).	If the critical threshold is exceeded, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0205	STATE_CHANGE	WARNING	Voltage of 1.0V power line of environmental service module "{0}" exceeded warning threshold ({1}).	If the warning threshold is exceeded, the device operation might get affected. However, this can also be because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0206	STATE_CHANGE	ERROR	Voltage of 1.0V power line of environmental service module "{0}" undercut the lower critical threshold ({1}).	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0207	STATE_CHANGE	WARNING	Voltage of 1.0V power line of environmental service module "{0}" undercut the lower warning threshold ({1}).	If the lower warning threshold is undercut, the device operation might get affected. However, this can also be because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0210	STATE_CHANGE	INFO	The state of the 1.2V power line of environmental service module "{0}" is unknown.	The state of the environmental service module's power line could not be determined.		
MS0211	STATE_CHANGE	INFO	The 1.2V power line of environmental service module "{0}" is ok ({1}).	The environmental service module reports no issues.		
MS0212	STATE_CHANGE	ERROR	The 1.2V power line of environmental service module "{0}" failed ({1}).	The environmental service module reports a failure of the 1.2V power line.		Follow the maintenance procedure for the enclosure.
MS0213	STATE_CHANGE	WARNING	Communication to sensor of 1.2V power line of environmental service module "{0}" lost.	The I2C bus used for reading the voltage sensor failed. Thus the voltage line cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0214	STATE_CHANGE	ERROR	Voltage of 1.2V power line of environmental service module "{0}" exceeded critical threshold ({1}).	If the critical threshold is exceeded, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0215	STATE_CHANGE	WARNING	Voltage of 1.2V power line of environmental service module "{0}" exceeded warning threshold ({1}).	If the warning threshold is exceeded, the device operation might get affected. However, this can also be because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0216	STATE_CHANGE	ERROR	Voltage of 1.2V power line of environmental service module "{0}" undercut the lower critical threshold ({1}).	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0217	STATE_CHANGE	WARNING	Voltage of 1.2V power line of environmental service module "{0}" undercut the lower warning threshold ({1}).	If the lower warning threshold is undercut, the device operation might get affected. However, this can also be because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0220	STATE_CHANGE	INFO	The state of the 3.3V power line of environmental service module "{0}" is unknown.	The state of the environmental service module's power line could not be determined.		
MS0221	STATE_CHANGE	INFO	The 3.3V power line of environmental service module "{0}" is ok ({1}).	The environmental service module reports no issues.		
MS0222	STATE_CHANGE	ERROR	The 3.3V power line of environmental service module "{0}" failed ({1}).	The environmental service module reports a failure of the 3.3V power line.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0223	STATE_CHANGE	WARNING	Communication to sensor of 3.3V power line of environmental service module "{0}" lost.	The I2C bus used for reading the voltage sensor failed. Thus the voltage line cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0224	STATE_CHANGE	ERROR	Voltage of 3.3V power line of environmental service module "{0}" exceeded critical threshold ({1}).	If the critical threshold is exceeded, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0225	STATE_CHANGE	WARNING	Voltage of 3.3V power line of environmental service module "{0}" exceeded warning threshold ({1}).	If the warning threshold is exceeded, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0226	STATE_CHANGE	ERROR	Voltage of 3.3V power line of environmental service module "{0}" undercut the lower critical threshold ({1}).	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0227	STATE_CHANGE	WARNING	Voltage of 3.3V power line of environmental service module "{0}" undercut the lower warning threshold ({1}).	If the lower warning threshold is undercut, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0230	STATE_CHANGE	INFO	The state of the 5.0V power line of environmental service module "{0}" is unknown.	The state of the environmental service module's power line could not be determined.		
MS0231	STATE_CHANGE	INFO	The 5.0V power line of environmental service module "{0}" is ok ({1}).	The environmental service module reports no issues.		
MS0232	STATE_CHANGE	ERROR	The 5.0V power line of environmental service module "{0}" failed ({1}).	The environmental service module reports a failure of the 5.0V power line.		Follow the maintenance procedure for the enclosure.
MS0233	STATE_CHANGE	WARNING	Communication to sensor of 5.0V power line of environmental service module "{0}" lost.	The I2C bus used for reading the voltage sensor failed. Thus, the voltage line cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0234	STATE_CHANGE	ERROR	Voltage of 5.0V power line of environmental service module "{0}" exceeded critical threshold ({1}).	If the critical threshold is exceeded, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0235	STATE_CHANGE	WARNING	Voltage of 5.0V power line of environmental service module "{0}" exceeded warning threshold ({1}).	If the warning threshold is exceeded, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0236	STATE_CHANGE	ERROR	Voltage of 5.0V power line of environmental service module "{0}" undercut the lower critical threshold ({1}).	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0237	STATE_CHANGE	WARNING	Voltage of 5.0V power line of environmental service module "{0}" undercut the lower warning threshold ({1}).	If the lower warning threshold is undercut, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0241	STATE_CHANGE	INFO	The temperature of drive carrier module "{0}" is unknown.	The temperature of the drive carrier module could not be determined.		
MS0242	STATE_CHANGE	INFO	The temperature of drive carrier module "{0}" is ok at {1}.	The drive carrier module reports no temperature issue.		
MS0243	STATE_CHANGE	ERROR	The temperature of drive carrier module "{0}" is too high at {1}.	The drive carrier module reports too high temperature.		Follow the maintenance procedure for the drive carrier module.
MS0244	STATE_CHANGE	WARNING	Communication to temperature sensor of drive carrier module "{0}" lost.	The I2C bus used for reading the temperature sensor failed. Thus the temperature cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0245	STATE_CHANGE	ERROR	Temperature of drive carrier module "{0}" exceeded critical threshold at {1}.	If the critical threshold is exceeded, the operation of the device might be affected or parts are damaged.		Follow the maintenance procedure for the enclosure.
MS0246	STATE_CHANGE	WARNING	Temperature of drive carrier module "{0}" exceeded warning threshold at {1}.	If the warning threshold is exceeded, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0247	STATE_CHANGE	ERROR	Temperature of drive carrier module "{0}" undercut the lower critical threshold at {1}.	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0248	STATE_CHANGE	WARNING	Temperature of drive carrier module "{0}" undercut the lower warning threshold at {1}.	If the lower warning threshold is undercut, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0251	STATE_CHANGE	INFO	The state of file system "{0}" is unknown.	The state of the file system could not be determined.		
MS0252	STATE_CHANGE	INFO	The state of file system "{0}" is OK.	The state of the file system is OK.		
MS0253	STATE_CHANGE	INFO	The file system "{0}" has errors. Failed NSDs: "{1}"	The file system is in error state, due to NSD failures.		Please run recommended fix procedure on the failing NSDs listed.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0261	STATE_CHANGE	INFO	The state of drive slot "{0}" is unknown.	The state of the drive slot could not be determined.		
MS0271	STATE_CHANGE	INFO	The SSH connection state of node "{0}" is unknown.	The SSH connection state of a node could not be determined.		
MS0272	STATE_CHANGE	INFO	SSH connection to node "{0}" established.	The GUI is able to communicate with the node through SSH.		
MS0273	STATE_CHANGE	ERROR	Connection failed to node "{0}". Network connection problem.	The GUI was unable to connect to a node because of a network problem (e.g. node route to host).		
MS0274	STATE_CHANGE	ERROR	Connection failed to node "{0}". SSH error.	The GUI received an SSH error when connecting to a node.		
MS0275	STATE_CHANGE	ERROR	Connection failed to node "{0}". Connection refused on SSH port.	The connection to the SSH port of a node was refused. Probably the SSH daemon is down or the firewall blocks access.		
MS0276	STATE_CHANGE	ERROR	Connection failed to node "{0}". The SSH version is unsupported.	The GUI could not connect to a node because the SSH version of client and daemon are incompatible.		
MS0277	STATE_CHANGE	ERROR	The node "{0}" runs on an unsupported operating system.	The GUI could connect to a node but as this node runs on a unsupported OS, it cannot manage it.		
MS0278	STATE_CHANGE	ERROR	Connection failed to node "{0}". Wrong credentials.	The GUI failed to connect to a node because the authentication is failed. Check that key-based access works for the root user.		
MS0279	STATE_CHANGE	ERROR	Connection failed to node "{0}". Wrong credentials.	The GUI failed to connect to a node because the authentication failed. Check that key-based access works for the root user.		
MS0281	STATE_CHANGE	INFO	The state of NSD "{0}" is unknown.	The state of an NSD could not be determined.		
MS0282	STATE_CHANGE	INFO	The NSD "{0}" is ready.	Everything fine with this NSD.		
MS0283	STATE_CHANGE	INFO	The NSD "{0}" is being emptied.	All data is migrated away from this NSD.		
MS0284	STATE_CHANGE	INFO	The NSD "{0}" is being replaced by another NSD. Data migration is ongoing.	All data is migrated from this NSD to another one.		
MS0285	STATE_CHANGE	INFO	The NSD "{0}" is replacing another NSD. Data migration is ongoing.	All data from another NSD is migrated to this NSD.		



Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0286	STATE_CHANGE	INFO	The NSD "{0}" is suspended.	The NSD is suspended. No data will be written to a suspended NSD.		
MS0287	STATE_CHANGE	ERROR	Errors while reading or writing data to NSD "{0}".	An error occurred while reading or writing data to the NSD.		
MS0288	STATE_CHANGE	INFO	The NSD "{0}" is in state to be emptied.	All data is migrated away from this NSD.		
MS0290	STATE_CHANGE	INFO	The ambient temperature of power supply "{0}" is unknown.	The ambient temperature of the power supply could not be determined.		
MS0291	STATE_CHANGE	INFO	The ambient temperature of power supply "{0}" is ok at {1}.	The power supply reports no ambient temperature issues.		
MS0292	STATE_CHANGE	ERROR	The ambient temperature of power supply "{0}" failed ({1}).	The power supply reports an ambient temperature problem.		Follow the maintenance procedure for the enclosure.
MS0293	STATE_CHANGE	WARNING	Communication to ambient temperature sensor power supply "{0}" lost.	The I2C bus used for reading the ambient temperature sensor failed. Thus the ambient temperature cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0294	STATE_CHANGE	ERROR	Ambient temperature of power supply "{0}" exceeded critical threshold at {1}.	If the critical threshold is exceeded, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0295	STATE_CHANGE	WARNING	Ambient temperature of power supply "{0}" exceeded warning threshold at {1}.	If the warning threshold is exceeded, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0296	STATE_CHANGE	ERROR	Ambient temperature of power supply "{0}" undercut the lower critical threshold at {1}.	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0297	STATE_CHANGE	WARNING	Ambient temperature of power supply "{0}" undercut the lower warning threshold at {1}.	If the lower warning threshold is undercut, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0300	STATE_CHANGE	INFO	The hot spot temperature of power supply "{0}" is unknown.	The hot spot temperature of the power supply could not be determined.		
MS0301	STATE_CHANGE	INFO	The hot spot temperature of power supply "{0}" is ok at {1}.	The power supply reports no hot spot temperature issues.		
MS0302	STATE_CHANGE	ERROR	The hot spot temperature of power supply "{0}" failed ({1}).	The power supply reports a hot spot temperature problem.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0303	STATE_CHANGE	WARNING	Communication to hot spot temperature sensor power supply "{0}" lost.	The I2C bus used for reading the hot spot temperature sensor failed. Thus the hot spot temperature cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0304	STATE_CHANGE	ERROR	Hot spot temperature of power supply "{0}" exceeded critical threshold at {1}.	If the critical threshold is exceeded, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0305	STATE_CHANGE	WARNING	Hot spot temperature of power supply "{0}" exceeded warning threshold at {1}.	If the warning threshold is exceeded, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0306	STATE_CHANGE	ERROR	Hot spot temperature of power supply "{0}" undercut the lower critical threshold at {1}.	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0307	STATE_CHANGE	WARNING	Hot spot temperature of power supply "{0}" undercut the lower warning threshold at {1}.	If the lower warning threshold is undercut, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0310	STATE_CHANGE	INFO	The primary heat sink temperature of power supply "{0}" is unknown.	The primary heat sink temperature of the power supply could not be determined.		
MS0311	STATE_CHANGE	INFO	The primary heat sink temperature of power supply "{0}" is ok at {1}.	The power supply reports no primary heat sink temperature issues.		
MS0312	STATE_CHANGE	ERROR	The primary heat sink temperature of power supply "{0}" failed ({1}).	The power supply reports a primary heat sink temperature problem.		Follow the maintenance procedure for the enclosure.
MS0313	STATE_CHANGE	WARNING	Communication to primary heat sink temperature sensor power supply "{0}" lost.	The I2C bus used for reading the primary heat sink temperature sensor failed. Thus the primary heat sink temperature cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0314	STATE_CHANGE	ERROR	Primary heat sink temperature of power supply "{0}" exceeded critical threshold at {1}.	If the critical threshold is exceeded, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0315	STATE_CHANGE	WARNING	Primary heat sink temperature of power supply "{0}" exceeded warning threshold at {1}.	If the warning threshold is exceeded, the device operation might get affected. However, this can also because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0316	STATE_CHANGE	ERROR	Primary heat sink temperature of power supply "{0}" undercut the lower critical threshold at {1}.	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0317	STATE_CHANGE	WARNING	Primary heat sink temperature of power supply "{0}" undercut the lower warning threshold at {1}.	If the lower warning threshold is undercut, the device operation might get affected. However, this can also be because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0320	STATE_CHANGE	INFO	The state of the 12V power line of power supply "{0}" is unknown.	The state of the power supply's power line could not be determined.		
MS0321	STATE_CHANGE	INFO	The 12V line of power supply "{0}" is ok ({1}).	The power supply reports no issues for the 12V line.		
MS0322	STATE_CHANGE	ERROR	The 12V power line of power supply "{0}" failed ({1}).	The power supply reports a failure of the 12V power line.		Follow the maintenance procedure for the enclosure.
MS0323	STATE_CHANGE	WARNING	Communication to sensor of 12V power line of power supply "{0}" lost.	The I2C bus used for reading the voltage sensor failed. Thus the voltage line cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0324	STATE_CHANGE	ERROR	Voltage of 12V power line of power supply "{0}" exceeded critical threshold ({1}).	If the critical threshold is exceeded, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0325	STATE_CHANGE	WARNING	Voltage of 12V power line of power supply "{0}" exceeded warning threshold ({1}).	If the warning threshold is exceeded, the device operation might get affected. However, this can also be because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0326	STATE_CHANGE	ERROR	Voltage of 12V power line of power supply "{0}" undercut the lower critical threshold ({1}).	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0327	STATE_CHANGE	WARNING	Voltage of 12V power line of power supply "{0}" undercut the lower warning threshold ({1}).	If the lower warning threshold is undercut, the device operation might get affected. However, this can also be because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0330	STATE_CHANGE	INFO	The state of a 5V power line of power supply "{0}" is unknown.	The state of the power supply's power line could not be determined.		
MS0331	STATE_CHANGE	INFO	A 5V line of power supply "{0}" is ok ({1}).	The power supply reports no issues for the 5V line.		
MS0332	STATE_CHANGE	ERROR	The 5V power line of power supply "{0}" failed ({1}).	The power supply reports a failure of the 5V power line.		Follow the maintenance procedure for the enclosure.
MS0333	STATE_CHANGE	WARNING	Communication to sensor of 5V power line of power supply "{0}" lost.	The I2C bus used for reading the voltage sensor failed. Thus the voltage line cannot be monitored.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0334	STATE_CHANGE	ERROR	Voltage of 5V power line of power supply "{0}" exceeded critical threshold ({1}).	If the critical threshold is exceeded, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0335	STATE_CHANGE	WARNING	Voltage of 5V power line of power supply "{0}" exceeded warning threshold ({1}).	If the warning threshold is exceeded, the device operation might get affected. However, this can also be because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0336	STATE_CHANGE	ERROR	Voltage of 5V power line of power supply "{0}" undercut the lower critical threshold ({1}).	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0337	STATE_CHANGE	WARNING	Voltage of 5V power line of power supply "{0}" undercut the lower warning threshold ({1}).	If the lower warning threshold is undercut, the device operation might get affected. However, this can also be because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0341	STATE_CHANGE	INFO	The firmware level of host adapter "{0}" is unknown.	The firmware level of a host adapter could not be determined.		
MS0342	STATE_CHANGE	INFO	The host adapter "{0}" runs the recommended firmware ({1}).	The reported firmware level of the host adapter is compliant with the recommendation.		
MS0343	STATE_CHANGE	WARNING	The host adapter "{0}" does not run the recommended firmware ({1} actual/ recommended).	The reported firmware level of the host adapter is not compliant with the recommendation.		Run 'mmchfirmware --type host-adapter' for the corresponding node to update the firmware
MS0344	STATE_CHANGE	WARNING	No supported firmware available for the host adapter "{0}" (part {1}).	The GPFS packages do not contain a firmware for this device. Most probably this device is not supported.		Check support status of device with IBM
MS0350	STATE_CHANGE	INFO	Availability of xCAT on cluster "{0}" is unknown.	The availability and state of xCAT could not be determined.		
MS0351	STATE_CHANGE	INFO	The availability of xCAT on cluster "{0}" is OK.	The availability and state of xCAT is OK.		
MS0352	STATE_CHANGE	WARNING	The xCAT host is not configured on cluster "{0}".	The host where xCAT is located is not specified.		Specify the host name or IP where xCAT is located.
MS0353	STATE_CHANGE	ERROR	Unable to connect to xCAT node "{1}" on cluster "{0}".	Cannot connect to the node specified as xCAT host.		Check that the IP address is correct and ensure that root does have key-based SSH set up to the xCAT node.
MS0354	STATE_CHANGE	ERROR	The xCAT on node "{1}" fail to operate properly on cluster "{0}".	The node specified as xCAT host is reachable but either xCAT is not installed on the node or not operating properly.		Check xCAT installation and try xCAT commands nodes, rin and rvitals for errors.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0355	STATE_CHANGE	WARNING	The xCAT service has not the recommended version ({1} actual/recommended)	The reported version of xCAT is not compliant with the recommendation.		Install the recommended xCAT version.
MS0360	STATE_CHANGE	INFO	State of performance monitoring collector service 'pmcollector' on cluster "{0}" is unknown.	The state of pmcollector service could not be determined.		
MS0361	STATE_CHANGE	INFO	The state of performance monitoring collector service 'pmcollector' on cluster "{0}" is OK.	The state of performance monitoring collector service pmcollector is OK and is ready to receive data from the performance monitoring sensor services pmsensor.		
MS0362	STATE_CHANGE	WARNING	The performance monitoring collector service 'pmcollector' on cluster "{0}", node "{1}" is not configured.	The performance monitoring collector service pmcollector could not be found on listed node.		Check performance monitoring collector service pmcollector is installed properly on listed node by issuing the <b>systemctl status pmcollector</b> command.
MS0363	STATE_CHANGE	ERROR	No connection to the performance monitoring collector service 'pmcollector' on cluster "{0}", node "{1}".	The performance monitoring collector service 'pmcollector' could not be connected on the node listed.		Check the pmcollector log files that are located at /var/log/zimon/ZIMonCollector.log. Issue the <b>systemctl status pmcollector</b> command to check the status of pmcollector service. If the pmcollector service is inactive, issue <b>systemctl restart pmcollector</b> command to restart the service.
MS0364	STATE_CHANGE	ERROR	The performance monitoring collector service 'pmcollector' on cluster "{0}", node "{1}" is inactive.	The performance monitoring collector service 'pmcollector' is in inactive state. The service might be down.		Issue the <b>systemctl status pmcollector</b> command to check the status of pmcollector service. If the pmcollector service is inactive, issue <b>systemctl restart pmcollector</b> command to restart the service.
MS0370	STATE_CHANGE	INFO	State of performance monitoring sensor service 'pmsensor' on node "{0}" is unknown.	The state of performance monitoring sensor service 'pmsensor' on the node could not be determined.		

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0371	STATE_CHANGE	INFO	The state of performance monitoring sensor service 'pmsensor' on node "{0}" is OK.	The state of performance monitoring sensor service 'pmsensor' is OK and it is sending data.		
MS0372	STATE_CHANGE	ERROR	The performance monitoring sensor service 'pmsensors' on node "{0}" is not sending any data.	The performance monitoring sensor service 'pmsensors' is not sending any data. The service might be down or the time of the node is more than 15 minutes away from the time on the node hosting the performance monitoring collector service 'pmcollector'.		Issue the <b>systemctl status pmsensors</b> command to check the status of the pmsensors service. If the pmsensors service is inactive, issue <b>systemctl restart pmsensors</b> command to restart the service.
MS0380	STATE_CHANGE	INFO	State of the node "{0}" in xCAT is unknown.	The state of the node within xCAT could not be determined.		
MS0381	STATE_CHANGE	INFO	The node "{0}" is known by xCAT.	xCAT knows about the node and manages it.		
MS0382	STATE_CHANGE	ERROR	The node "{0}" is unknown by xCAT.	The xCAT does not know about the node.		Add the node to xCAT and ensure that the host name used in xCAT matches the host name known by the node itself.
MS0390	STATE_CHANGE	INFO	The state of power supply "{0}" is unknown.	The state of the power supply could not be determined.		
MS0391	STATE_CHANGE	INFO	The power supply "{0}" is ok.	The power supply reports no issues.		
MS0392	STATE_CHANGE	ERROR	The power supply "{0}" failed.	The power supply reports a problem.		Follow the maintenance procedure for the enclosure.
MS0393	STATE_CHANGE	WARNING	Communication to power supply "{0}" lost.	Communication to power supply through the I2C bus failed. Thus the power supply cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0394	STATE_CHANGE	WARNING	Power supply "{0}" is absent.	The power supply was removed. The enclosure is still operating but lost power redundancy.		Reinsert the power supply.
MS0395	STATE_CHANGE	INFO	The power supply "{0}" is switched off.	The power supply was manually switched off. The enclosure is operational but lost power redundancy.		Switch on the power supply.
MS0396	STATE_CHANGE	WARNING	The power supply "{0}" is not providing power.	The power supply is off. The enclosure is still operational but lost power redundancy.		Follow the maintenance procedure for the enclosure.
MS0397	STATE_CHANGE	ERROR	The output voltage of power supply "{0}" is critically high.	The power supply has a critically high output voltage. This might damage electronic components inside the enclosure.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0398	STATE_CHANGE	ERROR	The output current of power supply "{0}" is critically high.	The power supply has a critically high output current. This might damage the power supply.		Follow the maintenance procedure for the enclosure.
MS0399	STATE_CHANGE	WARNING	The power supply "{0}" has no input power.	The power supply has no input power. The enclosure is still operational but has lost power redundancy.		Check the power cable of the power supply.
MS0400	STATE_CHANGE	INFO	Time synchronization on node "{0}" is unknown.	The time synchronization on the node could not be determined.		
MS0401	STATE_CHANGE	INFO	Time on node "{0}" is in sync with cluster.	The time on the node is synchronous with the time on the GUI node. A difference of a predefined number of seconds is allowed.		
MS0402	STATE_CHANGE	ERROR	Time of node "{0}" is out of sync. The time difference is {1}.	The time on the node is not synchronous with the median time in the cluster. It differs more than 1 minute.		Sync the time with the time of other cluster nodes.
MS0430	STATE_CHANGE	INFO	The state of a 5V power line of power supply "{0}" is unknown.	The state of the power supply's power line could not be determined.		
MS0431	STATE_CHANGE	INFO	A 5V line of power supply "{0}" is ok ({1}).	The power supply reports no issues for the 5V line.		
MS0432	STATE_CHANGE	ERROR	The 5V power line of power supply "{0}" failed ({1}).	The power supply reports a failure of the 5V power line.		Follow the maintenance procedure for the enclosure.
MS0433	STATE_CHANGE	WARNING	Communication to sensor of 5V power line of power supply "{0}" lost.	The I2C bus used for reading the voltage sensor failed. Thus the voltage line cannot be monitored.		Follow the maintenance procedure for the enclosure.
MS0434	STATE_CHANGE	ERROR	Voltage of 5V power line of power supply "{0}" exceeded critical threshold ({1}).	If the critical threshold is exceeded, the operation of the device might be affected or parts are damaged.		Follow the maintenance procedure for the enclosure.
MS0435	STATE_CHANGE	WARNING	Voltage of 5V power line of power supply "{0}" exceeded warning threshold ({1}).	If the warning threshold is exceeded, the device operation might get affected. However, this can also be because of a hardware defect.		Follow the maintenance procedure for the enclosure.
MS0436	STATE_CHANGE	ERROR	Voltage of 5V power line of power supply "{0}" undercut the lower critical threshold ({1}).	If the lower critical threshold is undercut, the operation of the device might be affected or parts damaged.		Follow the maintenance procedure for the enclosure.
MS0437	STATE_CHANGE	WARNING	Voltage of 5V power line of power supply "{0}" undercut the lower warning threshold ({1}).	If the lower warning threshold is undercut, the device operation might get affected. However, this can also be because of a hardware defect.		Follow the maintenance procedure for the enclosure.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0500	STATE_CHANGE	INFO	The overall CES state is healthy on node "{0}".	The state of a CES service changed.		
MS0501	STATE_CHANGE	WARNING	The overall CES state is degraded on node "{0}".	The state of a CES service changed.		
MS0502	STATE_CHANGE	INFO	The overall CES state is unknown on node "{0}".	The state of a CES service changed.		
MS0503	STATE_CHANGE	INFO	The overall CES state is disabled on node "{0}".	The state of a CES service changed.		
MS0504	STATE_CHANGE	INFO	The overall CES state is stopped on node "{0}".	The state of a CES service changed.		
MS0505	STATE_CHANGE	INFO	The overall CES state is suspended on node "{0}".	The state of a CES service changed.		
MS0506	STATE_CHANGE	ERROR	The overall CES state is failed on node "{0}".	The state of a CES service changed.		
MS0507	STATE_CHANGE	WARNING	The overall CES state is dependency failed on node "{0}".	The state of a CES service changed.		
MS0508	STATE_CHANGE	INFO	The overall CES state is starting on node "{0}".	The state of a CES service changed.		
MS0510	STATE_CHANGE	INFO	The NFS service is healthy on node "{0}".	The state of a CES service changed.		
MS0511	STATE_CHANGE	WARNING	The NFS service is degraded on node "{0}".	The state of a CES service changed.		
MS0512	STATE_CHANGE	INFO	The NFS service is unknown on node "{0}".	The state of a CES service changed.		
MS0513	STATE_CHANGE	INFO	The NFS service is disabled on node "{0}".	The state of a CES service changed.		
MS0514	STATE_CHANGE	INFO	The NFS service is stopped on node "{0}".	The state of a CES service changed.		
MS0515	STATE_CHANGE	INFO	The NFS service is suspended on node "{0}".	The state of a CES service changed.		
MS0516	STATE_CHANGE	ERROR	The NFS service is failed on node "{0}".	The state of a CES service changed.		
MS0517	STATE_CHANGE	WARNING	The NFS service is dependency failed on node "{0}".	The state of a CES service changed.		
MS0518	STATE_CHANGE	INFO	The NFS service is starting on node "{0}".	The state of a CES service changed.		
MS0520	STATE_CHANGE	INFO	The SMB service is healthy on node "{0}".	The state of a CES service changed.		
MS0521	STATE_CHANGE	WARNING	The SMB service is degraded on node "{0}".	The state of a CES service changed.		
MS0522	STATE_CHANGE	INFO	The SMB service is unknown on node "{0}".	The state of a CES service changed.		



Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0523	STATE_CHANGE	INFO	The SMB service is disabled on node "{0}".	The state of a CES service changed.		
MS0524	STATE_CHANGE	INFO	The SMB service is stopped on node "{0}".	The state of a CES service changed.		
MS0525	STATE_CHANGE	INFO	The SMB service is suspended on node "{0}".	The state of a CES service changed.		
MS0526	STATE_CHANGE	ERROR	The SMB service is failed on node "{0}".	The state of a CES service changed.		
MS0527	STATE_CHANGE	WARNING	The SMB service is dependency failed on node "{0}".	The state of a CES service changed.		
MS0528	STATE_CHANGE	INFO	The SMB service is starting on node "{0}".	The state of a CES service changed.		
MS0530	STATE_CHANGE	INFO	The object service is healthy on node "{0}".	The state of a CES service changed.		
MS0531	STATE_CHANGE	WARNING	The object service is degraded on node "{0}".	The state of a CES service changed.		
MS0532	STATE_CHANGE	INFO	The object service is unknown on node "{0}".	The state of a CES service changed.		
MS0533	STATE_CHANGE	INFO	The object service is disabled on node "{0}".	The state of a CES service changed.		
MS0534	STATE_CHANGE	INFO	The object service is stopped on node "{0}".	The state of a CES service changed.		
MS0535	STATE_CHANGE	INFO	The object service is suspended on node "{0}".	The state of a CES service changed.		
MS0536	STATE_CHANGE	ERROR	The object service is failed on node "{0}".	The state of a CES service changed.		
MS0537	STATE_CHANGE	WARNING	The object service is dependency failed on node "{0}".	The state of a CES service changed.		
MS0538	STATE_CHANGE	INFO	The object service is starting on node "{0}".	The state of a CES service changed.		
MS0540	STATE_CHANGE	INFO	The network state is healthy on node "{0}".	The state of a CES service changed.		
MS0541	STATE_CHANGE	WARNING	The network state is degraded on node "{0}".	The state of a CES service changed.		
MS0542	STATE_CHANGE	INFO	The network state is unknown on node "{0}".	The state of a CES service changed.		
MS0543	STATE_CHANGE	INFO	The network state is disabled on node "{0}".	The state of a CES service changed.		
MS0544	STATE_CHANGE	INFO	The network state is stopped on node "{0}".	The state of a CES service changed.		
MS0545	STATE_CHANGE	INFO	The network state is suspended on node "{0}".	The state of a CES service changed.		
MS0546	STATE_CHANGE	ERROR	The network state is failed on node "{0}".	The state of a CES service changed.		

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0547	STATE_CHANGE	WARNING	The network state is dependency failed on node "{0}".	The state of a CES service changed.		
MS0548	STATE_CHANGE	INFO	The network state is starting on node "{0}".	The state of a CES service changed.		
MS0550	STATE_CHANGE	INFO	The authentication service is healthy on node "{0}".	The state of a CES service changed.		
MS0551	STATE_CHANGE	WARNING	The authentication service is degraded on node "{0}".	The state of a CES service changed.		
MS0552	STATE_CHANGE	INFO	The authentication service is unknown on node "{0}".	The state of a CES service changed.		
MS0553	STATE_CHANGE	INFO	The authentication service is disabled on node "{0}".	The state of a CES service changed.		
MS0554	STATE_CHANGE	INFO	The authentication service is stopped on node "{0}".	The state of a CES service changed.		
MS0555	STATE_CHANGE	INFO	The authentication service is suspended on node "{0}".	The state of a CES service changed.		
MS0556	STATE_CHANGE	ERROR	The authentication service is failed on node "{0}".	The state of a CES service changed.		
MS0557	STATE_CHANGE	WARNING	The authentication service is dependency failed on node "{0}".	The state of a CES service changed.		
MS0558	STATE_CHANGE	INFO	The authentication service is starting on node "{0}".	The state of a CES service changed.		
MS0560	STATE_CHANGE	INFO	The object authentication service is healthy on node "{0}".	The state of a CES service changed.		
MS0561	STATE_CHANGE	WARNING	The object authentication service is degraded on node "{0}".	The state of a CES service changed.		
MS0562	STATE_CHANGE	INFO	The object authentication service is unknown on node "{0}".	The state of a CES service changed.		
MS0563	STATE_CHANGE	INFO	The object authentication service is disabled on node "{0}".	The state of a CES service changed.		
MS0564	STATE_CHANGE	INFO	The object authentication service is stopped on node "{0}".	The state of a CES service changed.		
MS0565	STATE_CHANGE	INFO	The object authentication service is suspended on node "{0}".	The state of a CES service changed.		
MS0566	STATE_CHANGE	ERROR	The object authentication service is failed on node "{0}".	The state of a CES service changed.		

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS0567	STATE_CHANGE	WARNING	The object authentication service is dependency failed on node "{0}".	The state of a CES service changed.		
MS0568	STATE_CHANGE	INFO	The object authentication service is starting on node "{0}".	The state of a CES service changed.		
MS1001	STATE_CHANGE	INFO	Server "{0}" state is unknown.	The last attempt to query the server state via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the server state on a regular basis.	Check if the server is operating.
MS1002	STATE_CHANGE	INFO	Server "{0}" is fully available.	The server reports no issue.		
MS1005	STATE_CHANGE	ERROR	Server "{0}" is missing.	The availability of a server changed to "unavailable".		Fix physical server issues and retry recovery.
MS1011	STATE_CHANGE	INFO	Overall CPU state of "{0}" is unknown.	The last attempt to query the CPU state via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the CPU state on a regular basis.	
MS1012	STATE_CHANGE	INFO	All CPUs of "{0}" are fully available.	The last query of overall CPU state reports no issue.		
MS1014	STATE_CHANGE	ERROR	At least one CPU of "{0}" failed.	The last query of overall CPU state reports error.		Check the other CPU related events, and act according to what they suggest.
MS1021	STATE_CHANGE	INFO	Overall DIMM state of "{0}" is unknown.	The last attempt to query the DIMM state via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the DIMM state on a regular basis.	
MS1022	STATE_CHANGE	INFO	All DIMMs of "{0}" are fully available.	The last query of overall DIMM state reports no issue.		
MS1024	STATE_CHANGE	ERROR	At least one DIMM of "{0}" failed.	The last query of overall DIMM state reports error.		Check the other DIMM related events, and act according to what they suggest.
MS1031	STATE_CHANGE	INFO	Overall PCI state of "{0}" is unknown.	The last attempt to query the PCI state via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the PCI state on a regular basis.	
MS1032	STATE_CHANGE	INFO	All PCIs of "{0}" are fully available.	The last query of overall PCI state reports no issue.		
MS1034	STATE_CHANGE	ERROR	At least one PCI of "{0}" failed.	The last query of overall PCI state reports error.		Check the other PCI related events, and act according to what they suggest.
MS1041	STATE_CHANGE	INFO	All power supply configurations state of "{0}" is unknown.	The last attempt to query the power supply state via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the power supply state on a regular basis.	

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS1042	STATE_CHANGE	INFO	Power supply configuration of "{0}" is correct.	The last query of power supply configuration state reports no issue.		
MS1044	STATE_CHANGE	ERROR	Power supply configuration of "{0}" is invalid.	The last query of power supply configuration state reports error.		Follow the maintenance procedure for replacing a power supply.
MS1051	STATE_CHANGE	INFO	Overall power supply load state of "{0}" is unknown.	The system will retry to query the power supply load state on a regular basis.		
MS1052	STATE_CHANGE	INFO	No heavy load on power supplies of "{0}".	The last query of power supply state reports normal load on the power supplies.		
MS1054	STATE_CHANGE	ERROR	Heavy load on power supplies of "{0}".	The last query of power supplies state reports heavy load on them.		Reduce the total power consumption by removing newly added or unused option like drives or adapters.
MS1061	STATE_CHANGE	INFO	Overall power supply resources state of "{0}" is unknown.	The last attempt to query the overall power supply resource state via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the overall power supply resource state on a regular basis.	
MS1062	STATE_CHANGE	INFO	All power supply resources of "{0}" are fully available.	The last query of overall power supply state reports that sufficient resources are available.		
MS1064	STATE_CHANGE	ERROR	At least one of the power supply of "{0}" has insufficient resources.	The last query of overall power supply state reports that there is insufficient resources. The redundancy lost.		
MS1071	STATE_CHANGE	INFO	Power supply unit state of "{0}" is unknown.	The last attempt to query the power supply unit state via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the power supply unit state on a regular basis.	
MS1072	STATE_CHANGE	INFO	Power supply unit of "{0}" is fully redundant.	The last query of power supply unit state reports that it is redundant.		
MS1074	STATE_CHANGE	ERROR	Power supply unit of "{0}" redundancy is lost.	The last query reports that Power Supply Redundancy has been lost.		Check if Power Supply is missing or unplugged; check the power supply LEDs. Follow the maintenance procedure for replacing a power supply.
MS1081	STATE_CHANGE	INFO	Ambient temperature of "{0}" is unknown.	The last attempt to query the ambient temperature via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the ambient temperature on a regular basis.	

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS1082	STATE_CHANGE	INFO	Ambient temperature of "{0}" is normal ("{1}").	The last query of ambient temperature reports that it is within the normal range.		
MS1084	STATE_CHANGE	ERROR	Ambient temperature of "{0}" is too high ("{1}").	The last query of the ambient temperature reports that it is too high.		Check the environmental conditions of the rack, especially the cooling.
MS1091	STATE_CHANGE	INFO	Planar state of "{0}" is unknown.	The last attempt to query the planar state via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the planar state on a regular basis.	
MS1092	STATE_CHANGE	INFO	Planar state of {0} is healthy, the voltage is normal ({1}).	The last query of overall planar state reports no issue. Also, the planar voltage is within normal range.		
MS1094	STATE_CHANGE	ERROR	Planar state of "{0}" is unhealthy, the voltage is too low or too high ("{10}").	The last query of planar state reports error regarding the power lines. The planar voltage is out of normal range, either too high or too low.		Follow the maintenance procedure for checking the planar voltage and replace the system board if needed.
MS1101	STATE_CHANGE	INFO	System board state of "{0}" is unknown.	The last attempt to query the system board state via xCAT failed. Therefore no status can be reported momentarily.		
MS1102	STATE_CHANGE	INFO	System board of "{0}" is healthy.	The last query of system board state reports no issue.		
MS1104	STATE_CHANGE	ERROR	System board of "{0}" is unhealthy.	The last query of system board state reports error.		Follow the maintenance procedure for replacing the system board.
MS1111	STATE_CHANGE	INFO	System boot status of "{0}" is unknown.	The last attempt to query the system boot state via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the system boot state on a regular basis.	
MS1112	STATE_CHANGE	INFO	System boot status of "{0}" is normal.	The last query of system boot state reports no issue.		
MS1114	STATE_CHANGE	ERROR	System boot failed on "{0}".	The last query of system boot state reports failure.		Check the error code for this event among the hardware event messages and act accordingly.
MS1121	STATE_CHANGE	INFO	System event log status of "{0}" is unknown.	The last attempt to query the system event log state via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the system event log state on a regular basis.	
MS1122	STATE_CHANGE	INFO	System event log of "{0}" operates normal.	The last query of system event log state reports no issue.		

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS1124	STATE_CHANGE	ERROR	System event log of "{0}" is full.	The last query of system event log state reports that it is full.	The system will automatically free the system event log area.	Nothing
MS1131	STATE_CHANGE	INFO	CPU "{0}" temperature is unknown.	The last attempt to query the CPU temperature via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the CPU temperature on a regular basis.	
MS1132	STATE_CHANGE	INFO	CPU "{0}" temperature is normal ("1").	The last query of CPU state reports that the temperature within normal range.		
MS1134	STATE_CHANGE	ERROR	CPU "{0}" temperature is too high (11).	The last query of CPU state reports that its temperature is too high.		Check the environmental conditions of the rack, especially the cooling.
MS1141	STATE_CHANGE	INFO	PECI state of "{0}" is unknown.	The last attempt to query the PEFI state of the CPU via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the PEFI state on a regular basis.	
MS1142	STATE_CHANGE	INFO	PECI state of "{0}" is healthy.	The last query of PEFI state of the CPU reports no issue.		
MS1144	STATE_CHANGE	ERROR	PECI state of "{0}" is unhealthy.	The last query of PEFI state of the CPU reports Fatal Bus Error.		Reset the microprocessor, and then restart the server or follow the maintenance procedure to replace the CPU.
MS1151	STATE_CHANGE	INFO	QPI Link state of "{0}" is unknown.	The last attempt to query the QPI Link state of the CPU via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the QPI Link state of the CPU on a regular basis.	
MS1152	STATE_CHANGE	INFO	QPI Link of "{0}" is healthy.	The last query of QPI Link state of the CPU reports no issue.		
MS1154	STATE_CHANGE	ERROR	QPI Link of "{0}" is unhealthy.	External QPI Link Failure detected.		Check IBM support site for an applicable service bulletin or firmware update that applies to this error. Inspect processor socket for foreign debris or damage. If debris is found remove debris. If error recurs, or socket damage is found, follow the maintenance procedure to replace the system board.
MS1161	STATE_CHANGE	INFO	DASD Backplane "{0}" state is unknown.	The last attempt to query the DASD Backplane state via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the DASD Backplane state on a regular basis.	

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS1162	STATE_CHANGE	INFO	DASD Backplane "{0}" is present.	The last query of DASD Backplane state reports no issue.		
MS1164	STATE_CHANGE	ERROR	DASD Backplane "{0}" is absent.	The last query of DASD Backplane state reports that it is absent.		Check the absence of DASD Backplane. Follow the maintenance procedure to replace DASD Backplane.
MS1171	STATE_CHANGE	INFO	DIMM "{0}" state is unknown.	The last attempt to query the DIMM state via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the DIMM state on a regular basis.	
MS1172	STATE_CHANGE	INFO	DIMM "{0}" is healthy.	The last query of DIMM state reports no issue.		
MS1174	STATE_CHANGE	ERROR	DIMM "{0}" is unhealthy.	The last query of DIMM state reports error.		Reset the DIMM, and then restart the server. Follow the maintenance procedure to replace DIMM.
MS1181	STATE_CHANGE	INFO	Drive "{0}" state is unknown.	The last attempt to query the Drive state via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the Drive state on a regular basis.	
MS1182	STATE_CHANGE	INFO	Drive "{0}" is healthy.	The last query of overall Drive state reports no issue.		
MS1184	STATE_CHANGE	ERROR	Drive "{0}" is unhealthy.	The last query of drive state reports error.		Follow the maintenance procedure and replace the failing drive.
MS1191	STATE_CHANGE	INFO	Fan Zone ("{0}") state is unknown.	The last attempt to query the Fan Zone state via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the Fan Zone state on a regular basis.	
MS1192	STATE_CHANGE	INFO	Fan Zone "{0}" is fully redundant.	The last query of Fan Zone state reports no issue it is fully redundant.		
MS1194	STATE_CHANGE	ERROR	Fan Zone "{0}" redundancy is lost.	The last query of Fan Zone state reports error which lead to redundancy loss.		Follow the maintenance procedure and replace the failing fan.
MS1201	STATE_CHANGE	INFO	PCI "{0}" state is unknown.	The last attempt to query the PCI state via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the CI state on a regular basis.	
MS1202	STATE_CHANGE	INFO	PCI "{0}" is healthy.	The last query of PCI state reports no issue.		
MS1204	STATE_CHANGE	ERROR	PCI "{0}" is unhealthy.	The last query of PCI state reports error.		Follow the maintenance procedure to replace the system board.

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS1211	STATE_CHANGE	INFO	PCI Riser "{0}" temperature state is unknown.	The last attempt to query the PCI Riser state via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the PCI Riser state on a regular basis.	
MS1212	STATE_CHANGE	INFO	PCI Riser "{0}" temperature is normal ("1").	The last query of PCI Riser state reports no issue; its temperature is within normal range.		
MS1214	STATE_CHANGE	ERROR	PCI Riser "{0}" temperature is too high ("1").	The last query of PCI Riser state reports that the temperature is out of normal range.		Check the environment temperature; the cooling of the system.
MS1221	STATE_CHANGE	INFO	Fan of Power Supply "{0}" state is unknown.	The last attempt to query the Fan state via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the Fan state on a regular basis.	
MS1222	STATE_CHANGE	INFO	Fan of Power Supply "{0}" is healthy.	The last query of fan state reports no issue.		
MS1224	STATE_CHANGE	ERROR	Fan of Power Supply "{0}" is unhealthy.	The last query of fan state reports fan failure.		Follow the maintenance procedure and replace the failing fan.
MS1231	STATE_CHANGE	INFO	12V OC Line of Power Supply "{0}" state is unknown.	The last attempt to query the 12V OC Line state via xCAT failed. Therefore no status can be reported momentarily.	The system will retry to query the 12V OC Line state on a regular basis.	
MS1232	STATE_CHANGE	INFO	12V OC Line of Power Supply "{0}" is healthy.	The last query of 12V OC Line state reports no issue.		
MS1234	STATE_CHANGE	ERROR	12V OC Line of Power Supply "{0}" is unhealthy.	The last query of 12V OC Line state reports failure.		
MS1241	STATE_CHANGE	INFO	12V OV Line of Power Supply "{0}" state is unknown.	The last attempt to query the 12V OV Line state via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the 12V OV Line state on a regular basis.	
MS1242	STATE_CHANGE	INFO	12V OV Line of Power Supply "{0}" is healthy.	The last query of 12V OV Line state reports no issue.		
MS1244	STATE_CHANGE	ERROR	12V OV Line of Power Supply "{0}" is unhealthy.	The last query of 12V OV Line state reports failure.		
MS1251	STATE_CHANGE	INFO	12V UV Line of Power Supply "{0}" state is unknown.	The last attempt to query the 12V UV Line state via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the 12V UV state on a regular basis.	
MS1252	STATE_CHANGE	INFO	12V UV Line of Power Supply "{0}" is healthy.	The last query of 12V UV Line state reports no issue.		
MS1254	STATE_CHANGE	ERROR	12V UV Line of Power Supply "{0}" is unhealthy.	The last query of 12V UV Line state reports failure.		



Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS1261	STATE_CHANGE	INFO	12V aux Line of Power Supply "{0}" state is unknown.	The last attempt to query the 12V aux Line state via xCAT failed. Therefore, no status can be reported momentarily.	The system will retry to query the 12V aux Line state on a regular basis.	
MS1262	STATE_CHANGE	INFO	12V aux Line of Power Supply "{0}" is healthy.	The last query of 12V aux Line state reports no issue.		
MS1264	STATE_CHANGE	ERROR	12V aux Line of Power Supply "{0}" is unhealthy.	The last query of 12V aux Line state reports failure.		
MS1270	STATE_CHANGE	INFO	The paths to the physical disk "{0}" ({1}) are okay.	All of the expected paths to the physical disk are available.		
MS1271	STATE_CHANGE	WARNING	Not enough active paths to the physical disk "{0}" ({1}).	One or more of the active paths to the physical disk are broken.		Check the cables to the physical disk for errors.
MS1272	STATE_CHANGE	WARNING	Not enough passive paths to the physical disk "{0}" ({1}).	One or more of the passive paths to the physical disk are broken.		Check the cables to the physical disk for errors.
MS1273	STATE_CHANGE	ERROR	No active path to the physical disk "{0}" ({1}).	All the active paths to the physical disk are broken.		Check the cables to the physical disk for errors.
MS1274	STATE_CHANGE	ERROR	No path to the physical disk "{0}" ({1}).	All the paths to the physical disk are broken.		Check the cables to the physical disk for errors.
MS1275	STATE_CHANGE	INFO	Can not determine the expected number of paths to the physical disk "{0}" ({1}).	Can not determine if the number of paths to the physical disk is as expected.		The properties "expectedPaths Active" and "expectedPaths Total" should have been specified in the pdisk stanza at creation time.
MS1301	STATE_CHANGE	INFO	The state of the GUI configuration for the cluster "{0}" is unknown.	The state of the GUI cluster configuration is unknown.		
MS1302	STATE_CHANGE	INFO	The GUI configuration for the cluster "{0}" is correct.	No problems regarding the current configuration of the GUI and the cluster were found.		
MS1303	STATE_CHANGE	INFO	Sudo wrappers were enabled on the cluster and the GUI configuration for the cluster "{0}" is correct.	No problems regarding the current configuration of the GUI and the cluster were found.		

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS1304	STATE_CHANGE	ERROR	Sudo wrappers are enabled on the cluster "{0}", but the GUI is not configured to use Sudo Wrappers.	Sudo wrappers are enabled on the cluster, but the value for GPFS_ADMIN in /usr/lpp/mmfs/gui/conf/gpfsgui.properties was either not set or is still set to root. The value of GPFS_ADMIN should be set to the user name for which sudo wrappers were configured on the cluster.		Make sure that sudo wrappers were correctly configured for a user that is available on the GUI node and all other nodes of the cluster. This user name should be set as the value of the GPFS_ADMIN option in /usr/lpp/mmfs/gui/conf/gpfsgui.properties. After that restart the GUI using <b>systemctl restart gpfsgui</b> .
MS1305	STATE_CHANGE	ERROR	Sudo wrappers are enabled on the cluster "{0}", but there is a misconfiguration regarding the user "{1}" that was set as GPFS_ADMIN in the GUI properties file.	Sudo wrappers are enabled on the cluster, but the user name that was set as GPFS_ADMIN in the GUI properties file at /usr/lpp/mmfs/gui/conf/gpfsgui.properties does not exist on the GUI node.		Make sure that sudo wrappers were correctly configured for a user that is available on the GUI node and all other nodes of the cluster. This user name should be set as the value of the GPFS_ADMIN option in /usr/lpp/mmfs/gui/conf/gpfsgui.properties. After that restart the GUI using <b>systemctl restart gpfsgui</b> .
MS1306	STATE_CHANGE	ERROR	Sudo wrappers are enabled on the cluster "{0}", but the GUI cannot connect to other nodes with the user name "{1}" that was defined as GPFS_ADMIN in the GUI properties file.	When sudo wrappers are configured and enabled on a cluster, the GUI does not execute commands as root, but as the user for which sudo wrappers were configured. This user should be set as GPFS_ADMIN in the GUI properties file at /usr/lpp/mmfs/gui/conf/gpfsgui.properties.		Make sure that sudo wrappers were correctly configured for a user that is available on the GUI node and all other nodes of the cluster. This user name should be set as the value of the GPFS_ADMIN option in /usr/lpp/mmfs/gui/conf/gpfsgui.properties. After that restart the GUI by using the <b>systemctl restart gpfsgui</b> command.
MS1307	STATE_CHANGE	ERROR	The cluster ID of the current cluster "{0}" and the cluster ID in the database do not match ("{1}"). It seems that the cluster was recreated.	When a cluster is deleted and created again, the cluster ID changes, but the GUI's database still references the old cluster ID in many ways.		Clear the GUI's database of the old cluster information by dropping all tables: <pre>psql postgres postgres -c "drop schema fsc cascade"</pre> . Then restart the GUI by using the <b>systemctl restart gpfsgui</b> .

Table 73. GUI events (continued)

Event	Event type	Severity	Message	Description	System action	User action
MS1308	STATE_CHANGE	WARNING	Sudo wrappers are not enabled on the cluster "{0}", but GPFS_ADMIN was set to a non-root user.	Sudo wrappers are not enabled on the cluster, but the value for GPFS_ADMIN in /usr/lpp/mmfs/gui/conf/gpfsgui.properties was set to a non-root user. The value of GPFS_ADMIN should be set to "root" when sudo wrappers are not enabled on the cluster.		Set GPFS_ADMIN in /usr/lpp/mmfs/gui/conf/gpfsgui.properties to "root". Restart the GUI by using the <b>systemctl restart gpfsgui</b> .

## Messages

This topic contains explanations for GPFS error messages.

Messages for IBM Spectrum Scale RAID in the ranges 6027-1850 – 6027-1899 and 6027-3000 – 6027-3099 are documented in *IBM Spectrum Scale RAID: Administration*.

## Message severity tags

GPFS has adopted a message severity tagging convention. This convention applies to some newer messages and to some messages that are being updated and adapted to be more usable by scripts or semi-automated management programs.

A severity tag is a one-character alphabetic code (**A** through **Z**), optionally followed by a colon (:) and a number, and surrounded by an opening and closing bracket ([ ]). For example:

[E] or [E:nnn]

If more than one substring within a message matches this pattern (for example, [A] or [A:nnn]), the severity tag is the first such matching string.

When the severity tag includes a numeric code (*nnn*), this is an error code associated with the message. If this were the only problem encountered by the command, the command return code would be *nnn*.

If a message does not have a severity tag, the message does not conform to this specification. You can determine the message severity by examining the text or any supplemental information provided in the message catalog, or by contacting the IBM Support Center.

Each message severity tag has an assigned priority that can be used to filter the messages that are sent to the error log on Linux. Filtering is controlled with the **mmchconfig** attribute **systemLogLevel**. The default for **systemLogLevel** is **error**, which means GPFS will send all error [E], critical [X], and alert [A] messages to the error log. The values allowed for **systemLogLevel** are: **alert**, **critical**, **error**, **warning**, **notice**, **configuration**, **informational**, **detail**, or **debug**. Additionally, the value **none** can be specified so no messages are sent to the error log.

Alert [A] messages have the highest priority, and debug [B] messages have the lowest priority. If the **systemLogLevel** default of **error** is changed, only messages with the specified severity and all those with a higher priority are sent to the error log. The following table lists the message severity tags in order of priority:

Table 74. Message severity tags ordered by priority

Severity tag	Type of message (systemLogLevel attribute)	Meaning
A	alert	Indicates a problem where action must be taken immediately. Notify the appropriate person to correct the problem.
X	critical	Indicates a critical condition that should be corrected immediately. The system discovered an internal inconsistency of some kind. Command execution might be halted or the system might attempt to continue despite the inconsistency. Report these errors to the IBM Support Center.
E	error	Indicates an error condition. Command execution might or might not continue, but this error was likely caused by a persistent condition and will remain until corrected by some other program or administrative action. For example, a command operating on a single file or other GPFS object might terminate upon encountering any condition of severity E. As another example, a command operating on a list of files, finding that one of the files has permission bits set that disallow the operation, might continue to operate on all other files within the specified list of files.
W	warning	Indicates a problem, but command execution continues. The problem can be a transient inconsistency. It can be that the command has skipped some operations on some objects, or is reporting an irregularity that could be of interest. For example, if a multipass command operating on many files discovers during its second pass that a file that was present during the first pass is no longer present, the file might have been removed by another command or program.
N	notice	Indicates a normal but significant condition. These events are unusual but not error conditions, and might be summarized in an email to developers or administrators for spotting potential problems. No immediate action is required.
C	configuration	Indicates a configuration change; such as, creating a file system or removing a node from the cluster.
I	informational	Indicates normal operation. This message by itself indicates that nothing is wrong; no action is required.
D	detail	Indicates verbose operational messages; no is action required.
B	debug	Indicates debug-level messages that are useful to application developers for debugging purposes. This information is not useful during operations.

---

**6027-000**     **Attention: A disk being removed reduces the number of failure groups to *nFailureGroups*, which is below the number required for replication: *nReplicas*.**

**Explanation:** Replication cannot protect data against disk failures when there are insufficient failure groups.

**User response:** Add more disks in new failure groups to the file system or accept the risk of data loss.

---

**6027-300 [N]**     **mmfsd ready**

**Explanation:** The **mmfsd** server is up and running.

**User response:** None. Informational message only.

---

**6027-301**     **File *fileName* could not be run with error *errno*.**

**Explanation:** The named shell script could not be executed. This message is followed by the error string that is returned by the **exec**.

**User response:** Check file existence and access permissions.

---

**6027-302 [E]**     **Could not execute *script***

**Explanation:** The **verifyGpfsReady=yes** configuration attribute is set, but the **/var/mmfs/etc/gpfsready** script could not be executed.

**User response:** Make sure **/var/mmfs/etc/gpfsready** exists and is executable, or disable the

verifyGpfsReady option via `mmchconfig verifyGpfsReady=no`.

---

6027-303 [N] *script killed by signal signal*

**Explanation:** The `verifyGpfsReady=yes` configuration attribute is set and `/var/mmfs/etc/gpfsready` script did not complete successfully.

**User response:** Make sure `/var/mmfs/etc/gpfsready` completes and returns a zero exit status, or disable the `verifyGpfsReady` option via `mmchconfig verifyGpfsReady=no`.

---

6027-304 [W] *script ended abnormally*

**Explanation:** The `verifyGpfsReady=yes` configuration attribute is set and `/var/mmfs/etc/gpfsready` script did not complete successfully.

**User response:** Make sure `/var/mmfs/etc/gpfsready` completes and returns a zero exit status, or disable the `verifyGpfsReady` option via `mmchconfig verifyGpfsReady=no`.

---

6027-305 [N] *script failed with exit code code*

**Explanation:** The `verifyGpfsReady=yes` configuration attribute is set and `/var/mmfs/etc/gpfsready` script did not complete successfully.

**User response:** Make sure `/var/mmfs/etc/gpfsready` completes and returns a zero exit status, or disable the `verifyGpfsReady` option via `mmchconfig verifyGpfsReady=no`.

---

6027-306 [E] **Could not initialize inter-node communication**

**Explanation:** The GPFS daemon was unable to initialize the communications required to proceed.

**User response:** User action depends on the return code shown in the accompanying message (`/usr/include/errno.h`). The communications failure that caused the failure must be corrected. One possibility is an `rc` value of 67, indicating that the required port is unavailable. This may mean that a previous version of the `mmfs` daemon is still running. Killing that daemon may resolve the problem.

---

6027-310 [I] *command initializing. {Version  
versionName: Built date time}*

**Explanation:** The `mmfsd` server has started execution.

**User response:** None. Informational message only.

---

6027-311 [N] *programName is shutting down.*

**Explanation:** The stated program is about to terminate.

**User response:** None. Informational message only.

---

6027-312 [E] **Unknown trace class 'traceClass'.**

**Explanation:** The trace class is not recognized.

**User response:** Specify a valid trace class.

---

6027-313 [X] **Cannot open configuration file *fileName*.**

**Explanation:** The configuration file could not be opened.

**User response:** The configuration file is `/var/mmfs/gen/mmfs.cfg`. Verify that this file and `/var/mmfs/gen/mmsdrfs` exist in your system.

---

6027-314 [E] *command requires SuperuserName authority to execute.*

**Explanation:** The `mmfsd` server was started by a user without superuser authority.

**User response:** Log on as a superuser and reissue the command.

---

6027-315 [E] **Bad config file entry in *fileName*, line *number*.**

**Explanation:** The configuration file has an incorrect entry.

**User response:** Fix the syntax error in the configuration file. Verify that you are not using a configuration file that was created on a release of GPFS subsequent to the one that you are currently running.

---

6027-316 [E] **Unknown config parameter "*parameter*" in *fileName*, line *number*.**

**Explanation:** There is an unknown parameter in the configuration file.

**User response:** Fix the syntax error in the configuration file. Verify that you are not using a configuration file that was created on a release of GPFS subsequent to the one you are currently running.

---

6027-317 [A] **Old server with PID *pid* still running.**

**Explanation:** An old copy of `mmfsd` is still running.

**User response:** This message would occur only if the user bypasses the SRC. The normal message in this case would be an SRC message stating that multiple instances are not allowed. If it occurs, stop the previous instance and use the SRC commands to restart the daemon.

---

**6027-318 [E] Watchdog: Some process appears stuck; stopped the daemon process.**

**Explanation:** A high priority process got into a loop.

**User response:** Stop the old instance of the `mmfs` server, then restart it.

---

**6027-319 Could not create shared segment**

**Explanation:** The shared segment could not be created.

**User response:** This is an error from the AIX operating system. Check the accompanying error indications from AIX.

---

**6027-320 Could not map shared segment**

**Explanation:** The shared segment could not be attached.

**User response:** This is an error from the AIX operating system. Check the accompanying error indications from AIX.

---

**6027-321 Shared segment mapped at wrong address (is *value*, should be *value*).**

**Explanation:** The shared segment did not get mapped to the expected address.

**User response:** Contact the IBM Support Center.

---

**6027-322 Could not map shared segment in kernel extension**

**Explanation:** The shared segment could not be mapped in the kernel.

**User response:** If an `EINVAL` error message is displayed, the kernel extension could not use the shared segment because it did not have the correct GPFS version number. Unload the kernel extension and restart the GPFS daemon.

---

**6027-323 [A] Error unmapping shared segment.**

**Explanation:** The shared segment could not be detached.

**User response:** Check reason given by error message.

---

**6027-324 Could not create message queue for main process**

**Explanation:** The message queue for the main process could not be created. This is probably an operating system error.

**User response:** Contact the IBM Support Center.

---



---

**6027-328 [W] Value '*value*' for '*parameter*' is out of range in *fileName*. Valid values are *value* through *value*. *value* used.**

**Explanation:** An error was found in the `/var/mmfs/gen/mmfs.cfg` file.

**User response:** Check the `/var/mmfs/gen/mmfs.cfg` file.

---

**6027-329 Cannot pin the main shared segment: *name***

**Explanation:** Trying to pin the shared segment during initialization.

**User response:** Check the `mmfs.cfg` file. The `pagepool` size may be too large. It cannot be more than 80% of real memory. If a previous `mmfsd` crashed, check for processes that begin with the name `mmfs` that may be holding on to an old pinned shared segment. Issue `mmchconfig` command to change the `pagepool` size.

---

**6027-334 [E] Error initializing internal communications.**

**Explanation:** The mailbox system used by the daemon for communication with the kernel cannot be initialized.

**User response:** Increase the size of available memory using the `mmchconfig` command.

---

**6027-335 [E] Configuration error: check *fileName*.**

**Explanation:** A configuration error is found.

**User response:** Check the `mmfs.cfg` file and other error messages.

---

**6027-336 [E] Value '*value*' for configuration parameter '*parameter*' is not valid. Check *fileName*.**

**Explanation:** A configuration error was found.

**User response:** Check the `mmfs.cfg` file.

---

**6027-337 [N] Waiting for resources to be reclaimed before exiting.**

**Explanation:** The `mmfsd` daemon is attempting to terminate, but cannot because data structures in the daemon shared segment may still be referenced by kernel code. This message may be accompanied by other messages that show which disks still have I/O in progress.

**User response:** None. Informational message only.

---

---

**6027-338 [N] Waiting for *number* user(s) of shared segment to release it.**

**Explanation:** The `mmfsd` daemon is attempting to terminate, but cannot because some process is holding the shared segment while in a system call. The message will repeat every 30 seconds until the count drops to zero.

**User response:** Find the process that is not responding, and find a way to get it out of its system call.

---

**6027-339 [E] Nonnumeric trace value '*value*' after class '*class*'.**

**Explanation:** The specified trace value is not recognized.

**User response:** Specify a valid trace integer value.

---

**6027-340 Child process *file* failed to start due to error *rc: errStr*.**

**Explanation:** A failure occurred when GPFS attempted to start a program.

**User response:** If the program was a user exit script, verify the script file exists and has appropriate permissions assigned. If the program was not a user exit script, then this is an internal GPFS error or the GPFS installation was altered.

---

**6027-341 [D] Node *nodeName* is incompatible because its maximum compatible version (*number*) is less than the version of this node (*number*). [*valuevalue*]**

**Explanation:** The GPFS daemon tried to make a connection with another GPFS daemon. However, the other daemon is not compatible. Its maximum compatible version is less than the version of the daemon running on this node. The numbers in square brackets are for use by the IBM Support Center.

**User response:** Verify your GPFS daemon version.

---

**6027-342 [E] Node *nodeName* is incompatible because its minimum compatible version is greater than the version of this node (*number*). [*valuevalue*]**

**Explanation:** The GPFS daemon tried to make a connection with another GPFS daemon. However, the other daemon is not compatible. Its minimum compatible version is greater than the version of the daemon running on this node. The numbers in square brackets are for use by the IBM Support Center.

**User response:** Verify your GPFS daemon version.

---

**6027-343 [E] Node *nodeName* is incompatible because its version (*number*) is less than the minimum compatible version of this node (*number*). [*valuevalue*]**

**Explanation:** The GPFS daemon tried to make a connection with another GPFS daemon. However, the other daemon is not compatible. Its version is less than the minimum compatible version of the daemon running on this node. The numbers in square brackets are for use by the IBM Support Center.

**User response:** Verify your GPFS daemon version.

---

**6027-344 [E] Node *nodeName* is incompatible because its version is greater than the maximum compatible version of this node (*number*). [*valuevalue*]**

**Explanation:** The GPFS daemon tried to make a connection with another GPFS daemon. However, the other daemon is not compatible. Its version is greater than the maximum compatible version of the daemon running on this node. The numbers in square brackets are for use by the IBM Support Center.

**User response:** Verify your GPFS daemon version.

---

**6027-345 Network error on *ipAddress*, check connectivity.**

**Explanation:** A TCP error has caused GPFS to exit due to a bad return code from an error. Exiting allows recovery to proceed on another node and resources are not tied up on this node.

**User response:** Follow network problem determination procedures.

---

**6027-346 [E] Incompatible daemon version. My version = *number*, repl.my\_version = *number***

**Explanation:** The GPFS daemon tried to make a connection with another GPFS daemon. However, the other GPFS daemon is not the same version and it sent a reply indicating its version number is incompatible.

**User response:** Verify your GPFS daemon version.

---

**6027-347 [E] Remote host *ipAddress* refused connection because IP address *ipAddress* was not in the node list file**

**Explanation:** The GPFS daemon tried to make a connection with another GPFS daemon. However, the other GPFS daemon sent a reply indicating it did not recognize the IP address of the connector.

**User response:** Add the IP address of the local host to the node list file on the remote host.

---

---

**6027-348 [E] Bad "subnets" configuration: invalid subnet *ipAddress*".**

**Explanation:** A subnet specified by the **subnets** configuration parameter could not be parsed.

**User response:** Run the **mmlsconfig** command and check the value of the **subnets** parameter. Each subnet must be specified as a dotted-decimal IP address. Run the **mmchconfig subnets** command to correct the value.

---

**6027-349 [E] Bad "subnets" configuration: invalid cluster name pattern *clusterNamePattern*".**

**Explanation:** A cluster name pattern specified by the **subnets** configuration parameter could not be parsed.

**User response:** Run the **mmlsconfig** command and check the value of the **subnets** parameter. The optional cluster name pattern following subnet address must be a shell-style pattern allowing '\*', '/' and '['...]' as wild cards. Run the **mmchconfig subnets** command to correct the value.

---

**6027-350 [E] Bad "subnets" configuration: primary IP address *ipAddress* is on a private subnet. Use a public IP address instead.**

**Explanation:** GPFS is configured to allow multiple IP addresses per node (**subnets** configuration parameter), but the primary IP address of the node (the one specified when the cluster was created or when the node was added to the cluster) was found to be on a private subnet. If multiple IP addresses are used, the primary address must be a public IP address.

**User response:** Remove the node from the cluster; then add it back using a public IP address.

---

**6027-358 Communication with **mmspsecserver** through socket *name* failed, err *value*: *errorString*, **msgType** *messageType*.**

**Explanation:** Communication failed between **spsecClient** (the daemon) and **spsecServer**.

**User response:** Verify both the communication socket and the **mmspsecserver** process.

---

**6027-359 The **mmspsecserver** process is shutting down. Reason: *explanation*.**

**Explanation:** The **mmspsecserver** process received a signal from the **mmfsd** daemon or encountered an error on execution.

**User response:** Verify the reason for shutdown.

---

**6027-360 Disk *name* must be removed from the */etc/filesystems* stanza before it can be deleted.**

**Explanation:** A disk being deleted is found listed in the **disks=** list for a file system.

**User response:** Remove the disk from list.

---

**6027-361 [E] Local access to *disk* failed with EIO, switching to access the disk remotely.**

**Explanation:** Local access to the disk failed. To avoid unmounting of the file system, the disk will now be accessed remotely.

**User response:** Wait until work continuing on the local node completes. Then determine why local access to the disk failed, correct the problem and restart the daemon. This will cause GPFS to begin accessing the disk locally again.

---

**6027-362 Attention: No disks were deleted, but some data was migrated. The file system may no longer be properly balanced.**

**Explanation:** The **mmdeldisk** command did not complete migrating data off the disks being deleted. The disks were restored to normal **ready**, status, but the migration has left the file system unbalanced. This may be caused by having too many disks unavailable or insufficient space to migrate all of the data to other disks.

**User response:** Check disk availability and space requirements. Determine the reason that caused the command to end before successfully completing the migration and disk deletion. Reissue the **mmdeldisk** command.

---

**6027-363 I/O error writing disk descriptor for disk *name*.**

**Explanation:** An I/O error occurred when the **mmaddisk** command was writing a disk descriptor on a disk. This could have been caused by either a configuration error or an error in the path to the disk.

**User response:** Determine the reason the disk is inaccessible for writing and reissue the **mmaddisk** command.

---

**6027-364 Error processing disks.**

**Explanation:** An error occurred when the **mmaddisk** command was reading disks in the file system.

**User response:** Determine the reason why the disks are inaccessible for reading, then reissue the **mmaddisk** command.



---

**6027-365 [I] Rediscovered local access to *disk*.**

**Explanation:** Rediscovered local access to disk, which failed earlier with **EIO**. For good performance, the disk will now be accessed locally.

**User response:** Wait until work continuing on the local node completes. This will cause GPFS to begin accessing the disk locally again.

---

**6027-369 I/O error writing file system descriptor for disk *name*.**

**Explanation:** **mmadddisk** detected an I/O error while writing a file system descriptor on a disk.

**User response:** Determine the reason the disk is inaccessible for writing and reissue the **mmadddisk** command.

---

**6027-370 mmdeldisk completed.**

**Explanation:** The **mmdeldisk** command has completed.

**User response:** None. Informational message only.

---

**6027-371 Cannot delete all disks in the file system**

**Explanation:** An attempt was made to delete all the disks in a file system.

**User response:** Either reduce the number of disks to be deleted or use the **mmdelfs** command to delete the file system.

---

**6027-372 Replacement disk must be in the same failure group as the disk being replaced.**

**Explanation:** An improper failure group was specified for **mmrpldisk**.

**User response:** Specify a failure group in the disk descriptor for the replacement disk that is the same as the failure group of the disk being replaced.

---

**6027-373 Disk *diskName* is being replaced, so status of disk *diskName* must be replacement.**

**Explanation:** The **mmrpldisk** command failed when retrying a replace operation because the new disk does not have the correct status.

**User response:** Issue the **mmlsdisk** command to display disk status. Then either issue the **mmchdisk** command to change the status of the disk to **replacement** or specify a new disk that has a status of **replacement**.

---

**6027-374 Disk *name* may not be replaced.**

**Explanation:** A disk being replaced with **mmrpldisk** does not have a status of **ready** or **suspended**.

**User response:** Use the **mmlsdisk** command to display disk status. Issue the **mmchdisk** command to change the status of the disk to be replaced to either **ready** or **suspended**.

---

**6027-375 Disk name *diskName* already in file system.**

**Explanation:** The replacement disk name specified in the **mmrpldisk** command already exists in the file system.

**User response:** Specify a different disk as the replacement disk.

---

**6027-376 Previous replace command must be completed before starting a new one.**

**Explanation:** The **mmrpldisk** command failed because the status of other disks shows that a replace command did not complete.

**User response:** Issue the **mmlsdisk** command to display disk status. Retry the failed **mmrpldisk** command or issue the **mmchdisk** command to change the status of the disks that have a status of **replacing** or **replacement**.

---

**6027-377 Cannot replace a disk that is in use.**

**Explanation:** Attempting to replace a disk in place, but the disk specified in the **mmrpldisk** command is still available for use.

**User response:** Use the **mmchdisk** command to stop GPFS's use of the disk.

---

**6027-378 [I] I/O still in progress near sector *number* on disk *diskName*.**

**Explanation:** The **mmfsd** daemon is attempting to terminate, but cannot because data structures in the daemon shared segment may still be referenced by kernel code. In particular, the daemon has started an I/O that has not yet completed. It is unsafe for the daemon to terminate until the I/O completes, because of asynchronous activity in the device driver that will access data structures belonging to the daemon.

**User response:** Either wait for the I/O operation to time out, or issue a device-dependent command to terminate the I/O.

---

---

**6027-379** Could not invalidate disk(s).

**Explanation:** Trying to delete a disk and it could not be written to in order to invalidate its contents.

**User response:** No action needed if removing that disk permanently. However, if the disk is ever to be used again, the **-v** flag must be specified with a value of **no** when using either the **mmcrfs** or **mmaddisk** command.

---

**6027-380** Disk name missing from disk descriptor list entry *name*.

**Explanation:** When parsing disk lists, no disks were named.

**User response:** Check the argument list of the command.

---

**6027-382** Value *value* for the 'sector size' option for disk *disk* is not a multiple of *value*.

**Explanation:** When parsing disk lists, the sector size given is not a multiple of the default sector size.

**User response:** Specify a correct sector size.

---

**6027-383** Disk name *name* appears more than once.

**Explanation:** When parsing disk lists, a duplicate name is found.

**User response:** Remove the duplicate name.

---

**6027-384** Disk name *name* already in file system.

**Explanation:** When parsing disk lists, a disk name already exists in the file system.

**User response:** Rename or remove the duplicate disk.

---

**6027-385** Value *value* for the 'sector size' option for disk *name* is out of range. Valid values are *number* through *number*.

**Explanation:** When parsing disk lists, the sector size given is not valid.

**User response:** Specify a correct sector size.

---

**6027-386** Value *value* for the 'sector size' option for disk *name* is invalid.

**Explanation:** When parsing disk lists, the sector size given is not valid.

**User response:** Specify a correct sector size.

---

**6027-387** Value *value* for the 'failure group' option for disk *name* is out of range. Valid values are *number* through *number*.

**Explanation:** When parsing disk lists, the failure group given is not valid.

**User response:** Specify a correct failure group.

---

**6027-388** Value *value* for the 'failure group' option for disk *name* is invalid.

**Explanation:** When parsing disk lists, the failure group given is not valid.

**User response:** Specify a correct failure group.

---

**6027-389** Value *value* for the 'has metadata' option for disk *name* is out of range. Valid values are *number* through *number*.

**Explanation:** When parsing disk lists, the 'has metadata' value given is not valid.

**User response:** Specify a correct 'has metadata' value.

---

**6027-390** Value *value* for the 'has metadata' option for disk *name* is invalid.

**Explanation:** When parsing disk lists, the 'has metadata' value given is not valid.

**User response:** Specify a correct 'has metadata' value.

---

**6027-391** Value *value* for the 'has data' option for disk *name* is out of range. Valid values are *number* through *number*.

**Explanation:** When parsing disk lists, the 'has data' value given is not valid.

**User response:** Specify a correct 'has data' value.

---

**6027-392** Value *value* for the 'has data' option for disk *name* is invalid.

**Explanation:** When parsing disk lists, the 'has data' value given is not valid.

**User response:** Specify a correct 'has data' value.

---

**6027-393** Either the 'has data' option or the 'has metadata' option must be '1' for disk *diskName*.

**Explanation:** When parsing disk lists the 'has data' or 'has metadata' value given is not valid.

**User response:** Specify a correct 'has data' or 'has metadata' value.

---

---

**6027-394**    **Too many disks specified for file system. Maximum = *number*.**

**Explanation:** Too many disk names were passed in the disk descriptor list.

**User response:** Check the disk descriptor list or the file containing the list.

---

**6027-399**    **Not enough items in disk descriptor list entry, need *fields*.**

**Explanation:** When parsing a disk descriptor, not enough fields were specified for one disk.

**User response:** Correct the disk descriptor to use the correct disk descriptor syntax.

---

**6027-416**    **Incompatible file system descriptor version or not formatted.**

**Explanation:** Possible reasons for the error are:

1. A file system descriptor version that is not valid was encountered.
2. No file system descriptor can be found.
3. Disks are not correctly defined on all active nodes.
4. Disks, logical volumes, network shared disks, or virtual shared disks were incorrectly re-configured after creating a file system.

**User response:** Verify:

1. The disks are correctly defined on all nodes.
  2. The paths to the disks are correctly defined and operational.
- 

**6027-417**    **Bad file system descriptor.**

**Explanation:** A file system descriptor that is not valid was encountered.

**User response:** Verify:

1. The disks are correctly defined on all nodes.
  2. The paths to the disks are correctly defined and operational.
- 

**6027-418**    **Inconsistent file system quorum. readQuorum=*value* writeQuorum=*value* quorumSize=*value*.**

**Explanation:** A file system descriptor that is not valid was encountered.

**User response:** Start any disks that have been stopped by the **mmchdisk** command or by hardware failures. If the problem persists, run offline **mmfsck**.

---

**6027-419**    **Failed to read a file system descriptor.**

**Explanation:** Not enough valid replicas of the file system descriptor could be read from the file system.

**User response:** Start any disks that have been stopped by the **mmchdisk** command or by hardware failures. Verify that paths to all disks are correctly defined and operational.

---

**6027-420**    **Inode size must be greater than zero.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-421**    **Inode size must be a multiple of logical sector size.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-422**    **Inode size must be at least as large as the logical sector size.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-423**    **Minimum fragment size must be a multiple of logical sector size.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-424**    **Minimum fragment size must be greater than zero.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-425**    **File system block size of *blockSize* is larger than *maxblocksize* parameter.**

**Explanation:** An attempt is being made to mount a file system whose block size is larger than the **maxblocksize** parameter as set by **mmchconfig**.

**User response:** Use the **mmchconfig maxblocksize=*xxx*** command to increase the maximum allowable block size.

---

---

**6027-426**     **Warning: mount detected unavailable disks. Use `mmlsdisk fileSystem` to see details.**

**Explanation:** The `mount` command detected that some disks needed for the file system are unavailable.

**User response:** Without file system replication enabled, the mount will fail. If it has replication, the mount may succeed depending on which disks are unavailable. Use `mmlsdisk` to see details of the disk status.

---

**6027-427**     **Indirect block size must be at least as large as the minimum fragment size.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-428**     **Indirect block size must be a multiple of the minimum fragment size.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-429**     **Indirect block size must be less than full data block size.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-430**     **Default metadata replicas must be less than or equal to default maximum number of metadata replicas.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-431**     **Default data replicas must be less than or equal to default maximum number of data replicas.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-432**     **Default maximum metadata replicas must be less than or equal to `value`.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-433**     **Default maximum data replicas must be less than or equal to `value`.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-434**     **Indirect blocks must be at least as big as inodes.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-435 [N]**     **The file system descriptor quorum has been overridden.**

**Explanation:** The `mmfsctl exclude` command was previously issued to override the file system descriptor quorum after a disaster.

**User response:** None. Informational message only.

---

**6027-438**     **Duplicate disk name `name`.**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-439**     **Disk name sector size `value` does not match sector size `value` of other disk(s).**

**Explanation:** An internal consistency check has found a problem with file system parameters.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-441**     **Unable to open disk '`name`' on node `nodeName`.**

**Explanation:** A disk name that is not valid was specified in a GPFS disk command.

**User response:** Correct the parameters of the executing GPFS disk command.

---

---

**6027-445** Value for option '-m' cannot exceed the number of metadata failure groups.

**Explanation:** The current number of replicas of metadata cannot be larger than the number of failure groups that are enabled to hold metadata.

**User response:** Use a smaller value for -m on the **mmchfs** command, or increase the number of failure groups by adding disks to the file system.

---

**6027-446** Value for option '-r' cannot exceed the number of data failure groups.

**Explanation:** The current number of replicas of data cannot be larger than the number of failure groups that are enabled to hold data.

**User response:** Use a smaller value for -r on the **mmchfs** command, or increase the number of failure groups by adding disks to the file system.

---

**6027-451** No disks= list found in mount options.

**Explanation:** No 'disks=' clause found in the mount options list when opening a file system.

**User response:** Check the operating system's file system database and local **mmsdrfs** file for this file system.

---

**6027-452** No disks found in disks= list.

**Explanation:** No disks listed when opening a file system.

**User response:** Check the operating system's file system database and local **mmsdrfs** file for this file system.

---

**6027-453** No disk name found in a clause of the list.

**Explanation:** No disk name found in a clause of the **disks=** list.

**User response:** Check the operating system's file system database and local **mmsdrfs** file for this file system.

---

**6027-461** Unable to find *name* device.

**Explanation:** Self explanatory.

**User response:** There must be a **/dev/sgname** special device defined. Check the error code. This could indicate a configuration error in the specification of disks, logical volumes, network shared disks, or virtual shared disks.

---

**6027-462** *name* must be a char or block special device.

**Explanation:** Opening a file system.

**User response:** There must be a **/dev/sgname** special device defined. This could indicate a configuration error in the specification of disks, logical volumes, network shared disks, or virtual shared disks.

---

**6027-463** SubblocksPerFullBlock was not 32.

**Explanation:** The value of the SubblocksPerFullBlock variable was not 32. This situation should never exist, and indicates an internal error.

**User response:** Record the above information and contact the IBM Support Center.

---

**6027-465** The average file size must be at least as large as the minimum fragment size.

**Explanation:** When parsing the command line of **tscrfs**, it was discovered that the average file size is smaller than the minimum fragment size.

**User response:** Correct the indicated command parameters.

---

**6027-468** Disk name listed in *fileName* or local **mmsdrfs** file, not found in device *name*.  
Run: **mmcommon recoverfs** *name*.

**Explanation:** Tried to access a file system but the disks listed in the operating system's file system database or the local **mmsdrfs** file for the device do not exist in the file system.

**User response:** Check the configuration and availability of disks. Run the **mmcommon recoverfs** *device* command. If this does not resolve the problem, configuration data in the SDR may be incorrect. If no user modifications have been made to the SDR, contact the IBM Support Center. If user modifications have been made, correct these modifications.

---

**6027-469** File system *name* does not match *descriptor*.

**Explanation:** The file system name found in the descriptor on disk does not match the corresponding device name in **/etc/filesystems**.

**User response:** Check the operating system's file system database.

---

**6027-470** Disk *name* may still belong to file system *filesystem*. Created on **IPandTime**.

**Explanation:** The disk being added by the **mmcrfs**, **mmadddisk**, or **mmrpldisk** command appears to still belong to some file system.

---

**User response:** Verify that the disks you are adding do not belong to an active file system, and use the **-v no** option to bypass this check. Use this option only if you are sure that no other file system has this disk configured because you may cause data corruption in both file systems if this is not the case.

---

**6027-471 Disk *diskName*: Incompatible file system descriptor version or not formatted.**

**Explanation:** Possible reasons for the error are:

1. A file system descriptor version that is not valid was encountered.
2. No file system descriptor can be found.
3. Disks are not correctly defined on all active nodes.
4. Disks, logical volumes, network shared disks, or virtual shared disks were incorrectly reconfigured after creating a file system.

**User response:** Verify:

1. The disks are correctly defined on all nodes.
2. The paths to the disks are correctly defined and operative.

---

**6027-472 [E] File system format version *versionString* is not supported.**

**Explanation:** The current file system format version is not supported.

**User response:** Verify:

1. The disks are correctly defined on all nodes.
2. The paths to the disks are correctly defined and operative.

---

**6027-473 [X] File System *fileSystem* unmounted by the system with return code *value* reason code *value***

**Explanation:** Console log entry caused by a forced unmount due to disk or communication failure.

**User response:** Correct the underlying problem and remount the file system.

---

**6027-474 [X] Recovery Log I/O failed, unmounting file system *fileSystem***

**Explanation:** I/O to the recovery log failed.

**User response:** Check the paths to all disks making up the file system. Run the **mmfsdisk** command to determine if GPFS has declared any disks unavailable. Repair any paths to disks that have failed. Remount the file system.

---

**6027-475 The option '--inode-limit' is not enabled. Use option '-V' to enable most recent features.**

**Explanation:** **mmchfs --inode-limit** is not enabled under the current file system format version.

**User response:** Run **mmchfs -V**, this will change the file system format to the latest format supported.

---

**6027-476 Restricted mount using only available file system descriptor.**

**Explanation:** Fewer than the necessary number of file system descriptors were successfully read. Using the best available descriptor to allow the restricted mount to continue.

**User response:** Informational message only.

---

**6027-477 The option -z is not enabled. Use the -V option to enable most recent features.**

**Explanation:** The file system format version does not support the **-z** option on the **mmchfs** command.

**User response:** Change the file system format version by issuing **mmchfs -V**.

---

**6027-478 The option -z could not be changed. *fileSystem* is still in use.**

**Explanation:** The file system is still mounted or another GPFS administration command (**mm...**) is running against the file system.

**User response:** Unmount the file system if it is mounted, and wait for any command that is running to complete before reissuing the **mmchfs -z** command.

---

**6027-479 [N] Mount of *fsName* was blocked by *fileName***

**Explanation:** The internal or external mount of the file system was blocked by the existence of the specified file.

**User response:** If the file system needs to be mounted, remove the specified file.

---

**6027-480 Cannot enable DMAPi in a file system with existing snapshots.**

**Explanation:** The user is not allowed to enable DMAPi for a file system with existing snapshots.

**User response:** Delete all existing snapshots in the file system and repeat the **mmchfs** command.

---

---

**6027-481 [E] Remount failed for mountid *id*:**  
*errnoDescription*

**Explanation:** `mmfsd` restarted and tried to remount any file systems that the VFS layer thinks are still mounted.

**User response:** Check the errors displayed and the `errno` description.

---

**6027-482 [E] Remount failed for device *name*:**  
*errnoDescription*

**Explanation:** `mmfsd` restarted and tried to remount any file systems that the VFS layer thinks are still mounted.

**User response:** Check the errors displayed and the `errno` description.

---

**6027-483 [N] Remounted *name***

**Explanation:** `mmfsd` restarted and remounted the specified file system because it was in the kernel's list of previously mounted file systems.

**User response:** Informational message only.

---

**6027-484 Remount failed for *device* after daemon restart.**

**Explanation:** A remount failed after daemon restart. This ordinarily occurs because one or more disks are unavailable. Other possibilities include loss of connectivity to one or more disks.

**User response:** Issue the `mmlsdisk` command and check for **down** disks. Issue the `mmchdisk` command to start any **down** disks, then remount the file system. If there is another problem with the disks or the connections to the disks, take necessary corrective actions and remount the file system.

---

**6027-485 Perform `mmchdisk` for any disk failures and re-mount.**

**Explanation:** Occurs in conjunction with 6027-484.

**User response:** Follow the *User response* for 6027-484.

---

**6027-486 No local device specified for *fileSystemName* in *clusterName*.**

**Explanation:** While attempting to mount a remote file system from another cluster, GPFS was unable to determine the local device name for this file system.

**User response:** There must be a `/dev/sgname` special device defined. Check the error code. This is probably a configuration error in the specification of a remote file system. Run `mmremotefs show` to check that the remote file system is properly configured.

---

**6027-487 Failed to write the file system descriptor to disk *diskName*.**

**Explanation:** An error occurred when `mmfsctl include` was writing a copy of the file system descriptor to one of the disks specified on the command line. This could have been caused by a failure of the corresponding disk device, or an error in the path to the disk.

**User response:** Verify that the disks are correctly defined on all nodes. Verify that paths to all disks are correctly defined and operational.

---

**6027-488 Error opening the exclusion disk file *fileName*.**

**Explanation:** Unable to retrieve the list of excluded disks from an internal configuration file.

**User response:** Ensure that GPFS executable files have been properly installed on all nodes. Perform required configuration steps prior to starting GPFS.

---

**6027-489 Attention: The desired replication factor exceeds the number of available *dataOrMetadata* failure groups. This is allowed, but the files will not be replicated and will therefore be at risk.**

**Explanation:** You specified a number of replicas that exceeds the number of failure groups available.

**User response:** Reissue the command with a smaller replication factor, or increase the number of failure groups.

---

**6027-490 [N] The descriptor replica on disk *diskName* has been excluded.**

**Explanation:** The file system descriptor quorum has been overridden and, as a result, the specified disk was excluded from all operations on the file system descriptor quorum.

**User response:** None. Informational message only.

---

**6027-492 The file system is already at file system version *number***

**Explanation:** The user tried to upgrade the file system format using `mmchfs -V --version=v`, but the specified version is smaller than the current version of the file system.

**User response:** Specify a different value for the `--version` option.

---

**6027-493 File system version *number* is not supported on *nodeName* nodes in the cluster.**

**Explanation:** The user tried to upgrade the file system format using `mmchfs -V`, but some nodes in the local

cluster are still running an older GPFS release that does support the new format version.

**User response:** Install a newer version of GPFS on those nodes.

**6027-494**     **File system version *number* is not supported on the following *nodeName* remote nodes mounting the file system:**

**Explanation:** The user tried to upgrade the file system format using `mmchfs -V`, but the file system is still mounted on some nodes in remote clusters that do not support the new format version.

**User response:** Unmount the file system on the nodes that do not support the new format version.

**6027-495**     **You have requested that the file system be upgraded to version *number*. This will enable new functionality but will prevent you from using the file system with earlier releases of GPFS. Do you want to continue?**

**Explanation:** Verification request in response to the `mmchfs -V full` command. This is a request to upgrade the file system and activate functions that are incompatible with a previous release of GPFS.

**User response:** Enter `yes` if you want the conversion to take place.

**6027-496**     **You have requested that the file system version for local access be upgraded to version *number*. This will enable some new functionality but will prevent local nodes from using the file system with earlier releases of GPFS. Remote nodes are not affected by this change. Do you want to continue?**

**Explanation:** Verification request in response to the `mmchfs -V` command. This is a request to upgrade the file system and activate functions that are incompatible with a previous release of GPFS.

**User response:** Enter `yes` if you want the conversion to take place.

**6027-497**     **The file system has already been upgraded to *number* using `-V full`. It is not possible to revert back.**

**Explanation:** The user tried to upgrade the file system format using `mmchfs -V compat`, but the file system has already been fully upgraded.

**User response:** Informational message only.

**6027-498**     **Incompatible file system format. Only file systems formatted with GPFS 3.2.1.5 or later can be mounted on this platform.**

**Explanation:** A user running GPFS on Microsoft Windows tried to mount a file system that was formatted with a version of GPFS that did not have Windows support.

**User response:** Create a new file system using current GPFS code.

**6027-499 [X]**     **An unexpected Device Mapper path *dmDevice (nsdId)* has been detected. The new path does not have a Persistent Reserve set up. File system *fileSystem* will be internally unmounted.**

**Explanation:** A new device mapper path is detected or a previously failed path is activated after the local device discovery has finished. This path lacks a Persistent Reserve, and can not be used. All device paths must be active at mount time.

**User response:** Check the paths to all disks making up the file system. Repair any paths to disks which have failed. Remount the file system.

**6027-500**     *name* loaded and configured.

**Explanation:** The kernel extension was loaded and configured.

**User response:** None. Informational message only.

**6027-501**     *name:module moduleName* unloaded.

**Explanation:** The kernel extension was unloaded.

**User response:** None. Informational message only.

**6027-502**     **Incorrect parameter: *name*.**

**Explanation:** `mmfsmnhelp` was called with an incorrect parameter.

**User response:** Contact the IBM Support Center.

**6027-504**     **Not enough memory to allocate internal data structure.**

**Explanation:** Self explanatory.

**User response:** Increase `ulimit` or paging space

**6027-505**     **Internal error, aborting.**

**Explanation:** Self explanatory.

**User response:** Contact the IBM Support Center.



---

**6027-506**     *program: loadFile is already loaded at address.*

**Explanation:** The program was already loaded at the address displayed.

**User response:** None. Informational message only.

---

**6027-507**     *program: loadFile is not loaded.*

**Explanation:** The program could not be loaded.

**User response:** None. Informational message only.

---

**6027-510**     **Cannot mount *fileSystem* on *mountPoint*:**  
*errorString*

**Explanation:** There was an error mounting the GPFS file system.

**User response:** Determine action indicated by the error messages and error log entries. Errors in the disk path often cause this problem.

---

**6027-511**     **Cannot unmount *fileSystem*:**  
*errorDescription*

**Explanation:** There was an error unmounting the GPFS file system.

**User response:** Take the action indicated by **errno** description.

---

**6027-512**     *name not listed in /etc/vfs*

**Explanation:** Error occurred while installing the GPFS kernel extension, or when trying to mount a file system.

**User response:** Check for the **mmfs** entry in **/etc/vfs**

---

**6027-514**     **Cannot mount *fileSystem* on *mountPoint*:**  
**Already mounted.**

**Explanation:** An attempt has been made to mount a file system that is already mounted.

**User response:** None. Informational message only.

---

**6027-515**     **Cannot mount *fileSystem* on *mountPoint***

**Explanation:** There was an error mounting the named GPFS file system. Errors in the disk path usually cause this problem.

**User response:** Take the action indicated by other error messages and error log entries.

---

**6027-516**     **Cannot mount *fileSystem***

**Explanation:** There was an error mounting the named GPFS file system. Errors in the disk path usually cause this problem.

---

**User response:** Take the action indicated by other error messages and error log entries.

---

**6027-517**     **Cannot mount *fileSystem*: *errorString***

**Explanation:** There was an error mounting the named GPFS file system. Errors in the disk path usually cause this problem.

**User response:** Take the action indicated by other error messages and error log entries.

---

**6027-518**     **Cannot mount *fileSystem*: Already mounted.**

**Explanation:** An attempt has been made to mount a file system that is already mounted.

**User response:** None. Informational message only.

---

**6027-519**     **Cannot mount *fileSystem* on *mountPoint*:**  
**File system table full.**

**Explanation:** An attempt has been made to mount a file system when the file system table is full.

**User response:** None. Informational message only.

---

**6027-520**     **Cannot mount *fileSystem*: File system table full.**

**Explanation:** An attempt has been made to mount a file system when the file system table is full.

**User response:** None. Informational message only.

---

**6027-530**     **Mount of *name* failed: cannot mount restorable file system for read/write.**

**Explanation:** A file system marked as **enabled** for restore cannot be mounted **read/write**.

**User response:** None. Informational message only.

---

**6027-531**     **The following disks of *name* will be formatted on node *nodeName*: *list*.**

**Explanation:** Output showing which disks will be formatted by the **mmcrfs** command.

**User response:** None. Informational message only.

---

**6027-532 [E]**   **The quota record *recordNumber* in file *fileName* is not valid.**

**Explanation:** A quota entry contained a checksum that is not valid.

**User response:** Remount the file system with quotas disabled. Restore the quota file from back up, and run **mmcheckquota**.

---

---

**6027-533 [W] Inode space *inodeSpace* in file system *fileSystem* is approaching the limit for the maximum number of inodes.**

**Explanation:** The number of files created is approaching the file system limit.

**User response:** Use the **mmchfileset** command to increase the maximum number of files to avoid reaching the inode limit and possible performance degradation.

---

**6027-534 Cannot create a snapshot in a DMAPI-enabled file system, *rc=returnCode*.**

**Explanation:** You cannot create a snapshot in a DMAPI-enabled file system.

**User response:** Use the **mmchfs** command to disable DMAPI, and reissue the command.

---

**6027-535 Disks up to size *size* can be added to storage pool *pool*.**

**Explanation:** Based on the parameters given to **mmcrfs** and the size and number of disks being formatted, GPFS has formatted its allocation maps to allow disks up the given size to be added to this storage pool by the **mmadddisk** command.

**User response:** None. Informational message only. If the reported maximum disk size is smaller than necessary, delete the file system with **mmdelfs** and rerun **mmcrfs** with either larger disks or a larger value for the **-n** parameter.

---

**6027-536 Insufficient system memory to run GPFS daemon. Reduce page pool memory size with the **mmchconfig** command or add additional RAM to system.**

**Explanation:** Insufficient memory for GPFS internal data structures with current system and GPFS configuration.

**User response:** Reduce page pool usage with the **mmchconfig** command, or add additional RAM to system.

---

**6027-537 Disks up to size *size* can be added to this file system.**

**Explanation:** Based on the parameters given to the **mmcrfs** command and the size and number of disks being formatted, GPFS has formatted its allocation maps to allow disks up the given size to be added to this file system by the **mmadddisk** command.

**User response:** None, informational message only. If the reported maximum disk size is smaller than necessary, delete the file system with **mmdelfs** and

reissue the **mmcrfs** command with larger disks or a larger value for the **-n** parameter.

---

**6027-538 Error accessing disks.**

**Explanation:** The **mmcrfs** command encountered an error accessing one or more of the disks.

**User response:** Verify that the disk descriptors are coded correctly and that all named disks exist and are online.

---

**6027-539 Unable to clear descriptor areas for *fileSystem*.**

**Explanation:** The **mmdelfs** command encountered an error while invalidating the file system control structures on one or more disks in the file system being deleted.

**User response:** If the problem persists, specify the **-p** option on the **mmdelfs** command.

---

**6027-540 Formatting file system.**

**Explanation:** The **mmcrfs** command began to write file system data structures onto the new disks.

**User response:** None. Informational message only.

---

**6027-541 Error formatting file system.**

**Explanation:** **mmcrfs** command encountered an error while formatting a new file system. This is often an I/O error.

**User response:** Check the subsystems in the path to the disk. Follow the instructions from other messages that appear with this one.

---

**6027-542 [N] Fileset in file system *fileSystem:filesetName* (*id filesetId*) has been incompletely deleted.**

**Explanation:** A fileset delete operation was interrupted, leaving this fileset in an incomplete state.

**User response:** Reissue the fileset delete command.

---

**6027-543 Error writing file system descriptor for *fileSystem*.**

**Explanation:** The **mmcrfs** command could not successfully write the file system descriptor in a particular file system. Check the subsystems in the path to the disk. This is often an I/O error.

**User response:** Check system error log, rerun **mmcrfs**.

---

---

**6027-544**     **Could not invalidate** *disk of fileSystem*.

**Explanation:** A disk could not be written to invalidate its contents. Check the subsystems in the path to the disk. This is often an I/O error.

**User response:** Ensure the indicated logical volume is writable.

---

**6027-545**     **Error processing fileset metadata file.**

**Explanation:** There is no I/O path to critical metadata or metadata has been corrupted.

**User response:** Verify that the I/O paths to all disks are valid and that all disks are either in the 'recovering' or 'up' availability states. If all disks are available and the problem persists, issue the **mmfsck** command to repair damaged metadata

---

**6027-546**     **Error processing allocation map for storage pool** *poolName*.

**Explanation:** There is no I/O path to critical metadata, or metadata has been corrupted.

**User response:** Verify that the I/O paths to all disks are valid, and that all disks are either in the 'recovering' or 'up' availability. Issue the **mmlsdisk** command.

---

**6027-547**     **Fileset** *filesetName* **was unlinked.**

**Explanation:** Fileset was already unlinked.

**User response:** None. Informational message only.

---

**6027-548**     **Fileset** *filesetName* **unlinked from** *filesetName*.

**Explanation:** A fileset being deleted contains junctions to other filesets. The cited fileset were unlinked.

**User response:** None. Informational message only.

---

**6027-549**     **Failed to open** *name*.

**Explanation:** The **mount** command was unable to access a file system. Check the subsystems in the path to the disk. This is often an I/O error.

**User response:** Follow the suggested actions for the other messages that occur with this one.

---

**6027-550 [X]**   **Allocation manager for** *fileSystem* **failed to revoke ownership from node** *nodeName*.

**Explanation:** An irrecoverable error occurred trying to revoke ownership of an allocation region. The allocation manager has panicked the file system to prevent corruption of on-disk data.

**User response:** Remount the file system.

---

**6027-551**     *fileSystem* **is still in use.**

**Explanation:** The **mmdelfs** or **mmcrfs** command found that the named file system is still mounted or that another GPFS command is running against the file system.

**User response:** Unmount the file system if it is mounted, or wait for GPFS commands in progress to terminate before retrying the command.

---

**6027-552**     **Scan completed successfully.**

**Explanation:** The scan function has completed without error.

**User response:** None. Informational message only.

---

**6027-553**     **Scan failed on** *number* **user or system files.**

**Explanation:** Data may be lost as a result of pointers that are not valid or unavailable disks.

**User response:** Some files may have to be restored from backup copies. Issue the **mmlsdisk** command to check the availability of all the disks that make up the file system.

---

**6027-554**     **Scan failed on** *number* **out of** *number* **user or system files.**

**Explanation:** Data may be lost as a result of pointers that are not valid or unavailable disks.

**User response:** Some files may have to be restored from backup copies. Issue the **mmlsdisk** command to check the availability of all the disks that make up the file system.

---

**6027-555**     **The desired replication factor exceeds the number of available failure groups.**

**Explanation:** You have specified a number of replicas that exceeds the number of failure groups available.

**User response:** Reissue the command with a smaller replication factor or increase the number of failure groups.

---

**6027-556**     **Not enough space for the desired number of replicas.**

**Explanation:** In attempting to restore the correct replication, GPFS ran out of space in the file system. The operation can continue but some data is not fully replicated.

**User response:** Make additional space available and reissue the command.

---

---

**6027-557**     **Not enough space or available disks to properly balance the file.**

**Explanation:** In attempting to stripe data within the file system, data was placed on a disk other than the desired one. This is normally not a problem.

**User response:** Run `mmrestripefs` to rebalance all files.

---

**6027-558**     **Some data are unavailable.**

**Explanation:** An I/O error has occurred or some disks are in the stopped state.

**User response:** Check the availability of all disks by issuing the `mmlsdisk` command and check the path to all disks. Reissue the command.

---

**6027-559**     **Some data could not be read or written.**

**Explanation:** An I/O error has occurred or some disks are in the stopped state.

**User response:** Check the availability of all disks and the path to all disks, and reissue the command.

---

**6027-560**     **File system is already suspended.**

**Explanation:** The `tsfsctl` command was asked to suspend a suspended file system.

**User response:** None. Informational message only.

---

**6027-561**     **Error migrating log.**

**Explanation:** There are insufficient available disks to continue operation.

**User response:** Restore the unavailable disks and reissue the command.

---

**6027-562**     **Error processing inodes.**

**Explanation:** There is no I/O path to critical metadata or metadata has been corrupted.

**User response:** Verify that the I/O paths to all disks are valid and that all disks are either in the recovering or up availability. Issue the `mmlsdisk` command.

---

**6027-563**     **File system is already running.**

**Explanation:** The `tsfsctl` command was asked to resume a file system that is already running.

**User response:** None. Informational message only.

---

**6027-564**     **Error processing inode allocation map.**

**Explanation:** There is no I/O path to critical metadata or metadata has been corrupted.

**User response:** Verify that the I/O paths to all disks

are valid and that all disks are either in the recovering or up availability. Issue the `mmlsdisk` command.

---

**6027-565**     **Scanning user file metadata ...**

**Explanation:** Progress information.

**User response:** None. Informational message only.

---

**6027-566**     **Error processing user file metadata.**

**Explanation:** Error encountered while processing user file metadata.

**User response:** None. Informational message only.

---

**6027-567**     **Waiting for pending file system scan to finish ...**

**Explanation:** Progress information.

**User response:** None. Informational message only.

---

**6027-568**     **Waiting for *number* pending file system scans to finish ...**

**Explanation:** Progress information.

**User response:** None. Informational message only.

---

**6027-569**     **Incompatible parameters. Unable to allocate space for file system metadata. Change one or more of the following as suggested and try again:**

**Explanation:** Incompatible file system parameters were detected.

**User response:** Refer to the details given and correct the file system parameters.

---

**6027-570**     **Incompatible parameters. Unable to create file system. Change one or more of the following as suggested and try again:**

**Explanation:** Incompatible file system parameters were detected.

**User response:** Refer to the details given and correct the file system parameters.

---

**6027-571**     **Logical sector size *value* must be the same as disk sector size.**

**Explanation:** This message is produced by the `mmcrfs` command if the sector size given by the `-l` option is not the same as the sector size given for disks in the `-d` option.

**User response:** Correct the options and reissue the command.

---

---

**6027-572**     **Completed creation of file system**  
*fileSystem.*

**Explanation:** The **mmcrfs** command has successfully completed.

**User response:** None. Informational message only.

---

**6027-573**     **All data on the following disks of**  
*fileSystem will be destroyed:*

**Explanation:** Produced by the **mmdelfs** command to list the disks in the file system that is about to be destroyed. Data stored on the disks will be lost.

**User response:** None. Informational message only.

---

**6027-574**     **Completed deletion of file system**  
*fileSystem.*

**Explanation:** The **mmdelfs** command has successfully completed.

**User response:** None. Informational message only.

---

**6027-575**     **Unable to complete low level format for**  
*fileSystem. Failed with error errorCode*

**Explanation:** The **mmcrfs** command was unable to create the low level file structures for the file system.

**User response:** Check other error messages and the error log. This is usually an error accessing disks.

---

**6027-576**     **Storage pools have not been enabled for**  
*file system fileSystem.*

**Explanation:** User invoked a command with a storage pool option (**-p** or **-P**) before storage pools were enabled.

**User response:** Enable storage pools with the **mmchfs -V** command, or correct the command invocation and reissue the command.

---

**6027-577**     **Attention: number user or system files**  
**are not properly replicated.**

**Explanation:** GPFS has detected files that are not replicated correctly due to a previous failure.

**User response:** Issue the **mmrestripefs** command at the first opportunity.

---

**6027-578**     **Attention: number out of number user or**  
**system files are not properly replicated:**

**Explanation:** GPFS has detected files that are not replicated correctly

---

**6027-579**     **Some unreplicated file system metadata**  
**has been lost. File system usable only in**  
**restricted mode.**

**Explanation:** A disk was deleted that contained vital file system metadata that was not replicated.

**User response:** Mount the file system in restricted mode (**-o rs**) and copy any user data that may be left on the file system. Then delete the file system.

---

**6027-580**     **Unable to access vital system metadata.**  
**Too many disks are unavailable.**

**Explanation:** Metadata is unavailable because the disks on which the data reside are stopped, or an attempt was made to delete them.

**User response:** Either start the stopped disks, try to delete the disks again, or recreate the file system.

---

**6027-581**     **Unable to access vital system metadata,**  
**file system corrupted.**

**Explanation:** When trying to access the files system, the metadata was unavailable due to a disk being deleted.

**User response:** Determine why a disk is unavailable.

---

**6027-582**     **Some data has been lost.**

**Explanation:** An I/O error has occurred or some disks are in the stopped state.

**User response:** Check the availability of all disks by issuing the **mmlsdisk** command and check the path to all disks. Reissue the command.

---

**6027-584**     **Incompatible parameters. Unable to**  
**allocate space for root directory. Change**  
**one or more of the following as**  
**suggested and try again:**

**Explanation:** Inconsistent parameters have been passed to the **mmcrfs** command, which would result in the creation of an inconsistent file system. Suggested parameter changes are given.

**User response:** Reissue the **mmcrfs** command with the suggested parameter changes.

---

**6027-585**     **Incompatible parameters. Unable to**  
**allocate space for ACL data. Change one**  
**or more of the following as suggested**  
**and try again:**

**Explanation:** Inconsistent parameters have been passed to the **mmcrfs** command, which would result in the creation of an inconsistent file system. The parameters entered require more space than is available. Suggested parameter changes are given.

**User response:** Reissue the `mmcrfs` command with the suggested parameter changes.

---

**6027-586 Quota server initialization failed.**

**Explanation:** Quota server initialization has failed. This message may appear as part of the detail data in the quota error log.

**User response:** Check status and availability of the disks. If quota files have been corrupted, restore them from the last available backup. Finally, reissue the command.

---

**6027-587 Unable to initialize quota client because there is no quota server. Please check error log on the file system manager node. The mmcheckquota command must be run with the file system unmounted before retrying the command.**

**Explanation:** `startQuotaClient` failed.

**User response:** If the quota file could not be read (check error log on file system manager. Issue the `mmlsmgr` command to determine which node is the file system manager), then the `mmcheckquota` command must be run with the file system unmounted.

---

**6027-588 No more than *number* nodes can mount a file system.**

**Explanation:** The limit of the number of nodes that can mount a file system was exceeded.

**User response:** Observe the stated limit for how many nodes can mount a file system.

---

**6027-589 Scanning file system metadata, phase *number* ...**

**Explanation:** Progress information.

**User response:** None. Informational message only.

---

**6027-590 [W] GPFS is experiencing a shortage of pagepool. This message will not be repeated for at least one hour.**

**Explanation:** Pool starvation occurs, buffers have to be continually stolen at high aggressiveness levels.

**User response:** Issue the `mmchconfig` command to increase the size of `pagepool`.

---

**6027-591 Unable to allocate sufficient inodes for file system metadata. Increase the value for *option* and try again.**

**Explanation:** Too few inodes have been specified on the `-N` option of the `mmcrfs` command.

**User response:** Increase the size of the `-N` option and reissue the `mmcrfs` command.

---

**6027-592 Mount of *fileSystem* is waiting for the mount disposition to be set by some data management application.**

**Explanation:** Data management utilizing DMAPI is enabled for the file system, but no data management application has set a disposition for the mount event.

**User response:** Start the data management application and verify that the application sets the mount disposition.

---

**6027-593 [E] The root quota entry is not found in its assigned record**

**Explanation:** On mount, the root entry is not found in the first record of the quota file.

**User response:** Issue the `mmcheckquota` command to verify that the use of root has not been lost.

---

**6027-594 Disk *diskName* cannot be added to storage pool *poolName*. Allocation map cannot accommodate disks larger than *size* MB.**

**Explanation:** The specified disk is too large compared to the disks that were initially used to create the storage pool.

**User response:** Specify a smaller disk or add the disk to a new storage pool.

---

**6027-595 [E] While creating quota files, file *fileName*, with no valid quota information was found in the root directory. Remove files with reserved quota file names (for example, `user.quota`) without valid quota information from the root directory by: - mounting the file system without quotas, - removing the files, and - remounting the file system with quotas to recreate new quota files. To use quota file names other than the reserved names, use the `mmcheckquota` command.**

**Explanation:** While mounting a file system, the state of the file system descriptor indicates that quota files do not exist. However, files that do not contain quota information but have one of the reserved names: `user.quota`, `group.quota`, or `filesset.quota` exist in the root directory.

**User response:** To mount the file system so that new quota files will be created, perform these steps:

1. Mount the file system without quotas.

2. Verify that there are no files in the root directory with the reserved names: **user.quota**, **group.quota**, or **fileset.quota**.
3. Remount the file system with quotas. To mount the file system with other files used as quota files, issue the **mmcheckquota** command.

---

**6027-596 [I]** While creating quota files, file *fileName* containing quota information was found in the root directory. This file will be used as *quotaType* quota file.

**Explanation:** While mounting a file system, the state of the file system descriptor indicates that quota files do not exist. However, files that have one of the reserved names **user.quota**, **group.quota**, or **fileset.quota** and contain quota information, exist in the root directory. The file with the reserved name will be used as the quota file.

**User response:** None. Informational message.

---

**6027-597 [E]** The quota command was requested to process quotas for a type (**user**, **group**, or **fileset**), which is not enabled.

**Explanation:** A quota command was requested to process quotas for a **user**, **group**, or **fileset** quota type, which is not enabled.

**User response:** Verify that the **user**, **group**, or **fileset** quota type is enabled and reissue the command.

---

**6027-598 [E]** The supplied file does not contain quota information.

**Explanation:** A file supplied as a quota file does not contain quota information.

**User response:** Change the file so it contains valid quota information and reissue the command.

To mount the file system so that new quota files are created:

1. Mount the file system without quotas.
2. Verify there are no files in the root directory with the reserved **user.quota** or **group.quota** name.
3. Remount the file system with quotas.

---

**6027-599 [E]** File supplied to the command does not exist in the root directory.

**Explanation:** The user-supplied name of a new quota file has not been found.

**User response:** Ensure that a file with the supplied name exists. Then reissue the command.

---

**6027-600** On node *nodeName* an earlier error may have caused some file system data to be inaccessible at this time. Check error log for additional information. After correcting the problem, the file system can be mounted again to restore normal data access.

**Explanation:** An earlier error may have caused some file system data to be inaccessible at this time.

**User response:** Check the error log for additional information. After correcting the problem, the file system can be mounted again.

---

**6027-601** Error changing pool size.

**Explanation:** The **mmchconfig** command failed to change the pool size to the requested value.

**User response:** Follow the suggested actions in the other messages that occur with this one.

---

**6027-602** ERROR: file system not mounted. Mount file system *fileSystem* and retry command.

**Explanation:** A GPFS command that requires the file system be mounted was issued.

**User response:** Mount the file system and reissue the command.

---

**6027-603** Current pool size: *valueK* = *valueM*, max block size: *valueK* = *valueM*.

**Explanation:** Displays the current pool size.

**User response:** None. Informational message only.

---

**6027-604 [E]** Parameter incompatibility. File system block size is larger than **maxblocksize** parameter.

**Explanation:** An attempt is being made to mount a file system whose block size is larger than the **maxblocksize** parameter as set by **mmchconfig**.

**User response:** Use the **mmchconfig maxblocksize=xxx** command to increase the maximum allowable block size.

---

**6027-605 [N]** File system has been renamed.

**Explanation:** Self-explanatory.

**User response:** None. Informational message only.

---

**6027-606 [E]** The node number *nodeNumber* is not defined in the node list

**Explanation:** A node matching *nodeNumber* was not found in the GPFS configuration file.

**User response:** Perform required configuration steps prior to starting GPFS on the node.

---

**6027-607** **mmcommon getEFOptions** *fileSystem* failed. Return code *value*.

**Explanation:** The **mmcommon getEFOptions** command failed while looking up the names of the disks in a file system. This error usually occurs during mount processing.

**User response:** Check the preceding messages. A frequent cause for such errors is lack of space in */var*.

---

**6027-608 [E]** File system manager takeover failed.

**Explanation:** An attempt to takeover as file system manager failed. The file system is unmounted to allow another node to try.

**User response:** Check the return code. This is usually due to network or disk connectivity problems. Issue the **mmlsdisk** command to determine if the paths to the disk are unavailable, and issue the **mmchdisk** if necessary.

---

**6027-609** File system *fileSystem* unmounted because it does not have a manager.

**Explanation:** The file system had to be unmounted because a file system manager could not be assigned. An accompanying message tells which node was the last manager.

**User response:** Examine error log on the last file system manager. Issue the **mmlsdisk** command to determine if a number of disks are down. Examine the other error logs for an indication of network, disk, or virtual shared disk problems. Repair the base problem and issue the **mmchdisk** command if required.

---

**6027-610** Cannot mount file system *fileSystem* because it does not have a manager.

**Explanation:** The file system had to be unmounted because a file system manager could not be assigned. An accompanying message tells which node was the last manager.

**User response:** Examine error log on the last file system manager node. Issue the **mmlsdisk** command to determine if a number of disks are down. Examine the other error logs for an indication of disk or network shared disk problems. Repair the base problem and issue the **mmchdisk** command if required.

---



---

**6027-611 [I]** Recovery: *fileSystem*, delay number sec. for safe recovery.

**Explanation:** Informational. When disk leasing is in use, wait for the existing lease to expire before performing log and token manager recovery.

**User response:** None.

---

**6027-612** Unable to run *command* while the file system is suspended.

**Explanation:** A command that can alter data in a file system was issued while the file system was suspended.

**User response:** Resume the file system and reissue the command.

---

**6027-613 [N]** Expel *node* request from *node*. Expelling: *node*

**Explanation:** One node is asking to have another node expelled from the cluster, usually because they have communications problems between them. The cluster manager node will decide which one will be expelled.

**User response:** Check that the communications paths are available between the two nodes.

---

**6027-614** Value *value* for option *name* is out of range. Valid values are *number* through *number*.

**Explanation:** The value for an option in the command line arguments is out of range.

**User response:** Correct the command line and reissue the command.

---

**6027-615** **mmcommon getContactNodes** *clusterName* failed. Return code *value*.

**Explanation:** **mmcommon getContactNodes** failed while looking up contact nodes for a remote cluster, usually while attempting to mount a file system from a remote cluster.

**User response:** Check the preceding messages, and consult the earlier chapters of this document. A frequent cause for such errors is lack of space in */var*.

---

**6027-616 [X]** Duplicate address *ipAddress* in node list

**Explanation:** The IP address appears more than once in the node list file.

**User response:** Check the node list shown by the **mmlscluster** command.

---



---

**6027-617 [I] Recovered *number* nodes for cluster *clusterName*.**

**Explanation:** The asynchronous part (phase 2) of node failure recovery has completed.

**User response:** None. Informational message only.

---

**6027-618 [X] Local host not found in node list (local ip interfaces: *interfaceList*)**

**Explanation:** The local host specified in the node list file could not be found.

**User response:** Check the node list shown by the **mmlscluster** command.

---

**6027-619 Negative grace times are not allowed.**

**Explanation:** The **mmedquota** command received a negative value for the **-t** option.

**User response:** Reissue the **mmedquota** command with a nonnegative value for grace time.

---

**6027-620 Hard quota limit must not be less than soft limit.**

**Explanation:** The hard quota limit must be greater than or equal to the soft quota limit.

**User response:** Reissue the **mmedquota** command and enter valid values when editing the information.

---

**6027-621 Negative quota limits are not allowed.**

**Explanation:** The quota value must be positive.

**User response:** Reissue the **mmedquota** command and enter valid values when editing the information.

---

**6027-622 [E] Failed to join remote cluster *clusterName***

**Explanation:** The node was not able to establish communication with another cluster, usually while attempting to mount a file system from a remote cluster.

**User response:** Check other console messages for additional information. Verify that contact nodes for the remote cluster are set correctly. Run **mmremotefs show** and **mmremotecluster show** to display information about the remote cluster.

---

**6027-623 All disks up and ready**

**Explanation:** Self-explanatory.

**User response:** None. Informational message only.

---

**6027-624 No disks**

**Explanation:** Self-explanatory.

**User response:** None. Informational message only.

---

**6027-625 File system manager takeover already pending.**

**Explanation:** A request to migrate the file system manager failed because a previous migrate request has not yet completed.

**User response:** None. Informational message only.

---

**6027-626 Migrate to node *nodeName* already pending.**

**Explanation:** A request to migrate the file system manager failed because a previous migrate request has not yet completed.

**User response:** None. Informational message only.

---

**6027-627 Node *nodeName* is already manager for *fileSystem*.**

**Explanation:** A request has been made to change the file system manager node to the node that is already the manager.

**User response:** None. Informational message only.

---

**6027-628 Sending migrate request to current manager node *nodeName*.**

**Explanation:** A request has been made to change the file system manager node.

**User response:** None. Informational message only.

---

**6027-629 [N] Node *nodeName* resigned as manager for *fileSystem*.**

**Explanation:** Progress report produced by the **mmchmgr** command.

**User response:** None. Informational message only.

---

**6027-630 [N] Node *nodeName* appointed as manager for *fileSystem*.**

**Explanation:** The **mmchmgr** command successfully changed the node designated as the file system manager.

**User response:** None. Informational message only.

---

**6027-631 Failed to appoint node *nodeName* as manager for *fileSystem*.**

**Explanation:** A request to change the file system manager node has failed.

**User response:** Accompanying messages will describe the reason for the failure. Also, see the `mmfs.log` file on the target node.

---

**6027-632 Failed to appoint new manager for *fileSystem*.**

**Explanation:** An attempt to change the file system manager node has failed.

**User response:** Accompanying messages will describe the reason for the failure. Also, see the `mmfs.log` file on the target node.

---

**6027-633 The best choice node *nodeName* is already the manager for *fileSystem*.**

**Explanation:** Informational message about the progress and outcome of a migrate request.

**User response:** None. Informational message only.

---

**6027-634 Node name or number *node* is not valid.**

**Explanation:** A node number, IP address, or host name that is not valid has been entered in the configuration file or as input for a command.

**User response:** Validate your configuration information and the condition of your network. This message may result from an inability to translate a node name.

---

**6027-635 [E] The current file system manager failed and no new manager will be appointed.**

**Explanation:** The file system manager node could not be replaced. This is usually caused by other system errors, such as disk or communication errors.

**User response:** See accompanying messages for the base failure.

---

**6027-636 [E] Disk marked as stopped or offline.**

**Explanation:** A disk continues to be marked **down** due to a previous error and was not opened again.

**User response:** Check the disk status by issuing the `mmlsdisk` command, then issue the `mmchdisk start` command to restart the disk.

---

**6027-637 [E] RVSD is not active.**

**Explanation:** The RVSD subsystem needs to be activated.

**User response:** See the appropriate IBM Reliable Scalable Cluster Technology (RSCT) document ([www.ibm.com/support/knowledgecenter/SGVKBA/welcome](http://www.ibm.com/support/knowledgecenter/SGVKBA/welcome)) and search on *diagnosing IBM Virtual Shared Disk problems*.

---

**6027-638 [E] File system *fileSystem* unmounted by node *nodeName***

**Explanation:** Produced in the console log on a forced unmount of the file system caused by disk or communication failures.

**User response:** Check the error log on the indicated node. Correct the underlying problem and remount the file system.

---

**6027-639 [E] File system cannot be mounted in restricted mode and ro or rw concurrently**

**Explanation:** There has been an attempt to concurrently mount a file system on separate nodes in both a normal mode and in 'restricted' mode.

**User response:** Decide which mount mode you want to use, and use that mount mode on both nodes.

---

**6027-640 [E] File system is mounted**

**Explanation:** A command has been issued that requires that the file system be unmounted.

**User response:** Unmount the file system and reissue the command.

---

**6027-641 [E] Unable to access vital system metadata. Too many disks are unavailable or the file system is corrupted.**

**Explanation:** An attempt has been made to access a file system, but the metadata is unavailable. This can be caused by:

1. The disks on which the metadata resides are either stopped or there was an unsuccessful attempt to delete them.
2. The file system is corrupted.

**User response:** To access the file system:

1. If the disks are the problem either start the stopped disks or try to delete them.
2. If the file system has been corrupted, you will have to recreate it from backup medium.

---

**6027-642 [N] File system has been deleted.**

**Explanation:** Self-explanatory.

**User response:** None. Informational message only.

---

**6027-643 [I] Node *nodeName* completed take over for *fileSystem*.**

**Explanation:** The `mmchmgr` command completed successfully.

**User response:** None. Informational message only.

---

**6027-644**    **The previous error was detected on node *nodeName*.**

**Explanation:** An unacceptable error was detected. This usually occurs when attempting to retrieve file system information from the operating system's file system database or the cached GPFS system control data. The message identifies the node where the error was encountered.

**User response:** See accompanying messages for the base failure. A common cause for such errors is lack of space in */var*.

---

**6027-645**    **Attention: mmcommon getEFOptions *fileSystem* failed. Checking *fileName*.**

**Explanation:** The names of the disks in a file system were not found in the cached GPFS system data, therefore an attempt will be made to get the information from the operating system's file system database.

**E! User response:** If the command fails, see "File system E! fails to mount" on page 245. A common cause for such errors is lack of space in */var*.

---

**6027-646 [E]**    **File system unmounted due to loss of cluster membership.**

**Explanation:** Quorum was lost, causing file systems to be unmounted.

**User response:** Get enough nodes running the GPFS daemon to form a quorum.

---

**6027-647 [E]**    **File *fileName* could not be run with error *errno*.**

**Explanation:** The specified shell script could not be run. This message is followed by the error string that is returned by the exec.

**User response:** Check file existence and access permissions.

---

**6027-648**    **EDITOR environment variable must be full pathname.**

**Explanation:** The value of the EDITOR environment variable is not an absolute path name.

**User response:** Change the value of the EDITOR environment variable to an absolute path name.

---

**6027-649**    **Error reading the mmpmon command file.**

**Explanation:** An error occurred when reading the **mmpmon** command file.

**User response:** Check file existence and access permissions.

---

**6027-650 [X]**    **The mmfs daemon is shutting down abnormally.**

**Explanation:** The GPFS daemon is shutting down as a result of an irrecoverable condition, typically a resource shortage.

**User response:** Review error log entries, correct a resource shortage condition, and restart the GPFS daemon.

---

**6027-660**    **Error displaying message from mmfsd.**

**Explanation:** GPFS could not properly display an output string sent from the **mmfsd** daemon due to some error. A description of the error follows.

**User response:** Check that GPFS is properly installed.

---

**6027-661**    **mmfsd waiting for primary node *nodeName*.**

**Explanation:** The **mmfsd** server has to wait during start up because **mmfsd** on the primary node is not yet ready.

**User response:** None. Informational message only.

---

**6027-662**    **mmfsd timed out waiting for primary node *nodeName*.**

**Explanation:** The **mmfsd** server is about to terminate.

**User response:** Ensure that the **mmfs.cfg** configuration file contains the correct host name or IP address of the primary node. Check **mmfsd** on the primary node.

---

**6027-663**    **Lost connection to file system daemon.**

**Explanation:** The connection between a GPFS command and the **mmfsd** daemon has broken. The daemon has probably crashed.

**User response:** Ensure that the **mmfsd** daemon is running. Check the error log.

---

**6027-664**    **Unexpected message from file system daemon.**

**Explanation:** The version of the **mmfsd** daemon does not match the version of the GPFS command.

**User response:** Ensure that all GPFS software components are at the same version.

---

**6027-665**    **Failed to connect to file system daemon: *errorString***

**Explanation:** An error occurred while trying to create a session with **mmfsd**.

**User response:** Ensure that the **mmfsd** daemon is running. Also, only root can run most GPFS

commands. The mode bits of the commands must be **set-user-id** to **root**.

---

**6027-666 Failed to determine file system manager.**

**Explanation:** While running a GPFS command in a multiple node configuration, the local file system daemon is unable to determine which node is managing the file system affected by the command.

**User response:** Check internode communication configuration and ensure that enough GPFS nodes are up to form a quorum.

---

**6027-667 Could not set up socket**

**Explanation:** One of the calls to create or bind the socket used for sending parameters and messages between the command and the daemon failed.

**User response:** Check additional error messages.

---

**6027-668 Could not send message to file system daemon**

**Explanation:** Attempt to send a message to the file system failed.

**User response:** Check if the file system daemon is up and running.

---

**6027-669 Could not connect to file system daemon.**

**Explanation:** The TCP connection between the command and the daemon could not be established.

**User response:** Check additional error messages.

---

**6027-670 Value for 'option' is not valid. Valid values are *list*.**

**Explanation:** The specified value for the given command option was not valid. The remainder of the line will list the valid keywords.

**User response:** Correct the command line.

---

**6027-671 Keyword missing or incorrect.**

**Explanation:** A missing or incorrect keyword was encountered while parsing command line arguments

**User response:** Correct the command line.

---

**6027-672 Too few arguments specified.**

**Explanation:** Too few arguments were specified on the command line.

**User response:** Correct the command line.

---

**6027-673 Too many arguments specified.**

**Explanation:** Too many arguments were specified on the command line.

**User response:** Correct the command line.

---

**6027-674 Too many values specified for option *name*.**

**Explanation:** Too many values were specified for the given option on the command line.

**User response:** Correct the command line.

---

**6027-675 Required value for *option* is missing.**

**Explanation:** A required value was not specified for the given option on the command line.

**User response:** Correct the command line.

---

**6027-676 Option *option* specified more than once.**

**Explanation:** The named option was specified more than once on the command line.

**User response:** Correct the command line.

---

**6027-677 Option *option* is incorrect.**

**Explanation:** An incorrect option was specified on the command line.

**User response:** Correct the command line.

---

**6027-678 Misplaced or incorrect parameter *name*.**

**Explanation:** A misplaced or incorrect parameter was specified on the command line.

**User response:** Correct the command line.

---

**6027-679 Device *name* is not valid.**

**Explanation:** An incorrect device name was specified on the command line.

**User response:** Correct the command line.

---

**6027-680 [E] Disk failure. Volume *name*. rc = *value*. Physical volume *name*.**

**Explanation:** An I/O request to a disk or a request to fence a disk has failed in such a manner that GPFS can no longer use the disk.

**User response:** Check the disk hardware and the software subsystems in the path to the disk.

---

**6027-681** Required option *name* was not specified.

**Explanation:** A required option was not specified on the command line.

**User response:** Correct the command line.

---

**6027-682** Device argument is missing.

**Explanation:** The device argument was not specified on the command line.

**User response:** Correct the command line.

---

**6027-683** Disk *name* is invalid.

**Explanation:** An incorrect disk name was specified on the command line.

**User response:** Correct the command line.

---

**6027-684** Value *value* for option is incorrect.

**Explanation:** An incorrect value was specified for the named option.

**User response:** Correct the command line.

---

**6027-685** Value *value* for option *option* is out of range. Valid values are *number* through *number*.

**Explanation:** An out of range value was specified for the named option.

**User response:** Correct the command line.

---

**6027-686** *option (value)* exceeds option *(value)*.

**Explanation:** The value of the first option exceeds the value of the second option. This is not permitted.

**User response:** Correct the command line.

---

**6027-687** Disk *name* is specified more than once.

**Explanation:** The named disk was specified more than once on the command line.

**User response:** Correct the command line.

---

**6027-688** Failed to read file system descriptor.

**Explanation:** The disk block containing critical information about the file system could not be read from disk.

**User response:** This is usually an error in the path to the disks. If there are associated messages indicating an I/O error such as **ENODEV** or **EIO**, correct that error and retry the operation. If there are no associated I/O errors, then run the **mmfsck** command with the file system unmounted.

---

**6027-689** Failed to update file system descriptor.

**Explanation:** The disk block containing critical information about the file system could not be written to disk.

**User response:** This is a serious error, which may leave the file system in an unusable state. Correct any I/O errors, then run the **mmfsck** command with the file system unmounted to make repairs.

---

**6027-690** Failed to allocate I/O buffer.

**Explanation:** Could not obtain enough memory (RAM) to perform an operation.

**User response:** Either retry the operation when the **mmfsd** daemon is less heavily loaded, or increase the size of one or more of the memory pool parameters by issuing the **mmchconfig** command.

---

**6027-691** Failed to send message to node *nodeName*.

**Explanation:** A message to another file system node could not be sent.

**User response:** Check additional error message and the internode communication configuration.

---

**6027-692** Value for *option* is not valid. Valid values are **yes**, **no**.

**Explanation:** An option that is required to be **yes** or **no** is neither.

**User response:** Correct the command line.

---

**6027-693** Cannot open disk *name*.

**Explanation:** Could not access the given disk.

**User response:** Check the disk hardware and the path to the disk.

---

**6027-694** Disk not started; disk *name* has a bad volume label.

**Explanation:** The volume label on the disk does not match that expected by GPFS.

**User response:** Check the disk hardware. For hot-pluggable drives, ensure that the proper drive has been plugged in.

---

**6027-695 [E]** File system is read-only.

**Explanation:** An operation was attempted that would require modifying the contents of a file system, but the file system is read-only.

**User response:** Make the file system R/W before retrying the operation.

---

---

**6027-696 [E] Too many disks are unavailable.**

**Explanation:** A file system operation failed because all replicas of a data or metadata block are currently unavailable.

**User response:** Issue the **mmlsdisk** command to check the availability of the disks in the file system; correct disk hardware problems, and then issue the **mmchdisk** command with the **start** option to inform the file system that the disk or disks are available again.

---

**6027-697 [E] No log available.**

**Explanation:** A file system operation failed because no space for logging metadata changes could be found.

**User response:** Check additional error message. A likely reason for this error is that all disks with available log space are currently unavailable.

---

**6027-698 [E] Not enough memory to allocate internal data structure.**

**Explanation:** A file system operation failed because no memory is available for allocating internal data structures.

**User response:** Stop other processes that may have main memory pinned for their use.

---

**6027-699 [E] Inconsistency in file system metadata.**

**Explanation:** File system metadata on disk has been corrupted.

**User response:** This is an extremely serious error that may cause loss of data. Issue the **mmfsck** command with the file system unmounted to make repairs. There will be a **POSSIBLE FILE CORRUPTION** entry in the system error log that should be forwarded to the IBM Support Center.

---

**6027-700 [E] Log recovery failed.**

**Explanation:** An error was encountered while restoring file system metadata from the log.

**User response:** Check additional error message. A likely reason for this error is that none of the replicas of the log could be accessed because too many disks are currently unavailable. If the problem persists, issue the **mmfsck** command with the file system unmounted.

---

**6027-701 [X] Some file system data are inaccessible at this time.**

**Explanation:** The file system has encountered an error that is serious enough to make some or all data inaccessible. This message indicates that an error occurred that left the file system in an unusable state.

**User response:** Possible reasons include too many

unavailable disks or insufficient memory for file system control structures. Check other error messages as well as the error log for additional information. Unmount the file system and correct any I/O errors. Then remount the file system and try the operation again. If the problem persists, issue the **mmfsck** command with the file system unmounted to make repairs.

---

**6027-702 [X] Some file system data are inaccessible at this time. Check error log for additional information. After correcting the problem, the file system must be unmounted and then mounted to restore normal data access.**

**Explanation:** The file system has encountered an error that is serious enough to make some or all data inaccessible. This message indicates that an error occurred that left the file system in an unusable state.

**User response:** Possible reasons include too many unavailable disks or insufficient memory for file system control structures. Check other error messages as well as the error log for additional information. Unmount the file system and correct any I/O errors. Then remount the file system and try the operation again. If the problem persists, issue the **mmfsck** command with the file system unmounted to make repairs.

---

**6027-703 [X] Some file system data are inaccessible at this time. Check error log for additional information.**

**Explanation:** The file system has encountered an error that is serious enough to make some or all data inaccessible. This message indicates that an error occurred that left the file system in an unusable state.

**User response:** Possible reasons include too many unavailable disks or insufficient memory for file system control structures. Check other error messages as well as the error log for additional information. Unmount the file system and correct any I/O errors. Then remount the file system and try the operation again. If the problem persists, issue the **mmfsck** command with the file system unmounted to make repairs.

---

**6027-704 Attention: Due to an earlier error normal access to this file system has been disabled. Check error log for additional information. After correcting the problem, the file system must be unmounted and then mounted again to restore normal data access.**

**Explanation:** The file system has encountered an error that is serious enough to make some or all data inaccessible. This message indicates that an error occurred that left the file system in an unusable state.

**User response:** Possible reasons include too many unavailable disks or insufficient memory for file system

control structures. Check other error messages as well as the error log for additional information. Unmount the file system and correct any I/O errors. Then remount the file system and try the operation again. If the problem persists, issue the **mmfsck** command with the file system unmounted to make repairs.

---

**6027-705**      **Error code** *value*.

**Explanation:** Provides additional information about an error.

**User response:** See accompanying error messages.

---

**6027-706**      **The device** *name* **has no corresponding entry in** *fileName* **or has an incomplete entry.**

**Explanation:** The command requires a device that has a file system associated with it.

**User response:** Check the operating system's file system database (the given file) for a valid device entry.

---

**6027-707**      **Unable to open file** *fileName*.

**Explanation:** The named file cannot be opened.

**User response:** Check that the file exists and has the correct permissions.

---

**6027-708**      **Keyword** *name* **is incorrect. Valid values are** *list*.

**Explanation:** An incorrect keyword was encountered.

**User response:** Correct the command line.

---

**6027-709**      **Incorrect response. Valid responses are** "yes", "no", or "noall"

**Explanation:** A question was asked that requires a **yes** or **no** answer. The answer entered was neither **yes**, **no**, nor **noall**.

**User response:** Enter a valid response.

---

**6027-710**      **Attention:**

**Explanation:** Precedes an attention messages.

**User response:** None. Informational message only.

---

**6027-711 [E]**    **Specified entity, such as a disk or file system, does not exist.**

**Explanation:** A file system operation failed because the specified entity, such as a disk or file system, could not be found.

**User response:** Specify existing disk, file system, etc.

---

**6027-712 [E]**    **Error in communications between mmfsd daemon and client program.**

**Explanation:** A message sent between the **mmfsd** daemon and the client program had an incorrect format or content.

**User response:** Verify that the **mmfsd** daemon is running.

---

**6027-713**      **Unable to start because conflicting program** *name* **is running. Waiting until it completes.**

**Explanation:** A program detected that it cannot start because a conflicting program is running. The program will automatically start once the conflicting program has ended, as long as there are no other conflicting programs running at that time.

**User response:** None. Informational message only.

---

**6027-714**      **Terminating because conflicting program** *name* **is running.**

**Explanation:** A program detected that it must terminate because a conflicting program is running.

**User response:** Reissue the command once the conflicting program has ended.

---

**6027-715**      *command* **is finished waiting. Starting execution now.**

**Explanation:** A program detected that it can now begin running because a conflicting program has ended.

**User response:** None. Information message only.

---

**6027-716 [E]**    **Some file system data or metadata has been lost.**

**Explanation:** Unable to access some piece of file system data that has been lost due to the deletion of disks beyond the replication factor.

**User response:** If the function did not complete, try to mount the file system in **restricted** mode.

---

**6027-717 [E]**    **Must execute mmfsck before mount.**

**Explanation:** An attempt has been made to mount a file system on which an incomplete **mmfsck** command was run.

**User response:** Reissue the **mmfsck** command to the repair file system, then reissue the **mount** command.

---

**6027-718**    **The mmfsd daemon is not ready to handle commands yet.**

**Explanation:** The mmfsd daemon is not accepting messages because it is restarting or stopping.

**User response:** None. Informational message only.

---

**6027-719 [E] Device type not supported.**

**Explanation:** A disk being added to a file system with the **mmadddisk** or **mmcrfs** command is not a character mode special file, or has characteristics not recognized by GPFS.

**User response:** Check the characteristics of the disk being added to the file system.

---

**6027-720 [E] Actual sector size does not match given sector size.**

**Explanation:** A disk being added to a file system with the **mmadddisk** or **mmcrfs** command has a physical sector size that differs from that given in the disk description list.

**User response:** Check the physical sector size of the disk being added to the file system.

---

**6027-721 [E] Host 'name' in fileName is not valid.**

**Explanation:** A host name or IP address that is not valid was found in a configuration file.

**User response:** Check the configuration file specified in the error message.

---

**6027-722**    **Attention: Due to an earlier error normal access to this file system has been disabled. Check error log for additional information. The file system must be mounted again to restore normal data access.**

**Explanation:** The file system has encountered an error that is serious enough to make some or all data inaccessible. This message indicates that an error occurred that left the file system in an unusable state. Possible reasons include too many unavailable disks or insufficient memory for file system control structures.

**User response:** Check other error messages as well as the error log for additional information. Correct any I/O errors. Then, remount the file system and try the operation again. If the problem persists, issue the **mmfsck** command with the file system unmounted to make repairs.

---

**6027-723**    **Attention: Due to an earlier error normal access to this file system has been disabled. Check error log for additional information. After correcting the problem, the file system must be mounted again to restore normal data access.**

**Explanation:** The file system has encountered an error that is serious enough to make some or all data inaccessible. This message indicates that an error occurred that left the file system in an unusable state. Possible reasons include too many unavailable disks or insufficient memory for file system control structures.

**User response:** Check other error messages as well as the error log for additional information. Correct any I/O errors. Then, remount the file system and try the operation again. If the problem persists, issue the **mmfsck** command with the file system unmounted to make repairs.

---

**6027-724 [E] Incompatible file system format.**

**Explanation:** An attempt was made to access a file system that was formatted with an older version of the product that is no longer compatible with the version currently running.

**User response:** To change the file system format version to the current version, issue the **-V** option on the **mmchfs** command.

---

**6027-725**    **The mmfsd daemon is not ready to handle commands yet. Waiting for quorum.**

**Explanation:** The GPFS mmfsd daemon is not accepting messages because it is waiting for quorum.

**User response:** Determine why insufficient nodes have joined the group to achieve quorum and rectify the problem.

---

**6027-726 [E] Quota initialization/start-up failed.**

**Explanation:** Quota manager initialization was unsuccessful. The file system manager finished without quotas. Subsequent client mount requests will fail.

**User response:** Check the error log and correct I/O errors. It may be necessary to issue the **mmcheckquota** command with the file system unmounted.

---

**6027-727**    **Specified driver type type does not match disk name driver type type.**

**Explanation:** The driver type specified on the **mmchdisk** command does not match the current driver type of the disk.

**User response:** Verify the driver type and reissue the command.

---



---

**6027-728**      **Specified sector size *value* does not match disk *name* sector size *value*.**

**Explanation:** The sector size specified on the **mmchdisk** command does not match the current sector size of the disk.

**User response:** Verify the sector size and reissue the command.

---

**6027-729**      **Attention: No changes for disk *name* were specified.**

**Explanation:** The disk descriptor in the **mmchdisk** command does not specify that any changes are to be made to the disk.

**User response:** Check the disk descriptor to determine if changes are needed.

---

**6027-730**      *command on fileSystem.*

**Explanation:** Quota was activated or deactivated as stated as a result of the **mmquotaon**, **mmquotaoff**, **mmdefquotaon**, or **mmdefquotaoff** commands.

**User response:** None, informational only. This message is enabled with the **-v** option on the **mmquotaon**, **mmquotaoff**, **mmdefquotaon**, or **mmdefquotaoff** commands.

---

**6027-731**      **Error *number* while performing *command* for *name* quota on *fileSystem***

**Explanation:** An error occurred when switching quotas of a certain type on or off. If errors were returned for multiple file systems, only the error code is shown.

**User response:** Check the error code shown by the message to determine the reason.

---

**6027-732**      **Error while performing *command* on *fileSystem*.**

**Explanation:** An error occurred while performing the stated command when listing or reporting quotas.

**User response:** None. Informational message only.

---

**6027-733**      **Edit quota: Incorrect format!**

**Explanation:** The format of one or more edited quota limit entries was not correct.

**User response:** Reissue the **mmedquota** command. Change only the values for the limits and follow the instructions given.

---



---

**6027-734 [W]**      **Quota check for '*fileSystem*' ended prematurely.**

**Explanation:** The user interrupted and terminated the command.

**User response:** If ending the command was not intended, reissue the **mmcheckquota** command.

---

**6027-735**      **Error editing string from *mmfsd*.**

**Explanation:** An internal error occurred in the **mmfsd** when editing a string.

**User response:** None. Informational message only.

---

**6027-736**      **Attention: Due to an earlier error normal access to this file system has been disabled. Check error log for additional information. The file system must be unmounted and then mounted again to restore normal data access.**

**Explanation:** The file system has encountered an error that is serious enough to make some or all data inaccessible. This message indicates that an error occurred that left the file system in an unusable state. Possible reasons include too many unavailable disks or insufficient memory for file system control structures.

**User response:** Check other error messages as well as the error log for additional information. Unmount the file system and correct any I/O errors. Then, remount the file system and try the operation again. If the problem persists, issue the **mmfsck** command with the file system unmounted to make repairs.

---

**6027-737**      **Attention: No metadata disks remain.**

**Explanation:** The **mmchdisk** command has been issued, but no metadata disks remain.

**User response:** None. Informational message only.

---

**6027-738**      **Attention: No data disks remain.**

**Explanation:** The **mmchdisk** command has been issued, but no data disks remain.

**User response:** None. Informational message only.

---

**6027-739**      **Attention: Due to an earlier configuration change the file system is no longer properly balanced.**

**Explanation:** The **mmlsdisk** command found that the file system is not properly balanced.

**User response:** Issue the **mmrestripefs -b** command at your convenience.

---

---

**6027-740**      **Attention: Due to an earlier configuration change the file system is no longer properly replicated.**

**Explanation:** The `mmfsdisk` command found that the file system is not properly replicated.

**User response:** Issue the `mmrestripefs -r` command at your convenience

---

**6027-741**      **Attention: Due to an earlier configuration change the file system may contain data that is at risk of being lost.**

**Explanation:** The `mmfsdisk` command found that critical data resides on disks that are suspended or being deleted.

**User response:** Issue the `mmrestripefs -m` command as soon as possible.

---

**6027-742**      **Error occurred while executing a command for *fileSystem*.**

**Explanation:** A quota command encountered a problem on a file system. Processing continues with the next file system.

**User response:** None. Informational message only.

---

**6027-743**      **Initial disk state was updated successfully, but another error may have changed the state again.**

**Explanation:** The `mmchdisk` command encountered an error after the disk status or availability change was already recorded in the file system configuration. The most likely reason for this problem is that too many disks have become unavailable or are still unavailable after the disk state change.

**User response:** Issue an `mmchdisk start` command when more disks are available.

---

**6027-744**      **Unable to run *command* while the file system is mounted in restricted mode.**

**Explanation:** A command that can alter the data in a file system was issued while the file system was mounted in restricted mode.

**User response:** Mount the file system in read-only or read-write mode or unmount the file system and then reissue the command.

---

**6027-745**      *fileSystem*: **no *quotaType* quota management enabled.**

**Explanation:** A quota command of the cited type was issued for the cited file system when no quota management was enabled.

---

**User response:** Enable quota management and reissue the command.

---

**6027-746**      **Editing quota limits for this user or group not permitted.**

**Explanation:** The `root` user or `system` group was specified for quota limit editing in the `mmedquota` command.

**User response:** Specify a valid user or group in the `mmedquota` command. Editing quota limits for the `root` user or `system` group is prohibited.

---

**6027-747 [E]**    **Too many nodes in cluster (*max number*) or file system (*max number*).**

**Explanation:** The operation cannot succeed because too many nodes are involved.

**User response:** Reduce the number of nodes to the applicable stated limit.

---

**6027-748**      *fileSystem*: **no quota management enabled**

**Explanation:** A quota command was issued for the cited file system when no quota management was enabled.

**User response:** Enable quota management and reissue the command.

---

**6027-749**      **Pool size changed to *number K* = *number M*.**

**Explanation:** Pool size successfully changed.

**User response:** None. Informational message only.

---

**6027-750 [E]**    **The node address *ipAddress* is not defined in the node list**

**Explanation:** An address does not exist in the GPFS configuration file.

**User response:** Perform required configuration steps prior to starting GPFS on the node.

---

**6027-751 [E]**    **Error code *value***

**Explanation:** Provides additional information about an error.

**User response:** See accompanying error messages.

---

**6027-752 [E]**    **Lost membership in cluster *clusterName*. Unmounting file systems.**

**Explanation:** This node has lost membership in the cluster. Either GPFS is no longer available on enough nodes to maintain quorum, or this node could not communicate with other members of the quorum. This

could be caused by a communications failure between nodes, or multiple GPFS failures.

**User response:** See associated error logs on the failed nodes for additional problem determination information.

**6027-753 [E] Could not run command *command***

**Explanation:** The GPFS daemon failed to run the specified command.

**User response:** Verify correct installation.

**6027-754 Error reading string for mmfsd.**

**Explanation:** GPFS could not properly read an input string.

**User response:** Check that GPFS is properly installed.

**6027-755 [I] Waiting for challenge *challengeValue* (node *nodeNumber*, sequence *sequenceNumber*) to be responded during disk election**

**Explanation:** The node has challenged another node, which won the previous election and is waiting for the challenger to respond.

**User response:** None. Informational message only.

**6027-756 [E] Configuration invalid or inconsistent between different nodes.**

**Explanation:** Self-explanatory.

**User response:** Check cluster and file system configuration.

**6027-757 *name* is not an excluded disk.**

**Explanation:** Some of the disks passed to the **mmfsctl include** command are not marked as **excluded** in the **mmsdrfs** file.

**User response:** Verify the list of disks supplied to this command.

**6027-758 Disk(s) not started; disk *name* has a bad volume label.**

**Explanation:** The volume label on the disk does not match that expected by GPFS.

**User response:** Check the disk hardware. For hot-pluggable drives, make sure the proper drive has been plugged in.

**6027-759 *fileSystem* is still in use.**

**Explanation:** The **mmfsctl include** command found that the named file system is still mounted, or another GPFS command is running against the file system.

**User response:** Unmount the file system if it is mounted, or wait for GPFS commands in progress to terminate before retrying the command.

**6027-760 [E] Unable to perform i/o to the disk. This node is either fenced from accessing the disk or this node's disk lease has expired.**

**Explanation:** A read or write to the disk failed due to either being fenced from the disk or no longer having a disk lease.

**User response:** Verify disk hardware fencing setup is correct if being used. Ensure network connectivity between this node and other nodes is operational.

**6027-761 [W] Attention: excessive timer drift between *node* and *node* (*number* over *number* sec).**

**Explanation:** GPFS has detected an unusually large difference in the rate of clock ticks (as returned by the **times()** system call) between two nodes. Another node's TOD clock and tick rate changed dramatically relative to this node's TOD clock and tick rate.

**User response:** Check error log for hardware or device driver problems that might cause timer interrupts to be lost or a recent large adjustment made to the TOD clock.

**6027-762 No quota enabled file system found.**

**Explanation:** There is no quota-enabled file system in this cluster.

**User response:** None. Informational message only.

**6027-763 uidInvalidate: Incorrect option *option*.**

**Explanation:** An incorrect option passed to the **uidinvalidate** command.

**User response:** Correct the command invocation.

**6027-764 Error invalidating UID remapping cache for *domain*.**

**Explanation:** An incorrect domain name passed to the **uidinvalidate** command.

**User response:** Correct the command invocation.

---

**6027-765 [W]** Tick value hasn't changed for nearly *number* seconds

**Explanation:** Clock ticks incremented by AIX have not been incremented.

**User response:** Check the error log for hardware or device driver problems that might cause timer interrupts to be lost.

---

**6027-766 [N]** This node will be expelled from cluster *cluster* due to expel msg from *node*

**Explanation:** This node is being expelled from the cluster.

**User response:** Check the network connection between this node and the node specified above.

---

**6027-767 [N]** Request sent to *node* to expel *node* from cluster *cluster*

**Explanation:** This node sent an expel request to the cluster manager node to expel another node.

**User response:** Check network connection between this node and the node specified above.

---

**6027-768** Wrong number of operands for **mmpmon** command '*command*'.

**Explanation:** The command read from the input file has the wrong number of operands.

**User response:** Correct the command invocation and reissue the command.

---

**6027-769** Malformed **mmpmon** command '*command*'.

**Explanation:** The command read from the input file is malformed, perhaps with an unknown keyword.

**User response:** Correct the command invocation and reissue the command.

---

**6027-770** Error writing **user.quota** file.

**Explanation:** An error occurred while writing the cited quota file.

**User response:** Check the status and availability of the disks and reissue the command.

---

**6027-771** Error writing **group.quota** file.

**Explanation:** An error occurred while writing the cited quota file.

**User response:** Check the status and availability of the disks and reissue the command.

---

**6027-772** Error writing **fileset.quota** file.

**Explanation:** An error occurred while writing the cited quota file.

**User response:** Check the status and availability of the disks and reissue the command.

---

**6027-774** *fileSystem*: **quota management is not enabled, or one or more quota clients are not available.**

**Explanation:** An attempt was made to perform quotas commands without quota management enabled, or one or more quota clients failed during quota check.

**User response:** Correct the cause of the problem, and then reissue the quota command.

---

**6027-775** During **mmcheckquota** processing, *number* **node(s)** failed. It is recommended that **mmcheckquota** be repeated.

**Explanation:** Nodes failed while an online quota check was running.

**User response:** Reissue the quota check command.

---

**6027-776** *fileSystem*: **There was not enough space for the report. Please repeat quota check!**

**Explanation:** The **vflag** is set in the **tscheckquota** command, but either no space or not enough space could be allocated for the differences to be printed.

**User response:** Correct the space problem and reissue the quota check.

---

**6027-777 [I]** Recovering nodes: *nodeList*

**Explanation:** Recovery for one or more nodes has begun.

**User response:** No response is needed if this message is followed by 'recovered nodes' entries specifying the nodes. If this message is not followed by such a message, determine why recovery did not complete.

---

**6027-778 [I]** Recovering nodes in cluster *cluster*: *nodeList*

**Explanation:** Recovery for one or more nodes in the cited cluster has begun.

**User response:** No response is needed if this message is followed by 'recovered nodes' entries on the cited cluster specifying the nodes. If this message is not followed by such a message, determine why recovery did not complete.

---

**6027-779**     **Incorrect fileset name** *filesetName*.

**Explanation:** The fileset name provided on the command line is incorrect.

**User response:** Correct the fileset name and reissue the command.

---

**6027-780**     **Incorrect path to fileset junction** *junctionName*.

**Explanation:** The path to the fileset junction is incorrect.

**User response:** Correct the junction path and reissue the command.

---

**6027-781**     **Storage pools have not been enabled for file system** *fileSystem*.

**Explanation:** The user invoked a command with a storage pool option (-p or -P) before storage pools were enabled.

**User response:** Enable storage pools with the **mmchfs -V** command, or correct the command invocation and reissue the command.

---

**6027-784 [E]**   **Device not ready.**

**Explanation:** A device is not ready for operation.

**User response:** Check previous messages for further information.

---

**6027-785 [E]**   **Cannot establish connection.**

**Explanation:** This node cannot establish a connection to another node.

**User response:** Check previous messages for further information.

---

**6027-786 [E]**   **Message failed because the destination node refused the connection.**

**Explanation:** This node sent a message to a node that refuses to establish a connection.

**User response:** Check previous messages for further information.

---

**6027-787 [E]**   **Security configuration data is inconsistent or unavailable.**

**Explanation:** There was an error configuring security on this node.

**User response:** Check previous messages for further information.

---

**6027-788 [E]**   **Failed to load or initialize security library.**

**Explanation:** There was an error loading or initializing the security library on this node.

**User response:** Check previous messages for further information.

---

**6027-789**     **Unable to read offsets** *offset to offset for inode inode snap snap, from disk diskName, sector sector*.

**Explanation:** The **mmdeldisk -c** command found that the cited addresses on the cited disk represent data that is no longer readable.

**User response:** Save this output for later use in cleaning up failing disks.

---

**6027-790**     **Specified storage pool** *poolName* **does not match disk** *diskName* **storage pool** *poolName*. **Use mmdeldisk and mmadddisk to change a disk's storage pool.**

**Explanation:** An attempt was made to change a disk's storage pool assignment using the **mmchdisk** command. This can only be done by deleting the disk from its current storage pool and then adding it to the new pool.

**User response:** Delete the disk from its current storage pool and then add it to the new pool.

---

**6027-792**     **Policies have not been enabled for file system** *fileSystem*.

**Explanation:** The cited file system must be upgraded to use policies.

**User response:** Upgrade the file system via the **mmchfs -V** command.

---

**6027-793**     **No policy file was installed for file system** *fileSystem*.

**Explanation:** No policy file was installed for this file system.

**User response:** Install a policy file.

---

**6027-794**     **Failed to read policy file for file system** *fileSystem*.

**Explanation:** Failed to read the policy file for the requested file system.

**User response:** Reinstall the policy file.

---

---

**6027-795** Failed to open *fileName: errorCode*.

**Explanation:** An incorrect file name was specified to **tschpolicy**.

**User response:** Correct the command invocation and reissue the command.

---

**6027-796** Failed to read *fileName: errorCode*.

**Explanation:** An incorrect file name was specified to **tschpolicy**.

**User response:** Correct the command invocation and reissue the command.

---

**6027-797** Failed to stat *fileName: errorCode*.

**Explanation:** An incorrect file name was specified to **tschpolicy**.

**User response:** Correct the command invocation and reissue the command.

---

**6027-798** Policy files are limited to *number* bytes.

**Explanation:** A user-specified policy file exceeded the maximum-allowed length.

**User response:** Install a smaller policy file.

---

**6027-799** Policy *'policyName'* installed and broadcast to all nodes.

**Explanation:** Self-explanatory.

**User response:** None. Informational message only.

---

**6027-850** Unable to issue this command from a non-root user.

**Explanation:** **tsiostat** requires root privileges to run.

**User response:** Get the system administrator to change the executable to set the UID to 0.

---

**6027-851** Unable to process interrupt received.

**Explanation:** An interrupt occurred that **tsiostat** cannot process.

**User response:** Contact the IBM Support Center.

---

**6027-852** interval and count must be positive integers.

**Explanation:** Incorrect values were supplied for **tsiostat** parameters.

**User response:** Correct the command invocation and reissue the command.

---

**6027-853** interval must be less than 1024.

**Explanation:** An incorrect value was supplied for the interval parameter.

**User response:** Correct the command invocation and reissue the command.

---

**6027-854** count must be less than 1024.

**Explanation:** An incorrect value was supplied for the count parameter.

**User response:** Correct the command invocation and reissue the command.

---

**6027-855** Unable to connect to server, **mmfsd** is not started.

**Explanation:** The **tsiostat** command was issued but the file system is not started.

**User response:** Contact your system administrator.

---

**6027-856** No information to report.

**Explanation:** The **tsiostat** command was issued but no file systems are mounted.

**User response:** Contact your system administrator.

---

**6027-857** Error retrieving values.

**Explanation:** The **tsiostat** command was issued and an internal error occurred.

**User response:** Contact the IBM Support Center.

---

**6027-858** File system not mounted.

**Explanation:** The requested file system is not mounted.

**User response:** Mount the file system and reattempt the failing operation.

---

**6027-859** Set **DIRECTIO** failed

**Explanation:** The **tsfattr** call failed.

**User response:** Check for additional error messages. Resolve the problems before reattempting the failing operation.

---

**6027-860** **-d** is not appropriate for an NFSv4 ACL

**Explanation:** Produced by the **mmgetacl** or **mmputacl** commands when the **-d** option was specified, but the object has an NFS Version 4 ACL (does not have a default).

**User response:** None. Informational message only.

---

---

**6027-861 Set afm ctl failed**

**Explanation:** The `tsfattr` call failed.

**User response:** Check for additional error messages. Resolve the problems before reattempting the failing operation.

---

**6027-862 Incorrect storage pool name *poolName*.**

**Explanation:** An incorrect storage pool name was provided.

**User response:** Determine the correct storage pool name and reissue the command.

---

**6027-863 File cannot be assigned to storage pool '*poolName*'.**

**Explanation:** The file cannot be assigned to the specified pool.

**User response:** Determine the correct storage pool name and reissue the command.

---

**6027-864 Set storage pool failed.**

**Explanation:** An incorrect storage pool name was provided.

**User response:** Determine the correct storage pool name and reissue the command.

---

**6027-865 Restripe file data failed.**

**Explanation:** An error occurred while restriping the file data.

**User response:** Check the error code and reissue the command.

---

**6027-866 [E] Storage pools have not been enabled for this file system.**

**Explanation:** The user invoked a command with a storage pool option (`-p` or `-P`) before storage pools were enabled.

**User response:** Enable storage pools via `mmchfs -V`, or correct the command invocation and reissue the command.

---

**6027-867 Change storage pool is not permitted.**

**Explanation:** The user tried to change a file's assigned storage pool but was not root or superuser.

**User response:** Reissue the command as root or superuser.

---

**6027-868 mmchattr failed.**

**Explanation:** An error occurred while changing a file's attributes.

**User response:** Check the error code and reissue the command.

---

**6027-869 File replication exceeds number of failure groups in destination storage pool.**

**Explanation:** The `tschattr` command received incorrect command line arguments.

**User response:** Correct the command invocation and reissue the command.

---

**6027-870 [E] Error on `getcwd()`: *errorString*. Try an absolute path instead of just *pathName***

**Explanation:** The `getcwd` system call failed.

**User response:** Specify an absolute path starting with `/'` on the command invocation, so that the command will not need to invoke `getcwd`.

---

**6027-871 [E] Error on `gpfs_get_pathname_from_fssnaphandle(pathName): errorString`.**

**Explanation:** An error occurred during a `gpfs_get_pathname_from_fssnaphandle` operation.

**User response:** Verify the invocation parameters and make sure the command is running under a user ID with sufficient authority (**root** or administrator privileges). Specify a GPFS file system device name or a GPFS directory path name as the first argument. Correct the command invocation and reissue the command.

---

**6027-872 [E] *pathName* is not within a mounted GPFS file system.**

**Explanation:** An error occurred while attempting to access the named GPFS file system or path.

**User response:** Verify the invocation parameters and make sure the command is running under a user ID with sufficient authority (**root** or administrator privileges). Mount the GPFS file system. Correct the command invocation and reissue the command.

---

**6027-873 [W] Error on `gpfs_stat_inode(pathName/fileName),inodeNumber,genNumber): errorString`**

**Explanation:** An error occurred during a `gpfs_stat_inode` operation.

**User response:** Reissue the command. If the problem persists, contact the IBM Support Center.

---

---

**6027-874 [E] Error: incorrect Date@Time (YYYY-MM-DD@HH:MM:SS) specification: specification**

**Explanation:** The *Date@Time* command invocation argument could not be parsed.

**User response:** Correct the command invocation and try again. The syntax should look similar to: 2005-12-25@07:30:00.

---

**6027-875 [E] Error on gpfs\_stat(pathName): errorString**

**Explanation:** An error occurred while attempting to **stat()** the cited path name.

**User response:** Determine whether the cited path name exists and is accessible. Correct the command arguments as necessary and reissue the command.

---

**6027-876 [E] Error starting directory scan(pathName): errorString**

**Explanation:** The specified path name is not a directory.

**User response:** Determine whether the specified path name exists and is an accessible directory. Correct the command arguments as necessary and reissue the command.

---

**6027-877 [E] Error opening pathName: errorString**

**Explanation:** An error occurred while attempting to open the named file. Its pool and replication attributes remain unchanged.

**User response:** Investigate the file and possibly reissue the command. The file may have been removed or locked by another application.

---

**6027-878 [E] Error on gpfs\_fcntl(pathName): errorString (offset=offset)**

**Explanation:** An error occurred while attempting **fcntl** on the named file. Its pool or replication attributes may not have been adjusted.

**User response:** Investigate the file and possibly reissue the command. Use the **mmlsattr** and **mmchattr** commands to examine and change the pool and replication attributes of the named file.

---

**6027-879 [E] Error deleting pathName: errorString**

**Explanation:** An error occurred while attempting to delete the named file.

**User response:** Investigate the file and possibly reissue the command. The file may have been removed or locked by another application.

---

**6027-880 Error on gpfs\_seek\_inode(inodeNumber): errorString**

**Explanation:** An error occurred during a **gpfs\_seek\_inode** operation.

**User response:** Reissue the command. If the problem persists, contact the contact the IBM Support Center

---

**6027-881 [E] Error on gpfs\_iopen([rootPath/pathName],inodeNumber): errorString**

**Explanation:** An error occurred during a **gpfs\_iopen** operation.

**User response:** Reissue the command. If the problem persists, contact the IBM Support Center.

---

**6027-882 [E] Error on gpfs\_ireaddir(rootPath/pathName): errorString**

**Explanation:** An error occurred during a **gpfs\_ireaddir()** operation.

**User response:** Reissue the command. If the problem persists, contact the IBM Support Center.

---

**6027-883 Error on gpfs\_next\_inode(maxInodeNumber): errorString**

**Explanation:** An error occurred during a **gpfs\_next\_inode** operation.

**User response:** Reissue the command. If the problem persists, contact the IBM Support Center.

---

**6027-884 [E:nnn] Error during directory scan**

**Explanation:** A terminal error occurred during the directory scan phase of the command.

**User response:** Verify the command arguments. Reissue the command. If the problem persists, contact the IBM Support Center.

---

**6027-885 [E:nnn] Error during inode scan: errorString**

**Explanation:** A terminal error occurred during the inode scan phase of the command.

**User response:** Verify the command arguments. Reissue the command. If the problem persists, contact the IBM Support Center.

---

**6027-886 [E:nnn] Error during policy decisions scan**

**Explanation:** A terminal error occurred during the policy decisions phase of the command.

**User response:** Verify the command arguments. Reissue the command. If the problem persists, contact the IBM Support Center.



---

**6027-887 [W] Error on gpfs\_igetstoragepool(*dataPoolId*):**  
*errorString*

**Explanation:** An error occurred during a `gpfs_igetstoragepool` operation. Possible inode corruption.

**User response:** Use `mmfsck` command. If the problem persists, contact the IBM Support Center.

---

**6027-888 [W] Error on gpfs\_igetfilesetName(*filesetId*):**  
*errorString*

**Explanation:** An error occurred during a `gpfs_igetfilesetName` operation. Possible inode corruption.

**User response:** Use `mmfsck` command. If the problem persists, contact the IBM Support Center.

---

**6027-889 [E] Error on**  
**gpfs\_get\_fssnaphandle(*rootPath*):**  
*errorString*.

**Explanation:** An error occurred during a `gpfs_get_fssnaphandle` operation.

**User response:** Reissue the command. If the problem persists, contact the IBM Support Center.

---

**6027-890 [E] Error on gpfs\_open\_inodescan(*rootPath*):**  
*errorString*

**Explanation:** An error occurred during a `gpfs_open_inodescan()` operation.

**User response:** Reissue the command. If the problem persists, contact the IBM Support Center.

---

**6027-891 [X] WEIGHT(*thresholdValue*) UNKNOWN**  
*pathName*

**Explanation:** The named file was assigned the indicated weight, but the rule type is UNKNOWN.

**User response:** Contact the IBM Support Center.

---

**6027-892 [E] Error on pthread\_create: where**  
**#threadNumber\_or\_portNumber\_or\_**  
**socketNumber: errorString**

**Explanation:** An error occurred while creating the thread during a `pthread_create` operation.

**User response:** Consider some of the command parameters that might affect memory usage. For further assistance, contact the IBM Support Center.

---



---

**6027-893 [X] Error on pthread\_mutex\_init: errorString**

**Explanation:** An error occurred during a `pthread_mutex_init` operation.

**User response:** Contact the IBM Support Center.

---

**6027-894 [X] Error on pthread\_mutex\_lock: errorString**

**Explanation:** An error occurred during a `pthread_mutex_lock` operation.

**User response:** Contact the IBM Support Center.

---

**6027-895 [X] Error on pthread\_mutex\_unlock:**  
*errorString*

**Explanation:** An error occurred during a `pthread_mutex_unlock` operation.

**User response:** Contact the IBM Support Center.

---

**6027-896 [X] Error on pthread\_cond\_init: errorString**

**Explanation:** An error occurred during a `pthread_cond_init` operation.

**User response:** Contact the IBM Support Center.

---

**6027-897 [X] Error on pthread\_cond\_signal: errorString**

**Explanation:** An error occurred during a `pthread_cond_signal` operation.

**User response:** Contact the IBM Support Center.

---

**6027-898 [X] Error on pthread\_cond\_broadcast:**  
*errorString*

**Explanation:** An error occurred during a `pthread_cond_broadcast` operation.

**User response:** Contact the IBM Support Center.

---

**6027-899 [X] Error on pthread\_cond\_wait: errorString**

**Explanation:** An error occurred during a `pthread_cond_wait` operation.

**User response:** Contact the IBM Support Center.

---

**6027-900 [E] Error opening work file *fileName*:**  
*errorString*

**Explanation:** An error occurred while attempting to open the named work file.

**User response:** Investigate the file and possibly reissue the command. Check that the path name is defined and accessible.

---

---

**6027-901 [E] Error writing to work file** *fileName: errorString*

**Explanation:** An error occurred while attempting to write to the named work file.

**User response:** Investigate the file and possibly reissue the command. Check that there is sufficient free space in the file system.

---

**6027-902 [E] Error parsing work file** *fileName. Service index: number*

**Explanation:** An error occurred while attempting to read the specified work file.

**User response:** Investigate the file and possibly reissue the command. Make sure that there is enough free space in the file system. If the error persists, contact the IBM Support Center.

---

**6027-903 [E:nnn] Error while loading policy rules.**

**Explanation:** An error occurred while attempting to read or parse the policy file, which may contain syntax errors. Subsequent messages include more information about the error.

**User response:** Read all of the related error messages and try to correct the problem.

---

**6027-904 [E] Error** *returnCode* **from PD writer for inode=inodeNumber pathname=pathName**

**Explanation:** An error occurred while writing the policy decision for the candidate file with the indicated inode number and path name to a work file. There probably will be related error messages.

**User response:** Read all the related error messages. Attempt to correct the problems.

---

**6027-905 [E] Error: Out of memory. Service index: number**

**Explanation:** The command has exhausted virtual memory.

**User response:** Consider some of the command parameters that might affect memory usage. For further assistance, contact the IBM Support Center.

---

**6027-906 [E:nnn] Error on system(command)**

**Explanation:** An error occurred during the system call with the specified argument string.

**User response:** Read and investigate related error messages.

---



---

**6027-907 [E:nnn] Error from sort\_file(inodeListname, sortCommand,sortInodeOptions,tempDir)**

**Explanation:** An error occurred while sorting the named work file using the named **sort** command with the given options and working directory.

**User response:** Check these:

- The **sort** command is installed on your system.
  - The **sort** command supports the given options.
  - The working directory is accessible.
  - The file system has sufficient free space.
- 

**6027-908 [W] Attention: In RULE 'ruleName' (ruleNumber), the pool named by "poolName 'poolType'" is not defined in the file system.**

**Explanation:** The cited pool is not defined in the file system.

**User response:** Correct the rule and reissue the command.

This is not an irrecoverable error; the command will continue to run. Of course it will not find any files in an incorrect **FROM POOL** and it will not be able to migrate any files to an incorrect **TO POOL**.

---

**6027-909 [E] Error on pthread\_join: where #threadNumber: errorString**

**Explanation:** An error occurred while reaping the thread during a **pthread\_join** operation.

**User response:** Contact the IBM Support Center.

---

**6027-910 [E:nnn] Error during policy execution**

**Explanation:** A terminating error occurred during the policy execution phase of the command.

**User response:** Verify the command arguments and reissue the command. If the problem persists, contact the IBM Support Center.

---

**6027-911 [E] Error on changeSpecification change for pathName. errorString**

**Explanation:** This message provides more details about a **gpfs\_fcntl()** error.

**User response:** Use the **mmlsattr** and **mmchattr** commands to examine the file, and then reissue the change command.

---

**6027-912 [E] Error on restriping of pathName. errorString**

**Explanation:** This provides more details on a **gpfs\_fcntl()** error.

**User response:** Use the **mmlsattr** and **mmchattr**

commands to examine the file and then reissue the restriping command.

---

**6027-913**      **Desired replication exceeds number of failure groups.**

**Explanation:** While restriping a file, the **tschattr** or **tsrestripefile** command found that the desired replication exceeded the number of failure groups.

**User response:** Reissue the command after adding or restarting file system disks.

---

**6027-914**      **Insufficient space in one of the replica failure groups.**

**Explanation:** While restriping a file, the **tschattr** or **tsrestripefile** command found there was insufficient space in one of the replica failure groups.

**User response:** Reissue the command after adding or restarting file system disks.

---

**6027-915**      **Insufficient space to properly balance file.**

**Explanation:** While restriping a file, the **tschattr** or **tsrestripefile** command found that there was insufficient space to properly balance the file.

**User response:** Reissue the command after adding or restarting file system disks.

---

**6027-916**      **Too many disks unavailable to properly balance file.**

**Explanation:** While restriping a file, the **tschattr** or **tsrestripefile** command found that there were too many disks unavailable to properly balance the file.

**User response:** Reissue the command after adding or restarting file system disks.

---

**6027-917**      **All replicas of a data block were previously deleted.**

**Explanation:** While restriping a file, the **tschattr** or **tsrestripefile** command found that all replicas of a data block were previously deleted.

**User response:** Reissue the command after adding or restarting file system disks.

---

**6027-918**      **Cannot make this change to a nonzero length file.**

**Explanation:** GPFS does not support the requested change to the replication attributes.

**User response:** You may want to create a new file with the desired attributes and then copy your data to that file and rename it appropriately. Be sure that there are sufficient disks assigned to the pool with different

failure groups to support the desired replication attributes.

---

**6027-919**      **Replication parameter range error (*value*, *value*).**

**Explanation:** Similar to message **6027-918**. The (a,b) numbers are the allowable range of the replication attributes.

**User response:** You may want to create a new file with the desired attributes and then copy your data to that file and rename it appropriately. Be sure that there are sufficient disks assigned to the pool with different failure groups to support the desired replication attributes.

---

**6027-920 [E]**    **Error on pthread\_detach(self): *where:* *errorString***

**Explanation:** An error occurred during a **pthread\_detach** operation.

**User response:** Contact the IBM Support Center.

---

**6027-921 [E]**    **Error on socket *socketName(hostName):* *errorString***

**Explanation:** An error occurred during a socket operation.

**User response:** Verify any command arguments related to interprocessor communication and then reissue the command. If the problem persists, contact the IBM Support Center.

---

**6027-922 [X]**    **Error in Mtconx - p\_accepts should not be empty**

**Explanation:** The program discovered an inconsistency or logic error within itself.

**User response:** Contact the IBM Support Center.

---

**6027-923 [W]**    **Error - command client is an incompatible version: *hostName* *protocolVersion***

**Explanation:** While operating in master/client mode, the command discovered that the client is running an incompatible version.

**User response:** Ensure the same version of the command software is installed on all nodes in the clusters and then reissue the command.

---

**6027-924 [X]**    **Error - unrecognized client response from *hostName:* *clientResponse***

**Explanation:** Similar to message **6027-923**, except this may be an internal logic error.

**User response:** Ensure the latest, same version

software is installed on all nodes in the clusters and then reissue the command. If the problem persists, contact the IBM Support Center.

---

**6027-925**     **Directory cannot be assigned to storage pool 'poolName'.**

**Explanation:** The file cannot be assigned to the specified pool.

**User response:** Determine the correct storage pool name and reissue the command.

---

**6027-926**     **Symbolic link cannot be assigned to storage pool 'poolName'.**

**Explanation:** The file cannot be assigned to the specified pool.

**User response:** Determine the correct storage pool name and reissue the command.

---

**6027-927**     **System file cannot be assigned to storage pool 'poolName'.**

**Explanation:** The file cannot be assigned to the specified pool.

**User response:** Determine the correct storage pool name and reissue the command.

---

**6027-928 [E]**     **Error: filesystem/device *fileSystem* has no snapshot with name *snapshotName*.**

**Explanation:** The specified file system does not have a snapshot with the specified snapshot name.

**User response:** Use the `mmssnapshot` command to list the snapshot names for the file system.

---

**6027-929 [W]**     **Attention: In RULE '*ruleName*' (*ruleNumber*), both pools '*poolName*' and '*poolName*' are EXTERNAL. This is not a supported migration.**

**Explanation:** The command does not support migration between two EXTERNAL pools.

**User response:** Correct the rule and reissue the command.

**Note:** This is not an unrecoverable error. The command will continue to run.

---

**6027-930 [W]**     **Attention: In RULE '*ruleName*' LIST name '*listName*' appears, but there is no corresponding EXTERNAL LIST '*listName*' EXEC ... OPTS ... rule to specify a program to process the matching files.**

**Explanation:** There should be an EXTERNAL LIST rule for every list named by your LIST rules.

**User response:** Add an "EXTERNAL LIST *listName* EXEC *scriptName* OPTS *opts*" rule.

**Note:** This is not an unrecoverable error. For execution with `-I defer`, file lists are generated and saved, so EXTERNAL LIST rules are not strictly necessary for correct execution.

---

**6027-931 [E]**     **Error - The policy evaluation phase did not complete.**

**Explanation:** One or more errors prevented the policy evaluation phase from examining all of the files.

**User response:** Consider other messages emitted by the command. Take appropriate action and then reissue the command.

---

**6027-932 [E]**     **Error - The policy execution phase did not complete.**

**Explanation:** One or more errors prevented the policy execution phase from operating on each chosen file.

**User response:** Consider other messages emitted by the command. Take appropriate action and then reissue the command.

---

**6027-933 [W]**     **EXEC '*wouldbeScriptPathname*' of EXTERNAL POOL or LIST '*PoolOrListName*' fails TEST with code *scriptReturnCode* on this node.**

**Explanation:** Each EXEC defined in an EXTERNAL POOL or LIST rule is run in TEST mode on each node. Each invocation that fails with a nonzero return code is reported. Command execution is terminated on any node that fails any of these tests.

**User response:** Correct the EXTERNAL POOL or LIST rule, the EXEC script, or do nothing because this is not necessarily an error. The administrator may suppress execution of the `mmapplypolicy` command on some nodes by deliberately having one or more EXECs return nonzero codes.

---

**6027-934 [W]**     **Attention: Specified snapshot: '*SnapshotName*' will be ignored because the path specified: '*PathName*' is not within that snapshot.**

**Explanation:** The command line specified both a path name to be scanned and a snapshot name, but the snapshot name was not consistent with the path name.

**User response:** If you wanted the entire snapshot, just specify the GPFS file system name or device name. If you wanted a directory within a snapshot, specify a path name within that snapshot (for example, `/gpfs/FileSystemName/.snapshots/SnapShotName/Directory`).

---

**6027-935 [W] Attention: In RULE 'ruleName' (ruleNumber) LIMIT or REPLICATE clauses are ignored; not supported for migration to EXTERNAL pool 'storagePoolName'.**

**Explanation:** GPFS does not support the **LIMIT** or **REPLICATE** clauses during migration to external pools.

**User response:** Correct the policy rule to avoid this warning message.

---

**6027-936 [W] Error - command master is an incompatible version.**

**Explanation:** While operating in master/client mode, the command discovered that the master is running an incompatible version.

**User response:** Upgrade the command software on all nodes and reissue the command.

---

**6027-937 [E] Error creating shared temporary sub-directory subDirName: subDirPath**

**Explanation:** The **mkdir** command failed on the named subdirectory path.

**User response:** Specify an existing writable shared directory as the shared temporary directory argument to the policy command. The policy command will create a subdirectory within that.

---

**6027-938 [E] Error closing work file fileName: errorString**

**Explanation:** An error occurred while attempting to close the named work file or socket.

**User response:** Record the above information. Contact the IBM Support Center.

---

**6027-939 [E] Error on gpfs\_quotactl(pathName,commandCode,resourceId): errorString**

**Explanation:** An error occurred while attempting **gpfs\_quotactl()**.

**User response:** Correct the policy rules and/or enable GPFS quota tracking. If problem persists contact the IBM Support Center.

---

**6027-940 Open failed.**

**Explanation:** The **open()** system call was not successful.

**User response:** Check additional error messages.

---

**6027-941 Set replication failed.**

**Explanation:** The **open()** system call was not successful.

**User response:** Check additional error messages.

---

**6027-943 -M and -R are only valid for zero length files.**

**Explanation:** The **mmchattr** command received command line arguments that were not valid.

**User response:** Correct command line and reissue the command.

---

**6027-944 -m value exceeds number of failure groups for metadata.**

**Explanation:** The **mmchattr** command received command line arguments that were not valid.

**User response:** Correct command line and reissue the command.

---

**6027-945 -r value exceeds number of failure groups for data.**

**Explanation:** The **mmchattr** command received command line arguments that were not valid.

**User response:** Correct command line and reissue the command.

---

**6027-946 Not a regular file or directory.**

**Explanation:** An **mmlsattr** or **mmchattr** command error occurred.

**User response:** Correct the problem and reissue the command.

---

**6027-947 Stat failed: A file or directory in the path name does not exist.**

**Explanation:** A file or directory in the path name does not exist.

**User response:** Correct the problem and reissue the command.

---

**6027-948 [E:nnn] fileName: get clone attributes failed: errorString**

**Explanation:** The **tsfattr** call failed.

**User response:** Check for additional error messages. Resolve the problems before reattempting the failing operation.

---

---

6027-949 [E] *fileName*: invalid clone attributes.

**Explanation:** Self explanatory.

**User response:** Check for additional error messages. Resolve the problems before reattempting the failing operation.

---

6027-950 [E:nnn] File cloning requires the 'fastea' feature to be enabled.

**Explanation:** The file system **fastea** feature is not enabled.

**User response:** Enable the **fastea** feature by issuing the **mmchfs -V** and **mmmigratefs --fastea** commands.

---

6027-951 [E] Error on *operationName* to work file *fileName*: *errorString*

**Explanation:** An error occurred while attempting to do a (write-like) operation on the named work file.

**User response:** Investigate the file and possibly reissue the command. Check that there is sufficient free space in the file system.

---

6027-953 Failed to get a handle for fileset *filesetName*, snapshot *snapshotName* in file system *fileSystem*. *errorMessage*.

**Explanation:** Failed to get a handle for a specific fileset snapshot in the file system.

**User response:** Correct the command line and reissue the command. If the problem persists, contact the IBM Support Center.

---

6027-954 Failed to get the maximum inode number in the active file system. *errorMessage*.

**Explanation:** Failed to get the maximum inode number in the current active file system.

**User response:** Correct the command line and reissue the command. If the problem persists, contact the IBM Support Center.

---

6027-955 Failed to set the maximum allowed memory for the specified *fileSystem* command.

**Explanation:** Failed to set the maximum allowed memory for the specified command.

**User response:** Correct the command line and reissue the command. If the problem persists, contact the IBM Support Center.

---

6027-956 Cannot allocate enough buffer to record different items.

**Explanation:** Cannot allocate enough buffer to record different items which are used in the next phase.

**User response:** Correct the command line and reissue the command. If the problem persists, contact the system administrator.

---

6027-957 Failed to get the root directory inode of fileset *filesetName*

**Explanation:** Failed to get the root directory inode of a fileset.

**User response:** Correct the command line and reissue the command. If the problem persists, contact the IBM Support Center.

---

6027-959 '*fileName*' is not a regular file.

**Explanation:** Only regular files are allowed to be clone parents.

**User response:** This file is not a valid target for **mmclone** operations.

---

6027-960 cannot access '*fileName*': *errorString*.

**Explanation:** This message provides more details about a **stat()** error.

**User response:** Correct the problem and reissue the command.

---

6027-961 Cannot execute *command*.

**Explanation:** The **mmeditACL** command cannot invoke the **mmgetACL** or **mmputACL** command.

**User response:** Contact your system administrator.

---

6027-962 Failed to list fileset *filesetName*.

**Explanation:** Failed to list specific fileset.

**User response:** None.

---

6027-963 EDITOR environment variable not set

**Explanation:** Self-explanatory.

**User response:** Set the EDITOR environment variable and reissue the command.

---

6027-964 EDITOR environment variable must be an absolute path name

**Explanation:** Self-explanatory.

**User response:** Set the EDITOR environment variable correctly and reissue the command.

---

---

**6027-965 Cannot create temporary file****Explanation:** Self-explanatory.**User response:** Contact your system administrator.

---

**6027-966 Cannot access *fileName*****Explanation:** Self-explanatory.**User response:** Verify file permissions.

---

**6027-967 Should the modified ACL be applied? (yes) or (no)****Explanation:** Self-explanatory.**User response:** Respond **yes** if you want to commit the changes, **no** otherwise.

---

**6027-971 Cannot find *fileName*****Explanation:** Self-explanatory.**User response:** Verify the file name and permissions.

---

**6027-972 *name* is not a directory (-d not valid).****Explanation:** Self-explanatory.**User response:** None, only directories are allowed to have default ACLs.

---

**6027-973 Cannot allocate *number* byte buffer for ACL.****Explanation:** There was not enough available memory to process the request.**User response:** Contact your system administrator.

---

**6027-974 Failure reading ACL (rc=*number*).****Explanation:** An unexpected error was encountered by `mmgetacl` or `mmeditACL`.**User response:** Examine the return code, contact the IBM Support Center if necessary.

---

**6027-976 Failure writing ACL (rc=*number*).****Explanation:** An unexpected error encountered by `mmpuACL` or `mmeditACL`.**User response:** Examine the return code, Contact the IBM Support Center if necessary.

---

**6027-977 Authorization failure****Explanation:** An attempt was made to create or modify the ACL for a file that you do not own.**User response:** Only the owner of a file or the root

user can create or change the access control list for a file.

---

**6027-978 Incorrect, duplicate, or missing access control entry detected.****Explanation:** An access control entry in the ACL that was created had incorrect syntax, one of the required access control entries is missing, or the ACL contains duplicate access control entries.**User response:** Correct the problem and reissue the command.

---

**6027-979 Incorrect ACL entry: *entry*.****Explanation:** Self-explanatory.**User response:** Correct the problem and reissue the command.

---

**6027-980 *name* is not a valid user name.****Explanation:** Self-explanatory.**User response:** Specify a valid user name and reissue the command.

---

**6027-981 *name* is not a valid group name.****Explanation:** Self-explanatory.**User response:** Specify a valid group name and reissue the command.

---

**6027-982 *name* is not a valid ACL entry type.****Explanation:** Specify a valid ACL entry type and reissue the command.**User response:** Correct the problem and reissue the command.

---

**6027-983 *name* is not a valid permission set.****Explanation:** Specify a valid permission set and reissue the command.**User response:** Correct the problem and reissue the command.

---

**6027-985 An error was encountered while deleting the ACL (rc=*value*).****Explanation:** An unexpected error was encountered by `tsdelACL`.**User response:** Examine the return code and contact the IBM Support Center, if necessary.

---

**6027-986**      **Cannot open *fileName*.**

**Explanation:** Self-explanatory.

**User response:** Verify the file name and permissions.

---

**6027-987**      ***name* is not a valid special name.**

**Explanation:** Produced by the **mmputacl** command when the NFS V4 'special' identifier is followed by an unknown special id string. *name* is one of the following: 'owner@', 'group@', 'everyone@'.

**User response:** Specify a valid NFS V4 special name and reissue the command.

---

**6027-988**      ***type* is not a valid NFS V4 type.**

**Explanation:** Produced by the **mmputacl** command when the type field in an ACL entry is not one of the supported NFS Version 4 type values. *type* is one of the following: 'allow' or 'deny'.

**User response:** Specify a valid NFS V4 type and reissue the command.

---

**6027-989**      ***name* is not a valid NFS V4 flag.**

**Explanation:** A flag specified in an ACL entry is not one of the supported values, or is not valid for the type of object (inherit flags are valid for directories only). Valid values are **FileInherit**, **DirInherit**, and **InheritOnly**.

**User response:** Specify a valid NFS V4 option and reissue the command.

---

**6027-990**      **Missing permissions (*value found*, *value are required*).**

**Explanation:** The permissions listed are less than the number required.

**User response:** Add the missing permissions and reissue the command.

---

**6027-991**      **Combining **FileInherit** and **DirInherit** makes the mask ambiguous.**

**Explanation:** Produced by the **mmputacl** command when **WRITE/CREATE** is specified without **MKDIR** (or the other way around), and both the **FILE\_INHERIT** and **DIR\_INHERIT** flags are specified.

**User response:** Make separate **FileInherit** and **DirInherit** entries and reissue the command.

---

**6027-992**      **Subdirectory *name* already exists. Unable to create snapshot.**

**Explanation:** **tsbackup** was unable to create a snapshot because the snapshot subdirectory already exists. This condition sometimes is caused by issuing a

IBM Spectrum Protect restore operation without specifying a different subdirectory as the target of the restore.

**User response:** Remove or rename the existing subdirectory and then retry the command.

---

**6027-993**      **Keyword *aclType* is incorrect. Valid values are: 'posix', 'nfs4', 'native'.**

**Explanation:** One of the **mm\*acl** commands specified an incorrect value with the **-k** option.

**User response:** Correct the *aclType* value and reissue the command.

---

**6027-994**      **ACL permissions cannot be denied to the file owner.**

**Explanation:** The **mmputacl** command found that the **READ\_ACL**, **WRITE\_ACL**, **READ\_ATTR**, or **WRITE\_ATTR** permissions are explicitly being denied to the file owner. This is not permitted, in order to prevent the file being left with an ACL that cannot be modified.

**User response:** Do not select the **READ\_ACL**, **WRITE\_ACL**, **READ\_ATTR**, or **WRITE\_ATTR** permissions on **deny** ACL entries for the **OWNER**.

---

**6027-995**      **This command will run on a remote node, *nodeName*.**

**Explanation:** The **mmputacl** command was invoked for a file that resides on a file system in a remote cluster, and UID remapping is enabled. To parse the user and group names from the ACL file correctly, the command will be run transparently on a node in the remote cluster.

**User response:** None. Informational message only.

---

**6027-996 [E:nnn]**      **Error reading policy text from: *fileName***

**Explanation:** An error occurred while attempting to open or read the specified policy file. The policy file may be missing or inaccessible.

**User response:** Read all of the related error messages and try to correct the problem.

---

**6027-997 [W]**      **Attention: **RULE** '*ruleName*' attempts to redefine **EXTERNAL POOL** or **LIST literal** '*poolName*', ignored.**

**Explanation:** Execution continues as if the specified rule was not present.

**User response:** Correct or remove the policy rule.

---



---

**6027-998 [E] Error in FLR/PDR serving for client**  
*clientHostNameAndPortNumber:*  
**FLRs=***numOfFileListRecords*  
**PDRs=***numOfPolicyDecisionResponses*  
**pdrs=***numOfPolicyDecisionResponseRecords*

**Explanation:** A protocol error has been detected among cooperating **mmapplypolicy** processes.

**User response:** Reissue the command. If the problem persists, contact the IBM Support Center.

---

**6027-999 [E] Authentication failed:**  
*myNumericNetworkAddress* **with**  
*partnersNumericNetworkAddress*  
**(code=***codeIndicatingProtocolStepSequence*  
**rc=***errnoStyleErrorCode*)

**Explanation:** Two processes at the specified network addresses failed to authenticate. The cooperating processes should be on the same network; they should not be separated by a firewall.

**User response:** Correct the configuration and try the operation again. If the problem persists, contact the IBM Support Center.

---

**6027-1004 Incorrect [nodelist] format in file:**  
*nodeListLine*

**Explanation:** A [nodelist] line in the input stream is not a comma-separated list of nodes.

**User response:** Fix the format of the [nodelist] line in the **mmfs.cfg** input file. This is usually the *NodeFile* specified on the **mmchconfig** command.

If no user-specified [nodelist] lines are in error, contact the IBM Support Center.

If user-specified [nodelist] lines are in error, correct these lines.

---

**6027-1005 Common is not sole item on [] line**  
*number.*

**Explanation:** A [nodelist] line in the input stream contains common plus any other names.

**User response:** Fix the format of the [nodelist] line in the **mmfs.cfg** input file. This is usually the *NodeFile* specified on the **mmchconfig** command.

If no user-specified [nodelist] lines are in error, contact the IBM Support Center.

If user-specified [nodelist] lines are in error, correct these lines.

---

**6027-1006 Incorrect custom [ ] line number.**

**Explanation:** A [nodelist] line in the input stream is not of the format: [nodelist]. This covers syntax errors not covered by messages 6027-1004 and 6027-1005.

**User response:** Fix the format of the list of nodes in the **mmfs.cfg** input file. This is usually the *NodeFile* specified on the **mmchconfig** command.

If no user-specified lines are in error, contact the IBM Support Center.

If user-specified lines are in error, correct these lines.

---

**6027-1007 attribute found in common multiple times: attribute.**

**Explanation:** The attribute specified on the command line is in the main input stream multiple times. This is occasionally legal, such as with the trace attribute. These attributes, however, are not meant to be repaired by **mmfixcfg**.

**User response:** Fix the configuration file (**mmfs.cfg** or **mmfscfg1** in the SDR). All attributes modified by GPFS configuration commands may appear only once in common sections of the configuration file.

---

**6027-1008 Attribute found in custom multiple times: attribute.**

**Explanation:** The attribute specified on the command line is in a custom section multiple times. This is occasionally legal. These attributes are not meant to be repaired by **mmfixcfg**.

**User response:** Fix the configuration file (**mmfs.cfg** or **mmfscfg1** in the SDR). All attributes modified by GPFS configuration commands may appear only once in custom sections of the configuration file.

---

**6027-1022 Missing mandatory arguments on command line.**

**Explanation:** Some, but not enough, arguments were specified to the **mmcrfsc** command.

**User response:** Specify all arguments as per the usage statement that follows.

---

**6027-1023 File system size must be an integer:**  
*value*

**Explanation:** The first two arguments specified to the **mmcrfsc** command are not integers.

**User response:** File system size is an internal argument. The **mmcrfsc** command should never call the **mmcrfsc** command without a valid file system size argument. Contact the IBM Support Center.

---

**6027-1028** Incorrect value for *-name* flag.

**Explanation:** An incorrect argument was specified with an option that requires one of a limited number of allowable options (for example, *-s* or any of the **yes** | **no** options).

**User response:** Use one of the valid values for the specified option.

---

**6027-1029** Incorrect characters in integer field for *-name* option.

**Explanation:** An incorrect character was specified with the indicated option.

**User response:** Use a valid integer for the indicated option.

---

**6027-1030** Value below minimum for *-optionLetter* option. Valid range is from *value* to *value*

**Explanation:** The value specified with an option was below the minimum.

**User response:** Use an integer in the valid range for the indicated option.

---

**6027-1031** Value above maximum for option *-optionLetter*. Valid range is from *value* to *value*.

**Explanation:** The value specified with an option was above the maximum.

**User response:** Use an integer in the valid range for the indicated option.

---

**6027-1032** Incorrect option *optionName*.

**Explanation:** An unknown option was specified.

**User response:** Use only the options shown in the syntax.

---

**6027-1033** Option *optionName* specified twice.

**Explanation:** An option was specified more than once on the command line.

**User response:** Use options only once.

---

**6027-1034** Missing argument after *optionName* option.

**Explanation:** An option was not followed by an argument.

**User response:** All options need an argument. Specify one.

---

**6027-1035** Option *-optionName* is mandatory.

**Explanation:** A mandatory input option was not specified.

**User response:** Specify all mandatory options.

---

**6027-1036** Option expected at *string*.

**Explanation:** Something other than an expected option was encountered on the latter portion of the command line.

**User response:** Follow the syntax shown. Options may not have multiple values. Extra arguments are not allowed.

---

**6027-1038** IndirectSize must be  $\leq$  BlockSize and must be a multiple of LogicalSectorSize (512).

**Explanation:** The IndirectSize specified was not a multiple of 512 or the IndirectSize specified was larger than BlockSize.

**User response:** Use valid values for IndirectSize and BlockSize.

---

**6027-1039** InodeSize must be a multiple of LocalSectorSize (512).

**Explanation:** The specified InodeSize was not a multiple of 512.

**User response:** Use a valid value for InodeSize.

---

**6027-1040** InodeSize must be less than or equal to Blocksize.

**Explanation:** The specified InodeSize was not less than or equal to Blocksize.

**User response:** Use a valid value for InodeSize.

---

**6027-1042** DefaultMetadataReplicas must be less than or equal to MaxMetadataReplicas.

**Explanation:** The specified DefaultMetadataReplicas was greater than MaxMetadataReplicas.

**User response:** Specify a valid value for DefaultMetadataReplicas.

---

**6027-1043** DefaultDataReplicas must be less than or equal MaxDataReplicas.

**Explanation:** The specified DefaultDataReplicas was greater than MaxDataReplicas.

**User response:** Specify a valid value for DefaultDataReplicas.

---

---

**6027-1055**    **LogicalSectorSize must be a multiple of 512**

**Explanation:** The specified *LogicalSectorSize* was not a multiple of 512.

**User response:** Specify a valid *LogicalSectorSize*.

---

**6027-1056**    **Blocksize must be a multiple of  $LogicalSectorSize \times 32$**

**Explanation:** The specified *Blocksize* was not a multiple of  $LogicalSectorSize \times 32$ .

**User response:** Specify a valid value for *Blocksize*.

---

**6027-1057**    **InodeSize must be less than or equal to Blocksize.**

**Explanation:** The specified *InodeSize* was not less than or equal to *Blocksize*.

**User response:** Specify a valid value for *InodeSize*.

---

**6027-1059**    **Mode must be M or S; mode**

**Explanation:** The first argument provided in the `mmcrfsc` command was not M or S.

**User response:** The `mmcrfsc` command should not be called by a user. If any other command produces this error, contact the IBM Support Center.

---

**6027-1084**    **The specified block size (*valueK*) exceeds the maximum allowed block size currently in effect (*valueK*). Either specify a smaller value for the -B parameter, or increase the maximum block size by issuing: `mmchconfig maxblocksize=valueK` and restart the GPFS daemon.**

**Explanation:** The specified value for block size was greater than the value of the `maxblocksize` configuration parameter.

**User response:** Specify a valid value or increase the value of the allowed block size by specifying a larger value on the `maxblocksize` parameter of the `mmchconfig` command.

---

**6027-1113**    **Incorrect option: *option*.**

**Explanation:** The specified command option is not valid.

**User response:** Specify a valid option and reissue the command.

---

**6027-1119**    **Obsolete option: *option*.**

**Explanation:** A command received an option that is not valid any more.

**User response:** Correct the command line and reissue the command.

---

**6027-1120**    **Interrupt received: No changes made.**

**Explanation:** A GPFS administration command (`mm...`) received an interrupt before committing any changes.

**User response:** None. Informational message only.

---

**6027-1123**    **Disk name must be specified in disk descriptor.**

**Explanation:** The disk name positional parameter (the first field) in a disk descriptor was empty. The bad disk descriptor is displayed following this message.

**User response:** Correct the input and rerun the command.

---

**6027-1124**    **Disk usage must be `dataOnly`, `metadataOnly`, `descOnly`, or `dataAndMetadata`.**

**Explanation:** The disk usage parameter has a value that is not valid.

**User response:** Correct the input and reissue the command.

---

**6027-1132**    **Interrupt received: changes not propagated.**

**Explanation:** An interrupt was received after changes were committed but before the changes could be propagated to all the nodes.

**User response:** All changes will eventually propagate as nodes recycle or other GPFS administration commands are issued. Changes can be activated now by manually restarting the GPFS daemons.

---

**6027-1133**    **Interrupt received. Only a subset of the parameters were changed.**

**Explanation:** An interrupt was received in `mmchfs` before all of the requested changes could be completed.

**User response:** Use `mmfsfs` to see what the currently active settings are. Reissue the command if you want to change additional parameters.

---

**6027-1135**    **Restripping may not have finished.**

**Explanation:** An interrupt occurred during restripping.

**User response:** Restart the restripe. Verify that the file system was not damaged by running the `mmfsck` command.

---

---

**6027-1136** *option option specified twice.*

**Explanation:** An option was specified multiple times on a command line.

**User response:** Correct the error on the command line and reissue the command.

---

**6027-1137** *option value must be yes or no.*

**Explanation:** A yes or no option was used with something other than **yes** or **no**.

**User response:** Correct the error on the command line and reissue the command.

---

**6027-1138** **Incorrect extra argument:** *argument*

**Explanation:** Non-option arguments followed the mandatory arguments.

**User response:** Unlike most POSIX commands, the main arguments come first, followed by the optional arguments. Correct the error and reissue the command.

---

**6027-1140** **Incorrect integer for option:** *number.*

**Explanation:** An option requiring an integer argument was followed by something that cannot be parsed as an integer.

**User response:** Specify an integer with the indicated option.

---

**6027-1141** **No disk descriptor file specified.**

**Explanation:** An **-F** flag was not followed by the path name of a disk descriptor file.

**User response:** Specify a valid disk descriptor file.

---

**6027-1142** **File *fileName* already exists.**

**Explanation:** The specified file already exists.

**User response:** Rename the file or specify a different file name and reissue the command.

---

**6027-1143** **Cannot open *fileName*.**

**Explanation:** A file could not be opened.

**User response:** Verify that the specified file exists and that you have the proper authorizations.

---

**6027-1144** **Incompatible cluster types. You cannot move file systems that were created by GPFS cluster type *sourceCluster* into GPFS cluster type *targetCluster*.**

**Explanation:** The source and target cluster types are incompatible.

---

**User response:** Contact the IBM Support Center for assistance.

---

**6027-1145** *parameter must be greater than 0: value*

**Explanation:** A negative value had been specified for the named parameter, which requires a positive value.

**User response:** Correct the input and reissue the command.

---

**6027-1147** **Error converting *diskName* into an NSD.**

**Explanation:** Error encountered while converting a disk into an NSD.

**User response:** Check the preceding messages for more information.

---

**6027-1148** **File system *fileSystem* already exists in the cluster. Use **mmchfs -W** to assign a new device name for the existing file system.**

**Explanation:** You are trying to import a file system into the cluster but there is already a file system with the same name in the cluster.

**User response:** Remove or rename the file system with the conflicting name.

---

**6027-1149** *fileSystem* **is defined to have mount point *mountpoint*. There is already such a mount point in the cluster. Use **mmchfs -T** to assign a new mount point to the existing file system.**

**Explanation:** The cluster into which the file system is being imported already contains a file system with the same mount point as the mount point of the file system being imported.

**User response:** Use the **-T** option of the **mmchfs** command to change the mount point of the file system that is already in the cluster and then rerun the **mmimportfs** command.

---

**6027-1150** **Error encountered while importing disk *diskName*.**

**Explanation:** The **mmimportfs** command encountered problems while processing the disk.

**User response:** Check the preceding messages for more information.

---

**6027-1151** **Disk *diskName* already exists in the cluster.**

**Explanation:** You are trying to import a file system that has a disk with the same name as some disk from a file system that is already in the cluster.

**User response:** Remove or replace the disk with the conflicting name.

---

**6027-1152**    **Block size must be 64K, 128K, 256K, 512K, 1M, 2M, 4M, 8M or 16M.**

**Explanation:** The specified block size value is not valid.

**User response:** Specify a valid block size value.

---

**6027-1153**    **At least one node in the cluster must be defined as a quorum node.**

**Explanation:** All nodes were explicitly designated or allowed to default to be nonquorum.

**User response:** Specify which of the nodes should be considered quorum nodes and reissue the command.

---

**6027-1154**    **Incorrect node *node* specified for command.**

**Explanation:** The user specified a node that is not valid.

**User response:** Specify a valid node.

---

**6027-1155**    **The NSD servers for the following disks from file system *fileSystem* were reset or not defined: *diskList***

**Explanation:** Either the **mmimportfs** command encountered disks with no NSD servers, or was forced to reset the NSD server information for one or more disks.

**User response:** After the **mmimportfs** command finishes, use the **mmchnsd** command to assign NSD server nodes to the disks as needed.

---

**6027-1156**    **The NSD servers for the following free disks were reset or not defined: *diskList***

**Explanation:** Either the **mmimportfs** command encountered disks with no NSD servers, or was forced to reset the NSD server information for one or more disks.

**User response:** After the **mmimportfs** command finishes, use the **mmchnsd** command to assign NSD server nodes to the disks as needed.

---

**6027-1157**    **Use the **mmchnsd** command to assign NSD servers as needed.**

**Explanation:** Either the **mmimportfs** command encountered disks with no NSD servers, or was forced to reset the NSD server information for one or more disks. Check the preceding messages for detailed information.

**User response:** After the **mmimportfs** command

finishes, use the **mmchnsd** command to assign NSD server nodes to the disks as needed.

---

**6027-1159**    **The following file systems were not imported: *fileSystemList***

**Explanation:** The **mmimportfs** command was not able to import the specified file systems. Check the preceding messages for error information.

**User response:** Correct the problems and reissue the **mmimportfs** command.

---

**6027-1160**    **The drive letters for the following file systems have been reset: *fileSystemList*.**

**Explanation:** The drive letters associated with the specified file systems are already in use by existing file systems and have been reset.

**User response:** After the **mmimportfs** command finishes, use the **-t** option of the **mmchfs** command to assign new drive letters as needed.

---

**6027-1161**    **Use the dash character (-) to separate multiple node designations.**

**Explanation:** A command detected an incorrect character used as a separator in a list of node designations.

**User response:** Correct the command line and reissue the command.

---

**6027-1162**    **Use the semicolon character (;) to separate the disk names.**

**Explanation:** A command detected an incorrect character used as a separator in a list of disk names.

**User response:** Correct the command line and reissue the command.

---

**6027-1163**    **GPFS is still active on *nodeName*.**

**Explanation:** The GPFS daemon was discovered to be active on the specified node during an operation that requires the daemon to be stopped.

**User response:** Stop the daemon on the specified node and rerun the command.

---

**6027-1164**    **Use **mmchfs -t** to assign drive letters as needed.**

**Explanation:** The **mmimportfs** command was forced to reset the drive letters associated with one or more file systems. Check the preceding messages for detailed information.

**User response:** After the **mmimportfs** command finishes, use the **-t** option of the **mmchfs** command to assign new drive letters as needed.

---

**6027-1165** The PR attributes for the following disks from file system *fileSystem* were reset or not yet established: *diskList*

**Explanation:** The **mmimportfs** command disabled the Persistent Reserve attribute for one or more disks.

**User response:** After the **mmimportfs** command finishes, use the **mmchconfig** command to enable Persistent Reserve in the cluster as needed.

---

**6027-1166** The PR attributes for the following free disks were reset or not yet established: *diskList*

**Explanation:** The **mmimportfs** command disabled the Persistent Reserve attribute for one or more disks.

**User response:** After the **mmimportfs** command finishes, use the **mmchconfig** command to enable Persistent Reserve in the cluster as needed.

---

**6027-1167** Use **mmchconfig** to enable Persistent Reserve in the cluster as needed.

**Explanation:** The **mmimportfs** command disabled the Persistent Reserve attribute for one or more disks.

**User response:** After the **mmimportfs** command finishes, use the **mmchconfig** command to enable Persistent Reserve in the cluster as needed.

---

**6027-1168** Inode size must be 512, 1K or 4K.

**Explanation:** The specified inode size is not valid.

**User response:** Specify a valid inode size.

---

**6027-1169** *attribute* must be *value*.

**Explanation:** The specified value of the given attribute is not valid.

**User response:** Specify a valid value.

---

**6027-1178** *parameter* must be from *value* to *value: valueSpecified*

**Explanation:** A parameter value specified was out of range.

**User response:** Keep the specified value within the range shown.

---

**6027-1188** Duplicate disk specified: *disk*

**Explanation:** A disk was specified more than once on the command line.

**User response:** Specify each disk only once.

---



---

**6027-1189** You cannot delete all the disks.

**Explanation:** The number of disks to delete is greater than or equal to the number of disks in the file system.

**User response:** Delete only some of the disks. If you want to delete them all, use the **mmdelfs** command.

---

**6027-1197** *parameter* must be greater than *value: value*.

**Explanation:** An incorrect value was specified for the named parameter.

**User response:** Correct the input and reissue the command.

---

**6027-1200** **tscrfs** failed. Cannot create *device*

**Explanation:** The internal **tscrfs** command failed.

**User response:** Check the error message from the command that failed.

---

**6027-1201** Disk *diskName* does not belong to file system *fileSystem*.

**Explanation:** The specified disk was not found to be part of the cited file system.

**User response:** If the disk and file system were specified as part of a GPFS command, reissue the command with a disk that belongs to the specified file system.

---

**6027-1203** Attention: File system *fileSystem* may have some disks that are in a non-ready state. Issue the command: **mmcommon recoverfs** *fileSystem*

**Explanation:** The specified file system may have some disks that are in a non-ready state.

**User response:** Run **mmcommon recoverfs** *fileSystem* to ensure that the GPFS configuration data for the file system is current, and then display the states of the disks in the file system using the **mmlsdisk** command.

If any disks are in a non-ready state, steps should be taken to bring these disks into the ready state, or to remove them from the file system. This can be done by mounting the file system, or by using the **mmchdisk** command for a mounted or unmounted file system. When maintenance is complete or the failure has been repaired, use the **mmchdisk** command with the **start** option. If the failure cannot be repaired without loss of data, you can use the **mmdeldisk** command to delete the disks.

---

---

**6027-1204** *command failed.*

**Explanation:** An internal command failed. This is usually a call to the GPFS daemon.

**User response:** Check the error message from the command that failed.

---

**6027-1205** **Failed to connect to remote cluster** *clusterName.*

**Explanation:** Attempt to establish a connection to the specified cluster was not successful. This can be caused by a number of reasons: GPFS is down on all of the contact nodes, the contact node list is obsolete, the owner of the remote cluster revoked authorization, and so forth.

**User response:** If the error persists, contact the administrator of the remote cluster and verify that the contact node information is current and that the authorization key files are current as well.

---

**6027-1206** **File system *fileSystem* belongs to cluster** *clusterName*. **Command is not allowed for remote file systems.**

**Explanation:** The specified file system is not local to the cluster, but belongs to the cited remote cluster.

**User response:** Choose a local file system, or issue the command on a node in the remote cluster.

---

**6027-1207** **There is already an existing file system using *value*.**

**Explanation:** The mount point or device name specified matches that of an existing file system. The device name and mount point must be unique within a GPFS cluster.

**User response:** Choose an unused name or path.

---

**6027-1208** **File system *fileSystem* not found in cluster** *clusterName*.

**Explanation:** The specified file system does not belong to the cited remote cluster. The local information about the file system is not current. The file system may have been deleted, renamed, or moved to a different cluster.

**User response:** Contact the administrator of the remote cluster that owns the file system and verify the accuracy of the local information. Use the **mmremotefs show** command to display the local information about the file system. Use the **mmremotefs update** command to make the necessary changes.

---

**6027-1209** **GPFS is down on this node.**

**Explanation:** GPFS is not running on this node.

**User response:** Ensure that GPFS is running and reissue the command.

---

**6027-1210** **GPFS is not ready to handle commands yet.**

**Explanation:** GPFS is in the process of initializing or waiting for quorum to be reached.

**User response:** Reissue the command.

---

**6027-1211** *fileSystem* **refers to file system *fileSystem* in cluster** *clusterName*.

**Explanation:** Informational message.

**User response:** None.

---

**6027-1212** **File system *fileSystem* does not belong to cluster** *clusterName*.

**Explanation:** The specified file system refers to a file system that is remote to the cited cluster. Indirect remote file system access is not allowed.

**User response:** Contact the administrator of the remote cluster that owns the file system and verify the accuracy of the local information. Use the **mmremotefs show** command to display the local information about the file system. Use the **mmremotefs update** command to make the necessary changes.

---

**6027-1213** *command failed. Error code* *errorCode*.

**Explanation:** An internal command failed. This is usually a call to the GPFS daemon.

**User response:** Examine the error code and other messages to determine the reason for the failure. Correct the problem and reissue the command.

---

**6027-1214** **Unable to enable Persistent Reserve on the following disks:** *diskList*

**Explanation:** The command was unable to set up all of the disks to use Persistent Reserve.

**User response:** Examine the disks and the additional error information to determine if the disks should have supported Persistent Reserve. Correct the problem and reissue the command.

---

**6027-1215** **Unable to reset the Persistent Reserve attributes on one or more disks on the following nodes:** *nodeList*

**Explanation:** The command could not reset Persistent Reserve on at least one disk on the specified nodes.

**User response:** Examine the additional error

information to determine whether nodes were down or if there was a disk error. Correct the problems and reissue the command.

---

**6027-1216** File *fileName* contains additional error information.

**Explanation:** The command generated a file containing additional error information.

**User response:** Examine the additional error information.

---

**6027-1217** A disk descriptor contains an incorrect separator character.

**Explanation:** A command detected an incorrect character used as a separator in a disk descriptor.

**User response:** Correct the disk descriptor and reissue the command.

---

**6027-1218** Node *nodeName* does not have a GPFS server license designation.

**Explanation:** The function that you are assigning to the node requires the node to have a GPFS server license.

**User response:** Use the `mmchlicense` command to assign a valid GPFS license to the node or specify a different node.

---

**6027-1219** NSD discovery on node *nodeName* failed with return code *value*.

**Explanation:** The NSD discovery process on the specified node failed with the specified return code.

**User response:** Determine why the node cannot access the specified NSDs. Correct the problem and reissue the command.

---

**6027-1220** Node *nodeName* cannot be used as an NSD server for Persistent Reserve disk *diskName* because it is not an AIX node.

**Explanation:** The node shown was specified as an NSD server for *diskName*, but the node does not support Persistent Reserve.

**User response:** Specify a node that supports Persistent Reserve as an NSD server.

---

**6027-1221** The number of NSD servers exceeds the maximum (*value*) allowed.

**Explanation:** The number of NSD servers in the disk descriptor exceeds the maximum allowed.

**User response:** Change the disk descriptor to specify no more NSD servers than the maximum allowed.

---

**6027-1222** Cannot assign a minor number for file system *fileSystem* (major number *deviceMajorNumber*).

**Explanation:** The command was not able to allocate a minor number for the new file system.

**User response:** Delete unneeded `/dev` entries for the specified major number and reissue the command.

---

**6027-1223** *ipAddress* cannot be used for NFS serving; it is used by the GPFS daemon.

**Explanation:** The IP address shown has been specified for use by the GPFS daemon. The same IP address cannot be used for NFS serving because it cannot be failed over.

**User response:** Specify a different IP address for NFS use and reissue the command.

---

**6027-1224** There is no file system with drive letter *driveLetter*.

**Explanation:** No file system in the GPFS cluster has the specified drive letter.

**User response:** Reissue the command with a valid file system.

---

**6027-1225** Explicit drive letters are supported only in a Windows environment. Specify a mount point or allow the default settings to take effect.

**Explanation:** An explicit drive letter was specified on the `mmmount` command but the target node does not run the Windows operating system.

**User response:** Specify a mount point or allow the default settings for the file system to take effect.

---

**6027-1226** Explicit mount points are not supported in a Windows environment. Specify a drive letter or allow the default settings to take effect.

**Explanation:** An explicit mount point was specified on the `mmmount` command but the target node runs the Windows operating system.

**User response:** Specify a drive letter or allow the default settings for the file system to take effect.

---

**6027-1227** The main GPFS cluster configuration file is locked. Retrying ...

**Explanation:** Another GPFS administration command has locked the cluster configuration file. The current process will try to obtain the lock a few times before giving up.

**User response:** None. Informational message only.



---

**6027-1228 Lock creation successful.**

**Explanation:** The holder of the lock has released it and the current process was able to obtain it.

**User response:** None. Informational message only. The command will now continue.

---

**6027-1229 Timed out waiting for lock. Try again later.**

**Explanation:** Another GPFS administration command kept the main GPFS cluster configuration file locked for over a minute.

**User response:** Try again later. If no other GPFS administration command is presently running, see E! "GPFS cluster configuration data file issues" on page 229.

---

**6027-1230 *diskName* is a tiebreaker disk and cannot be deleted.**

**Explanation:** A request was made to GPFS to delete a node quorum tiebreaker disk.

**User response:** Specify a different disk for deletion.

---

**6027-1231 GPFS detected more than eight quorum nodes while node quorum with tiebreaker disks is in use.**

**Explanation:** A GPFS command detected more than eight quorum nodes, but this is not allowed while node quorum with tiebreaker disks is in use.

**User response:** Reduce the number of quorum nodes to a maximum of eight, or use the normal node quorum algorithm.

---

**6027-1232 GPFS failed to initialize the tiebreaker disks.**

**Explanation:** A GPFS command unsuccessfully attempted to initialize the node quorum tiebreaker disks.

**User response:** Examine prior messages to determine why GPFS was unable to initialize the tiebreaker disks and correct the problem. After that, reissue the command.

---

**6027-1233 Incorrect keyword: *value*.**

**Explanation:** A command received a keyword that is not valid.

**User response:** Correct the command line and reissue the command.

---

---

**6027-1234 Adding node *node* to the cluster will exceed the quorum node limit.**

**Explanation:** An attempt to add the cited node to the cluster resulted in the quorum node limit being exceeded.

**User response:** Change the command invocation to not exceed the node quorum limit, and reissue the command.

---

---

**6027-1235 The *fileName* kernel extension does not exist. Use the `mmbuildgpl` command to create the needed kernel extension for your kernel or copy the binaries from another node with the identical environment.**

**Explanation:** The cited kernel extension does not exist.

**User response:** Create the needed kernel extension by compiling a custom `mmfslinux` module for your kernel (see steps in `/usr/lpp/mmfs/src/README`), or copy the binaries from another node with the identical environment.

---

---

**6027-1236 Unable to verify kernel/module configuration.**

**Explanation:** The `mmfslinux` kernel extension does not exist.

**User response:** Create the needed kernel extension by compiling a custom `mmfslinux` module for your kernel (see steps in `/usr/lpp/mmfs/src/README`), or copy the binaries from another node with the identical environment.

---

---

**6027-1237 The GPFS daemon is still running; use the `mmshutdown` command.**

**Explanation:** An attempt was made to unload the GPFS kernel extensions while the GPFS daemon was still running.

**User response:** Use the `mmshutdown` command to shut down the daemon.

---

---

**6027-1238 Module *fileName* is still in use. Unmount all GPFS file systems and issue the command: `mmfsadm cleanup`**

**Explanation:** An attempt was made to unload the cited module while it was still in use.

**User response:** Unmount all GPFS file systems and issue the command `mmfsadm cleanup`. If this does not solve the problem, reboot the machine.

---

---

**6027-1239** Error unloading module *moduleName*.

**Explanation:** GPFS was unable to unload the cited module.

**User response:** Unmount all GPFS file systems and issue the command **mmfsadm cleanup**. If this does not solve the problem, reboot the machine.

---

**6027-1240** Module *fileName* is already loaded.

**Explanation:** An attempt was made to load the cited module, but it was already loaded.

**User response:** None. Informational message only.

---

**6027-1241** *diskName* was not found in **/proc/partitions**.

**Explanation:** The cited disk was not found in **/proc/partitions**.

**User response:** Take steps to cause the disk to appear in **/proc/partitions**, and then reissue the command.

---

**6027-1242** GPFS is waiting for *requiredCondition*

**Explanation:** GPFS is unable to come up immediately due to the stated required condition not being satisfied yet.

**User response:** This is an informational message. As long as the required condition is not satisfied, this message will repeat every five minutes. You may want to stop the GPFS daemon after a while, if it will be a long time before the required condition will be met.

---

**6027-1243** *command*: Processing user configuration file *fileName*

**Explanation:** Progress information for the **mmcrcluster** command.

**User response:** None. Informational message only.

---

**6027-1244** *configParameter* is set by the **mmcrcluster** processing. Line in error: *configLine*. The line will be ignored; processing continues.

**Explanation:** The specified parameter is set by the **mmcrcluster** command and cannot be overridden by the user.

**User response:** None. Informational message only.

---

**6027-1245** *configParameter* must be set with the *command* command. Line in error: *configLine*. The line is ignored; processing continues.

**Explanation:** The specified parameter has additional dependencies and cannot be specified prior to the

completion of the **mmcrcluster** command.

**User response:** After the cluster is created, use the specified command to establish the desired configuration parameter.

---

**6027-1246** *configParameter* is an obsolete parameter. Line in error: *configLine*. The line is ignored; processing continues.

**Explanation:** The specified parameter is not used by GPFS anymore.

**User response:** None. Informational message only.

---

**6027-1247** *configParameter* cannot appear in a node-override section. Line in error: *configLine*. The line is ignored; processing continues.

**Explanation:** The specified parameter must have the same value across all nodes in the cluster.

**User response:** None. Informational message only.

---

**6027-1248** Mount point can not be a relative path name: *path*

**Explanation:** The mount point does not begin with **/**.

**User response:** Specify the absolute path name for the mount point.

---

**6027-1249** *operand* can not be a relative path name: *path*.

**Explanation:** The specified path name does not begin with **'/'**.

**User response:** Specify the absolute path name.

---

**6027-1250** Key file is not valid.

**Explanation:** While attempting to establish a connection to another node, GPFS detected that the format of the public key file is not valid.

**User response:** Use the **mmremoteccluster** command to specify the correct public key.

---

**6027-1251** Key file mismatch.

**Explanation:** While attempting to establish a connection to another node, GPFS detected that the public key file does not match the public key file of the cluster to which the file system belongs.

**User response:** Use the **mmremoteccluster** command to specify the correct public key.

---

---

**6027-1252** Node *nodeName* already belongs to the GPFS cluster.

**Explanation:** A GPFS command found that a node to be added to a GPFS cluster already belongs to the cluster.

**User response:** Specify a node that does not already belong to the GPFS cluster.

---

**6027-1253** Incorrect value for *option* option.

**Explanation:** The provided value for the specified option is not valid.

**User response:** Correct the error and reissue the command.

---

**6027-1254** Warning: Not all nodes have proper GPFS license designations. Use the `mmchlicense` command to designate licenses as needed.

**Explanation:** Not all nodes in the cluster have valid license designations.

**User response:** Use `mmlslicense` to see the current license designations. Use `mmchlicense` to assign valid GPFS licenses to all nodes as needed.

---

**6027-1255** There is nothing to commit. You must first run: *command*.

**Explanation:** You are attempting to commit an SSL private key but such a key has not been generated yet.

**User response:** Run the specified command to generate the public/private key pair.

---

**6027-1256** The current authentication files are already committed.

**Explanation:** You are attempting to commit public/private key files that were previously generated with the `mmauth` command. The files have already been committed.

**User response:** None. Informational message.

---

**6027-1257** There are uncommitted authentication files. You must first run: *command*.

**Explanation:** You are attempting to generate new public/private key files but previously generated files have not been committed yet.

**User response:** Run the specified command to commit the current public/private key pair.

---

**6027-1258** You must establish a cipher list first. Run: *command*.

**Explanation:** You are attempting to commit an SSL private key but a cipher list has not been established yet.

**User response:** Run the specified command to specify a cipher list.

---

**6027-1259** *command* not found. Ensure the OpenSSL code is properly installed.

**Explanation:** The specified command was not found.

**User response:** Ensure the OpenSSL code is properly installed and reissue the command.

---

**6027-1260** File *fileName* does not contain any *typeOfStanza* stanzas.

**Explanation:** The input file should contain at least one specified stanza.

**User response:** Correct the input file and reissue the command.

---

**6027-1261** *descriptorField* must be specified in *descriptorType* descriptor.

**Explanation:** A required field of the descriptor was empty. The incorrect descriptor is displayed following this message.

**User response:** Correct the input and reissue the command.

---

**6027-1262** Unable to obtain the GPFS configuration file lock. Retrying ...

**Explanation:** A command requires the lock for the GPFS system data but was not able to obtain it.

**User response:** None. Informational message only.

---

**6027-1263** Unable to obtain the GPFS configuration file lock.

**Explanation:** A command requires the lock for the GPFS system data but was not able to obtain it.

**User response:** Check the preceding messages, if any. E! Follow the procedure in "GPFS cluster configuration E! data file issues" on page 229, and then reissue the command.

---

**6027-1268** Missing arguments.

**Explanation:** A GPFS administration command received an insufficient number of arguments.

**User response:** Correct the command line and reissue the command.

---

---

**6027-1269** The device name *device* starts with a slash, but not `/dev/`.

**Explanation:** The device name does not start with `/dev/`.

**User response:** Correct the device name.

---

**6027-1270** The device name *device* contains a slash, but not as its first character.

**Explanation:** The specified device name contains a slash, but the first character is not a slash.

**User response:** The device name must be an unqualified device name or an absolute device path name, for example: `fs0` or `/dev/fs0`.

---

**6027-1271** Unexpected error from *command*. Return code: *value*

**Explanation:** A GPFS administration command (`mm...`) received an unexpected error code from an internally called command.

**User response:** Perform problem determination. See “GPFS commands are unsuccessful” on page 237.

---

**6027-1272** Unknown user name *userName*.

**Explanation:** The specified value cannot be resolved to a valid user ID (UID).

**User response:** Reissue the command with a valid user name.

---

**6027-1273** Unknown group name *groupName*.

**Explanation:** The specified value cannot be resolved to a valid group ID (GID).

**User response:** Reissue the command with a valid group name.

---

**6027-1274** Unexpected error obtaining the *lockName* lock.

**Explanation:** GPFS cannot obtain the specified lock.

**User response:** Examine any previous error messages. Correct any problems and reissue the command. If the problem persists, perform problem determination and contact the IBM Support Center.

---

**6027-1275** Daemon node adapter *Node* was not found on admin node *Node*.

**Explanation:** An input node descriptor was found to be incorrect. The node adapter specified for GPFS daemon communications was not found to exist on the cited GPFS administrative node.

**User response:** Correct the input node descriptor and reissue the command.

---

**6027-1276** Command failed for disks: *diskList*.

**Explanation:** A GPFS command was unable to complete successfully on the listed disks.

**User response:** Correct the problems and reissue the command.

---

**6027-1277** No contact nodes were provided for cluster *clusterName*.

**Explanation:** A GPFS command found that no contact nodes have been specified for the cited cluster.

**User response:** Use the `mmremotecoluster` command to specify some contact nodes for the cited cluster.

---

**6027-1278** None of the contact nodes in cluster *clusterName* can be reached.

**Explanation:** A GPFS command was unable to reach any of the contact nodes for the cited cluster.

**User response:** Determine why the contact nodes for the cited cluster cannot be reached and correct the problem, or use the `mmremotecoluster` command to specify some additional contact nodes that can be reached.

---

**6027-1287** Node *nodeName* returned ENODEV for disk *diskName*.

**Explanation:** The specified node returned ENODEV for the specified disk.

**User response:** Determine the cause of the ENODEV error for the specified disk and rectify it. The ENODEV may be due to disk fencing or the removal of a device that previously was present.

---

**6027-1288** Remote cluster *clusterName* was not found.

**Explanation:** A GPFS command found that the cited cluster has not yet been identified to GPFS as a remote cluster.

**User response:** Specify a remote cluster known to GPFS, or use the `mmremotecoluster` command to make the cited cluster known to GPFS.

---

**6027-1289** Name *name* is not allowed. It contains the following invalid special character: *char*

**Explanation:** The cited name is not allowed because it contains the cited invalid special character.

**User response:** Specify a name that does not contain an invalid special character, and reissue the command.

---

---

**6027-1290** GPFS configuration data for file system *fileSystem* may not be in agreement with the on-disk data for the file system. Issue the command: `mmcommon recoverfs fileSystem`

**Explanation:** GPFS detected that the GPFS configuration database data for the specified file system may not be in agreement with the on-disk data for the file system. This may be caused by a GPFS disk command that did not complete normally.

**User response:** Issue the specified command to bring the GPFS configuration database into agreement with the on-disk data.

---

**6027-1291** Options *name* and *name* cannot be specified at the same time.

**Explanation:** Incompatible options were specified on the command line.

**User response:** Select one of the options and reissue the command.

---

**6027-1292** The `-N` option cannot be used with attribute *name*.

**Explanation:** The specified configuration attribute cannot be changed on only a subset of nodes. This attribute must be the same on all nodes in the cluster.

**User response:** Certain attributes, such as `autoload`, may not be customized from node to node. Change the attribute for the entire cluster.

---

**6027-1293** There are no remote file systems.

**Explanation:** A value of `all` was specified for the remote file system operand of a GPFS command, but no remote file systems are defined.

**User response:** None. There are no remote file systems on which to operate.

---

**6027-1294** Remote file system *fileSystem* is not defined.

**Explanation:** The specified file system was used for the remote file system operand of a GPFS command, but the file system is not known to GPFS.

**User response:** Specify a remote file system known to GPFS.

---

**6027-1295** The GPFS configuration information is incorrect or not available.

**Explanation:** A problem has been encountered while verifying the configuration information and the execution environment.

**User response:** Check the preceding messages for

more information. Correct the problem and restart GPFS.

---

**6027-1296** Device name cannot be 'all'.

**Explanation:** A device name of `all` was specified on a GPFS command.

**User response:** Reissue the command with a valid device name.

---

**6027-1297** Each device specifies `metadataOnly` for disk usage. This file system could not store data.

**Explanation:** All disk descriptors specify `metadataOnly` for disk usage.

**User response:** Change at least one disk descriptor in the file system to indicate the usage of `dataOnly` or `dataAndMetadata`.

---

**6027-1298** Each device specifies `dataOnly` for disk usage. This file system could not store metadata.

**Explanation:** All disk descriptors specify `dataOnly` for disk usage.

**User response:** Change at least one disk descriptor in the file system to indicate a usage of `metadataOnly` or `dataAndMetadata`.

---

**6027-1299** Incorrect value *value* specified for failure group.

**Explanation:** The specified failure group is not valid.

**User response:** Correct the problem and reissue the command.

---

**6027-1300** No file systems were found.

**Explanation:** A GPFS command searched for file systems, but none were found.

**User response:** Create a GPFS file system before reissuing the command.

---

**6027-1301** The NSD servers specified in the disk descriptor do not match the NSD servers currently in effect.

**Explanation:** The set of NSD servers specified in the disk descriptor does not match the set that is currently in effect.

**User response:** Specify the same set of NSD servers in the disk descriptor as is currently in effect or omit it from the disk descriptor and then reissue the command. Use the `mmchnsd` command to change the NSD servers as needed.

---

**6027-1302** *clusterName* is the name of the local cluster.

**Explanation:** The cited cluster name was specified as the name of a remote cluster, but it is already being used as the name of the local cluster.

**User response:** Use the **mmchcluster** command to change the name of the local cluster, and then reissue the command that failed.

---

**6027-1303** This function is not available in the GPFS Express Edition.

**Explanation:** The requested function is not part of the GPFS Express Edition.

**User response:** Install the GPFS Standard Edition on all nodes in the cluster, and then reissue the command.

---

**6027-1304** Missing argument after *option* option.

**Explanation:** The specified command option requires a value.

**User response:** Specify a value and reissue the command.

---

**6027-1305** Prerequisite libraries not found or correct version not installed. Ensure *productName* is properly installed.

**Explanation:** The specified software product is missing or is not properly installed.

**User response:** Verify that the product is installed properly.

---

**6027-1306** Command *command* failed with return code *value*.

**Explanation:** A command was not successfully processed.

**User response:** Correct the failure specified by the command and reissue the command.

---

**6027-1307** Disk *disk* on node *nodeName* already has a volume group *vgName* that does not appear to have been created by this program in a prior invocation. Correct the descriptor file or remove the volume group and retry.

**Explanation:** The specified disk already belongs to a volume group.

**User response:** Either remove the volume group or remove the disk descriptor and retry.

---

**6027-1308** *feature* is not available in the GPFS Express Edition.

**Explanation:** The specified function or feature is not part of the GPFS Express Edition.

**User response:** Install the GPFS Standard Edition on all nodes in the cluster, and then reissue the command.

---

**6027-1309** Storage pools are not available in the GPFS Express Edition.

**Explanation:** Support for multiple storage pools is not part of the GPFS Express Edition.

**User response:** Install the GPFS Standard Edition on all nodes in the cluster, and then reissue the command.

---

**6027-1332** Cannot find *disk* with *command*.

**Explanation:** The specified disk cannot be found.

**User response:** Specify a correct disk name.

---

**6027-1333** The following nodes could not be restored: *nodeList*. Correct the problems and use the **mmsdrrestore** command to recover these nodes.

**Explanation:** The **mmsdrrestore** command was unable to restore the configuration information for the listed nodes.

**User response:** Correct the problems and reissue the **mmsdrrestore** command for these nodes.

---

**6027-1334** Incorrect value for option *option*. Valid values are: *validValues*.

**Explanation:** An incorrect argument was specified with an option requiring one of a limited number of legal options.

**User response:** Use one of the legal values for the indicated option.

---

**6027-1335** Command completed: Not all required changes were made.

**Explanation:** Some, but not all, of the required changes were made.

**User response:** Examine the preceding messages, correct the problems, and reissue the command.

---

**6027-1338** Command is not allowed for remote file systems.

**Explanation:** A command for which a remote file system is not allowed was issued against a remote file system.

**User response:** Choose a local file system, or issue the

command on a node in the cluster that owns the file system.

---

**6027-1339** Disk usage *value* is incompatible with storage pool *name*.

**Explanation:** A disk descriptor specified a disk usage involving metadata and a storage pool other than system.

**User response:** Change the descriptor's disk usage field to **dataOnly**, or do not specify a storage pool name.

---

**6027-1340** File *fileName* not found. Recover the file or run **mmauth genkey**.

**Explanation:** The cited file was not found.

**User response:** Recover the file or run the **mmauth genkey** command to recreate it.

---

**6027-1341** Starting force unmount of GPFS file systems

**Explanation:** Progress information for the **mmshutdown** command.

**User response:** None. Informational message only.

---

**6027-1342** Unmount not finished after *value* seconds. Waiting *value* more seconds.

**Explanation:** Progress information for the **mmshutdown** command.

**User response:** None. Informational message only.

---

**6027-1343** Unmount not finished after *value* seconds.

**Explanation:** Progress information for the **mmshutdown** command.

**User response:** None. Informational message only.

---

**6027-1344** Shutting down GPFS daemons

**Explanation:** Progress information for the **mmshutdown** command.

**User response:** None. Informational message only.

---

**6027-1345** Finished

**Explanation:** Progress information for the **mmshutdown** command.

**User response:** None. Informational message only.

---

**6027-1347** Disk with NSD volume id *NSD volume id* no longer exists in the GPFS cluster configuration data but the NSD volume id was not erased from the disk. To remove the NSD volume id, issue: **mmdeinsd -p NSD volume id**

**Explanation:** A GPFS administration command (**mm...**) successfully removed the disk with the specified NSD volume id from the GPFS cluster configuration data but was unable to erase the NSD volume id from the disk.

**User response:** Issue the specified command to remove the NSD volume id from the disk.

---

**6027-1348** Disk with NSD volume id *NSD volume id* no longer exists in the GPFS cluster configuration data but the NSD volume id was not erased from the disk. To remove the NSD volume id, issue: **mmdeinsd -p NSD volume id -N nodeNameList**

**Explanation:** A GPFS administration command (**mm...**) successfully removed the disk with the specified NSD volume id from the GPFS cluster configuration data but was unable to erase the NSD volume id from the disk.

**User response:** Issue the specified command to remove the NSD volume id from the disk.

---

**6027-1352** *fileSystem* is not a remote file system known to GPFS.

**Explanation:** The cited file system is not the name of a remote file system known to GPFS.

**User response:** Use the **mmremotefs** command to identify the cited file system to GPFS as a remote file system, and then reissue the command that failed.

---

**6027-1357** An internode connection between GPFS nodes was disrupted.

**Explanation:** An internode connection between GPFS nodes was disrupted, preventing its successful completion.

**User response:** Reissue the command. If the problem recurs, determine and resolve the cause of the disruption. If the problem persists, contact the IBM Support Center.

---

**6027-1358** No clusters are authorized to access this cluster.

**Explanation:** Self-explanatory.

**User response:** This is an informational message.

---

**6027-1359** Cluster *clusterName* is not authorized to access this cluster.

**Explanation:** Self-explanatory.

**User response:** This is an informational message.

---

**6027-1361** Attention: There are no available valid VFS type values for mmfs in /etc/vfs.

**Explanation:** An out of range number was used as the vfs number for GPFS.

**User response:** The valid range is 8 through 32. Check /etc/vfs and remove unneeded entries.

---

**6027-1362** There are no remote cluster definitions.

**Explanation:** A value of **all** was specified for the remote cluster operand of a GPFS command, but no remote clusters are defined.

**User response:** None. There are no remote clusters on which to operate.

---

**6027-1363** Remote cluster *clusterName* is not defined.

**Explanation:** The specified cluster was specified for the remote cluster operand of a GPFS command, but the cluster is not known to GPFS.

**User response:** Specify a remote cluster known to GPFS.

---

**6027-1364** No disks specified

**Explanation:** There were no disks in the descriptor list or file.

**User response:** Specify at least one disk.

---

**6027-1365** Disk *diskName* already belongs to file system *fileSystem*.

**Explanation:** The specified disk name is already assigned to a GPFS file system. This may be because the disk was specified more than once as input to the command, or because the disk was assigned to a GPFS file system in the past.

**User response:** Specify the disk only once as input to the command, or specify a disk that does not belong to a file system.

---

**6027-1366** File system *fileSystem* has some disks that are in a non-ready state.

**Explanation:** The specified file system has some disks that are in a non-ready state.

**User response:** Run **mmcommon recoverfs** *fileSystem* to ensure that the GPFS configuration data for the file system is current. If some disks are still in a non-ready

state, display the states of the disks in the file system using the **mmlsdisk** command. Any disks in an undesired non-ready state should be brought into the ready state by using the **mmchdisk** command or by mounting the file system. If these steps do not bring the disks into the ready state, use the **mmdeldisk** command to delete the disks from the file system.

---

**6027-1367** Attention: Not all disks were marked as available.

**Explanation:** The process of marking the disks as available could not be completed.

**User response:** Before adding these disks to a GPFS file system, you should either reformat them, or use the **-v no** option on the **mmcrfs** or **mmadddisk** command.

---

**6027-1368** This GPFS cluster contains declarations for remote file systems and clusters. You cannot delete the last node. First use the **delete** option of the **mmremoteclass** and **mmremotefs** commands.

**Explanation:** An attempt has been made to delete a GPFS cluster that still has declarations for remote file systems and clusters.

**User response:** Before deleting the last node of a GPFS cluster, delete all remote cluster and file system information. Use the **delete** option of the **mmremoteclass** and **mmremotefs** commands.

---

**6027-1370** The following nodes could not be reached:

**Explanation:** A GPFS command was unable to communicate with one or more nodes in the cluster. A list of the nodes that could not be reached follows.

**User response:** Determine why the reported nodes could not be reached and resolve the problem.

---

**6027-1371** Propagating the cluster configuration data to all affected nodes. This is an asynchronous process.

**Explanation:** A process is initiated to distribute the cluster configuration data to other nodes in the cluster.

**User response:** This is an informational message. The command does not wait for the distribution to finish.

---

**6027-1373** There is no file system information in input file *fileName*.

**Explanation:** The cited input file passed to the **mmimportfs** command contains no file system information. No file system can be imported.

**User response:** Reissue the **mmimportfs** command while specifying a valid input file.



---

**6027-1374** File system *fileSystem* was not found in input file *fileName*.

**Explanation:** The specified file system was not found in the input file passed to the **mmimportfs** command. The file system cannot be imported.

**User response:** Reissue the **mmimportfs** command while specifying a file system that exists in the input file.

---

**6027-1375** The following file systems were not imported: *fileSystem*.

**Explanation:** The **mmimportfs** command was unable to import one or more of the file systems in the input file. A list of the file systems that could not be imported follows.

**User response:** Examine the preceding messages, rectify the problems that prevented the importation of the file systems, and reissue the **mmimportfs** command.

---

**6027-1377** Attention: Unknown attribute specified: *name*. Press the ENTER key to continue.

**Explanation:** The **mmchconfig** command received an unknown attribute.

**User response:** Unless directed otherwise by the IBM Support Center, press any key to bypass this attribute.

---

**6027-1378** Incorrect record found in the **mmsdrfs** file (code *value*):

**Explanation:** A line that is not valid was detected in the main GPFS cluster configuration file **/var/mmfs/gen/mmsdrfs**.

**User response:** The data in the cluster configuration file is incorrect. If no user modifications have been made to this file, contact the IBM Support Center. If user modifications have been made, correct these modifications.

---

**6027-1379** There is no file system with mount point *mountpoint*.

**Explanation:** No file system in the GPFS cluster has the specified mount point.

**User response:** Reissue the command with a valid file system.

---

**6027-1380** File system *fileSystem* is already mounted at *mountpoint*.

**Explanation:** The specified file system is mounted at a mount point different than the one requested on the **mmmount** command.

**User response:** Unmount the file system and reissue the command.

---

**6027-1381** Mount point cannot be specified when mounting all file systems.

**Explanation:** A device name of **all** and a mount point were specified on the **mmmount** command.

**User response:** Reissue the command with a device name for a single file system or do not specify a mount point.

---

**6027-1382** This node does not belong to a GPFS cluster.

**Explanation:** The specified node does not appear to belong to a GPFS cluster, or the GPFS configuration information on the node has been lost.

**User response:** Informational message. If you suspect that there is corruption of the GPFS configuration information, recover the data following the procedures outlined in "Recovery from loss of GPFS cluster configuration data file" on page 230.

---

**6027-1383** There is no record for this node in file *fileName*. Either the node is not part of the cluster, the file is for a different cluster, or not all of the node's adapter interfaces have been activated yet.

**Explanation:** The **mmsdrrestore** command cannot find a record for this node in the specified cluster configuration file. The search of the file is based on the currently active IP addresses of the node as reported by the **ifconfig** command.

**User response:** Ensure that all adapter interfaces are properly functioning. Ensure that the correct GPFS configuration file is specified on the command line. If the node indeed is not a member of the cluster, use the **mmaddnode** command instead.

---

**6027-1386** Unexpected value for Gpfs object: *value*.

**Explanation:** A function received a value that is not allowed for the Gpfs object.

**User response:** Perform problem determination.

---

**6027-1388** File system *fileSystem* is not known to the GPFS cluster.

**Explanation:** The file system was not found in the GPFS cluster.

**User response:** If the file system was specified as part of a GPFS command, reissue the command with a valid file system.

---

**6027-1390** Node *node* does not belong to the GPFS cluster, or was specified as input multiple times.

**Explanation:** Nodes that are not valid were specified.

**User response:** Verify the list of nodes. All specified nodes must belong to the GPFS cluster, and each node can be specified only once.

---

**6027-1393** Incorrect node designation specified: *type*.

**Explanation:** A node designation that is not valid was specified. Valid values are **client** or **manager**.

**User response:** Correct the command line and reissue the command.

---

**6027-1394** Operation not allowed for the local cluster.

**Explanation:** The requested operation cannot be performed for the local cluster.

**User response:** Specify the name of a remote cluster.

---

**6027-1450** Could not allocate storage.

**Explanation:** Sufficient memory cannot be allocated to run the **mmsanrepairfs** command.

**User response:** Increase the amount of memory available.

---

**6027-1500 [E]** Open *devicetype device* failed with error:

**Explanation:** The "open" of a device failed. Operation of the file system may continue unless this device is needed for operation. If this is a replicated disk device, it will often not be needed. If this is a block or character device for another subsystem (such as */dev/VSD0*) then GPFS will discontinue operation.

**User response:** Problem diagnosis will depend on the subsystem that the device belongs to. For instance device *"/dev/VSD0"* belongs to the IBM Virtual Shared Disk subsystem and problem determination should follow guidelines in that subsystem's documentation. If this is a normal disk device then take needed repair action on the specified disk.

---

**6027-1501 [X]** Volume label of disk *name* is *name*, should be *uid*.

**Explanation:** The UID in the disk descriptor does not match the expected value from the file system descriptor. This could occur if a disk was overwritten by another application or if the IBM Virtual Shared Disk subsystem incorrectly identified the disk.

**User response:** Check the disk configuration.

---



---

**6027-1502 [X]** Volume label of disk *diskName* is corrupt.

**Explanation:** The disk descriptor has a bad magic number, version, or checksum. This could occur if a disk was overwritten by another application or if the IBM Virtual Shared Disk subsystem incorrectly identified the disk.

**User response:** Check the disk configuration.

---

**6027-1503** Completed adding disks to file system *fileSystem*.

**Explanation:** The **mmadddisk** command successfully completed.

**User response:** None. Informational message only.

---

**6027-1504** File *name* could not be run with **err error**.

**Explanation:** A failure occurred while trying to run an external program.

**User response:** Make sure the file exists. If it does, check its access permissions.

---

**6027-1505** Could not get minor number for *name*.

**Explanation:** Could not obtain a minor number for the specified block or character device.

**User response:** Problem diagnosis will depend on the subsystem that the device belongs to. For example, device */dev/VSD0* belongs to the IBM Virtual Shared Disk subsystem and problem determination should follow guidelines in that subsystem's documentation.

---

**6027-1507** **READ\_KEYS ioctl** failed with **errno=returnCode**, tried *timesTried* times. Related values are **scsi\_status=scsiStatusValue**, **sense\_key=senseKeyValue**, **scsi\_asc=scsiAscValue**, **scsi\_ascq=scsiAscqValue**.

**Explanation:** A **READ\_KEYS ioctl** call failed with the **errno=** and related values shown.

**User response:** Check the reported **errno=** value and try to correct the problem. If the problem persists, contact the IBM Support Center.

---

**6027-1508** **Registration** failed with **errno=returnCode**, tried *timesTried* times. Related values are **scsi\_status=scsiStatusValue**, **sense\_key=senseKeyValue**, **scsi\_asc=scsiAscValue**, **scsi\_ascq=scsiAscqValue**.

**Explanation:** A **REGISTER ioctl** call failed with the **errno=** and related values shown.

---

**User response:** Check the reported **errno=** value and try to correct the problem. If the problem persists, contact the IBM Support Center.

---

**6027-1509**    **READRES ioctl failed with**  
**errno=returnCode, tried timesTried times.**  
**Related values are**  
**scsi\_status=scsiStatusValue,**  
**sense\_key=senseKeyValue,**  
**scsi\_asc=scsiAscValue,**  
**scsi\_ascq=scsiAscqValue.**

**Explanation:** A READRES ioctl call failed with the **errno=** and related values shown.

**User response:** Check the reported **errno=** value and try to correct the problem. If the problem persists, contact the IBM Support Center.

---

**6027-1510 [E]** **Error mounting file system** *stripeGroup*  
**on mountPoint; errorQualifier (gpfsErrno)**

**Explanation:** An error occurred while attempting to mount a GPFS file system on Windows.

**User response:** Examine the error details, previous errors, and the GPFS message log to identify the cause.

---

**6027-1511 [E]** **Error unmounting file system**  
*stripeGroup; errorQualifier (gpfsErrno)*

**Explanation:** An error occurred while attempting to unmount a GPFS file system on Windows.

**User response:** Examine the error details, previous errors, and the GPFS message log to identify the cause.

---

**6027-1512 [E]** **WMI query for** *queryType* **failed;**  
*errorQualifier (gpfsErrno)*

**Explanation:** An error occurred while running a WMI query on Windows.

**User response:** Examine the error details, previous errors, and the GPFS message log to identify the cause.

---

**6027-1513**    *DiskName* **is not an sg device, or sg**  
**driver is older than sg3**

**Explanation:** The disk is not a SCSI disk, or supports SCSI standard older than SCSI 3.

**User response:** Correct the command invocation and try again.

---

**6027-1514**    **ioctl failed with rc=returnCode. Related**  
**values are SCSI status=scsiStatusValue,**  
**host\_status=hostStatusValue,**  
**driver\_status=driverStatsValue.**

**Explanation:** An ioctl call failed with stated return code, **errno** value, and related values.

**User response:** Check the reported **errno** and correct the problem if possible. Otherwise, contact the IBM Support Center.

---

**6027-1515**    **READ KEY ioctl failed with**  
**rc=returnCode. Related values are SCSI**  
**status=scsiStatusValue,**  
**host\_status=hostStatusValue,**  
**driver\_status=driverStatsValue.**

**Explanation:** An ioctl call failed with stated return code, **errno** value, and related values.

**User response:** Check the reported **errno** and correct the problem if possible. Otherwise, contact the IBM Support Center.

---

**6027-1516**    **REGISTER ioctl failed with**  
**rc=returnCode. Related values are SCSI**  
**status=scsiStatusValue,**  
**host\_status=hostStatusValue,**  
**driver\_status=driverStatsValue.**

**Explanation:** An ioctl call failed with stated return code, **errno** value, and related values.

**User response:** Check the reported **errno** and correct the problem if possible. Otherwise, contact the IBM Support Center.

---

**6027-1517**    **READ RESERVE ioctl failed with**  
**rc=returnCode. Related values are SCSI**  
**status=scsiStatusValue,**  
**host\_status=hostStatusValue,**  
**driver\_status=driverStatsValue.**

**Explanation:** An ioctl call failed with stated return code, **errno** value, and related values.

**User response:** Check the reported **errno** and correct the problem if possible. Otherwise, contact the IBM Support Center.

---

**6027-1518**    **RESERVE ioctl failed with rc=returnCode.**  
**Related values are SCSI**  
**status=scsiStatusValue,**  
**host\_status=hostStatusValue,**  
**driver\_status=driverStatsValue.**

**Explanation:** An ioctl call failed with stated return code, **errno** value, and related values.

**User response:** Check the reported **errno** and correct the problem if possible. Otherwise, contact the IBM Support Center.

---

**6027-1519**    **INQUIRY ioctl failed with rc=returnCode.**  
**Related values are SCSI**  
**status=scsiStatusValue,**  
**host\_status=hostStatusValue,**  
**driver\_status=driverStatsValue.**

**Explanation:** An `ioctl` call failed with stated return code, `errno` value, and related values.

**User response:** Check the reported `errno` and correct the problem if possible. Otherwise, contact the IBM Support Center.

**6027-1520**    **PREEMPT ABORT ioctl failed with `rc=returnCode`. Related values are SCSI `status=scsiStatusValue`, `host_status=hostStatusValue`, `driver_status=driverStatsValue`.**

**Explanation:** An `ioctl` call failed with stated return code, `errno` value, and related values.

**User response:** Check the reported `errno` and correct the problem if possible. Otherwise, contact the IBM Support Center.

**6027-1521**    **Can not find register key `registerKeyValue` at device `diskName`.**

**Explanation:** Unable to find given register key at the disk.

**User response:** Correct the problem and reissue the command.

**6027-1522**    **CLEAR ioctl failed with `rc=returnCode`. Related values are SCSI `status=scsiStatusValue`, `host_status=hostStatusValue`, `driver_status=driverStatsValue`.**

**Explanation:** An `ioctl` call failed with stated return code, `errno` value, and related values.

**User response:** Check the reported `errno` and correct the problem if possible. Otherwise, contact the IBM Support Center.

**6027-1523**    **Disk name longer than `value` is not allowed.**

**Explanation:** The specified disk name is too long.

**User response:** Reissue the command with a valid disk name.

**6027-1524**    **The `READ_KEYS` ioctl data does not contain the key that was passed as input.**

**Explanation:** A `REGISTER` `ioctl` call apparently succeeded, but when the device was queried for the key, the key was not found.

**User response:** Check the device subsystem and try to correct the problem. If the problem persists, contact the IBM Support Center.

**6027-1530**    **Attention: `parameter` is set to `value`.**

**Explanation:** A configuration parameter is temporarily assigned a new value.

**User response:** Check the `mmfs.cfg` file. Use the `mmchconfig` command to set a valid value for the parameter.

**6027-1531**    *parameter value*

**Explanation:** The configuration parameter was changed from its default value.

**User response:** Check the `mmfs.cfg` file.

**6027-1532**    **Attention: `parameter (value)` is not valid in conjunction with `parameter (value)`.**

**Explanation:** A configuration parameter has a value that is not valid in relation to some other parameter. This can also happen when the default value for some parameter is not sufficiently large for the new, user set value of a related parameter.

**User response:** Check the `mmfs.cfg` file.

**6027-1533**    *parameter cannot be set dynamically.*

**Explanation:** The `mmchconfig` command encountered a configuration parameter that cannot be set dynamically.

**User response:** Check the `mmchconfig` command arguments. If the parameter must be changed, use the `mmshutdown`, `mmchconfig`, and `mmstartup` sequence of commands.

**6027-1534**    *parameter must have a value.*

**Explanation:** The `tsctl` command encountered a configuration parameter that did not have a specified value.

**User response:** Check the `mmchconfig` command arguments.

**6027-1535**    **Unknown config name: `parameter`**

**Explanation:** The `tsctl` command encountered an unknown configuration parameter.

**User response:** Check the `mmchconfig` command arguments.

**6027-1536**    *parameter must be set using the `tschpool` command.*

**Explanation:** The `tsctl` command encountered a configuration parameter that must be set using the `tschpool` command.

**User response:** Check the `mmchconfig` command arguments.

---

**6027-1537 [E] Connect failed to *ipAddress: reason***

**Explanation:** An attempt to connect sockets between nodes failed.

**User response:** Check the reason listed and the connection to the indicated IP address.

---

**6027-1538 [I] Connect in progress to *ipAddress***

**Explanation:** Connecting sockets between nodes.

**User response:** None. Information message only.

---

**6027-1539 [E] Connect progress select failed to *ipAddress: reason***

**Explanation:** An attempt to connect sockets between nodes failed.

**User response:** Check the reason listed and the connection to the indicated IP address.

---

**6027-1540 [A] Try and buy license has expired!**

**Explanation:** Self explanatory.

**User response:** Purchase a GPFS license to continue using GPFS.

---

**6027-1541 [N] Try and buy license expires in *number* days.**

**Explanation:** Self-explanatory.

**User response:** When the **Try and Buy** license expires, you will need to purchase a GPFS license to continue using GPFS.

---

**6027-1542 [A] Old shared memory exists but it is not valid nor cleanable.**

**Explanation:** A new GPFS daemon started and found existing shared segments. The contents were not recognizable, so the GPFS daemon could not clean them up.

**User response:**

1. Stop the GPFS daemon from trying to start by issuing the **mmshutdown** command for the nodes having the problem.
  2. Find the owner of the shared segments with keys from 0x9283a0ca through 0x9283a0d1. If a non-GPFS program owns these segments, GPFS cannot run on this node.
  3. If these segments are left over from a previous GPFS daemon:
    - a. Remove them by issuing:
 

```
ipcrm -m shared_memory_id
```
    - b. Restart GPFS by issuing the **mmstartup** command on the affected nodes.
- 

**6027-1543 error propagating *parameter*.**

**Explanation:** **mmfsd** could not propagate a configuration parameter value to one or more nodes in the cluster.

**User response:** Contact the IBM Support Center.

---

**6027-1544 [W] Sum of **prefetchthreads**(*value*), **worker1threads**(*value*) and **nsdMaxWorkerThreads** (*value*) exceeds *value*. Reducing them to *value*, *value* and *value*.**

**Explanation:** The sum of **prefetchthreads**, **worker1threads**, and **nsdMaxWorkerThreads** exceeds the permitted value.

**User response:** Accept the calculated values or reduce the individual settings using **mmchconfig prefetchthreads=newvalue** or **mmchconfig worker1threads=newvalue**. or **mmchconfig nsdMaxWorkerThreads=newvalue**. After using **mmchconfig**, the new settings will not take affect until the GPFS daemon is restarted.

---

**6027-1545 [A] The GPFS product that you are attempting to run is not a fully functioning version. This probably means that this is an update version and not the full product version. Install the GPFS full product version first, then apply any applicable update version before attempting to start GPFS.**

**Explanation:** GPFS requires a fully licensed GPFS installation.

**User response:** Verify installation of licensed GPFS, or purchase and install a licensed version of GPFS.

---

**6027-1546 [W] Attention: *parameter* size of *value* is too small. New *value* is *value*.**

**Explanation:** A configuration parameter is temporarily assigned a new value.

**User response:** Check the **mmfs.cfg** file. Use the **mmchconfig** command to set a valid value for the parameter.

---

**6027-1547 [A] Error initializing daemon: performing shutdown**

**Explanation:** GPFS kernel extensions are not loaded, and the daemon cannot initialize. GPFS may have been started incorrectly.

**User response:** Check GPFS log for errors resulting from kernel extension loading. Ensure that GPFS is started with the **mmstartup** command.

---

---

**6027-1548 [A] Error: daemon and kernel extension do not match.**

**Explanation:** The GPFS kernel extension loaded in memory and the daemon currently starting do not appear to have come from the same build.

**User response:** Ensure that the kernel extension was reloaded after upgrading GPFS. See “GPFS modules cannot be loaded on Linux” on page 232 for details.

---

**6027-1549 [A] Attention: custom-built kernel extension; the daemon and kernel extension do not match.**

**Explanation:** The GPFS kernel extension loaded in memory does not come from the same build as the starting daemon. The kernel extension appears to have been built from the kernel open source package.

**User response:** None.

---

**6027-1550 [W] Error: Unable to establish a session with an Active Directory server. ID remapping via Microsoft Identity Management for Unix will be unavailable.**

**Explanation:** GPFS tried to establish an LDAP session with an Active Directory server (normally the domain controller host), and has been unable to do so.

**User response:** Ensure the domain controller is available.

---

**6027-1555 Mount point and device name cannot be equal: *name***

**Explanation:** The specified mount point is the same as the absolute device name.

**User response:** Enter a new device name or absolute mount point path name.

---

**6027-1556 Interrupt received.**

**Explanation:** A GPFS administration command received an interrupt.

**User response:** None. Informational message only.

---

**6027-1557 You must first generate an authentication key file. Run: `mmauth genkey new`.**

**Explanation:** Before setting a cipher list, you must generate an authentication key file.

**User response:** Run the specified command to establish an authentication key for the nodes in the cluster.

---



---

**6027-1559 The -i option failed. Changes will take effect after GPFS is restarted.**

**Explanation:** The -i option on the `mmchconfig` command failed. The changes were processed successfully, but will take effect only after the GPFS daemons are restarted.

**User response:** Check for additional error messages. Correct the problem and reissue the command.

---

**6027-1560 This GPFS cluster contains file systems. You cannot delete the last node.**

**Explanation:** An attempt has been made to delete a GPFS cluster that still has one or more file systems associated with it.

**User response:** Before deleting the last node of a GPFS cluster, delete all file systems that are associated with it. This applies to both local and remote file systems.

---

**6027-1561 Attention: Failed to remove node-specific changes.**

**Explanation:** The internal `mmfixcfg` routine failed to remove node-specific configuration settings, if any, for one or more of the nodes being deleted. This is of consequence only if the `mmchconfig` command was indeed used to establish node specific settings and these nodes are later added back into the cluster.

**User response:** If you add the nodes back later, ensure that the configuration parameters for the nodes are set as desired.

---

**6027-1562 *command* command cannot be executed. Either none of the nodes in the cluster are reachable, or GPFS is down on all of the nodes.**

**Explanation:** The command that was issued needed to perform an operation on a remote node, but none of the nodes in the cluster were reachable, or GPFS was not accepting commands on any of the nodes.

**User response:** Ensure that the affected nodes are available and all authorization requirements are met. Correct any problems and reissue the command.

---

**6027-1563 Attention: The file system may no longer be properly balanced.**

**Explanation:** The restripe phase of the `mmaddisk` or `mmdeldisk` command failed.

**User response:** Determine the cause of the failure and run the `mmrestripefs -b` command.

---

---

**6027-1564** To change the authentication key for the local cluster, run: `mmauth genkey`.

**Explanation:** The authentication keys for the local cluster must be created only with the specified command.

**User response:** Run the specified command to establish a new authentication key for the nodes in the cluster.

---

**6027-1565** *disk not found in file system `fileSystem`.*

**Explanation:** A disk specified for deletion or replacement does not exist.

**User response:** Specify existing disks for the indicated file system.

---

**6027-1566** Remote cluster *clusterName* is already defined.

**Explanation:** A request was made to add the cited cluster, but the cluster is already known to GPFS.

**User response:** None. The cluster is already known to GPFS.

---

**6027-1567** *fileSystem* from cluster *clusterName* is already defined.

**Explanation:** A request was made to add the cited file system from the cited cluster, but the file system is already known to GPFS.

**User response:** None. The file system is already known to GPFS.

---

**6027-1568** *command* **command failed. Only `parameterList` changed.**

**Explanation:** The `mmchfs` command failed while making the requested changes. Any changes to the attributes in the indicated parameter list were successfully completed. No other file system attributes were changed.

**User response:** Reissue the command if you want to change additional attributes of the file system. Changes can be undone by issuing the `mmchfs` command with the original value for the affected attribute.

---

**6027-1570** virtual shared disk support is not installed.

**Explanation:** The command detected that IBM Virtual Shared Disk support is not installed on the node on which it is running.

**User response:** Install IBM Virtual Shared Disk support.

---

**6027-1571** *commandName* does not exist or failed; automount mounting may not work.

**Explanation:** One or more of the GPFS file systems were defined with the automount attribute but the requisite automount command is missing or failed.

**User response:** Correct the problem and restart GPFS. Or use the `mount` command to explicitly mount the file system.

---

**6027-1572** The command must run on a node that is part of the cluster.

**Explanation:** The node running the `mmcluster` command (this node) must be a member of the GPFS cluster.

**User response:** Issue the command from a node that will belong to the cluster.

---

**6027-1573** Command completed: No changes made.

**Explanation:** Informational message.

**User response:** Check the preceding messages, correct any problems, and reissue the command.

---

**6027-1574** Permission failure. The command requires root authority to execute.

**Explanation:** The command, or the specified command option, requires root authority.

**User response:** Log on as `root` and reissue the command.

---

**6027-1578** File *fileName* does not contain node names.

**Explanation:** The specified file does not contain valid node names.

**User response:** Node names must be specified one per line. The name `localhost` and lines that start with '#' character are ignored.

---

**6027-1579** File *fileName* does not contain data.

**Explanation:** The specified file does not contain data.

**User response:** Verify that you are specifying the correct file name and reissue the command.

---

**6027-1587** Unable to determine the local device name for disk *nsdName* on node *nodeName*.

**Explanation:** GPFS was unable to determine the local device name for the specified GPFS disk.

**User response:** Determine why the specified disk on the specified node could not be accessed and correct the problem. Possible reasons include: connectivity

problems, authorization problems, fenced disk, and so forth.

---

**6027-1588**    **Unknown GPFS execution environment:**  
*value*

**Explanation:** A GPFS administration command (prefixed by **mm**) was asked to operate on an unknown GPFS cluster type. The only supported GPFS cluster type is **lc**. This message may also be generated if there is corruption in the GPFS system files.

**User response:** Verify that the correct level of GPFS is installed on the node. If this is a cluster environment, make sure the node has been defined as a member of the GPFS cluster with the help of the **mmcluster** or the **mmaddnode** command. If the problem persists, contact the IBM Support Center.

---

**6027-1590**    *nodeName* **cannot be reached.**

**Explanation:** A command needs to issue a remote function on a particular node but the node is not reachable.

**User response:** Determine why the node is unreachable, correct the problem, and reissue the command.

---

**6027-1591**    **Attention: Unable to retrieve GPFS cluster files from node** *nodeName*

**Explanation:** A command could not retrieve the GPFS cluster files from a particular node. An attempt will be made to retrieve the GPFS cluster files from a backup node.

**User response:** None. Informational message only.

---

**6027-1592**    **Unable to retrieve GPFS cluster files from node** *nodeName*

**Explanation:** A command could not retrieve the GPFS cluster files from a particular node.

**User response:** Correct the problem and reissue the command.

---

**6027-1594**    **Run the *command* command until successful.**

**Explanation:** The command could not complete normally. The GPFS cluster data may be left in a state that precludes normal operation until the problem is corrected.

**User response:** Check the preceding messages, correct the problems, and issue the specified command until it completes successfully.

---

**6027-1595**    **No nodes were found that matched the input specification.**

**Explanation:** No nodes were found in the GPFS cluster that matched those specified as input to a GPFS command.

**User response:** Determine why the specified nodes were not valid, correct the problem, and reissue the GPFS command.

---

**6027-1596**    **The same node was specified for both the primary and the secondary server.**

**Explanation:** A command would have caused the primary and secondary GPFS cluster configuration server nodes to be the same.

**User response:** Specify a different primary or secondary node.

---

**6027-1597**    **Node *node* is specified more than once.**

**Explanation:** The same node appears more than once on the command line or in the input file for the command.

**User response:** All specified nodes must be unique. Note that even though two node identifiers may appear different on the command line or in the input file, they may still refer to the same node.

---

**6027-1598**    **Node *nodeName* was not added to the cluster. The node appears to already belong to a GPFS cluster.**

**Explanation:** A GPFS cluster command found that a node to be added to a cluster already has GPFS cluster files on it.

**User response:** Use the **mmclscluster** command to verify that the node is in the correct cluster. If it is not, follow the procedure in "Node cannot be added to the GPFS cluster" on page 227.

---

**6027-1599**    **The level of GPFS on node *nodeName* does not support the requested action.**

**Explanation:** A GPFS command found that the level of the GPFS code on the specified node is not sufficient for the requested action.

**User response:** Install the correct level of GPFS.

---

**6027-1600**    **Make sure that the following nodes are available:** *nodeList*

**Explanation:** A GPFS command was unable to complete because nodes critical for the success of the operation were not reachable or the command was interrupted.

**User response:** This message will normally be



followed by a message telling you which command to issue as soon as the problem is corrected and the specified nodes become available.

---

**6027-1602** *nodeName* is not a member of this cluster.

**Explanation:** A command found that the specified node is not a member of the GPFS cluster.

**User response:** Correct the input or add the node to the GPFS cluster and reissue the command.

---

**6027-1603** The following nodes could not be added to the GPFS cluster: *nodeList*. Correct the problems and use the **mmaddnode** command to add these nodes to the cluster.

**Explanation:** The **mmcrcluster** or the **mmaddnode** command was unable to add the listed nodes to a GPFS cluster.

**User response:** Correct the problems and add the nodes to the cluster using the **mmaddnode** command.

---

**6027-1604** Information cannot be displayed. Either none of the nodes in the cluster are reachable, or GPFS is down on all of the nodes.

**Explanation:** The command needed to perform an operation on a remote node, but none of the nodes in the cluster were reachable, or GPFS was not accepting commands on any of the nodes.

**User response:** Ensure that the affected nodes are available and all authorization requirements are met. Correct any problems and reissue the command.

---

**6027-1610** Disk *diskName* is the only disk in file system *fileSystem*. You cannot replace a disk when it is the only remaining disk in the file system.

**Explanation:** The **mmrpldisk** command was issued, but there is only one disk in the file system.

**User response:** Add a second disk and reissue the command.

---

**6027-1613** WCOLL (working collective) environment variable not set.

**Explanation:** The **mmdsh** command was invoked without explicitly specifying the nodes on which the command is to run by means of the **-F** or **-L** options, and the WCOLL environment variable has not been set.

**User response:** Change the invocation of the **mmdsh** command to use the **-F** or **-L** options, or set the WCOLL environment variable before invoking the **mmdsh** command.

---

**6027-1614** Cannot open file *fileName*. Error string was: *errorString*.

**Explanation:** The **mmdsh** command was unable to successfully open a file.

**User response:** Determine why the file could not be opened and correct the problem.

---

**6027-1615** *nodeName* remote shell process had return code *value*.

**Explanation:** A child remote shell process completed with a nonzero return code.

**User response:** Determine why the child remote shell process failed and correct the problem.

---

**6027-1616** Caught SIG *signal* - terminating the child processes.

**Explanation:** The **mmdsh** command has received a signal causing it to terminate.

**User response:** Determine what caused the signal and correct the problem.

---

**6027-1617** There are no available nodes on which to run the command.

**Explanation:** The **mmdsh** command found that there are no available nodes on which to run the specified command. Although nodes were specified, none of the nodes were reachable.

**User response:** Determine why the specified nodes were not available and correct the problem.

---

**6027-1618** Unable to pipe. Error string was: *errorString*.

**Explanation:** The **mmdsh** command attempted to open a pipe, but the pipe command failed.

**User response:** Determine why the call to pipe failed and correct the problem.

---

**6027-1619** Unable to redirect *outputStream*. Error string was: *string*.

**Explanation:** The **mmdsh** command attempted to redirect an output stream using **open**, but the **open** command failed.

**User response:** Determine why the call to **open** failed and correct the problem.

---

**6027-1623** *command*: Mounting file systems ...

**Explanation:** This message contains progress information about the **mmmount** command.

**User response:** None. Informational message only.

---

**6027-1625** *option cannot be used with attribute name.*

**Explanation:** An attempt was made to change a configuration attribute and requested the change to take effect immediately (-i or -I option). However, the specified attribute does not allow the operation.

**User response:** If the change must be made now, leave off the -i or -I option. Then recycle the nodes to pick up the new value.

---

**6027-1626** **Command is not supported in the *type* environment.**

**Explanation:** A GPFS administration command (mm...) is not supported in the specified environment.

**User response:** Verify if the task is needed in this environment, and if it is, use a different command.

---

**6027-1627** **The following nodes are not aware of the configuration server change: *nodeList*. Do not start GPFS on the above nodes until the problem is resolved.**

**Explanation:** The mmchcluster command could not propagate the new cluster configuration servers to the specified nodes.

**User response:** Correct the problems and run the mmchcluster -p LATEST command before starting GPFS on the specified nodes.

---

**6027-1628** **Cannot determine basic environment information. Not enough nodes are available.**

**Explanation:** The mmchcluster command was unable to retrieve the GPFS cluster data files. Usually, this is due to too few nodes being available.

**User response:** Correct any problems and ensure that as many of the nodes in the cluster are available as possible. Reissue the command. If the problem persists, record the above information and contact the IBM Support Center.

---

**6027-1629** **Error found while checking node descriptor *descriptor***

**Explanation:** A node descriptor was found to be unsatisfactory in some way.

**User response:** Check the preceding messages, if any, and correct the condition that caused the disk descriptor to be rejected.

---



---

**6027-1630** **The GPFS cluster data on *nodeName* is back level.**

**Explanation:** A GPFS command attempted to commit changes to the GPFS cluster configuration data, but the data on the server is already at a higher level. This can happen if the GPFS cluster configuration files were altered outside the GPFS environment, or if the mmchcluster command did not complete successfully.

**User response:** Correct any problems and reissue the command. If the problem persists, issue the mmrefresh -f -a command.

---

**6027-1631** **The commit process failed.**

**Explanation:** A GPFS administration command (mm...) cannot commit its changes to the GPFS cluster configuration data.

**User response:** Examine the preceding messages, correct the problem, and reissue the command. If the problem persists, perform problem determination and contact the IBM Support Center.

---

**6027-1632** **The GPFS cluster configuration data on *nodeName* is different than the data on *nodeName*.**

**Explanation:** The GPFS cluster configuration data on the primary cluster configuration server node is different than the data on the secondary cluster configuration server node. This can happen if the GPFS cluster configuration files were altered outside the GPFS environment or if the mmchcluster command did not complete successfully.

**User response:** Correct any problems and issue the mmrefresh -f -a command. If the problem persists, perform problem determination and contact the IBM Support Center.

---

**6027-1633** **Failed to create a backup copy of the GPFS cluster data on *nodeName*.**

**Explanation:** Commit could not create a correct copy of the GPFS cluster configuration data.

**User response:** Check the preceding messages, correct any problems, and reissue the command. If the problem persists, perform problem determination and contact the IBM Support Center.

---

**6027-1634** **The GPFS cluster configuration server node *nodeName* cannot be removed.**

**Explanation:** An attempt was made to delete a GPFS cluster configuration server node.

**User response:** You cannot remove a cluster configuration server node unless all nodes in the GPFS cluster are being deleted. Before deleting a cluster configuration server node, you must use the

**mmchcluster** command to transfer its function to another node in the GPFS cluster.

---

**6027-1636**    **Error found while checking disk descriptor** *descriptor*

**Explanation:** A disk descriptor was found to be unsatisfactory in some way.

**User response:** Check the preceding messages, if any, and correct the condition that caused the disk descriptor to be rejected.

---

**6027-1637**    *command* **quitting. None of the specified nodes are valid.**

**Explanation:** A GPFS command found that none of the specified nodes passed the required tests.

**User response:** Determine why the nodes were not accepted, fix the problems, and reissue the command.

---

**6027-1638**    *Command:* **There are no unassigned nodes in the cluster.**

**Explanation:** A GPFS command in a cluster environment needs unassigned nodes, but found there are none.

**User response:** Verify whether there are any unassigned nodes in the cluster. If there are none, either add more nodes to the cluster using the **mmaddnode** command, or delete some nodes from the cluster using the **mmdelnode** command, and then reissue the command.

---

**6027-1639**    **Command failed. Examine previous error messages to determine cause.**

**Explanation:** A GPFS command failed due to previously-reported errors.

**User response:** Check the previous error messages, fix the problems, and then reissue the command. If no other messages are shown, examine the GPFS log files in the **/var/adm/ras** directory on each node.

---

**6027-1642**    *command:* **Starting GPFS ...**

**Explanation:** Progress information for the **mmstartup** command.

**User response:** None. Informational message only.

---

**6027-1643**    **The number of quorum nodes exceeds the maximum (number) allowed.**

**Explanation:** An attempt was made to add more quorum nodes to a cluster than the maximum number allowed.

**User response:** Reduce the number of quorum nodes, and reissue the command.

---

**6027-1644**    **Attention: The number of quorum nodes exceeds the suggested maximum (number).**

**Explanation:** The number of quorum nodes in the cluster exceeds the maximum suggested number of quorum nodes.

**User response:** Informational message. Consider reducing the number of quorum nodes to the maximum suggested number of quorum nodes for improved performance.

---

**6027-1645**    **Node *nodeName* is fenced out from disk *diskName*.**

**Explanation:** A GPFS command attempted to access the specified disk, but found that the node attempting the operation was fenced out from the disk.

**User response:** Check whether there is a valid reason why the node should be fenced out from the disk. If there is no such reason, unfence the disk and reissue the command.

---

**6027-1647**    **Unable to find disk with NSD volume id *NSD volume id*.**

**Explanation:** A disk with the specified NSD volume id cannot be found.

**User response:** Specify a correct disk NSD volume id.

---

**6027-1648**    **GPFS was unable to obtain a lock from node *nodeName*.**

**Explanation:** GPFS failed in its attempt to get a lock from another node in the cluster.

**User response:** Verify that the reported node is reachable. Examine previous error messages, if any. Fix the problems and then reissue the command.

---

**6027-1661**    **Failed while processing disk descriptor *descriptor* on node *nodeName*.**

**Explanation:** A disk descriptor was found to be unsatisfactory in some way.

**User response:** Check the preceding messages, if any, and correct the condition that caused the disk descriptor to be rejected.

---

**6027-1662**    **Disk device *deviceName* refers to an existing NSD *name***

**Explanation:** The specified disk device refers to an existing NSD.

**User response:** Specify another disk that is not an existing NSD.

---

**6027-1663** Disk descriptor *descriptor* should refer to an existing NSD. Use `mmcrnsd` to create the NSD.

**Explanation:** An NSD disk given as input is not known to GPFS.

**User response:** Create the NSD. Then rerun the command.

---

**6027-1664** *command: Processing node nodeName*

**Explanation:** Progress information.

**User response:** None. Informational message only.

---

**6027-1665** Issue the command from a node that remains in the cluster.

**Explanation:** The nature of the requested change requires the command be issued from a node that will remain in the cluster.

**User response:** Run the command from a node that will remain in the cluster.

---

**6027-1666 [I]** No disks were found.

**Explanation:** A command searched for disks but found none.

**User response:** If disks are desired, create some using the `mmcrnsd` command.

---

**6027-1670** Incorrect or missing remote shell command: *name*

**Explanation:** The specified remote command does not exist or is not executable.

**User response:** Specify a valid command.

---

**6027-1671** Incorrect or missing remote file copy command: *name*

**Explanation:** The specified remote command does not exist or is not executable.

**User response:** Specify a valid command.

---

**6027-1672** *option value* parameter must be an absolute path name.

**Explanation:** The mount point does not begin with '/'.

**User response:** Specify the full path for the mount point.

---

**6027-1674** *command: Unmounting file systems ...*

**Explanation:** This message contains progress information about the `mmumount` command.

**User response:** None. Informational message only.

---

**6027-1677** Disk *diskName* is of an unknown type.

**Explanation:** The specified disk is of an unknown type.

**User response:** Specify a disk whose type is recognized by GPFS.

---

**6027-1680** Disk name *diskName* is already registered for use by GPFS.

**Explanation:** The cited disk name was specified for use by GPFS, but there is already a disk by that name registered for use by GPFS.

**User response:** Specify a different disk name for use by GPFS and reissue the command.

---

**6027-1681** Node *nodeName* is being used as an NSD server.

**Explanation:** The specified node is defined as a server node for some disk.

**User response:** If you are trying to delete the node from the GPFS cluster, you must either delete the disk or define another node as its server.

---

**6027-1685** Processing continues without lock protection.

**Explanation:** The command will continue processing although it was not able to obtain the lock that prevents other GPFS commands from running simultaneously.

**User response:** Ensure that no other GPFS command is running. See the command documentation for additional details.

---

**6027-1688** Command was unable to obtain the lock for the GPFS system data. Unable to reach the holder of the lock *nodeName*. Check the preceding messages, if any. Follow the procedure outlined in the GPFS: Problem Determination Guide.

**Explanation:** A command requires the lock for the GPFS system data but was not able to obtain it.

**User response:** Check the preceding messages, if any. Follow the procedure in the *IBM Spectrum Scale: Problem Determination Guide* for what to do when the GPFS system data is locked. Then reissue the command.

---

**6027-1689** vpath disk *diskName* is not recognized as an IBM SDD device.

**Explanation:** The `mmvsdhelper` command found that the specified disk is a vpath disk, but it is not recognized as an IBM SDD device.

**User response:** Ensure the disk is configured as an IBM SDD device. Then reissue the command.

---

**6027-1699**    **Remount failed for file system**  
*fileSystem. Error code errorCode.*

**Explanation:** The specified file system was internally unmounted. An attempt to remount the file system failed with the specified error code.

**User response:** Check the daemon log for additional error messages. Ensure that all file system disks are available and reissue the **mount** command.

---

**6027-1700**    **Failed to load LAPI library. functionName**  
**not found. Changing communication**  
**protocol to TCP.**

**Explanation:** The GPFS daemon failed to load **liblapi\_r.a** dynamically.

**User response:** Verify installation of **liblapi\_r.a**.

---

**6027-1701**    **mmfsd waiting to connect to**  
**mmspsecserver. Setting up to retry every**  
**number seconds for number minutes.**

**Explanation:** The GPFS daemon failed to establish a connection with the **mmspsecserver** process.

**User response:** None. Informational message only.

---

**6027-1702**    **Process pid failed at functionName call,**  
**socket socketName, errno value**

**Explanation:** Either The **mmfsd** daemon or the **mmspsecserver** process failed to create or set up the communication socket between them.

**User response:** Determine the reason for the error.

---

**6027-1703**    **The processName process encountered**  
**error: errorString.**

**Explanation:** Either the **mmfsd** daemon or the **mmspsecserver** process called the error log routine to log an incident.

**User response:** None. Informational message only.

---

**6027-1704**    **mmspsecserver (pid number) ready for**  
**service.**

**Explanation:** The **mmspsecserver** process has created all the service threads necessary for **mmfsd**.

**User response:** None. Informational message only.

---

**6027-1705**    **command: incorrect number of**  
**connections (number), exiting...**

**Explanation:** The **mmspsecserver** process was called with an incorrect number of connections. This will happen only when the **mmspsecserver** process is run as an independent program.

**User response:** Retry with a valid number of connections.

---

**6027-1706**    **mmspsecserver: parent program is not**  
**"mmfsd", exiting...**

**Explanation:** The **mmspsecserver** process was invoked from a program other than **mmfsd**.

**User response:** None. Informational message only.

---

**6027-1707**    **mmfsd connected to mmspsecserver**

**Explanation:** The **mmfsd** daemon has successfully connected to the **mmspsecserver** process through the communication socket.

**User response:** None. Informational message only.

---

**6027-1708**    **The mmfsd daemon failed to fork**  
**mmspsecserver. Failure reason**  
*explanation*

**Explanation:** The **mmfsd** daemon failed to fork a child process.

**User response:** Check the GPFS installation.

---

**6027-1709 [I]** **Accepted and connected to ipAddress**

**Explanation:** The local **mmfsd** daemon has successfully accepted and connected to a remote daemon.

**User response:** None. Informational message only.

---

**6027-1710 [N]** **Connecting to ipAddress**

**Explanation:** The local **mmfsd** daemon has started a connection request to a remote daemon.

**User response:** None. Informational message only.

---

**6027-1711 [I]** **Connected to ipAddress**

**Explanation:** The local **mmfsd** daemon has successfully connected to a remote daemon.

**User response:** None. Informational message only.

---

**6027-1712**    **Unexpected zero bytes received from**  
**name. Continuing.**

**Explanation:** This is an informational message. A socket read resulted in zero bytes being read.

## 6027-1715 • 6027-1731 [E]

**User response:** If this happens frequently, check IP connections.

---

**6027-1715** **EINVAL trap from connect call to *ipAddress* (socket name)**

**Explanation:** The connect call back to the requesting node failed.

**User response:** This is caused by a bug in AIX socket support. Upgrade AIX kernel and TCP client support.

---

**6027-1716 [N]** **Close connection to *ipAddress***

**Explanation:** Connection socket closed.

**User response:** None. Informational message only.

---

**6027-1717 [E]** **Error initializing the configuration server, err *value***

**Explanation:** The configuration server module could not be initialized due to lack of system resources.

**User response:** Check system memory.

---

**6027-1718 [E]** **Could not run command *name*, err *value***

**Explanation:** The GPFS daemon failed to run the specified command.

**User response:** Verify correct installation.

---

**6027-1724 [E]** **The key used by the cluster named *clusterName* has changed. Contact the administrator to obtain the new key and register it using "mmremotecluster update".**

**Explanation:** The administrator of the cluster has changed the key used for authentication.

**User response:** Contact the administrator to obtain the new key and register it using **mmremotecluster update**.

---

**6027-1725 [E]** **The key used by the cluster named *clusterName* has changed. Contact the administrator to obtain the new key and register it using "mmauth update".**

**Explanation:** The administrator of the cluster has changed the key used for authentication.

**User response:** Contact the administrator to obtain the new key and register it using **mmauth update**.

---

**6027-1726 [E]** **The administrator of the cluster named *clusterName* requires authentication. Contact the administrator to obtain the clusters key and register the key using "mmremotecluster update".**

**Explanation:** The administrator of the cluster requires authentication.

**User response:** Contact the administrator to obtain the cluster's key and register it using: **mmremotecluster update**.

---

**6027-1727 [E]** **The administrator of the cluster named *clusterName* does not require authentication. Unregister the clusters key using "mmremotecluster update".**

**Explanation:** The administrator of the cluster does not require authentication.

**User response:** Unregister the clusters key using: **mmremotecluster update**.

---

**6027-1728 [E]** **Remote mounts are not enabled within the cluster named *clusterName*. Contact the administrator and request that they enable remote mounts.**

**Explanation:** The administrator of the cluster has not enabled remote mounts.

**User response:** Contact the administrator and request remote mount access.

---

**6027-1729 [E]** **The cluster named *clusterName* has not authorized this cluster to mount file systems. Contact the cluster administrator and request access.**

**Explanation:** The administrator of the cluster has not authorized this cluster to mount file systems.

**User response:** Contact the administrator and request access.

---

**6027-1730 [E]** **Unsupported cipherList *cipherList* requested.**

**Explanation:** The target cluster requested a **cipherList** not supported by the installed version of OpenSSL.

**User response:** Install a version of OpenSSL that supports the required **cipherList** or contact the administrator of the target cluster and request that a supported **cipherList** be assigned to this remote cluster.

---

**6027-1731 [E]** **Unsupported cipherList *cipherList* requested.**

**Explanation:** The target cluster requested a **cipherList** that is not supported by the installed version of OpenSSL.

**User response:** Either install a version of OpenSSL that supports the required **cipherList** or contact the administrator of the target cluster and request that a supported **cipherList** be assigned to this remote cluster.

---

**6027-1732 [X] Remote mounts are not enabled within this cluster.**

**Explanation:** Remote mounts cannot be performed in this cluster.

**User response:** See the *IBM Spectrum Scale: Administration Guide* for instructions about enabling remote mounts. In particular, make sure the keys have been generated and a **cipherlist** has been set.

---

**6027-1733 OpenSSL dynamic lock support could not be loaded.**

**Explanation:** One of the functions required for dynamic lock support was not included in the version of the OpenSSL library that GPFS is configured to use.

**User response:** If this functionality is required, shut down the daemon, install a version of OpenSSL with the desired functionality, and configure GPFS to use it. Then restart the daemon.

---

**6027-1734 [E] OpenSSL engine support could not be loaded.**

**Explanation:** One of the functions required for engine support was not included in the version of the OpenSSL library that GPFS is configured to use.

**User response:** If this functionality is required, shut down the daemon, install a version of OpenSSL with the desired functionality, and configure GPFS to use it. Then restart the daemon.

---

**6027-1735 [E] Close connection to *ipAddress*. Attempting reconnect.**

**Explanation:** Connection socket closed. The GPFS daemon will attempt to reestablish the connection.

**User response:** None. Informational message only.

---

**6027-1736 [N] Reconnected to *ipAddress***

**Explanation:** The local **mmfsd** daemon has successfully reconnected to a remote daemon following an unexpected connection break.

**User response:** None. Informational message only.

---

**6027-1737 [N] Close connection to *ipAddress* (*errorString*).**

**Explanation:** Connection socket closed.

**User response:** None. Informational message only.

---



---

**6027-1738 [E] Close connection to *ipAddress* (*errorString*). Attempting reconnect.**

**Explanation:** Connection socket closed.

**User response:** None. Informational message only.

---

**6027-1739 [X] Accept socket connection failed: *err value*.**

**Explanation:** The Accept socket connection received an unexpected error.

**User response:** None. Informational message only.

---

**6027-1740 [E] Timed out waiting for a reply from node *ipAddress*.**

**Explanation:** A message that was sent to the specified node did not receive a response within the expected time limit.

**User response:** None. Informational message only.

---

**6027-1741 [E] Error code *value* received from node *ipAddress*.**

**Explanation:** When a message was sent to the specified node to check its status, an error occurred and the node could not handle the message.

**User response:** None. Informational message only.

---

**6027-1742 [E] Message ID *value* was lost by node *ipAddress*.**

**Explanation:** During a periodic check of outstanding messages, a problem was detected where the destination node no longer has any knowledge of a particular message.

**User response:** None. Informational message only.

---

**6027-1743 [W] Failed to load GSKit library *path*: (*dlerror*) *errorMessage***

**Explanation:** The GPFS daemon could not load the library required to secure the node-to-node communications.

**User response:** Verify that the **gpfs.gskit** package was properly installed.

---

**6027-1744 [I] GSKit library loaded and initialized.**

**Explanation:** The GPFS daemon successfully loaded the library required to secure the node-to-node communications.

**User response:** None. Informational message only.

---

---

**6027-1745 [E] Unable to resolve symbol for routine:**  
*functionName (dlerror) errorMessage*

**Explanation:** An error occurred while resolving a symbol required for transport-level security.

**User response:** Verify that the gpfs.gskit package was properly installed.

---

**6027-1746 [E] Failed to load or initialize GSKit library: error value**

**Explanation:** An error occurred during the initialization of the transport-security code.

**User response:** Verify that the gpfs.gskit package was properly installed.

---

**6027-1747 [W] The TLS handshake with node ipAddress failed with error value (handshakeType).**

**Explanation:** An error occurred while trying to establish a secure connection with another GPFS node.

**User response:** Examine the error messages to obtain information about the error. Under normal circumstances, the retry logic will ensure that the connection is re-established. If this error persists, record the error code and contact the IBM Support Center.

---

**6027-1748 [W] A secure receive from node ipAddress failed with error value.**

**Explanation:** An error occurred while receiving an encrypted message from another GPFS node.

**User response:** Examine the error messages to obtain information about the error. Under normal circumstances, the retry logic will ensure that the connection is re-established and the message is received. If this error persists, record the error code and contact the IBM Support Center.

---

**6027-1749 [W] A secure send to node ipAddress failed with error value.**

**Explanation:** An error occurred while sending an encrypted message to another GPFS node.

**User response:** Examine the error messages to obtain information about the error. Under normal circumstances, the retry logic will ensure that the connection is re-established and the message is sent. If this error persists, record the error code and contact the IBM Support Center.

---

**6027-1750 [N] The handshakeType TLS handshake with node ipAddress was cancelled: connection reset by peer (return code value).**

**Explanation:** A secure connection could not be

established because the remote GPFS node closed the connection.

**User response:** None. Informational message only.

---

**6027-1751 [N] A secure send to node ipAddress was cancelled: connection reset by peer (return code value).**

**Explanation:** Securely sending a message failed because the remote GPFS node closed the connection.

**User response:** None. Informational message only.

---

**6027-1752 [N] A secure receive to node ipAddress was cancelled: connection reset by peer (return code value).**

**Explanation:** Securely receiving a message failed because the remote GPFS node closed the connection.

**User response:** None. Informational message only.

---

**6027-1753 [E] The crypto library with FIPS support is not available for this architecture. Disable FIPS mode and reattempt the operation.**

**Explanation:** GPFS is operating in FIPS mode, but the initialization of the cryptographic library failed because FIPS mode is not yet supported on this architecture.

**User response:** Disable FIPS mode and attempt the operation again.

---

**6027-1754 [E] Failed to initialize the crypto library in FIPS mode. Ensure that the crypto library package was correctly installed.**

**Explanation:** GPFS is operating in FIPS mode, but the initialization of the cryptographic library failed.

**User response:** Ensure that the packages required for encryption are properly installed on each node in the cluster.

---

**6027-1803 [E] Global NSD disk, name, not found.**

**Explanation:** A client tried to open a globally-attached NSD disk, but a scan of all disks failed to find that NSD.

**User response:** Ensure that the globally-attached disk is available on every node that references it.

---

**6027-1804 [E] I/O to NSD disk, name, fails. No such NSD locally found.**

**Explanation:** A server tried to perform I/O on an NSD disk, but a scan of all disks failed to find that NSD.

**User response:** Make sure that the NSD disk is accessible to the client. If necessary, break a reservation.



---

**6027-1805 [N]** Rediscovered nsd server access to *name*.

**Explanation:** A server rediscovered access to the specified disk.

**User response:** None.

---

**6027-1806 [X]** A Persistent Reserve could not be established on device name (*deviceName*): *errorLine*.

**Explanation:** GPFS is using Persistent Reserve on this disk, but was unable to establish a reserve for this node.

**User response:** Perform disk diagnostics.

---

**6027-1807 [E]** NSD *nsdName* is using Persistent Reserve, this will require an NSD server on an *osName* node.

**Explanation:** A client tried to open a globally-attached NSD disk, but the disk is using Persistent Reserve. An *osName* NSD server is needed. GPFS only supports Persistent Reserve on certain operating systems.

**User response:** Use the `mmchnsd` command to add an *osName* NSD server for the NSD.

---

**6027-1808 [A]** Unable to reserve space for NSD buffers. Increase pagepool size to at least *requiredPagePoolSize* MB. Refer to the *IBM Spectrum Scale: Administration Guide* for more information on selecting an appropriate pagepool size.

**Explanation:** The `pagepool` usage for an NSD buffer ( $4 * \text{maxblocksize}$ ) is limited by factor `nsdBufSpace`. The value of `nsdBufSpace` can be in the range of 10-70. The default value is 30.

**User response:** Use the `mmchconfig` command to decrease the value of `maxblocksize` or to increase the value of `pagepool` or `nsdBufSpace`.

---

**6027-1809 [E]** The defined server *serverName* for NSD *NsdName* couldn't be resolved.

**Explanation:** The host name of the NSD server could not be resolved by `gethostbyname()`.

**User response:** Fix the host name resolution.

---

**6027-1810 [I]** Vdisk server recovery: delay *number sec*. for safe recovery.

**Explanation:** Wait for the existing disk lease to expire before performing vdisk server recovery.

**User response:** None.

---



---

**6027-1811 [I]** Vdisk server recovery: delay complete.

**Explanation:** Done waiting for existing disk lease to expire before performing vdisk server recovery.

**User response:** None.

---

**6027-1812 [E]** Rediscovery failed for *name*.

**Explanation:** A server failed to rediscover access to the specified disk.

**User response:** Check the disk access issues and run the command again.

---

**6027-1813 [A]** Error reading volume identifier (for *objectName name*) from configuration file.

**Explanation:** The volume identifier for the named recovery group or vdisk could not be read from the `mmsdrfs` file. This should never occur.

**User response:** Check for damage to the `mmsdrfs` file.

---

**6027-1814 [E]** Vdisk *vdiskName* cannot be associated with its recovery group *recoveryGroupName*. This vdisk will be ignored.

**Explanation:** The named vdisk cannot be associated with its recovery group.

**User response:** Check for damage to the `mmsdrfs` file.

---

**6027-1815 [A]** Error reading volume identifier (for NSD *name*) from configuration file.

**Explanation:** The volume identifier for the named NSD could not be read from the `mmsdrfs` file. This should never occur.

**User response:** Check for damage to the `mmsdrfs` file.

---

**6027-1816 [E]** The defined server *serverName* for recovery group *recoveryGroupName* could not be resolved.

**Explanation:** The hostname of the NSD server could not be resolved by `gethostbyname()`.

**User response:** Fix hostname resolution.

---

**6027-1817 [E]** Vdisks are defined, but no recovery groups are defined.

**Explanation:** There are vdisks defined in the `mmsdrfs` file, but no recovery groups are defined. This should never occur.

**User response:** Check for damage to the `mmsdrfs` file.

---

---

**6027-1818 [I] Relinquished recovery group**  
*recoveryGroupName (err errorCode).*

**Explanation:** This node has relinquished serving the named recovery group.

**User response:** None.

---

**6027-1819 Disk descriptor for *name* refers to an existing pdisk.**

**Explanation:** The **mmcrrecoverygroup** command or **mmaddpdisk** command found an existing pdisk.

**User response:** Correct the input file, or use the **-v** option.

---

**6027-1820 Disk descriptor for *name* refers to an existing NSD.**

**Explanation:** The **mmcrrecoverygroup** command or **mmaddpdisk** command found an existing NSD.

**User response:** Correct the input file, or use the **-v** option.

---

**6027-1821 Error *errno* writing disk descriptor on *name*.**

**Explanation:** The **mmcrrecoverygroup** command or **mmaddpdisk** command got an error writing the disk descriptor.

**User response:** Perform disk diagnostics.

---

**6027-1822 Error *errno* reading disk descriptor on *name*.**

**Explanation:** The **tspreparedpdisk** command got an error reading the disk descriptor.

**User response:** Perform disk diagnostics.

---

**6027-1823 Path error, *name* and *name* are the same disk.**

**Explanation:** The **tspreparedpdisk** command got an error during path verification. The pdisk descriptor file is miscoded.

**User response:** Correct the pdisk descriptor file and reissue the command.

---

**6027-1824 [X] An unexpected Device Mapper path *dmDevice (nsdId)* has been detected. The new path does not have a Persistent Reserve set up. Server disk *diskName* will be put offline**

**Explanation:** A new device mapper path is detected or a previously failed path is activated after the local device discovery has finished. This path lacks a Persistent Reserve, and cannot be used. All device

paths must be active at mount time.

**User response:** Check the paths to all disks making up the file system. Repair any paths to disks which have failed. Rediscover the paths for the NSD.

---

**6027-1825 [A] Unrecoverable NSD checksum error on I/O to NSD disk *nsdName*, using server *serverName*. Exceeds retry limit *number*.**

**Explanation:** The allowed number of retries was exceeded when encountering an NSD checksum error on I/O to the indicated disk, using the indicated server.

**User response:** There may be network issues that require investigation.

---

**6027-1826 [W] The host name of the server *serverName* that is defined for NSD local cache *NsdName* could not be resolved.**

**Explanation:** The host name of NSD server could not be resolved by **gethostbyname()**.

**User response:** Fix host name resolution.

---

**6027-1900 Failed to stat *pathName*.**

**Explanation:** A **stat()** call failed for the specified object.

**User response:** Correct the problem and reissue the command.

---

**6027-1901 *pathName* is not a GPFS file system object.**

**Explanation:** The specified path name does not resolve to an object within a mounted GPFS file system.

**User response:** Correct the problem and reissue the command.

---

**6027-1902 The policy file cannot be determined.**

**Explanation:** The command was not able to retrieve the policy rules associated with the file system.

**User response:** Examine the preceding messages and correct the reported problems. Establish a valid policy file with the **mmchpolicy** command or specify a valid policy file on the command line.

---

**6027-1903 *path* must be an absolute path name.**

**Explanation:** The path name did not begin with a **/**.

**User response:** Specify the absolute path name for the object.

---

---

**6027-1904**    **Device with major/minor numbers  
number and number already exists.**

**Explanation:** A device with the cited major and minor numbers already exists.

**User response:** Check the preceding messages for detailed information.

---

**6027-1905**    **name was not created by GPFS or could  
not be refreshed.**

**Explanation:** The attributes (device type, major/minor number) of the specified file system device name are not as expected.

**User response:** Check the preceding messages for detailed information on the current and expected values. These errors are most frequently caused by the presence of `/dev` entries that were created outside the GPFS environment. Resolve the conflict by renaming or deleting the offending entries. Reissue the command letting GPFS create the `/dev` entry with the appropriate parameters.

---

**6027-1906**    **There is no file system with drive letter  
driveLetter.**

**Explanation:** No file system in the GPFS cluster has the specified drive letter.

**User response:** Reissue the command with a valid file system.

---

**6027-1908**    **The option option is not allowed for  
remote file systems.**

**Explanation:** The specified option can be used only for locally-owned file systems.

**User response:** Correct the command line and reissue the command.

---

**6027-1909**    **There are no available free disks. Disks  
must be prepared prior to invoking  
command. Define the disks using the  
command command.**

**Explanation:** The currently executing command (`mmcrfs`, `mmadddisk`, `mmrpldisk`) requires disks to be defined for use by GPFS using one of the GPFS disk creation commands: `mmcrnsd`, `mmcrvsd`.

**User response:** Create disks and reissue the failing command.

---

**6027-1910**    **Node nodeName is not a quorum node.**

**Explanation:** The `mmchmgr` command was asked to move the cluster manager to a nonquorum node. Only one of the quorum nodes can be a cluster manager.

**User response:** Designate the node to be a quorum

node, specify a different node on the command line, or allow GPFS to choose the new cluster manager node.

---

**6027-1911**    **File system fileSystem belongs to cluster  
clusterName. The option option is not  
allowed for remote file systems.**

**Explanation:** The specified option can be used only for locally-owned file systems.

**User response:** Correct the command line and reissue the command.

---

**6027-1922**    **IP aliasing is not supported (node).  
Specify the main device.**

**Explanation:** IP aliasing is not supported.

**User response:** Specify a node identifier that resolves to the IP address of a main device for the node.

---

**6027-1927**    **The requested disks are not known to  
GPFS.**

**Explanation:** GPFS could not find the requested NSDs in the cluster.

**User response:** Reissue the command, specifying known disks.

---

**6027-1929**    **cipherlist is not a valid cipher list.**

**Explanation:** The cipher list must be set to a value supported by GPFS. All nodes in the cluster must support a common cipher.

**User response:** Use `mmauth show ciphers` to display a list of the supported ciphers.

---

**6027-1930**    **Disk diskName belongs to file system  
fileSystem.**

**Explanation:** A GPFS administration command (`mm...`) found that the requested disk to be deleted still belongs to a file system.

**User response:** Check that the correct disk was requested. If so, delete the disk from the file system before proceeding.

---

**6027-1931**    **The following disks are not known to  
GPFS: diskNames.**

**Explanation:** A GPFS administration command (`mm...`) found that the specified disks are not known to GPFS.

**User response:** Verify that the correct disks were requested.

---

---

**6027-1932** No disks were specified that could be deleted.

**Explanation:** A GPFS administration command (**mm...**) determined that no disks were specified that could be deleted.

**User response:** Examine the preceding messages, correct the problems, and reissue the command.

---

**6027-1933** Disk *diskName* has been removed from the GPFS cluster configuration data but the NSD volume id was not erased from the disk. To remove the NSD volume id, issue **mmdelnsd -p NSDvolumeid**.

**Explanation:** A GPFS administration command (**mm...**) successfully removed the specified disk from the GPFS cluster configuration data, but was unable to erase the NSD volume id from the disk.

**User response:** Issue the specified command to remove the NSD volume id from the disk.

---

**6027-1934** Disk *diskName* has been removed from the GPFS cluster configuration data but the NSD volume id was not erased from the disk. To remove the NSD volume id, issue: **mmdelnsd -p NSDvolumeid -N nodeList**.

**Explanation:** A GPFS administration command (**mm...**) successfully removed the specified disk from the GPFS cluster configuration data but was unable to erase the NSD volume id from the disk.

**User response:** Issue the specified command to remove the NSD volume id from the disk.

---

**6027-1936** Node *nodeName* cannot support Persistent Reserve on disk *diskName* because it is not an AIX node. The disk will be used as a non-PR disk.

**Explanation:** A non-AIX node was specified as an NSD server for the disk. The disk will be used as a non-PR disk.

**User response:** None. Informational message only.

---

**6027-1937** A node was specified more than once as an NSD server in disk descriptor *descriptor*.

**Explanation:** A node was specified more than once as an NSD server in the disk descriptor shown.

**User response:** Change the disk descriptor to eliminate any redundancies in the list of NSD servers.

---



---

**6027-1938** *configParameter* is an incorrect parameter. Line in error: *configLine*. The line is ignored; processing continues.

**Explanation:** The specified parameter is not valid and will be ignored.

**User response:** None. Informational message only.

---

**6027-1939** Line in error: *line*.

**Explanation:** The specified line from a user-provided input file contains errors.

**User response:** Check the preceding messages for more information. Correct the problems and reissue the command.

---

**6027-1940** Unable to set reserve policy *policy* on disk *diskName* on node *nodeName*.

**Explanation:** The specified disk should be able to support Persistent Reserve, but an attempt to set up the registration key failed.

**User response:** Correct the problem and reissue the command.

---

**6027-1941** Cannot handle multiple interfaces for host *hostName*.

**Explanation:** Multiple entries were found for the given hostname or IP address either in **/etc/hosts** or by the **host** command.

**User response:** Make corrections to **/etc/hosts** and reissue the command.

---

**6027-1942** Unexpected output from the 'host -t a *name*' command:

**Explanation:** A GPFS administration command (**mm...**) received unexpected output from the **host -t a** command for the given host.

**User response:** Issue the **host -t a** command interactively and carefully review the output, as well as any error messages.

---

**6027-1943** Host *name* not found.

**Explanation:** A GPFS administration command (**mm...**) could not resolve a host from **/etc/hosts** or by using the **host** command.

**User response:** Make corrections to **/etc/hosts** and reissue the command.

---

---

**6027-1945** Disk name *diskName* is not allowed. Names beginning with **gpfs** are reserved for use by GPFS.

**Explanation:** The cited disk name is not allowed because it begins with **gpfs**.

**User response:** Specify a disk name that does not begin with **gpfs** and reissue the command.

---

**6027-1947** Use **mmauth genkey** to recover the file *fileName*, or to generate and commit a new key.

**Explanation:** The specified file was not found.

**User response:** Recover the file, or generate a new key by running: **mmauth genkey propagate** or generate a new key by running **mmauth genkey new**, followed by the **mmauth genkey commit** command.

---

**6027-1948** Disk *diskName* is too large.

**Explanation:** The specified disk is too large.

**User response:** Specify a smaller disk and reissue the command.

---

**6027-1949** Propagating the cluster configuration data to all affected nodes.

**Explanation:** The cluster configuration data is being sent to the rest of the nodes in the cluster.

**User response:** This is an informational message.

---

**6027-1950** Local update lock is busy.

**Explanation:** More than one process is attempting to update the GPFS environment at the same time.

**User response:** Repeat the command. If the problem persists, verify that there are no blocked processes.

---

**6027-1951** Failed to obtain the local environment update lock.

**Explanation:** GPFS was unable to obtain the local environment update lock for more than 30 seconds.

**User response:** Examine previous error messages, if any. Correct any problems and reissue the command. If the problem persists, perform problem determination and contact the IBM Support Center.

---

**6027-1962** Permission denied for disk *diskName*

**Explanation:** The user does not have permission to access disk *diskName*.

**User response:** Correct the permissions and reissue the command.

---



---

**6027-1963** Disk *diskName* was not found.

**Explanation:** The specified disk was not found.

**User response:** Specify an existing disk and reissue the command.

---

**6027-1964** I/O error on *diskName*

**Explanation:** An I/O error occurred on the specified disk.

**User response:** Check for additional error messages. Check the error log for disk hardware problems.

---

**6027-1967** Disk *diskName* belongs to back-level file system *fileSystem* or the state of the disk is not ready. Use **mmchfs -V** to convert the file system to the latest format. Use **mmchdisk** to change the state of a disk.

**Explanation:** The specified disk cannot be initialized for use as a tiebreaker disk. Possible reasons are suggested in the message text.

**User response:** Use the **mmlsfs** and **mmlsdisk** commands to determine what action is needed to correct the problem.

---

**6027-1968** Failed while processing disk *diskName*.

**Explanation:** An error was detected while processing the specified disk.

**User response:** Examine prior messages to determine the reason for the failure. Correct the problem and reissue the command.

---

**6027-1969** Device *device* already exists on node *nodeName*

**Explanation:** This device already exists on the specified node.

**User response:** None.

---

**6027-1970** Disk *diskName* has no space for the quorum data structures. Specify a different disk as tiebreaker disk.

**Explanation:** There is not enough free space in the file system descriptor for the tiebreaker disk data structures.

**User response:** Specify a different disk as a tiebreaker disk.

---

**6027-1974** None of the quorum nodes can be reached.

**Explanation:** Ensure that the quorum nodes in the cluster can be reached. At least one of these nodes is required for the command to succeed.

**User response:** Ensure that the quorum nodes are available and reissue the command.

---

**6027-1975**    **The descriptor file contains more than one descriptor.**

**Explanation:** The descriptor file must contain only one descriptor.

**User response:** Correct the descriptor file.

---

**6027-1976**    **The descriptor file contains no descriptor.**

**Explanation:** The descriptor file must contain only one descriptor.

**User response:** Correct the descriptor file.

---

**6027-1977**    **Failed validating disk *diskName*. Error code *errorCode*.**

**Explanation:** GPFS control structures are not as expected.

**User response:** Contact the IBM Support Center.

---

**6027-1984**    **Name *name* is not allowed. It is longer than the maximum allowable length (*length*).**

**Explanation:** The cited name is not allowed because it is longer than the cited maximum allowable length.

**User response:** Specify a name whose length does not exceed the maximum allowable length, and reissue the command.

---

**6027-1985**    **mmfskxload: The format of the GPFS kernel extension is not correct for this version of AIX.**

**Explanation:** This version of AIX is incompatible with the current format of the GPFS kernel extension.

**User response:** Contact your system administrator to check the AIX version and GPFS kernel extension.

---

**6027-1986**    ***junctionName* does not resolve to a directory in *deviceName*. The junction must be within the specified file system.**

**Explanation:** The cited junction path name does not belong to the specified file system.

**User response:** Correct the junction path name and reissue the command.

---

**6027-1987**    **Name *name* is not allowed.**

**Explanation:** The cited name is not allowed because it is a reserved word or a prohibited character.

**User response:** Specify a different name and reissue the command.

---

**6027-1988**    **File system *fileSystem* is not mounted.**

**Explanation:** The cited file system is not currently mounted on this node.

**User response:** Ensure that the file system is mounted and reissue the command.

---

**6027-1993**    **File *fileName* either does not exist or has an incorrect format.**

**Explanation:** The specified file does not exist or has an incorrect format.

**User response:** Check whether the input file specified actually exists.

---

**6027-1994**    **Did not find any match with the input disk address.**

**Explanation:** The `mmfileid` command returned without finding any disk addresses that match the given input.

**User response:** None. Informational message only.

---

**6027-1995**    **Device *deviceName* is not mounted on node *nodeName*.**

**Explanation:** The specified device is not mounted on the specified node.

**User response:** Mount the specified device on the specified node and reissue the command.

---

**6027-1996**    **Command was unable to determine whether file system *fileSystem* is mounted.**

**Explanation:** The command was unable to determine whether the cited file system is mounted.

**User response:** Examine any prior error messages to determine why the command could not determine whether the file system was mounted, resolve the problem if possible, and then reissue the command. If you cannot resolve the problem, reissue the command with the daemon down on all nodes of the cluster. This will ensure that the file system is not mounted, which may allow the command to proceed.

---

**6027-1998** Line *lineNumber* of file *fileName* is incorrect:

**Explanation:** A line in the specified file passed to the command had incorrect syntax. The line with the incorrect syntax is displayed next, followed by a description of the correct syntax for the line.

**User response:** Correct the syntax of the line and reissue the command.

---

**6027-1999** Syntax error. The correct syntax is: *string*.

**Explanation:** The specified input passed to the command has incorrect syntax.

**User response:** Correct the syntax and reissue the command.

---

**6027-2000** Could not clear fencing for disk *physicalDiskName*.

**Explanation:** The fencing information on the disk could not be cleared.

**User response:** Make sure the disk is accessible by this node and retry.

---

**6027-2002** Disk *physicalDiskName* of type *diskType* is not supported for fencing.

**Explanation:** This disk is not a type that supports fencing.

**User response:** None.

---

**6027-2004** None of the specified nodes belong to this GPFS cluster.

**Explanation:** The nodes specified do not belong to the GPFS cluster.

**User response:** Choose nodes that belong to the cluster and try the command again.

---

**6027-2007** Unable to display fencing for disk *physicalDiskName*.

**Explanation:** Cannot retrieve fencing information for this disk.

**User response:** Make sure that this node has access to the disk before retrying.

---

**6027-2008** For the logical volume specification **-l *lvName* to be valid *lvName* must be the only logical volume in the volume group. However, volume group *vgName* contains logical volumes.**

**Explanation:** The command is being run on a logical

volume that belongs to a volume group that has more than one logical volume.

**User response:** Run this command only on a logical volume where it is the only logical volume in the corresponding volume group.

---

**6027-2009** *logicalVolume* is not a valid logical volume.

**Explanation:** *logicalVolume* does not exist in the ODM, implying that logical name does not exist.

**User response:** Run the command on a valid logical volume.

---

**6027-2010** *vgName* is not a valid volume group name.

**Explanation:** *vgName* passed to the command is not found in the ODM, implying that *vgName* does not exist.

**User response:** Run the command on a valid volume group name.

---

**6027-2011** For the hdisk specification **-h *physicalDiskName* to be valid *physicalDiskName* must be the only disk in the volume group. However, volume group *vgName* contains disks.**

**Explanation:** The hdisk specified belongs to a volume group that contains other disks.

**User response:** Pass an hdisk that belongs to a volume group that contains only this disk.

---

**6027-2012** *physicalDiskName* is not a valid physical volume name.

**Explanation:** The specified name is not a valid physical disk name.

**User response:** Choose a correct physical disk name and retry the command.

---

**6027-2013** *pvid* is not a valid physical volume id.

**Explanation:** The specified value is not a valid physical volume ID.

**User response:** Choose a correct physical volume ID and retry the command.

---

**6027-2014** Node *node* does not have access to disk *physicalDiskName*.

**Explanation:** The specified node is not able to access the specified disk.

**User response:** Choose a different node or disk (or both), and retry the command. If both the node and

disk name are correct, make sure that the node has access to the disk.

---

**6027-2015** Node *node* does not hold a reservation for disk *physicalDiskName*.

**Explanation:** The node on which this command is run does not have access to the disk.

**User response:** Run this command from another node that has access to the disk.

---

**6027-2016** SSA fencing support is not present on this node.

**Explanation:** This node does not support SSA fencing.

**User response:** None.

---

**6027-2017** Node ID *nodeId* is not a valid SSA node ID. SSA node IDs must be a number in the range of 1 to 128.

**Explanation:** You specified a node ID outside of the acceptable range.

**User response:** Choose a correct node ID and retry the command.

---

**6027-2018** The SSA node id is not set.

**Explanation:** The SSA node ID has not been set.

**User response:** Set the SSA node ID.

---

**6027-2019** Unable to retrieve the SSA node id.

**Explanation:** A failure occurred while trying to retrieve the SSA node ID.

**User response:** None.

---

**6027-2020** Unable to set fencing for disk *physicalDiskName*.

**Explanation:** A failure occurred while trying to set fencing for the specified disk.

**User response:** None.

---

**6027-2021** Unable to clear PR reservations for disk *physicalDiskName*.

**Explanation:** Failed to clear Persistent Reserve information on the disk.

**User response:** Make sure the disk is accessible by this node before retrying.

---

**6027-2022** Could not open disk *physicalDiskName*, **errno** *value*.

**Explanation:** The specified disk cannot be opened.

**User response:** Examine the **errno** value and other messages to determine the reason for the failure. Correct the problem and reissue the command.

---

**6027-2023** **retVal = value, errno = value for key value.**

**Explanation:** An **ioctl** call failed with stated return code, **errno** value, and related values.

**User response:** Check the reported **errno** and correct the problem if possible. Otherwise, contact the IBM Support Center.

---

**6027-2024** **ioctl failed with rc=returnCode, errno=errnoValue. Related values are scsi\_status=scsiStatusValue, sense\_key=senseKeyValue, scsi\_asc=scsiAscValue, scsi\_ascq=scsiAscqValue.**

**Explanation:** An **ioctl** call failed with stated return code, **errno** value, and related values.

**User response:** Check the reported **errno** and correct the problem if possible. Otherwise, contact the IBM Support Center.

---

**6027-2025** **READ\_KEYS ioctl failed with errno=returnCode, tried timesTried times. Related values are scsi\_status=scsiStatusValue, sense\_key=senseKeyValue, scsi\_asc=scsiAscValue, scsi\_ascq=scsiAscqValue.**

**Explanation:** A **READ\_KEYS ioctl** call failed with stated **errno** value, and related values.

**User response:** Check the reported **errno** and correct the problem if possible. Otherwise, contact the IBM Support Center.

---

**6027-2026** **READRES ioctl failed with errno=returnCode, tried timesTried times. Related values are: scsi\_status=scsiStatusValue, sense\_key=senseKeyValue, scsi\_asc=scsiAscValue, scsi\_ascq=scsiAscqValue.**

**Explanation:** A **REGISTER ioctl** call failed with stated **errno** value, and related values.

**User response:** Check the reported **errno** and correct the problem if possible. Otherwise, contact the IBM Support Center.



---

**6027-2027**    **READRES ioctl failed with errno=returnCode, tried timesTried times.**  
**Related values are:**  
*scsi\_status=scsiStatusValue,*  
*sense\_key=senseKeyValue,*  
*scsi\_asc=scsiAscValue,*  
*scsi\_ascq=scsiAscqValue.*

**Explanation:** A READRES ioctl call failed with stated **errno** value, and related values.

**User response:** Check the reported **errno** and correct the problem if possible. Otherwise, contact the IBM Support Center.

---

**6027-2028**    **could not open disk device**  
*diskDeviceName*

**Explanation:** A problem occurred on a disk open.

**User response:** Ensure the disk is accessible and not fenced out, and then reissue the command.

---

**6027-2029**    **could not close disk device**  
*diskDeviceName*

**Explanation:** A problem occurred on a disk close.

**User response:** None.

---

**6027-2030**    **ioctl failed with DSB=value and result=value reason: explanation**

**Explanation:** An ioctl call failed with stated return code, *errno* value, and related values.

**User response:** Check the reported *errno* and correct the problem, if possible. Otherwise, contact the IBM Support Center.

---

**6027-2031**    **ioctl failed with non-zero return code**

**Explanation:** An ioctl failed with a non-zero return code.

**User response:** Correct the problem, if possible. Otherwise, contact the IBM Support Center.

---

**6027-2049 [X]** **Cannot pin a page pool of size *value* bytes.**

**Explanation:** A GPFS page pool cannot be pinned into memory on this machine.

**User response:** Increase the physical memory size of the machine.

---

**6027-2050 [E]** **Pagepool has size *actualValue* bytes instead of the requested *requestedValue* bytes.**

**Explanation:** The configured GPFS page pool is too large to be allocated or pinned into memory on this

machine. GPFS will work properly, but with reduced capacity for caching user data.

**User response:** To prevent this message from being generated when the GPFS daemon starts, reduce the page pool size using the **mmchconfig** command.

---

**6027-2100**    **Incorrect range *value-value* specified.**

**Explanation:** The range specified to the command is incorrect. The first parameter value must be less than or equal to the second parameter value.

**User response:** Correct the address range and reissue the command.

---

**6027-2101**    **Insufficient free space in *fileSystem* (*storage* minimum required).**

**Explanation:** There is not enough free space in the specified file system or directory for the command to successfully complete.

**User response:** Correct the problem and reissue the command.

---

**6027-2102**    **Node *nodeName* is not mmremotefs to run the command.**

**Explanation:** The specified node is not available to run a command. Depending on the command, a different node may be tried.

**User response:** Determine why the specified node is not available and correct the problem.

---

**6027-2103**    **Directory *dirName* does not exist**

**Explanation:** The specified directory does not exist.

**User response:** Reissue the command specifying an existing directory.

---

**6027-2104**    **The GPFS release level could not be determined on nodes: *nodeList*.**

**Explanation:** The command was not able to determine the level of the installed GPFS code on the specified nodes.

**User response:** Reissue the command after correcting the problem.

---

**6027-2105**    **The following nodes must be upgraded to GPFS release *productVersion* or higher: *nodeList***

**Explanation:** The command requires that all nodes be at the specified GPFS release level.

**User response:** Correct the problem and reissue the command.

---

---

**6027-2106**    **Ensure the nodes are available and run:**  
*command.*

**Explanation:** The command could not complete normally.

**User response:** Check the preceding messages, correct the problems, and issue the specified command until it completes successfully.

---

**6027-2107**    **Upgrade the lower release level nodes and run:** *command.*

**Explanation:** The command could not complete normally.

**User response:** Check the preceding messages, correct the problems, and issue the specified command until it completes successfully.

---

**6027-2108**    **Error found while processing stanza**

**Explanation:** A stanza was found to be unsatisfactory in some way.

**User response:** Check the preceding messages, if any, and correct the condition that caused the stanza to be rejected.

---

**6027-2109**    **Failed while processing disk stanza on node** *nodeName.*

**Explanation:** A disk stanza was found to be unsatisfactory in some way.

**User response:** Check the preceding messages, if any, and correct the condition that caused the stanza to be rejected.

---

**6027-2110**    **Missing required parameter** *parameter*

**Explanation:** The specified parameter is required for this command.

**User response:** Specify the missing information and reissue the command.

---

**6027-2111**    **The following disks were not deleted:**  
*diskList*

**Explanation:** The command could not delete the specified disks. Check the preceding messages for error information.

**User response:** Correct the problems and reissue the command.

---

**6027-2112**    **Permission failure. Option** *option*  
**requires root authority to run.**

**Explanation:** The specified command option requires root authority.

**User response:** Log on as root and reissue the command.

---

**6027-2113**    **Not able to associate** *diskName* **on node**  
*nodeName* **with any known GPFS disk.**

**Explanation:** A command could not find a GPFS disk that matched the specified disk and node values passed as input.

**User response:** Correct the disk and node values passed as input and reissue the command.

---

**6027-2114**    **The** *subsystem* **subsystem is already active.**

**Explanation:** The user attempted to start a subsystem that was already active.

**User response:** None. Informational message only.

---

**6027-2115**    **Unable to resolve address range for disk**  
*diskName* **on node** *nodeName.*

**Explanation:** A command could not perform address range resolution for the specified disk and node values passed as input.

**User response:** Correct the disk and node values passed as input and reissue the command.

---

**6027-2116 [E]** **The GPFS daemon must be active on the recovery group server nodes.**

**Explanation:** The command requires that the GPFS daemon be active on the recovery group server nodes.

**User response:** Ensure GPFS is running on the recovery group server nodes and reissue the command.

---

**6027-2117 [E]** *object name* **already exists.**

**Explanation:** The user attempted to create an object with a name that already exists.

**User response:** Correct the name and reissue the command.

---

**6027-2118 [E]** **The** *parameter* **is invalid or missing in the pdisk descriptor.**

**Explanation:** The pdisk descriptor is not valid. The bad descriptor is displayed following this message.

**User response:** Correct the input and reissue the command.

---

**6027-2119 [E]** **Recovery group** *name* **not found.**

**Explanation:** The specified recovery group was not found.

**User response:** Correct the input and reissue the command.

---

**6027-2120 [E] Unable to delete recovery group *name* on nodes *nodeNames*.**

**Explanation:** The recovery group could not be deleted on the specified nodes.

**User response:** Perform problem determination.

---

**6027-2121 [I] Recovery group *name* deleted on node *nodeName*.**

**Explanation:** The recovery group has been deleted.

**User response:** This is an informational message.

---

**6027-2122 [E] The number of spares (*numberOfSpares*) must be less than the number of pdisks (*numberOfPdisks*) being created.**

**Explanation:** The number of spares specified must be less than the number of pdisks that are being created.

**User response:** Correct the input and reissue the command.

---

**6027-2123 [E] The GPFS daemon is down on the *vdiskName* servers.**

**Explanation:** The GPFS daemon was down on the vdisk servers when **mmdelvdisk** was issued.

**User response:** Start the GPFS daemon on the specified nodes and issue the specified **mmdelvdisk** command.

---

**6027-2124 [E] Vdisk *vdiskName* is still NSD *nsdName*. Use the **mmdelnsd** command.**

**Explanation:** The specified vdisk is still an NSD.

**User response:** Use the **mmdelnsd** command.

---

**6027-2125 [E] *nsdName* is a vdisk-based NSD and cannot be used as a tiebreaker disk.**

**Explanation:** Vdisk-based NSDs cannot be specified as tiebreaker disks.

**User response:** Correct the input and reissue the command.

---

**6027-2126 [I] No recovery groups were found.**

**Explanation:** A command searched for recovery groups but found none.

**User response:** None. Informational message only.

---

**6027-2127 [E] Disk descriptor *descriptor* refers to an existing pdisk.**

**Explanation:** The specified disk descriptor refers to an existing pdisk.

---

**User response:** Specify another disk that is not an existing pdisk.

---

**6027-2128 [E] The *attribute* attribute must be configured to use *hostname* as a recovery group server.**

**Explanation:** The specified GPFS configuration attributes must be configured to use the node as a recovery group server.

**User response:** Use the **mmchconfig** command to set the attributes, then reissue the command.

---

**6027-2129 [E] Vdisk block size (*blockSize*) must match the file system block size (*blockSize*).**

**Explanation:** The specified NSD is a vdisk with a block size that does not match the block size of the file system.

**User response:** Reissue the command using block sizes that match.

---

**6027-2130 [E] Could not find an active server for recovery group *name*.**

**Explanation:** A command was issued that acts on a recovery group, but no active server was found for the specified recovery group.

**User response:** Perform problem determination.

---

**6027-2131 [E] Cannot create an NSD on a log vdisk.**

**Explanation:** The specified disk is a log vdisk; it cannot be used for an NSD.

**User response:** Specify another disk that is not a log vdisk.

---

**6027-2132 [E] Log vdisk *vdiskName* cannot be deleted while there are other vdisks in recovery group *name*.**

**Explanation:** The specified disk is a log vdisk; it must be the last vdisk deleted from the recovery group.

**User response:** Delete the other vdisks first.

---

**6027-2133 [E] Unable to delete recovery group *name*; vdisks are still defined.**

**Explanation:** Cannot delete a recovery group while there are still vdisks defined.

**User response:** Delete all the vdisks first.

---

---

**6027-2134** Node *nodeName* cannot be used as an NSD server for Persistent Reserve disk *diskName* because it is not a Linux node.

**Explanation:** There was an attempt to enable Persistent Reserve for a disk, but not all of the NSD server nodes are running Linux.

**User response:** Correct the configuration and enter the command again.

---

**6027-2135** All nodes in the cluster must be running AIX to enable Persistent Reserve for SAN attached disk *diskName*.

**Explanation:** There was an attempt to enable Persistent Reserve for a SAN-attached disk, but not all nodes in the cluster are running AIX.

**User response:** Correct the configuration and run the command again.

---

**6027-2136** All NSD server nodes must be running AIX to enable Persistent Reserve for disk *diskName*.

**Explanation:** There was an attempt to enable Persistent Reserve for the specified disk, but not all NSD servers are running AIX.

**User response:** Correct the configuration and enter the command again.

---

**6027-2137** An attempt to clear the Persistent Reserve reservations on disk *diskName* failed.

**Explanation:** You are importing a disk into a cluster in which Persistent Reserve is disabled. An attempt to clear the Persistent Reserve reservations on the disk failed.

**User response:** Correct the configuration and enter the command again.

---

**6027-2138** The cluster must be running either all AIX or all Linux nodes to change Persistent Reserve disk *diskName* to a SAN-attached disk.

**Explanation:** There was an attempt to redefine a Persistent Reserve disk as a SAN attached disk, but not all nodes in the cluster were running either all AIX or all Linux nodes.

**User response:** Correct the configuration and enter the command again.

---

**6027-2139** NSD server nodes must be running either all AIX or all Linux to enable Persistent Reserve for disk *diskName*.

**Explanation:** There was an attempt to enable Persistent Reserve for a disk, but not all NSD server nodes were running all AIX or all Linux nodes.

**User response:** Correct the configuration and enter the command again.

---

**6027-2140** All NSD server nodes must be running AIX or all running Linux to enable Persistent Reserve for disk *diskName*.

**Explanation:** Attempt to enable Persistent Reserve for a disk while not all NSD server nodes are running AIX or all running Linux.

**User response:** Correct the configuration first.

---

**6027-2141** Disk *diskName* is not configured as a regular hdisk.

**Explanation:** In an AIX only cluster, Persistent Reserve is supported for regular hdisks only.

**User response:** Correct the configuration and enter the command again.

---

**6027-2142** Disk *diskName* is not configured as a regular generic disk.

**Explanation:** In a Linux only cluster, Persistent Reserve is supported for regular generic or device mapper virtual disks only.

**User response:** Correct the configuration and enter the command again.

---

**6027-2143** Mount point *mountPoint* can not be part of automount directory *automountDir*.

**Explanation:** The mount point cannot be the parent directory of the automount directory.

**User response:** Specify a mount point that is not the parent of the automount directory.

---

**6027-2144 [E]** The *lockName* lock for file system *fileSystem* is busy.

**Explanation:** More than one process is attempting to obtain the specified lock.

**User response:** Repeat the command. If the problem persists, verify that there are no blocked processes.

---

---

**6027-2145 [E] Internal remote command 'mmremote command' no longer supported.**

**Explanation:** A GPFS administration command invoked an internal remote command which is no longer supported. Backward compatibility for remote commands are only supported for release 3.4 and newer.

**User response:** All nodes within the cluster must be at release 3.4 or newer. If all the cluster nodes meet this requirement, contact the IBM Support Center.

---

**6027-2147 [E] BlockSize must be specified in disk descriptor.**

**Explanation:** The blockSize positional parameter in a vdisk descriptor was empty. The bad disk descriptor is displayed following this message.

**User response:** Correct the input and reissue the command.

---

**6027-2148 [E] nodeName is not a valid recovery group server for recoveryGroupName.**

**Explanation:** The server name specified is not one of the defined recovery group servers.

**User response:** Correct the input and reissue the command.

---

**6027-2149 [E] Could not get recovery group information from an active server.**

**Explanation:** A command that needed recovery group information failed; the GPFS daemons may have become inactive or the recovery group is temporarily unavailable.

**User response:** Reissue the command.

---

**6027-2150 The archive system client backupProgram could not be found or is not executable.**

**Explanation:** TSM dsmc or other specified backup or archive system client could not be found.

**User response:** Verify that TSM is installed, dsmc can be found in the installation location or that the archiver client specified is executable.

---

**6027-2151 The path directoryPath is not contained in the snapshot snapshotName.**

**Explanation:** The directory path supplied is not contained in the snapshot named with the -S parameter.

**User response:** Correct the directory path or snapshot name supplied, or omit -S and the snapshot name in the command.

---

**6027-2152 The path directoryPath containing image archives was not found.**

**Explanation:** The directory path supplied does not contain the expected image files to archive into TSM.

**User response:** Correct the directory path name supplied.

---

**6027-2153 The archiving system backupProgram exited with status return code. Image backup files have been preserved in globalWorkDir**

**Explanation:** Archiving system executed and returned a non-zero exit status due to some error.

**User response:** Examine archiver log files to discern the cause of the archiver's failure. Archive the preserved image files from the indicated path.

---

**6027-2154 Unable to create a policy file for image backup in policyFilePath.**

**Explanation:** A temporary file could not be created in the global shared directory path.

**User response:** Check or correct the directory path name supplied.

---

**6027-2155 File system fileSystem must be mounted read only for restore.**

**Explanation:** The empty file system targeted for restoration must be mounted in read only mode during restoration.

**User response:** Unmount the file system on all nodes and remount it read only, then try the command again.

---

**6027-2156 The image archive index ImagePath could not be found.**

**Explanation:** The archive image index could be found in the specified path

**User response:** Check command arguments for correct specification of image path, then try the command again.

---

**6027-2157 The image archive index ImagePath is corrupt or incomplete.**

**Explanation:** The archive image index specified is damaged.

**User response:** Check the archive image index file for corruption and remedy.

---

---

**6027-2158** Disk usage must be `dataOnly`, `metadataOnly`, `descOnly`, `dataAndMetadata`, `vdiskLog`, `vdiskLogTip`, `vdiskLogTipBackup`, or `vdiskLogReserved`.

**Explanation:** The disk usage positional parameter in a vdisk descriptor has a value that is not valid. The bad disk descriptor is displayed following this message.

**User response:** Correct the input and reissue the command.

---

**6027-2159 [E]** *parameter is not valid or missing in the vdisk descriptor.*

**Explanation:** The vdisk descriptor is not valid. The bad descriptor is displayed following this message.

**User response:** Correct the input and reissue the command.

---

**6027-2160 [E]** Vdisk *vdiskName* is already mapped to NSD *nsdName*.

**Explanation:** The command cannot create the specified NSD because the underlying vdisk is already mapped to a different NSD.

**User response:** Correct the input and reissue the command.

---

**6027-2161 [E]** NAS servers cannot be specified when creating an NSD on a vdisk.

**Explanation:** The command cannot create the specified NSD because servers were specified and the underlying disk is a vdisk.

**User response:** Correct the input and reissue the command.

---

**6027-2162 [E]** Cannot set `nsdRAIDTracks` to zero; *nodeName* is a recovery group server.

**Explanation:** `nsdRAIDTracks` cannot be set to zero while the node is still a recovery group server.

**User response:** Modify or delete the recovery group and reissue the command.

---

**6027-2163 [E]** Vdisk *name* not found in the daemon. Recovery may be occurring. The disk will not be deleted.

**Explanation:** GPFS cannot find the specified vdisk. This can happen if recovery is taking place and the recovery group is temporarily inactive.

**User response:** Reissue the command. If the recovery group is damaged, specify the `-p` option.

---

**6027-2164 [E]** Disk descriptor for *name* refers to an existing pdisk.

**Explanation:** The specified pdisk already exists.

**User response:** Correct the command invocation and try again.

---

**6027-2165 [E]** Node *nodeName* cannot be used as a server of both vdisks and non-vdisk NSDs.

**Explanation:** The command specified an action that would have caused vdisks and non-vdisk NSDs to be defined on the same server. This is not a supported configuration.

**User response:** Correct the command invocation and try again.

---

**6027-2166 [E]** IBM Spectrum Scale RAID is not configured.

**Explanation:** IBM Spectrum Scale RAID is not configured on this node.

**User response:** Reissue the command on the appropriate node.

---

**6027-2167 [E]** Device *deviceName* does not exist or is not active on this node.

**Explanation:** The specified device does not exist or is not active on the node.

**User response:** Reissue the command on the appropriate node.

---

**6027-2168 [E]** The GPFS cluster must be shut down before downloading firmware to port cards.

**Explanation:** The GPFS daemon must be down on all nodes in the cluster before attempting to download firmware to a port card.

**User response:** Stop GPFS on all nodes and reissue the command.

---

**6027-2169** Unable to disable Persistent Reserve on the following disks: *diskList*

**Explanation:** The command was unable to disable Persistent Reserve on the specified disks.

**User response:** Examine the disks and additional error information to determine if the disks should support Persistent Reserve. Correct the problem and reissue the command.

---

---

**6027-2170 [E] Recovery group *recoveryGroupName* does not exist or is not active.**

**Explanation:** A command was issued to a recovery group that does not exist or is not in the active state.

**User response:** Reissue the command with a valid recovery group name or wait for the recovery group to become active.

---

**6027-2171 [E] *objectType objectName* already exists in the cluster.**

**Explanation:** The file system being imported contains an object with a name that conflicts with the name of an existing object in the cluster.

**User response:** If possible, remove the object with the conflicting name.

---

**6027-2172 [E] Errors encountered while importing IBM Spectrum Scale RAID objects.**

**Explanation:** Errors were encountered while trying to import a IBM Spectrum Scale RAID based file system. No file systems will be imported.

**User response:** Check the previous error messages and if possible, correct the problems.

---

**6027-2173 [I] Use *mmchrecoverygroup* to assign and activate servers for the following recovery groups (automatically assigns NSD servers as well): *recoveryGroupList***

**Explanation:** The *mmimportfs* command imported the specified recovery groups. These must have servers assigned and activated.

**User response:** After the *mmimportfs* command finishes, use the *mmchrecoverygroup* command to assign NSD server nodes as needed.

---

**6027-2174 Option *option* can be specified only in conjunction with *option*.**

**Explanation:** The cited option cannot be specified by itself.

**User response:** Correct the input and reissue the command.

---

**6027-2175 [E] Exported path *exportPath* does not exist**

**Explanation:** The directory or one of the components in the directory path to be exported does not exist.

**User response:** Correct the input and reissue the command.

---



---

**6027-2176 [E] *mmchattr* for *fileName* failed.**

**Explanation:** The command to change the attributes of the file failed.

**User response:** Check the previous error messages and correct the problems.

---

**6027-2177 [E] Cannot create file *fileName*.**

**Explanation:** The command to create the specified file failed.

**User response:** Check the previous error messages and correct the problems.

---

**6027-2178 File *fileName* does not contain any NSD descriptors or stanzas.**

**Explanation:** The input file should contain at least one NSD descriptor or stanza.

**User response:** Correct the input file and reissue the command.

---

**6027-2181 [E] Failover is allowed only for single-writer, independent-writer filesets.**

**Explanation:** The fileset AFM mode is not compatible with the requested operation.

**User response:** Check the previous error messages and correct the problems.

---

**6027-2182 [E] Resync is allowed only for single-writer filesets.**

**Explanation:** The fileset AFM mode is not compatible with the requested operation.

**User response:** Check the previous error messages and correct the problems.

---

**6027-2183 [E] Peer snapshots using *mmpsnap* are allowed only for single-writer or primary filesets.**

**Explanation:** The fileset AFM mode is not compatible with the requested operation.

**User response:** Check the previous error messages and correct the problems.

---

**6027-2184 [E] If the recovery group is damaged, issue *mmdelrecoverygroup name -p*.**

**Explanation:** No active servers were found for the recovery group that is being deleted. If the recovery group is damaged the *-p* option is needed.

**User response:** Perform diagnosis and reissue the command.

---

---

**6027-2185 [E]** There are no `pdisk` stanzas in the input file *fileName*.

**Explanation:** The `mmcrrecoverygroup` input stanza file has no `pdisk` stanzas.

**User response:** Correct the input file and reissue the command.

---

**6027-2186 [E]** There were no valid `vdisk` stanzas in the input file *fileName*.

**Explanation:** The `mmcrvdisk` input stanza file has no valid `vdisk` stanzas.

**User response:** Correct the input file and reissue the command.

---

**6027-2187 [E]** Could not get `pdisk` information for the following recovery groups:  
*recoveryGroupList*

**Explanation:** An `mmlspdisk all` command could not query all of the recovery groups because some nodes could not be reached.

**User response:** None.

---

**6027-2188** Unable to determine the local node identity.

**Explanation:** The command is not able to determine the identity of the local node. This can be the result of a disruption in the network over which the GPFS daemons communicate.

**User response:** Ensure the GPFS daemon network (as identified in the output of the `mmlscluster` command on a good node) is fully operational and reissue the command.

---

**6027-2189 [E]** Action *action* is allowed only for read-only filesets.

**Explanation:** The specified action is only allowed for read-only filesets.

**User response:** None.

---

**6027-2190 [E]** Cannot prefetch file *fileName*. The file does not belong to fileset *fileset*.

**Explanation:** The requested file does not belong to the fileset.

**User response:** None.

---

**6027-2191 [E]** `Vdisk` *vdiskName* not found in recovery group *recoveryGroupName*.

**Explanation:** The `mmdelvdisk` command was invoked with the `--recovery-group` option to delete one or more `vdisks` from a specific recovery group. The specified

`vdisk` does not exist in this recovery group.

**User response:** Correct the input and reissue the command.

---

**6027-2193 [E]** Recovery group *recoveryGroupName* must be active on the primary server *serverName*.

**Explanation:** The recovery group must be active on the specified node.

**User response:** Use the `mmchrecoverygroup` command to activate the group and reissue the command.

---

**6027-2194 [E]** The state of fileset *filesetName* is Expired; prefetch cannot be performed.

**Explanation:** The prefetch operation cannot be performed on filesets that are in the Expired state.

**User response:** None.

---

**6027-2195 [E]** Error getting snapshot ID for *snapshotName*.

**Explanation:** The command was unable to obtain the resync snapshot ID.

**User response:** Examine the preceding messages, correct the problem, and reissue the command. If the problem persists, perform problem determination and contact the IBM Support Center.

---

**6027-2196 [E]** Resync is allowed only when the fileset queue is in active state.

**Explanation:** This operation is allowed only when the fileset queue is in active state.

**User response:** None.

---

**6027-2197 [E]** Empty file encountered when running the `mmafmctl flushPending` command.

**Explanation:** The `mmafmctl flushPending` command did not find any entries in the file specified with the `--list-file` option.

**User response:** Correct the input file and reissue the command.

---

**6027-2198 [E]** Cannot run the `mmafmctl flushPending` command on directory *dirName*.

**Explanation:** The `mmafmctl flushPending` command cannot be issued on this directory.

**User response:** Correct the input and reissue the command.



---

**6027-2199 [E] No enclosures were found.**

**Explanation:** A command searched for disk enclosures but none were found.

**User response:** None.

---

**6027-2200 [E] Cannot have multiple nodes updating firmware for the same enclosure. Enclosure *serialNumber* is already being updated by node *nodeName*.**

**Explanation:** The `mmchenclosure` command was called with multiple nodes updating the same firmware.

**User response:** Correct the node list and reissue the command.

---

**6027-2201 [E] The `mmafmctl flushPending` command completed with errors.**

**Explanation:** An error occurred while flushing the queue.

**User response:** Examine the GPFS log to identify the cause.

---

**6027-2202 [E] There is a SCSI-3 PR reservation on disk *diskname*. `mmcrnsd` cannot format the disk because the cluster is not configured as PR enabled.**

**Explanation:** The specified disk has a SCSI-3 PR reservation, which prevents the `mmcrnsd` command from formatting it.

**User response:** Clear the PR reservation by following the instructions in "Clearing a leftover Persistent Reserve reservation" on page 288.

---

**6027-2203 Node *nodeName* is not a gateway node.**

**Explanation:** The specified node is not a gateway node.

**User response:** Designate the node as a gateway node or specify a different node on the command line.

---

**6027-2204 AFM target map *mapName* is already defined.**

**Explanation:** A request was made to create an AFM target map with the cited name, but that map name is already defined.

**User response:** Specify a different name for the new AFM target map or first delete the current map definition and then recreate it.

---

**6027-2205 There are no AFM target map definitions.**

**Explanation:** A command searched for AFM target map definitions but found none.

**User response:** None. Informational message only.

---

**6027-2206 AFM target map *mapName* is not defined.**

**Explanation:** The cited AFM target map name is not known to GPFS.

**User response:** Specify an AFM target map known to GPFS.

---

**6027-2207 Node *nodeName* is being used as a gateway node for the AFM cluster *clusterName*.**

**Explanation:** The specified node is defined as a gateway node for the specified AFM cluster.

**User response:** If you are trying to delete the node from the GPFS cluster or delete the gateway node role, you must remove it from the export server map.

---

**6027-2208 [E] *commandName* is already running in the cluster.**

**Explanation:** Only one instance of the specified command is allowed to run.

**User response:** None.

---

**6027-2209 [E] Unable to list *objectName* on node *nodeName*.**

**Explanation:** A command was unable to list the specific object that was requested.

**User response:** None.

---

**6027-2210 [E] Unable to build a storage enclosure inventory file on node *nodeName*.**

**Explanation:** A command was unable to build a storage enclosure inventory file. This is a temporary file that is required to complete the requested command.

**User response:** None.

---

**6027-2211 [E] Error collecting firmware information on node *nodeName*.**

**Explanation:** A command was unable to gather firmware information from the specified node.

**User response:** Ensure the node is active and retry the command.

---

---

**6027-2212 [E] Firmware update file *updateFile* was not found.**

**Explanation:** The `mmchfirmware` command could not find the specified firmware update file to load.

**User response:** Locate the firmware update file and retry the command.

---

**6027-2213 [E] Pdisk path redundancy was lost while updating enclosure firmware.**

**Explanation:** The `mmchfirmware` command lost paths after loading firmware and rebooting the Enclosure Services Module.

**User response:** Wait a few minutes and then retry the command. GPFS might need to be shut down to finish updating the enclosure firmware.

---

**6027-2214 [E] Timeout waiting for firmware to load.**

**Explanation:** A storage enclosure firmware update was in progress, but the update did not complete within the expected time frame.

**User response:** Wait a few minutes, and then use the `mmlsfirmware` command to ensure the operation completed.

---

**6027-2215 [E] Storage enclosure *serialNumber* not found.**

**Explanation:** The specified storage enclosure was not found.

**User response:** None.

---

**6027-2216 Quota management is disabled for file system *fileSystem*.**

**Explanation:** Quota management is disabled for the specified file system.

**User response:** Enable quota management for the file system.

---

**6027-2217 [E] Error *errno* updating firmware for drives *driveList*.**

**Explanation:** The firmware load failed for the specified drives. Some of the drives may have been updated.

**User response:** None.

---

**6027-2218 [E] Storage enclosure *serialNumber* component *componentType* component ID *componentId* not found.**

**Explanation:** The `mmchenclousure` command could not find the component specified for replacement.

**User response:** Use the `mmlsenclousure` command to

determine valid input and then retry the command.

---

**6027-2219 [E] Storage enclosure *serialNumber* component *componentType* component ID *componentId* did not fail. Service is not required.**

**Explanation:** The component specified for the `mmchenclousure` command does not need service.

**User response:** Use the `mmlsenclousure` command to determine valid input and then retry the command.

---

**6027-2220 [E] Recovery group *name* has pdisks with missing paths. Consider using the `-v no` option of the `mmchrecoverygroup` command.**

**Explanation:** The `mmchrecoverygroup` command failed because all the servers could not see all the disks, and the primary server is missing paths to disks.

**User response:** If the disks are cabled correctly, use the `-v no` option of the `mmchrecoverygroup` command.

---

**6027-2221 [E] Error determining redundancy of enclosure *serialNumber* **ESM** *esmName*.**

**Explanation:** The `mmchrecoverygroup` command failed. Check the following error messages.

**User response:** Correct the problem and retry the command.

---

**6027-2222 [E] Storage enclosure *serialNumber* already has a newer firmware version: *firmwareLevel*.**

**Explanation:** The `mmchfirmware` command found a newer level of firmware on the specified storage enclosure.

**User response:** If the intent is to force on the older firmware version, use the `-v no` option.

---

**6027-2223 [E] Storage enclosure *serialNumber* is not redundant. Shutdown GPFS in the cluster and retry the `mmchfirmware` command.**

**Explanation:** The `mmchfirmware` command found a non-redundant storage enclosure. Proceeding could cause loss of data access.

**User response:** Shutdown GPFS in the cluster and retry the `mmchfirmware` command.

---

**6027-2224 [E] Peer snapshot creation failed. Error code *errorCode*.**

**Explanation:** For an active fileset, check the AFM target configuration for peer snapshots. Ensure there is at least one gateway node configured for the cluster.

Examine the preceding messages and the GPFS log for additional details.

**User response:** Correct the problems and reissue the command.

**6027-2225 [E] Peer snapshot successfully deleted at cache. The delete snapshot operation failed at home. Error code *errorCode*.**

**Explanation:** For an active fileset, check the AFM target configuration for peer snapshots. Ensure there is at least one gateway node configured for the cluster. Examine the preceding messages and the GPFS log for additional details.

**User response:** Correct the problems and reissue the command.

**6027-2226 [E] Invalid firmware update file.**

**Explanation:** An invalid firmware update file was specified for the `mmchfirmware` command.

**User response:** Reissue the command with a valid update file.

**6027-2227 [E] Failback is allowed only for independent-writer filesets.**

**Explanation:** Failback operation is allowed only for independent-writer filesets.

**User response:** Check the fileset mode.

**6027-2228 [E] The daemon version (*daemonVersion*) on node *nodeName* is lower than the daemon version (*daemonVersion*) on node *nodeName*.**

**Explanation:** A command was issued that requires nodes to be at specific levels, but the affected GPFS servers are not at compatible levels to support this operation.

**User response:** Update the GPFS code on the specified servers and retry the command.

**6027-2229 [E] Cache Eviction/Prefetch is not allowed for Primary and Secondary mode filesets.**

**Explanation:** Cache eviction/prefetch is not allowed for primary and secondary mode filesets.

**User response:** None.

**6027-2230 [E] `afmTarget=newTargetString` is not allowed. To change the AFM target, use `mmafmctl failover` with the `--target-only` option. For primary filesets, use `mmafmctl changeSecondary`.**

**Explanation:** The `mmchfileset` command cannot be used to change the NFS server or IP address of the home cluster.

**User response:** To change the AFM target, use the `mmafmctl failover` command and specify the `--target-only` option. To change the AFM target for primary filesets, use the `mmafmctl changeSecondary` command.

**6027-2231 [E] The specified block size *blockSize* is smaller than the system page size *pageSize*.**

**Explanation:** The file system block size cannot be smaller than the system memory page size.

**User response:** Specify a block size greater than or equal to the system memory page size.

**6027-2232 [E] Peer snapshots are allowed only for targets using the NFS protocol.**

**Explanation:** The `mmpsnap` command can be used to create snapshots only for filesets that are configured to use the NFS protocol.

**User response:** Specify a valid fileset target.

**6027-2233 [E] Fileset *filesetName* in file system *filesystemName* does not contain peer snapshot *snapshotName*. The delete snapshot operation failed at cache. Error code *errorCode*.**

**Explanation:** The specified snapshot name was not found. The command expects the name of an existing peer snapshot of the active fileset in the specified file system.

**User response:** Reissue the command with a valid peer snapshot name.

**6027-2234 [E] Use the `mmafmctl converttoprimary` command for converting to primary fileset.**

**Explanation:** Converting to a primary fileset is not allowed directly.

**User response:** Check the previous error messages and correct the problems.

**6027-2235 [E] Only independent filesets can be converted to secondary filesets.**

**Explanation:** Converting to secondary filesets is allowed only for independent filesets.

**User response:** None.

---

**6027-2236 [E]** The CPU architecture on this node does not support tracing in *traceMode* mode. Switching to *traceMode* mode.

**Explanation:** The CPU does not have constant time stamp counter capability, which is required for overwrite trace mode. The trace has been enabled in blocking mode.

**User response:** Update the configuration parameters to use the trace facility in blocking mode or replace this node with modern CPU architecture.

---

**6027-2237 [W]** An image backup made from the live file system may not be usable for image restore. Specify a valid global snapshot for image backup.

**Explanation:** The `mmimgbackup` command should always be used with a global snapshot to make a consistent image backup of the file system.

**User response:** Correct the command invocation to include the `-S` option to specify either a global snapshot name or a directory path that includes the snapshot root directory for the file system and a valid global snapshot name.

---

**6027-2238 [E]** Use the `mmafmctl convertToSecondary` command for converting to secondary.

**Explanation:** Converting to secondary is allowed by using the `mmafmctl convertToSecondary` command.

**User response:** None.

---

**6027-2239 [E]** Drive `serialNumber` *serialNumber* is being managed by server *nodeName*. Reissue the `mmchfirmware` command for server *nodeName*.

**Explanation:** The `mmchfirmware` command was issued to update a specific disk drive which is not currently being managed by this node.

**User response:** Reissue the command specifying the active server.

---

**6027-2240 [E]** Option is not supported for a secondary fileset.

**Explanation:** This option cannot be set for a secondary fileset.

**User response:** None.

---

**6027-2241 [E]** Node *nodeName* is not a CES node.

**Explanation:** A Cluster Export Service command specified a node that is not defined as a CES node.

**User response:** Reissue the command specifying a CES node.

---

**6027-2242 [E]** Error in configuration file.

**Explanation:** The `mmnfs export load loadCfgFile` command found an error in the NFS configuration files.

**User response:** Correct the configuration file error.

---

**6027-2245 [E]** To change the AFM target, use `mmafmctl changeSecondary` for the primary.

**Explanation:** Failover with the `targetonly` option can be run on a primary fileset.

**User response:** None.

---

**6027-2246 [E]** Timeout executing function: *functionName* (`return code=returnCode`).

**Explanation:** The `executeCommandWithTimeout` function was called but it timed out.

**User response:** Correct the problem and issue the command again.

---

**6027-2247 [E]** Creation of *exchangeDir* failed.

**Explanation:** A Cluster Export Service command was unable to create the CCR exchange directory.

**User response:** Correct the problem and issue the command again.

---

**6027-2248 [E]** CCR command failed: *command*

**Explanation:** A CCR update command failed.

**User response:** Correct the problem and issue the command again.

---

**6027-2249 [E]** Error getting next *nextName* from CCR.

**Explanation:** An expected value from CCR was not obtained.

**User response:** Issue the command again.

---

**6027-2250 [E]** Error putting next *nextName* to CCR, new ID: *newExpid* version: *version*

**Explanation:** A CCR value update failed.

**User response:** Issue the command again.

---

**6027-2251 [E]** Error retrieving configuration file: *configFile*

**Explanation:** Error retrieving configuration file from CCR.

**User response:** Issue the command again.

---

---

**6027-2252 [E] Error reading export configuration file (return code: *returnCode*).**

**Explanation:** A CES command was unable to read the export configuration file.

**User response:** Correct the problem and issue the command again.

---

**6027-2253 [E] Error creating the internal export data objects (return code *returnCode*).**

**Explanation:** A CES command was unable to create an export data object.

**User response:** Correct the problem and issue the command again.

---

**6027-2254 [E] Error creating single export output, export *exportPath* not found (return code *returnCode*).**

**Explanation:** A CES command was unable to create a single export print output.

**User response:** Correct the problem and reissue the command.

---

**6027-2255 [E] Error creating export output (return code: *returnCode*).**

**Explanation:** A CES command was unable to create the export print output.

**User response:** Correct the problem and issue the command again.

---

**6027-2256 [E] Error creating the internal export output file string array (return code: *returnCode*).**

**Explanation:** A CES command was unable to create the array for print output.

**User response:** Correct the problem and issue the command again.

---

**6027-2257 [E] Error deleting export, export *exportPath* not found (return code: *returnCode*).**

**Explanation:** A CES command was unable to delete an export. The *exportPath* was not found.

**User response:** Correct the problem and issue the command again.

---

**6027-2258 [E] Error writing export configuration file to CCR (return code: *returnCode*).**

**Explanation:** A CES command was unable to write configuration file to CCR.

**User response:** Correct the problem and issue the command again.

---



---

**6027-2259 [E] The path *exportPath* to create the export does not exist (return code: *returnCode*).**

**Explanation:** A CES command was unable to create an export because the path does not exist.

**User response:** Correct the problem and issue the command again.

---

**6027-2260 [E] The path *exportPath* to create the export is invalid (return code: *returnCode*).**

**Explanation:** A CES command was unable to create an export because the path is invalid.

**User response:** Correct the problem and issue the command again.

---

**6027-2261 [E] Error creating new export object, invalid data entered (return code: *returnCode*).**

**Explanation:** A CES command was unable to add an export because the input data is invalid.

**User response:** Correct the problem and issue the command again.

---

**6027-2262 [E] Error creating new export object; getting new export ID (return code: *returnCode*).**

**Explanation:** A CES command was unable to add an export. A new export ID was not obtained.

**User response:** Correct the problem and issue the command again.

---

**6027-2263 [E] Error adding export; new export path *exportPath* already exists.**

**Explanation:** A CES command was unable to add an export because the path already exists.

**User response:** Correct the problem and issue the command again.

---

**6027-2264 [E] The --servers option is only used to provide names for primary and backup server configurations. Provide a maximum of two server names.**

**Explanation:** An input node list has too many nodes specified.

**User response:** Verify the list of nodes and shorten the list to the supported number.

---

**6027-2265 [E] Cannot convert fileset to secondary fileset.**

**Explanation:** Fileset cannot be converted to a secondary fileset.

**User response:** None.

---

---

**6027-2266 [E]** The snapshot names that start with `psnap-rpo` or `psnap0-rpo` are reserved for RPO.

**Explanation:** The specified snapshot name starts with `psnap-rpo` or `psnap0-rpo`, which are reserved for RPO snapshots.

**User response:** Use a different snapshot name for the `mmcrsnapshot` command.

---

**6027-2267 [I]** Fileset *filesetName* in file system *fileSystem* is either unlinked or being deleted. Home delete-snapshot operation was not queued.

**Explanation:** The command expects that the peer snapshot at home is not deleted because the fileset at cache is either unlinked or being deleted.

**User response:** Delete the snapshot at home manually.

---

**6027-2268 [E]** This is already a secondary fileset.

**Explanation:** The fileset is already a secondary fileset.

**User response:** None.

---

**6027-2269 [E]** Adapter *adapterIdentifier* was not found.

**Explanation:** The specified adapter was not found.

**User response:** Specify an existing adapter and reissue the command.

---

**6027-2270 [E]** Error *errno* updating firmware for adapter *adapterIdentifier*.

**Explanation:** The firmware load failed for the specified adapter.

**User response:** None.

---

**6027-2271 [E]** Error locating the reference client in *inputStringContainingClient* (return code: *returnCode*).

**Explanation:** The reference client for reordering a client could not be found for the given export path.

**User response:** Correct the problem and try again.

---

**6027-2272 [E]** Error removing the requested client in *inputStringContainingClient* from a client declaration, return code: *returnCode*

**Explanation:** One of the specified clients to remove could not be found in any client declaration for the given export path.

**User response:** Correct the problem and try again.

---

**6027-2273 [E]** Error adding the requested client in *inputStringContainingClient* to a client declaration, return code: *returnCode*

**Explanation:** One of the specified clients to add could not be applied for the given export path.

**User response:** Correct the problem and try again.

---

**6027-2274 [E]** Error locating the reference client in *inputStringContainingClient* (return code: *returnCode*).

**Explanation:** The reference client for reordering a client could not be applied for the given export path.

**User response:** Correct the problem and try again.

---

**6027-2275 [E]** Unable to determine the status of DASD device *dasdDevice*

**Explanation:** The `dasdview` command failed.

**User response:** Examine the preceding messages, correct the problem, and reissue the command.

---

**6027-2276 [E]** The specified DASD device *dasdDevice* is not properly formatted. It is not an ECKD-type device, or it has a format other than CDL or LDL, or it has a block size other than 4096.

**Explanation:** The specified device is not properly formatted.

**User response:** Correct the problem and reissue the command.

---

**6027-2277 [E]** Unable to determine if DASD device *dasdDevice* is partitioned.

**Explanation:** The `fdasd` command failed.

**User response:** Examine the preceding messages, correct the problem, and reissue the command.

---

**6027-2278 [E]** Cannot partition DASD device *dasdDevice*; it is already partitioned.

**Explanation:** The specified DASD device is already partitioned.

**User response:** Remove the existing partitions, or reissue the command using the desired partition name.

---

**6027-2279 [E]** Unable to partition DASD device *dasdDevice*

**Explanation:** The `fdasd` command failed.

**User response:** Examine the preceding messages, correct the problem, and reissue the command.

---

---

**6027-2280 [E]** The DASD device with bus ID *busID* cannot be found or it is in use.

**Explanation:** The `chccwdev` command failed.

**User response:** Examine the preceding messages, correct the problem, and reissue the command.

---

**6027-2281 [E]** Error *errno* updating firmware for enclosure *enclosureIdentifier*.

**Explanation:** The firmware load failed for the specified enclosure.

**User response:** None.

---

**6027-2282 [E]** Action *action* is not allowed for secondary filesets.

**Explanation:** The specified action is not allowed for secondary filesets.

**User response:** None.

---

**6027-2283 [E]** Node *nodeName* is already a CES node.

**Explanation:** An `mmchnode` command attempted to enable CES services on a node that is already part of the CES cluster.

**User response:** Reissue the command specifying a node that is not a CES node.

---

**6027-2284 [E]** The fileset `afmshowhomesnapshot` value is 'yes'. The fileset mode cannot be changed.

**Explanation:** The fileset `afmshowhomesnapshot` attribute value is `yes`. The fileset mode change is not allowed.

**User response:** First change the attribute `afmshowhomesnapshot` value to `no`, and then issue the command again to change the mode.

---

**6027-2285 [E]** Deletion of initial snapshot *snapshotName* of fileset *filesetName* in file system *fileSystem* failed. The delete fileset operation failed at cache. Error code *errorCode*.

**Explanation:** The deletion of the initial snapshot `psnap0` of *filesetName* failed. The primary and secondary filesets cannot be deleted without deleting the initial snapshot.

**User response:** None.

---



---

**6027-2286 [E]** RPO peer snapshots using `mmpsnap` are allowed only for primary filesets.

**Explanation:** RPO snapshots can be created only for primary filesets.

**User response:** Reissue the command with a valid primary fileset or without the `--rpo` option.

---

**6027-2287** The fileset needs to be linked to change `afmShowHomeSnapshot` to 'no'.

**Explanation:** The `afmShowHomeSnapshot` value cannot be changed to `no` if the fileset is unlinked.

**User response:** Link the fileset and reissue the command.

---

**6027-2288 [E]** Option *optionName* is not supported for AFM filesets.

**Explanation:** IAM modes are not supported for AFM filesets.

**User response:** None.

---

**6027-2289 [E]** Peer snapshot creation failed while running *subCommand*. Error code *errorCode*

**Explanation:** For an active fileset, check the AFM target configuration for peer snapshots. Ensure there is at least one gateway node configured for the cluster. Examine the preceding messages and the GPFS log for additional details.

**User response:** Correct the problems and reissue the command.

---

**6027-2290 [E]** The comment string should be less than 50 characters long.

**Explanation:** The comment/prefix string of the snapshot is longer than 50 characters.

**User response:** Reduce the comment string size and reissue the command.

---

**6027-2291 [E]** Peer snapshot creation failed while generating snapshot name. Error code *errorCode*

**Explanation:** For an active fileset, check the AFM target configuration for peer snapshots. Ensure there is at least one gateway node configured for the cluster. Examine the preceding messages and the GPFS log for additional details.

**User response:** Correct the problems and reissue the command.

---

---

**6027-2292 [E]** The initial snapshot *psnap0Name* does not exist. The peer snapshot creation failed. Error code *errorCode*

**Explanation:** For an active fileset, check the AFM target configuration for peer snapshots. Ensure the initial peer snapshot exists for the fileset. Examine the preceding messages and the GPFS log for additional details.

**User response:** Verify that the fileset is a primary fileset and that it has *psnap0* created and try again.

---

**6027-2293 [E]** The peer snapshot creation failed because fileset *filesetName* is in *filesetState* state.

**Explanation:** For an active fileset, check the AFM target configuration for peer snapshots. Ensure there is at least one gateway node configured for the cluster. Examine the preceding messages and the GPFS log for additional details.

**User response:** None. The fileset needs to be in active or dirty state.

---

**6027-2294 [E]** Removing older peer snapshots failed while obtaining snap IDs. Error code *errorCode*

**Explanation:** Ensure the fileset exists. Examine the preceding messages and the GPFS log for additional details.

**User response:** Verify that snapshots exist for the given fileset.

---

**6027-2295 [E]** Removing older peer snapshots failed while obtaining old snap IDs. Error code *errorCode*

**Explanation:** Ensure the fileset exists. Examine the preceding messages and the GPFS log for additional details.

**User response:** Verify that snapshots exist for the given fileset.

---

**6027-2296 [E]** Need a target to convert to the primary fileset.

**Explanation:** Need a target to convert to the primary fileset.

**User response:** Specify a target to convert to the primary fileset.

---

**6027-2297 [E]** The check-metadata and nocheck-metadata options are not supported for a non-AFM fileset.

**Explanation:** The check-metadata and

nocheck-metadata options are not supported for a non-AFM fileset.

**User response:** None.

---

**6027-2298 [E]** Only independent filesets can be converted to primary or secondary.

**Explanation:** Only independent filesets can be converted to primary or secondary.

**User response:** Specify an independent fileset.

---

**6027-2299 [E]** Issue the `mmafmctl getstate` command to check fileset state and if required issue `mmafmctl convertToPrimary`.

**Explanation:** Issue the `mmafmctl getstate` command to check fileset state and if required issue `mmafmctl convertToPrimary`.

**User response:** Issue the `mmafmctl getstate` command to check fileset state and if required issue `mmafmctl convertToPrimary`.

---

**6027-2300 [E]** The check-metadata and nocheck-metadata options are not supported for the primary fileset.

**Explanation:** The check-metadata and nocheck-metadata options are not supported for the primary fileset.

**User response:** None.

---

**6027-2301 [E]** The inband option is not supported for the primary fileset.

**Explanation:** The inband option is not supported for the primary fileset.

**User response:** None.

---

**6027-2302 [E]** AFM target cannot be changed for the primary fileset.

**Explanation:** AFM target cannot be changed for the primary fileset.

**User response:** None.

---

**6027-2303 [E]** The inband option is not supported for an AFM fileset.

**Explanation:** The inband option is not supported for an AFM fileset.

**User response:** None.



---

**6027-2304 [E]** Target cannot be changed for an AFM fileset.

**Explanation:** Target cannot be changed for an AFM fileset.

**User response:** None.

---

**6027-2305 [E]** The `mmafmctl convertToPrimary` command is not allowed for this primary fileset.

**Explanation:** The `mmafmctl convertToPrimary` command is not allowed for the primary fileset because it is not in `PrimInitFail` state.

**User response:** None.

---

**6027-2306 [E]** Failed to check for cached files while doing primary conversion from `filesetMode` mode.

**Explanation:** Failed to check for cached files while doing primary conversion.

**User response:** None.

---

**6027-2307 [E]** Uncached files present, run `prefetch` first.

**Explanation:** Uncached files present.

**User response:** Run `prefetch` and then do the conversion.

---

**6027-2308 [E]** Uncached files present, run `prefetch` first using policy output: `nodeDirFileOut`.

**Explanation:** Uncached files present.

**User response:** Run `prefetch` first using policy output.

---

**6027-2309 [E]** Conversion to primary not allowed for `filesetMode` mode.

**Explanation:** Conversion to primary not allowed for this mode.

**User response:** None.

---

**6027-2310 [E]** This option is available only for a primary fileset.

**Explanation:** This option is available only for a primary fileset.

**User response:** None.

---

**6027-2311 [E]** The `target-only` option is not allowed for a promoted primary without a target.

**Explanation:** The `target-only` option is not allowed for a promoted primary without a target.

**User response:** None.

---

**6027-2312 [E]** Need a target to setup the new secondary.

**Explanation:** Target is required to setup the new secondary.

**User response:** None.

---

**6027-2313 [E]** The `target-only` and `inband` options are not allowed together.

**Explanation:** The `target-only` and `inband` options are not allowed together.

**User response:** None.

---

**6027-2314 [E]** Could not run `commandName`. Verify that the Object protocol was installed.

**Explanation:** The `mmcesobjlscfg` command cannot find a prerequisite command on the system.

**User response:** Install the missing command and try again.

---

**6027-2315 [E]** Could not determine CCR file for service `serviceName`

**Explanation:** For the given service name, there is not a corresponding file in the CCR.

**User response:** None.

---

**6027-2316 [E]** Unable to retrieve file `fileName` from CCR using `command` command. Verify that the Object protocol is correctly installed.

**Explanation:** There was an error downloading a file from the CCR repository.

**User response:** Correct the error and try again.

---

**6027-2317 [E]** Unable to parse version number of file `fileName` from `mmccr` output

**Explanation:** The current version should be printed by `mmccr` when a file is extracted. The command could not read the version number from the output and failed.

**User response:** Investigate the failure in the CCR and fix the problem.

---

**6027-2318 [E]** Could not put `localFilePath` into the CCR as `ccrName`

**Explanation:** There was an error when trying to do an `fput` of a file into the CCR.

**User response:** Investigate the error and fix the problem.

---

**6027-2319 [I] Version mismatch during upload of *fileName* (*version*). Retrying.**

**Explanation:** The file could not be uploaded to the CCR because another process updated it in the meantime. The file will be downloaded, modified, and uploaded again.

**User response:** None. The upload will automatically be tried again.

---

**6027-2320 *directoryName* does not resolve to a directory in *deviceName*. The directory must be within the specified file system.**

**Explanation:** The cited directory does not belong to the specified file system.

**User response:** Correct the directory name and reissue the command.

---

**6027-2321 [E] AFM primary or secondary filesets cannot be created for file system *fileSystem* because version is less than *supportedVersion*.**

**Explanation:** The AFM primary or secondary filesets are not supported for a file system version that is less than 14.20.

**User response:** Upgrade the file system and reissue the command.

---

**6027-2322 [E] The OBJ service cannot be enabled because it is not installed. The file *fileName* was not found.**

**Explanation:** The node could not enable the CES OBJ service because of a missing binary or configuration file.

**User response:** Install the required software and retry the command.

---

**6027-2323 [E] The OBJ service cannot be enabled because the number of CES IPs below the minimum of *minValue* expected.**

**Explanation:** The value of CES IPs was below the minimum.

**User response:** Add at least *minValue* CES IPs to the cluster.

---

**6027-2324 [E] The object store for *serviceName* is either not a GPFS type or *mountPoint* does not exist.**

**Explanation:** The object store is not available at this time.

**User response:** Verify that *serviceName* is a GPFS type.

Verify that the *mountPoint* exists, the file system is mounted, or the fileset is linked.

---

**6027-2325 [E] File *fileName* does not exist in CCR. Verify that the Object protocol is correctly installed.**

**Explanation:** There was an error verifying Object config and ring files in the CCR repository.

**User response:** Correct the error and try again.

---

**6027-2326 [E] The OBJ service cannot be enabled because attribute *attributeName* for a CES IP has not been defined. Verify that the Object protocol is correctly installed.**

**Explanation:** There was an error verifying *attributeName* on CES IPs.

**User response:** Correct the error and try again.

---

**6027-2327 The snapshot *snapshotName* is the wrong scope for use in *targetType* backup**

**Explanation:** The snapshot specified is the wrong scope.

**User response:** Please provide a valid snapshot name for this backup type.

---

**6027-2329 [E] The fileset attributes cannot be set for the primary fileset with caching disabled.**

**Explanation:** The fileset attributes cannot be set for the primary fileset with caching disabled.

**User response:** None.

---

**6027-2330 [E] The outband option is not supported for AFM filesets.**

**Explanation:** The outband option is not supported for AFM filesets.

**User response:** None.

---

**6027-2331 [E] CCR value *ccrValue* not defined. The OBJ service cannot be enabled if identity authentication is not configured.**

**Explanation:** Object authentication type was not found.

**User response:** Configure identity authentication and try again.

---

**6027-2332 [E] Only regular independent filesets are converted to secondary filesets.**

**Explanation:** Only regular independent filesets can be converted to secondary filesets.

**User response:** Specify a regular independent fileset and run the command again.

---

**6027-2333 [E] Failed to disable *serviceName* service. Ensure *authType* authentication is removed.**

**Explanation:** Disable CES service failed because authentication was not removed.

**User response:** Remove authentication and retry.

---

**6027-2334 [E] Fileset *indFileset* cannot be changed because it has a dependent fileset *depFileset***

**Explanation:** Filesets with dependent filesets cannot be converted to primary or secondary.

**User response:** This operation cannot proceed until all the dependent filesets are unlinked.

---

**6027-2335 [E] Failed to convert fileset, because the policy to detect special files is failing.**

**Explanation:** The policy to detect special files is failing.

**User response:** Retry the command later.

---

**6027-2336 [E] Immutable/append-only files or clones copied from a snapshot are present, hence conversion is disallowed**

**Explanation:** Conversion is disallowed if immutable/append-only files or clones copied from a snapshot are present.

**User response:** Files should not be immutable/append-only.

---

**6027-2337 [E] Conversion to primary is not allowed at this time. Retry the command later.**

**Explanation:** Conversion to primary is not allowed at this time.

**User response:** Retry the command later.

---

**6027-2338 [E] Conversion to primary is not allowed because the state of the fileset is *filesetState*.**

**Explanation:** Conversion to primary is not allowed with the current state of the fileset.

**User response:** Retry the command later.

---



---

**6027-2339 [E] Orphans are present, run prefetch first.**

**Explanation:** Orphans are present.

**User response:** Run prefetch on the fileset and then do the conversion.

---

**6027-2340 [E] Fileset was left in PrimInitFail state. Take the necessary actions.**

**Explanation:** The fileset was left in PrimInitFail state.

**User response:** Take the necessary actions.

---

**6027-2341 [E] This operation can be done only on a primary fileset**

**Explanation:** This is not a primary fileset.

**User response:** None.

---

**6027-2342 [E] Failover/resync is currently running so conversion is not allowed**

**Explanation:** Failover/resync is currently running so conversion is not allowed.

**User response:** Retry the command later after failover/resync completes.

---

**6027-2343 [E] DR Setup cannot be done on a fileset with mode *filesetMode*.**

**Explanation:** Setup cannot be done on a fileset with this mode.

**User response:** None.

---

**6027-2344 [E] The GPFS daemon must be active on the node from which the *mmcmd* is executed with option *--inode-criteria* or *-o*.**

**Explanation:** The GPFS daemon needs to be active on the node where the command is issued with *--inode-criteria* or *-o* options.

**User response:** Run the command where the daemon is active.

---

**6027-2345 [E] The provided snapshot name must be unique to list filesets in a specific snapshot**

**Explanation:** The *mmlsfileset* command received a snapshot name that is not unique.

**User response:** Correct the command invocation or remove the duplicate named snapshots and try again.

---

---

**6027-2346 [E] The local node is not a CES node.**

**Explanation:** A local Cluster Export Service command was invoked on a node that is not defined as a Cluster Export Service node.

**User response:** Reissue the command on a CES node.

---

**6027-2347 [E] Error changing export, export *exportPath* not found.**

**Explanation:** A CES command was unable to change an export. The *exportPath* was not found.

**User response:** Correct problem and issue the command again.

---

**6027-2348 [E] A device for *directoryName* does not exist or is not active on this node.**

**Explanation:** The device containing the specified directory does not exist or is not active on the node.

**User response:** Reissue the command with a correct directory or on an appropriate node.

---

**6027-2349 [E] The fileset for *junctionName* does not exist in the *targetType* specified.**

**Explanation:** The fileset to back up cannot be found in the file system or snapshot specified.

**User response:** Reissue the command with a correct name for the fileset, snapshot, or file system.

---

**6027-2350 [E] The fileset for *junctionName* is not linked in the *targetType* specified.**

**Explanation:** The fileset to back up is not linked in the file system or snapshot specified.

**User response:** Relink the fileset in the file system. Optionally create a snapshot and reissue the command with a correct name for the fileset, snapshot, and file system.

---

**6027-2351 [E] One or more unlinked filesets (*filesetName*s) exist in the *targetType* specified. Check your filesets and try again.**

**Explanation:** The file system to back up contains one or more filesets that are unlinked in the file system or snapshot specified.

**User response:** Relink the fileset in the file system. Optionally create a snapshot and reissue the command with a correct name for the fileset, snapshot, and file system.

---



---

**6027-2352 The snapshot *snapshotName* could not be found for use by *commandName***

**Explanation:** The snapshot specified could not be located.

**User response:** Please provide a valid snapshot name.

---

**6027-2353 [E] The snapshot name cannot be generated.**

**Explanation:** The snapshot name cannot be generated.

**User response:** None.

---

**6027-2354 Node *nodeName* must be disabled as a CES node before trying to remove it from the GPFS cluster.**

**Explanation:** The specified node is defined as a CES node.

**User response:** Disable the CES node and try again.

---

**6027-2355 [E] Unable to reload *moduleName*. Node *hostname* should be rebooted.**

**Explanation:** Host adapter firmware was updated so the specified module needs to be unloaded and reloaded. Linux does not display the new firmware level until the module is reloaded.

**User response:** Reboot the node.

---

**6027-2356 [E] Node *nodeName* is being used as a recovery group server.**

**Explanation:** The specified node is defined as a server node for some disk.

**User response:** If you are trying to delete the node from the GPFS cluster, you must either delete the disk or define another node as its server.

---

**6027-2357 [E] Root fileset cannot be converted to primary fileset.**

**Explanation:** Root fileset cannot be converted to the primary fileset.

**User response:** None.

---

**6027-2358 [E] Root fileset cannot be converted to secondary fileset.**

**Explanation:** Root fileset cannot be converted to the secondary fileset.

**User response:** None.

---

---

**6027-2359 [I] Attention:** *command* is now enabled. This attribute can no longer be modified.

**Explanation:** Indefinite retention protection is enabled. This value can not be changed in the future.

**User response:** None.

---

**6027-2360 [E] The current value of *command* is *attrName*. This value cannot be changed.**

**Explanation:** Indefinite retention protection is enabled for this cluster and this attribute cannot be changed.

**User response:** None.

---

**6027-2361 [E] *command* is enabled. File systems cannot be deleted.**

**Explanation:** When indefinite retention protection is enabled the file systems cannot be deleted.

**User response:** None.

---

**6027-2362 [E] The current value of *command* is *attrName*. No changes made.**

**Explanation:** The current value and the request value are the same. No changes made.

**User response:** None.

---

**6027-2363 [E] Operation is not permitted as state of the fileset is *filesetState*.**

**Explanation:** This operation is not allowed with the current state of the fileset.

**User response:** Retry the command later.

---

**6027-2364 [E] Fileset name is missing.**

**Explanation:** This operation needs to be run for a particular fileset.

**User response:** Retry the command with a fileset name.

---

**6027-2365 [E] Firmware loader *filename* not executable.**

**Explanation:** The listed firmware loader is not executable.

**User response:** Make the firmware loader executable and retry the command.

---

**6027-2366** Node *nodeName* is being used as an NSD server. This may include Local Read Only Cache (LROC) storage. Review these details and determine the NSD type by running the `mm1nsd` command. For standard NSDs, you must either delete the disk or define another node as its server. For nodes that include LROC NSDs (local cache) must have all the LROC NSDs removed before the node can be deleted. Fully review the `mmde1nsd` command documentation before making any changes.

**Explanation:** The specified node is defined as a server node for some disk.

**User response:** If you are trying to delete the node from the GPFS cluster, you must either delete the disk or define another node as its server.

---

**6027-2367 [E] Fileset having *iammode* mode cannot be converted to primary fileset.**

**Explanation:** Fileset with Integrated Archive Manager (IAM) mode cannot be converted to primary fileset.

**User response:** None.

---

**6027-2368 [E] Unable to find information for Hypervisor.**

**Explanation:** The `lscpu` command failed.

**User response:** Examine the preceding messages, correct the problem, and reissue the command.

---

**6027-2369 [E] Unable to list DASD devices**

**Explanation:** The `lsdasd` command failed.

**User response:** Examine the preceding messages, correct the problem, and reissue the command.

---

**6027-2370 [E] Unable to flush buffer for DASD device *name1***

**Explanation:** The `blockdev --flushbufs` command failed.

**User response:** Examine the preceding messages, correct the problem, and reissue the command.

---

**6027-2371 [E] Unable to read the partition table for DASD device *dasdDevice*.**

**Explanation:** The `blockdev --rereadpt` command failed.

**User response:** Examine the preceding messages, correct the problem, and reissue the command.

---

---

6027-2372 [E] Unable to find information to DASD device *dasdDevice*.

**Explanation:** The **dasdinfo** command failed.

**User response:** Examine the preceding messages, correct the problem, and reissue the command.

---

6027-2373 *feature is only available in the IBM Spectrum Scale Advanced Edition.*

**Explanation:** The specified function or feature is only part of the IBM Spectrum Scale Advanced Edition.

**User response:** Install the IBM Spectrum Scale Advanced Edition on all nodes in the cluster, and then reissue the command.

---

6027-2374 [E] Unable to delete recovery group *name*; associated VDisk sets are still defined.

**Explanation:** Cannot delete a recovery group while there are still VDisk sets defined.

**User response:** Delete all the VDisk sets that use the recovery group first.

---

6027-2376 [E] Node class *nodeclass* cannot be action. It is marked for use by Transparent Cloud Tiering. To remove this node class, first disable all the nodes with **mmchnode --cloud-gateway-disable**.

**Explanation:** Cannot delete a node class that has cloud gateway enabled.

**User response:** Disable the nodes first with **mmchnode --cloud-gateway-disable**.

---

6027-2377 [E] Node *nodeclass* cannot be deleted. It is marked for use by Transparent Cloud Tiering. To remove this node, first disable it with **mmchnode --cloud-gateway-disable**.

**Explanation:** Cannot delete a node that has cloud gateway enabled.

**User response:** Disable the node first with **mmchnode --cloud-gateway-disable**.

---

6027-2378 [E] To enable Transparent Cloud Tiering nodes, you must first enable the Transparent Cloud Tiering feature. This feature provides a new level of storage tiering capability to the Spectrum Scale customer. Therefore, the IBM development team wants to proactively work with clients who are going to be implementing this feature. Please contact Spectrum Scale Support at [scale@us.ibm.com](mailto:scale@us.ibm.com) to have your use case reviewed. IBM will evaluate your use case, work with you to ensure your environment is appropriate for Transparent Cloud Tiering, and will provide guidance on enabling the feature. Please include this message when contacting IBM Support.

**Explanation:** The Transparent Cloud Tiering feature must be enabled with assistance from IBM.

**User response:** Contact IBM support for more information.

---

6027-2500 **mmsanrepairfs already in progress for "*name*"**

**Explanation:** This is an output from **mmsanrepairfs** when another **mmsanrepairfs** command is already running.

**User response:** Wait for the currently running command to complete and reissue the command.

---

6027-2501 **Could not allocate storage.**

**Explanation:** Sufficient memory could not be allocated to run the **mmsanrepairfs** command.

**User response:** Increase the amount of memory available.

---

6027-2576 [E] **Error: Daemon *value* kernel *value* PAGE\_SIZE mismatch.**

**Explanation:** The GPFS kernel extension loaded in memory does not have the same **PAGE\_SIZE** value as the GPFS daemon **PAGE\_SIZE** value that was returned from the POSIX **sysconf** API.

**User response:** Verify that the kernel header files used to build the GPFS portability layer are the same kernel header files used to build the running kernel.

---

6027-2600 **Cannot create a new snapshot until an existing one is deleted. File system *fileSystem* has a limit of *number* online snapshots.**

**Explanation:** The file system has reached its limit of online snapshots

**User response:** Delete an existing snapshot, then issue the create snapshot command again.

---

**6027-2601** Snapshot name *dirName* already exists.

**Explanation:** by the `tscrsnapshot` command.

**User response:** Delete existing file/directory and reissue the command.

---

**6027-2602** Unable to delete snapshot *snapshotName* from file system *fileSystem*. *rc=returnCode*.

**Explanation:** This message is issued by the `tscrsnapshot` command.

**User response:** Delete the snapshot using the `tsdelsnapshot` command.

---

**6027-2603** Unable to get permission to create snapshot, *rc=returnCode*.

**Explanation:** This message is issued by the `tscrsnapshot` command.

**User response:** Reissue the command.

---

**6027-2604** Unable to quiesce all nodes, *rc=returnCode*.

**Explanation:** This message is issued by the `tscrsnapshot` command.

**User response:** Restart failing nodes or switches and reissue the command.

---

**6027-2605** Unable to resume all nodes, *rc=returnCode*.

**Explanation:** This message is issued by the `tscrsnapshot` command.

**User response:** Restart failing nodes or switches.

---

**6027-2606** Unable to sync all nodes, *rc=returnCode*.

**Explanation:** This message is issued by the `tscrsnapshot` command.

**User response:** Restart failing nodes or switches and reissue the command.

---

**6027-2607** Cannot create new snapshot until an existing one is deleted. Fileset *filesetName* has a limit of *number* snapshots.

**Explanation:** The fileset has reached its limit of snapshots.

**User response:** Delete an existing snapshot, then issue the create snapshot command again.

---

**6027-2608** Cannot create new snapshot: state of fileset *filesetName* is inconsistent (*badState*).

**Explanation:** An operation on the cited fileset is incomplete.

**User response:** Complete pending fileset actions, then issue the create snapshot command again.

---

**6027-2609** Fileset named *filesetName* does not exist.

**Explanation:** One of the filesets listed does not exist.

**User response:** Specify only existing fileset names.

---

**6027-2610** File system *fileSystem* does not contain snapshot *snapshotName* *err = number*.

**Explanation:** An incorrect snapshot name was specified.

**User response:** Select a valid snapshot and issue the command again.

---

**6027-2611** Cannot delete snapshot *snapshotName* which is in state *snapshotState*.

**Explanation:** The snapshot cannot be deleted while it is in the cited transition state because of an in-progress snapshot operation.

**User response:** Wait for the in-progress operation to complete and then reissue the command.

---

**6027-2612** Snapshot named *snapshotName* does not exist.

**Explanation:** A snapshot to be listed does not exist.

**User response:** Specify only existing snapshot names.

---

**6027-2613** Cannot restore snapshot. *fileSystem* is mounted on *number* node(s) and in use on *number* node(s).

**Explanation:** This message is issued by the `tsressnapshot` command.

**User response:** Unmount the file system and reissue the restore command.

---

**6027-2614** File system *fileSystem* does not contain snapshot *snapshotName* *err = number*.

**Explanation:** An incorrect snapshot name was specified.

**User response:** Specify a valid snapshot and issue the command again.

---

**6027-2615** Cannot restore snapshot *snapshotName* which is *snapshotState*, **err = number**.

**Explanation:** The specified snapshot is not in a valid state.

**User response:** Specify a snapshot that is in a valid state and issue the command again.

---

**6027-2616** Restoring snapshot *snapshotName* requires *quotaTypes* **quotas to be enabled**.

**Explanation:** The snapshot being restored requires quotas to be enabled, since they were enabled when the snapshot was created.

**User response:** Issue the recommended **mmchfs** command to enable quotas.

---

**6027-2617** You must run: **mmchfs** *fileSystem* **-Q yes**.

**Explanation:** The snapshot being restored requires quotas to be enabled, since they were enabled when the snapshot was created.

**User response:** Issue the cited **mmchfs** command to enable quotas.

---

**6027-2618 [N]** Restoring snapshot *snapshotName* in file system *fileSystem* requires *quotaTypes* **quotas to be enabled**.

**Explanation:** The snapshot being restored in the cited file system requires quotas to be enabled, since they were enabled when the snapshot was created.

**User response:** Issue the **mmchfs** command to enable quotas.

---

**6027-2619** Restoring snapshot *snapshotName* requires *quotaTypes* **quotas to be disabled**.

**Explanation:** The snapshot being restored requires quotas to be disabled, since they were not enabled when the snapshot was created.

**User response:** Issue the cited **mmchfs** command to disable quotas.

---

**6027-2620** You must run: **mmchfs** *fileSystem* **-Q no**.

**Explanation:** The snapshot being restored requires quotas to be disabled, since they were not enabled when the snapshot was created.

**User response:** Issue the cited **mmchfs** command to disable quotas.

---

**6027-2621 [N]** Restoring snapshot *snapshotName* in file system *fileSystem* requires *quotaTypes* **quotas to be disabled**.

**Explanation:** The snapshot being restored in the cited file system requires quotas to be disabled, since they were disabled when the snapshot was created.

**User response:** Issue the **mmchfs** command to disable quotas.

---

**6027-2623 [E]** Error deleting snapshot *snapshotName* in file system *fileSystem* **err number**

**Explanation:** The cited snapshot could not be deleted during file system recovery.

**User response:** Run the **mmfsck** command to recover any lost data blocks.

---

**6027-2624** Previous snapshot *snapshotName* is not valid and must be deleted before a new snapshot may be created.

**Explanation:** The cited previous snapshot is not valid and must be deleted before a new snapshot may be created.

**User response:** Delete the previous snapshot using the **mmdelnsnapshot** command, and then reissue the original snapshot command.

---

**6027-2625** Previous snapshot *snapshotName* must be restored before a new snapshot may be created.

**Explanation:** The cited previous snapshot must be restored before a new snapshot may be created.

**User response:** Run **mmrestorefs** on the previous snapshot, and then reissue the original snapshot command.

---

**6027-2626** Previous snapshot *snapshotName* is not valid and must be deleted before another snapshot may be deleted.

**Explanation:** The cited previous snapshot is not valid and must be deleted before another snapshot may be deleted.

**User response:** Delete the previous snapshot using the **mmdelnsnapshot** command, and then reissue the original snapshot command.

---

**6027-2627** Previous snapshot *snapshotName* is not valid and must be deleted before another snapshot may be restored.

**Explanation:** The cited previous snapshot is not valid and must be deleted before another snapshot may be restored.



**User response:** Delete the previous snapshot using the `mmdelnsnapshot` command, and then reissue the original snapshot command.

---

**6027-2628**    **More than one snapshot is marked for restore.**

**Explanation:** More than one snapshot is marked for restore.

**User response:** Restore the previous snapshot and then reissue the original snapshot command.

---

**6027-2629**    **Offline snapshot being restored.**

**Explanation:** An offline snapshot is being restored.

**User response:** When the restore of the offline snapshot completes, reissue the original snapshot command.

---

**6027-2630**    *Program failed, error number.*

**Explanation:** The `tssnaplatest` command encountered an error and `printErrnoMsg` failed.

**User response:** Correct the problem shown and reissue the command.

---

**6027-2631**    **Attention: Snapshot *snapshotName* was being restored to *fileSystem*.**

**Explanation:** A file system in the process of a snapshot restore cannot be mounted except under a restricted mount.

**User response:** None. Informational message only.

---

**6027-2633**    **Attention: Disk configuration for *fileSystem* has changed while `tsdf` was running.**

**Explanation:** The disk configuration for the cited file system changed while the `tsdf` command was running.

**User response:** Reissue the `mmdf` command.

---

**6027-2634**    **Attention: *number of number* regions in *fileSystem* were unavailable for free space.**

**Explanation:** Some regions could not be accessed during the `tsdf` run. Typically, this is due to utilities such `mmdefragfs` or `mmfsck` running concurrently.

**User response:** Reissue the `mmdf` command.

---

**6027-2635**    **The free space data is not available. Reissue the command without the `-q` option to collect it.**

**Explanation:** The existing free space information for the file system is currently unavailable.

**User response:** Reissue the `mmdf` command.

---

**6027-2636**    **Disks in storage pool *storagePool* must have disk usage type `dataOnly`.**

**Explanation:** A non-system storage pool cannot hold metadata or descriptors.

**User response:** Modify the command's disk descriptors and reissue the command.

---

**6027-2637**    **The file system must contain at least one disk for metadata.**

**Explanation:** The disk descriptors for this command must include one and only one storage pool that is allowed to contain metadata.

**User response:** Modify the command's disk descriptors and reissue the command.

---

**6027-2638**    **Maximum of *number* storage pools allowed.**

**Explanation:** The cited limit on the number of storage pools that may be defined has been exceeded.

**User response:** Modify the command's disk descriptors and reissue the command.

---

**6027-2639**    **Incorrect fileset name *filesetName*.**

**Explanation:** The fileset name provided in the command invocation is incorrect.

**User response:** Correct the fileset name and reissue the command.

---

**6027-2640**    **Incorrect path to fileset junction *filesetJunction*.**

**Explanation:** The path to the cited fileset junction is incorrect.

**User response:** Correct the junction path and reissue the command.

---

**6027-2641**    **Incorrect fileset junction name *filesetJunction*.**

**Explanation:** The cited junction name is incorrect.

**User response:** Correct the junction name and reissue the command.

---

**6027-2642**    **Specify one and only one of `FilesetName` or `-J JunctionPath`.**

**Explanation:** The change fileset and unlink fileset commands accept either a fileset name or the fileset's junction path to uniquely identify the fileset. The user failed to provide either of these, or has tried to provide both.

**User response:** Correct the command invocation and reissue the command.

---

**6027-2643**    **Cannot create a new fileset until an existing one is deleted. File system *fileSystem* has a limit of *maxNumber* filesets.**

**Explanation:** An attempt to create a fileset for the cited file system failed because it would exceed the cited limit.

**User response:** Remove unneeded filesets and reissue the command.

---

**6027-2644**    **Comment exceeds maximum length of *maxNumber* characters.**

**Explanation:** The user-provided comment for the new fileset exceeds the maximum allowed length.

**User response:** Shorten the comment and reissue the command.

---

**6027-2645**    **Fileset *filesetName* already exists.**

**Explanation:** An attempt to create a fileset failed because the specified fileset name already exists.

**User response:** Select a unique name for the fileset and reissue the command.

---

**6027-2646**    **Unable to sync all nodes while quiesced, rc=*returnCode***

**Explanation:** This message is issued by the `tscrsnapshot` command.

**User response:** Restart failing nodes or switches and reissue the command.

---

**6027-2647**    **Fileset *filesetName* must be unlinked to be deleted.**

**Explanation:** The cited fileset must be unlinked before it can be deleted.

**User response:** Unlink the fileset, and then reissue the delete command.

---

**6027-2648**    **Filesets have not been enabled for file system *fileSystem*.**

**Explanation:** The current file system format version does not support filesets.

**User response:** Change the file system format version by issuing `mmchfs -V`.

---

**6027-2649**    **Fileset *filesetName* contains user files and cannot be deleted unless the `-f` option is specified.**

**Explanation:** An attempt was made to delete a non-empty fileset.

**User response:** Remove all files and directories from the fileset, or specify the `-f` option to the `mmdelfileset` command.

---

**6027-2650**    **Fileset information is not available.**

**Explanation:** A fileset command failed to read file system metadata file. The file system may be corrupted.

**User response:** Run the `mmfsck` command to recover the file system.

---

**6027-2651**    **Fileset *filesetName* cannot be unlinked.**

**Explanation:** The user tried to unlink the root fileset, or is not authorized to unlink the selected fileset.

**User response:** None. The fileset cannot be unlinked.

---

**6027-2652**    **Fileset at *junctionPath* cannot be unlinked.**

**Explanation:** The user tried to unlink the root fileset, or is not authorized to unlink the selected fileset.

**User response:** None. The fileset cannot be unlinked.

---

**6027-2653**    **Failed to unlink fileset *filesetName* from *filesetName*.**

**Explanation:** An attempt was made to unlink a fileset that is linked to a parent fileset that is being deleted.

**User response:** Delete or unlink the children, and then delete the parent fileset.

---

**6027-2654**    **Fileset *filesetName* cannot be deleted while other filesets are linked to it.**

**Explanation:** The fileset to be deleted has other filesets linked to it, and cannot be deleted without using the `-f` flag, or unlinking the child filesets.

**User response:** Delete or unlink the children, and then delete the parent fileset.

---

**6027-2655**    **Fileset *filesetName* cannot be deleted.**

**Explanation:** The user is not allowed to delete the root fileset.

**User response:** None. The fileset cannot be deleted.

---

**6027-2656**    **Unable to quiesce fileset at all nodes.**

**Explanation:** An attempt to quiesce the fileset at all nodes failed.

**User response:** Check communication hardware and reissue the command.

---

**6027-2657**    **Fileset *filesetName* has open files. Specify -f to force unlink.**

**Explanation:** An attempt was made to unlink a fileset that has open files.

**User response:** Close the open files and then reissue command, or use the **-f** option on the unlink command to force the open files to close.

---

**6027-2658**    **Fileset *filesetName* cannot be linked into a snapshot at *pathName*.**

**Explanation:** The user specified a directory within a snapshot for the junction to a fileset, but snapshots cannot be modified.

**User response:** Select a directory within the active file system, and reissue the command.

---

**6027-2659**    **Fileset *filesetName* is already linked.**

**Explanation:** The user specified a fileset that was already linked.

**User response:** Unlink the fileset and then reissue the link command.

---

**6027-2660**    **Fileset *filesetName* cannot be linked.**

**Explanation:** The fileset could not be linked. This typically happens when the fileset is in the process of being deleted.

**User response:** None.

---

**6027-2661**    **Fileset junction *pathName* already exists.**

**Explanation:** A file or directory already exists at the specified junction.

**User response:** Select a new junction name or a new directory for the link and reissue the link command.

---

**6027-2662**    **Directory *pathName* for junction has too many links.**

**Explanation:** The directory specified for the junction has too many links.

**User response:** Select a new directory for the link and reissue the command.

---



---

**6027-2663**    **Fileset *filesetName* cannot be changed.**

**Explanation:** The user specified a fileset to **tschfileset** that cannot be changed.

**User response:** None. You cannot change the attributes of the root fileset.

---

**6027-2664**    **Fileset at *pathName* cannot be changed.**

**Explanation:** The user specified a fileset to **tschfileset** that cannot be changed.

**User response:** None. You cannot change the attributes of the root fileset.

---

**6027-2665**    **mmfileid already in progress for *name*.**

**Explanation:** An **mmfileid** command is already running.

**User response:** Wait for the currently running command to complete, and issue the new command again.

---

**6027-2666**    **mmfileid can only handle a maximum of *diskAddresses* disk addresses.**

**Explanation:** Too many disk addresses specified.

**User response:** Provide less than 256 disk addresses to the command.

---

**6027-2667 [I]**    **Allowing block allocation for file system *fileSystem* that makes a file ill-replicated due to insufficient *resource* and puts data at risk.**

**Explanation:** The **partialReplicaAllocation** file system option allows allocation to succeed even when all replica blocks cannot be allocated. The file was marked as not replicated correctly and the data may be at risk if one of the remaining disks fails.

**User response:** None. Informational message only.

---

**6027-2670**    **Fileset name *filesetName* not found.**

**Explanation:** The fileset name that was specified with the command invocation was not found.

**User response:** Correct the fileset name and reissue the command.

---

**6027-2671**    **Fileset command on *fileSystem* failed; snapshot *snapshotName* must be restored first.**

**Explanation:** The file system is being restored either from an offline backup or a snapshot, and the restore operation has not finished. Fileset commands cannot be run.

**User response:** Run the **mmrestorefs** command to

---

complete the snapshot restore operation or to finish the offline restore, then reissue the fileset command.

---

**6027-2672**    **Junction parent directory inode number *inodeNumber* is not valid.**

**Explanation:** An inode number passed to `tslinkfileset` is not valid.

**User response:** Check the `mmlinkfileset` command arguments for correctness. If a valid junction path was provided, contact the IBM Support Center.

---

**6027-2673 [X]**    **Duplicate owners of an allocation region (index *indexNumber*, region *regionNumber*, pool *poolNumber*) were detected for file system *fileSystem*: nodes *nodeName* and *nodeName*.**

**Explanation:** The allocation region should not have duplicate owners.

**User response:** Contact the IBM Support Center.

---

**6027-2674 [X]**    **The owner of an allocation region (index *indexNumber*, region *regionNumber*, pool *poolNumber*) that was detected for file system *fileSystem*: node *nodeName* is not valid.**

**Explanation:** The file system had detected a problem with the ownership of an allocation region. This may result in a corrupted file system and loss of data. One or more nodes may be terminated to prevent any further damage to the file system.

**User response:** Unmount the file system and run the `kwdfmfsck` command to repair the file system.

---

**6027-2675**    **Only file systems with NFSv4 ACL semantics enabled can be mounted on this platform.**

**Explanation:** A user is trying to mount a file system on Microsoft Windows, but the ACL semantics disallow NFSv4 ACLs.

**User response:** Enable NFSv4 ACL semantics using the `mmchfs -k` option).

---

**6027-2676**    **Only file systems with NFSv4 locking semantics enabled can be mounted on this platform.**

**Explanation:** A user is trying to mount a file system on Microsoft Windows, but the POSIX locking semantics are in effect.

**User response:** Enable NFSv4 locking semantics using the `mmchfs -D` option).

---

**6027-2677**    **Fileset *filesetName* has pending changes that need to be synced.**

**Explanation:** A user is trying to change a caching option for a fileset while it has local changes that are not yet synced with the home server.

**User response:** Perform AFM recovery before reissuing the command.

---

**6027-2678**    **File system *fileSystem* is mounted on nodes *nodes* or fileset *filesetName* is not unlinked.**

**Explanation:** A user is trying to change a caching feature for a fileset while the file system is still mounted or the fileset is still linked.

**User response:** Unmount the file system from all nodes or unlink the fileset before reissuing the command.

---

**6027-2679**    **Mount of *fileSystem* failed because mount event not handled by any data management application.**

**Explanation:** The mount failed because the file system is enabled for DMAPI events (`-z yes`), but there was no data management application running to handle the event.

**User response:** Make sure the DM application (for example HSM or HPSS) is running before the file system is mounted.

---

**6027-2680**    **AFM filesets cannot be created for file system *fileSystem*.**

**Explanation:** The current file system format version does not support AFM-enabled filesets; the `-p` option cannot be used.

**User response:** Change the file system format version by issuing `mmchfs -V`.

---

**6027-2681**    **Snapshot *snapshotName* has linked independent filesets**

**Explanation:** The specified snapshot is not in a valid state.

**User response:** Correct the problem and reissue the command.

---

**6027-2682 [E]**    **Set quota file attribute error (*reasonCode*)*explanation***

**Explanation:** While mounting a file system a new quota file failed to be created due to inconsistency with the current degree of replication or the number of failure groups.

**User response:** Disable quotas. Check and correct the

degree of replication and the number of failure groups.  
Re-enable quotas.

---

**6027-2683** Fileset *filesetName* in file system *fileSystem* does not contain snapshot *snapshotName*, err = *number*

**Explanation:** An incorrect snapshot name was specified.

**User response:** Select a valid snapshot and issue the command again.

---

**6027-2684** File system *fileSystem* does not contain global snapshot *snapshotName*, err = *number*

**Explanation:** An incorrect snapshot name was specified.

**User response:** Select a valid snapshot and issue the command again.

---

**6027-2685** Total file system capacity allows *minMaxInodes* inodes in *fileSystem*. Currently the total inode limits used by all the inode spaces in *inodeSpace* is *inodeSpaceLimit*. There must be at least *number* inodes available to create a new inode space. Use the `mmlsfileset -L` command to show the maximum inode limits of each fileset. Try reducing the maximum inode limits for some of the inode spaces in *fileSystem*.

**Explanation:** The number of inodes available is too small to create a new inode space.

**User response:** Reduce the maximum inode limits and issue the command again.

---

**6027-2688** Only independent filesets can be configured as AFM filesets. The `--inode-space=new` option is required.

**Explanation:** Only independent filesets can be configured for caching.

**User response:** Specify the `--inode-space=new` option.

---

**6027-2689** The value for `--block-size` must be the keyword `auto` or the value must be of the form `[n]K`, `[n]M`, `[n]G` or `[n]T`, where `n` is an optional integer in the range 1 to 1023.

**Explanation:** An invalid value was specified with the `--block-size` option.

**User response:** Reissue the command with a valid option.

---

**6027-2690** Fileset *filesetName* can only be linked within its own inode space.

**Explanation:** A dependent fileset can only be linked within its own inode space.

**User response:** Correct the junction path and reissue the command.

---

**6027-2691** The `fastea` feature needs to be enabled for file system *fileSystem* before creating AFM filesets.

**Explanation:** The current file system on-disk format does not support storing of extended attributes in the file's inode. This is required for AFM-enabled filesets.

**User response:** Use the `mmmigratefs` command to enable the fast extended-attributes feature.

---

**6027-2692** Error encountered while processing the input file.

**Explanation:** The `tscrsnapshot` command encountered an error while processing the input file.

**User response:** Check and validate the fileset names listed in the input file.

---

**6027-2693** Fileset junction name *junctionName* conflicts with the current setting of `mmsnapdir`.

**Explanation:** The fileset junction name conflicts with the current setting of `mmsnapdir`.

**User response:** Select a new junction name or a new directory for the link and reissue the `mmlinkfileset` command.

---

**6027-2694 [I]** The requested maximum number of inodes is already at *number*.

**Explanation:** The specified number of nodes is already in effect.

**User response:** This is an informational message.

---

**6027-2695 [E]** The number of inodes to preallocate cannot be higher than the maximum number of inodes.

**Explanation:** The specified number of nodes to preallocate is not valid.

**User response:** Correct the `--inode-limit` argument then retry the command.

---

**6027-2696 [E]** The number of inodes to preallocate cannot be lower than the *number* inodes already allocated.

**Explanation:** The specified number of nodes to preallocate is not valid.

**User response:** Correct the `--inode-limit` argument then retry the command.

---

**6027-2697** Fileset at *junctionPath* has pending changes that need to be synced.

**Explanation:** A user is trying to change a caching option for a fileset while it has local changes that are not yet synced with the home server.

**User response:** Perform AFM recovery before reissuing the command.

---

**6027-2698** File system *fileSystem* is mounted on nodes *nodes* or fileset at *junctionPath* is not unlinked.

**Explanation:** A user is trying to change a caching feature for a fileset while the file system is still mounted or the fileset is still linked.

**User response:** Unmount the file system from all nodes or unlink the fileset before reissuing the command.

---

**6027-2699** Cannot create a new independent fileset until an existing one is deleted. File system *fileSystem* has a limit of *maxNumber* independent filesets.

**Explanation:** An attempt to create an independent fileset for the cited file system failed because it would exceed the cited limit.

**User response:** Remove unneeded independent filesets and reissue the command.

---

**6027-2700 [E]** A node join was rejected. This could be due to incompatible daemon versions, failure to find the node in the configuration database, or no configuration manager found.

**Explanation:** A request to join nodes was explicitly rejected.

**User response:** Verify that compatible versions of GPFS are installed on all nodes. Also, verify that the joining node is in the configuration database.

---

**6027-2701** The `mmpmon` command file is empty.

**Explanation:** The `mmpmon` command file is empty.

**User response:** Check file size, existence, and access permissions.

---



---

**6027-2702** Unexpected `mmpmon` response from file system daemon.

**Explanation:** An unexpected response was received to an `mmpmon` request.

**User response:** Ensure that the `mmfsd` daemon is running. Check the error log. Ensure that all GPFS software components are at the same version.

---

**6027-2703** Unknown `mmpmon` command *command*.

**Explanation:** An unknown `mmpmon` command was read from the input file.

**User response:** Correct the command and rerun.

---

**6027-2704** Permission failure. The command requires root authority to execute.

**Explanation:** The `mmpmon` command was issued with a nonzero UID.

**User response:** Log on as root and reissue the command.

---

**6027-2705** Could not establish connection to file system daemon.

**Explanation:** The connection between a GPFS command and the `mmfsd` daemon could not be established. The daemon may have crashed, or never been started, or (for `mmpmon`) the allowed number of simultaneous connections has been exceeded.

**User response:** Ensure that the `mmfsd` daemon is running. Check the error log. For `mmpmon`, ensure that the allowed number of simultaneous connections has not been exceeded.

---

**6027-2706 [I]** Recovered *number* nodes.

**Explanation:** The asynchronous part (phase 2) of node failure recovery has completed.

**User response:** None. Informational message only.

---

**6027-2707 [I]** Node join protocol waiting *value* seconds for node recovery

**Explanation:** Node join protocol is delayed until phase 2 of previous node failure recovery protocol is complete.

**User response:** None. Informational message only.

---

**6027-2708 [E]** Rejected node join protocol. Phase two of node failure recovery appears to still be in progress.

**Explanation:** Node join protocol is rejected after a number of internal delays and phase two node failure protocol is still in progress.

---

**User response:** None. Informational message only.

---

**6027-2709 Configuration manager node *nodeName* not found in the node list.**

**Explanation:** The specified node was not found in the node list.

**User response:** Add the specified node to the node list and reissue the command.

---

**6027-2710 [E] Node *nodeName* is being expelled due to expired lease.**

**Explanation:** The nodes listed did not renew their lease in a timely fashion and will be expelled from the cluster.

**User response:** Check the network connection between this node and the node specified above.

---

**6027-2711 [E] File system table full.**

**Explanation:** The `mmfsd` daemon cannot add any more file systems to the table because it is full.

**User response:** None. Informational message only.

---

**6027-2712 Option '*optionName*' has been deprecated.**

**Explanation:** The option that was specified with the command is no longer supported. A warning message is generated to indicate that the option has no effect.

**User response:** Correct the command line and then reissue the command.

---

**6027-2713 Permission failure. The command requires *SuperuserName* authority to execute.**

**Explanation:** The command, or the specified command option, requires administrative authority.

**User response:** Log on as a user with administrative privileges and reissue the command.

---

**6027-2714 Could not appoint node *nodeName* as cluster manager. *errorString***

**Explanation:** The `mmchmgr -c` command generates this message if the specified node cannot be appointed as a new cluster manager.

**User response:** Make sure that the specified node is a quorum node and that GPFS is running on that node.

---

**6027-2715 Could not appoint a new cluster manager. *errorString***

**Explanation:** The `mmchmgr -c` command generates this message when a node is not available as a cluster manager.

**User response:** Make sure that GPFS is running on a sufficient number of quorum nodes.

---

**6027-2716 [I] Challenge response received; canceling disk election.**

**Explanation:** The node has challenged another node, which won the previous election, and detected a response to the challenge.

**User response:** None. Informational message only.

---

**6027-2717 Node *nodeName* is already a cluster manager or another node is taking over as the cluster manager.**

**Explanation:** The `mmchmgr -c` command generates this message if the specified node is already the cluster manager.

**User response:** None. Informational message only.

---

**6027-2718 Incorrect port range: GPFS\_CMDPORT\_RANGE=*range*'. Using default.**

**Explanation:** The GPFS command port range format is `lllll[-hhhhh]`, where `lllll` is the low port value and `hhhhh` is the high port value. The valid range is 1 to 65535.

**User response:** None. Informational message only.

---

**6027-2719 The files provided do not contain valid quota entries.**

**Explanation:** The quota file provided does not have valid quota entries.

**User response:** Check that the file being restored is a valid GPFS quota file.

---

**6027-2722 [E] Node limit of *number* has been reached. Ignoring *nodeName*.**

**Explanation:** The number of nodes that have been added to the cluster is greater than some cluster members can handle.

**User response:** Delete some nodes from the cluster using the `mmdelnode` command, or shut down GPFS on nodes that are running older versions of the code with lower limits.

---

**6027-2723 [N] This node (*nodeName*) is now Cluster Manager for *clusterName*.**

**Explanation:** This is an informational message when a new cluster manager takes over.

**User response:** None. Informational message only.

---

**6027-2724 [I] *reasonString*. Probing cluster *clusterName***

**Explanation:** This is an informational message when a lease request has not been renewed.

**User response:** None. Informational message only.

---

**6027-2725 [N] Node *nodeName* lease renewal is overdue. Pinging to check if it is alive**

**Explanation:** This is an informational message on the cluster manager when a lease request has not been renewed.

**User response:** None. Informational message only.

---

**6027-2726 [I] Recovered *number* nodes for file system *fileSystem*.**

**Explanation:** The asynchronous part (phase 2) of node failure recovery has completed.

**User response:** None. Informational message only.

---

**6027-2727 *fileSystem*: quota manager is not available.**

**Explanation:** An attempt was made to perform a quota command without a quota manager running. This could be caused by a conflicting offline **mmfsck** command.

**User response:** Reissue the command once the conflicting program has ended.

---

**6027-2728 [N] Connection from *node* rejected because it does not support IPv6**

**Explanation:** A connection request was received from a node that does not support Internet Protocol Version 6 (IPv6), and at least one node in the cluster is configured with an IPv6 address (not an IPv4-mapped one) as its primary address. Since the connecting node will not be able to communicate with the IPv6 node, it is not permitted to join the cluster.

**User response:** Upgrade the connecting node to a version of GPFS that supports IPv6, or delete all nodes with IPv6-only addresses from the cluster.

---



---

**6027-2729 Value *value* for option *optionName* is out of range. Valid values are *value* through *value*.**

**Explanation:** An out of range value was specified for the specified option.

**User response:** Correct the command line.

---

**6027-2730 [E] Node *nodeName* failed to take over as cluster manager.**

**Explanation:** An attempt to takeover as cluster manager failed.

**User response:** Make sure that GPFS is running on a sufficient number of quorum nodes.

---

**6027-2731 Failed to locate a working cluster manager.**

**Explanation:** The cluster manager has failed or changed. The new cluster manager has not been appointed.

**User response:** Check the internode communication configuration and ensure enough GPFS nodes are up to make a quorum.

---

**6027-2732 Attention: No data disks remain in the system pool. Use **mmapplypolicy** to migrate all data left in the system pool to other storage pool.**

**Explanation:** The **mmchdisk** command has been issued but no data disks remain in the system pool. Warn user to use **mmapplypolicy** to move data to other storage pool.

**User response:** None. Informational message only.

---

**6027-2733 The file system name (*fsname*) is longer than the maximum allowable length (*maxLength*).**

**Explanation:** The file system name is invalid because it is longer than the maximum allowed length of 255 characters.

**User response:** Specify a file system name whose length is 255 characters or less and reissue the command.

---

**6027-2734 [E] Disk failure from node *nodeName* Volume *name*. Physical volume *name*.**

**Explanation:** An I/O request to a disk or a request to fence a disk has failed in such a manner that GPFS can no longer use the disk.

**User response:** Check the disk hardware and the software subsystems in the path to the disk.

---



---

**6027-2735 [E] Not a manager**

**Explanation:** This node is not a manager or no longer a manager of the type required to proceed with the operation. This could be caused by the change of manager in the middle of the operation.

**User response:** Retry the operation.

---

**6027-2736** The value for `--block-size` must be the keyword `auto` or the value must be of the form `nK`, `nM`, `nG` or `nT`, where `n` is an optional integer in the range 1 to 1023.

**Explanation:** An invalid value was specified with the `--block-size` option.

**User response:** Reissue the command with a valid option.

---

**6027-2738** Editing quota limits for the root user is not permitted

**Explanation:** The `root` user was specified for quota limits editing in the `mmedquota` command.

**User response:** Specify a valid user or group in the `mmedquota` command. Editing quota limits for the `root` user or `system` group is prohibited.

---

**6027-2739** Editing quota limits for `groupName` group not permitted.

**Explanation:** The `system` group was specified for quota limits editing in the `mmedquota` command.

**User response:** Specify a valid user or group in the `mmedquota` command. Editing quota limits for the `root` user or `system` group is prohibited.

---

**6027-2740 [I]** Starting new election as previous `clmgr` is expelled

**Explanation:** This node is taking over as `clmgr` without challenge as the old `clmgr` is being expelled.

**User response:** None. Informational message only.

---

**6027-2741 [W]** This node can not continue to be cluster manager

**Explanation:** This node invoked the user-specified callback handler for event `tiebreakerCheck` and it returned a non-zero value. This node cannot continue to be the cluster manager.

**User response:** None. Informational message only.

---

**6027-2742 [I]** `CallExitScript`: exit script `exitScript` on event `eventName` returned code `returnCode`, `quorumloss`.

**Explanation:** This node invoked the user-specified callback handler for the `tiebreakerCheck` event and it returned a non-zero value. The user-specified action with the error is `quorumloss`.

**User response:** None. Informational message only.

---

**6027-2743** Permission denied.

**Explanation:** The command is invoked by an unauthorized user.

**User response:** Retry the command with an authorized user.

---

**6027-2744 [D]** Invoking `tiebreaker` callback script

**Explanation:** The node is invoking the callback script due to change in quorum membership.

**User response:** None. Informational message only.

---

**6027-2745 [E]** File system is not mounted.

**Explanation:** A command was issued, which requires that the file system be mounted.

**User response:** Mount the file system and reissue the command.

---

**6027-2746 [E]** Too many disks unavailable for this server to continue serving a `RecoveryGroup`.

**Explanation:** `RecoveryGroup` panic: Too many disks unavailable to continue serving this `RecoveryGroup`. This server will resign, and failover to an alternate server will be attempted.

**User response:** Ensure the alternate server took over. Determine what caused this event and address the situation. Prior messages may help determine the cause of the event.

---

**6027-2747 [E]** Inconsistency detected between the local node number retrieved from '`mmsdrfs`' (`nodeNumber`) and the node number retrieved from '`mmfs.cfg`' (`nodeNumber`).

**Explanation:** The node number retrieved by obtaining the list of nodes in the `mmsdrfs` file did not match the node number contained in `mmfs.cfg`. There may have been a recent change in the IP addresses being used by network interfaces configured at the node.

**User response:** Stop and restart GPFS daemon.

---

---

**6027-2748** Terminating because a conflicting program on the same inode space *inodeSpace* is running.

**Explanation:** A program detected that it must terminate because a conflicting program is running.

**User response:** Reissue the command after the conflicting program ends.

---

**6027-2749** Specified locality group '*number*' does not match disk '*name*' locality group '*number*'. To change locality groups in an SNC environment, please use the **mmdeldisk** and **mmadddisk** commands.

**Explanation:** The locality group specified on the **mmchdisk** command does not match the current locality group of the disk.

**User response:** To change locality groups in an SNC environment, use the **mmdeldisk** and **mmaddisk** commands.

---

**6027-2750 [I]** Node *nodeName* is now the Group Leader.

**Explanation:** A new cluster Group Leader has been assigned.

**User response:** None. Informational message only.

---

**6027-2751 [I]** Starting new election: Last elected: *NodeNumber* Sequence: *SequenceNumber*

**Explanation:** A new disk election will be started. The disk challenge will be skipped since the last elected node was either none or the local node.

**User response:** None. Informational message only.

---

**6027-2752 [I]** This node got elected. Sequence: *SequenceNumber*

**Explanation:** Local node got elected in the disk election. This node will become the cluster manager.

**User response:** None. Informational message only.

---

**6027-2753 [N]** Responding to disk challenge: response: *ResponseValue*. Error code: *ErrorCode*.

**Explanation:** A disk challenge has been received, indicating that another node is attempting to become a Cluster Manager. Issuing a challenge response, to confirm the local node is still alive and will remain the Cluster Manager.

**User response:** None. Informational message only.

---

**6027-2754 [X]** Challenge thread did not respond to challenge in time: took *TimeIntervalSecs* seconds.

**Explanation:** Challenge thread took too long to respond to a disk challenge. Challenge thread will exit, which will result in the local node losing quorum.

**User response:** None. Informational message only.

---

**6027-2755 [N]** Another node committed disk election with sequence *CommittedSequenceNumber* (our sequence was *OurSequenceNumber*).

**Explanation:** Another node committed a disk election with a sequence number higher than the one used when this node used to commit an election in the past. This means that the other node has become, or is becoming, a Cluster Manager. To avoid having two Cluster Managers, this node will lose quorum.

**User response:** None. Informational message only.

---

**6027-2756** Attention: In file system *FileSystemName*, *FileSetName* (Default) *QuotaLimitType*(*QuotaLimit*) for *QuotaTypeUser/GroupName/FilesetName* is too small. Suggest setting it higher than *minQuotaLimit*.

**Explanation:** Users set too low quota limits. It will cause unexpected quota behavior. *MinQuotaLimit* is computed through:

1. for block:  $QUOTA\_THRESHOLD * MIN\_SHARE\_BLOCKS * subblocksize$
2. for inode:  $QUOTA\_THRESHOLD * MIN\_SHARE\_INODES$

**User response:** Users should reset quota limits so that they are more than *MinQuotaLimit*. It is just a warning. Quota limits will be set anyway.

---

**6027-2757 [E]** The peer snapshot is in progress. Queue cannot be flushed now.

**Explanation:** The Peer Snapshot is in progress. Queue cannot be flushed now.

**User response:** Reissue the command once the peer snapshot has ended.

---

**6027-2758 [E]** The AFM target does not support this operation. Run **mmafmconfig** on the AFM target cluster.

**Explanation:** The **.afmctl** file is probably not present on the AFM target cluster.

**User response:** Run **mmafmconfig** on the AFM target cluster to configure the AFM target cluster.

---

---

**6027-2759 [N] Disk lease period expired in cluster**  
*ClusterName*. Attempting to reacquire lease.

**Explanation:** The disk lease period expired, which will prevent the local node from being able to perform disk I/O. This can be caused by a temporary communication outage.

**User response:** If message is repeated then the communication outage should be investigated.

---

**6027-2760 [N] Disk lease reacquired in cluster**  
*ClusterName*.

**Explanation:** The disk lease has been reacquired, and disk I/O will be resumed.

**User response:** None. Informational message only.

---

**6027-2761 Unable to run command on 'fileSystem' while the file system is mounted in restricted mode.**

**Explanation:** A command that can alter data in a file system was issued while the file system was mounted in restricted mode.

**User response:** Mount the file system in read-only or read-write mode or unmount the file system and then reissue the command.

---

**6027-2762 Unable to run command on 'fileSystem' while the file system is suspended.**

**Explanation:** A command that can alter data in a file system was issued while the file system was suspended.

**User response:** Resume the file system and reissue the command.

---

**6027-2763 Unable to start command on 'fileSystem' because conflicting program *name* is running. Waiting until it completes.**

**Explanation:** A program detected that it cannot start because a conflicting program is running. The program will automatically start once the conflicting program has ended as long as there are no other conflicting programs running at that time.

**User response:** None. Informational message only.

---

**6027-2764 Terminating command on *fileSystem* because a conflicting program *name* is running.**

**Explanation:** A program detected that it must terminate because a conflicting program is running.

**User response:** Reissue the command after the conflicting program ends.

---

**6027-2765 command on 'fileSystem' is finished waiting. Processing continues ... *name***

**Explanation:** A program detected that it can now continue the processing since a conflicting program has ended.

**User response:** None. Informational message only.

---

**6027-2766 [I] User script has chosen to expel node *nodeName* instead of node *nodeName*.**

**Explanation:** User has specified a callback script that is invoked whenever a decision is about to be taken on what node should be expelled from the active cluster. As a result of the execution of the script, GPFS will reverse its decision on what node to expel.

**User response:** None.

---

**6027-2767 [E] Error *errorNumber* while accessing tiebreaker devices.**

**Explanation:** An error was encountered while reading from or writing to the tiebreaker devices. When such error happens while the cluster manager is checking for challenges, it will cause the cluster manager to lose cluster membership.

**User response:** Verify the health of the tiebreaker devices.

---

**6027-2770 Disk *diskName* belongs to a write-affinity enabled storage pool. Its failure group cannot be changed.**

**Explanation:** The failure group specified on the **mmchdisk** command does not match the current failure group of the disk.

**User response:** Use the **mmdeldisk** and **mmaddisk** commands to change failure groups in a write-affinity enabled storage pool.

---

**6027-2771 *fileSystem*: Default per-fileset quotas are disabled for *quotaType*.**

**Explanation:** A command was issued to modify default fileset-level quota, but default quotas are not enabled.

**User response:** Ensure the **--perfileset-quota** option is in effect for the file system, then use the **mmdefquotaon** command to enable default fileset-level quotas. After default quotas are enabled, issue the failed command again.

---

**6027-2772 Cannot close disk *name*.**

**Explanation:** Could not access the specified disk.

**User response:** Check the disk hardware and the path

---

to the disk. Refer to “Unable to access disks” on page 279.

---

**6027-2773** *fileSystem:filesetName:* **default quota for quotaType is disabled.**

**Explanation:** A command was issued to modify default quota, but default quota is not enabled.

**User response:** Ensure the **-Q yes** option is in effect for the file system, then enable default quota with the **mmdefquotaon** command.

---

**6027-2774** *fileSystem:* **Per-fileset quotas are not enabled.**

**Explanation:** A command was issued to modify fileset-level quota, but per-fileset quota management is not enabled.

**User response:** Ensure that the **--perfileset-quota** option is in effect for the file system and reissue the command.

---

**6027-2775** **Storage pool named *poolName* does not exist.**

**Explanation:** The **mmlspool** command was issued, but the specified storage pool does not exist.

**User response:** Correct the input and reissue the command.

---

**6027-2776** **Attention: A disk being stopped reduces the degree of system metadata replication (*value*) or data replication (*value*) to lower than tolerable.**

**Explanation:** The **mmchdisk stop** command was issued, but the disk cannot be stopped because of the current file system metadata and data replication factors.

**User response:** Make more disks available, delete unavailable disks, or change the file system metadata replication factor. Also check the current value of the **unmountOnDiskFail** configuration parameter.

---

**6027-2777 [E]** **Node *nodeName* is being expelled because of an expired lease. Pings sent: *pingsSent*. Replies received: *pingRepliesReceived*.**

**Explanation:** The node listed did not renew its lease in a timely fashion and is being expelled from the cluster.

**User response:** Check the network connection between this node and the node listed in the message.

---

**6027-2778 [I]** **Node *nodeName*: ping timed out. Pings sent: *pingsSent*. Replies received: *pingRepliesReceived*.**

**Explanation:** Ping timed out for the node listed, which should be the cluster manager. A new cluster manager will be chosen while the current cluster manager is expelled from the cluster.

**User response:** Check the network connection between this node and the node listed in the message.

---

**6027-2779 [E]** **Challenge thread stopped.**

**Explanation:** A tiebreaker challenge thread stopped because of an error. Cluster membership will be lost.

**User response:** Check for additional error messages. File systems will be unmounted, then the node will rejoin the cluster.

---

**6027-2780 [E]** **Not enough quorum nodes reachable: *reachableNodes*.**

**Explanation:** The cluster manager cannot reach a sufficient number of quorum nodes, and therefore must resign to prevent cluster partitioning.

**User response:** Determine if there is a network outage or if too many nodes have failed.

---

**6027-2781 [E]** **Lease expired for *numSecs* seconds (shutdownOnLeaseExpiry).**

**Explanation:** Disk lease expired for too long, which results in the node losing cluster membership.

**User response:** None. The node will attempt to rejoin the cluster.

---

**6027-2782 [E]** **This node is being expelled from the cluster.**

**Explanation:** This node received a message instructing it to leave the cluster, which might indicate communication problems between this node and some other node in the cluster.

**User response:** None. The node will attempt to rejoin the cluster.

---

**6027-2783 [E]** **New leader elected with a higher ballot number.**

**Explanation:** A new group leader was elected with a higher ballot number, and this node is no longer the leader. Therefore, this node must leave the cluster and rejoin.

**User response:** None. The node will attempt to rejoin the cluster.

---

**6027-2784 [E] No longer a cluster manager or lost quorum while running a group protocol.**

**Explanation:** Cluster manager no longer maintains quorum after attempting to run a group protocol, which might indicate a network outage or node failures.

**User response:** None. The node will attempt to rejoin the cluster.

---

**6027-2785 [X] A severe error was encountered during cluster probe.**

**Explanation:** A severe error was encountered while running the cluster probe to determine the state of the nodes in the cluster.

**User response:** Examine additional error messages. The node will attempt to rejoin the cluster.

---

**6027-2786 [E] Unable to contact any quorum nodes during cluster probe.**

**Explanation:** This node has been unable to contact any quorum nodes during cluster probe, which might indicate a network outage or too many quorum node failures.

**User response:** Determine whether there was a network outage or whether quorum nodes failed.

---

**6027-2787 [E] Unable to contact enough other quorum nodes during cluster probe.**

**Explanation:** This node, a quorum node, was unable to contact a sufficient number of quorum nodes during cluster probe, which might indicate a network outage or too many quorum node failures.

**User response:** Determine whether there was a network outage or whether quorum nodes failed.

---

**6027-2788 [E] Attempt to run leader election failed with error *errorNumber*.**

**Explanation:** This node attempted to run a group leader election but failed to get elected. This failure might indicate that two or more quorum nodes attempted to run the election at the same time. As a result, this node will lose cluster membership and then attempt to rejoin the cluster.

**User response:** None. The node will attempt to rejoin the cluster.

---

**6027-2789 [E] Tiebreaker script returned a non-zero value.**

**Explanation:** The tiebreaker script, invoked during group leader election, returned a non-zero value, which results in the node losing cluster membership and then attempting to rejoin the cluster.

**User response:** None. The node will attempt to rejoin the cluster.

---

**6027-2790 Attention: Disk parameters were changed. Use the `mmrestripefs` command with the `-r` option to relocate data and metadata.**

**Explanation:** The `mmchdisk` command with the `change` option was issued.

**User response:** Issue the `mmrestripefs -r` command to relocate data and metadata.

---

**6027-2791 Disk *diskName* does not belong to file system *deviceName*.**

**Explanation:** The input disk name does not belong to the specified file system.

**User response:** Correct the command line.

---

**6027-2792 The current file system version does not support default per-fileset quotas.**

**Explanation:** The current version of the file system does not support default fileset-level quotas.

**User response:** Use the `mmchfs -V` command to activate the new function.

---

**6027-2793 [E] Contents of local *fileName* file are invalid. Node may be unable to be elected group leader.**

**Explanation:** In an environment where tie-breaker disks are used, the contents of the ballot file have become invalid, possibly because the file has been overwritten by another application. This node will be unable to be elected group leader.

**User response:** Run `mmcommon resetTiebreaker`, which will ensure the GPFS daemon is down on all quorum nodes and then remove the given file on this node. After that, restart the cluster on this and on the other nodes.

---

**6027-2794 [E] Invalid content of disk paxos sector for disk *diskName*.**

**Explanation:** In an environment where tie-breaker disks are used, the contents of either one of the tie-breaker disks or the ballot files became invalid, possibly because the file has been overwritten by another application.

**User response:** Examine `mmfs.log` file on all quorum nodes for indication of a corrupted ballot file. If 6027-2793 is found then follow instructions for that message. If problem cannot be resolved, shut down GPFS across the cluster, undefine, and then redefine the `tiebreakerdisks` configuration variable, and finally restart the cluster.

---

**6027-2795** An error occurred while executing command for *fileSystem*.

**Explanation:** A quota command encountered a problem on a file system. Processing continues with the next file system.

**User response:** None. Informational message only.

---

**6027-2796 [W]** Callback event *eventName* is not supported on this node; processing continues ...

**Explanation:** informational

**User response:**

---

**6027-2797 [I]** Node *nodeName*: lease request received late. Pings sent: *pingsSent*. Maximum pings missed: *maxPingsMissed*.

**Explanation:** The cluster manager reports that the lease request from the given node was received late, possibly indicating a network outage.

**User response:** Check the network connection between this node and the node listed in the message.

---

**6027-2798 [E]** The node *nodeName* does not have a valid Extended License to run the requested command.

**Explanation:** The file system manager node does not have a valid extended license to run ILM, AFM, or CNFS commands.

**User response:** Make sure *gpfs.ext* package is installed correctly on file system manager node and try again.

---

**6027-2799** Option '*option*' is incompatible with option '*option*'.

**Explanation:** The options specified on the command are incompatible.

**User response:** Do not specify these two options together.

---

**6027-2800** Available memory exceeded on request to allocate *number* bytes. Trace point *sourceFile-tracePoint*.

**Explanation:** The available memory was exceeded during an allocation request made from the cited source file and trace point.

**User response:** Try shutting down and then restarting GPFS. If the problem recurs, contact the IBM Support Center.

---

**6027-2801** Policy set syntax version *versionString* not supported.

**Explanation:** The policy rules do not comply with the supported syntax.

**User response:** Rewrite the policy rules, following the documented, supported syntax and keywords.

---

**6027-2802** Object name '*poolName\_or\_filesetName*' is not valid.

**Explanation:** The cited name is not a valid GPFS object, names an object that is not valid in this context, or names an object that no longer exists.

**User response:** Correct the input to identify a GPFS object that exists and is valid in this context.

---

**6027-2803** Policy set must start with VERSION.

**Explanation:** The policy set does not begin with VERSION as required.

**User response:** Rewrite the policy rules, following the documented, supported syntax and keywords.

---

**6027-2804** Unexpected SQL result code - *sqlResultCode*.

**Explanation:** This could be an IBM programming error.

**User response:** Check that your SQL expressions are correct and supported by the current release of GPFS. If the error recurs, contact the IBM Support Center.

---

**6027-2805 [I]** Loaded policy '*policyFileName* or *filesystemName*': *summaryOfPolicyRules*

**Explanation:** The specified loaded policy has the specified policy rules.

**User response:** None. Informational message only.

---

**6027-2806 [E]** Error while validating policy '*policyFileName* or *filesystemName*': *rc=errorCode: errorDetailsString*

**Explanation:** An error occurred while validating the specified policy.

**User response:** Correct the policy rules, heeding the error details in this message and other messages issued immediately before or after this message. Use the **mmchpolicy** command to install a corrected policy rules file.

---

---

**6027-2807 [W] Error in evaluation of placement policy for file *fileName: errorDetailsString***

**Explanation:** An error occurred while evaluating the installed placement policy for a particular new file. Although the policy rules appeared to be syntactically correct when the policy was installed, evidently there is a problem when certain values of file attributes occur at runtime.

**User response:** Determine which file names and attributes trigger this error. Correct the policy rules, heeding the error details in this message and other messages issued immediately before or after this message. Use the **mmchpolicy** command to install a corrected policy rules file.

---

**6027-2808 In rule '*ruleName*' (*ruleNumber*), '*wouldBePoolName*' is not a valid pool name.**

**Explanation:** The cited name that appeared in the cited rule is not a valid pool name. This may be because the cited name was misspelled or removed from the file system.

**User response:** Correct or remove the rule.

---

**6027-2809 Validated policy '*policyFileName* or *filesystemName*': *summaryOfPolicyRules***

**Explanation:** The specified validated policy has the specified policy rules.

**User response:** None. Informational message only.

---

**6027-2810 [W] There are *numberOfPools* storage pools but the policy file is missing or empty.**

**Explanation:** The cited number of storage pools are defined, but the policy file is missing or empty.

**User response:** You should probably install a policy with placement rules using the **mmchpolicy** command, so that at least some of your data will be stored in your nonsystem storage pools.

---

**6027-2811 Policy has no storage pool placement rules!**

**Explanation:** The policy has no storage pool placement rules.

**User response:** You should probably install a policy with placement rules using the **mmchpolicy** command, so that at least some of your data will be stored in your nonsystem storage pools.

---

**6027-2812 Keyword '*keywordValue*' begins a second *clauseName* clause - only one is allowed.**

**Explanation:** The policy rule should only have one clause of the indicated type.

**User response:** Correct the rule and reissue the policy command.

---

**6027-2813 This '*ruleName*' rule is missing a *clauseType* required clause.**

**Explanation:** The policy rule must have a clause of the indicated type.

**User response:** Correct the rule and reissue the policy command.

---

**6027-2814 This '*ruleName*' rule is of unknown type or not supported.**

**Explanation:** The policy rule set seems to have a rule of an unknown type or a rule that is unsupported by the current release of GPFS.

**User response:** Correct the rule and reissue the policy command.

---

**6027-2815 The value '*value*' is not supported in a '*clauseType*' clause.**

**Explanation:** The policy rule clause seems to specify an unsupported argument or value that is not supported by the current release of GPFS.

**User response:** Correct the rule and reissue the policy command.

---

**6027-2816 Policy rules employ features that would require a file system upgrade.**

**Explanation:** One or more policy rules have been written to use new features that cannot be installed on a back-level file system.

**User response:** Install the latest GPFS software on all nodes and upgrade the file system or change your rules. (Note that **LIMIT** was introduced in GPFS Release 3.2.)

---

**6027-2817 Error on **popen/pclose** (*command\_string*): *rc=return\_code\_from\_popen\_or\_pclose***

**Explanation:** The execution of the *command\_string* by **popen/pclose** resulted in an error.

**User response:** To correct the error, do one or more of the following:

Check that the standard **m4** macro processing command is installed on your system as **/usr/bin/m4**.

Or:

Set the **MM\_M4\_CMD** environment variable.

Or:

Correct the macro definitions in your policy rules file.

If the problem persists, contact the IBM Support Center.

**6027-2818**    **A problem occurred during m4 processing of policy rules. rc = *return\_code\_from\_popen\_pclose\_or\_m4***

**Explanation:** An attempt to expand the policy rules with an **m4** subprocess yielded some warnings or errors or the **m4** macro wrote some output to standard error. Details or related messages may follow this message.

**User response:** To correct the error, do one or more of the following:

Check that the standard **m4** macro processing command is installed on your system as **/usr/bin/m4**.

Or:

Set the **MM\_M4\_CMD** environment variable.

Or:

Correct the macro definitions in your policy rules file.

If the problem persists, contact the IBM Support Center.

**6027-2819**    **Error opening temp file *temp\_file\_name: errorString***

**Explanation:** An error occurred while attempting to open the specified temporary work file.

**User response:** Check that the path name is defined and accessible. Check the file and then reissue the command.

**6027-2820**    **Error reading temp file *temp\_file\_name: errorString***

**Explanation:** An error occurred while attempting to read the specified temporary work file.

**User response:** Check that the path name is defined and accessible. Check the file and then reissue the command.

**6027-2821**    **Rule '*ruleName*' (*ruleNumber*) specifies a THRESHOLD for EXTERNAL POOL '*externalPoolName*'. This is not supported.**

**Explanation:** GPFS does not support the **THRESHOLD** clause within a migrate rule that names an external pool in the **FROM POOL** clause.

**User response:** Correct or remove the rule.

**6027-2822**    **This file system does not support fast extended attributes, which are needed for encryption.**

**Explanation:** Fast extended attributes need to be supported by the file system for encryption to be activated.

**User response:** Enable the fast extended attributes feature in this file system.

**6027-2823 [E]**    **Encryption activated in the file system, but node not enabled for encryption.**

**Explanation:** The file system is enabled for encryption, but this node is not.

**User response:** Ensure the GPFS encryption packages are installed. Verify if encryption is supported on this node architecture.

**6027-2824**    **This file system version does not support encryption rules.**

**Explanation:** This file system version does not support encryption.

**User response:** Update the file system to a version which supports encryption.

**6027-2825**    **Duplicate encryption set name '*setName*'.**

**Explanation:** The given set name is duplicated in the policy file.

**User response:** Ensure each set name appears only once in the policy file.

**6027-2826**    **The encryption set '*setName*' requested by rule '*rule*' could not be found.**

**Explanation:** The given set name used in the rule cannot be found.

**User response:** Verify if the set name is correct. Add the given set if it is missing from the policy.

**6027-2827 [E]**    **Error in evaluation of encryption policy for file *fileName: %s***

**Explanation:** An error occurred while evaluating the encryption rules in the given policy file.

**User response:** Examine the other error messages produced while evaluating the policy file.

**6027-2828 [E]**    **Encryption not supported on Windows. Encrypted file systems are not allowed when Windows nodes are present in the cluster.**

**Explanation:** Self-explanatory.

**User response:** To activate encryption, ensure there are



no Windows nodes in the cluster.

---

**6027-2950 [E] Trace value '*value*' after class '*class*' must be from 0 to 14.**

**Explanation:** The specified trace value is not recognized.

**User response:** Specify a valid trace integer value.

---

**6027-2951 [W] Value *value* for worker1Threads must be <= than the original setting *value***

**Explanation:** An attempt to dynamically set **worker1Threads** found the value out of range. The dynamic value must be  $2 \leq \textit{value} \leq$  the original setting when the GPFS daemon was started.

---

**6027-2952 [E] Unknown assert class '*assertClass*'.**

**Explanation:** The assert class is not recognized.

**User response:** Specify a valid assert class.

---

**6027-2953 [E] Non-numeric assert value '*value*' after class '*class*'.**

**Explanation:** The specified assert value is not recognized.

**User response:** Specify a valid assert integer value.

---

**6027-2954 [E] Assert value '*value*' after class '*class*' must be from 0 to 127.**

**Explanation:** The specified assert value is not recognized.

**User response:** Specify a valid assert integer value.

---

**6027-2955 [W] Time-of-day may have jumped back. Late by *delaySeconds* seconds to wake certain threads.**

**Explanation:** Time-of-day may have jumped back, which has resulted in some threads being awakened later than expected. It is also possible that some other factor has caused a delay in waking up the threads.

**User response:** Verify if there is any problem with network time synchronization, or if time-of-day is being incorrectly set.

---

**6027-2956 [E] Invalid crypto engine type (**encryptionCryptoEngineType**): *cryptoEngineType*.**

**Explanation:** The specified value for **encryptionCryptoEngineType** is incorrect.

**User response:** Specify a valid value for **encryptionCryptoEngineType**.

---

**6027-2957 [E] Invalid cluster manager selection choice (**clusterManagerSelection**): *clusterManagerSelection*.**

**Explanation:** The specified value for **clusterManagerSelection** is incorrect.

**User response:** Specify a valid value for **clusterManagerSelection**.

---

**6027-2958 [E] Invalid NIST compliance type (**nistCompliance**): *nistComplianceValue*.**

**Explanation:** The specified value for **nistCompliance** is incorrect.

**User response:** Specify a valid value for **nistCompliance**.

---

**6027-2959 [E] The CPU architecture on this node does not support tracing in *traceMode* mode. Switching to *traceMode* mode.**

**Explanation:** The CPU does not have constant time stamp counter capability required for overwrite trace mode. The trace has been enabled in blocking mode.

**User response:** Update configuration parameters to use trace facility in blocking mode or replace this node with modern CPU architecture.

---

**6027-3101 Pdisk rotation rate invalid in option '*option*'.**

**Explanation:** When parsing disk lists, the pdisk rotation rate is not valid.

**User response:** Specify a valid rotation rate (SSD, NVRAM, or 1025 through 65535).

---

**6027-3102 Pdisk FRU number too long in option '*option*', maximum length *length*.**

**Explanation:** When parsing disk lists, the pdisk FRU number is too long.

**User response:** Specify a valid FRU number that is shorter than or equal to the maximum length.

---

**6027-3103 Pdisk location too long in option '*option*', maximum length *length*.**

**Explanation:** When parsing disk lists, the pdisk location is too long.

**User response:** Specify a valid location that is shorter than or equal to the maximum length.

---

**6027-3105** Pdisk `nPathActive` invalid in option '*option*'.

**Explanation:** When parsing disk lists, the `nPathActive` value is not valid.

**User response:** Specify a valid `nPathActive` value (0 to 255).

---

**6027-3106** Pdisk `nPathTotal` invalid in option '*option*'.

**Explanation:** When parsing disk lists, the `nPathTotal` value is not valid.

**User response:** Specify a valid `nPathTotal` value (0 to 255).

---

**6027-3107** Pdisk `nsdFormatVersion` invalid in option '*name1name2*'.

**Explanation:** The `nsdFormatVersion` that is entered while parsing the disk is invalid.

**User response:** Specify valid `nsdFormatVersion`, 1 or 2.

---

**6027-3108** Declustered array name *name1* appears more than once in the declustered array stanzas.

**Explanation:** when parsing declustered array lists a duplicate name is found.

**User response:** Remove duplicate `MSG_PARSE_DUPNAME` which is not documented.

---

**6027-3200** AFM ERROR: command `pCacheCmd` fileset *filesetName* fileids [*parentId.childId.tParentId.targetId,ReqCmd*] original error *oerr* application error *aerr* remote error *remoteError*

**Explanation:** AFM operations on a particular file failed.

**User response:** For asynchronous operations that are requeued, run the `mmafmctl` command with the `resumeRequeued` option after fixing the problem at the home cluster.

---

**6027-3201** AFM ERROR DETAILS: type: *remoteCmdType* snapshot name *snapshotName* snapshot ID *snapshotId*

**Explanation:** Peer snapshot creation or deletion failed.

**User response:** Fix snapshot creation or deletion error.

---

**6027-3204** AFM: Failed to set `xattr` on inode *inodeNum* error *err*, ignoring.

**Explanation:** Setting extended attributes on an inode failed.

**User response:** None.

---

**6027-3205** AFM: Failed to get `xattrs` for inode *inodeNum*, ignoring.

**Explanation:** Getting extended attributes on an inode failed.

**User response:** None.

---

**6027-3209** Home NFS mount of *host:path* failed with error *err*

**Explanation:** NFS mounting of path from the home cluster failed.

**User response:** Make sure the exported path can be mounted over NFSv3.

---

**6027-3210** Cannot find AFM control file for fileset *filesetName* in the exported file system at home. ACLs and extended attributes will not be synchronized. Sparse files will have zeros written for holes.

**Explanation:** Either home path does not belong to GPFS, or the AFM control file is not present in the exported path.

**User response:** If the exported path belongs to a GPFS file system, run the `mmafmconfig` command with the `enable` option on the export path at home.

---

**6027-3211** Change in home export detected. Caching will be disabled.

**Explanation:** A change in home export was detected or the home path is stale.

**User response:** Ensure the exported path is accessible.

---

**6027-3212** AFM ERROR: Cannot enable AFM for fileset *filesetName* (error *err*)

**Explanation:** AFM was not enabled for the fileset because the root file handle was modified, or the remote path is stale.

**User response:** Ensure the remote export path is accessible for NFS mount.

---

**6027-3213** Cannot find snapshot link directory name for exported file system at home for fileset *filesetName*. Snapshot directory at home will be cached.

**Explanation:** Unable to determine the snapshot

directory at the home cluster.

**User response:** None.

**6027-3214 [E] AFM: Unexpiration of fileset *filesetName* failed with error *err*. Use `mmafmctl` to manually unexpire the fileset.**

**Explanation:** Unexpiration of fileset failed after a home reconnect.

**User response:** Run the `mmafmctl` command with the `unexpire` option on the fileset.

**6027-3215 [W] AFM: Peer snapshot delayed due to long running execution of operation to remote cluster for fileset *filesetName*. Peer snapshot continuing to wait.**

**Explanation:** Peer snapshot command timed out waiting to flush messages.

**User response:** None.

**6027-3216 Fileset *filesetName* encountered an error synchronizing with the remote cluster. Cannot synchronize with the remote cluster until AFM recovery is executed.**

**Explanation:** Cache failed to synchronize with home because of an out of memory or conflict error. Recovery, resynchronization, or both will be performed by GPFS to synchronize cache with the home.

**User response:** None.

**6027-3217 AFM ERROR Unable to unmount NFS export for fileset *filesetName***

**Explanation:** NFS unmount of the path failed.

**User response:** None.

**6027-3220 AFM: Home NFS mount of *host:path* failed with error *err* for file system *fileSystem* fileset id *filesetName*. Caching will be disabled and the mount will be tried again after *mountRetryTime* seconds, on next request to gateway**

**Explanation:** NFS mount of the home cluster failed. The mount will be tried again after *mountRetryTime* seconds.

**User response:** Make sure the exported path can be mounted over NFSv3.

**6027-3221 AFM: Home NFS mount of *host:path* succeeded for file system *fileSystem* fileset *filesetName*. Caching is enabled.**

**Explanation:** NFS mount of the path from the home cluster succeeded. Caching is enabled.

**User response:** None.

**6027-3224 [I] AFM: Failed to set extended attributes on file system *fileSystem* inode *inodeNum* error *err*, ignoring.**

**Explanation:** Setting extended attributes on an inode failed.

**User response:** None.

**6027-3225 [I] AFM: Failed to get extended attributes for file system *fileSystem* inode *inodeNum*, ignoring.**

**Explanation:** Getting extended attributes on an inode failed.

**User response:** None.

**6027-3226 [I] AFM: Cannot find control file for file system *fileSystem* fileset *filesetName* in the exported file system at home. ACLs and extended attributes will not be synchronized. Sparse files will have zeros written for holes.**

**Explanation:** Either the home path does not belong to GPFS, or the AFM control file is not present in the exported path.

**User response:** If the exported path belongs to a GPFS file system, run the `mmafmconfig` command with the `enable` option on the export path at home.

**6027-3227 [E] AFM: Cannot enable AFM for file system *fileSystem* fileset *filesetName* (error *err*)**

**Explanation:** AFM was not enabled for the fileset because the root file handle was modified, or the remote path is stale.

**User response:** Ensure the remote export path is accessible for NFS mount.

**6027-3228 [E] AFM: Unable to unmount NFS export for file system *fileSystem* fileset *filesetName***

**Explanation:** NFS unmount of the path failed.

**User response:** None.

**6027-3229 [E] AFM: File system *fileSystem* fileset *filesetName* encountered an error synchronizing with the remote cluster. Cannot synchronize with the remote cluster until AFM recovery is executed.**

**Explanation:** The cache failed to synchronize with home because of an out of memory or conflict error. Recovery, resynchronization, or both will be performed

by GPFS to synchronize the cache with the home.

**User response:** None.

**6027-3230 [I] AFM: Cannot find snapshot link directory name for exported file system at home for file system *fileSystem* fileset *filesetName*. Snapshot directory at home will be cached.**

**Explanation:** Unable to determine the snapshot directory at the home cluster.

**User response:** None.

**6027-3232** *type* **AFM: pCacheCmd file system *fileSystem* fileset *filesetName* file IDs [*parentId.childId.tParentId.targetId.flag*] name *sourceName* origin error *err***

**Explanation:** AFM operations on a particular file failed.

**User response:** For asynchronous operations that are requeued, run the **mmafmctl** command with the **resumeRequeued** option after fixing the problem at the home cluster.

**6027-3233 [I] AFM: Previous error repeated *repeatNum* times.**

**Explanation:** Multiple AFM operations have failed.

**User response:** None.

**6027-3234 [E] AFM: Unable to start thread to unexpire filesets.**

**Explanation:** Failed to start thread for unexpiration of fileset.

**User response:** None.

**6027-3235 [I] AFM: Stopping recovery for the file system *fileSystem* fileset *filesetName***

**Explanation:** AFM recovery terminated because the current node is no longer MDS for the fileset.

**User response:** None.

**6027-3236 [E] AFM: Recovery on file system *fileSystem* fileset *filesetName* failed with error *err*. Recovery will be retried on next access after recovery retry interval (*timeout* seconds) or manually resolve known problems and recover the fileset.**

**Explanation:** Recovery failed to complete on the fileset. The next access will restart recovery.

**Explanation:** AFM recovery failed. Fileset will be temporarily put into dropped state and will be recovered on accessing fileset after timeout mentioned

in the error message. User can recover the fileset manually by running **mmafmctl** command with **recover** option after rectifying any known errors leading to failure.

**User response:** None.

**6027-3239 [E] AFM: Remote command *remoteCmdType* on file system *fileSystem* snapshot *snapshotName* snapshot ID *snapshotId* failed.**

**Explanation:** A failure occurred when creating or deleting a peer snapshot.

**User response:** Examine the error details and retry the operation.

**6027-3240 [E] AFM: pCacheCmd file system *fileSystem* fileset *filesetName* file IDs [*parentId.childId.tParentId.targetId.flag*] error *err***

**Explanation:** Operation failed to execute on home in independent-writer mode.

**User response:** None.

**6027-3241 [I] AFM: GW queue transfer started for file system *fileSystem* fileset *filesetName*. Transferring to *nodeAddress*.**

**Explanation:** An old GW initiated the queue transfer because a new GW node joined the cluster, and the fileset now belongs to the new GW node.

**User response:** None.

**6027-3242 [I] AFM: GW queue transfer started for file system *fileSystem* fileset *filesetName*. Receiving from *nodeAddress*.**

**Explanation:** An old MDS initiated the queue transfer because this node joined the cluster as GW and the fileset now belongs to this node.

**User response:** None.

**6027-3243 [I] AFM: GW queue transfer completed for file system *fileSystem* fileset *filesetName*. error *error***

**Explanation:** A GW queue transfer completed.

**User response:** None.

**6027-3244 [I] AFM: Home mount of *afmTarget* succeeded for file system *fileSystem* fileset *filesetName*. Caching is enabled.**

**Explanation:** A mount of the path from the home cluster succeeded. Caching is enabled.

**User response:** None.

---

**6027-3245 [E] AFM: Home mount of *afmTarget* failed with error *error* for file system *fileSystem* fileset ID *filesetName*. Caching will be disabled and the mount will be tried again after *mountRetryTime* seconds, on the next request to the gateway.**

**Explanation:** A mount of the home cluster failed. The mount will be tried again after *mountRetryTime* seconds.

**User response:** Verify that the *afmTarget* can be mounted using the specified protocol.

---

**6027-3246 [I] AFM: Prefetch recovery started for the file system *fileSystem* fileset *filesetName*.**

**Explanation:** Prefetch recovery started.

**User response:** None.

---

**6027-3247 [I] AFM: Prefetch recovery completed for the file system *fileSystem* fileset *filesetName*. error *error***

**Explanation:** Prefetch recovery completed.

**User response:** None.

---

**6027-3248 [E] AFM: Cannot find the control file for fileset *filesetName* in the exported file system at home. This file is required to operate in primary mode. The fileset will be disabled.**

**Explanation:** Either the home path does not belong to GPFS, or the AFM control file is not present in the exported path.

**User response:** If the exported path belongs to a GPFS file system, run the `mmafmconfig` command with the `enable` option on the export path at home.

---

**6027-3249 [E] AFM: Target for fileset *filesetName* is not a secondary-mode fileset or file system. This is required to operate in primary mode. The fileset will be disabled.**

**Explanation:** The AFM target is not a secondary fileset or file system.

**User response:** The AFM target fileset or file system should be converted to secondary mode.

---

**6027-3250 [E] AFM: Refresh intervals cannot be set for fileset.**

**Explanation:** Refresh intervals are not supported on primary and secondary-mode filesets.

**User response:** None.

---



---

**6027-3252 [I] AFM: Home has been restored for cache *filesetName*. Synchronization with home will be resumed.**

**Explanation:** A change in home export was detected that caused the home to be restored. Synchronization with home will be resumed.

**User response:** None.

---

**6027-3253 [E] AFM: Change in home is detected for cache *filesetName*. Synchronization with home is suspended until the problem is resolved.**

**Explanation:** A change in home export was detected or the home path is stale.

**User response:** Ensure the exported path is accessible.

---

**6027-3254 [W] AFM: Home is taking longer than expected to respond for cache *filesetName*. Synchronization with home is temporarily suspended.**

**Explanation:** A pending message from gateway node to home is taking longer than expected to respond. This could be the result of a network issue or a problem at the home site.

**User response:** Ensure the exported path is accessible.

---

**6027-3255 [E] AFM: Target for fileset *filesetName* is a secondary-mode fileset or file system. Only a primary-mode, read-only or local-update mode fileset can operate on a secondary-mode fileset. The fileset will be disabled.**

**Explanation:** The AFM target is a secondary fileset or file system. Only a primary-mode, read-only, or local-update fileset can operate on a secondary-mode fileset.

**User response:** Use a secondary-mode fileset as the target for the primary-mode, read-only or local-update mode fileset.

---

**6027-3256 [I] AFM: The RPO peer snapshot was missed for file system *fileSystem* fileset *filesetName*.**

**Explanation:** The periodic RPO peer snapshot was not taken in time for the primary fileset.

**User response:** None.

---

**6027-3257 [E] AFM: Unable to start thread to verify primary filesets for RPO.**

**Explanation:** Failed to start thread for verification of primary filesets for RPO.

**User response:** None.

---

**6027-3300**    **Attribute `afmShowHomeSnapshot` cannot be changed for a single-writer fileset.**

**Explanation:** Changing `afmShowHomeSnapshot` is not supported for single-writer filesets.

**User response:** None.

---

**6027-3301**    **Unable to quiesce all nodes; some processes are busy or holding required resources.**

**Explanation:** A timeout occurred on one or more nodes while trying to quiesce the file system during a snapshot command.

**User response:** Check the GPFS log on the file system manager node.

---

**6027-3302**    **Attribute `afmShowHomeSnapshot` cannot be changed for a `afmMode` fileset.**

**Explanation:** Changing `afmShowHomeSnapshot` is not supported for single-writer or independent-writer filesets.

**User response:** None.

---

**6027-3303**    **Cannot restore snapshot; quota management is active for `fileSystem`.**

**Explanation:** File system quota management is still active. The file system must be unmounted when restoring global snapshots.

**User response:** Unmount the file system and reissue the restore command.

---

**6027-3304**    **Attention: Disk space reclaim on `number` of `number` regions in `fileSystem` returned errors.**

**Explanation:** Free disk space reclaims on some regions failed during `tsreclaim` run. Typically this is due to the lack of space reclaim support by the disk controller or operating system. It may also be due to utilities such as `mmdefragfs` or `mmfsck` running concurrently.

**User response:** Verify that the disk controllers and the operating systems in the cluster support thin-provisioning space reclaim. Or, rerun the `mmfsctl reclaimSpace` command after `mmdefragfs` or `mmfsck` completes.

---

**6027-3305**    **AFM Fileset `filesetName` cannot be changed as it is in `beingDeleted` state**

**Explanation:** The user specified a fileset to `tschfileset` that cannot be changed.

**User response:** None. You cannot change the attributes of the root fileset.

---

**6027-3306**    **Fileset cannot be changed because it is unlinked.**

**Explanation:** The fileset cannot be changed when it is unlinked.

**User response:** Link the fileset and then try the operation again.

---

**6027-3307**    **Fileset cannot be changed.**

**Explanation:** Fileset cannot be changed.

**User response:** None.

---

**6027-3308**    **This AFM option cannot be set for a secondary fileset.**

**Explanation:** This AFM option cannot be set for a secondary fileset. The fileset cannot be changed.

**User response:** None.

---

**6027-3309**    **The AFM attribute specified cannot be set for a primary fileset.**

**Explanation:** This AFM option cannot be set for a primary fileset. The fileset cannot be changed.

**User response:** None.

---

**6027-3310**    **A secondary fileset cannot be changed.**

**Explanation:** A secondary fileset cannot be changed.

**User response:** None.

---

**6027-3311**    **A primary fileset cannot be changed.**

**Explanation:** A primary fileset cannot be changed.

**User response:** None.

---

**6027-3312**    **No inode was found matching the criteria.**

**Explanation:** No inode was found matching the criteria.

**User response:** None.

---

**6027-3313**    **File system scan RESTARTED due to resume of all disks being emptied.**

**Explanation:** The parallel inode traversal (PIT) phase is restarted with a file system restrripe.

**User response:** None.

---

**6027-3314** File system scan **RESTARTED** due to new disks to be emptied.

**Explanation:** The file system restripe was restarted after a new disk was suspended.

**User response:** None.

---

**6027-3315** File system scan **CANCELLED** due to new disks to be emptied or resume of all disks being emptied.

**Explanation:** The parallel inode traversal (PIT) phase is cancelled during the file system restripe.

**User response:** None.

---

**6027-3316** Unable to create file system because there is not enough space for the log files. Number of log files: *numberOfLogFiles*. Log file size: *logFileSize*. Change one or more of the following as suggested and try again:

**Explanation:** There is not enough space available to create all the required log files. This can happen when the storage pool is not large enough.

**User response:** Refer to the details given and correct the file system parameters.

---

**6027-3317** Warning: file system is not 4K aligned due to small *reasonString*. Native 4K sector disks cannot be added to this file system unless the disk that is used is **dataOnly** and the data block size is at least 128K.

**Explanation:** The file system is created with a small inode or block size. Native 4K sector disk cannot be added to the file system, unless the disk that is used is **dataOnly** and the data block size is at least 128K.

**User response:** None.

---

**6027-3318** Fileset *filesetName* cannot be deleted as it is in compliant mode and it contains user files.

**Explanation:** An attempt was made to delete a non-empty fileset that is in compliant mode.

**User response:** None.

---

**6027-3319** The AFM attribute *optionName* cannot be set for a primary fileset.

**Explanation:** This AFM option cannot be set for a primary fileset. Hence, the fileset cannot be changed.

**User response:** None.

---

**6027-3320** *commandName*: **indefiniteRetentionProtection** is enabled. File system cannot be deleted.

**Explanation:** Indefinite retention is enabled for the file system so it cannot be deleted.

**User response:** None.

---

**6027-3400** Attention: The file system is at risk. The specified replication factor does not tolerate unavailable metadata disks.

**Explanation:** The default metadata replication was reduced to one while there were unavailable, or stopped, metadata disks. This condition prevents future file system manager takeover.

**User response:** Change the default metadata replication, or delete unavailable disks if possible.

---

**6027-3401** Failure group *value* for disk *diskName* is not valid.

**Explanation:** An explicit failure group must be specified for each disk that belongs to a write affinity enabled storage pool.

**User response:** Specify a valid failure group.

---

**6027-3402 [X]** An unexpected device mapper path *dmDevice (nsdId)* was detected. The new path does not have **Persistent Reserve** enabled. The local access to disk *diskName* will be marked as down.

**Explanation:** A new device mapper path was detected, or a previously failed path was activated after the local device discovery was finished. This path lacks a Persistent Reserve and cannot be used. All device paths must be active at mount time.

**User response:** Check the paths to all disks in the file system. Repair any failed paths to disks then rediscover the local disk access.

---

**6027-3404 [E]** The current file system version does not support write caching.

**Explanation:** The current file system version does not allow the write caching option.

**User response:** Use **mmchfs -V** to convert the file system to version 14.04 (4.1.0.0) or higher and reissue the command.

---

**6027-3405 [E]** Cannot change the rapid repair, *"fileSystemName"* is mounted on *number* node(s).

**Explanation:** Rapid repair can only be changed on unmounted file systems.

---

**User response:** Unmount the file system before running this command.

---

**6027-3406 Error: Cannot add 4K native dataOnly disk *diskName* to non-4K aligned file system unless the file system version is at least 4.1.1.4.**

**Explanation:** An attempt was made through the `mmadddisk` command to add a 4K native disk to a non-4K aligned file system while the file system version is not at 4.1.1.4 or later.

**User response:** Upgrade the file system to 4.1.1.4 or later, and then retry the command.

---

**6027-3450 Error *errorNumber* when purging key (file system *fileSystem*). Key name format possibly incorrect.**

**Explanation:** An error was encountered when purging a key from the key cache. The specified key name might have been incorrect, or an internal error was encountered.

**User response:** Ensure that the key name specified in the command is correct.

---

**6027-3451 Error *errorNumber* when emptying cache (file system *fileSystem*).**

**Explanation:** An error was encountered when purging all the keys from the key cache.

**User response:** Contact the IBM Support Center.

---

**6027-3452 [E] Unable to create encrypted file *fileName* (inode *inodeNumber*, fileset *filesetNumber*, file system *fileSystem*).**

**Explanation:** Unable to create a new encrypted file. The key required to encrypt the file might not be available.

**User response:** Examine the error message following this message for information on the specific failure.

---

**6027-3453 [E] Unable to open encrypted file: inode *inodeNumber*, fileset *filesetNumber*, file system *fileSystem*.**

**Explanation:** Unable to open an existing encrypted file. The key used to encrypt the file might not be available.

**User response:** Examine the error message following this message for information on the specific failure.

---

**6027-3457 [E] Unable to rewrap key with name *KeyName* (inode *inodeNumber*, fileset *filesetNumber*, file system *fileSystem*).**

**Explanation:** Unable to rewrap the key for a specified file because of an error with the key name.

**User response:** Examine the error message following this message for information on the specific failure.

---

**6027-3458 [E] Invalid length for the *KeyName* string.**

**Explanation:** The *KeyName* string has an incorrect length. The length of the specified string was either zero or it was larger than the maximum allowed length.

**User response:** Verify the *KeyName* string.

---

**6027-3459 [E] Not enough memory.**

**Explanation:** Unable to allocate memory for the *KeyName* string.

**User response:** Restart GPFS. Contact the IBM Support Center.

---

**6027-3460 [E] Incorrect format for the *KeyName* string.**

**Explanation:** An incorrect format was used when specifying the *KeyName* string.

**User response:** Verify the format of the *KeyName* string.

---

**6027-3461 [E] Error code: *errorNumber*.**

**Explanation:** An error occurred when processing a key ID.

**User response:** Contact the IBM Support Center.

---

**6027-3462 [E] Unable to rewrap key: original key name: *originalKeyName*, new key name: *newKeyName* (inode *inodeNumber*, fileset *filesetNumber*, file system *fileSystem*).**

**Explanation:** Unable to rewrap the key for a specified file, possibly because the existing key or the new key cannot be retrieved from the key server.

**User response:** Examine the error message following this message for information on the specific failure.

---

**6027-3463 [E] Rewrap error.**

**Explanation:** An internal error occurred during key rewrap.

**User response:** Examine the error messages surrounding this message. Contact the IBM Support Center.



---

**6027-3464 [E] New key is already in use.**

**Explanation:** The new key specified in a key rewrap is already being used.

**User response:** Ensure that the new key specified in the key rewrap is not being used by the file.

---

**6027-3465 [E] Cannot retrieve original key.**

**Explanation:** Original key being used by the file cannot be retrieved from the key server.

**User response:** Verify that the key server is available, the credentials to access the key server are correct, and that the key is defined on the key server.

---

**6027-3466 [E] Cannot retrieve new key.**

**Explanation:** Unable to retrieve the new key specified in the rewrap from the key server.

**User response:** Verify that the key server is available, the credentials to access the key server are correct, and that the key is defined on the key server.

---

**6027-3468 [E] Rewrap error code *errorNumber*.**

**Explanation:** Key rewrap failed.

**User response:** Record the error code and contact the IBM Support Center.

---

**6027-3469 [E] Encryption is enabled but the crypto module could not be initialized. Error code: *number*. Ensure that the GPFS crypto package was installed.**

**Explanation:** Encryption is enabled, but the cryptographic module required for encryption could not be loaded.

**User response:** Ensure that the packages required for encryption are installed on each node in the cluster.

---

**6027-3470 [E] Cannot create file *fileName*: extended attribute is too large: *numBytesRequired* bytes (*numBytesAvailable* available) (fileset *filesetName*, file system *fileSystem*).**

**Explanation:** Unable to create an encryption file because the extended attribute required for encryption is too large.

**User response:** Change the encryption policy so that the file key is wrapped fewer times, reduce the number of keys used to wrap a file key, or create a file system with a larger inode size.

---

**6027-3471 [E] At least one key must be specified.**

**Explanation:** No key name was specified.

**User response:** Specify at least one key name.

---

**6027-3472 [E] Could not combine the keys.**

**Explanation:** Unable to combine the keys used to wrap a file key.

**User response:** Examine the keys being used. Contact the IBM Support Center.

---

**6027-3473 [E] Could not locate the RKM.conf file.**

**Explanation:** Unable to locate the RKM.conf configuration file.

**User response:** Contact the IBM Support Center.

---

**6027-3474 [E] Could not open *fileType* file (*fileName* was specified).**

**Explanation:** Unable to open the specified configuration file. Encryption files will not be accessible.

**User response:** Ensure that the specified configuration file is present on all nodes.

---

**6027-3475 [E] Could not read file *fileName*.**

**Explanation:** Unable to read the specified file.

**User response:** Ensure that the specified file is accessible from the node.

---

**6027-3476 [E] Could not seek through file *fileName*.**

**Explanation:** Unable to seek through the specified file. Possible inconsistency in the local file system where the file is stored.

**User response:** Ensure that the specified file can be read from the local node.

---

**6027-3477 [E] Could not wrap the FEK.**

**Explanation:** Unable to wrap the file encryption key.

**User response:** Examine other error messages. Verify that the encryption policies being used are correct.

---

**6027-3478 [E] Insufficient memory.**

**Explanation:** Internal error: unable to allocate memory.

**User response:** Restart GPFS. Contact the IBM Support Center.

---

---

**6027-3479 [E] Missing combine parameter string.**

**Explanation:** The combine parameter string was not specified in the encryption policy.

**User response:** Verify the syntax of the encryption policy.

---

**6027-3480 [E] Missing encryption parameter string.**

**Explanation:** The encryption parameter string was not specified in the encryption policy.

**User response:** Verify the syntax of the encryption policy.

---

**6027-3481 [E] Missing wrapping parameter string.**

**Explanation:** The wrapping parameter string was not specified in the encryption policy.

**User response:** Verify the syntax of the encryption policy.

---

**6027-3482 [E] 'combineParameter' could not be parsed as a valid combine parameter string.**

**Explanation:** Unable to parse the combine parameter string.

**User response:** Verify the syntax of the encryption policy.

---

**6027-3483 [E] 'encryptionParameter' could not be parsed as a valid encryption parameter string.**

**Explanation:** Unable to parse the encryption parameter string.

**User response:** Verify the syntax of the encryption policy.

---

**6027-3484 [E] 'wrappingParameter' could not be parsed as a valid wrapping parameter string.**

**Explanation:** Unable to parse the wrapping parameter string.

**User response:** Verify the syntax of the encryption policy.

---

**6027-3485 [E] The Keyname string cannot be longer than *number* characters.**

**Explanation:** The specified *Keyname* string has too many characters.

**User response:** Verify that the specified *Keyname* string is correct.

---

---

**6027-3486 [E] The KMIP library could not be initialized.**

**Explanation:** The KMIP library used to communicate with the key server could not be initialized.

**User response:** Restart GPFS. Contact the IBM Support Center.

---

**6027-3487 [E] The RKM ID cannot be longer than *number* characters.**

**Explanation:** The remote key manager ID cannot be longer than the specified length.

**User response:** Use a shorter remote key manager ID.

---

**6027-3488 [E] The length of the key ID cannot be zero.**

**Explanation:** The length of the specified key ID string cannot be zero.

**User response:** Specify a key ID string with a valid length.

---

**6027-3489 [E] The length of the RKM ID cannot be zero.**

**Explanation:** The length of the specified RKM ID string cannot be zero.

**User response:** Specify an RKM ID string with a valid length.

---

**6027-3490 [E] The maximum size of the RKM.conf file currently supported is *number* bytes.**

**Explanation:** The RKM.conf file is larger than the size that is currently supported.

**User response:** User a smaller RKM.conf configuration file.

---

**6027-3491 [E] The string 'Keyname' could not be parsed as a valid key name.**

**Explanation:** The specified string could not be parsed as a valid key name.

**User response:** Specify a valid *Keyname* string.

---

**6027-3493 [E] *numKeys* keys were specified but a maximum of *numKeysMax* is supported.**

**Explanation:** The maximum number of specified key IDs was exceeded.

**User response:** Change the encryption policy to use fewer keys.

---

---

**6027-3494 [E] Unrecognized cipher mode.**

**Explanation:** Unable to recognize the specified cipher mode.

**User response:** Specify one of the valid cipher modes.

---

**6027-3495 [E] Unrecognized cipher.**

**Explanation:** Unable to recognize the specified cipher.

**User response:** Specify one of the valid ciphers.

---

**6027-3496 [E] Unrecognized combine mode.**

**Explanation:** Unable to recognize the specified combine mode.

**User response:** Specify one of the valid combine modes.

---

**6027-3497 [E] Unrecognized encryption mode.**

**Explanation:** Unable to recognize the specified encryption mode.

**User response:** Specify one of the valid encryption modes.

---

**6027-3498 [E] Invalid key length.**

**Explanation:** An invalid key length was specified.

**User response:** Specify a valid key length for the chosen cipher mode.

---

**6027-3499 [E] Unrecognized wrapping mode.**

**Explanation:** Unable to recognize the specified wrapping mode.

**User response:** Specify one of the valid wrapping modes.

---

**6027-3500 [E] Duplicate Keyname string '*keyIdentifier*'.**

**Explanation:** A given *Keyname* string has been specified twice.

**User response:** Change the encryption policy to eliminate the duplicate.

---

**6027-3501 [E] Unrecognized combine mode ('*combineMode*').**

**Explanation:** The specified combine mode was not recognized.

**User response:** Specify a valid combine mode.

---

**6027-3502 [E] Unrecognized cipher mode ('*cipherMode*').**

**Explanation:** The specified cipher mode was not recognized.

**User response:** Specify a valid cipher mode.

---

**6027-3503 [E] Unrecognized cipher ('*cipher*').**

**Explanation:** The specified cipher was not recognized.

**User response:** Specify a valid cipher.

---

**6027-3504 [E] Unrecognized encryption mode ('*mode*').**

**Explanation:** The specified encryption mode was not recognized.

**User response:** Specify a valid encryption mode.

---

**6027-3505 [E] Invalid key length ('*keyLength*').**

**Explanation:** The specified key length was incorrect.

**User response:** Specify a valid key length.

---

**6027-3506 [E] Mode '*mode1*' is not compatible with mode '*mode2*', aborting.**

**Explanation:** The two specified encryption parameters are not compatible.

**User response:** Change the encryption policy and specify compatible encryption parameters.

---

**6027-3509 [E] Key '*keyID:RK MID*' could not be fetched (RKM reported error *errorNumber*).**

**Explanation:** The key with the specified name cannot be fetched from the key server.

**User response:** Examine the error messages to obtain information about the failure. Verify connectivity to the key server and that the specified key is present at the server.

---

**6027-3510 [E] Could not bind symbol *symbolName* (*errorDescription*).**

**Explanation:** Unable to find the location of a symbol in the library.

**User response:** Contact the IBM Support Center.

---

**6027-3512 [E] The specified type '*type*' for backend '*backend*' is invalid.**

**Explanation:** An incorrect type was specified for a key server backend.

**User response:** Specify a correct backend type in RKM.conf.

---

---

**6027-3513 [E] Duplicate backend 'backend'.**

**Explanation:** A duplicate backend name was specified in RKM.conf.

**User response:** Specify unique RKM backends in RKM.conf.

---

**6027-3517 [E] Could not open library (libName).**

**Explanation:** Unable to open the specified library.

**User response:** Verify that all required packages are installed for encryption. Contact the IBM Support Center.

---

**6027-3518 [E] The length of the RKM ID string is invalid (must be between 0 and length characters).**

**Explanation:** The length of the RKM backend ID is invalid.

**User response:** Specify an RKM backend ID with a valid length.

---

**6027-3519 [E] 'numAttempts' is not a valid number of connection attempts.**

**Explanation:** The value specified for the number of connection attempts is incorrect.

**User response:** Specify a valid number of connection attempts.

---

**6027-3520 [E] 'sleepInterval' is not a valid sleep interval.**

**Explanation:** The value specified for the sleep interval is incorrect.

**User response:** Specify a valid sleep interval value (in microseconds).

---

**6027-3521 [E] 'timeout' is not a valid connection timeout.**

**Explanation:** The value specified for the connection timeout is incorrect.

**User response:** Specify a valid connection timeout (in seconds).

---

**6027-3522 [E] 'url' is not a valid URL.**

**Explanation:** The specified string is not a valid URL for the key server.

**User response:** Specify a valid URL for the key server.

---

**6027-3524 [E] 'tenantName' is not a valid tenantName.**

**Explanation:** An incorrect value was specified for the tenant name.

**User response:** Specify a valid tenant name.

---

**6027-3527 [E] Backend 'backend' could not be initialized (error errorNumber).**

**Explanation:** Key server backend could not be initialized.

**User response:** Examine the error messages. Verify connectivity to the server. Contact the IBM Support Center.

---

**6027-3528 [E] Unrecognized wrapping mode ('wrapMode').**

**Explanation:** The specified key wrapping mode was not recognized.

**User response:** Specify a valid key wrapping mode.

---

**6027-3529 [E] An error was encountered while processing file 'fileName':**

**Explanation:** An error was encountered while processing the specified configuration file.

**User response:** Examine the error messages that follow and correct the corresponding conditions.

---

**6027-3530 [E] Unable to open encrypted file: key retrieval not initialized (inode inodeNumber, fileset filesetNumber, file system fileSystem).**

**Explanation:** File is encrypted but the infrastructure required to retrieve encryption keys was not initialized, likely because processing of RKM.conf failed.

**User response:** Examine error messages at the time the file system was mounted.

---

**6027-3533 [E] Invalid encryption key derivation function.**

**Explanation:** An incorrect key derivation function was specified.

**User response:** Specify a valid key derivation function.

---

**6027-3534 [E] Unrecognized encryption key derivation function ('keyDerivation').**

**Explanation:** The specified key derivation function was not recognized.

**User response:** Specify a valid key derivation function.

---

**6027-3535 [E] Incorrect client certificate label**  
*'clientCertLabel' for backend 'backend'.*

**Explanation:** The specified client keypair certificate label is incorrect for the backend.

**User response:** Ensure that the correct client certificate label is used in RKM.conf.

---

**6027-3537 [E] Setting default encryption parameters requires empty combine and wrapping parameter strings.**

**Explanation:** A non-empty combine or wrapping parameter string was used in an encryption policy rule that also uses the default parameter string.

**User response:** Ensure that neither the combine nor the wrapping parameter is set when the default parameter string is used in the encryption rule.

---

**6027-3540 [E] The specified RKM backend type**  
*(rkmType) is invalid.*

**Explanation:** The specified RKM type in RKM.conf is incorrect.

**User response:** Ensure that only supported RKM types are specified in RKM.conf.

---

**6027-3541 [E] Encryption is not supported on**  
**Windows.**

**Explanation:** Encryption cannot be activated if there are Windows nodes in the cluster.

**User response:** Ensure that encryption is not activated if there are Windows nodes in the cluster.

---

**6027-3543 [E] The integrity of the file encrypting key**  
**could not be verified after unwrapping;**  
**the operation was cancelled.**

**Explanation:** When opening an existing encrypted file, the integrity of the file encrypting key could not be verified. Either the cryptographic extended attributes were damaged, or the master key(s) used to unwrap the FEK have changed.

**User response:** Check for other symptoms of data corruption, and verify that the configuration of the key server has not changed.

---

**6027-3545 [E] Encryption is enabled but there is no**  
**valid license. Ensure that the GPFS**  
**crypto package was installed properly.**

**Explanation:** The required license is missing for the GPFS encryption package.

**User response:** Ensure that the GPFS encryption package was installed properly.

---

**6027-3546 [E] Key 'keyID:rkmID' could not be fetched.**  
**The specified RKM ID does not exist;**  
**check the RKM.conf settings.**

**Explanation:** The specified RKM ID part of the key name does not exist, and therefore the key cannot be retrieved. The corresponding RKM might have been removed from RKM.conf.

**User response:** Check the set of RKMs specified in RKM.conf.

---

**6027-3547 [E] Key 'keyID:rkmID' could not be fetched.**  
**The connection was reset by the peer**  
**while performing the TLS handshake.**

**Explanation:** The specified key could not be retrieved from the server, because the connection with the server was reset while performing the TLS handshake.

**User response:** Check connectivity to the server. Check credentials to access the server. Contact the IBM Support Center.

---

**6027-3548 [E] Key 'keyID:rkmID' could not be fetched.**  
**The IP address of the RKM could not be**  
**resolved.**

**Explanation:** The specified key could not be retrieved from the server because the IP address of the server could not be resolved.

**User response:** Ensure that the hostname of the key server is correct. Verify whether there are problems with name resolutions.

---

**6027-3549 [E] Key 'keyID:rkmID' could not be fetched.**  
**The TCP connection with the RKM**  
**could not be established.**

**Explanation:** Unable to establish a TCP connection with the key server.

**User response:** Check the connectivity to the key server.

---

**6027-3550 Error when retrieving encryption**  
**attribute: errorDescription.**

**Explanation:** Unable to retrieve or decode the encryption attribute for a given file.

**User response:** File could be damaged and may need to be removed if it cannot be read.

---

**6027-3551 Error flushing work file** *fileName:*  
*errorString*

**Explanation:** An error occurred while attempting to flush the named work file or socket.

**User response:** None.

---

---

**6027-3552** Failed to fork a new process to *operationString* file system.

**Explanation:** Failed to fork a new process to suspend/resume file system.

**User response:** None.

---

**6027-3553** Failed to sync fileset *filesetName*.

**Explanation:** Failed to sync fileset.

**User response:** None.

---

**6027-3554** The restore command encountered an out-of-memory error.

**Explanation:** The fileset snapshot restore command encountered an out-of-memory error.

**User response:** None.

---

**6027-3555** *name* must be combined with FileInherit, DirInherit or both.

**Explanation:** NoPropagateInherit must be accompanied by other inherit flags. Valid values are FileInherit and DirInherit.

**User response:** Specify a valid NFSv4 option and reissue the command.

---

**6027-3556** *cmdName* error: insufficient memory.

**Explanation:** The command exhausted virtual memory.

**User response:** Consider some of the command parameters that might affect memory usage. Contact the IBM Support Center.

---

**6027-3557** *cmdName* error: could not create a temporary file.

**Explanation:** A temporary file could not be created in the current directory.

**User response:** Ensure that the file system is not full and that files can be created. Contact the IBM Support Center.

---

**6027-3558** *cmdName* error: could not initialize the key management subsystem (error *returnCode*).

**Explanation:** An internal component of the cryptographic library could not be properly initialized.

**User response:** Ensure that the gpfs.gskit package was installed properly. Contact the IBM Support Center.

---

**6027-3559** *cmdName* error: could not create the key database (error *returnCode*).

**Explanation:** The key database file could not be created.

**User response:** Ensure that the file system is not full and that files can be created. Contact the IBM Support Center.

---

**6027-3560** *cmdName* error: could not create the new self-signed certificate (error *returnCode*).

**Explanation:** A new certificate could not be successfully created.

**User response:** Ensure that the supplied canonical name is valid. Contact the IBM Support Center.

---

**6027-3561** *cmdName* error: could not extract the key item (error *returnCode*).

**Explanation:** The public key item could not be extracted successfully.

**User response:** Contact the IBM Support Center.

---

**6027-3562** *cmdName* error: base64 conversion failed (error *returnCode*).

**Explanation:** The conversion from or to the BASE64 encoding could not be performed successfully.

**User response:** Contact the IBM Support Center.

---

**6027-3563** *cmdName* error: could not extract the private key (error *returnCode*).

**Explanation:** The private key could not be extracted successfully.

**User response:** Contact the IBM Support Center.

---

**6027-3564** *cmdName* error: could not initialize the ICC subsystem (error *returnCode* *returnCode*).

**Explanation:** An internal component of the cryptographic library could not be properly initialized.

**User response:** Ensure that the gpfs.gskit package was installed properly. Contact the IBM Support Center.

---

**6027-3565** *cmdName* error: I/O error.

**Explanation:** A terminal failure occurred while performing I/O.

**User response:** Contact the IBM Support Center.

---

---

**6027-3566** *cmdName* **error: could not open file 'fileName'.**

**Explanation:** The specified file could not be opened.

**User response:** Ensure that the specified path and file name are correct and that you have sufficient permissions to access the file.

---

**6027-3567** *cmdName* **error: could not convert the private key.**

**Explanation:** The private key material could not be converted successfully.

**User response:** Contact the IBM Support Center.

---

**6027-3568** *cmdName* **error: could not extract the private key information structure.**

**Explanation:** The private key could not be extracted successfully.

**User response:** Contact the IBM Support Center.

---

**6027-3569** *cmdName* **error: could not convert the private key information to DER format.**

**Explanation:** The private key material could not be converted successfully.

**User response:** Contact the IBM Support Center.

---

**6027-3570** *cmdName* **error: could not encrypt the private key information structure (error returnCode).**

**Explanation:** The private key material could not be encrypted successfully.

**User response:** Contact the IBM Support Center.

---

**6027-3571** *cmdName* **error: could not insert the key in the keystore, check your system's clock (error returnCode).**

**Explanation:** Insertion of the new keypair into the keystore failed because the local date and time are not properly set on your system.

**User response:** Synchronize the local date and time on your system and try this command again.

---

**6027-3572** *cmdName* **error: could not insert the key in the keystore (error returnCode).**

**Explanation:** Insertion of the new keypair into the keystore failed.

**User response:** Contact the IBM Support Center.

---

**6027-3573** *cmdName* **error: could not insert the certificate in the keystore (error returnCode).**

**Explanation:** Insertion of the new certificate into the keystore failed.

**User response:** Contact the IBM Support Center.

---

**6027-3574** *cmdName* **error: could not initialize the digest algorithm.**

**Explanation:** Initialization of a cryptographic algorithm failed.

**User response:** Contact the IBM Support Center.

---

**6027-3575** *cmdName* **error: error while computing the digest.**

**Explanation:** Computation of the certificate digest failed.

**User response:** Contact the IBM Support Center.

---

**6027-3576** *cmdName* **error: could not initialize the SSL environment (error returnCode).**

**Explanation:** An internal component of the cryptographic library could not be properly initialized.

**User response:** Ensure that the `gpfs.gskit` package was installed properly. Contact the IBM Support Center.

---

**6027-3577** **Failed to sync fileset *filesetName*. *errString*.**

**Explanation:** Failed to sync fileset.

**User response:** Check the error message and try again. If the problem persists, contact the IBM Support Center.

---

**6027-3578 [E]** *pathName* **is not a valid argument for this command. You must specify a path name within a single GPFS snapshot.**

**Explanation:** This message is similar to message number 6027-872, but the *pathName* does not specify a path that can be scanned. The value specified for *pathName* might be a `.snapdir` or similar object.

**User response:** Correct the command invocation and reissue the command.

---

**6027-3579** *cmdName* **error: the cryptographic library could not be initialized in FIPS mode.**

**Explanation:** The cluster is configured to operate in FIPS mode but the cryptographic library could not be initialized in that mode.

**User response:** Verify that the `gpfs.gskit` package has been installed properly and that GPFS supports FIPS

mode on your platform. Contact the IBM Support Center.

---

**6027-3580**    **Failed to sync file system:** *fileSystem*  
**Error:** *errString*.

**Explanation:** Failed to sync file system.

**User response:** Check the error message and try again. If the problem persists, contact the IBM Support Center.

---

**6027-3581**    **Failed to create the operation list file.**

**Explanation:** Failed to create the operation list file.

**User response:** Verify that the file path is correct and check the additional error messages.

---

**6027-3582 [E]**    **Compression is not supported for clone or clone-parent files.**

**Explanation:** File compression is not supported as the file being compressed is a clone or a clone parent file.

**User response:** None.

---

**6027-3583 [E]**    **Compression is not supported for snapshot files.**

**Explanation:** The file being compressed is within a snapshot and snapshot file compression is not supported.

**User response:** None.

---

**6027-3584 [E]**    **Current file system version does not support compression.**

**Explanation:** The current file system version is not recent enough for file compression support.

**User response:** Upgrade the file system to the latest version and retry the command.

---

**6027-3585 [E]**    **Compression is not supported for AFM cached files.**

**Explanation:** The file being compressed is cached in an AFM cache fileset and compression is not supported for such files.

**User response:** None.

---

**6027-3586 [E]**    **Compression/uncompression failed.**

**Explanation:** Compression or uncompression failed.

**User response:** Refer to the error message below this line for the cause of the compression failure.

---

**6027-3587 [E]**    **Aborting compression as the file is opened in hyper allocation mode.**

**Explanation:** Compression operation is not performed because the file is opened in hyper allocation mode.

**User response:** Compress this file after the file is closed.

---

**6027-3588 [E]**    **Aborting compression as the file is currently memory mapped, opened in direct I/O mode, or stored in a horizontal storage pool.**

**Explanation:** Compression operation is not performed because it is inefficient or unsafe to compress the file at this time.

**User response:** Compress this file after the file is no longer memory mapped, opened in direct I/O mode, or stored in a horizontal storage pool.

---

**6027-3589**    *cmdName* **error: Cannot set the password twice.**

**Explanation:** An attempt was made to set the password by using different available options.

**User response:** Set the password either through the CLI or by specifying a file that contains it.

---

**6027-3590**    *cmdName* **error: Could not access file *fileName* (error *errorCode*).**

**Explanation:** The specified file could not be accessed.

**User response:** Check whether the file name is correct and verify whether you have required access privileges to access the file.

---

**6027-3591**    *cmdName* **error: The password specified in file *fileName* exceeds the maximum length of *length* characters.**

**Explanation:** The password stored in the specified file is too long.

**User response:** Pick a shorter password and retry the operation.

---

**6027-3592**    *cmdName* **error: Could not read the password from file *fileName*.**

**Explanation:** The password could not be read from the specified file.

**User response:** Ensure that the file can be read.



---

**6027-3593 [E]** Compression is supported only for regular files.

**Explanation:** The file is not compressed because compression is supported only for regular files.

**User response:** None.

---

**6027-3594 [E] [E]** Failed to synchronize the being restored fileset:*filesetName*. [I] Please stop the activities in the fileset and rerun the command.

**Explanation:** Failed to synchronize the being restored fileset due to some conflicted activities in the fileset.

**User response:** Stop the activities in the fileset and try the command again. If the problem persists, contact the IBM Support Center.

---

**6027-3595 [E] [E]** Failed to synchronize the being restored file system:*fileSystem*. [I] Please stop the activities in the file system and rerun the command.

**Explanation:** Failed to synchronize the being restored file system due to some conflicted activities in the file system.

**User response:** Stop the activities in the file system and try the command again. If the problem persists, contact the IBM Support Center.

---

**6027-3596** *cmdName* error: could not read/write file from/to directory '*pathName*'. This path does not exist.

**Explanation:** A file could not be read from/written to the specified directory.

**User response:** Ensure that the path exists.

---

**6027-3700 [E]** Key '*keyID:rkmID*' was not found on RKM ID '*rkmID*'.

**Explanation:** The specified key could not be retrieved from the key server.

**User response:** Verify that the key is present at the server. Verify that the name of the keys used in the encryption policy is correct.

---

**6027-3701 [E]** Key '*keyID:rkmID*' could not be fetched. The authentication with the RKM was not successful.

**Explanation:** Unable to authenticate with the key server.

**User response:** Verify that the credentials used to authenticate with the key server are correct.

---

**6027-3702 [E]** Key '*keyID:rkmID*' could not be fetched. Permission denied.

**Explanation:** Unable to authenticate with the key server.

**User response:** Verify that the credentials used to authenticate with the key server are correct.

---

**6027-3703 [E]** I/O error while accessing the keystore file '*keystoreFileName*'.

**Explanation:** An error occurred while accessing the keystore file.

**User response:** Verify that the name of the keystore file in RKM.conf is correct. Verify that the keystore file can be read on each node.

---

**6027-3704 [E]** The keystore file '*keystoreFileName*' has an invalid format.

**Explanation:** The specified keystore file has an invalid format.

**User response:** Verify that the format of the keystore file is correct.

---

**6027-3705 [E]** Incorrect FEK length after unwrapping; the operation was cancelled.

**Explanation:** When opening an existing encrypted file, the size of the FEK that was unwrapped did not correspond to the one recorded in the file's extended attributes. Either the cryptographic extended attributes were damaged, or the master key(s) used to unwrap the FEK have changed.

**User response:** Check for other symptoms of data corruption, and verify that the configuration of the key server has not changed.

---

**6027-3706 [E]** The crypto library with FIPS support is not available for this architecture. Disable FIPS mode and reattempt the operation.

**Explanation:** GPFS is operating in FIPS mode, but the initialization of the cryptographic library failed because FIPS mode is not yet supported on this architecture.

**User response:** Disable FIPS mode and attempt the operation again.

---

**6027-3707 [E]** The crypto library could not be initialized in FIPS mode. Ensure that the crypto library package was correctly installed.

**Explanation:** GPFS is operating in FIPS mode, but the initialization of the cryptographic library failed.

**User response:** Ensure that the packages required for

encryption are properly installed on each node in the cluster.

---

**6027-3708 [E] Incorrect passphrase for backend**  
*lineNumber: 'backend'.*

**Explanation:** The specified passphrase is incorrect for the backend.

**User response:** Ensure that the correct passphrase is used for the backend in RKM.conf.

---

**6027-3709 [E] Error encountered when parsing line**  
*lineNumber: expected a new RKM backend stanza.*

**Explanation:** An error was encountered when parsing a line in RKM.conf. Parsing of the previous backend is complete, and the stanza for the next backend is expected.

**User response:** Correct the syntax in RKM.conf.

---

**6027-3710 [E] Error encountered when parsing line**  
*lineNumber: invalid key 'keyIdentifier'.*

**Explanation:** An error was encountered when parsing a line in RKM.conf.

**User response:** Specify a well-formed stanza in RKM.conf.

---

**6027-3711 [E] Error encountered when parsing line**  
*lineNumber: invalid key-value pair.*

**Explanation:** An error was encountered when parsing a line in RKM.conf: an invalid key-value pair was found.

**User response:** Correct the specification of the RKM backend in RKM.conf.

---

**6027-3712 [E] Error encountered when parsing line**  
*lineNumber: incomplete RKM backend stanza 'backend'.*

**Explanation:** An error was encountered when parsing a line in RKM.conf. The specification of the backend stanza was incomplete.

**User response:** Correct the specification of the RKM backend in RKM.conf.

---

**6027-3713 [E] An error was encountered when parsing line**  
*lineNumber: duplicate key 'key'.*

**Explanation:** A duplicate keyword was found in RKM.conf.

**User response:** Eliminate duplicate entries in the backend specification.

---

**6027-3714 [E] Incorrect permissions for the**  
*/var/mmfs/etc/RKM.conf configuration file on node nodeName: the file must be owned by the root user and be in the root group, must be a regular file and be readable and writable by the owner only.*

**Explanation:** The permissions for the /var/mmfs/etc/RKM.conf configuration file are incorrect. The file must be owned by the root user, must be in the root group, must be a regular file, and must be readable and writable by the owner only.

**User response:** Fix the permissions on the file and retry the operation.

---

**6027-3715 [E] Error encountered when parsing line**  
*lineNumber: RKM ID 'RKMID' is too long, it cannot exceed length characters.*

**Explanation:** The RKMID chosen at the specified line of /var/mmfs/etc/RKM.conf contains too many characters.

**User response:** Choose a shorter string for the RKMID.

---

**6027-3716 [E] Key 'keyID:rkmID' could not be fetched.**  
**The TLS handshake could not be completed successfully.**

**Explanation:** The specified key could not be retrieved from the server because the TLS handshake did not complete successfully.

**User response:** Ensure that the configurations of GPFS and the remote key management (RKM) server are compatible when it comes to the version of the TLS protocol used upon key retrieval (GPFS uses the **nistCompliance** configuration variable to control that). In particular, if **nistCompliance=SP800-131A** is set in GPFS, ensure that the TLS v1.2 protocol is enabled in the RKM server. If this does not resolve the issue, contact the IBM Support Center.

---

**6027-3717 [E] Key 'keyID:rkmID' could not be fetched.**  
**The RKM is in quarantine after experiencing a fatal error.**

**Explanation:** GPFS has quarantined the remote key management (RKM) server and will refrain from initiating further connections to it for a limited amount of time.

**User response:** Examine the error messages that precede this message to determine the cause of the quarantine.

---

**6027-3718 [E] Key 'keyID:rkmID' could not be fetched. Invalid request.**

**Explanation:** The key could not be fetched because the remote key management (RKM) server reported that the request was invalid.

**User response:** Ensure that the RKM server trusts the client certificate that was used for this request. If this does not resolve the issue, contact the IBM Support Center.

---

**6027-3719 [W] Wrapping parameter string 'oldWrappingParameter' is not safe and will be replaced with 'newWrappingParameter'.**

**Explanation:** The wrapping parameter specified by the policy should no longer be used since it may cause data corruption or weaken the security of the system. For this reason, the wrapping parameter specified in the message will be used instead.

**User response:** Change the policy file and replace the specified wrapping parameter with a more secure one. Consult the *IBM Spectrum Scale: Administration Guide* for a list of supported wrapping parameters.

---

**6027-3720 [E] binaryName error: Invalid command type 'command'.**

**Explanation:** The command supplied to the specified binary is invalid.

**User response:** Specify a valid command. Refer to the documentation for a list of supported commands.

---

**6027-3721 [E] binaryName error: Invalid arguments.**

**Explanation:** The arguments supplied to the specified binary are invalid.

**User response:** Supply valid arguments. Refer to the documentation for a list of valid arguments.

---

**6027-3722 [E] An error was encountered while processing file 'fileName': errorString**

**Explanation:** An error was encountered while processing the specified configuration file.

**User response:** Examine the error message and correct the corresponding conditions.

---

**6027-3723 [E] Incorrect permissions for the configuration file fileName on node nodeName.**

**Explanation:** The permissions for the specified configuration file are incorrect. The file must be owned by the root user, must be in the root group, must be a regular file, and must be readable and writeable by the owner only.

**User response:** Fix the permissions on the file and retry the operation.

---

**6027-3726 [E] Key 'keyID:rkmID' could not be fetched. Bad certificate.**

**Explanation:** The key could not be fetched from the remote key management (RKM) server because of a problem with the validation of the certificate.

**User response:** Verify the steps used to generate the server and client certificates. Check whether the NIST settings are correct on the server. If this does not resolve the issue, contact the IBM Support Center.

---

**6027-3727 [E] Key 'keyID:rkmID' could not be fetched. Invalid tenantName.**

**Explanation:** The key could not be fetched from the remote key management (RKM) server because the tenantName specified in the RKM.conf file stanza was invalid.

**User response:** Verify that the tenantName specified in the RKM.conf file stanza is valid, and corresponds to an existing Device Group in the RKM server.

---

**6027-3728 [E] The keyStore permissions are incorrect for fileName. Access should be only granted to root, and no execute permission is allowed for the file.**

**Explanation:** The specified file allows access from a non-root user, or has execute permission, which is not allowed.

**User response:** Ensure the specified file is not granted access to non root. **Explanation:** The specified file allows access from a non-root user, or has execute permission, which is not allowed.

---

**6027-3900 Invalid flag 'flagName' in the criteria file.**

**Explanation:** An invalid flag was found in the criteria file.

**User response:** None.

---

**6027-3901 Failed to receive inode list: listName.**

**Explanation:** A failure occurred while receiving an inode list.

**User response:** None.

---

**6027-3902 Check file 'fileName' on fileSystem for inodes that were found matching the criteria.**

**Explanation:** The named file contains the inodes generated by parallel inode traversal (PIT) with interesting flags; for example, **dataUpdateMiss** or **BROKEN**.

## 6027-3903 [W] • 6027-3915

User response: None.

---

6027-3903 [W] *quotaType* quota is disabled or quota file is invalid.

**Explanation:** The corresponding quota type is disabled or invalid, and cannot be copied.

**User response:** Verify that the corresponding quota type is enabled.

---

6027-3904 [W] *quotaType* quota file is not a metadata file. File was not copied.

**Explanation:** The quota file is not a metadata file, and it cannot be copied in this way.

**User response:** Copy quota files directly.

---

6027-3905 [E] Specified directory does not exist or is invalid.

**Explanation:** The specified directory does not exist or is invalid.

**User response:** Check the spelling or validity of the directory.

---

6027-3906 [W] *backupQuotaFile* already exists.

**Explanation:** The destination file for a metadata quota file backup already exists.

**User response:** Move or delete the specified file and retry.

---

6027-3907 [E] No other quorum node found during cluster probe.

**Explanation:** The node could not renew its disk lease and there was no other quorum node available to contact.

**User response:** Determine whether there was a network outage, and also ensure the cluster is configured with enough quorum nodes. The node will attempt to rejoin the cluster.

---

6027-3908 Check file '*fileName*' on *fileSystem* for inodes with broken disk addresses or failures.

**Explanation:** The named file contains the inodes generated by parallel inode traversal (PIT) with interesting flags; for example, `dataUpdateMiss` or `BROKEN`.

**User response:** None.

---

6027-3909 The file (*backupQuotaFile*) is a quota file in *fileSystem* already.

**Explanation:** The file is a quota file already. An incorrect file name might have been specified.

**User response:** None.

---

6027-3910 [I] Delay number seconds for safe recovery.

**Explanation:** When disk lease is in use, wait for the existing lease to expire before performing log and token manager recovery.

**User response:** None.

---

6027-3911 Error reading message from the file system daemon: *errorString* : The system ran out of memory buffers or memory to expand the memory buffer pool.

**Explanation:** The system ran out of memory buffers or memory to expand the memory buffer pool. This prevented the client from receiving a message from the file system daemon.

**User response:** Try again later.

---

6027-3912 [E] File *fileName* cannot run with error *errorCode*: *errorString*.

**Explanation:** The named shell script cannot run.

**User response:** Verify that the file exists and that the access permissions are correct.

---

6027-3913 Attention: disk *diskName* is a 4K native dataOnly disk and it is used in a non-4K aligned file system. Its usage is not allowed to change from dataOnly.

**Explanation:** An attempt was made through the `mmchdisk` command to change the usage of a 4K native disk in a non-4K aligned file system from `dataOnly` to something else.

**User response:** None.

---

6027-3914 [E] Current file system version does not support compression.

**Explanation:** File system version is not recent enough for file compression support.

**User response:** Upgrade the file system to the latest version, then retry the command.

---

| 6027-3915 Invalid file system name provided:  
| '*FileSystemName*'.

| **Explanation:** The specified file system name contains  
| invalid characters.

| **User response:** Specify an existing file system name or

one which only contains valid characters.

---

**6027-3916 [E]** *FileSystemName* is a clone of *FileSystemName*, which is mounted already.

**Explanation:** A cloned file system is already mounted internally or externally. The file system with the same stripe group ID fails to mount and receive this message.

**User response:** Unmount the cloned file system and remount this file system.

---

**6027-4000 [I]** *descriptorType* descriptor on this NSD can be updated by running the following command from the node physically connected to NSD *nsdName*:

**Explanation:** This message is displayed when a descriptor validation thread finds a valid NSD, or disk, or stripe group descriptor but with a different ID. This can happen if a device is reused for another NSD.

**User response:** None. After this message, another message is displayed with a command to fix the problem.

---

**6027-4001 [I]** `'mmfsadm writeDesc <device> descriptorType descriptorId:descriptorId nsdFormatVersion pdiskStatus'`, where device is the device name of that NSD.

**Explanation:** This message displays the command that must run to fix the NSD or disk descriptor on that device. The *deviceName* must be supplied by system administrator or obtained from `mmlnsd -m` command. The *descriptorId* is a hexadecimal value.

**User response:** Run the command that is displayed on that NSD server node and replace *deviceName* with the device name of that NSD.

---

**6027-4002 [I]** Before running this command, check both NSDs. You might have to delete one of the NSDs.

**Explanation:** Informational message.

**User response:** The system administrator should decide which NSD to keep before running the command to fix it. If you want to keep the NSD found on disk, then you do not run the command. Instead, delete the other NSD found in cache (the NSD ID shown in the command).

---

**6027-4003 [E]** The on-disk *descriptorType* descriptor of *nsdName* *descriptorIdName* *descriptorId:descriptorId* is not valid because of bad *corruptionType*:

**Explanation:** The descriptor validation thread found

an on-disk descriptor that is corrupted. GPFS will automatically fix it.

**User response:** None.

---

**6027-4004 [D]** On-disk NSD descriptor: *nsdId* *nsdId* *nsdMagic* *nsdMagic* *nsdFormatVersion* *nsdFormatVersion* on disk *nsdChecksum* *nsdChecksum* calculated checksum *calculatedChecksum* *nsdDescSize* *nsdDescSize* *firstPaxosSector* *firstPaxosSector* *nPaxosSectors* *nPaxosSectors* *nsdIsPdisk* *nsdIsPdisk*

**Explanation:** Description of an on-disk NSD descriptor.

**User response:** None.

---

**6027-4005 [D]** Local copy of NSD descriptor: *nsdId* *nsdId* *nsdMagic* *nsdMagic* *formatVersion* *formatVersion* *nsdDescSize* *nsdDescSize* *firstPaxosSector* *firstPaxosSector* *nPaxosSectors* *nPaxosSectors*

**Explanation:** Description of the cached NSD descriptor.

**User response:** None.

---

**6027-4006 [I]** Writing NSD descriptor of *nsdName* with local copy: *nsdId* *nsdId* *nsdFormatVersion* *formatVersion* *firstPaxosSector* *firstPaxosSector* *nPaxosSectors* *nPaxosSectors* *nsdDescSize* *nsdDescSize* *nsdIsPdisk* *nsdIsPdisk* *nsdChecksum* *nsdChecksum*

**Explanation:** Description of the NSD descriptor that was written.

**User response:** None.

---

**6027-4007** *errorType* descriptor on *descriptorType* *nsdId* *nsdId:nsdId* error *error*

**Explanation:** This message is displayed after reading and writing NSD, disk and stripe group descriptors.

**User response:** None.

---

**6027-4008 [E]** On-disk *descriptorType* descriptor of *nsdName* is valid but has a different UID: *uid* *descriptorId:descriptorId* on-disk *uid* *descriptorId:descriptorId* *nsdId* *nsdId:nsdId*

**Explanation:** While verifying an on-disk descriptor, a valid descriptor was found but with a different ID. This can happen if a device is reused for another NSD with the `mmcrnsd -v no` command.

**User response:** After this message there are more

messages displayed that describe the actions to follow.

---

**6027-4009 [E] On-disk NSD descriptor of *nsdName* is valid but has a different ID. ID in cache is *cachedId* and ID on-disk is *ondiskId***

**Explanation:** While verifying an on-disk NSD descriptor, a valid descriptor was found but with a different ID. This can happen if a device is reused for another NSD with the `mmcrnsd -v no` command.

**User response:** After this message, there are more messages displayed that describe the actions to follow.

---

**6027-4010 [I] This corruption can happen if the device is reused by another NSD with the -v option and a file system is created with that reused NSD.**

**Explanation:** Description of a corruption that can happen when an NSD is reused.

**User response:** Verify that the NSD was not reused to create another NSD with the `-v` option and that the NSD was not used for another file system.

---

**6027-4011 [D] On-disk disk descriptor: uid *descriptorID:descriptorID* magic *descMagic* formatVersion *formatVersion* descSize *descSize* checksum on disk *diskChecksum* calculated checksum *calculatedChecksum* firstSGDescSector *firstSGDescSector* nSGDescSectors *nSGDescSectors* lastUpdateTime *lastUpdateTime***

**Explanation:** Description of the on-disk disk descriptor.

**User response:** None.

---

**6027-4012 [D] Local copy of disk descriptor: uid *descriptorID:descriptorID* firstSGDescSector *firstSGDescSector* nSGDescSectors *nSGDescSectors***

**Explanation:** Description of the cached disk descriptor.

**User response:** None.

---

**6027-4013 [I] Writing disk descriptor of *nsdName* with local copy: uid *descriptorID:descriptorID*, magic *magic*, formatVersion *formatVersion* firstSGDescSector *firstSGDescSector* nSGDescSectors *nSGDescSectors* descSize *descSize***

**Explanation:** Writing disk descriptor to disk with local information.

**User response:** None.

---

**6027-4014 [D] Local copy of StripeGroup descriptor: uid *descriptorID:descriptorID* curFmtVersion *curFmtVersion* configVersion *configVersion***

**Explanation:** Description of the cached stripe group descriptor.

**User response:** None.

---

**6027-4015 [D] On-disk StripeGroup descriptor: uid *sgUid:sgUid* magic *magic* curFmtVersion *curFmtVersion* descSize *descSize* on-disk checksum *diskChecksum* calculated checksum *calculatedChecksum* configVersion *configVersion* lastUpdateTime *lastUpdateTime***

**Explanation:** Description of the on-disk stripe group descriptor.

**User response:** None.

---

**6027-4016 [E] Data buffer checksum mismatch during write. File system *fileSystem* tag *tag1 tag2* nBytes *nBytes* diskAddresses**

**Explanation:** GPFS detected a mismatch in the checksum of the data buffer content which means content of data buffer was changing while a direct I/O write operation was in progress.

**User response:** None.

---

## Accessibility features for IBM Spectrum Scale

Accessibility features help users who have a disability, such as restricted mobility or limited vision, to use information technology products successfully.

---

### Accessibility features

The following list includes the major accessibility features in IBM Spectrum Scale:

- Keyboard-only operation
- Interfaces that are commonly used by screen readers
- Keys that are discernible by touch but do not activate just by touching them
- Industry-standard devices for ports and connectors
- The attachment of alternative input and output devices

IBM Knowledge Center, and its related publications, are accessibility-enabled. The accessibility features are described in IBM Knowledge Center ([www.ibm.com/support/knowledgecenter](http://www.ibm.com/support/knowledgecenter)).

---

### Keyboard navigation

This product uses standard Microsoft Windows navigation keys.

---

### IBM and accessibility

See the IBM Human Ability and Accessibility Center ([www.ibm.com/able](http://www.ibm.com/able)) for more information about the commitment that IBM has to accessibility.





---

## Notices

This information was developed for products and services offered in the US. This material might be available from IBM in other languages. However, you may be required to own a copy of the product or product version in that language in order to access it.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not grant you any license to these patents. You can send license inquiries, in writing, to:

*IBM Director of Licensing IBM Corporation North Castle Drive, MD-NC119 Armonk, NY 10504-1785 US*

For license inquiries regarding double-byte character set (DBCS) information, contact the IBM Intellectual Property Department in your country or send inquiries, in writing, to:

*Intellectual Property Licensing Legal and Intellectual Property Law IBM Japan Ltd. 19-21, Nihonbashi-Hakozakicho, Chuo-ku Tokyo 103-8510, Japan*

INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you provide in any way it believes appropriate without incurring any obligation to you.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact:

*IBM Director of Licensing IBM Corporation North Castle Drive, MD-NC119 Armonk, NY 10504-1785 US*

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The licensed program described in this document and all licensed material available for it are provided by IBM under terms of the IBM Customer Agreement, IBM International Program License Agreement or any equivalent agreement between us.

The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Statements regarding IBM's future direction or intent are subject to change or withdrawal without notice, and represent goals and objectives only.

All IBM prices shown are IBM's suggested retail prices, are current and are subject to change without notice. Dealer prices may vary.

This information is for planning purposes only. The information herein is subject to change before the products described become available.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to actual people or business enterprises is entirely coincidental.

#### COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. The sample programs are provided "AS IS", without warranty of any kind. IBM shall not be liable for any damages arising out of your use of the sample programs.

Each copy or any portion of these sample programs or any derivative work must include a copyright notice as follows:

© (your company name) (year).

Portions of this code are derived from IBM Corp.

Sample Programs. © Copyright IBM Corp. \_enter the year or years\_.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

---

## Trademarks

IBM, the IBM logo, and [ibm.com](http://ibm.com) are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the Web at Copyright and trademark information at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

Intel is a trademark of Intel Corporation or its subsidiaries in the United States and other countries.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft and Windows are trademarks of Microsoft Corporation in the United States, other countries, or both.

UNIX is a registered trademark of the Open Group in the United States and other countries.

---

## **Terms and conditions for product documentation**

Permissions for the use of these publications are granted subject to the following terms and conditions.

### **Applicability**

These terms and conditions are in addition to any terms of use for the IBM website.

### **Personal use**

You may reproduce these publications for your personal, noncommercial use provided that all proprietary notices are preserved. You may not distribute, display or make derivative work of these publications, or any portion thereof, without the express consent of IBM.

### **Commercial use**

You may reproduce, distribute and display these publications solely within your enterprise provided that all proprietary notices are preserved. You may not make derivative works of these publications, or reproduce, distribute or display these publications or any portion thereof outside your enterprise, without the express consent of IBM.

### **Rights**

Except as expressly granted in this permission, no other permissions, licenses or rights are granted, either express or implied, to the publications or any information, data, software or other intellectual property contained therein.

IBM reserves the right to withdraw the permissions granted herein whenever, in its discretion, the use of the publications is detrimental to its interest or, as determined by IBM, the above instructions are not being properly followed.

You may not download, export or re-export this information except in full compliance with all applicable laws and regulations, including all United States export laws and regulations.

IBM MAKES NO GUARANTEE ABOUT THE CONTENT OF THESE PUBLICATIONS. THE PUBLICATIONS ARE PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, AND FITNESS FOR A PARTICULAR PURPOSE.

---

## **IBM Online Privacy Statement**

IBM Software products, including software as a service solutions, ("Software Offerings") may use cookies or other technologies to collect product usage information, to help improve the end user experience, to tailor interactions with the end user or for other purposes. In many cases no personally identifiable information is collected by the Software Offerings. Some of our Software Offerings can help enable you to

collect personally identifiable information. If this Software Offering uses cookies to collect personally identifiable information, specific information about this offering's use of cookies is set forth below.

This Software Offering does not use cookies or other technologies to collect personally identifiable information.

If the configurations deployed for this Software Offering provide you as customer the ability to collect personally identifiable information from end users via cookies and other technologies, you should seek your own legal advice about any laws applicable to such data collection, including any requirements for notice and consent.

For more information about the use of various technologies, including cookies, for these purposes, See IBM's Privacy Policy at <http://www.ibm.com/privacy> and IBM's Online Privacy Statement at <http://www.ibm.com/privacy/details> the section entitled "Cookies, Web Beacons and Other Technologies" and the "IBM Software Products and Software-as-a-Service Privacy Statement" at <http://www.ibm.com/software/info/product-privacy>.

---

## Glossary

This glossary provides terms and definitions for IBM Spectrum Scale.

The following cross-references are used in this glossary:

- *See* refers you from a nonpreferred term to the preferred term or from an abbreviation to the spelled-out form.
- *See also* refers you to a related or contrasting term.

For other terms and definitions, see the IBM Terminology website ([www.ibm.com/software/globalization/terminology](http://www.ibm.com/software/globalization/terminology)) (opens in new window).

### B

#### block utilization

The measurement of the percentage of used subblocks per allocated blocks.

### C

#### cluster

A loosely-coupled collection of independent systems (nodes) organized into a network for the purpose of sharing resources and communicating with each other. See also *GPFS cluster*.

#### cluster configuration data

The configuration data that is stored on the cluster configuration servers.

#### Cluster Export Services (CES) nodes

A subset of nodes configured within a cluster to provide a solution for exporting GPFS file systems by using the Network File System (NFS), Server Message Block (SMB), and Object protocols.

#### cluster manager

The node that monitors node status using disk leases, detects failures, drives recovery, and selects file system managers. The cluster manager must be a quorum node. The selection of the cluster manager node favors the quorum-manager node with the lowest node number among the nodes that are operating at that particular time.

**Note:** The cluster manager role is not moved to another node when a node with a lower node number becomes active.

#### control data structures

Data structures needed to manage file data and metadata cached in memory. Control data structures include hash tables and link pointers for finding cached data; lock states and tokens to implement distributed locking; and various flags and sequence numbers to keep track of updates to the cached data.

### D

#### Data Management Application Program Interface (DMAPI)

The interface defined by the Open Group's XDSM standard as described in the publication *System Management: Data Storage Management (XDSM) API Common Application Environment (CAE) Specification C429*, The Open Group ISBN 1-85912-190-X.

#### deadman switch timer

A kernel timer that works on a node that has lost its disk lease and has outstanding I/O requests. This timer ensures that the node cannot complete the outstanding I/O requests (which would risk causing file system corruption), by causing a panic in the kernel.

#### dependent fileset

A fileset that shares the inode space of an existing independent fileset.

#### disk descriptor

A definition of the type of data that the disk contains and the failure group to which this disk belongs. See also *failure group*.

#### disk leasing

A method for controlling access to storage devices from multiple host systems. Any host that wants to access a storage device configured to use disk leasing registers for a lease; in the event of a perceived failure, a host system can deny access,

preventing I/O operations with the storage device until the preempted system has reregistered.

**disposition**

The session to which a data management event is delivered. An individual disposition is set for each type of event from each file system.

**domain**

A logical grouping of resources in a network for the purpose of common management and administration.

**E**

**ECKD** See *extended count key data (ECKD)*.

**ECKD device**

See *extended count key data device (ECKD device)*.

**encryption key**

A mathematical value that allows components to verify that they are in communication with the expected server. Encryption keys are based on a public or private key pair that is created during the installation process. See also *file encryption key*, *master encryption key*.

**extended count key data (ECKD)**

An extension of the count-key-data (CKD) architecture. It includes additional commands that can be used to improve performance.

**extended count key data device (ECKD device)**

A disk storage device that has a data transfer rate faster than some processors can utilize and that is connected to the processor through use of a speed matching buffer. A specialized channel program is needed to communicate with such a device. See also *fixed-block architecture disk device*.

**F**

**failback**

Cluster recovery from failover following repair. See also *failover*.

**failover**

(1) The assumption of file system duties by another node when a node fails. (2) The process of transferring all control of the ESS to a single cluster in the ESS when the other clusters in the ESS fails.

See also *cluster*. (3) The routing of all transactions to a second controller when the first controller fails. See also *cluster*.

**failure group**

A collection of disks that share common access paths or adapter connection, and could all become unavailable through a single hardware failure.

**FEK** See *file encryption key*.

**fileset** A hierarchical grouping of files managed as a unit for balancing workload across a cluster. See also *dependent fileset*, *independent fileset*.

**fileset snapshot**

A snapshot of an independent fileset plus all dependent filesets.

**file clone**

A writable snapshot of an individual file.

**file encryption key (FEK)**

A key used to encrypt sectors of an individual file. See also *encryption key*.

**file-management policy**

A set of rules defined in a policy file that GPFS uses to manage file migration and file deletion. See also *policy*.

**file-placement policy**

A set of rules defined in a policy file that GPFS uses to manage the initial placement of a newly created file. See also *policy*.

**file system descriptor**

A data structure containing key information about a file system. This information includes the disks assigned to the file system (*stripe group*), the current state of the file system, and pointers to key files such as quota files and log files.

**file system descriptor quorum**

The number of disks needed in order to write the file system descriptor correctly.

**file system manager**

The provider of services for all the nodes using a single file system. A file system manager processes changes to the state or description of the file system, controls the regions of disks that are allocated to each node, and controls token management and quota management.

**fixed-block architecture disk device (FBA disk device)**

A disk device that stores data in blocks of fixed size. These blocks are addressed by block number relative to the beginning of the file. See also *extended count key data device*.

**fragment**

The space allocated for an amount of data too small to require a full block. A fragment consists of one or more subblocks.

**G**

**global snapshot**

A snapshot of an entire GPFS file system.

**GPFS cluster**

A cluster of nodes defined as being available for use by GPFS file systems.

**GPFS portability layer**

The interface module that each installation must build for its specific hardware platform and Linux distribution.

**GPFS recovery log**

A file that contains a record of metadata activity, and exists for each node of a cluster. In the event of a node failure, the recovery log for the failed node is replayed, restoring the file system to a consistent state and allowing other nodes to continue working.

**I**

**ill-placed file**

A file assigned to one storage pool, but having some or all of its data in a different storage pool.

**ill-replicated file**

A file with contents that are not correctly replicated according to the desired setting for that file. This situation occurs in the interval between a change in the file's replication settings or suspending one of its disks, and the restripe of the file.

**independent fileset**

A fileset that has its own inode space.

**indirect block**

A block containing pointers to other blocks.

**inode** The internal structure that describes the

individual files in the file system. There is one inode for each file.

**inode space**

A collection of inode number ranges reserved for an independent fileset, which enables more efficient per-fileset functions.

**ISKLM**

IBM Security Key Lifecycle Manager. For GPFS encryption, the ISKLM is used as an RKM server to store MEKs.

**J**

**journaled file system (JFS)**

A technology designed for high-throughput server environments, which are important for running intranet and other high-performance e-business file servers.

**junction**

A special directory entry that connects a name in a directory of one fileset to the root directory of another fileset.

**K**

**kernel** The part of an operating system that contains programs for such tasks as input/output, management and control of hardware, and the scheduling of user tasks.

**M**

**master encryption key (MEK)**

A key used to encrypt other keys. See also *encryption key*.

**MEK** See *master encryption key*.

**metadata**

Data structures that contain information that is needed to access file data. Metadata includes inodes, indirect blocks, and directories. Metadata is not accessible to user applications.

**metanode**

The one node per open file that is responsible for maintaining file metadata integrity. In most cases, the node that has had the file open for the longest period of continuous time is the metanode.

**mirroring**

The process of writing the same data to multiple disks at the same time. The

mirroring of data protects it against data loss within the database or within the recovery log.

### **Microsoft Management Console (MMC)**

A Windows tool that can be used to do basic configuration tasks on an SMB server. These tasks include administrative tasks such as listing or closing the connected users and open files, and creating and manipulating SMB shares.

### **multi-tailed**

A disk connected to multiple nodes.

## **N**

### **namespace**

Space reserved by a file system to contain the names of its objects.

### **Network File System (NFS)**

A protocol, developed by Sun Microsystems, Incorporated, that allows any host in a network to gain access to another host or netgroup and their file directories.

### **Network Shared Disk (NSD)**

A component for cluster-wide disk naming and access.

### **NSD volume ID**

A unique 16 digit hex number that is used to identify and access all NSDs.

**node** An individual operating-system image within a cluster. Depending on the way in which the computer system is partitioned, it may contain one or more nodes.

### **node descriptor**

A definition that indicates how GPFS uses a node. Possible functions include: manager node, client node, quorum node, and nonquorum node.

### **node number**

A number that is generated and maintained by GPFS as the cluster is created, and as nodes are added to or deleted from the cluster.

### **node quorum**

The minimum number of nodes that must be running in order for the daemon to start.

### **node quorum with tiebreaker disks**

A form of quorum that allows GPFS to run with as little as one quorum node

available, as long as there is access to a majority of the quorum disks.

### **non-quorum node**

A node in a cluster that is not counted for the purposes of quorum determination.

## **P**

**policy** A list of file-placement, service-class, and encryption rules that define characteristics and placement of files. Several policies can be defined within the configuration, but only one policy set is active at one time.

### **policy rule**

A programming statement within a policy that defines a specific action to be performed.

**pool** A group of resources with similar characteristics and attributes.

### **portability**

The ability of a programming language to compile successfully on different operating systems without requiring changes to the source code.

### **primary GPFS cluster configuration server**

In a GPFS cluster, the node chosen to maintain the GPFS cluster configuration data.

### **private IP address**

A IP address used to communicate on a private network.

### **public IP address**

A IP address used to communicate on a public network.

## **Q**

### **quorum node**

A node in the cluster that is counted to determine whether a quorum exists.

**quota** The amount of disk space and number of inodes assigned as upper limits for a specified user, group of users, or fileset.

### **quota management**

The allocation of disk blocks to the other nodes writing to the file system, and comparison of the allocated space to quota limits at regular intervals.



## R

### **Redundant Array of Independent Disks (RAID)**

A collection of two or more disk physical drives that present to the host an image of one or more logical disk drives. In the event of a single physical device failure, the data can be read or regenerated from the other disk drives in the array due to data redundancy.

### **recovery**

The process of restoring access to file system data when a failure has occurred. Recovery can involve reconstructing data or providing alternative routing through a different server.

### **remote key management server (RKM server)**

A server that is used to store master encryption keys.

### **replication**

The process of maintaining a defined set of data in more than one location. Replication involves copying designated changes for one location (a source) to another (a target), and synchronizing the data in both locations.

### **RKM server**

See *remote key management server*.

**rule** A list of conditions and actions that are triggered when certain conditions are met. Conditions include attributes about an object (file name, type or extension, dates, owner, and groups), the requesting client, and the container name associated with the object.

## S

### **SAN-attached**

Disks that are physically attached to all nodes in the cluster using Serial Storage Architecture (SSA) connections or using Fibre Channel switches.

### **Scale Out Backup and Restore (SOBAR)**

A specialized mechanism for data protection against disaster only for GPFS file systems that are managed by IBM Spectrum Protect Hierarchical Storage Management (HSM).

### **secondary GPFS cluster configuration server**

In a GPFS cluster, the node chosen to maintain the GPFS cluster configuration

data in the event that the primary GPFS cluster configuration server fails or becomes unavailable.

### **Secure Hash Algorithm digest (SHA digest)**

A character string used to identify a GPFS security key.

### **session failure**

The loss of all resources of a data management session due to the failure of the daemon on the session node.

### **session node**

The node on which a data management session was created.

### **Small Computer System Interface (SCSI)**

An ANSI-standard electronic interface that allows personal computers to communicate with peripheral hardware, such as disk drives, tape drives, CD-ROM drives, printers, and scanners faster and more flexibly than previous interfaces.

### **snapshot**

An exact copy of changed data in the active files and directories of a file system or fileset at a single point in time. See also *fileset snapshot*, *global snapshot*.

### **source node**

The node on which a data management event is generated.

### **stand-alone client**

The node in a one-node cluster.

### **storage area network (SAN)**

A dedicated storage network tailored to a specific environment, combining servers, storage products, networking products, software, and services.

### **storage pool**

A grouping of storage space consisting of volumes, logical unit numbers (LUNs), or addresses that share a common set of administrative characteristics.

### **stripe group**

The set of disks comprising the storage assigned to a file system.

### **striping**

A storage process in which information is split into blocks (a fixed amount of data) and the blocks are written to (or read from) a series of disks in parallel.

### **subblock**

The smallest unit of data accessible in an I/O operation, equal to one thirty-second of a data block.

### **system storage pool**

A storage pool containing file system control structures, reserved files, directories, symbolic links, special devices, as well as the metadata associated with regular files, including indirect blocks and extended attributes. The **system storage pool** can also contain user data.

## **T**

### **token management**

A system for controlling file access in which each application performing a read or write operation is granted some form of access to a specific block of file data. Token management provides data consistency and controls conflicts. Token management has two components: the token management server, and the token management function.

### **token management function**

A component of token management that requests tokens from the token management server. The token management function is located on each cluster node.

### **token management server**

A component of token management that controls tokens relating to the operation of the file system. The token management server is located at the file system manager node.

### **transparent cloud tiering (TCT)**

A separately installable add-on feature of IBM Spectrum Scale that provides a native cloud storage tier. It allows data center administrators to free up on-premise storage capacity, by moving out cooler data to the cloud storage, thereby reducing capital and operational expenditures. .

### **twin-tailed**

A disk connected to two nodes.

## **U**

### **user storage pool**

A storage pool containing the blocks of data that make up user files.

## **V**

**VFS** See *virtual file system*.

### **virtual file system (VFS)**

A remote file system that has been mounted so that it is accessible to the local user.

### **virtual node (vnode)**

The structure that contains information about a file system object in a virtual file system (VFS).

---

# Index

## Special characters

- /etc/filesystems 246
- /etc/fstab 246
- /etc/hosts 227
- /etc/resolv.conf 243
- /tmp/mmfs 134, 367
- /usr/lpp/mmfs/bin 232
- /usr/lpp/mmfs/bin/runmmfs 163
- /usr/lpp/mmfs/samples/gatherlogs.samples.sh file 141
- /var/adm/ras/mmfs.log.previous 237
- /var/mmfs/etc/mmlock 229
- /var/mmfs/gen/mmsdrfs 230
- .ptrash directory 136
- .rhosts 228
- .snapshots 268, 269, 270

## A

- access
  - to disk 279
- ACCESS\_TIME attribute 198, 199
- accessibility features for IBM Spectrum Scale 591
- active file management, questions related to 135
- administration commands
  - failure 229
- administration, collector node 107
- AFM 135
  - fileset states 93
  - issues 353
  - monitoring commands 100
  - monitoring policies 102
- AFM DR
  - fileset states 96
  - issues 357
  - monitoring commands 100
- AFM fileset, changing mode of 135
- AFM, extended attribute size supported by 136
- AFM, messages queuing 274
- AIX
  - kernel debugger 206
- AIX error logs
  - MMFS\_DISKFAIL 279
  - MMFS\_QUOTA 251
  - unavailable disks 251
- AIX logical volume
  - down 282
- AIX platform
  - gpfs.snap command 177
- application program errors 240
- application programs
  - errors 158, 159, 236, 239
- audit messages 141
- authentication 180
  - problem determination 294
- Authentication error events 295
- Authentication errors
  - SSD process not running (sssd\_down) 295
  - sssd\_down 295
  - Winbind process not running (wnbd\_down) 295
  - wnbd\_down 295

- Authentication errors (*continued*)
  - yp\_down 295
  - YPBIND process not running (yp\_down) 295
- authorization error 228
- autofs 254
- autofs mount 253
- autoload option
  - on mmchconfig command 233
  - on mmcrcluster command 233
- automatic backup of cluster data 231
- automount 248, 253
- automount daemon 253
- automount failure 253, 254, 255
- Availability 371

## B

- back up data 130
- backup
  - automatic backup of cluster data 231
- best practices for troubleshooting 129
- block allocation 334

## C

- call home 117
  - configuring 119
  - mmcallhome 117
  - monitoring IBM Spectrum Scale system remotely 117
  - prerequisites 124
  - sharing collected data 124
- callback events
  - AFM 100
- candidate file 194, 196
  - attributes 198
- CES
  - monitoring 155
  - troubleshooting 155
- CES administration 155
- CES collection 167
- ces configuration issues 239
- CES logs 146
- CES monitoring 155
- CES tracing 167
- changing mode of AFM fileset 135
- checking, Persistent Reserve 288
- chosen file 194, 195
- CIFS serving, Windows SMB2 protocol 241
- cipherList 258
- Clearing a leftover Persistent Reserve reservation 288
- client node 248
- clock synchronization 141, 263
- cluster
  - deleting a node 238
- cluster configuration information
  - displaying 188
- cluster configuration information, SNMP 109
- Cluster Export Services
  - administration 155
  - issue collection 167

- Cluster Export Services *(continued)*
  - monitoring 155
  - tracing 167
- cluster file systems
  - displaying 189
- cluster security configuration 256
- cluster state information 187
- cluster status information, SNMP 109
- collecting details 117
- collector
  - performance monitoring tool 41
    - configuring 43, 60, 61
- collector node
  - installing MIB files 107
- commands
  - cluster state information 187
  - conflicting invocation 245
  - errpt 367
  - file system and disk information 191
  - gpfs.snap 176, 177, 178, 179, 367
  - grep 157
  - lslpp 368
  - lslv 133
  - lsdf 192, 249, 250
  - lspv 282
  - lsvg 282
  - lxtrace 162, 186
  - mmadddisk 260, 266, 281, 283, 284
  - mmaddnode 133, 227, 250
  - mmafmctl 274
  - mmafmctl Device getstate 187
  - mmapplypolicy 193, 263, 264, 266, 294
  - mmauth 204, 256
  - mmbackup 273
  - mmchcluster 228
  - mmchconfig 189, 233, 250, 258
  - mmchdisk 246, 260, 266, 275, 278, 279, 284, 287
  - mmcheckquota 159, 199, 240, 251
  - mmchfs 160, 231, 238, 246, 248, 251, 253, 305, 345
  - mmchnode 107, 108, 133
  - mmchnsd 275
  - mmcommon recoverfs 261
  - mmcommon showLocks 230
  - mmcrcluster 133, 189, 227, 228, 233
  - mmcrfs 238, 275, 283, 305
  - mmcrnsd 275, 278
  - mmcrsnapshot 268, 269
  - mmdeldisk 260, 266, 281, 284
  - mmdelfileset 265
  - mmdelfs 286
  - mmdelnode 227, 238
  - mmdelnsd 278, 286
  - mmdelsnapshot 269
  - mmdf 261, 281, 344
  - mmdiag 187
  - mmdsh 229
  - mmdumpperfdata 184
  - mmexpelnode 190
  - mmfileid 201, 274, 284
  - mmfsadm 166, 186, 234, 274, 284, 345
  - mmfsck 135, 192, 245, 246, 266, 274, 281, 285
  - mmgetstate 187, 234, 237
  - mmlsattr 264, 265
  - mmlscluster 107, 133, 188, 227, 257
  - mmlsconfig 162, 189, 253
  - mmlsdisk 238, 245, 246, 251, 260, 275, 278, 279, 283, 287, 368

- commands *(continued)*
  - mmlsfileset 265
  - mmlsfs 247, 284, 286, 368
  - mmlsmgr 186, 246
  - mmlsmount 192, 233, 240, 245, 249, 250, 275
  - mmlsnsd 200, 276, 282
  - mmlspolicy 264
  - mmlsquota 240
  - mmlssnapshot 267, 268, 269, 270
  - mmmount 192, 245, 251, 283
  - mmperfmon 64
  - mmpmon 3, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 29, 30, 31, 32, 34, 35, 41, 206, 271, 273
  - mmquotaoff 240
  - mmquotaon 240
  - mmrefresh 189, 246, 253
  - mmremotefilesystem 204, 256, 257
  - mmremotefs 253, 257
  - mmrepquota 240
  - mmrestorefs 269, 270, 271
  - mmrestripefile 264, 267
  - mmrestripefs 267, 281, 284
  - mmrpldisk 260, 266, 283
  - mmsdrrestore 190
  - mmsshutdown 108, 188, 190, 233, 234, 236, 254, 259
  - mmsnapdir 267, 268, 270
  - mmstartup 233, 254
  - mmumount 249, 251, 281
  - mmunlinkfileset 265
  - mmwindisk 201
  - mount 245, 246, 248, 283, 285
  - ping 229
  - rcp 228
  - rpm 367
  - rsh 228, 237
  - scp 228
  - ssh 228
  - umount 250, 251
  - varyonvg 283
- commands, administration
  - failure 229
- common issues and workarounds
  - transparent cloud tiering 359
- communication paths
  - unavailable 246
- compiling mmfslinux module 232
- configuration
  - hard loop ID 228
  - performance tuning 228
- configuration data 260
- configuration parameters
  - kernel 232
- configuration problems 226
- configuration variable settings
  - displaying 189
- configuring call home 119
- configuring Net-SNMP 106
- configuring SNMP for use with IBM Spectrum Scale 107
- configuring SNMP-based management applications 107
- connectivity problems 229
- contact node address 256
- contact node failure 257
- creating a file, failure 293
- cron 134

## D

- data
  - replicated 284
- data always gathered by gpfs.snap 176
  - for a master snapshot 179
  - on AIX 177
  - on all platforms 176
  - on Linux 178
  - on Windows 179
- Data always gathered for an Object on Linux 181
- Data always gathered for authentication on Linux 182
- Data always gathered for CES on Linux 182
- Data always gathered for NFS on Linux 180
- Data always gathered for performance on Linux 183
- Data always gathered for SMB on Linux 180
- data collection 180
- data file issues
  - cluster configuration 229
- data gathered by
  - gpfs.snap on Linux 180
- data integrity 159, 273
- Data Management API (DMAPI)
  - file system will not mount 247
- data replication 280
- data structure 158
- data update
  - file system with snapshot 344
  - fileset with snapshot 344
- dataOnly attribute 266
- dataStructureDump 162
- dead man switch timer 291
- deadlock
  - automated breakup 210
  - breakup on demand 211
- deadlocks 344, 345
  - automated detection 208, 210
  - debug data 207
  - information about 207
  - log 207
- debug data
  - deadlocks 207
- debug data collection
  - CES tracing 167
- delays 344, 345
- DELETE rule 194, 196
- deleting a node
  - from a cluster 238
- descOnly 252
- diagnostic data
  - deadlock diagnostics 174
  - standard diagnostics 174
- directed maintenance procedure 362
  - increase fileset space 363
  - start cluster 362
  - start gpfs daemon 363
  - start NSD 362
  - start performance monitoring collector service 364
  - start performance monitoring sensor service 365
  - synchronize node clocks 363
- directories
  - /tmp/mmfs 134, 367
  - .snapshots 268, 269, 270
- directory that has not been cached, traversing 136
- disabling IPv6
  - for SSH connection delays 243
- disabling Persistent Reserve manually 289

- disaster recovery
  - other problems 326
  - setup problems 325
- disk access 279
- disk commands
  - hang 283
- disk configuration information, SNMP 113
- disk connectivity failure 281
- disk descriptor replica 251
- disk failover 281
- disk failure 286
- disk leasing 291
- disk performance information, SNMP 114
- disk recovery 281
- disk status information, SNMP 113
- disk subsystem
  - failure 275
- disks
  - damaged files 201
  - declared down 278
  - define for GPFS use 282
  - displaying information of 200
  - failure 157, 159, 275
  - media failure 283
  - partial failure 281
  - replacing 261
  - usage 252
- disks down 282
- disks, viewing 201
- displaying disk information 200
- displaying NSD information 276
- DMP 362
- DNS server failure 256

## E

- enabling Persistent Reserve manually 289
- encryption issues 293
  - issues with adding encryption policy 293
  - permission denied message 293
- ERRNO I/O error code 237
- error codes
  - EINVAL 263
  - EIO 158, 275, 285, 286
  - ENODEV 236
  - ENOENT 250
  - ENOSPC 261, 285
  - ERRNO I/O 237
  - ESTALE 159, 236, 250
  - NO SUCH DIRECTORY 236
  - NO SUCH FILE 236
- error log
  - MMFS\_LONGDISKIO 158
  - MMFS\_QUOTA 159
- error logs
  - example 160
  - MMFS\_ABNORMAL\_SHUTDOWN 157
  - MMFS\_DISKFAIL 157
  - MMFS\_ENVIRON 158
  - MMFS\_FSSTRUCT 158
  - MMFS\_GENERIC 158
  - MMFS\_LONGDISKIO 158
  - MMFS\_QUOTA 159, 199
  - MMFS\_SYSTEM\_UNMOUNT 159
  - MMFS\_SYSTEM\_WARNING 160
  - operating system 157

## error messages

0516-1339 278  
0516-1397 278  
0516-862 278  
6027-1209 236  
6027-1242 230  
6027-1290 261  
6027-1598 227  
6027-1615 229  
6027-1617 229  
6027-1627 239  
6027-1628 230  
6027-1630 230  
6027-1631 230  
6027-1632 230  
6027-1633 230  
6027-1636 276  
6027-1661 276  
6027-1662 278  
6027-1995 273  
6027-1996 259  
6027-2108 276  
6027-2109 276  
6027-300 233  
6027-306 235  
6027-319 234, 235  
6027-320 235  
6027-321 235  
6027-322 235  
6027-341 232, 235  
6027-342 232, 235  
6027-343 232, 235  
6027-344 232, 235  
6027-361 281  
6027-418 252, 286  
6027-419 247, 252  
6027-435 325  
6027-473 252  
6027-474 252  
6027-482 247, 286  
6027-485 286  
6027-490 325  
6027-506 240  
6027-533 345  
6027-538 238  
6027-549 247  
6027-580 247  
6027-631 260  
6027-632 260  
6027-635 260  
6027-636 260, 286  
6027-638 260  
6027-645 247  
6027-650 236  
6027-663 238  
6027-665 233, 238, 239  
6027-695 241  
6027-953 271  
ANS1312E 273  
cluster configuration data file issues 230  
descriptor replica 325  
disk media failures 286  
failed to connect 233, 281  
file system forced unmount problems 252  
file system mount problems 247  
GPFS cluster data recovery 230  
IBM Spectrum Protect 273

## error messages (continued)

incompatible version number 234  
mmbackup 273  
mmfsd ready 233  
multiple file system manager failures 260  
network problems 235  
quorum 325  
rsh problems 229  
shared segment problems 234, 235  
snapshot 268, 269, 270, 271

## error numbers

application calls 248  
configuration problems 231  
data corruption 274  
EALL\_UNAVAIL = 218 252  
ECONFIG = 208 231  
ECONFIG = 215 231, 235  
ECONFIG = 218 231  
ECONFIG = 237 231  
ENO\_MGR = 212 260, 287  
ENO\_QUOTA\_INST = 237 248  
EOFFLINE = 208 286  
EPANIC = 666 252  
EVALIDATE = 214 274  
file system forced unmount 252  
GPFS application calls 286  
GPFS daemon will not come up 235  
installation problems 231  
multiple file system manager failures 260

errors, application program 240  
errors, Persistent Reserve 287  
errpt command 367

## event notifications

call back 129  
call home 129  
email 129  
emails 80  
simple network management protocol (SNMP) 129

## events 80

Availability 371  
notifications 79  
snmp 80  
Reliability 371  
Serviceability 371

## example

error logs 160

## examples for GUI issues 347

EXCLUDE rule 197, 198

excluded file 197  
attributes 198

extended attribute size supported by AFM 136

## F

facility  
Linux kernel crash dump (LKCD) 206

## failure

disk 278  
mmccr command 135  
mmfsck command 135  
of disk media 283  
snapshot 267

failure creating a file 293

## failure group 251

## failure groups

loss of 251  
use of 251

- failure, key rewrap 294
- failure, mount 293
- failures
  - mmbackup 273
- File Authentication
  - setup problems 294
- file creation failure 293
- file migration
  - problems 264
- file placement policy 264
- file system 245
  - mount status 259
  - space 261
- File system
  - high utilization 337
- file system descriptor 251, 252
  - failure groups 251
  - inaccessible 251
- file system manager
  - cannot appoint 250
  - contact problems
    - communication paths unavailable 246
    - multiple failures 259, 260
- file system mount failure 293
- file system or fileset getting full 136
- file system performance information, SNMP 112
- file system status information, SNMP 111
- file systems
  - cannot be unmounted 192
  - creation failure 238
  - determining if mounted 259
  - discrepancy between configuration data and on-disk data 260
  - displaying statistics 5, 6
  - do not mount 245
  - does not mount 245
  - does not unmount 249
  - forced unmount 159, 250, 259, 260
  - free space shortage 270, 271
  - listing mounted 192
  - loss of access 239
  - remote 255
  - reset statistics 15
  - state after restore 270, 271
  - unable to determine if mounted 259
  - will not mount 192
- FILE\_SIZE attribute 198, 199
- files
  - /etc/filesystems 246
  - /etc/fstab 246
  - /etc/group 159
  - /etc/hosts 227
  - /etc/passwd 159
  - /etc/resolv.conf 243
  - /usr/lpp/mmfs/bin/runmmfs 163
  - /usr/lpp/mmfs/samples/gatherlogs.samples.sh 141
  - /var/adm/ras/mmfs.log.previous 237
  - /var/mmfs/etc/mmlock 229
  - /var/mmfs/gen/mmsdrfs 230
  - .rhosts 228
  - detecting damage 201
  - mmfs.log 233, 234, 236, 245, 248, 250, 254, 256, 257, 258, 367
  - protocol authentication log 153
- FILESET\_NAME attribute 198, 199
- filesets
  - child 265

- filesets (*continued*)
  - deleting 265
  - emptying 265
  - errors 265
  - lost+found 266
  - moving contents 265
  - performance 265
  - problems 261
  - snapshots 265
  - unlinking 265
  - usage errors 265
- FSDesc structure 251
- full file system or fileset 136

## G

- generate
  - trace reports 162
- generating GPFS trace reports
  - mmtracectl command 162
- getting started with troubleshooting 129
- GPFS
  - /tmp/mmfs directory 134
  - abnormal termination in mmpmon 272
  - active file management 135
  - AFM 274
  - AIX 254
  - application program errors 240
  - authentication issues 294
  - automount 253
  - automount failure 254
  - automount failure in Linux 253
  - Availability 371
  - checking Persistent Reserve 288
  - cipherList option has not been set properly 258
  - clearing a leftover Persistent Reserve reservation 288
  - client nodes 248
  - cluster configuration
    - issues 230
  - cluster name 256
  - cluster security configurations 256
  - cluster state information commands 187, 188, 189, 190
  - command 176, 177, 178, 179, 187
  - configuration data 260
  - contact node address 256
  - contact nodes down 257
  - core dumps 161
  - corrupted data integrity 274
  - data gathered for protocol on Linux 180, 181, 182, 183
  - data integrity 273, 274
  - data integrity may be corrupted 274
  - deadlocks 207, 210, 211
  - delays and deadlocks 344
  - determine if a file system is mounted 259
  - determining the health of integrated SMB server 304
  - disaster recovery issues 325
  - discrepancy between GPFS configuration data and the on-disk data for a file system 260
  - disk accessing command failure 283
  - disk connectivity failure 281
  - disk failure 280, 286
  - disk information commands 191, 192, 193, 201
  - disk issues 275, 291
  - disk media failure 283, 286
  - disk recovery 281
  - disk subsystem failures 275
  - displaying NSD information 276

GPFS (*continued*)

- encryption rules 293
- error creating internal storage 135
- error encountered while creating NSD disks 275
- error encountered while using NSD disks 275
- error messages for file system 247, 248
- error message 135, 260
- error messages 269, 270, 271, 273, 286
- error messages for file system forced unmount problems 252
- error messages for file system mount status 259
- error messages for indirect snapshot errors 268
- error messages not directly related to snapshots 268
- error messages related to snapshots 268, 269
- error message "Function not implemented" 255
- error numbers 252, 260, 286
- error numbers specific to GPFS application calls 274
- errors 264, 265, 287
- errors associated with filesets 261
- errors associated with policies 261
- errors associated with storage pools, 261
- errors encountered 267
- errors encountered with filesets 265
- failure group considerations 251
- failures using the mmbackup command 273
- file system 249, 250, 293
- file system commands 191, 192, 193, 201
- file system failure 245
- file system has adequate free space 261
- file system is forced to unmount 252
- file system is mounted 259
- file system issues 245
- file system manager appointment fails 260
- file system manager failures 260
- file system mount problems 247, 248, 253
- file system mount status 259
- file system mounting 135
- file systems manager failure 259, 260
- filesets usage 265
- forced unmount 250
- gpfs.snap 176, 177, 178, 179
- guarding against disk failures 280
- GUI logs 174
- hang in mmpmon 272
- health of integrated SMB server 304
- IBM Spectrum Protect error messages 273
- ill-placed files 263
- incorrect output from mmpmon 272
- indirect issues with snapshot 268
- installation and configuration issues 213, 226, 227, 229, 232, 233, 234, 235, 236, 237, 239, 241
- integrated SMB server 304
- issues while working with Samba 305
- issues with snapshot 268, 269
- key rewrap 294
- limitations 137
- local node failure 257
- locating snapshot 268
- logical volumes are properly defined 282
- logs 140
- manually disabling Persistent Reserve 289
- manually enabling Persistent Reserve 289
- mapping 255
- message 6027-648 134
- message referring to an existing NSD 278
- message requeuing 274
- message requeuing in AFM 274

GPFS (*continued*)

- message severity tags 443
- messages 443
- mmafmctl Device getstate 187
- mmapplypolicy -L command 194, 195, 196, 197, 198, 199
- mmbackup command 273
- mmbackup errors 273
- mmdumpperfdata command 184
- mmexpelnode command 190
- mmfsadm command 186
- mmpmon 272
- mmpmon command 273
- mmpmon output 272
- mmremotecluster command 256
- mount 134, 253, 255
- mount failure 248, 293
- mounting cluster 258
- mounting cluster does not have direct access to the disks 258
- multipath device 290
- multiple file system manager failures 259, 260
- negative values in the 'predicted pool utilizations', 263
- network issues 243
- NFS client 297
- NFS problems 297
- NFS V4 297
- NFS V4 issues 297
- NFS V4 problem 297
- no replication 285
- NO\_SPACE error 261
- nodes will not start 234
- NSD creation failure 278
- NSD disk does not have an NSD server specified 258
- NSD information 276
- NSD is down 278
- NSD server 248
- NSD subsystem failures 275
- NSDs built on top of AIX logical volume is down 282
- offline mmfsck command failure 135
- old inode data 297
- on-disk data 260
- Operating system error logs 157
- partial disk failure 281
- permission denied error message 259
- permission denied failure 294
- Persistent Reserve errors 287
- physical disk association 133
- physical disk association with logical volume 133
- policies 263, 264
- predicted pool utilizations 263
- problem determination 213, 218, 222, 225
- problem determination hints 132
- problem determination tips 132
- problems not directly related to snapshots 268
- problems while working with Samba in 305
- problems with locating a snapshot 268
- problems with non-IBM disks 283
- protocol cluster disaster recovery issues 326
- protocol service logs 152, 155, 167
- quorum nodes in cluster 133
- RAS events 371
- Reliability 371
- remote cluster name 256
- remote command issues 228, 229
- remote file system 255, 256
- remote file system does not mount 255, 256
- remote file system I/O failure 255



- GPFS (*continued*)
    - remote mount failure 259
    - replicated data 284
    - replicated metadata 284, 285
    - replication 280, 285
    - Requeuing message 274
    - requeuing of messages in AFM 274
    - restoring a snapshot 270, 271
    - Samba 305
    - security issues 294
    - Serviceability 371
    - set up 161
    - setup issues 271
    - SMB server health 304
    - snapshot directory name conflict 269
    - snapshot problems 267, 268
    - snapshot status errors 269
    - snapshot usage errors 268, 269
    - some files are 'ill-placed' 263
    - stale inode data 297
    - storage pools 266, 267
    - strict replication 285
    - system load increase in night 134
    - timeout executing function error message 135
    - trace facility 162
    - tracing the mmpmon command 273
    - troubleshooting 213, 218, 222, 225
    - UID mapping 255
    - unable to access disks 279
    - unable to determine if a file system is mounted 259
    - unable to start 226
    - underlying disk subsystem failures 275
    - understanding Persistent Reserve 287
    - unmount failure 249
    - unused underlying multipath device 290
    - upgrade failure 361
    - upgrade recovery 361
    - usage errors 263, 266
    - using mmpmon 271
    - value to large failure 293
    - value to large failure while creating a file 293
    - varyon problems 283
    - volume group 283
    - volume group on each node 282
    - Windows file system 134
    - Windows issues 241
    - working with Samba 305
  - GPFS cluster
    - problems adding nodes 227
    - recovery from loss of GPFS cluster configuration data files 230
  - GPFS cluster data
    - locked 229
  - GPFS cluster data files storage 230
  - GPFS command
    - failed 237
    - return code 237
    - unsuccessful 237
  - GPFS commands
    - mmpmon 4
    - unsuccessful 237
  - GPFS configuration data 260
  - GPFS configuration parameters
    - low values 330
  - GPFS daemon 228, 232, 233, 245, 249
    - crash 236
    - fails to start 233
  - GPFS daemon (*continued*)
    - went down 158, 236
    - will not start 232
  - GPFS daemon went down 236
  - GPFS failure
    - network failure 243
  - GPFS GUI logs 174
  - GPFS is not using the underlying multipath device 290
  - GPFS kernel extension 232
  - GPFS local node failure 257
  - GPFS log 233, 234, 236, 245, 248, 250, 254, 256, 257, 258, 367
  - GPFS logs
    - master GPFS log file 141
  - GPFS messages 444
  - GPFS modules
    - cannot be loaded 232
    - unable to load on Linux 232
  - GPFS problems 213, 245, 275
  - GPFS startup time 141
  - GPFS support for SNMP 105
  - GPFS trace facility 162
  - GPFS Windows SMB2 protocol (CIFS serving) 241
  - gpfs.snap 180
  - gpfs.snap command 367
    - data always gathered for a master snapshot 179
    - data always gathered on AIX 177
    - data always gathered on all platforms 176
    - data always gathered on Linux 178
    - data always gathered on Windows 179
    - using 176
  - grep command 157
  - Group Services
    - verifying quorum 234
  - GROUP\_ID attribute 198, 199
  - gui 80
    - event notifications 79
    - snmp 80
  - GUI
    - capacity information is not available 352
    - directed maintenance procedure 362
    - displaying outdated information 350
    - DMP 362
    - GUI fails to start 347
    - GUI issues 347
    - limitations 347
    - login page does not open 348
    - logs 174
    - performance monitoring 69
    - performance monitoring issues 348
    - server was unable to process the request error 350
    - support matrix 347
    - system health
      - overview 79
  - GUI login page does not open 348
  - GUI logs 174
  - GUI refresh tasks 350
- ## H
- hard loop ID 228
  - HDFS
    - transparency log 152
  - health
    - transparent cloud tiering service 127
  - health monitoring
    - features 85

- Health monitoring
  - features 187
- hints and tips for GPFS problems 132
- Home and .ssh directory ownership and permissions 241

## I

- I/O failure
  - remote file system 255
- I/O hang 291
- I/O operations slow 158
- IBM Spectrum Protect client 273
- IBM Spectrum Protect server 273
  - MAXNUMMP 273
- IBM Spectrum Protectd server 273
- IBM Spectrum Scale 3, 5, 6, 7, 9, 10, 11, 15, 16, 17, 18, 19, 20, 21, 22, 23, 25, 26, 28, 29, 30, 31, 32, 34, 35, 40, 41, 43, 44, 45, 60, 61, 62, 77, 107, 108, 109, 110, 111, 112, 113, 114, 115, 187, 190, 298, 307, 308, 335, 340
  - /tmp/mmfs directory 134
  - automount fails to mount on Linux 253
  - abnormal termination in mmpmon 272
  - active file management 135
  - active tracing 342
  - add new nodes 330
  - AIX 254
  - AIX platform 177
  - application calls 231
  - application program errors 239, 240
  - audit messages 141
  - Authentication
    - error events 295
    - errors 295
  - authentication issues 294
  - authentication on Linux 182
  - authorization issues 228
  - automount fails to mount on AIX 254
  - automount failure 254
  - automount failure in Linux 253
  - Automount file system 253
  - Automount file system will not mount 253
  - back up data 130
  - best practices for troubleshooting 129
  - CES NFS
    - failure 297
    - network failure 297
  - CES tracing
    - debug data collection 167
  - checking Persistent Reserve 288
  - cipherList option has not been set properly 258
  - clearing a leftover Persistent Reserve reservation 288
  - client nodes 248
  - cluster configuration
    - issues 229, 230
    - recovery 230
  - cluster crash 226
  - cluster name 256
  - cluster state information 187, 188, 189, 190
  - collecting details of issues 160
  - command 187
  - commands 187, 188, 189, 190
  - connectivity problems 229
  - contact node address 256
  - contact nodes down 257
  - core dumps 161
  - corrupted data integrity 274
  - creating a file 293

- IBM Spectrum Scale (*continued*)
  - data always gathered 176
  - data gathered 177, 179, 180, 182
    - Object on Linux 181
  - data gathered for CES on Linux 182
  - data gathered for core dumps on Linux 183
  - data gathered for performance 183
  - data gathered for protocols on Linux 180, 181, 182, 183
  - data gathered for SMB on Linux 180
  - data integrity may be corrupted 274
  - deadlock breakup
    - on demand 211
  - deadlocks 207, 210, 211
  - default parameter value 330
  - deployment problem determination 213, 218, 222, 225
  - deployment troubleshooting 213, 218, 222, 225
  - determining the health of integrated SMB server 304
  - disaster recovery issues 325
  - discrepancy between GPFS configuration data and the on-disk data for a file system 260
  - disk accessing commands fail to complete 283
  - disk connectivity failure 281
    - failover to secondary server 336
  - disk failure 286
  - disk information commands 191, 192, 193, 201, 204
  - disk media failure 284, 285
  - disk media failures 283
  - disk recovery 281
  - displaying NSD information 276
  - dumps 139
  - encryption issues 293
  - encryption rules 293
  - error creating internal storage 135
  - error encountered while creating and using NSD disks 275
  - error log 157, 158
  - error message for file system 247, 248
  - error messages 213
  - error message "Function not implemented" 255
  - error numbers 248
  - error numbers for GPFS application calls 260
  - error numbers specific to GPFS application calls 248, 274, 286
  - Error numbers specific to GPFS application calls 252
  - error numbers specific to GPFS application calls when data integrity may be corrupted 274
  - error numbers when a file system mount is unsuccessful 248
  - errors associated with filesets 261
  - errors associated with policies 261
  - errors associated with storage pools 261
  - errors encountered 267
  - errors encountered while restoring a snapshot 270, 271
  - errors encountered with filesets 265
  - errors encountered with policies 264
  - errors encountered with storage pools 267
  - event notifications 129
  - failure analysis 213
  - failure group considerations 251
  - failures using the mmbackup command 273
  - file system block allocation type 334
  - file system commands 191, 192, 193, 201, 204
  - file system does not mount 255
  - file system fails to mount 245
  - file system fails to unmount 249
  - file system forced unmount 250
  - file system is forced to unmount 252

IBM Spectrum Scale (continued)

- file system is known to have adequate free space 261
- file system is mounted 259
- file system manager appointment fails 260
- file system manager failures 260
- file system mount problems 247, 248
- file system mount status 259
- file system mounting on wrong drive 135
- file system update 344
- File system utilization 337
- file systems manager failure 259, 260
- fileset update 344
- filesets usage errors 265
- Ganesha NFSD process not running (nfsd\_down) 301
- GPFS cluster security configurations 256
- GPFS commands unsuccessful 238
- GPFS configuration parameters 330
- GPFS daemon does not start 235
- GPFS daemon issues 232, 233, 234, 235, 236
- GPFS declared NSD is down 278
- GPFS disk issues 275, 291
- GPFS down on contact nodes 257
- GPFS error message 247
- GPFS error messages 260, 269
- GPFS error messages for disk media failures 286
- GPFS error messages for file system forced unmount problems 252
- GPFS error messages for file system mount status 259
- GPFS error messages for mmbakup errors 273
- GPFS failure
  - network issues 243
- GPFS file system issues 245
- GPFS has declared NSDs built on top of AIX logical volume as down 282
- GPFS is not running on the local node 257
- GPFS modules
  - unable to load on Linux 232
- gpfs.snap 176, 177, 178, 179
  - gpfs.snap command 178
  - Linux platform 178
- gpfs.snap command
  - usage 176
- guarding against disk failures 280
- GUI logs 174
- hang in mmpmon 272
- HDFS transparency log 152
- health monitoring 85
- hints and tips for problem determination 132
- hosts file issue 227
- IBM Spectrum Protect error messages 273
- incorrect output from mmpmon 272, 371
- installation and configuration issues 213, 226, 227, 229, 231, 232, 233, 234, 235, 236, 237, 238, 239, 241, 250
- key rewrap 294
- limitations 137
- logical volume 133
- logical volumes are properly defined for GPFS use 282
- logs 139, 140
  - GPFS log 140, 141
  - NFS logs 148
    - protocol service logs 146
    - syslog 141
- lsof command 192
- maintenance commands 340
- maintenance commands execution 341
- manually disabling Persistent Reserve 289
- manually enabling Persistent Reserve 289

IBM Spectrum Scale (continued)

- master snapshot 179
- message 6027-648 134
- message referring to an existing NSD 278
- message requeuing in AFM 274
- message severity tags 443
- messages 443
- mixed OS levels 222
- mmafmctl Device getstate 187
- mmapplypolicy -L 0 command 194
- mmapplypolicy -L 1 command 195
- mmapplypolicy -L 2 command 195
- mmapplypolicy -L 3 command 196
- mmapplypolicy -L 4 command 197
- mmapplypolicy -L 5 command 198
- mmapplypolicy -L 6 command 199
- mmapplypolicy -L command 194, 195, 196, 197, 198, 199
- mmapplypolicy command 193
- mmdumpperfdata command 184
- mmfileid command 201
- MMFS\_DISKFAIL 157
- MMFS\_ENVIRON
  - error log 158
- MMFS\_FSSTRUCT error log 158
- MMFS\_GENERIC error log 158
- MMFS\_LONGDISKIO 158
- mmfsadm command 186
- mmhealth 85
- mmlscluster command 188
- mmlsconfig command 189
- mmlsmount command 192
- mmrefresh command 189
- mmremotecluster command 256
- mmsdrrestore command 190
- mmwindisk command 201
- mount 253, 255
- mount failure 248
- mount failure as the client nodes joined before NSD servers 248
- mount failure for a file system 293
- mounting cluster does not have direct access to the disks 258
- multiple file system manager failures 259, 260
- negative values occur in the 'predicted pool utilizations', 263
- net use on Windows fails 308
- network issues 243
- newly mounted windows file system is not displayed 134
- NFS
  - client access exported data 303
  - client cannot mount NFS exports 303
  - client I/O temporarily stalled 303
  - error events 301
  - error scenarios 303
  - errors 301, 303
- NFS client 297
  - client access exported data 303
  - client cannot mount NFS exports 303
  - client I/O temporarily stalled 303
- NFS is not active (nfs\_not\_active) 301
- NFS on Linux 180
- NFS problems 297
- NFS V4 issues 297
- nfs\_not\_active 301
- nfsd\_down 301
- no replication 285
- NO\_SPACE error 261

## IBM Spectrum Scale *(continued)*

- NSD and underlying disk subsystem failures 275
- NSD creation fails 278
- NSD disk does not have an NSD server specified 258
- NSD server 248
- NSD server failure 335
- NSD-to-server mapping 332
- offline mmfsck command failure 135
- old NFS inode data 297
- operating system error logs 157, 158, 159, 160
- operating system logs 157, 158, 159, 160
- other problem determination tools 206
- partial disk failure 281
- Password invalid 306
- performance issues 344
- permission denied error message 259
- permission denied failure 294
- Persistent Reserve errors 287
- physical disk association 133
- policies 263
- Portmapper port 111 is not active (portmapper\_down) 301
- portmapper\_down 301
- prerequisites 218
- problem determination 132, 213, 218, 222, 225
- problems while working with Samba 305
- problems with locating a snapshot 268
- problems with non-IBM disks 283
- protocol cluster disaster recovery issues 326
- protocol service logs 152, 155, 167
  - object logs 149
- QoSIO operation classes 331
- quorum loss 239
- quorum nodes 133
- quorum nodes in cluster 133
- recovery procedures 361
- remote cluster name 256
- remote cluster name does not match with the cluster name 256
- remote command issues 228, 229
- remote file system 255, 256
- remote file system does not mount 255, 256
- remote file system does not mount due to differing GPFS cluster security configurations 256
- remote file system I/O fails with "Function not implemented" error 255
- remote file system I/O failure 255
- remote mounts fail with the "permission denied" error 259
- remote node expelled from cluster 250
- replicated metadata 285
- replicated metadata and data 284
- replication setting 343
- requeuing of messages in AFM 274
- RPC statd process is not running (statd\_down) 301
- security issues 294
- set up 161
- setup issues while using mmpmon 271
- SHA digest 204
- SMB
  - access issues 310
  - error events 309
  - errors 309
- SMB client on Linux fails 306
- SMB service logs 146
- snapshot 344
- snapshot directory name conflict 269
- snapshot problems 267
- snapshot status errors 269

## IBM Spectrum Scale *(continued)*

- snapshot usage errors 268, 269
- some files are 'ill-placed' 263
- SSSD process not running (sssd\_down) 295
- stale inode data 297
- statd\_down 301
- storage pools usage errors 266
- strict replication 285
- support for troubleshooting 367
- System error 86 308
- system load 134
- timeout executing function error message 135
- trace facility 162
- trace reports 162
- traces 139
- tracing the mmpmon command 273
- transparent cloud tiering service
  - managing 127
- troubleshooting 130, 213, 218, 222, 225
  - best practices 131
  - collecting issue details 139, 140
  - getting started 129
  - report problems 131
  - resolve events 131
  - support notifications 131
  - update software 131
- UID mapping 255
- unable to access disks 279
- unable to determine if a file system is mounted 259
- unable to resolve contact node address 256
- understanding Persistent Reserve 287
- unused underlying multipath device by GPFS 290
- upgrade failure 361
- upgrade recovery 361
- usage errors 263
- user does not exist 306
- value to large failure 293
- VERBS RDMA
  - inactive 338
- volume group on each node 282
- volume group varyon problems 283
- warranty and maintenance 131
- Winbind process not running (wnbd\_down) 295
- Windows 179
- Windows issues 241
- Wrong version of SMB client 306
- YPBIND process not running (yp\_down) 295
- IBM Spectrum Scale commands
  - mmpmon 4
- IBM Spectrum Scale information units xi
- IBM Spectrum Scale support for SNMP 105
- IBM Spectrum Scale time stamps 139
- IBM Spectrum Scale collector node
  - administration 107
- IBM Spectrum Scale command
  - mmapfctl Device getstate 187
- IBM Spectrum Scale mmdiag command 187
- ill-placed files 263, 267
- ILM
  - problems 261
- improper mapping 332
- inode data
  - stale 297
- inode limit 160
- installation and configuration issues 213
- installation problems 226

- installing MIB files on the collector and management node,
  - SNMP 107
- installing Net-SNMP 105
- investigationperformance problems 64
- io\_s 7
- issues
  - mmprotocoltrace 174

## J

- junctions
  - deleting 265

## K

- KB\_ALLOCATED attribute 198, 199
- kdb 206
- KDB kernel debugger 206
- kernel module
  - mmfslinux 232
- kernel panic 291
- kernel threads
  - at time of system hang or panic 206
- key rewrap failure 294

## L

- Linux kernel
  - configuration considerations 227
  - crash dump facility 206
- Linux on z Systems 227, 368
- logical volume 282
  - location 133
- Logical Volume Manager (LVM) 280
- logical volumes 282
- logs 140
  - GPFS log 140, 141
  - NFS logs 148
  - object logs 149
  - protocol service logs 146
  - SMB logs 146
  - syslog 141
- logsIBM Spectrum Scale
  - performance monitoring logs 161
- long waiters
  - increasing the number of inodes 344
- lspp command 368
- lslv command 133
- lsdf command 192, 249, 250
- lspv command 282
- lsvg command 282
- lxtrace command 162, 186

## M

- maintenance commands
  - mmadddisk 340
  - mmapplypolicy 340
  - mmdeldisk 340
  - mmrestripefs 340
- maintenance operations
  - execution 341
- management and monitoring, SNMP subagent 108
- management node, installing MIB files 107

- managing
  - transparent cloud tiering service 127
- manually enabling or disabling Persistent Reserve 289
- maxblocksize parameter 247
- MAXNUMMP 273
- memory shortage 157, 228
- message 6027-648 134
- message severity tags 443
- messages 443, 444
  - 6027-1941 227
- metadata
  - replicated 284, 285
- metrics
  - performance monitoring 46
  - performance monitoring tool
    - defining 64
- MIB files, installing on the collector and management node 107
- MIB objects, SNMP 109
- MIGRATE rule 194, 196
- migration
  - file system will not mount 246
  - new commands will not run 238
- mmadddisk command 260, 266, 281, 283, 284
- mmaddnode command 133, 227
- mmafmctl command 274
- mmafmctl Device getstate command 187
- mmapplypolicy -L 0 194
- mmapplypolicy -L 1 195
- mmapplypolicy -L 2 195
- mmapplypolicy -L 3 196
- mmapplypolicy -L 4 197
- mmapplypolicy -L 5 198
- mmapplypolicy -L 6 199
- mmapplypolicy command 193, 263, 264, 266, 294
- mmauth command 204, 256
- mmbackup command 273
- mmccr command
  - failure 135
- mmchcluster command 228
- mmchconfig command 189, 233, 250, 258
- mmchdisk command 246, 260, 266, 275, 278, 279, 284, 287
- mmcheckquota command 159, 199, 240, 251
- mmchfs command 160, 231, 238, 246, 248, 251, 253, 305, 345
- mmchnode 107, 108
- mmchnode command 133
- mmchnsd command 275
- mmchpolicy
  - issues with adding encryption policy 293
- mmcommon 253, 254
- mmcommon breakDeadlock 211
- mmcommon recoversfs command 261
- mmcommon showLocks command 230
- mmcrcluster command 133, 189, 227, 228, 233
- mmcrfs command 238, 275, 283, 305
- mmcrnsd command 275, 278
- mmcrsnapshot command 268, 269
- mmdefedquota command fails 134
- mmdeldisk command 260, 266, 281, 284
- mmdelfileset command 265
- mmdelfs command 286
- mmdelnode command 227, 238
- mmdelnsd command 278, 286
- mmdelsnapshot command 269
- mmdf command 261, 281, 344
- mmdiag command 187
- mmdsh command 229

- mmdumperpdata 184
- mmedquota command fails 134
- mmexpelnode command 190
- mmfileid command 201, 274, 284
- MMFS\_ABNORMAL\_SHUTDOWN
  - error logs 157
- MMFS\_DISKFAIL
  - error logs 157
- MMFS\_ENVIRON
  - error logs 158
- MMFS\_FSSTRUCT
  - error logs 158
- MMFS\_GENERIC
  - error logs 158
- MMFS\_LONGDISKIO
  - error logs 158
- MMFS\_QUOTA
  - error log 159
  - error logs 159, 199
- MMFS\_SYSTEM\_UNMOUNT
  - error logs 159
- MMFS\_SYSTEM\_WARNING
  - error logs 160
- mmfs.log 233, 234, 236, 245, 248, 250, 254, 256, 257, 258, 367
- mmfsadm command 166, 186, 234, 274, 284, 345
- mmfsck command 192, 245, 246, 266, 274, 281, 285
  - failure 135
- mmfsd 232, 233, 245, 249
  - will not start 232
- mmfslinux
  - kernel module 232
- mmgetstate command 187, 234, 237
- mmlock directory 229
- mmlsattr command 264, 265
- mmlscluster 107
- mmlscluster command 133, 188, 227, 257
- mmlsconfig command 162, 189, 253
- mmlsdisk command 238, 245, 246, 251, 260, 275, 278, 279, 283, 287, 368
- mmlsfileset command 265
- mmlsfs command 247, 284, 286, 368
- mmlsmgr command 186, 246
- mmlsmount command 192, 233, 240, 245, 249, 250, 275
- mmlsnsd command 200, 276, 282
- mmlspolicy command 264
- mmlsquota command 240
- mmlssnapshot command 267, 268, 269, 270
- mmmount command 192, 245, 251, 283
- mmperfmon 64
- mmpmon 3
  - abend 272
  - adding nodes to a node list 8
  - altering input file 272
  - concurrent processing 5
  - concurrent usage 272
  - counters 41
  - counters wrap 272
  - deleting a node list 10
  - deleting nodes from node list 12
  - disabling the request histogram facility 20
  - displaying a node list 11
  - displaying statistics 5, 6, 25, 28, 30
  - dump 272
  - enabling the request histogram facility 21
  - examples 32
    - failure 14
    - fs\_io\_s 5, 12
- mmpmon (*continued*)
  - examples (*continued*)
    - io\_s 7
    - nlist add 9
    - nlist del 10
    - nlist s 11
    - node shutdown and quorum loss 15
    - once 35
    - reset 15
    - rhist nr 19
    - rhist off 20
    - rhist on 21
    - rhist p 23
    - rhist r 26
    - rhist reset 25
    - rpc\_s 29
    - rpc\_s size 30
    - source 35
    - ver 31
  - failure 14
  - fs\_io\_s 5, 12
    - aggregate and analyze results 32
  - hang 272
  - histogram reset 25
  - I/O histograms 16
  - I/O latency ranges 17
  - incorrect input 271
  - incorrect output 272
  - input 3
  - interpreting results 32
  - interpreting rhist results 34
  - io\_s 6
    - aggregate and analyze results 32
    - latency ranges 17
    - miscellaneous information 40
    - multiple node processing 4
    - new node list 10
    - nlist 4, 8, 12
    - nlist add 8, 9
    - nlist del 10
    - nlist failures 15
    - nlist new 10
    - nlist s 11
    - nlist sub 12
    - node list facility 8, 12
    - node list failure values 15
    - node list show 11
    - node shutdown and quorum loss 15
    - once 34, 35
    - output considerations 40
    - overview 3
    - request histogram 16
    - request histogram facility pattern 22
    - reset 15
    - restrictions 271
    - return codes 41
    - rhist 16
    - rhist nr 18, 19
    - rhist off 20
    - rhist on 21
    - rhist p 22, 23
    - rhist r 26
    - rhist reset 25
    - rhist s 25
    - rpc\_s 28, 29
    - rpc\_s size 30
    - setup problems 271

- mmpmon (*continued*)
  - size ranges 16
  - source 34, 35
  - specifying new ranges 18
  - unsupported features 272
  - version 31
- mmpmon command 206, 271
  - trace 273
- mmprotocoltrace command
  - issues 174
- mmquotaoff command 240
- mmquotaon command 240
- mmrefresh command 189, 246, 253
- mmremotecluster command 204, 256, 257
- mmremotefs command 253, 257
- mmrepquota command 240
- mmrestorefs command 269, 270, 271
- mmrestripefile command 264, 267
- mmrestripefs command 267, 281, 284
- mmrpldisk command 260, 266, 283
- mmsdrrestore command 190
- mmshutdown 108
- mmshutdown command 188, 190, 233, 234, 236, 254
- mmsnapdir command 267, 268, 270
- mmstartup command 233, 254
- mmtracectl command
  - generating GPFS trace reports 162
- mmumount command 249, 251, 281
- mmunlinkfileset command 265
- mmwindisk command 201
- mode of AFM fileset, changing 135
- MODIFICATION\_TIME attribute 198, 199
- module is incompatible 232
- monitor prefetch 101
- monitoring
  - performance 1
- mount
  - problems 248
- mount command 245, 246, 248, 283, 285
- mount error (127)
  - Permission denied 308
- mount error (13)
  - Permission denied 308
- mount failure 293
- mount on Mac fails with authentication error
  - mount\_smbfs: server rejected the connection:
    - Authentication error 308
- mount.cifs on Linux fails with mount error (13)
  - Permission denied 308
- mounting cluster 258
- Mounting file system
  - error messages 247
- Multi-Media LAN Server 139

## N

- Net-SNMP
  - configuring 106
  - installing 105
  - running under SNMP master agent 108
  - traps 115
- network
  - performance
    - Remote Procedure Calls (RPCs) 1
  - network failure 243
  - network problems 158
- NFS 180, 297

- NFS (*continued*)
  - problems 297
- NFS client
  - with stale inode data 297
- NFS client cannot mount exports
  - mount exports, NFS client cannot mount 298
- NFS error events 301
- NFS error scenarios 303
- NFS errors 303
  - Ganesha NFSD process not running (nfsd\_down) 301
  - NFS is not active (nfs\_not\_active) 301
  - nfs\_not\_active 301
  - nfsd\_down 301
  - Portmapper port 111 is not active (portmapper\_down) 301
  - portmapper\_down 301
- NFS logs 148
- NFS mount on client times out
  - NFS mount on server fails 298
  - Permission denied 298
  - time out error 298
- NFS mount on server fails
  - access type is one 298
  - NFS mount on server fails
    - no such file or directory 298
    - protocol version not supported by server 298
- NFS V4
  - problems 297
- NFS, SMB, and Object logs 146
- no replication 285
- NO SUCH DIRECTORY error code 236
- NO SUCH FILE error code 236
- NO\_SPACE
  - error 261
- node
  - crash 369
  - hang 369
  - rejoin 249
- node configuration information, SNMP 110
- node crash 226
- node failure 291
- node reinstall 226
- node status information, SNMP 110
- nodes
  - cannot be added to GPFS cluster 227
- non-quorum node 133
- NSD 282
  - creating 278
  - deleting 278
  - displaying information of 276
  - extended information 277
  - failure 275
- NSD build 282
- NSD disks 258
  - creating 275
  - using 275
- NSD failure 275
- NSD server 248, 249, 258
  - failover to secondary server 335, 336
- nsdServerWaitTimeForMount
  - changing 248
- nsdServerWaitTimeWindowOnMount
  - changing 249
- NT STATUS LOGON FAILURE
  - SMB client on Linux fails 306

## O

- object
  - logs 149
- Object 180
- object IDs
  - SNMP 109
- object metrics
  - proxy server 43, 60, 62
- OpenSSH connection delays
  - Windows 243
- orphaned file 266

## P

- partitioning information, viewing 201
- password must change 307
- performance 180, 228
  - monitoring 1, 3
  - network
    - Remote Procedure Calls (RPCs) 1
- performance issues 327
  - caused by the low-level system components 327
  - due to high utilization of the system-level components 327
  - due to improper system level settings 329
  - due to long waiters 327
  - due to networking issues 328
  - due to suboptimal setup or configuration 329
- performance monitoring
  - GUI performance monitoring issues 348
  - Limitations 77
  - log 161
  - metrics 46
  - performance monitoring through GUI 69
  - queries 64
- performance monitoring tool
  - configuring 43, 60, 61, 62
    - Automated configuration 44
    - Manual configuration 45
  - manual restart 63
  - metrics 46
    - defining 64
  - overview 41
  - queries 64
- performance problemsinvestigation 64
- permission denied
  - remote mounts failure 259
- permission denied failure (key rewrap) 294
- Persistent Reserve
  - checking 288
  - clearing a leftover reservation 288
  - errors 287
  - manually enabling or disabling 289
  - understanding 287
- ping command 229
- PMR 369
- policies
  - DEFAULT clause 263
  - deleting referenced objects 263
  - errors 264
  - file placement 263
  - incorrect file placement 264
  - LIMIT clause 263
  - long runtime 264
  - MIGRATE rule 263
  - problems 261

- policies (*continued*)
  - rule evaluation 263
  - usage errors 263
  - verifying 193
- policy file
  - detecting errors 194
  - size limit 263
  - totals 195
- policy rules
  - runtime problems 264
- POOL\_NAME attribute 198, 199
- possible GPFS problems 213, 245, 275
- predicted pool utilization
  - incorrect 263
- primary NSD server 248
- problem
  - locating a snapshot 268
  - not directly related to snapshot 268
  - snapshot 267, 268
    - snapshot directory name 269
    - snapshot restore 270, 271
    - snapshot status 269
    - snapshot usage 268
    - snapshot usage errors 268
- problem determination
  - cluster state information 187
  - remote file system I/O fails with the "Function not implemented" error message when UID mapping is enabled 255
  - tools 191
  - tracing 162
- Problem Management Record 369
- problems
  - configuration 226
  - installation 226
  - mmbackup 273
- problems running as administrator, Windows 241
- protocol (CIFS serving), Windows SMB2 241
- protocol authentication log 153
- protocol service logs
  - NFS logs 148
  - object logs 149
  - SMB logs 146
- Protocols 294
- proxies
  - performance monitoring tool 41
- proxy server
  - object metrics 43, 60, 62

## Q

- QoSIO operation classes
  - low values 331
- queries
  - performance monitoring 64
- quorum 133, 234
  - disk 239
  - loss 239
- quorum node 133
- quota
  - cannot write to quota file 251
  - denied 240
  - error number 231
- quota files 199
- quota problems 159



## R

- RAID controller 280
- rcp command 228
- read-only mode mount 192
- recovery
  - cluster configuration data 230
- recovery log 291
- recovery procedure
  - restore data and system configuration 361
- recovery procedures 361
- recreation of GPFS storage file
  - mmchcluster -p LATEST 230
- Reliability 371
- remote command problems 228
- remote file copy command
  - default 228
- remote file system
  - mount 256
- remote file system I/O fails with "Function not implemented" error 255
- remote mounts fail with permission denied 259
- remote node
  - expelled 250
- remote node expelled 250
- Remote Procedure Calls (RPCs)
  - network performance 1
- remote shell
  - default 228
- removing the setuid bit 236
- replicated
  - metadata 285
- replicated data 284
- replicated metadata 284
- replication 266
  - of data 280
- replication setting 343
- replication, none 285
- report problems 131
- reporting a problem to IBM 186
- request returns the current values for all latency ranges which have a nonzero count.IBM Spectrum Scale 26
- resetting of setuid/setgits at AFM home 135
- resolve events 131
- restore data and system configuration 361
- restricted mode mount 192
- return codes, mmpmon 41
- RPC statistics
  - aggregation of execution time 28
  - RPC execution time 30
- rpc\_s size 28
- RPCs (Remote Procedure Calls)
  - network performance 1
- rpm command 367
- rsh
  - problems using 228
- rsh command 228, 237

## S

- Samba
  - client failure 305
- scp command 228
- Secure Hash Algorithm digest 204
- sensors
  - performance monitoring tool 41
  - Automated configuration 43

- sensors (*continued*)
  - performance monitoring tool (*continued*)
    - configuring 43, 60
    - configuring, 43
    - File-managed configuration 43
    - GPFS-managed configuration 43
    - Manual configuration 43
- Serviceability 371
- servicing (CIFS), Windows SMB2 protocol 241
- set up
  - core dumps 161
- setuid bit, removing 236
- setuid/setgid bits at AFM home, resetting of 135
- severity tags
  - messages 443
- SHA digest 204, 256
- shared segments 234
  - problems 235
- SMB 180
  - logs 146
- SMB access issues 310
- SMB client on Linux fails 307
- SMB error events 309
- SMB errors 309
- SMB on Linux 180
- SMB server 304
- SMB2 protocol (CIFS servicing), Windows 241
- snapshot
  - directory name conflict 269
  - error messages 268, 269
  - invalid state 269
  - restoring 270, 271
  - status error 269
  - usage error 268
  - valid 268
- snapshot problems 267
- SNMP
  - cluster configuration information 109
  - cluster status information 109
  - collector node administration 107
  - configuring Net-SNMP to work with IBM Spectrum Scale 106
  - configuring SNMP-based management applications 107
  - disk configuration information 113
  - disk performance information 114
  - disk status information 113
  - file system performance information 112
  - file system status information 111
  - GPFS support 105
  - IBM Spectrum Scale support 105
  - installing MIB files on the collector and management node 107
  - installing Net-SNMP on the collector node of the IBM Spectrum Scale 105
  - management and monitoring subagent 108
  - MIB objects 109
  - Net-SNMP traps 115
  - node configuration information 110
  - node status information 110
  - object IDs 109
  - starting and stopping the SNMP subagent 108
  - storage pool information 112
- spectrumscale installation toolkit
  - failure analysis 213
  - prerequisites 218
  - supported configurations 222, 225
  - supported setups 222

- spectrumscale installation toolkit *(continued)*
  - upgrade support 225
- ssh command 228
- statistics
  - network performance 1
- steps to follow
  - GPFS daemon does not come up 233
- storage pool information, SNMP 112
- storage pools
  - deleting 263, 266
  - errors 267
  - failure groups 266
  - problems 261
  - slow access time 267
  - usage errors 266
- strict replication 285
- subnets attribute 250
- suboptimal performance 330, 331, 332, 334, 335, 340, 343
- suboptimal system performance 335, 336, 337, 338, 340, 341, 342, 344
- support for troubleshooting 367
  - call home 370
  - contacting IBM support center 367
    - how to contact IBM support center 369
  - information to be collected before contacting IBM support center 367
- support notifications 131
- syslog 141
- syslog facility
  - Linux 157
- syslogd 254
- system health
  - GUI
    - overview 79
- system load 134
- system snapshots 176
- system storage pool 263, 266
- System z 227, 368

## T

- threads
  - tuning 228
  - waiting 345
- time stamps 139
- trace
  - active file management 163
  - allocation manager 163
  - basic classes 163
  - behaviorals 166
  - byte range locks 163
  - call to routines in SharkMsg.h 165
  - checksum services 163
  - cleanup routines 163
  - cluster security 165
  - concise vnop description 166
  - daemon routine entry/exit 164
  - daemon specific code 165
  - data shipping 164
  - defragmentation 164
  - dentry operations 164
  - disk lease 164
  - disk space allocation 163
  - DMAPI 164
  - error logging 164
  - events exporter 164
  - file operations 164

- trace *(continued)*
  - file system 164
  - generic kernel vfs information 164
  - inode allocation 164
  - interprocess locking 164
  - kernel operations 164
  - kernel routine entry/exit 164
  - low-level vfs locking 164
  - mailbox message handling 164
  - malloc/free in shared segment 164
  - miscellaneous tracing and debugging 165
  - mmpmon 164
  - mnode operations 165
  - mutexes and condition variables 165
  - network shared disk 165
  - online multinode fsck 164
  - operations in Thread class 165
  - page allocator 165
  - parallel inode tracing 165
  - performance monitors 165
  - physical disk I/O 164
  - physical I/O 164
  - pinning to real memory 165
  - quota management 165
  - rdma 165
  - recovery log 164
  - SANergy 165
  - scsi services 165
  - shared segments 165
  - SMB locks 165
  - SP message handling 165
  - super operations 165
  - tasking system 165
  - token manager 165
  - ts commands 164
  - vdisk 166
  - vdisk debugger 166
  - vdisk hospital 166
  - vnode layer 166
- trace classes 163
- trace facility 162, 163
  - mmfsadm command 186
- trace level 166
- trace reports, generating 162
- tracing
  - active 342
- transparent cloud tiering
  - common issues and workarounds 359
  - troubleshooting 359
- transparent cloud tiering logs
  - collecting 160
- transparent cloud tiering service
  - health monitoring 127
- traps, Net-SNMP 115
- traversing a directory that has not been cached 136
- troubleshooting
  - AFM DR issues 357
  - best practices 129, 130
    - report problems 131
    - resolve events 131
    - support notifications 131
    - update software 131
  - capacity information is not available in GUI pages 352
  - CES 155
  - collecting issue details 139
  - disaster recovery issues 325
    - setup problems 325

- troubleshooting (*continued*)
  - event notifications 129
  - getting started 129
  - GUI fails to start 347
  - GUI is displaying outdated information 350
  - GUI issues 347
  - GUI login page does not open 348
  - GUI logs 174
  - GUI performance monitoring issues 348
  - logs 140
    - GPFS log 140, 141
    - syslog 141
  - performance issues 327
    - caused by the low-level system components 327
    - due to high utilization of the system-level components 327
    - due to improper system level settings 329
    - due to long waiters 327
    - due to networking issues 328
    - due to suboptimal setup or configuration 329
  - protocol cluster disaster recovery issues 326
  - recovery procedures 361
  - server was unable to process the request 350
  - support for troubleshooting 367
    - call home 370
  - transparent cloud tiering 359
  - warranty and maintenance 131
- troubleshooting errors 241
- troubleshooting Windows errors 241
- tuning 228

## U

- UID mapping 255
- umount command 250, 251
- unable to start GPFS 234
- underlying multipath device 290
- understanding, Persistent Reserve 287
- unsuccessful GPFS commands 237
- usage errors
  - policies 263
- useNSDserver attribute 281
- USER\_ID attribute 198, 199
- using the gpfs.snap command 176

## V

- v 228
- value too large failure 293
- varyon problems 283
- varyonvg command 283
- VERBS RDMA
  - inactive 338
- viewing disks and partitioning information 201
- volume group 282

## W

- warranty and maintenance 131
- Windows 241
  - data always gathered 179
  - file system mounted on the wrong drive letter 135
  - gpfs.snap 179
  - Home and .ssh directory ownership and permissions 241
  - mounted file systems, Windows 134
  - OpenSSH connection delays 243

- Windows (*continued*)
  - problem seeing newly mounted file systems 134
  - problem seeing newly mounted Windows file systems 134
  - problems running as administrator 241
  - Windows 135
- Windows issues 241
- Windows SMB2 protocol (CIFS serving) 241

## Z

- z Systems 227, 368







Product Number: 5725-Q01  
5641-GPF  
5725-S28

Printed in USA

GA76-0443-08

