Multiprotocol File Sharing on IBM Spectrum Scale without an AD or LDAP server

Implementation Guide

Abstract

This document provides steps to set up local authentication on IBM® Spectrum Scale. The setup can be used to provide NFS and CIFS file shares to small or isolated environments without using external discrete AD or LDAP authentication.

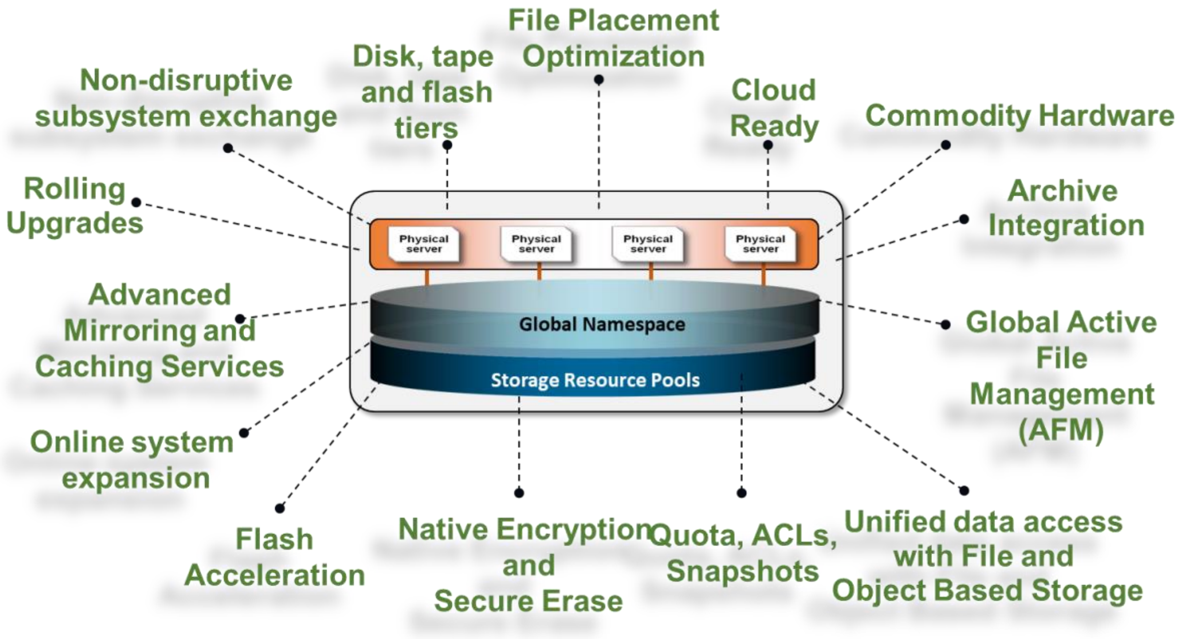
Rakesh Chutke, Kiran Ghag

rachutke@in.ibm.com, kiran.ghag@in.ibm.com

Table of Contents

1 Introduction to IBM Spectrum Scale protocol services

IBM Spectrum Scale is feature-rich clustered file system (formerly GPFS) that has evolved over the years. GPFS initially facilitated clustered shared file system for multimedia and/or HPC environments. These environments required parallel access to a large number of files, in terms of both size and count. Over the years, additional features were added to base GPFS to support a variety of workloads and applications.


File Placement Optimization
Disk, tape and flash tiers
Non-disruptive subsystem exchange
Cloud Ready
Commodity Hardware
Rolling Upgrades
Archive Integration
Physical server
Physical server
Physical server
Physical server
Global Namespace
Storage Resource Pools
Advanced Mirroring and Caching Services
Global Active File Management (AFM)
Online system expansion
Flash Acceleration
Native Encryption and Secure Erase
Quota, ACLs, Snapshots
Unified data access with File and Object Based Storage

IBM Spectrum Scale protocol services can export an existing GPFS file system to end clients using CIFS and NFS protocols. These are typical file sharing options available on Windows and UNIX servers, respectively.

Protocol services provide a great way to move data in and out of the Spectrum Scale file system. Remote or local clients can access the file system without being part of the cluster. Cluster nodes can also access the file system, and protocol services transform the file system into a unified storage system for file data.

2 Protocol services using local authentication

Protocol services requires a central authentication system to authenticate the users. Once a user is authenticated, it can provide authorization to restrict/provide access to files using Access Control Lists.

Authentication is supported through Microsoft Active Directory (AD) or Lightweight Directory Access Protocol (LDAP).

Using AD with RFC 2307 support or LDAP, protocol services can export data to UNIX and Windows clients. It also allows a given share to be accessed simultaneously through a UNIX host and a Windows host.

This document covers the procedure to set up Spectrum Scale protocol services for UNIX and Windows users.

2.1 Authenticating NAS users and sharing files without existing AD/LDAP

Local authentication can be set up on an IBM Spectrum Scale cluster if there aren’t any existing AD or LDAP servers. This eliminates the need to set up a Windows AD server or an LDAP server just for providing authentication services.

2.2 Authenticating NAS users and sharing files along with existing AD/LDAP

Spectrum Scale local authentication can be used in addition to the existing AD or LDAP authentication. NAS access can be provided to selected users without any configuration change on the existing infrastructure.

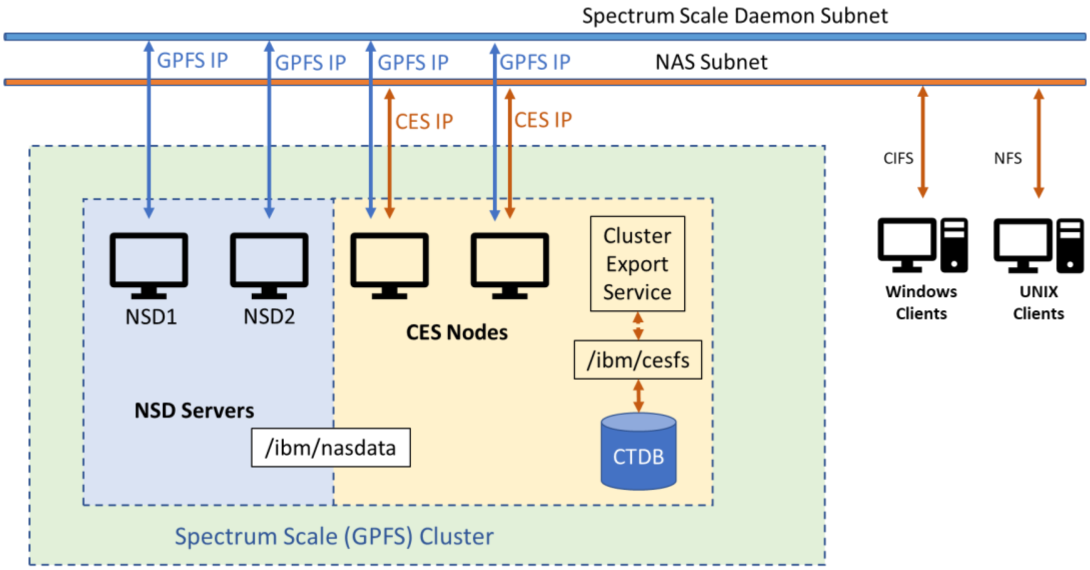
3 Intended use cases

The solution described in the paper is useful for small environments and proof of concept (POC) deployments. The solution can be used to provide NAS sharing using NFS and CIFS for few application servers. The solution can be used for a small to medium sized workgroup environments to share data between users.

This can be deployed quickly, without the need to set up and maintain a central authentication separately.

4 Test system architecture

A functional Spectrum Scale cluster is required to run protocol services. For smaller setups, protocol services can be run on the NSD servers directly. Separate CES nodes and NSD servers provide flexibility in system tuning, and it's recommended to keep them separate.



Our test setup is comprised of four Spectrum Scale nodes — two NSD servers and two CES nodes.

NSD servers will have direct attached connectivity to NAS file systems, and protocol nodes can access it over the NSD network/Spectrum Scale daemon subnet.

Each node has the following configuration. However, it is recommended to go for a higher configuration for a production workload.

	PoC Configuration	Typically Recommended
vCPUs	1	8 or more
System RAM	4 GB	32 GB or more
Disk space for OS	20 GB	100 GB or more
Operating System	RHEL 7.6	(Consult the latest Scale FAQ)
Spectrum Scale version	5.0.3.0 DA Edition	(Consult the latest Scale FAQ)
Network Adapter	1 x 1 Gbps shared port	Separate port(s) for cluster and NAS communication
Shared Root File system	/dev/sdd (8 GB)	8 GB or more
NAS File system	20 GB - /dev/sdb (10GB), /dev/sdc (10 GB)	Depending on users
NSD server nodes	nsd1 (192.168.175.128), nsd2 (192.168.175.129)	-
Protocol nodes	protocol1 (192.168.175.130), protocol2 (192.168.175.131)	-
CES (NAS) IPs	gpfsnas 192.168.175.201,192.168.175.202	

5 Spectrum Scale Cluster Implementation Using Toolkit

The IBM Spectrum Scale Installation Toolkit enables quick deployment of a Spectrum Scale cluster, especially if protocol and GUI services are required. It can also be used to add these services to an existing cluster.

The toolkit is a set of scripts that come with Spectrum Scale binaries. After extracting scale binaries, the toolkit is extracted at /usr/lpp/mmfs/x.x.x.x/Installer (x.x.x.x is the version of Scale binaries being extracted).

5.1 Node preparation for Installation Toolkit

Perform the following steps on each node before using the installation toolkit:

- Disable SELinux
- Disable firewall
- Add /usr/lpp/mmfs/bin in PATH for root user
- Set "StrictHostKeyChecking no" in /etc/ssh/ssh_config
- Setup passwordless SSH access between all nodes
- Create a REPO using RHEL installation ISO
- Install pre-requisite packages (net-tools, ntp, libc.so.6, ksh, kernel-devel, cpp, gcc, gcc-c++, kernel-devel and m4)
- Configure NTP client and sync time with server
- Update /etc/hosts to include IP address and hostname for all Spectrum Scale nodes (including shortname and FQDN)

A snippet of /etc/hosts file is given below. Note that NAS IPs use same hostname.

```
192.168.175.128 nsd1 nsd1.testdomain.com
192.168.175.129 nsd2 nsd2.testdomain.com
192.168.175.130 protocol1 protocol1.testdomain.com
192.168.175.131 protocol2 protocol2.testdomain.com
192.168.175.200 gpfsnas gpfsnas.testdomain.com
192.168.175.201 gpfsnas gpfsnas.testdomain.com
```

All host IP addresses and names can be resolved through DNS too, but it can cause disruption to service if DNS servers are not reachable.

A reboot may be required at the end to apply SELinux change.

5.2 Set up Control Node

The installation toolkit needs to be configured on one node first. This node is called Control Node. The toolkit needs to be provided cluster configuration using the “spectrumscale” script. This script then pushes binaries on other nodes and completes configuration.

Let’s set up node nsd1 as a control node and prepare it for configuration:

```
./spectrumscale setup -i ~/.ssh/id_rsa -s 192.168.175.128
```

The SSH key provided here must be set up on all nodes for password-less access.

Sample command output is given below:

```
[root@nsd1 installer]# ./spectrumscale setup -i ~/.ssh/id_rsa -s 192.168.175.128
[ INFO  ] Installing prerequisites for install node
[ INFO  ] Your control node has been configured to use the IP 192.168.175.128 to communicate with other nodes.
[ INFO  ] The key located at /root/.ssh/id_rsa, will be used during install.
[ INFO  ] Port 8889 will be used for chef communication.
[ INFO  ] Port 10080 will be used for package distribution.
[ INFO  ] Install Toolkit setup type is set to Spectrum Scale (default). If an ESS is in the cluster, run this command to set ESS mode:
./spectrumscale setup -s server_ip -st ess
[ INFO  ] SUCCESS
[ INFO  ] Tip : Designate protocol, nsd and admin nodes in your environment to use during install:./spectrumscale -v node add
<node> -p  -a -n
```

5.3 Disable callhome (unless needed)

Callhome is not being configured since this is a PoC system.

```
[root@nsd1 installer]# ./spectrumscale callhome disable
[ INFO  ] Disabling the callhome.
[ INFO  ] Configuration updated.
```

5.4 Define cluster name and daemon ports

The following commands will set the cluster name and provide the TCP port range used for cluster communication.

```
[root@nsd1 installer]# ./spectrumscale config gpfs -c gpfsnas -e 60000-61000
[ INFO  ] Setting GPFS cluster name to gpfsnas
[ INFO  ] Setting GPFS Daemon communication port range to 60000-61000
```

5.5 Define nodes

The following commands are used to add nodes to the Spectrum Scale cluster.

```
./spectrumscale node add -m -n -a -g -q nsd1
./spectrumscale node add -m -n -a -g -q nsd2
./spectrumscale node add -p -q protocol1
./spectrumscale node add -p protocol2
```

Various options define node properties based on cluster designs:

- m – node can do manager role and be a file system or token manager
- n – node can be NSD owner and has direct connections to NSDs through SAN
- a – node can perform admin functions
- g – node will run GUI service
- q – node will perform as a quorum node

Sample command output is given below:

```
[root@nsd1 installer]# ./spectrumscale node add -m -n -a -g -q nsd1
[ INFO  ] Adding node nsd1 as a GPFS node.
[ INFO  ] Adding node nsd1 as a quorum node.
[ INFO  ] Adding node nsd1 as a manager node.
[ INFO  ] Adding node nsd1 as an NSD server.
[ INFO  ] Configuration updated.
[ INFO  ] Tip :If all node designations are complete, add NSDs to your cluster definition and define required
filessytems:./spectrumscale nsd add <device> -p <primary node> -s <secondary node> -fs <file system>
[ INFO  ] Setting nsd1 as an admin node.
[ INFO  ] Configuration updated.
[ INFO  ] Tip : Designate protocol or nsd nodes in your environment to use during install:./spectrumscale node add <node> -p -n
[ INFO  ] Setting nsd1 as a GUI server.
[root@nsd1 installer]# ./spectrumscale node add -m -n -a -g -q nsd2
[ INFO  ] Adding node nsd2 as a GPFS node.
[ INFO  ] Adding node nsd2 as a quorum node.
[ INFO  ] Adding node nsd2 as a manager node.
[ INFO  ] Adding node nsd2 as an NSD server.
[ INFO  ] Configuration updated.
[ INFO  ] Tip :If all node designations are complete, add NSDs to your cluster definition and define required
filessytems:./spectrumscale nsd add <device> -p <primary node> -s <secondary node> -fs <file system>
[ INFO  ] Setting nsd2 as an admin node.
[ INFO  ] Configuration updated.
[ INFO  ] Tip : Designate protocol or nsd nodes in your environment to use during install:./spectrumscale node add <node> -p -n
[ INFO  ] Setting nsd2 as a GUI server.
[root@nsd1 installer]# ./spectrumscale node add -p protocol1
[ INFO  ] Adding node protocol1 as a GPFS node.
[ INFO  ] Adding node protocol1 as a quorum node.
[ INFO  ] Setting protocol1 as a protocol node.
[ INFO  ] Configuration updated.
[ INFO  ] Tip : If all node designations are complete, configure the protocol environment as needed: ./spectrumscale config
protocols -f cesSharedRoot -m /ibm/cesSharedRoot
[root@nsd1 installer]# ./spectrumscale node add -p protocol2
[ INFO  ] Adding node protocol2 as a GPFS node.
[ INFO  ] Setting protocol2 as a protocol node.
[ INFO  ] Configuration updated.
[ INFO  ] Tip : If all node designations are complete, configure the protocol environment as needed: ./spectrumscale config
protocols -f cesSharedRoot -m /ibm/cesSharedRoot
```

5.6 Create NSD definitions and file systems

Assign shared drives to the NSD servers and use the installer toolkit to prepare definitions, using the “nsd add” option.

```
./spectrumscale nsd add -p nsd1 -s nsd2 -fs nasdata -u dataAndMetadata -name nasdatansd01 /dev/sdb
./spectrumscale nsd add -p nsd2 -s nsd1 -fs nasdata -u dataAndMetadata -name nasdatansd02 /dev/sdc
./spectrumscale nsd add -p nsd1 -s nsd2 -fs cesfs -u dataAndMetadata -name cesfsnsd01 /dev/sdd
```

Two disks are being added in the nasdata file system, each 10 GB in size. One 8 GB disk is being added in the cesfs file system.

Sample command output is given below:

```
[root@nsd1 installer]# ./spectrumscale nsd add -p nsd1 -s nsd2 -fs nasdata -u dataAndMetadata -name nasdatansd01 /dev/sdb
[ INFO  ] Connecting to nsd1 to check devices and expand wildcards.
[ INFO  ] Looking up details of /dev/sdb.
[ INFO  ] The installer will create the new file system nasdata if it does not exist.
[ INFO  ] Adding NSD nasdatansd01 on nsd1 using device /dev/sdb.
[ INFO  ] Configuration updated
[ INFO  ] Tip : If all node designations and any required protocol configurations are complete, proceed to check the installation
configuration: ./spectrumscale install --precheck
[root@nsd1 installer]# ./spectrumscale nsd add -p nsd2 -s nsd1 -fs nasdata -u dataAndMetadata -name nasdatansd02 /dev/sdc
[ INFO  ] Connecting to nsd2 to check devices and expand wildcards.
[ INFO  ] Looking up details of /dev/sdc.
[ INFO  ] Adding NSD nasdatansd02 on nsd2 using device /dev/sdc.
[ INFO  ] Configuration updated
[ INFO  ] Tip : If all node designations and any required protocol configurations are complete, proceed to check the installation
configuration: ./spectrumscale install --precheck
[root@nsd1 installer]# ./spectrumscale nsd add -p nsd1 -s nsd2 -fs cesfs -u dataAndMetadata -name cesfsnsd01 /dev/sdd
[ INFO  ] Connecting to nsd1 to check devices and expand wildcards.
[ INFO  ] Looking up details of /dev/sdd.
[ INFO  ] The installer will create the new file system cesfs if it does not exist.
[ INFO  ] Adding NSD cesfsnsd01 on nsd1 using device /dev/sdd.
[ INFO  ] Configuration updated
[ INFO  ] Tip : If all node designations and any required protocol configurations are complete, proceed to check the installation
configuration: ./spectrumscale install –precheck
```

5.7 Verify file system parameters

The toolkit automatically builds file system definitions based on details provided for NSD creation. Verify the details by the following command. You can make changes to mountpoints and other details if required.

```
[root@nsd1 installer]# ./spectrumscale filesystem list
[ INFO  ] Name    BlockSize   Mountpoint   NSDs Assigned  Default Data Replicas    Max Data Replicas     Default Metadata Replicas
Max Metadata Replicas
[ INFO  ] nasdata Default (4M)/ibm/nasdata 2            1                    2                1                      2
[ INFO  ] cesfs   Default (4M)/ibm/cesfs   1            1                    2                1                      2
```

5.8 CES service definitions

We need to instruct the toolkit to configure SMB and NFS sharing on the cluster. This requires the cesSharedRoot file system to be made available to the CES service and virtual CES IPs to be provided.

```
./spectrumscale config protocols -f cesfs -m /ibm/cesfs/ -e 192.168.175.200,192.168.175.201
./spectrumscale enable smb
./spectrumscale enable nfs
```

The CES service will automatically distribute CES IPs to each protocol node and use the cesSharedRoot file system for storing cluster export information.

Sample command output is given below:

```
[root@nsd1 installer]# ./spectrumscale config protocols -f cesfs -m /ibm/cesfs/ -e 192.168.175.200,192.168.175.201
[ INFO  ] Setting Shared File System Name to cesfs
[ INFO  ] Setting Shared File System Mountpoint or Path to /ibm/cesfs
[ INFO  ] Setting Export IP Pool to 192.168.175.200,192.168.175.201
[ INFO  ] Tip :Enable NFS, Object or SMB protocols as appropriate:./spectrumscale enable nfs|object|smb
[root@nsd1 installer]# ./spectrumscale enable smb
[ INFO  ] Enabling SMB on all protocol nodes.
[ INFO  ] Tip :If all node designations and any required protocol configurations are complete, proceed to check the installation
configuration:./spectrumscale deploy --precheck
[root@nsd1 installer]# ./spectrumscale enable nfs
[ INFO  ] Enabling NFS on all protocol nodes.
[ INFO  ] Tip :If all node designations and any required protocol configurations are complete, proceed to check the installation
configuration:./spectrumscale deploy –precheck
```

5.9 Run installation toolkit

The installation toolkit commands executed earlier do not make any configuration changes by themselves. The commands only store the definitions in a file called configuration\clusterdefinition.ext.

The installation script needs to be run with install and deploy options to initiate the cluster creation.

5.9.1 Install steps

Install option creates the cluster, NSDs and installs the GUI.

```
[root@nsd1 installer]# ./spectrumscale install
[ INFO  ] Logging to file: /usr/lpp/mmfs/5.0.3.0/installer/logs/INSTALL-03-12-2019_21:58:11.log
[ INFO  ] Validating configuration
[ INFO  ] Running pre-install checks
[ INFO  ] Running environment checks
[ INFO  ] GPFS precheck OK
[ INFO  ] Running environment checks for Performance Monitoring
[ INFO  ] Running environment checks for file  Audit logging
[ INFO  ] Preparing nodes for install
[ INFO  ] Installing Chef (deploy tool)
[ INFO  ] Installing Chef Client on nodes
[ INFO  ] Installing GPFS
[ INFO  ] GPFS Packages to be installed: gpfs.base, gpfs.gpl, gpfs.msg.en_US, gpfs.docs, and gpfs.gskit
[ INFO  ] [nsd1 03-12-2019 22:02:57] IBM SPECTRUM SCALE: Generating node description file for cluster configuration (SS03)
[ INFO  ] [nsd1 03-12-2019 22:02:57] IBM SPECTRUM SCALE: Creating GPFS cluster with default profile (SS04)
[ INFO  ] [nsd1 03-12-2019 22:03:13] IBM SPECTRUM SCALE: Setting client licenses (SS09)
[ INFO  ] [nsd1 03-12-2019 22:03:13] IBM SPECTRUM SCALE: Setting server licenses on manager nodes (SS12)
[ INFO  ] [nsd1 03-12-2019 22:04:06] IBM SPECTRUM SCALE: Starting GPFS (SS05)
[ INFO  ] Installing NSDs
[ INFO  ] [nsd1 03-12-2019 22:06:18] IBM SPECTRUM SCALE: Generating stanza file for NSD creation (SS14)
[ INFO  ] [nsd1 03-12-2019 22:06:21] IBM SPECTRUM SCALE: Creating NSDs (SS16)
[ INFO  ] Installing GUI
[ INFO  ] GUI packages to be installed: gpfs.gui
[ INFO  ] Checking state of GPFS on all nodes
[ INFO  ] GPFS active on all nodes
[ INFO  ] GPFS ACTIVE
[ INFO  ] Checking state of NSDs
[ INFO  ] NSDs ACTIVE
[ INFO  ] Checking state of Performance Monitoring
[ INFO  ] Running Performance Monitoring post-install checks
[ INFO  ] pmcollector running on all nodes
[ INFO  ] pmsensors running on all nodes
[ INFO  ] Performance Monitoring ACTIVE
[ INFO  ] Checking state of GUI
[ INFO  ] Running Graphical User Interface  post-install checks
[ INFO  ] Graphical User Interface running on all GUI servers
[ INFO  ] Enter one of the following addresses into a web browser to access the Graphical User Interface: nsd1, nsd2
[ INFO  ] GUI ACTIVE
[ INFO  ] SUCCESS
[ INFO  ] All services running
[ INFO  ] StanzaFile and NodeDesc file for NSD, filesystem, and cluster setup have been saved to /usr/lpp/mmfs folder on node:
nsd1
[ INFO  ] Installation successful. 4 GPFS nodes active in cluster gpfsnas.nsd1. Completed in 21 minutes 15 seconds.
[ INFO  ] Tip :If all node designations and any required protocol configurations are complete, proceed to check the deploy
configuration:./spectrumscale deploy –precheck
```

5.9.2 Deploy steps

Deploy options creates the file system and also configures CES services.

```
[root@nsd1 installer]# ./spectrumscale deploy
[ INFO  ] Logging to file: /usr/lpp/mmfs/5.0.3.0/installer/logs/DEPLOY-03-12-2019_22:24:07.log
[ INFO  ] Validating configuration
[ WARN  ] File Authentication will not be configured on your nodes. Run ./spectrumscale auth --help for information on auth.
[ INFO  ] Running pre-install checks
[ INFO  ] NSDs are in a valid state
[ INFO  ] Running environment checks for protocols
[ INFO  ] Checking state of GPFS on all nodes
[ INFO  ] GPFS active on all nodes
[ INFO  ] Checking state of GPFS on all nodes
[ INFO  ] GPFS active on all nodes
[ INFO  ] protocol precheck OK
[ INFO  ] Running environment checks for SMB
[ INFO  ] SMB precheck OK
[ INFO  ] Running environment checks for NFS
[ INFO  ] NFS precheck OK
[ INFO  ] [nsd1 03-12-2019 22:28:29] IBM SPECTRUM SCALE: Enabling CES (SS26)
[ INFO  ] [nsd1 03-12-2019 22:29:00] IBM SPECTRUM SCALE: Check CES enabled on all nodes (SS145)
[ INFO  ] [nsd1 03-12-2019 22:30:21] IBM SPECTRUM SCALE: Assigning additional Export IPs (SS144)
[ INFO  ] [nsd1 03-12-2019 22:30:37] IBM SPECTRUM SCALE: Rebalance CES IPs (SS145)
[ INFO  ] Installing SMB
[ INFO  ] [protocol1 03-12-2019 22:32:26] IBM SPECTRUM SCALE: Enabling SMB service (SS41)
[ INFO  ] Installing NFS
[ INFO  ] [protocol1 03-12-2019 22:34:52] IBM SPECTRUM SCALE: Enabling NFS service (SS63)
[ INFO  ] Checking for a successful install
[ INFO  ] Checking state of Chef (deploy tool)
[ INFO  ] Chef (deploy tool) ACTIVE
[ INFO  ] Checking state of Filesystem
[ INFO  ] File systems have been created successfully
[ INFO  ] Filesystem ACTIVE
[ INFO  ] Checking state of Cluster Export Services
[ INFO  ] Checking state of CES on all nodes
[ INFO  ] CES healthy on all nodes
[ INFO  ] Cluster Export Services ACTIVE
[ INFO  ] Checking state of SMB
[ INFO  ] Running SMB post-install checks
[ INFO  ] Checking state of SMB on all nodes
[ INFO  ] SMB healthy on all nodes
[ INFO  ] SMB ACTIVE
[ INFO  ] Checking state of NFS
[ INFO  ] Running NFS post-install checks
[ INFO  ] Checking state of NFS on all nodes
[ INFO  ] NFS healthy on all nodes
[ INFO  ] NFS ACTIVE
[ INFO  ] Checking state of Performance Monitoring
[ INFO  ] Running Performance Monitoring post-install checks
[ INFO  ] pmcollector running on all nodes
[ INFO  ] pmsensors running on all nodes
[ INFO  ] Performance Monitoring ACTIVE
[ INFO  ] Checking state of GUI
[ INFO  ] Running Graphical User Interface  post-install checks
[ INFO  ] Graphical User Interface running on all GUI servers
[ INFO  ] Enter one of the following addresses into a web browser to access the Graphical User Interface: nsd1, nsd2
[ INFO  ] GUI ACTIVE
[ INFO  ] SUCCESS
[ INFO  ] All services running
[ INFO  ] Successfully installed and configured protocols. 2 protocol nodes were enabled. Components installed: Chef (deploy tool),
Filesystem, Cluster Export Services, SMB, NFS, Performance Monitoring, GUI, FILE AUDIT LOGGING. It took 17 minutes 9 seconds.
```

Base cluster setup here is complete. Verify that cluster nodes are in active state.

```
[root@nsd1 ~]# mmgetstate -aL

 Node number  Node name       Quorum  Nodes up  Total nodes  GPFS state   Remarks
-------------------------------------------------------------------------------------
       1      nsd1               2        2          4       active       quorum node
       2      nsd2               2        2          4       active       quorum node
       3      protocol1          2        2          4       active       quorum node
       4      protocol2          2        2          4       active
```

Verify that both file systems are mounted.

```
[root@nsd1 installer]# mmlsmount all -L

File system cesfs is mounted on 4 nodes:
  192.168.175.129 nsd2
  192.168.175.128 nsd1
  192.168.175.130 protocol1
  192.168.175.131 protocol2

File system nasdata is mounted on 4 nodes:
  192.168.175.128 nsd1
  192.168.175.129 nsd2
  192.168.175.131 protocol2
  192.168.175.130 protocol1
```

5.10 Set up local authentication database

CES uses the clustered version of Samba and Ganesha services. Samba provides CIFS or SMB shares for Windows users. Ganesha provides NFS exports for UNIX users. The clustered version relies on the CTDB backend to store state information required for IP failover and user authentication. When a central authentication schema is used, the authentication is performed externally.

With local authentication, the Samba daemon acts like a Windows Workgroup server to the clients. Local authentication uses the underlying OS's user database (/etc/passwd). Local authentication database setup is simple and can be done by the following command:

```
[root@protocol1 ~]# mmuserauth service create --data-access-method file --type userdefined
File authentication configuration completed successfully.
```

5.10.1 NFS ID mapping

For NFS users, the user authentication is provided through client IP based filtering. Client provides user ID and Group ID whenever it tries to read/write file data. Spectrum Scale file system accepts this UID+GID and creates all files with the same credentials.

Local authentication hence requires UID and GID to be unique across all servers and clients. The user may not be present on the protocol nodes, but it needs to be present on client nodes. If the user is not present on the protocol node, the directory listing output (ls) on the protocol node cannot resolve the names and only shows numerical UID/GID.

This whole arrangement can be troublesome once users and clients grow, but the only way out is to avoid access permissions issues. Hence for a larger setup, central authentication is recommended.

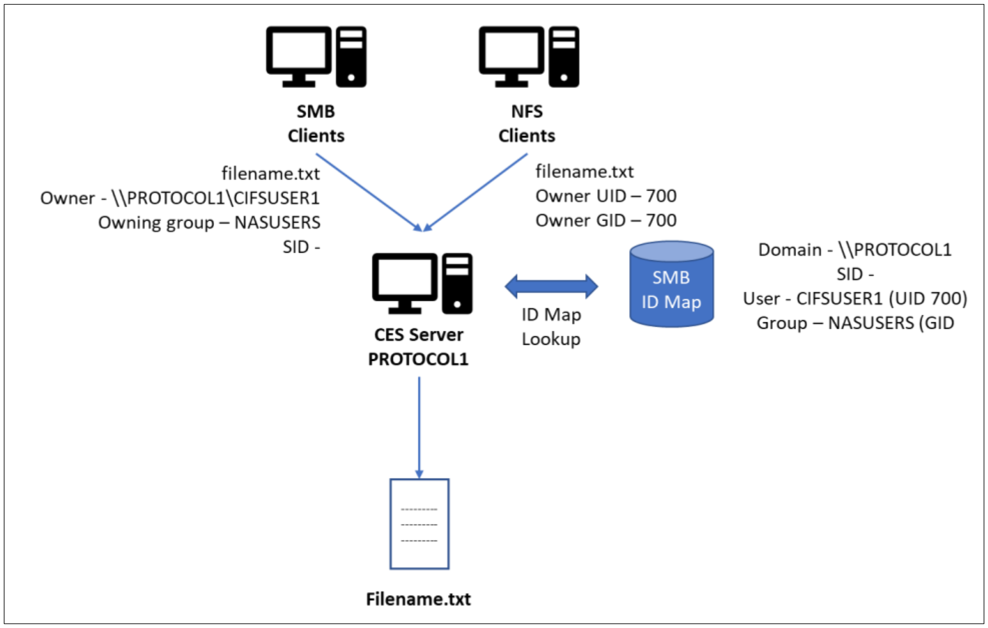
In addition to this, NFS option root_squash is recommended while configuring NFS exports.

5.10.2 SMB and multiprotocol ID mapping

Each local user is considered as a Samba user, and clients need to provide a username in the format “SERVERNAME\USERNAME” when they try to connect. SERVERNAME here can be any CES IP address or the hostname of a protocol node itself (should be resolvable on client using HOSTS file or DNS).

Internally, each SMB user is attached to the UNIX UID, UNIX GID and locally generated Windows SID. This UID and GID is visible to NFS users. SMB users see the Windows SID. This allows multiprotocol access to each file.

The following diagram explains this ID mapping:



5.10.3 Adding users and groups in Samba locally

SMB authentication relies on the local UNIX user database; however, it maintains the passwords in a separate location (/usr/lpp/mmfs/bin/smbpasswd).

Clients can get connected to any protocol node since the CES IPs can be served by any node. Hence a user needs to be created on all protocol nodes. While creating users, UID and GID needs to be specified manually to ensure the same UID and GID is attached to each user. Use mmdsh to add a user and group, as shown below.

```
[root@protocol1 ~]# mmdsh -N cesNodes "groupadd -g 700 nasusers"
[root@protocol1 ~]# mmdsh -N cesNodes "useradd -u 701 -g 700 cifsuser1"
```

Multiple groups can be created to create granular ACLs similar to the way expected on a central server. Just make sure that the same IDs are used across all nodes and deletions/changes are done on all of them.

The Samba password needs to be assigned manually using the smbpasswd command. This command needs to be run manually on each protocol node.

```
[root@protocol1 ~]# /usr/lpp/mmfs/bin/smbpasswd -a cifsuser1
New SMB password:
Retype new SMB password:
Added user cifsuser1.
[root@protocol1 ~]# ssh protocol2
[root@protocol2 ~]# /usr/lpp/mmfs/bin/smbpasswd -a cifsuser1
New SMB password:
Retype new SMB password:
```

User cifsuser1 can now connect to shares created by CES.

5.10.4 User ID creation summary

The following guidelines summarize the care to be taken for seamless multiprotocol access:

- Create user and group on all nodes manually as UNIX user (protocol nodes and NFS clients)
- Assign identical UID and GID to UNIX/NFS users on protocol nodes
- Assign identical UID and GID to UNIX/NFS users on all NFS client nodes
- Update SMB password for UNIX/NFS users on all protocol nodes if SMB access is needed

6 File system sharing

Once users are created, file system sharing can be completed by performing the following steps:

- Prepare file system for data sharing by setting up quotas and ACLS
- Create file set for the share
- Enable NFS or CIFS share on the file set
- Configure access to clients (NFS)
- Map export from NFS client or map drive letter from Windows client

6.1 File system preparation for sharing

The file system should have nfs4 ACLs enabled for sharing data. This is a one-time change and can be done while there is existing data on the file system.

It is also recommended to enable quotas, as they can help control and track capacity usage by clients.

```
[root@protocol1 New folder]#  mmchfs nasdata -Q yes
mmchfs: Propagating the cluster configuration data to all
  affected nodes.  This is an asynchronous process.
[root@protocol1 New folder]#  mmchfs nasdata --filesetdf --perfileset-quota
mmchfs: Propagating the cluster configuration data to all
  affected nodes.  This is an asynchronous process.
[root@protocol1 New folder]#  mmchfs nasdata -k nfs4
```

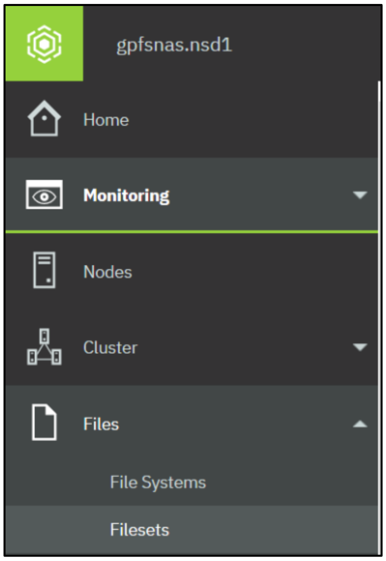
The option perfileset-quota allow us to set different quotas for same user across different file sets. The filesetdf option is useful to show correct free space to users; otherwise the system reports total file system free capacity to users irrespective of set quota limit.

6.2 File set creation

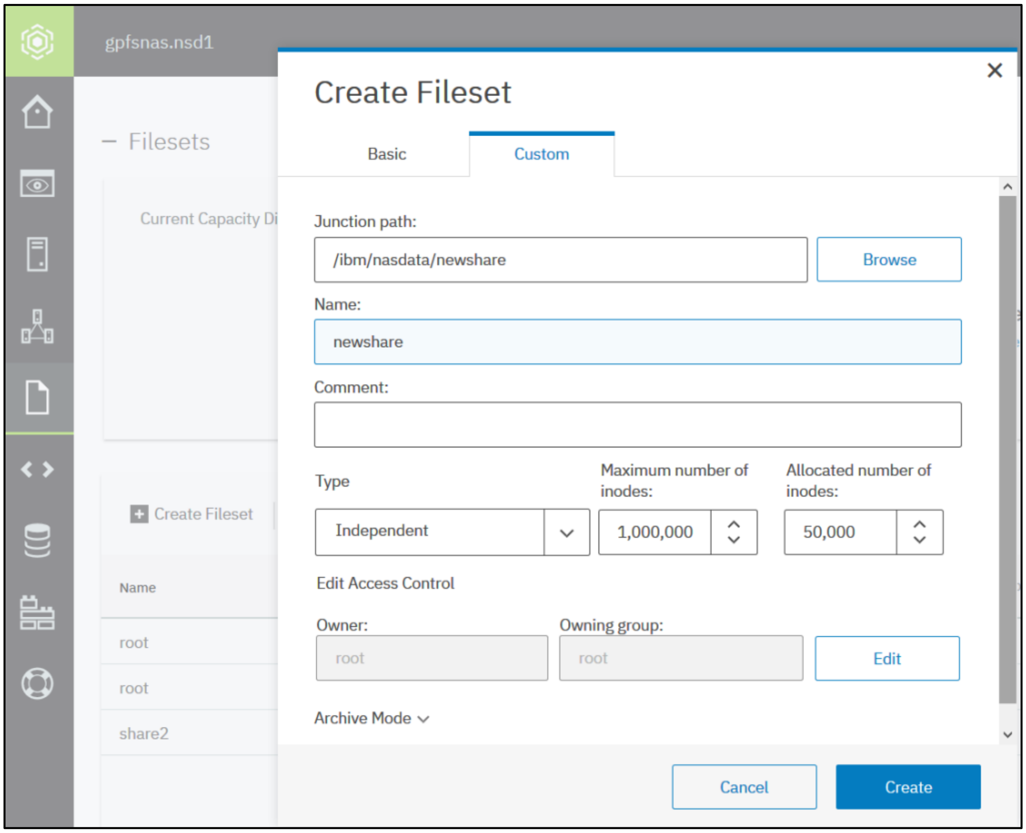
Any folder inside the file system can be shared, but it is recommended to create a new file set and export it. This helps set up quotas for each export and limit the maximum capacity in each.

Connect to the GUI for the following steps:

- Navigate to “**Filesets”** in the “**Files”** section



- Click “**Create Fileset**” button to invoke the file set creation dialog box



Junction path: /ibm/nasdata/newshare

Name: newshare

Comment:

Type: Independent

Maximum number of inodes: 1,000,000

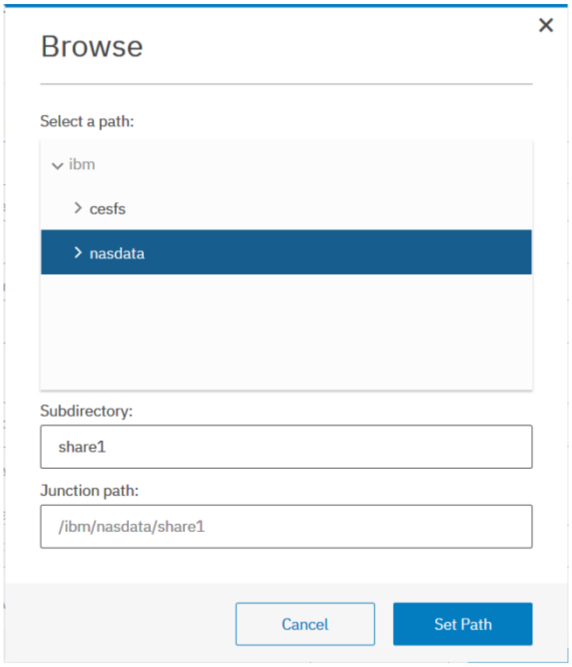
Allocated number of inodes: 50,000

Owner: root

Owning group: root

- Select Type as “**Independent**”

- Click “**Browse**” and select the path for the new file set, and type file set name in “**Subdirectory**” box. It will be automatically created:



Browse

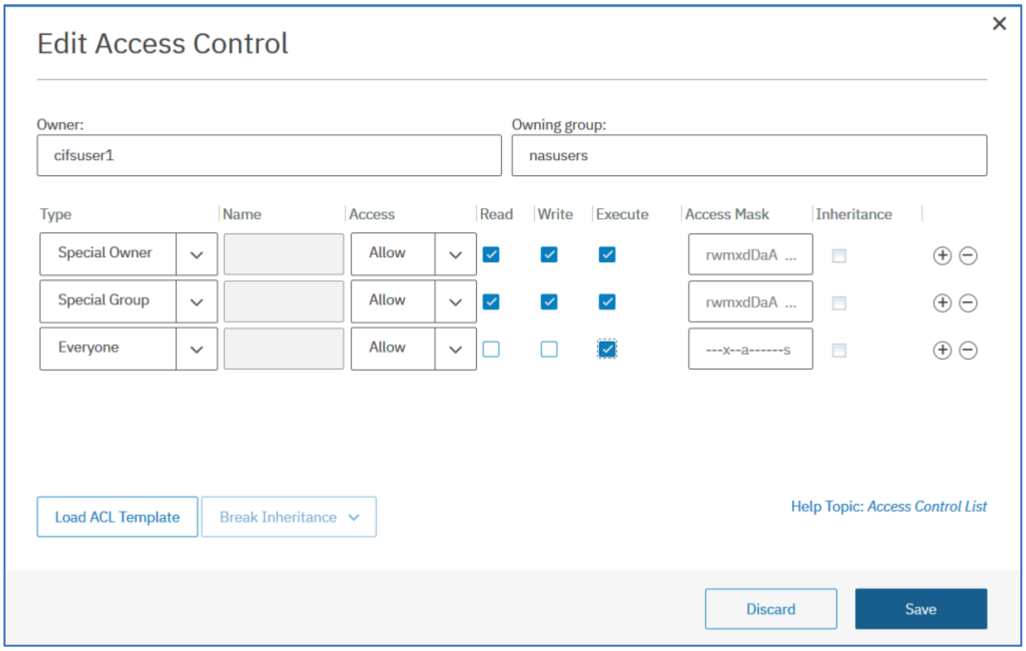
Select a path:

- ibm
  - cesfs
  - nasdata

Subdirectory: share1

Junction path: /ibm/nasdata/share1

- Click the “**Edit**” button to update the owner and group for the file set folder



Edit Access Control

Owner: cifsuser1

Owning group: nasusers

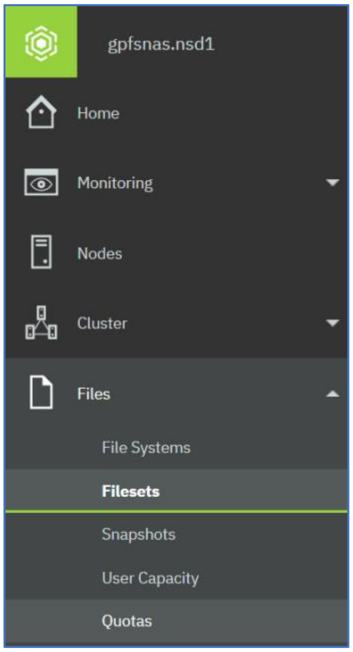
Type	Name	Access	Read	Write	Execute	Access Mask	Inheritance
Special Owner		Allow	☑	☑	☑	rwmxdDaA ...	□
Special Group		Allow	☑	☑	☑	rwmxdDaA ...	□
Everyone		Allow	□	□	☑	---x--a------s	□

- Select Type as “**Independent”** and click “**Create”** to continue

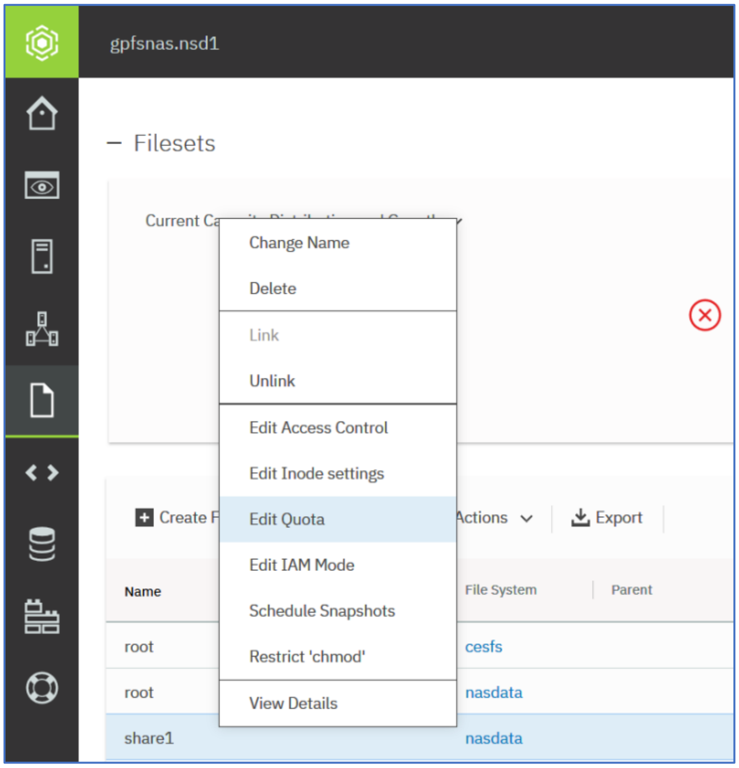
6.3 Assign quota to new file set

Perform the following steps to assign quota to the newly created file set.

- Navigate to "**Quotas"** in the "**Files"** section

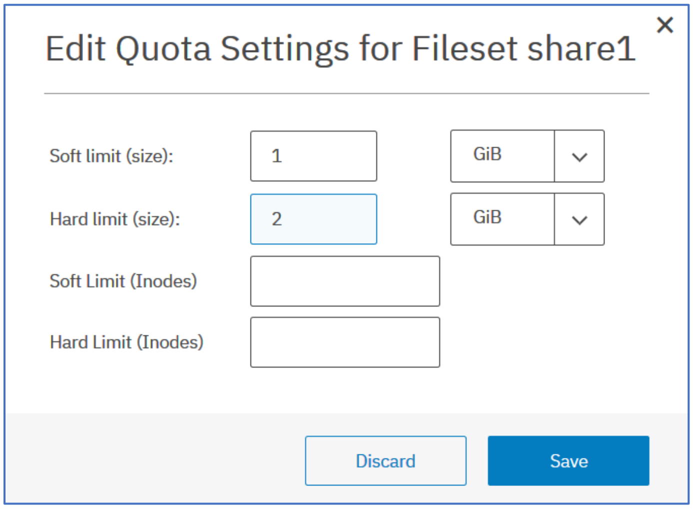


- Right-click the newly created file set and select **Edit Quota**



Name	File System	Parent
root	cesfs	
root	nasdata	
share1	nasdata	

- Enter size of the share to be created in “**Hard Limit”** and click “**OK”** to continue. You can assign a “**Soft Limit”** that is lower than the maximum capacity to receive proactive alerts.



Edit Quota Settings for Fileset share1

Soft limit (size): 1 GiB

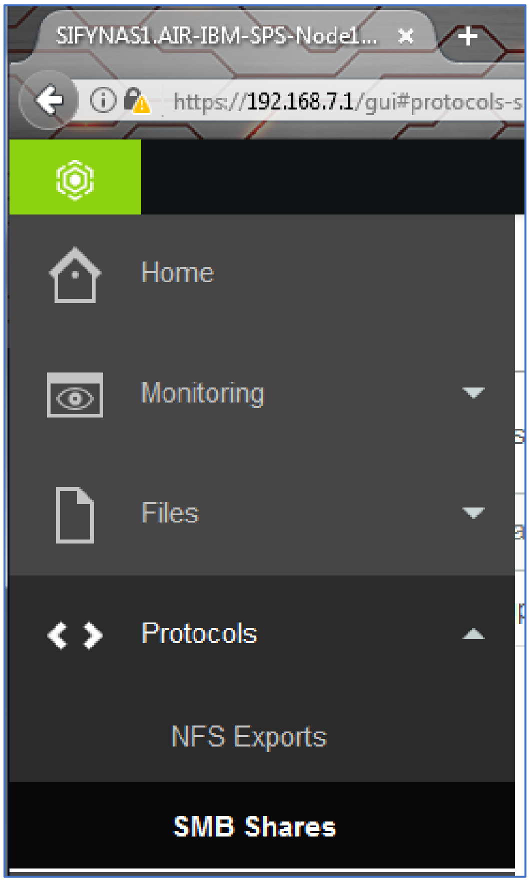
Hard limit (size): 2 GiB

Soft Limit (Inodes)

Hard Limit (Inodes)

6.4 Create share for SMB users

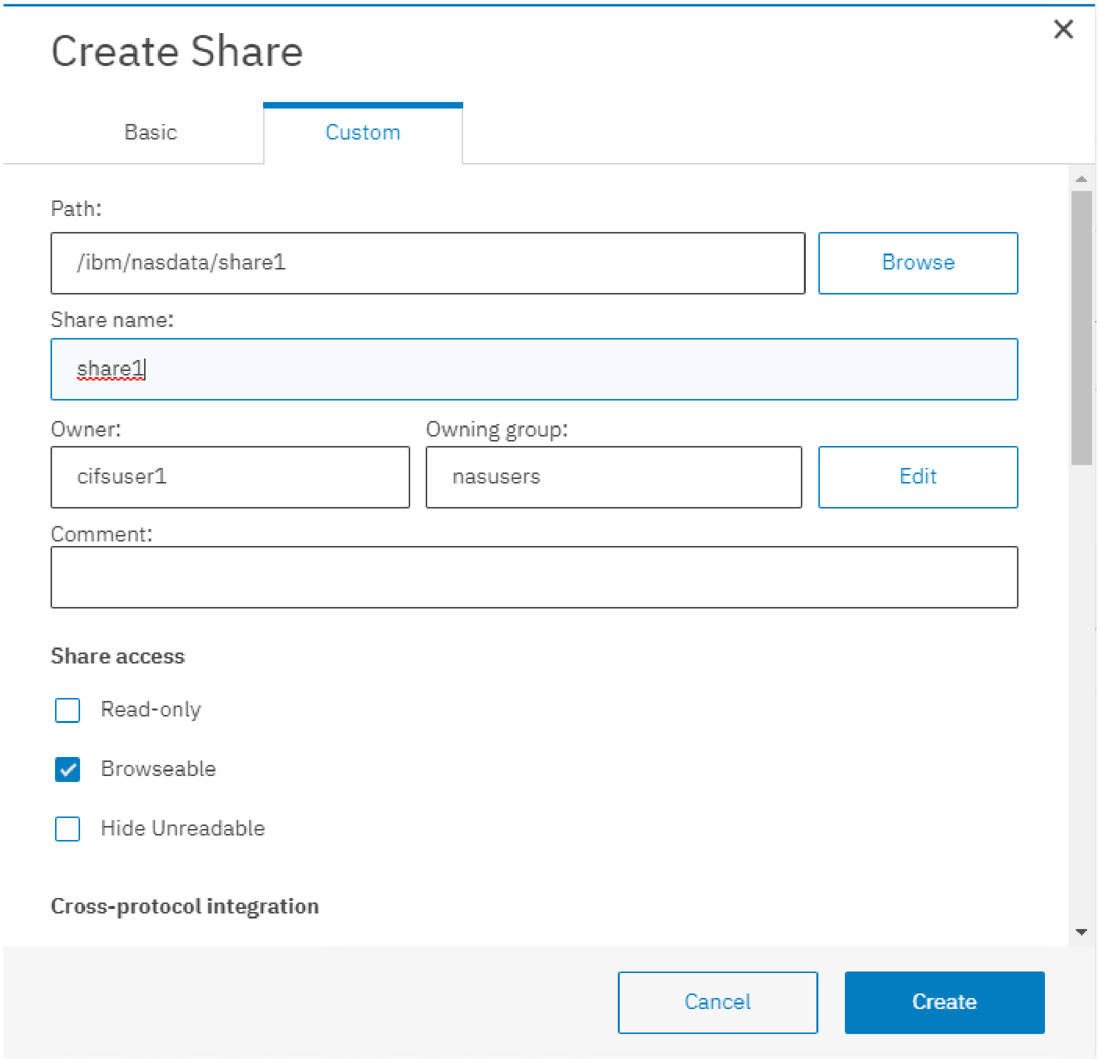
- Navigate to “**SMB Shares”** in the “**Protocols”** section



SIFYNAS1.AIR-IBM-SPS-Node1...
https://192.168.7.1/gui#protocols-s
Home
Monitoring
Files
Protocols
NFS Exports
SMB Shares

- Click “**Create Share**” and select “**Custom”**

- Select newly created file set path from the “**Browse”** button
- Edit ownership and permission information using “**Edit”** and click “**OK**” to complete share creation



Create Share

Path: /ibm/nasdata/share1

Share name: share1

Owner: cifsuser1

Owning group: nasusers

Comment:

Share access

☐ Read-only

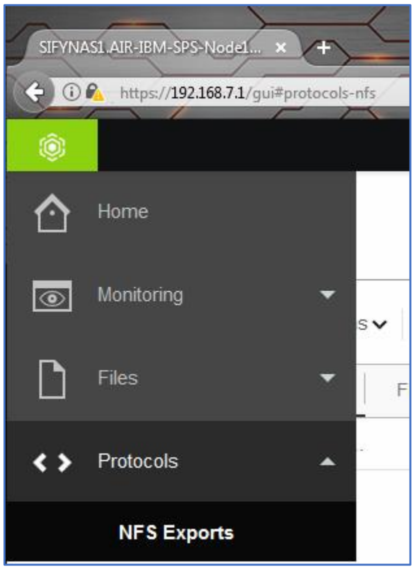
☑ Browseable

☐ Hide Unreadable

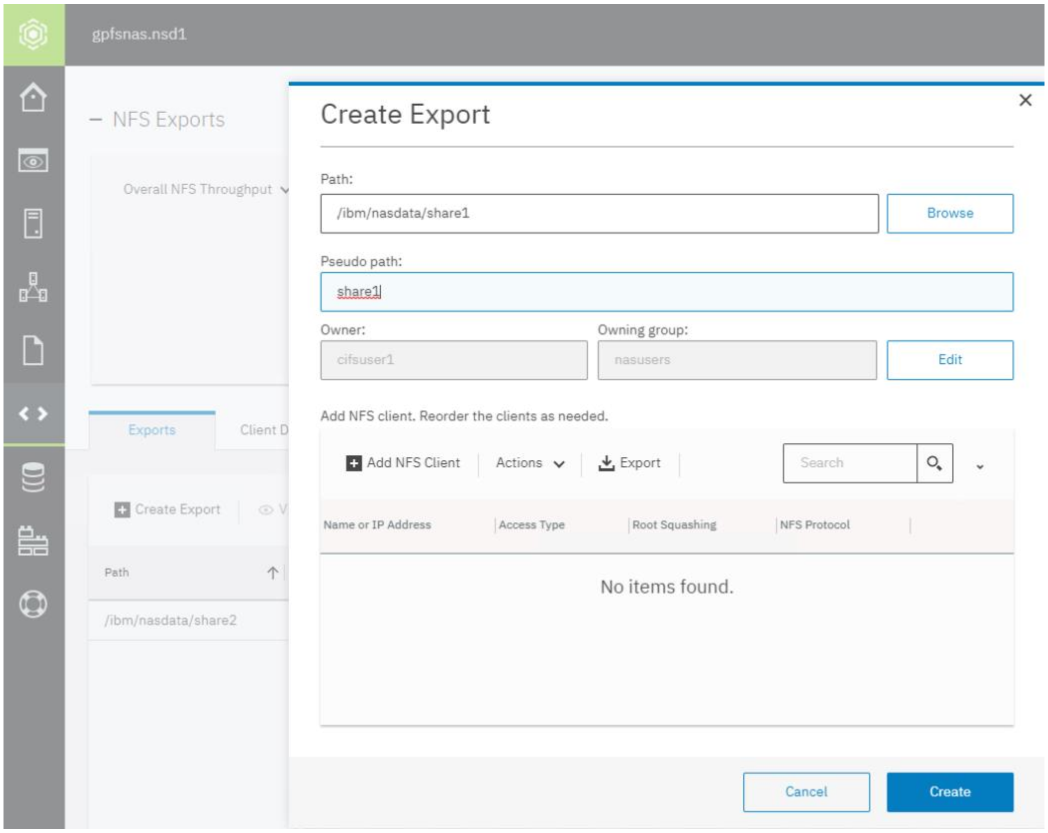
Cross-protocol integration

6.5 Create exports for NFS users

- Navigate to "**NFS Exports**" in the "**Protocols**" section


SIFYNAS1.AIR-IBM-SPS-Node1...
https://192.168.7.1/gui#protocols-nfs
Home
Monitoring
Files
Protocols
NFS Exports

- Click "**Create Export**"
- Select newly created file set path from the "**Browse"** button



Create Export

Path: /ibm/nasdata/share1

Pseudo path: share1

Owner: cifsuser1

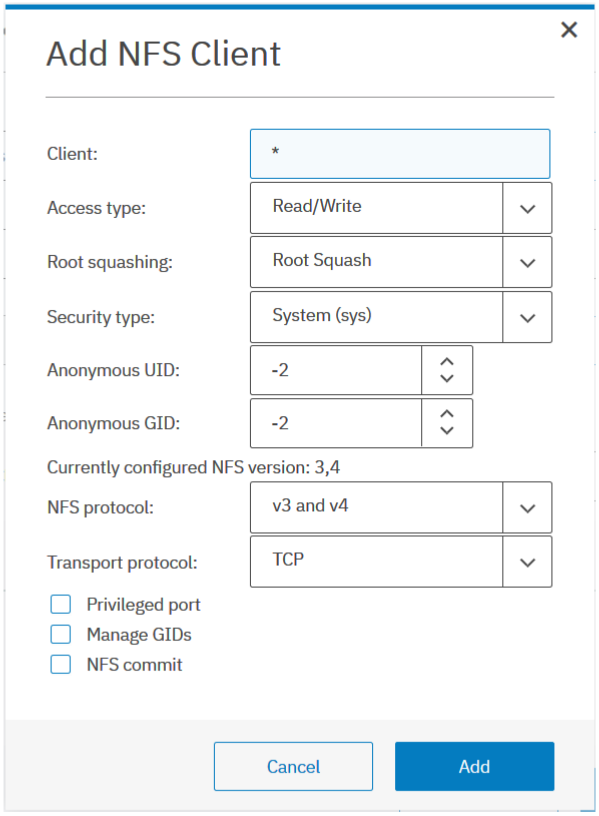
Owning group: nasusers

Add NFS client. Reorder the clients as needed.

Name or IP Address	Access Type	Root Squashing	NFS Protocol

No items found.

- **“Add client”** using the IP address of the client



Add NFS Client

Client:	*
Access type:	Read/Write
Root squashing:	Root Squash
Security type:	System (sys)
Anonymous UID:	-2
Anonymous GID:	-2

Currently configured NFS version: 3,4

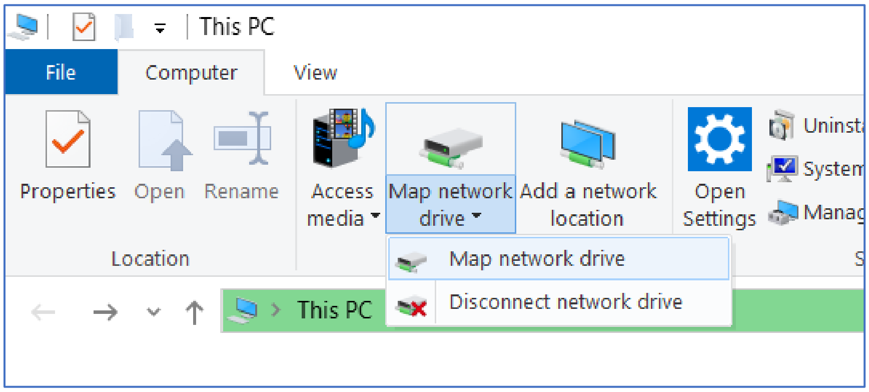
NFS protocol:	v3 and v4
Transport protocol:	TCP

□ Privileged port
□ Manage GIDs
□ NFS commit

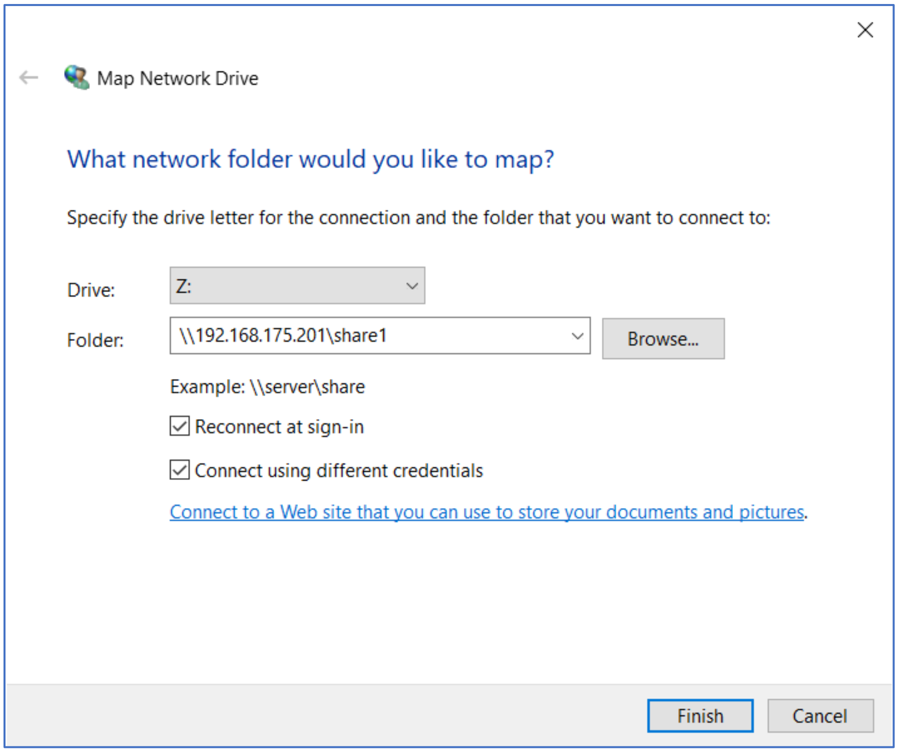
- If required, edit ownership and permission information, and click “**OK”** to complete export creation

6.6 Windows client-side mapping

- Use “**Map Network Drive**” from Windows Explorer to map the share on an unused drive letter



- Check the box “Connect using different credentials.” This will allow you to enter credentials as defined on the Spectrum Scale cluster.
- While providing the folder path, CES IPs configured on protocol nodes can be used. All CES IPs can be added in the DNS server to provide round-robin or similar load balancing. This will help to distribute client connections over all protocol nodes.



Map Network Drive

What network folder would you like to map?

Specify the drive letter for the connection and the folder that you want to connect to:

Drive: Z:

Folder: \\192.168.175.201\share1

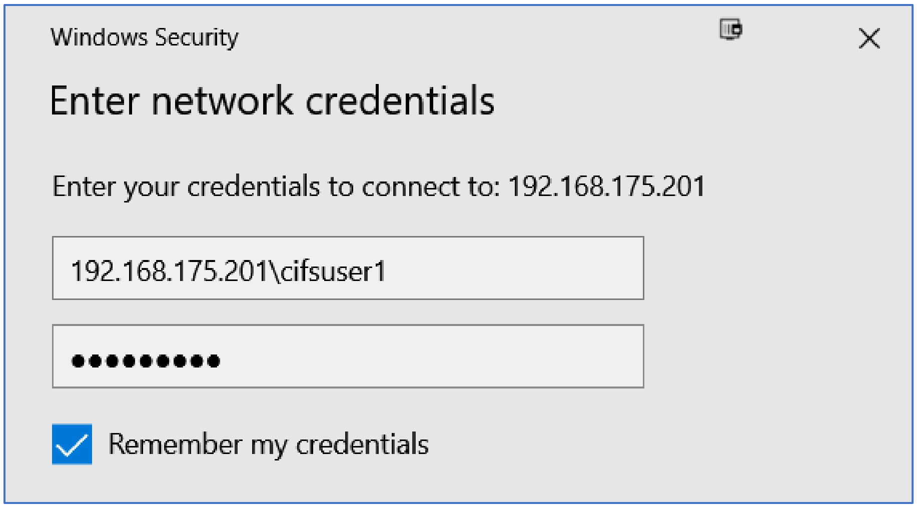
Example: \\server\share

☑ Reconnect at sign-in

☑ Connect using different credentials

Connect to a Web site that you can use to store your documents and pictures.

- Provide username and password for share



Windows Security

Enter network credentials

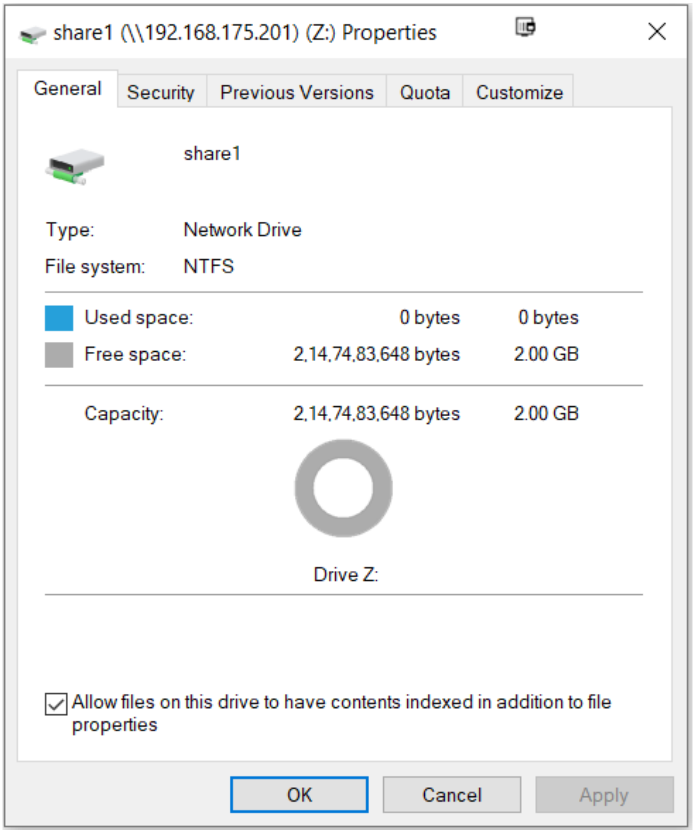
Enter your credentials to connect to: 192.168.175.201

192.168.175.201\cifsuser1

●●●●●●●●●●

☑ Remember my credentials

- The drive will get mapped, and its capacity will be set as per quota set on the underlying file set



share1 (\\192.168.175.201) (Z:) Properties

share1

Type: Network Drive

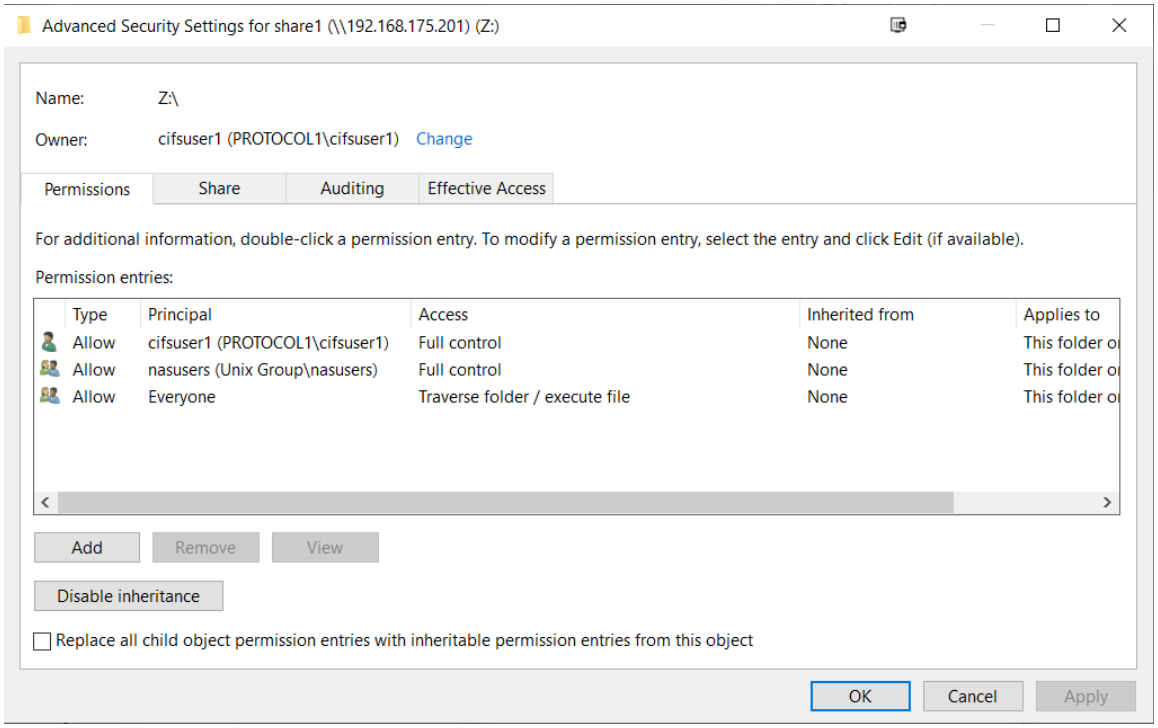
File system: NTFS

Used space:	0 bytes	0 bytes
Free space:	2,14,74,83,648 bytes	2.00 GB
Capacity:	2,14,74,83,648 bytes	2.00 GB

Drive Z:

☑ Allow files on this drive to have contents indexed in addition to file properties

- User permissions on the root directory can be verified. They would be same as defined while the share was created.



Advanced Security Settings for share1 (\\192.168.175.201) (Z:)

Name: Z:\

Owner: cifsuser1 (PROTOCOL1\cifsuser1) Change

For additional information, double-click a permission entry. To modify a permission entry, select the entry and click Edit (if available).

Permission entries:

Type	Principal	Access	Inherited from	Applies to
Allow	cifsuser1 (PROTOCOL1\cifsuser1)	Full control	None	This folder o
Allow	nasusers (Unix Group\nasusers)	Full control	None	This folder o
Allow	Everyone	Traverse folder / execute file	None	This folder o

□ Replace all child object permission entries with inheritable permission entries from this object

- The security permissions can also be seen on the Spectrum Scale cluster

```
[root@protocol1 ~]# ls -al /ibm/nasdata/
total 257
drwxr-xr-x 3 root     root     262144 Dec  9 23:54 .
drwxr-xr-x 4 root     root         34 Dec  3 22:25 ..
drwxrwx--x 2 cifsuser1 nasusers   4096 Dec  9 23:54 share1
```

6.7 UNIX/NFS client-side mapping

NFS export can be mapped on a UNIX client with the following steps:

- Create a local mount folder
- Use CES IP to mount the remote export onto the local mount folder

```
[user@client1 ~]# mount -t nfs 192.168.175.200:/ibm/nasdata/share1 /mnt/share1
[user@client1 ~]# ls -al /mnt/share1
drwxr-xr-x 3 root     root     262144 Dec  9 23:54 .
drwxr-xr-x 4 root     root         34 Dec  3 22:25 ..
drwxrwx--x 2 cifsuser1 nasusers   4096 Dec  9 23:54 testfile
```

7 Troubleshooting

This section provides some useful commands to verify and troubleshoot the setup.

7.1 System health check

“mmlscluster --ces" shows the CES IP distribution on protocol nodes

```
[root@protocol1 ~]# mmlscluster --ces

GPFS cluster information
========================
  GPFS cluster name:         gpfsnas.nsd1
  GPFS cluster id:           8831718549143064122

Cluster Export Services global parameters
-----------------------------------------
  Shared root directory:                /ibm/cesfs
  Enabled Services:                     SMB NFS
  Log level:                            0
  Address distribution policy:          even-coverage

 Node  Daemon node name            IP address       CES IP address list
-----------------------------------------------------------------------
   3   protocol1                   192.168.175.130  192.168.175.201
   4   protocol2                   192.168.175.131  192.168.175.200
```

“**mmces service list -a**” shows the status of SMB and NFS services on all nodes

```
[root@protocol1 ~]# mmces service list -a
Enabled services: SMB NFS
protocol2:  SMB is running, NFS is running
protocol1:  SMB is running, NFS is running
```

“**mmces state show -a**” shows the status of associated CES services on all nodes

```
[root@protocol1 ~]# mmces state show -a
NODE        AUTH       BLOCK      NETWORK   AUTH_OBJ   NFS       OBJ        SMB       CES
----------- ---------- ---------- --------- ---------- --------- ---------- --------- -----
protocol1   DISABLED   DISABLED   HEALTHY   DISABLED   HEALTHY   DISABLED   HEALTHY   HEALTHY
protocol2   DISABLED   DISABLED   HEALTHY   DISABLED   HEALTHY   DISABLED   HEALTHY   HEALTHY
```

Specific service issues can be checked using the mmces subcommand — “**mmces events list NETWORK -N cesnodes**”

```
[root@protocol1 ~]# mmces events list NETWORK -N cesnodes
NODE       TIMESTAMP                          EVENT NAME              SEVERITY   DETAILS
protocol1   2019-12-03 04:10:51.526533 IST       ces_enable_node_network   INFO       Network was enabled
protocol1   2019-12-03 04:11:02.506966 IST       ces_network_ips_not_defined WARNING    No CES IP addresses have been defined
protocol1   2019-12-03 04:57:34.340756 IST       ces_network_ips_up        INFO       CES-relevant IPs are served by found NICs
protocol1   2019-12-03 04:57:34.439925 IST       ces_network_found         INFO       CES NIC ens33 was found
protocol1   2019-12-03 04:57:48.684995 IST       move_cesip_to            INFO       Address 192.168.175.201 was moved from
unassigned to this node
protocol1   2019-12-03 05:14:04.324865 IST       ces_network_ips_down      WARNING    No CES IPs were assigned to this node
protocol1   2019-12-03 05:14:04.413803 IST       move_cesip_to            INFO       Address 192.168.175.200 was moved from
protocol2 to this node
protocol1   2019-12-03 05:14:04.443781 IST       move_cesip_from          INFO       Address 192.168.175.200 was moved from this
node to protocol2
protocol1   2019-12-03 05:14:04.473921 IST       move_cesip_from          INFO       Address 192.168.175.201 was moved from this
node to protocol2
protocol1   2019-12-03 22:28:57.289117 IST       ces_enable_node_network   INFO       Network was enabled
```

7.2 Check Samba configuration options

"**net conf list**" shows the Samba internal configuration

```
[root@protocol1 ~]# net conf list
[global]
        disable netbios = yes
        disable spoolss = yes
        printcap cache time = 0
        fileid:algorithm = fsname
        fileid:fstype allow = gpfs
        syncops:onmeta = no
        passdb backend = tdbsam
        preferred master = no
        client NTLMv2 auth = yes
        kernel oplocks = no
        level2 oplocks = yes
        debug hires timestamp = yes
        max log size = 100000
        host msdfs = yes
        notify:inotify = yes
        wide links = no
        log writeable files on exit = yes
        ctdb locktime warn threshold = 5000
        smbd:backgroundqueue = False
        read only = no
        use sendfile = no
        strict locking = auto
        posix locking = no
        large readwrite = yes
        aio read size = 1
        aio write size = 1
        force unknown acl user = yes
        store dos attributes = yes
        map readonly = yes
        map archive = yes
        map system = yes
        map hidden = yes
        ea support = yes
        groupdb:backend = tdb
        winbind:online check timeout = 30
        winbind:request profile threshold = 5
        winbind max domain connections = 5
        winbind max clients = 10000
        dmapi support = no
        UNIX extensions = no
        socket options = TCP_NODELAY SO_KEEPALIVE TCP_KEEPCNT=4 TCP_KEEPIDLE=240 TCP_KEEPINTVL=15
        strict allocate = yes
        tdbsam:map builtin = no
        password server = *
        aio_pthread:aio open = yes
        dfree cache time = 100
        change notify = yes
        max open files = 20000
        time_audit:timeout = 5000
        gencache:stabilize_count = 100000
        gencache:stabilize_interval = 3600
        server min protocol = SMB2_02
```

```
	server max protocol = SMB3_11
	vfs objects = shadow_copy2 syncops gpfs fileid time_audit
	smbd profiling level = on
	log level = 1
	logging = syslog@0 file
	smbd exit on ip drop = yes
	durable handles = no
	ctdb:smbxsrv_open_global.tdb = false
	mangled names = illegal
	include system krb5 conf = no
	smbd:async search ask sharemode = yes
	restrict anonymous = 2
	gpfs:sharemodes = yes
	gpfs:leases = yes
	gpfs:dfreequota = yes
	gpfs:prealloc = yes
	gpfs:hsm = yes
	gpfs:winattr = yes
	gpfs:merge_writeappend = no
	fruit:copyfile = yes
	fruit:metadata = stream
	fruit:nfs_aces = no
	fruit:veto_appledouble = no
	readdir_attr:aapl_max_access = false
	shadow:snapdir = .snapshots
	shadow:fixinodes = yes
	shadow:snapdirseverywhere = yes
	shadow:sort = desc
	idmap:cache = no
	idmap config * : read only = no
	idmap config * : backend = autorid
	idmap config * : range = 10000000-299999999
	idmap config * : rangesize = 1000000
	nfs4:mode = simple
	nfs4:chown = yes
	nfs4:acedup = merge
	add share command = /usr/lpp/mmfs/bin/mmcesmmccrexport
	change share command = /usr/lpp/mmfs/bin/mmcesmmcchexport
	delete share command = /usr/lpp/mmfs/bin/mmcesmmcdelexport
	server string = IBM NAS
	security = user

[share1]
	path = /ibm/nasdata/share1
	guest ok = no
	smb encrypt = auto
	browseable = yes
```

Note that some of these options can be changed, but avoid doing so. The Samba service used by Spectrum Scale CES will reject most changes, but changes that can go through may have undefined consequences.

7.3 Increase Samba debug log level

Samba service uses the default Linux log (/var/log/messages) for important messages. Detailed protocol logs can also be enabled to troubleshoot any issue.

Current Samba logging details can be checked and reset using the “log level” option in the Samba configuration. It is recommended that you change the log level back to its default value “1” after troubleshooting is done.

```
[root@protocol1 ~]# net conf list | grep "log level"
        log level = 1
[root@protocol1 ~]# net conf setparm global "log level" "2"
[root@protocol1 ~]# net conf list | grep "log level"
        log level = 2
```

This default log can be accessed from the file – “/var/adm/ras/log.smbd”

8 Conclusion

Spectrum Scale protocol services can provide access to the GPFS file system for external clients without integrating them into the cluster. Protocol services also provide granular access to the file system with user authentication and access control.

Through this paper, we have tried to provide a concise implementation guide to set up a cluster based on local authentication configuration for Spectrum Scale protocol services. The reference architecture discussed here can be used to plan and deploy similar solutions with less effort.

9 Authors

Kiran Ghag is a storage consultant for IBM Systems Labs Services India. Kiran joined IBM in 2013 as a consultant with 10 years of experience in storage systems. His current interests include providing solutions for unstructured data storage using IBM Spectrum family, NAS implementations and storage infrastructure optimization.

Rakesh Chutke has been an IBM Storage and Spectrum Scale consultant with IBM Systems Lab Services for the last 13 years. He has more than 18 years of IT industry experience working with clients across industries, including banking, insurance, telecommunications, media and entertainment, oil and gas.

Contact IBM Systems Lab Services at ibmsls@us.ibm.com.

10 References

10.1 IBM Spectrum Scale Knowledge Center

https://www.ibm.com/support/knowledgecenter/en/STXKQY_5.0.4/ibmspectrumscale504_welcome.html

10.2 Planning for IBM Spectrum Scale – Planning for Protocols

https://www.ibm.com/support/knowledgecenter/en/STXKQY_5.0.4/com.ibm.spectrum.scale.v5r04.doc/bl1in_PlanningForProtocols.htm

10.3 Installing IBM Spectrum Scale on Linux nodes and deploying protocols

https://www.ibm.com/support/knowledgecenter/en/STXKQY_5.0.4/com.ibm.spectrum.scale.v5r04.doc/bl1ins_loosein.htm

© Copyright IBM Corporation 2020

IBM Corporation
Route 100
Somers, NY 10589

Produced in the United States of America
May 2020

IBM, the IBM logo, ibm.com and IBM Spectrum are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. The registered trademark Linux is used pursuant to a sublicense from the Linux Foundation, the exclusive licensee of Linus Torvalds, owner of the mark on a world-wide basis. Microsoft is a trademark 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. 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.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

Please Recycle