Elastic Storage Server
Version 5.3.0
(Updated for 5.3.0.1)

Quick Deployment Guide

IBM

Elastic Storage Server
Version 5.3.0
(Updated for 5.3.0.1)

Quick Deployment Guide

IBM

> **Note**
>
> Before using this information and the product it supports, read the information in "Notices" on page 87.

Contents

Tables

About this information

This information guides you in installing, or upgrading to, version 5.3.x of the Elastic Storage Server (ESS).

Who should read this information

This information is intended for experienced system installers and upgraders who are familiar with ESS systems.

Related information

ESS information

The ESS 5.3 library consists of these information units:
- *Elastic Storage Server: Quick Deployment Guide*, SC27-9205
- *Elastic Storage Server: Problem Determination Guide*, SC27-9208
- *Elastic Storage Server: Command Reference*, SC27-9246
- *IBM Spectrum Scale RAID: Administration*, SC27-9206
- *IBM ESS Expansion: Quick Installation Guide (Model 084)*, SC27-4627
- *IBM ESS Expansion: Installation and User Guide (Model 084)*, SC27-4628
- *IBM ESS Expansion: Hot Swap Side Card - Quick Installation Guide (Model 084)*, GC27-9210
- *Installing the Model 024, ESLL, or ESLS storage enclosure*, GI11-9921
- *Removing and replacing parts in the 5147-024, ESLL, and ESLS storage enclosure*
- *Disk drives or solid-state drives for the 5147-024, ESLL, or ESLS storage enclosure*

For more information, see IBM® Knowledge Center:

http://www-01.ibm.com/support/knowledgecenter/SSYSP8_5.3.0/sts53_welcome.html

For the latest support information about IBM Spectrum Scale™ RAID, see the IBM Spectrum Scale RAID FAQ in IBM Knowledge Center:

http://www.ibm.com/support/knowledgecenter/SSYSP8/sts_welcome.html

Switch information

ESS release updates are independent of switch updates. Therefore, it is recommended that Ethernet and Infiniband switches used with the ESS cluster be at their latest switch firmware levels. Customers are responsible for upgrading their switches to the latest switch firmware. If switches were purchased through IBM, review the minimum switch firmware used in validation of this ESS release available in *Customer networking considerations* section in *Elastic Storage Server: Quick Deployment Guide*.

Other related information

For information about:
- IBM Spectrum Scale, see IBM Knowledge Center:
  http://www.ibm.com/support/knowledgecenter/STXKQY/ibmspectrumscale_welcome.html
- IBM POWER8® servers, see IBM Knowledge Center:
  http://www.ibm.com/support/knowledgecenter/POWER8/p8hdx/POWER8welcome.htm

- The DCS3700 storage enclosure, see:
  - *System Storage® DCS3700 Quick Start Guide*, GA32-0960-03:

    http://www.ibm.com/support/docview.wss?uid=ssg1S7004915
  - *IBM System Storage DCS3700 Storage Subsystem and DCS3700 Storage Subsystem with Performance Module Controllers: Installation, User's, and Maintenance Guide*, GA32-0959-07:

    http://www.ibm.com/support/docview.wss?uid=ssg1S7004920
- The IBM Power Systems™ EXP24S I/O Drawer (FC 5887), see IBM Knowledge Center :

  http://www.ibm.com/support/knowledgecenter/8247-22L/p8ham/p8ham_5887_kickoff.htm
- Extreme Cluster/Cloud Administration Toolkit (xCAT), go to the xCAT website :

  http://sourceforge.net/p/xcat/wiki/Main_Page/
- Mellanox OFED Release Notes®, go to https://www.mellanox.com/related-docs/prod_software/ Mellanox_OFED_Linux_Release_Notes_4_1-1_0_2_0.pdf

Conventions used in this information

Table 1 describes the typographic conventions used in this information. UNIX file name conventions are used throughout this information.

Table 1. Conventions

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

How to submit your comments

Your feedback is important in helping us to produce accurate, high-quality information. You can add comments about this information in IBM Knowledge Center:

http://www.ibm.com/support/knowledgecenter/SSYSP8/sts_welcome.html

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

scale@us.ibm.com

Change history

Version	PDF form number	Summary
3	SC27-9205-05	Fix for Genesis discovery failure issue
2	SC27-9205-04	Changes for ESS 5.3.0.1
1	SC27-9205-03	Initial version for ESS 5.3.0

Chapter 1. Installation and upgrade related information and checklists

Review the following installation and upgrade related information before starting with the installation or the upgrade of Elastic Storage Server (ESS).

- "Component versions for this release"
- "Supported editions on each architecture" on page 2
- "ESS best practices and support statements" on page 2
- "Obtaining the required Red Hat Enterprise Linux and ESS code" on page 2
- "Customer networking considerations" on page 3
- "Supported upgrade paths" on page 5
- "Support for hardware call home" on page 5
- "Pre-installation checklist" on page 5
- "Post-installation checklist" on page 6
- "Other topics" on page 6
- "Sample installation and upgrade flow" on page 6

New features and enhancements in ESS 5.3.0

- **gssgennetworks** support for IB bonding
- **gssdeploy** improvements (PPC64LE discovery / genesis)
- **gssutils** overhaul
- Support for setting MTU in **gssgennetworks**

Component versions for this release

Note: Your version might be slightly different from the version indicated in this document. Refer to the release notes document that comes with the installation package for the exact version of the installation package and the component version.

The respective versions for the core components in this release of ESS are as follows:

- Supported architectures: PPC64BE and PPC64LE
- IBM Spectrum Scale: 5.0.0.1
- xCAT: 2.13.9
- HMC: 860 SP2
- System firmware: SV860_138(FW860.42)
- Red Hat Enterprise Linux: 7.3
- Kernel: 3.10.0-514.44.1
- Systemd: 219-42.el7_4.10
- Network Manager: 1.8.0-11.el7_4
- OFED: MLNX_OFED_LINUX-4.1-4.1.6.1
- IPR: 17518300
- ESA: 4.2.0-9

Supported editions on each architecture

The following are the ESS editions supported on the available architectures.

PPC64BE
- Standard Edition
- Advanced Edition
- Data Management Edition

PPC64LE
- Standard Edition
- Data Management Edition

ESS best practices and support statements

- It is advised that when performing normal maintenance operations (or upgrades) that you disable autoload first.

  ```
  mmchconfig autoload=no
  ```

  Once the maintenance operation (or upgrade) is complete, re-enable autoload.

  ```
  mmchconfig autoload=yes
  ```

- By default, file systems must only be mounted on the management server node (EMS). Do not mount the file system on any other ESS nodes besides the EMS (where the primary GUI runs) which is mandatory for the GUI to function correctly.
- It is advised that you disable automount for file systems when performing an upgrade to ESS 5.3.0.1 or later.

  ```
  mmchfs Device -A no
  ```

  Device is the device name of the file system.
- Do not configure more than 5 failure groups in a single file system.
- All Infiniband devices must be set to `CONNECTED_MODE=no`.
- If you have 40Gb adapters, enable flow control on your switch
- RDMA over Ethernet (RoCE) is not supported.
- Sudo on the ESS nodes is not supported.
- Enabling the firewall on any ESS node is not supported.
- Enabling SELinux on any ESS node is not supported.
- Running any additional service or protocols on any ESS node is not supported.
- Move quorum and cluster or file system management function off of the ESS nodes where possible.
- All nodes must be at the same level prior to adding a building block. Therefore, upgrade existing ESS building block before adding the new one in.
- You must take down the GPFS cluster to run firmware updates in parallel.
- Do not independently update IBM Spectrum Scale (or any component) on any ESS node unless specifically advised from the L2 service. Normally this is only needed to resolve an issue. Under normal scenarios it is advised to only upgrade in our tested bundles.

Obtaining the required Red Hat Enterprise Linux and ESS code

The required Red Hat components are:
- Red Hat Enterprise Linux 7.3 ISO

  ```
  18402 3370316 rhel-server-7.3-ppc64-dvd.iso
  45464 3114204 rhel-server-7.3-ppc64le-dvd.iso
  ```
- Network manager version : 1.8.0-11.el7_4

```
|    40869 8769 netmgr-RHBA-2017-2925-BE.tar.gz
|    31336 8660 netmgr-RHBA-2017-2925-LE.tar.gz
```
| • Systemd version: 219-42.el7_4.10
```
|    56649 7212 systemd-530-RHBA-2018-0416-BE.tar.gz
|    24625 7149 systemd-530-RHBA-2018-0416-LE.tar.gz
```
| • Kernel version: 3.10.0-514.44.1
```
|    2013 63700 kernel-530-RHSA-2018-0399-BE.tar.gz
|    5436 63544 kernel-530-RHSA-2018-0399-LE.tar.gz
```

On ESS 5.3.0.x systems shipped from manufacturing, these items can be found on the management server node in the /home/deploy directory.

Customers or business partners can download the required Red Hat components from Red Hat Network using the customer license. For more information, see:
- Appendix I, "Obtaining kernel for system upgrades," on page 71
- Appendix J, "Obtaining systemd update for system upgrades," on page 75
- Appendix K, "Obtaining Network Manager updates for system upgrades," on page 79

If you are a member of IBM, you must contact ESS development or L2 service to obtain the code directly.

ESS 5.3.0.x can be downloaded from IBM FixCentral.

The ESS software archive that is available in different versions for both PPC64BE and PPC64LE architectures.

PPC64BE packages available:
```
ESS_STD_BASEIMAGE-5.3.0.1-ppc64-Linux.tgz
ESS_ADV_BASEIMAGE-5.3.0.1-ppc64-Linux.tgz
ESS_DM_BASEIMAGE-5.3.0.1-ppc64-Linux.tgz
```

PPC64LE packages available:
```
ESS_STD_BASEIMAGE-5.3.0.1-ppc64le-Linux.tgz
ESS_DM_BASEIMAGE-5.3.0.1-ppc64le-Linux.tgz
```

You can download these packages from FixCentral at https://www-945.ibm.com/support/fixcentral/ swg/selectFixes?parent=Software%20defined%20storage&product=ibm/StorageSoftware/ IBM+Elastic+Storage+Server+(ESS)&release=All&platform=All&function=all.

Once downloaded and placed in /home/deploy, untar and uncompress the package to view the contents. For example, for the standard edition PPC64BE package, use the following command:
```
tar -xvf ESS_STD_BASEIMAGE-5.3.0.1-ppc64-Linux.tgz
```

The BASEIMAGE tar file contains the following files that get extracted with the preceding command:
- ESS_5.3.0.1_ppc64_Release_note_Standard.txt: This file contains the release notes for the latest code.
- gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z.tgz: This .tgz file contains the ESS code.
- gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z.md5: This .md5 file to check the integrity of the tgz file.

Customer networking considerations

Review the information about switches and switch firmware that were used to validate this ESS release. For information about available IBM networking switches, see the IBM networking switches page on IBM Knowledge Center.

The following switches and switch firmware were used to validate this ESS release.

```
Switch MTM - Switch description - Switch FW <<<<< This is just an example.
8828-E36/E37 - Mellanox SB7700 36port EDR - 3.6.5011
8831-F36 / F37 - Mellanox SX6036 36port FDR - 3.6.5011
8831-NF2 - Mellanox SX1710 36Port 40GbE - 3.6.5011
```

If you are using the this switch(es) schedule, a switch maintenance downtime to upgrade the switches is recommended prior to using this release of ESS. Also, it is recommended that if two switches are used in a high availability (HA) configuration, both switches be at the same firmware level.

To check the firmware version, do the following:

1. SSH to the switch.

2. Issue the following commands.

   ```
   # en
   # show version
   ```

   For example:

   ```
   login as: admin
   Mellanox MLNX-OS Switch Management
   Using keyboard-interactive authentication.
   Password:
   Last login: Mon Mar 5 12:03:14 2018 from 9.3.17.119
   Mellanox Switch
   io232 [master] >
   io232 [master] > en
   io232 [master] # show version
   ```

   Example output:

   ```
   Product name: MLNX-OS
   Product release: 3.4.3002
   Build ID: #1-dev
   Build date: 2015-07-30 20:13:19
   Target arch: x86_64
   Target hw: x86_64
   Built by: jenkins@fit74
   Version summary: X86_64 3.4.3002 2015-07-30 20:13:19 x86_64
   Product model: x86
   Host ID: E41D2D52A040
   System serial num: Defined in system VPD
   System UUID: 03000200-0400-0500-0006-000700080009
   ```

Infiniband with multiple fabric

In a multiple fabric network, the Infiniband Fabric ID might not be properly appended in the verbsPorts configuration statement during the cluster creation. It is advised to do the following to ensure that the verbsPorts setting is accurate:

1. Use **gssgennetworks** to properly set up IB or Ethernet bonds on the ESS system.

2. Create a cluster.

3. Run **mmfsadm test verbs config | grep verbsPorts**

These steps show the Fabric ID found for each link.

For example:

```
mmfs verbsPorts:  mlx5_0/1/4  mlx5_1/1/7
Adapter mlx5_0, port 1 connected to fabric 4 and adapter mlx5_1 port 1 connected to fabric 7.
```

Now using **mmchconfig**, modify the verbsPorts for each node or node class to take the subnet into account.

Supported upgrade paths

The following upgrade paths are supported:
- ESS version 5.1.x and 5.2.x to version 5.3.x on PPC64BE.
- ESS version 5.1.x and 5.2.x to version 5.3.x on PPC64LE.

Note: For upgrading to ESS 5.3.x from version 5.0.x or earlier (PPC64BE), you must contact IBM Support because direct upgrade to version 5.3.x from these versions is not supported. The available indirect upgrade paths are as follows.
- **3.5.5 (or earlier) > 4.5.2 > 5.1.x > 5.3.x**
- **4.0.x > 5.0.x > 5.1.x (or 5.2.x) > 5.3.x**
- **4.5.x (or 4.6.x) > 5.1.x > 5.3.x**

Important: If you are not upgrading to ESS 5.3.x, it is recommended that you install ESS 5.2.1 to avoid system stability or functional issues.

Support for hardware call home

	PPC64BE	PPC64LE
Call home when disk needs to be replaced	X	X
Enclosure call home	Unsupported	Unsupported
Server call home	Through HMC	Unsupported

Pre-installation checklist

Before you arrive at a customer site, it is advised that you perform the following tasks:

	Obtain access to the required RHEL components (Contact IBM ESS development or L2 Service for access).
	Obtain the kernel, systemd, networkmanager, RHEL ISO (Provided by ESS development or L2 Service), and ESS tarball (FixCentral). Verify that the checksum match with what is listed in this document. Also ensure that you have the correct architecture packages (PPC64LE or PPC64BE).
	Download and read the latest ESS Quick Deployment Guide and browse the related ESS 5.3 documentation in IBM Knowledge Center.
	Obtain the customer RHEL license.
	Contact the local SSR and ensure that all hardware checks have been completed. Make sure all hardware found to have any issues has been replaced.
	If the 1Gb switch is not included in the order, contact the local network administrator to ensure isolated xCAT and FSP VLANs are in place.
	Develop an inventory and plan for how to upgrade, install, or tune the client nodes.
	Upgrade the HMC to SP2 if doing a PPC64BE installation. This can be done concurrently.
	Consider talking to the local network administrator regarding ESS switch best practices, especially the prospect of upgrading the high-speed switch firmware at some point prior to moving the system into production, or before an upgrade is complete. For more information, see "Customer networking considerations" on page 3.
	Review *Elastic Storage Server: Command Reference*.
	Review ESS FAQ and ESS best practices.
	Review the ESS 5.3.0 known issues.

	Ensure that all client node levels are compatible with the ESS version. If needed, prepare to update the client node software on site and possibly other items such as the kernel and the network firmware or driver.
	Power down the storage enclosures, or remove the SAS cables, until the `gssdeploy -x` operation is complete.

Post-installation checklist

After the installation is completed, it is advised that you verify the following:

	Call home has been set up and tested.
	• For more information, see *Drive call home in 5146 and 5148 systems* in *Elastic Storage Server: Problem Determination Guide*. • For information about HMC call home (Server PPC64BE Only), see Configuring HMC Version 8.8.3 and Later for Call Home. • For call home support information, see "Support for hardware call home" on page 5.
	GUI has been set up and demonstrated to the customer.
	GUI SNMP alerts have been set up, if desired.
	The customer RHEL license is registered and active.
	No issues have been found with `mmhealth`, GUI, `gnrhealthcheck`, `gssinstallcheck`, serviceable events.
	No SAS width/speed issues have been found.
	Client nodes are properly tuned.
	It is advised that you turn on autoload to enable GPFS to recover automatically in case of a daemon problem. `mmchconfig autoload=yes`
	Connect all nodes to Red Hat Network (RHN).
	Update any security related erratas from RHN if the customer desires (`yum -y security`).
	Ensure that you have saved a copy of the xCAT database off to a secure location.
	Install or upgrade the protocols. For more information, see Appendix A, "Upgrading a cluster containing ESS and protocol nodes," on page 39.
	Ensure (if possible) that all network switches have had the firmware updated.

Other topics

For help with the following topics, and many others that are unlisted, contact L2 Service.
• Adding a building block (same architecture or LE<->BE)
• Restoring a management server
• Part upgrades or replacements
• VLAN reconfiguration on the 1Gb switch

Sample installation and upgrade flow

New installations go through manufacturing CSC. The system is fully installed with ESS 5.3.0, tested, malfunctioning parts replaced, and required RHEL pieces shipped in `/home/deploy`.

Installation

To install an ESS 5.3.0 system at the customer site, it is recommended that you use the Fusion mode available with `gssutils`. For more information, see "Elastic Storage Server 5.2 or later: Fusion Mode" on page 28 and Chapter 2, "`gssutils` - ESS Installation and Deployment Toolkit," on page 9.

Note: All operations must use `gssutils`.

- SSR checkout complete
- LBS arrival on site
- Plug-n-Play mode demonstrated
- Decisions made on block size, host names, IP addresses (`/etc/hosts` generated)
- Check high speed switch settings or firmware
- Firmware updated on ESS nodes
- Fusion mode used to bring system to cluster creation
- Network bonds created
- Cluster created
- Recovery groups, NSDs, file system created
- Stress test performed
- Final checks performed
- GUI setup (w/SNMP alerts if desired)
- Call home setup
- Nodes attached to RHN and security updates applied

Upgrade

To upgrade to an ESS 5.3.0 system at the customer site, it is recommended that you use `gssutils`. For more information, see Chapter 2, "`gssutils` - ESS Installation and Deployment Toolkit," on page 9.

Note: All operations must use `gssutils`.

- SSR checkout complete
- Check high speed switch settings or firmware
- Ensure that there are no hardware issues
- Ensure client / protocol node compatibility
- Ensure no heavy IO operations are being performed
- Upgrade ESS (rolling upgrade or with cluster down)
  - Always ensure you have quorum (if rolling upgrade)
  - Always carefully balance the recovery groups and scale management functions as you upgrade each node (if rolling upgrade)
- Final checks performed
- Determine if any `mmperfmon` changes required
- Ensure GUI SNMP alerts and call home still working

Chapter 2. `gssutils` - ESS Installation and Deployment Toolkit

Currently, LBS team members or customers use ESS command line tools to perform SSR prechecks, installation, deployment, or upgrade tasks on the ESS server by following the ESS: Quick Deployment Guide (QDG).

The ESS Installation and Deployment Toolkit (**gssutils**) is designed to facilitate these tasks by using a menu driven tool.

The ESS Installation and Deployment Toolkit (**gssutils**) is designed to eliminate these problems.

- "gssutils introduction"
- "gssutils prerequisites" on page 11
- "gssutils defaults" on page 12
- "gssutils usage" on page 12
- "gssutils customization" on page 14
- "gssutils restrictions" on page 16

`gssutils` introduction

gssutils is a text-based, menu-driven utility that facilitates SSR prechecks, installation, deployment, and upgrade tasks. It can be started with or without any optional arguments. It provides a set of task menus that are related to installation and deployment activities. When a task is selected from the menu, a command is issued to the system for that task. **gssutils** requires a minimum of 80 x 24 character window to operate.

When running **gssutils**, one can pass the EMS server, I/O server nodes, prefix, and suffix while initiating the tool and all the commands are automatically adjusted to use the provided value while running the actual commands. For example, the SSR check, Plug N Play mode, Fusion mode, installation, deployment, or upgrade commands are customized to run for user provided EMS server, I/O server nodes, prefix, and suffix.

gssutils can be customized depending on environment and it can be run by just pressing the Enter key without doing any command line changes such as changing EMS host name, I/O server node names, prefix or suffix for CLI and there is no need to remember the CLI options.

gssutils also guides users to run the command in the correct order as it is a menu-driven, step wise user interface. This allows the user to not having to remember the next command in the sequence.

gssutils also has a built-in help system available for each menu. All menus in **gssutils** have an option called **Help**. Help menu brings up a Linux manual page which explains the use of the menu and also explains the commands that are invoked from that menu. Users must read the help from each menu before running any tasks.

The **gssutils** console is as follows.

```
        ESS INSTALLATION AND DEPLOYMENT TOOLKIT


        1. Help
        2. SSR Tools                        >
        3. Plug n Play and Hybrid           >
        4. Install                          >
        5. Upgrade                          >
        6. Validation checks                >
        7. View/Collect service data (snaps) >
        8. Exit




  man gssutils_panel_1
  Help
```
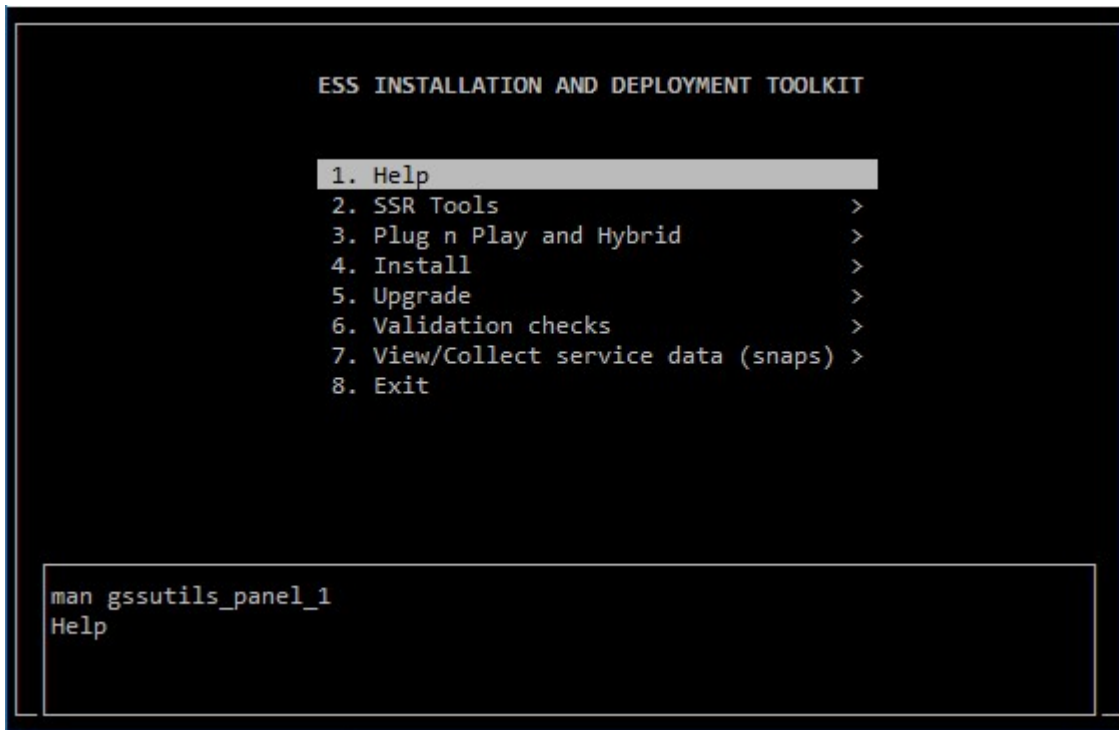
The **gssutils** command line help is as follows:

gssutils -h

usage: **gssutils** [-h] [-N NODE-LIST | -G NODE-GROUP] [--prefix PREFIX][--suffix SUFFIX]
[--config CONFIG_FILE] [--customize][--ems-node-name EMS_NODE]
[--io-node-one-name IO_NODE1][--io-node-two-name IO_NODE2]

Optional arguments:

-N *NODE-LIST*
> Provides a list of nodes. If node list or group name is not provided, -N localhost is assumed.

-G *NODE-GROUP*
> Provides the name of node group. Nodes in the *NODE-LIST* are members of the *NODE-GROUP*.

--prefix *PREFIX*
> Provides the host name prefix. Use = between --prefix and value if the value starts with -.

--suffix *SUFFIX*
> Provides the host name suffix. Use = between --suffix and value if the value starts with -.

--config *CONFIG_FILE*
> Provides the configuration file for **gssutils** for a specific environment.

--customize
> Customizes the EMS host name and I/O node name and generates the **gssutils** configuration file.
> This file can be used with --config to run **gssutils** specific to an environment.

--ems-node-name *EMS_NODE*
> Specifies EMS host name to populate the **gssutils** configuration file.

--io-node-one-name *IO_NODE1*
> Specifies I/O node 1 host name to populate the **gssutils** configuration file.

--io-node-two-name *IO_NODE2*
> Specifies I/O node 2 host name to populate the **gssutils** configuration file.

-h | --help
   Displays usage information about this script and exits.

This tool has six major sections:
* **SSR Tools**
* **Plug n Play and Fusion**
* **Install/Deploy**
* **Upgrade**
* **Validation checks**
* **View/Collect service data (snaps)**

The **SSR Tools** menu option can be used to validate the system after it has arrived to the customer location and LBS teams have plugged in all servers and components together and the GPFS™ cluster is ready to be deployed. SSR tools help you to validate the system interconnect including whether the network is correct or not, the enclosures are connected to the I/O server nodes or not, etc. It is the primary menu option that must be used by LBS or the customer before creating the GPFS cluster.

The **Plug n Play and Fusion** menu option provides two modes: **Plug N Play** and **Fusion**. The Plug N Play mode allows customers to build a cluster, file system and begin sampling the GUI as soon as possible. The stated goal is for this to be achieved in under an hour after LBS starts working on the system. Manufacturing now ships EMS with xCAT preconfigured with default settings. The Fusion mode no longer requires that Elastic Storage Server (ESS) systems be rediscovered or re-deployed at a customer site. The end goal of this mode is to greatly reduce the time and the complexity in bringing up an ESS system.

The **Install/Deploy** menu option allows you to re-deploy the EMS along with I/O server nodes in case the customer has cleaned up (re-imaged) the ESS nodes and they want to restart the EMS and I/O server nodes deployment from scratch. This menu option is a menu-driven representation of the installation section of the ESS: Quick Deployment Guide.

The **Upgrade** menu option allows you to upgrade the existing ESS server to the latest version of the ESS software. This menu item is a menu-driven representation of the upgrade section of the ESS: Quick Deployment Guide.

The **Validation checks** menu option can be used once the ESS system has been deployed and the GPFS cluster is up and running. Using this option, users can collect firmware level, run stress test on network and file system, etc. It can be used as a post deployment utility to validate the ESS server.

The **View/Collect service data** option can be used to collect the `gpfs.snap` SOS reports.

gssutils prerequisites

`gssutils` is a part of `gpfs.gss.tools-<version>.el7.noarch.rpm`. The `gpfs.gss.tools` RPM must be installed either manually (using **yum** or **rpm -ivh**) or using the **gssinstall_<arch> -u** command. Once the `gpfs.gss.tools` RPM is installed, users can type **gssutils** on the command prompt to use this utility, in case the ESS systems were wiped and re-imaged for fresh deployment.

If the system has been shipped from manufacturing, the ESS server is already deployed and the `gpfs.gss.tools` RPM is already installed. Users can type **gssutils** on the command prompt to use it for activities such as doing SSR checks or creating GPFS cluster as GPFS cluster needs to be created on newly shipped ESS systems.

If you want to use **gssutils** for upgrading to ESS 5.3.x, you must do these steps before you can use **gssutils** for upgrade.

1. Make the **gssdeploy** script executable:

   `chmod +x /opt/ibm/gss/install/rhel7/<ARCH>/samples/gssdeploy`

2. Clean the current xCAT installation and associated configuration to remove any preexisting xCAT configuration, and then address any errors before proceeding:

   `/opt/ibm/gss/install/rhel7/<ARCH>/samples/gssdeploy -c`

   Note: If you are doing an upgrade, use the following command to clean up the current xCAT configuration and to save a backup copy of the xCAT database.

   `/opt/ibm/gss/install/rhel7/<ARCH>/samples/gssdeploy -c -r /var/tmp/xcatdb`

3. Run one of the following commands depending on the architecture:
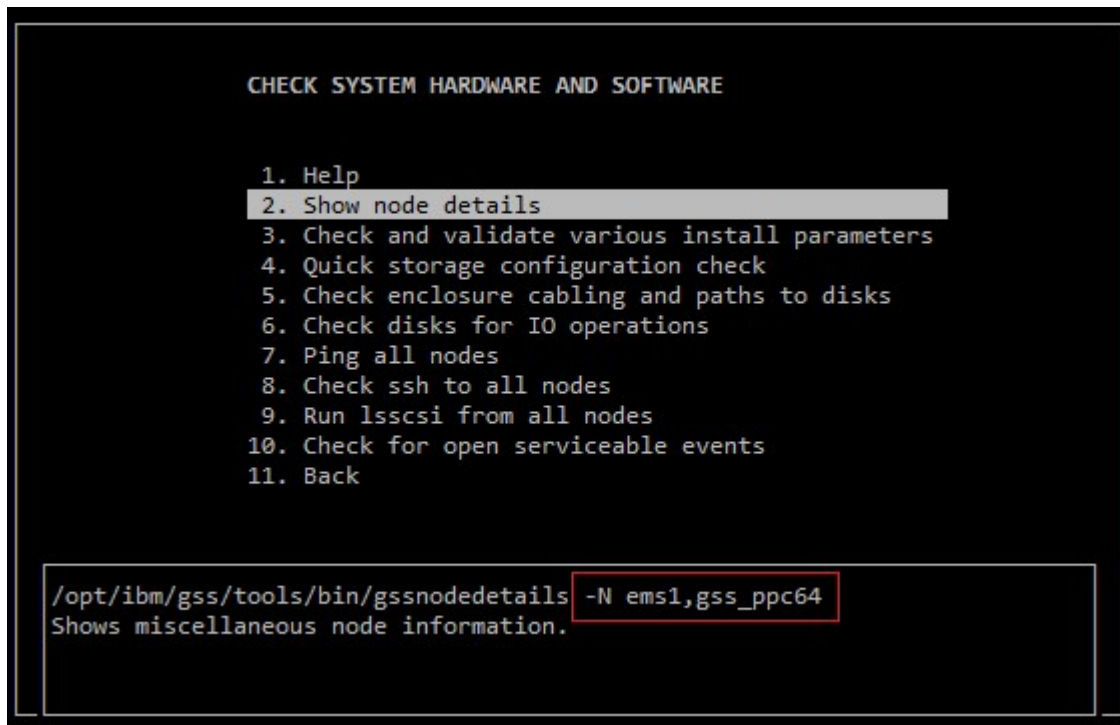
   For PPC64BE:

   `cd /var/tmp ; ./gssinstall_ppc64 -u`

   For PPC64LE:

   `cd /var/tmp ; ./gssinstall_ppc64le -u`

gssutils defaults

gssutils comes with some defaults such as the default EMS host name, I/O server node names, prefix, and suffix. By default, all **gssutils** commands assume that the default EMS host name is ems1, default I/O server node name group gss_ppc64, I/O server node names are gssio1 and gssio2, and prefix and suffix are empty.

gssutils usage

The -N and -G options:

gssutils has `-N` and `-G` options which can be used to replace the default node list or node group in ESS commands. The default is node list is **ems1,gss_ppc64**. If you want to run the **gssutils** with a different node list option, do it as follows.

$ **gssutils -N ems2**

```
                 CHECK SYSTEM HARDWARE AND SOFTWARE


             1. Help
             2. Show node details
             3. Check and validate various install parameters
             4. Quick storage configuration check
             5. Check enclosure cabling and paths to disks
             6. Check disks for IO operations
             7. Ping all nodes
             8. Check ssh to all nodes
             9. Run lsscsi from all nodes
            10. Check for open serviceable events
            11. Back



  /opt/ibm/gss/tools/bin/gssnodedetails -N ems2
  Shows miscellaneous node information.
```

With this command, the default node list is changed from ems1, gss_ppc64 (in default section) to ems2.
The changes specified using the -N and -G options are only applicable until the **gssutils** instance is
running.

Restriction with -N and -G options:
- If -N and -G options have been used with **gssutils**, none of the other **gssutils** options can be used
  along with them.
- Some of the ESS server deployment commands do not require -N or -G option, however they still need
  EMS or I/O server node names. Those commands are not a part of core ESS deployment however they
  are a part of other binaries in the ESS deployment toolkit. For example, the **updatenode** command
  belongs to xCAT however this command used to update the node and needs EMS or I/O server node
  name without -N and -G. For example:

  $ updatenode ems1 -P gss_updatenode

```
updatenode ems1 -P gss_updatenode
Update the ESS Management Server (EMS) node and installs required rpms.
This must be run twice with a reboot in between. See QDG for details.
Press c to change nodename and options of the command.
```

  In this example, ems1 is an argument to **updatenode** (xCAT command) without -N. **gssutils** -N and -G
  do not change the values for such command. To change these default values, user must use the
  gssutils customization option, which is described in the **gssutils** *customization* section.

--prefix and --suffix options:

These options can be used in conjunction with any switches with **gssutils** to provide prefix and suffix
names to host name for high speed network. For example:

$ **gssutils** --prefix=hr --suffix=-10g

```
                CHECK SYSTEM HARDWARE AND SOFTWARE


           1. Help
           2. Show node details
           3. Check and validate various install parameters
           4. Quick storage configuration check
           5. Check enclosure cabling and paths to disks
           6. Check disks for IO operations
           7. Ping all nodes
           8. Check ssh to all nodes
           9. Run lsscsi from all nodes
          10. Check for open serviceable events
          11. Back



   /opt/ibm/gss/tools/bin/gssnodedetails -N ems2 --prefix hr --suffix=-10g
   Shows miscellaneous node information.
```

When user provides `--prefix` and `--suffix` option then now all deployment commands also get prefix and suffix.

gssutils customization

gssutils allows customization for specific targeted environment. As a result, you do not need to use the `-N` or `-G` or `--prefix` or `--suffix` option every time while using **gssutils**. One can generate the **gssutils** configuration file specific to their environment using the `--customize` option and the generate a configuration file that can be used subsequently. For example, in an environment where the EMS node is ems2 and I/O server node names are io3 and io4, and suffix -40g is to be specified, the following command needs to be issued.

```
$ gssutils --customize --config /var/tmp/env2 --ems-node-name ems2 \
--io-node-one-name io3 --io-node-two-name io4 --suffix=-40g

Successfully generated the customization configuration file.
```

This command generates a **gssutils** customization file at /var/tmp/env2. Thereafter, **gssutils** can be invoked using the customization file as follows.

```
$ gssutils --config /var/tmp/env2
```

Here, running **gssutils** with a customized configuration file replaced **gssutils** defaults to user provided defaults used while generating the configuration file.

The custom configuration file is capable of replacing the defaults of those commands also which are not a part of the ESS core deployment.

```
updatenode ems2 -P gss_updatenode
Update the ESS Management Server (EMS) node and installs required rpms. This must be run
twice with a reboot in between. See QDG for details. Press c to change nodename and options
of the command.
```

Note: Running **gssutils** with the custom configuration file is the recommended way of using **gssutils**.

The "C" button customization in gssutils

gssutils supports a special kind of customization called the "C" button customization. At any point of time in **gssutils**, users can press the "C" button and change the command inline. This customization allows users to modify the command inline without persisting the customization forever. For example:

$ **gssutils**

```
                    CHECK SYSTEM HARDWARE AND SOFTWARE


                  1. Help
                  2. Show node details
    /opt/ibm/gss/tools/bin/gssnodedetails -N ems2 hr -10g


    Change command option(s) below. Press Ctrl+G to return.
    Nodelist/Group:-N ems2
            Prefix:hr
            Suffix:-10g
       cmd options:█



    /opt/ibm/gss/tools/bin/gssnodedetails -N ems2 --prefix hr --suffix=-10g
    Shows miscellaneous node information.
```

When user presses the "C" button, a dialog opens which allows user to do changes and commit the changes of command using Ctrl+G. Another example is as follows:

```
                  UPDATE EMS NODE TO LATEST FIXES


                     1. Help
                     2. Update Kernel Errata Repo (Optional)

 ┌──────────────────────────────────────────────────────────────────┐
 │ updatenode ems1 -P gss_updatenode                                  │
 │                                                                    │
 │                                                                    │
 │ Change command option(s) below. Press Ctrl+G to return.            │
 │    cmd options:ems1 -P gss_updatenode█                             │
 │                                                                    │
 │                                                                    │
 │                                                                    │
 └──────────────────────────────────────────────────────────────────┘

 ┌──────────────────────────────────────────────────────────────────┐
 │ updatenode ems1 -P gss_updatenode                                  │
 │ Update the ESS Management Server (EMS) node and installs required rpms. │
 │ This must be run twice with a reboot in between. See QDG for details. │
 └──────────────────────────────────────────────────────────────────┘
```

`gssutils` restrictions

- **gssutils** must be invoked from the EMS node and not from I/O server nodes.
- Make sure that you use a console of adequate height and width to specify a large file name while extracting the tar ball. 80 x 24 console size might lead to an unexpected error. If you experience an error, this task can be performed from the shell prompt also.
- **gssutils** must only be invoked from a PuTTY session or an equivalent standard Linux terminal.
  gssutils must not be invoked from an IPMI SOL connection or from an xCAT Remote Console (`rcons`) session. It might not work because of the different terminal type used by IPMI SOL and xCAT Remote Console connections.

Chapter 3. Installing Elastic Storage Server

New installations from manufacturing provide all the necessary pieces to start the deployment.

Inside the /home/deploy directory on the management server node (EMS), the following items are available for the architecture you are deploying:

* Kernel
* Systemd
* Network Manager
* RHEL ISO
* README containing critical xCAT information such as MAC addresses

The xCAT database is intact using the default host name and IP addresses. All nodes are installed and do not need to be re-installed.

Note: It is recommended to use **gssutils** for all aspects of deployment or upgrade.

1. Assuming that the SSR has completed the full check of the system (no bad hardware, device paths, basic networking verified), you have the option to start by using the Plug N Play mode to demonstrate to the customer how fast and easy creating a file system can be and to provide an overview of the GUI. For more information, see "Elastic Storage Server 5.2 or later: Plug-N-Play Mode" on page 26.
2. After Plug-N-Play mode has been demonstrated, use Fusion mode to move the system into the production state. For more information, see "Elastic Storage Server 5.2 or later: Fusion Mode" on page 28.
3. After Fusion mode is complete, you should be ready to setup the GUI and perform the post installation action items needed to complete deployment. For more information, see "Post-installation checklist" on page 6.

ESS installation flow

The following are the legacy installation steps required to complete a new ESS deployment. These steps must be issued from within **gssutils**.

Install the management server software

These steps are mandatory for installation of an ESS system.

Note: The package name depends on the platform and the edition on which you are installing the software.

1. Unpack the ESS software archive (This is contained in the ESS_STD_BASEIMAGE-5.3.0.1-ppc64-Linux.tgz file.

   ```
   tar -zxvf gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z.tgz
   ```
2. Check the MD5 checksum:

   ```
   md5sum -c gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z.md5
   ```
3. Make sure the /opt/ibm/gss/install/rhel7/<ARCH> directory is clean:

   ```
   /bin/sh gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z --remove
   ```

   Depending on the architecture, replace <ARCH> with ppc64 or ppc64le.

17

Note: If you are upgrading to 5.3.x from an earlier release, you might need to clean up the directory structure used in earlier releases. To do so, issue the following command:

```
/bin/sh gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z —-remove —-dir /opt/ibm/gss/install
```

4. Extract the ESS packages and accept the license as follows. By default, it is extracted to the /opt/ibm/gss/install directory:

```
/bin/sh gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z --text-only
```

Install the ESS system

Before proceeding with the following steps, ensure that you have completed all the steps in "Install the management server software " on page 17.

Follow these steps to perform a new installation of the ESS software on a management server node and I/O server nodes. Node host names **ems1**, **gssio1**, and **gssio2** are examples. Each environment could have its own unique naming conventions. For an xCAT command such as **updatenode**, use an xCAT host name. For the IBM Spectrum Scale commands (those start with mm), use an IBM Spectrum Scale host name. For example, **ems1** is an xCAT host name (typically a hostname associated with the management interface) and **ems1-hs** is the corresponding IBM Spectrum Scale host name (typically a host name associated with the high speed interface).

1. Make the **gssdeploy** script executable, if it is already not executable:

   ```
   chmod +x /opt/ibm/gss/install/rhel7/<ARCH>/samples/gssdeploy
   ```

2. Clean the current xCAT installation and associated configuration to remove any preexisting xCAT configuration, and then address any errors before proceeding:

   ```
   /opt/ibm/gss/install/rhel7/<ARCH>/samples/gssdeploy -c
   ```

3. Run one of the following commands depending on the architecture:

   For PPC64BE:

   ```
   cd /var/tmp ; ./gssinstall_ppc64 -u
   ```

   For PPC64LE:

   ```
   cd /var/tmp ; ./gssinstall_ppc64le -u
   ```

4. Run the following command to copy the gssdeploy.cfg.default and customize it for your environment by editing it:

   ```
   cp /var/tmp/gssdeploy.cfg.default /var/tmp/gssdeploy.cfg
   ```

   Note: The directory from which you execute the **gssinstall** script determines where the gssdeploy.cfg.default is stored. It is recommended that you run **gssinstall** script from /var/tmp, but not mandatory.
   Do not copy the **gssdeploy.cfg** configuration file to the /tmp directory because the **gssdeploy** script uses the /tmp/gssdeploy directory and the /tmp directory might get cleaned up in case of a system reboot.

5. If deploying on the **PPC64LE** platform, gather information for the gssdeploy.cfg configuration file using the following commands when you are in close proximity with the rack containing the nodes:

   a. Scan the nodes in the FSP subnet range:

   ```
   /var/tmp/gssdeploy -f FSP_Subnet_Range
   ```

   FSP_Subnet_Range is the FSP management node interface subnet range. For example, 10.0.0.0/24.

   Note:
   - It is recommended to use the IP address 10.0.0.1 for the management interface, if possible.
   - It is highly recommended that you use the /24 netmask because scanning of the subnet takes a considerable duration of time if a wider network range is used.

- The **gssdeploy -f** command first determines if a DHCP server is running on the network. If the DHCP sever is not running, it prompts you to start one so that the I/O server nodes can obtain addresses. Select Y to start the DHCP server when prompted.

  Note:

  This command scans the specified subnet range to ensure that only the nodes on which you want to deploy are available. These include I/O server nodes and management server node (EMS).

  This command also returns the following:

  - Chassis serial numbers and FSP IPMI IP addresses of I/O server nodes in the building block
  - Serial numbers and IP addresses of I/O server nodes in the building block

  Note: Do not proceed to the next step until FSP IP addresses and serial numbers of all known nodes are visible using the **gssdeploy -f** script.

b. Physically identify the nodes in the rack:

   `/var/tmp/gssdeploy -i`

   With the -i option, *Node_IP*, *Default_Password*, and *Duration* need to be provided as input, where:

   - *Node_IP* is the returned FSP IPMI IP address of the node obtained by using the **gssdeploy -f** command.
   - *Default_Password* is the FSP IPMI default password of the node, which is PASSW0RD
   - *Duration* is the time duration in seconds for which the LED on the node should blink.

   After you issue this command, the LED blinks on the specified node for the specified duration. You can identify the node in the rack using the blinking LED.

   Depending on the order of a node in the rack, its corresponding entry is made in the gssdeploy.cfg file. For example, for the bottommost node in the rack, its corresponding entry is put first in gssdeploy.cfg.

6. Update the gssdeploy.cfg file according to your requirements and the gathered information.

   The options that you can specify in the gssdeploy.cfg file include:

   - Whether use DVD for installation: RHEL_USE_DVD

     The default option is to use ISO.
   - If DVD, then device location: RHEL_DVD
   - Mount point to use for RHEL media: RHEL_MNT
   - ISO location: RHEL_ISODIR

     The default location is /opt/ibm/gss/iso.
   - ISO file name: RHEL_ISO
   - EMS host name: EMS_HOSTNAME
   - Network interface for xCAT management network: EMS_MGTNETINTERFACE
   - Network interface for FSP network: FSP_MGTNETINTERFACE [**Not applicable for PPC64BE**]
   - FSP default IPMI password: FSP_PASSWD [**Not applicable for PPC64BE**]
   - HMC host name: HMC_HOSTNAME [**Not applicable for PPC64LE**]
   - HMC default user ID: HMC_ROOTUID [**Not applicable for PPC64LE**]
   - HMC default password: HMC_PASSWD[**Not applicable for PPC64LE**]
   - I/O server user ID: IOSERVERS_UID
   - I/O server default password: IOSERVERS_PASSWD
   - I/O server serial numbers: IOSERVERS_SERIAL [**Not applicable for PPC64BE**]
   - I/O server node names: IOSERVERS_NODES

     For example, gssio1 gssio2
   - Deployment OS image: DEPLOY_OSIMAGE

Note: For PPC64LE, there must be a one-to-one relationship between serial number and node in gssdeploy.cfg and for every node specified in gssdeploy.cfg, there must be a matching entry in /etc/hosts.

7. Copy the RHEL 7.3 ISO file to the directory specified in the gssdeploy.cfg file.

8. Perform precheck to detect any errors and address them before proceeding further:

```
/opt/ibm/gss/tools/samples/gssprecheck -N ems1 --pre --install --file /var/tmp/gssdeploy.cfg
```

Note: gssprecheck gives hints on ways to fix any discovered issues. It is recommended to review each found issue carefully though resolution of all might not be mandatory.

Attention: Power down the storage enclosures, or remove the SAS cables, before running **gssdeploy -x**.

9. Verify that the ISO is placed in the location specified in the gssdeploy.cfg configuration file and then run the **gssdeploy** script:

```
/var/tmp/gssdeploy -x
```

Note: To perform I/O server discovery task this step will power cycle the I/O server nodes specified in the gssdeploy.cfg file.

10. Log out and then log back in to acquire the environment updates.

11. Back up the xCAT database and save it to a location not on the management server node:

```
dumpxCATdb -p /var/tmp/db
tar -zcvf xCATDB-backup.tar.gz /var/tmp/db
```

12. Set up the kernel, systemd, and Network Manager errata repositories. For example, use the following command on PPC64BE systems:

```
/var/tmp/gssdeploy -k /home/deploy/kernel-RHSA-2018-0399-73-BE.tar.gz -p \
/home/deploy/systemd-RHBA-2018-0416-73-BE.tar.gz,/home/deploy/netmgr-RHBA-2017-2925-73-BE.tar.gz --silent
```

Note: This command extracts the supplied tar zip files and builds the associated repository.

- -k option: Set up the kernel repository
- -p option: Set up the patch repository (For example: systemd, network manager). One or more patches might be specified at the same time separated by comma.
- Directory structure:

  Kernel repository

  /install/gss/otherpkgs/rhels7/*<arch>*/kernel

  Patch repository

  /install/gss/otherpkgs/rhels7/*<arch>*/patch

Important: Make sure that all RPMs in the /install directory including the extracted files in the kernel directory (/install/gss/otherpkgs/rhels7/<arch>/kernel), the patch directory (/install/gss/otherpkgs/rhels7/<arch>/patch), and xCAT RPMs, etc. have the correct read permission for user, group, and others (chmod 644 files). For example:

```
/install/gss/otherpkgs/rhels7/<arch>/kernel
-rw-r--r-- 1 nobody nobody 39315448 Nov 29 17:59 kernel-3.10.0-514.44.1.el7.ppc64.rpm
```

```
/install/gss/otherpkgs/rhels7/<arch>/patch
-rw-r--r-- 1 nobody nobody  5412240 Nov 29 12:02 systemd-219-42.el7_4.10.ppc64.rpm
-rw-r--r-- 1 nobody nobody 1785872 Nov 29 12:35 NetworkManager-1.8.0-11.el7_4.ppc64.rpm
```

Wrong file permission will lead to node deployment failure.

13. Update the management server node. Here **ems1** is the xCAT host name. This step installs the kernel, uninstalls OFED, installs IBM Spectrum Scale, and applies the IBM Spectrum Scale profile.

```
updatenode ems1 -P gss_updatenode
```

Use **systemctl reboot** to reboot the management server node and run this step again as shown below. This additional step rebuilds OFED for new kernel and builds GPFS portability layer (GPL) for IBM Spectrum Scale.

```
updatenode ems1 -P gss_updatenode
```

Note: You can use the **-V** option with the **updatenode** command for a more verbose output on the screen for a better understanding of failures, if any.

14. Update OFED on the management server node:

```
updatenode ems1 -P gss_ofed
```

15. Update the IP RAID Adapter firmware on the management server node:

```
updatenode ems1 -P gss_ipraid
```

16. Use **systemctl reboot** to reboot the management server node.

Deploy the I/O server nodes

1. Before initiating the deployment of the I/O server nodes, do the following:
   a. Verify that the running kernel level is at the desired level (for example, 514.44.1) using the **uname -a** command.
   b. Verify that there are no repository errors using the **yum repolist** command.
   c. Ensure that the attached storage enclosures are powered off.
2. Run the **gssinstallcheck** script:

```
gssinstallcheck -N ems1
```

By default, **gssinstallcheck** runs on all nodes sequentially. You can run **gssinstallcheck** parallely from the management server node as follows.

```
# xdsh ems1,gss_ppc64 "/opt/ibm/gss/tools/bin/gssinstallcheck -N localhost" | xcoll -n
```

For more information, see Appendix L, "Running gssinstallcheck in parallel," on page 83.

This script is used to verify IBM Spectrum Scale profile, OFED, and kernel, etc.
   a. Check for any error with the following:
      1) Installed packages
      2) Linux kernel release
      3) OFED level
      4) IPR SAS FW
      5) IPR SAS queue depth
      6) System firmware
      7) System profile setting
      8) Host adapter driver

Ignore other errors that may be flagged by the **gssinstallcheck** script. They will go away after the remaining installation steps are completed.

3. Run the **gssprecheck** script in full install mode and address any errors:

```
/opt/ibm/gss/tools/samples/gssprecheck -N ems1 --install --file /var/tmp/gssdeploy.cfg
```

Note: gssprecheck gives hints on ways to fix any discovered issues. It is recommended to review each found issue carefully though resolution of all might not be mandatory.

4. Deploy on the I/O server nodes using the customized deploy script:

```
./gssdeploy -d
```

5. After a duration of about five minutes, run the following command:

```
nodestat gss_ppc64
```

After running the command, the output displays the OS image name or packages being installed. For example:

PPC64LE installations:

```
node: rhels7.3-ppc64le-install-gss
node: rhels7.3-ppc64le-install-gss
```

PPC64BE installations:

```
node: rhels7.3-ppc64-install-gss
node: rhels7.3-ppc64-install-gss
```

After about 30 minutes, the following output displays:

```
node: sshd
node: sshd
```

The installation is complete when **nodestat** displays sshd for all I/O server nodes. Here **gss_ppc64** is the xCAT node group containing I/O server nodes. To follow the progress of a node installation, you can tail the console log by using the following command:

```
tailf /var/log/consoles/NodeName
```

where *NodeName* is the node name.

Note: Make sure the xCAT post-installation script is complete before rebooting the nodes. You can check xCAT post process running on the I/O server nodes as follows:

```
xdsh gss_ppc64 "ps -eaf | grep -v grep | grep xcatpost"
```

If there are any processes still running, wait for them to complete.

6. At the end of the deployment, wait for approximately five minutes and reboot the node:

```
xdsh gss_ppc64 systemctl reboot
```

7. Once rebooted, verify the installation by running **gssinstallcheck:**

```
gssinstallcheck -G ems1,gss_ppc64
```

Check for any error with the following:

a. Installed packages

b. Linux kernel release

c. OFED level

d. IPR SAS FW

e. IPR SAS queue depth

f. System firmware

g. System profile setting

h. Host adapter driver

Ignore other errors that may be flagged by the **gssinstallcheck** script. They will go away after the remaining installation steps are completed.

Check the system hardware

After the I/O server nodes have been installed successfully, power on the storage enclosures and then wait for at least 10 minutes from power on for discovery to complete before moving on to the next step. Here is the list of key log files that should be reviewed for possible problem resolution during deployment.

- By default /var/log/message log from all I/O server nodes are directed to the message log in the EMS node.

- The **gssdeploy** log is located at /var/log/gss

- The xCAT log is located at /var/log/xcat

- Console outputs from the I/O server node during deployment are located at `/var/log/consoles`

1. Update the `/etc/hosts` file with high-speed hostname entries in the management server node and copy the modified `/etc/hosts` file to the I/O server nodes of the cluster as follows:

   ```
   xdcp gss_ppc64 /etc/hosts /etc/hosts
   ```

2. Run **gssstoragequickcheck**:

   ```
   gssstoragequickcheck -G gss_ppc64
   ```

3. Run **gssfindmissingdisks**:

   ```
   gssfindmissingdisks -G gss_ppc64
   ```

   If **gssfindmissingdisks** displays an error, run **mmgetpdisktopology** and save the output. Run **topsummary** using the saved output on each I/O server node to obtain more information about the error:

   ```
   mmgetpdisktopology > /var/tmp/NODE_top.out
   topsummary /var/tmp/NODE_top.out
   ```

4. Run **gsscheckdisks**:

   ```
   GSSENV=INSTALL gsscheckdisks -G gss_ppc64 --encl all --iotest a --write-enable
   ```

   Attention: When run with **--iotest w** (write) or **--iotest a** (all), **gsscheckdisks** will perform write I/O to the disks attached through the JBOD. This will overwrite the disks and will result in the loss of any configuration or user data stored on the attached disks. **gsscheckdisks** should be run only during the installation of a building block to validate that read and write operations can be performed to the attached drives without any error. The **GSSENV** environment variable must be set to **INSTALL** to indicate that **gsscheckdisks** is being run during installation.

5. Check for any hardware serviceable events and address them as needed. To view the serviceable events, issue the following command:

   ```
   gssinstallcheck -N ems1,gss_ppc64 --srv-events
   ```

   If any serviceable events are displayed, you can obtain more information by using the `--platform-events EVENTLIST` flag.

   Note: During the initial deployment of the nodes on the PPC64BE platform, SRC BA15D001 might be logged as a serviceable event by Partition Firmware. This is normal and should be cleared after the initial deployment. For more information, see Appendix B, "Known issues," on page 47.

Note: Configure the node to connect to the Red Hat Network and apply the latest security patches, if needed.

Set up the high-speed network

Customer networking requirements are site-specific. The use of bonding to increase fault-tolerance and performance is advised but guidelines for doing this have not been provided in this document. Consult with your local network administrator before proceeding further.

- To set up bond over IB, run the following command.

  ```
  gssgennetworks -G ems,gss_ppc64 --create --ipoib --suffix=-hs --mtu 4092
  ```

  In this example, MTU is set to 4092. Consult your network administrator for the proper MTU setting.

- To set up bond over Ethernet, run the following command.

  ```
  gssgennetworks -N ems1,gss_ppc64 --suffix=-hs --create-bond
  ```

Create the cluster, recovery groups, and file system

1. Create the GPFS cluster:

```
gssgencluster -C test01 -G gss_ppc64 --suffix=-hs --accept-license
```

In this example, `test01` is used as the cluster name and `-hs` is used as the suffix of the host name.

2. Verify healthy network connectivity:

```
xdsh gss_ppc64 /usr/lpp/mmfs/bin/mmnetverify
```

3. Create the recovery groups:

```
gssgenclusterrgs -G gss_ppc64 --suffix=-hs
```

4. Create the vdisks, NSDs, and file system:

```
gssgenvdisks --create-vdisk --create-nsds --create-filesystem --contact-node gssio1
```

Note: gssgenvdisk, by default, creates data vdisk with 8+2p RAID code and 8MB block size, and metadata vdisk with 3WayReplication and 1MB block size. These default values can be changed to suitable values for the customer environment.

5. Add the management server node to the cluster:

```
gssaddnode -N ems1 --cluster-node gssio1 --suffix=-hs --accept-license --no-fw-update
```

In this example, the management server hostname is `ems1` with a suffix of `-hs` (`ems1-hs`) in the high-speed network. The **--no-fw-update** option is used because the management server node does not contain a SAS adapter or attached drives.

Check the installed software and system health

1. Run **gssinstallcheck** parallely from the management server node.

```
# xdsh ems1,gss_ppc64 "/opt/ibm/gss/tools/bin/gssinstallcheck -N localhost" | xcoll -n
```

By default, **gssinstallcheck** runs on all nodes sequentially. For more information, see Appendix L, "Running gssinstallcheck in parallel," on page 83.

Note: When **gssinstallcheck** in run in parallel, you might get an error for the system firmware.

2. Shut down GPFS in all nodes and reboot all nodes.

   a. Shut down GPFS all nodes:
   ```
   mmshutdown -a
   ```

   b. Reboot all server nodes:
   ```
   xdsh gss_ppc64 "systemctl reboot"
   ```

   c. Reboot the management server node:
   ```
   systemctl reboot
   ```

3. After reboots, run the following command (**Not applicable for PPC64LE**):

```
gssinstallcheck -G gss_ppc64 --phy-mapping
```

Ensure that the phy mapping check is OK.

4. Restart GPFS in all nodes and wait for all nodes to become active:

```
mmstartup -a
```

5. Mount the filesystem and perform a stress test. For example, run:

```
mmmount gpfs0 -a
gssstress /gpfs/gpfs0 gssio1 gssio2
```

In this example, **gssstress** is invoked on the management server node. It is run on I/O server nodes `gssio1` and `gssio2` with `/gpfs/gpfs0` as the target path. By default gssstress runs for 20 iterations and can be adjusted using the -i option (type **gssstress** and press Enter to see the available options). During the I/O stress test, check for network error by running from another console:

```
gssinstallcheck -N ems1,gss_ppc64 -—net-errors
```
6. Perform a health check. Run:

```
gnrhealthcheck
/usr/lpp/mmfs/bin/mmhealth node show -N all --verbose
```

Address any issues that are identified.
7. Check for any open hardware serviceable events and address them as needed. The serviceable events can be viewed as follows:

```
gssinstallcheck -N ems1,gss_ppc64 --srv-events
```

If any serviceable events are displayed, you can obtain more information by using the `--platform-events EVENTLIST` flag.

Note: During initial deployment of the nodes, SRC BA15D001 may be logged as serviceable event by Partition Firmware. This is normal and should be cleared after the initial deployment. For more information, see Appendix B, "Known issues," on page 47.
8. Verify that NTP is set up and enabled.
   a. On the management server node verify that `/etc/ntp.conf` is pointing to the management server node itself over the management interface.
   b. Restart NTP daemon on each node.
   ```
   xdsh <ems>,gss_ppc64 "systemctl restart ntpd"
   ```
   c. Verify that NTP is setup correctly by running the following checks:
      • Verify that offset is 0.
      ```
      xdsh ems1,gss_ppc64 "ntpq -p"
      ```
      • Verify that NTP is enabled and synchronized.
      ```
      xdsh ems1,gss_ppc64 "timedatectl status" | grep -i NTP
      ```
      • Verify that the timezone is set correctly on each node.
      ```
      xdsh ems1,gss_ppc64 "timedatectl status"  | grep -i zone
      ```

Install the ESS GUI

Important: Complete all of the following steps carefully including the steps for configuring **mmperfmon** and restricting certain sensors to the management server node (EMS) only.
1. Generate performance collector on the management server node by running the following command. The management server node must be part of the ESS cluster and the node name must be the node name used in the cluster (e.g., ems1-hs).

```
mmperfmon config generate --collectors ems1-hs
```
2. Set up the nodes in the *ems nodeclass* and *gss_ppc64 nodeclass* for performance monitoring by running the following command.
```
mmchnode --perfmon -N ems,gss_ppc64
```
3. Start the performance monitoring sensors by running the following command.
```
xdsh ems1,gss_ppc64 "systemctl start pmsensors"
```
4. Capacity and fileset quota monitoring is not enabled in the GUI by default. You must correctly update the values and restrict collection to the management server node only.
   a. To modify the GPFS Disk Capacity collection interval, run the following command:
   ```
   mmperfmon config update GPFSDiskCap.restrict=EMSNodeName
             GPFSDiskCap.period=PeriodInSeconds
   ```
   The recommended period is 86400 so that the collection is done once per day.

b. To restrict GPFS Fileset Quota to run on the management server node only, run the following command:

```
mmperfmon config update GPFSFilesetQuota.period=600 GPFSFilesetQuota.restrict=EMSNodeName
```

Here the *EMSNodeName* must be the name shown in the **mmlscluster** output.

Note: To enable quota, the filesystem quota checking must be enabled. Refer **mmchfs -Q** and **mmcheckquota** commands in the *IBM Spectrum Scale: Command and Programming Reference*.

5. Verify that the values are set correctly in the performance monitoring configuration by running the **mmperfmon config show** command on the management server node. Make sure that GPFSDiskCap.period is properly set, and GPFSFilesetQuota and GPFSDiskCap are both restricted to the management server node only.

Note: If you are moving from manual configuration to auto configuration then all sensors are set to default. Make the necessary changes using the **mmperfmon** command to customize your environment accordingly. For information on how to configure various sensors using **mmperfmon**, see Manually installing IBM Spectrum Scale GUI.

6. Start the performance collector on the management server node:

```
systemctl start pmcollector
```

7. Enable and start gpfsgui:

```
systemctl enable gpfsgui.service
systemctl start gpfsgui
```

8. To launch the ESS GUI in a browser, go to: `https://EssGuiNode` where ESSGuiNode is the host name or IP address of the management server node for GUI access. To log in, type admin in the User Name field and your password in the Password field on the login page. The default password for admin is admin001. Walk through each panel and complete the GUI Setup Wizard.

This completes the installation task of the ESS system. After completing the installation, apply security updates available from Red Hat.

For information on applying optimized configuration settings to a set of client nodes or a node class, see "Adding IBM Spectrum Scale nodes to an ESS cluster" on page 65.

Elastic Storage Server 5.2 or later: Plug-N-Play Mode

Enabling the Plug N Play using **gssutils** is highly encouraged. For more information, see Chapter 2, "**gssutils** - ESS Installation and Deployment Toolkit," on page 9.

The goal of the Plug-N-Play mode is to allow customers to build a cluster, file system and begin sampling the GUI as soon as possible. The stated goal is for this to be achieved in under an hour after lab-based services (LBS) starts working on the system. Manufacturing now ships EMS with xCAT preconfigured with default settings.

Prerequisites
- Unpacking and basic power connectivity are completed.
- FSP and xCAT networks are set up in documented ports and they are connected to proper VLANs.
- SSRs have done validation checks using **gssutils** to ensure correct disk placement, cabling, networking, and server health.
- Access to EMS is available over SSH for LBS.

Option #1

The primary option is to build a very generic environment to allow the customer to preview their working Elastic Storage Server (ESS) system as fast as possible with the assumption that the final

customizations are coming later. This gives the customers an opportunity to see their storage subsystem working right away. They start to get familiar with the installation process and the available file system space, start deciding on file system and block sizes, and become familiar with the GUI.

Some basic health checks are also run in this mode that give LBS confidence that the actual installation will go smoothly:
- Default manufacturing host name, IPs, user IDs, passwords
- Networking over the 1Gb (provisioning) only. For more information, see "Option #2."
- Basic hardware checking:
  - `gsssstoragequickcheck`
  - `gssfindmissingdisks`
  - `gsscheckdisks`
- Basic file system creation (use of entire space, 8M/1M block size, 8+2p RAID code)
- GUI and performance monitoring setup

Option #2

The secondary option is to start the process quickly to move the system into an actual installation state. There are several upfront items that need to be decided upon to choose this option. The result is a system that already has the actual host names, IPs, domain, netmasks, and potentially the high-speed connections. The disadvantage of going with option #2 is that you might not have all this information. Since the main goal of the Plug-N-Play mode is speed, the primary mode must be option #1 which allows the customer to start using ESS as fast as possible.

Requirements for option #2
- All customer host name, IPs, netmasks, domain name must be known
- Optional: The high-speed network items must be known and connected properly to the switch. The switch must be configured correctly for bonding.

Work flow

1. System arrives at customer site; Basic unpacking and connectivity established; All nodes powered on to the operating system.
2. SSRs arrive and do full hardware check using `gssutils`. They replace any bad components prior to LBS arrival.
3. Prior to arrival, LBS asks the following questions in association with the customer:
   - **Option 1**: Do I want to bring this system up as fast as possible with defaults (1Gb network, default host name, default cluster/FS settings) to show the customer how fast we can bring the system and begin using it (play with the GUI, look at capacity, etc)? I may or may not have the true host name and IPs.
   - **Option 2**: Do I have the actual host name and IPs including confidence that the high-speed network is cabled up and ready to go?

   Both options are previews to the customer. The only difference is that how much upfront information and confidence do you have in the information and environment at an early stage.

   Basic assumptions:
   - EMS has xCAT connection in T3 (1Gb card).
   - All nodes have FSP connections in the HMC 1 port.
   - On PPC64BE, HMC is properly configured with connections to the FSP and xCAT networks.
   - On PPC64LE, EMS has an extra FSP connection in the T2 port (1Gb card).
   - All standard VLANS (xCAT, FSP) are set up properly.
4. LBS logs in to EMS through SSH.

5. If customer is ready to change the xCAT VLAN IP information, use the following commands:

   a. Copy the gsschenv.cfg from /opt/ibm/gss/tools/conf to /opt/ibm/gss/tools/bin.

   b. Modify the **gsschenv.cfg** file.

   ```
   [root@ems2 conf]# cat gsschenv.cfg
   # Modify the following
   # HOSTNAME_ORIG = Original host name in your xCAT ESS environment
   # HOSTNAME_NEW = The new hostname (1 to 1 with the HOSTNAME_ORIG)
   # IP_ADDR_NEW = The new IPs you want (1 to 1 with HOSTNAME_NEW/ORIG)
   # NETMASK = The new netmask associated with the IPs
   # DOMAIN = The new domain associated with the IPs
   HOSTNAME_ORIG=(ems1 gssio1 gssio2)
   HOSTNAME_NEW=(modems1 modgssio1 modgssio2)
   IP_ADDR_NEW=(192.168.202.20 192.168.202.21 192.168.202.22)
   NETMASK="255.255.0.0"
   DOMAIN="gpfs.net"
   ```

6. Run **gsschenv** to modify your environment.

   ```
   cd /opt/ibm/gss/tools/bin ; ./gsschenv --modify
   ```

7. After the environment is updated, a default /etc/hosts file is created on EMS. If you have the high-speed host name and IPs, add them to this file. After updating, copy /etc/hosts to all the I/O nodes.

   ```
   xdcp gss_ppc64 /etc/hosts
   ```

8. Proceed to running the standard set of ESS verification checks.

   - **gssstoragequickcheck**
   - **gssfindmissingdisks**
   - **gsscheckdisks**

   For more information, see man pages of these commands.

9. Create your network bonds (if going this route) using **gssgennetworks** and test through **gssnettest**. If simply using the 1Gb network at this point then continue.

10. Create your cluster using **gssgencluster**, either over low or high-speed network. Use the **--no-fw-update** option.

11. Create your recovery groups.

12. Create your file system:

    - If customer is using the high-speed network, now is a good opportunity to have them create multiple file systems of different block sizes. This way they can start running workload and deciding what works best for them when the production environment is actually built.
    - If using the 1Gb network for pure speed purposes, it is best to use the default values.

13. Add EMS using **gssaddnode**.

14. Set up the performance monitoring collector and sensors. For more information, see this section.

15. Start the GUI.

Conclusion

At this point, the customer must be able to do several tasks with their new ESS system. At a minimum, they should be able to mount the file system, view free space, and use the GUI. The best case scenario is that they already have the host names and IPs set up for xCAT and they are able to do estimates of proper block size and file system sizes. This mode shows how fast an ESS system can be brought up and used at a customer site.

Elastic Storage Server 5.2 or later: Fusion Mode

Enabling the Fusion mode using **gssutils** is highly encouraged. For more informatin, see Chapter 2, "**gssutils** - ESS Installation and Deployment Toolkit," on page 9.

The goal of the Fusion mode is to no longer require that Elastic Storage Server (ESS) systems be rediscovered or redeployed at a customer site.

Prerequisites

All of the prerequisites for any ESS installation apply here.

End Goal

The end goal of this mode is to greatly reduce the time and the complexity in bringing up an ESS system. There are several tasks that you no longer have to perform:

- No need for **gssdeploy -x**: No need to install and rediscover the nodes through xCAT
- No need for **gssdeploy -d**: No need to reinstall the I/O nodes with Red Hat Enterprise Linux

Everything is treated as an upgrade and the amount of time saved significantly goes up if the system was shipped with the latest levels. This is achieved by shipping xCAT preconfigured out of manufacturing and providing a new tool (**gsschenv**) which automatically changes your IPs, host names, domain, and netmask.

This mode is called Fusion because it mixes parts of the upgrade and installation flows. The flow is all upgrade until the cluster creation. After cluster creation, it turns into installation because the cluster, file system, and GUI etc. need to be set up.

Plug-N-Play mode considerations

The Plug-N-Play mode can be used in conjunction with the Fusion mode. The best combination is to use Plug-N-Play to quickly bring up a system for the customer to experiment with. This shows how fast a cluster can be created, and file system, and GUI can be set up. This also allows the customer to potentially make decisions early. For example. the number and size of the file systems and the block size. After Plug-N-Play, LBS can begin using the Fusion mode to quickly bring the system into production after all final decisions are made.

Work Flow

1. Stop the GUI on EMS using **systemctl stop gpfsgui**.
2. Wipe the GUI database clean.
   ```
   su -l postgres -c 'psql -d postgres -c' "drop schema fscc cascade"'
   ```
3. Unmount the file systems.
   ```
   mmumount all -a
   ```
4. SSH to one of the I/O nodes and delete the data and metadata vdisks.
   ```
   /opt/ibm/gss/tools/samples/gssdelvdisks
   ```
5. Delete the log vdisks using **mmdelvdisk**.
   You can query the log vdisks with **mmdelvdisk**.
6. Delete the recovery groups using **mmdelrecoverygroup**.
   You can query the recovery groups using **mmdelrecoverygroup**.
7. Shut down GPFS.
   ```
   mmshutdown -a
   ```
8. Delete the cluster.
   ```
   mmdelnode -a
   ```
9. Break the network bonds on each node.
   ```
   cd /etc/sysconfig/network-scripts ; rm -f *bond*
   nmcli c reload
   ```
   If host names were already changed during Plug-N-Play, skip the next step (**gsschenv**).

10. Change xCAT IPs, host names, domain, and netmasks.
    a. Copy the gsschenv.cfg from /opt/ibm/gss/tools/conf to /opt/ibm/gss/tools/bin.
    b. Modify the **gsschenv.cfg**.

    ```
    [root@ems2 conf]# cat gsschenv.cfg
    # Modify the following
    # HOSTNAME_ORIG = Original hostnames in your xCAT ESS environment
    # HOSTNAME_NEW = The new hostname (1 to 1 with the HOSTNAME_ORIG)
    # IP_ADDR_NEW = The new IPs you want (1 to 1 with HOSTNAME_NEW/ORIG)
    # NETMASK = The new netmask associated with the IPs
    #DOMAIN = The new domain associated with the IPs
    HOSTNAME_ORIG=(ems1 gssio1 gssio2)
    HOSTNAME_NEW=(modems1 modgssio1 modgssio2)
    IP_ADDR_NEW=(192.168.202.20 192.168.202.21 192.168.202.22)
    NETMASK="255.255.0.0"
    DOMAIN="gpfs.net"
    ```

11. Run **gsschenv** to modify your environment.

    cd /opt/ibm/gss/tools/bin; ./gsschenv --modify

12. After the environment is updated, a default /etc/hosts file is created on EMS. If you have the high-speed host name and IPs, add them to this file. After updating, copy /etc/hosts to all the I/O nodes.

    `xdcp gss_ppc64 /etc/hosts`

13. Compare the installed ESS version to the version from Fix Central you are attempting to install. In case of a new system, it should be the same.

    `gssinstallcheck -G ems1,gss_ppc64 --get-version`

    - **If the versions matched first**, do a verification using **gssinstallcheck**.

      `gssinstallcheck -G ems1,gss_ppc64`

      Note: There is no GPFS cluster at this point so cluster configuration checks will fail. Assuming this check is clean, proceed to the steps in "Check the system hardware" on page 22. Continue from this point and complete the rest of the installation steps.

    - **If the versions did not match:**
      – Perform the steps in "Install the management server software " on page 17.
      – Proceed with the steps in Chapter 4, "Upgrading Elastic Storage Server," on page 31. Do all the steps till the "Update the management server node" on page 33 procedure (including this procedure), apart from **gssdeploy -c -r** or **gssdeploy -x -r**.

    - After EMS is updated, do a verification using **gssinstallcheck**.

      `gssinstallcheck -G ems1,gss_ppc64`

      Note: There is no GPFS cluster at this point so cluster configuration checks will fail. Assuming this check is clean, proceed to the steps in "Check the system hardware" on page 22. Continue from this point and complete the rest of the installation steps. This is because you still need to create the network links, cluster, etc. as if this were an installation.

Conclusion

The Fusion mode is a way of reducing a few pain points in ESS. No longer requiring LBS to discover the nodes or reinstall should be a significant help when setting up new systems.

Chapter 4. Upgrading Elastic Storage Server

These are the legacy steps required to complete a new ESS upgrade. These steps must be issued from within **gssutils**.

Note:
- It is recommended that you upgrade the Power 8 firmware before you upgrade each node.
- The HMC (PPC64BE) can be upgraded at anytime without disrupting a production system.

During the upgrade process if a step fails, it must be addressed before moving to the next step. Follow these steps to perform an upgrade of the ESS system.

Note: For considerations and instructions to upgrade a cluster that contains ESS and protocol nodes, see Appendix A, "Upgrading a cluster containing ESS and protocol nodes," on page 39. You can decide when to upgrade the ESS system in such a cluster. You can either upgrade protocol nodes first and then the ESS system or you can upgrade the ESS system first, followed by the protocol nodes.

Prerequisites

Before you begin the upgrade procedure, do the following:
- Obtain the ESS tarball, kernel, systemd, RHEL ISO, and network manager packages for the architecture being used.
- Archive the current contents of `/home/deploy` and move the 5.3.0.1 packages there.
- Make sure that the RHEL ISO is moved to `/opt/ibm/gss/iso` or the location specified in the `gssdeploy.cfg` file).
- Disable the subscription manager and any external repositories by issuing the following commands on each node that you want to upgrade:

  ```
  subscription-manager config --rhsm.manage_repos=0
  yum clean all
  ```
- Understand the implications of upgrading the release level to LATEST and upgrading the file system format version. After you complete the upgrade to the latest code level, you cannot revert to the previous code level. For more information, see Completing the migration to a new level of IBM Spectrum Scale.

Prepare the system for upgrade

1. Perform a health check by issuing the following command:

   ```
   gnrhealthcheck
   ```

   Address any issues that are identified.
2. Verify network connectivity and node health by issuing the following commands:

   ```
   xdsh ems1,gss_ppc64 /usr/lpp/mmfs/bin/mmnetverify
   /usr/lpp/mmfs/bin/mmhealth node show -N all
   ```
3. Wait for any of these commands that are performing file system maintenance tasks to complete:
   mmadddisk
   mmapplypolicy
   mmcheckquota
   mmdeldisk
   mmfsck
   mmlssnapshot

> **mmrestorefs**
> **mmrestripefile**
> **mmrestripefs**
> **mmrpldisk**

For information on upgrade considerations specific to functions used in a cluster containing ESS and protocol nodes, see "Planning upgrade in a cluster containing ESS and protocol nodes" on page 39.

4. Stop the creation and deletion of snapshots using **mmcrsnapshot** and **mmdelsnapshot** during the upgrade window.

Upgrading from ESS 5.1.x, ESS 5.2.x, or ESS 5.3.0

Perform the following steps if you are upgrading from ESS 5.1.x, ESS 5.2.x, or ESS 5.3.0:

1. Check for any hardware serviceable events:

   ```
   gssinstallcheck -G ems1,gss_ppc64 --srv-events
   ```

   Address any hardware issues identified in the serviceable events. If any serviceable events are displayed, you can obtain more information by using the `--platform-events EVENTLIST` flag.

2. Check for any deployment errors:

   ```
   gssinstallcheck -G ems1,gss_ppc64
   ```

3. Unpack the ESS software archive (This is contained in the `ESS_STD_BASEIMAGE-5.3.0.1-ppc64-Linux.tgz` file.

   ```
   tar -zxvf gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z.tgz
   ```

4. Check the MD5 checksum:

   ```
   md5sum -c gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z.md5
   ```

5. Make sure the `/opt/ibm/gss/install/rhel7/<ARCH>` directory is clean:

   ```
   /bin/sh gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z --remove
   ```

   Depending on the architecture, replace *<ARCH>* with `ppc64` or `ppc64le`.

   Note: If you are upgrading to 5.3.x from an earlier release, you might need to clean up the directory structure used in earlier releases. To do so, issue the following command:

   ```
   /bin/sh gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z --remove --dir /opt/ibm/gss/install
   ```

6. Extract the ESS packages and accept the license as follows. By default, it is extracted to the `/opt/ibm/gss/install` directory:

   ```
   /bin/sh gss_install-5.3.0.1_ppc64le_standard_20180412T022648Z --text-only
   ```

7. Make the **gssdeploy** script executable:

   ```
   chmod +x /opt/ibm/gss/install/rhel7/<arch>/samples/gssdeploy
   ```

8. If upgrading from ESS 5.1.x or ESS 5.2.x, perform cleanup and save a backup copy of the xCAT database:

   ```
   /opt/ibm/gss/install/rhel7/<arch>/samples/gssdeploy -c -r /var/tmp/xcatdb
   ```

9. Run one of the following commands depending on the architecture.

   For PPC64BE:

   ```
   cd /var/tmp ; ./gssinstall_ppc64 -u
   ```

   For PPC64LE:

   ```
   cd /var/tmp ; ./gssinstall_ppc64le -u
   ```

10. Run the following command to copy the `gssdeploy.cfg.default` and customize it for your environment by editing it:

    ```
    cp /var/tmp/gssdeploy.cfg.default /var/tmp/gssdeploy.cfg
    ```

Note: The directory from which you execute the **gssinstall** script determines where the gssdeploy.cfg.default is stored. It is recommended that you run **gssinstall** script from /var/tmp, but not mandatory.

Do not copy the gssdeploy.cfg configuration file to the /tmp directory because the **gssdeploy** script uses the /tmp/gssdeploy directory and the /tmp directory might get cleaned up in case of a system reboot.

11. Customize the gssdeploy.cfg configuration file according to your environment. For information about the contents of gssdeploy.cfg, see "Install the ESS system" on page 18.

Update the management server node

1. On the management server node, stop GUI services, and save performance monitoring collector and sensor configuration files:

   ```
   systemctl stop gpfsgui
   ```

2. Copy the RHEL 7.3 ISO file to the directory specified in the gssdeploy.cfg file.

3. If upgrading from ESS 5.1.x or ESS 5.2.x, install tools and xCAT and restore the xCAT database:

   ```
   /var/tmp/gssdeploy -x -r /var/tmp/xcatdb
   ```

4. Perform precheck to detect any errors and address them before proceeding further:

   ```
   /opt/ibm/gss/tools/samples/gssprecheck -N ems1 --upgrade --file /var/tmp/gssdeploy.cfg
   ```

   Note: gssprecheck gives hints on ways to fix any discovered issues. It is recommended to review each found issue carefully though resolution of all might not be mandatory.

5. Shut down IBM Spectrum Scale on the management server node while making sure quorum is still maintained:

   ```
   mmshutdown
   ```

6. Set up the kernel, systemd, and Network Manager errata repositories. For example, use the following command on PPC64BE systems:

   ```
   /var/tmp/gssdeploy -k /home/deploy/kernel-RHSA-2018-0399-73-BE.tar.gz -p \
   /home/deploy/systemd-RHBA-2018-0416-73-BE.tar.gz,/home/deploy/netmgr-RHBA-2017-2925-73-BE.tar.gz --silent
   ```

   Note: This command extracts the supplied tar zip files and builds the associated repository.

   - -k option: Set up the kernel repository
   - -p option: Set up the patch repository (For example: systemd, network manager). One or more patches might be specified at the same time separated by comma.
   - Directory structure:

     Kernel repository

     /install/gss/otherpkgs/rhels7/*<arch>*/kernel

     Patch repository

     /install/gss/otherpkgs/rhels7/*<arch>*/patch

   Important: Make sure that all RPMs in the /install directory including the extracted files in the kernel directory (/install/gss/otherpkgs/rhels7/<arch>/kernel), the patch directory (/install/gss/otherpkgs/rhels7/<arch>/patch), and xCAT RPMs, etc. have the correct read permission for user, group, and others (chmod 644 files). For example:

   ```
   /install/gss/otherpkgs/rhels7/<arch>/kernel
   -rw-r--r-- 1 nobody nobody 39315448 Nov 29 17:59 kernel-3.10.0-514.44.1.el7.ppc64.rpm
   ```

   ```
   /install/gss/otherpkgs/rhels7/<arch>/patch
   -rw-r--r-- 1 nobody nobody  5412240 Nov 29 12:02 systemd-219-42.el7_4.10.ppc64.rpm
   -rw-r--r-- 1 nobody nobody 1785872 Nov 29 12:35 NetworkManager-1.8.0-11.el7_4.ppc64.rpm
   ```

   Wrong file permission will lead to node deployment failure.

7. Update the management server node:

```
updatenode ems1 -P gss_updatenode
```

Use **systemctl reboot** to reboot the management server node and complete this step again as follows:

```
updatenode ems1 -P gss_updatenode
```

This additional step rebuilds OFED for the new kernel and builds GPFS Portability Layer (GPL) for IBM Spectrum Scale, if required.

8. Update OFED on the management server node:

```
updatenode ems1 -P gss_ofed
```

9. Update IP RAID Adapter firmware on the management server node:

```
updatenode ems1 -P gss_ipraid
```

10. Use **systemctl reboot** to reboot the management server node.

11. Perform the following steps to upgrade IBM Spectrum Scale RAID configuration parameters.

```
/opt/ibm/gss/tools/samples/gssupg530.sh -b ems1-hs,gss_ppc64
/opt/ibm/gss/tools/samples/gssupg530.sh -c
```

12. Start IBM Spectrum Scale on the management server node:

```
mmstartup
```

13. Verify that IBM Spectrum Scale is in the active state before upgrading the I/O server nodes:

```
mmgetstate
```

Do not proceed if the system is not active.

14. Ensure that the management server node is fully updated and active:

```
gssinstallcheck -N ems1
```

Update the I/O server nodes

Repeat the following steps for each I/O server node, one node at a time.

1. Before shutting down GPFS on any I/O server node, run precheck from the management server node:

```
/opt/ibm/gss/tools/samples/gssprecheck -N IO_NODE --upgrade --file /var/tmp/gssdeploy.cfg
```

Note: gssprecheck gives hints on ways to fix any discovered issues. It is recommended to review each found issue carefully though resolution of all might not be mandatory.

2. Move the cluster and file system manager role to another node if the current node is a cluster manager or file system manager.

a. To find the cluster and file system managers, run:

```
mmlsmgr
```

b. To change the file system manager, run:

```
mmchmgr gpfs0 gssio2-hs
```

In this example, gssio2-hs is the new file system manager of file system gpfs0.

c. To change the cluster manager, run:

```
mmchmgr -c gssio2-hs
```

In this example, gssio2-hs is the new cluster manager.

3. Move the recovery group in the current I/O server node to the peer I/O server node in the same building block.

a. To list the recovery groups, run:

```
mmlsrecoverygroup
```

b. To list the active server, primary server, and secondary server, run:

```
mmlsrecoverygroup rg_gssio1-hs -L | grep active -A2
```

c. To move the recovery group from the current active I/O server node (`rg_gssio1-hs`) to the peer I/O server node (`gssio2-hs`) in the same building block, run the following commands in the shown order:

```
mmchrecoverygroup rg_gssio1-hs --active gssio2-hs
```

```
mmchrecoverygroup rg_gssio1-hs --servers gssio2-hs,gssio1-hs
```

4. After confirming that the recovery group has been successfully moved to the peer I/O server node, unmount all GPFS file systems if mounted, and shut down IBM Spectrum Scale on the current I/O server node while maintaining quorum:

```
mmunmount all -N CurrentIoServer-hs
```
```
mmshutdown -N CurrentIoServer-hs
```

5. Run **updatenode**:

```
updatenode CurrentIoServer -P gss_updatenode
```

6. Reboot the I/O server node and complete this step again if you are instructed to do so in the **updatenode** output. Reboot the I/O server node as follows:

```
xdsh CurrentIoServer "systemctl reboot"
```

7. Run **updatenode** again (if instructed to do so):

```
updatenode CurrentIoServer -P gss_updatenode
```

8. Update OFED.

```
updatenode CurrentIoServer -P gss_ofed
```

9. Update IP RAID FW in the I/O Server node that is being upgraded.

```
updatenode CurrentIoServer -P gss_ipraid
```

10. Reboot the I/O server node as follows:

```
xdsh CurrentIoServer "systemctl reboot"
```

11. Update the SAS host adapter firmware on *CurrentIoServer*:

```
CurrentIoServer$ mmchfirmware --type host-adapter
```

Here CurrentIOServer is an I/O server node and the command is run on the I/O server node.

12. Update the node configuration:

```
/opt/ibm/gss/tools/samples/gssupg530.sh -s CurrentIoServer-hs
```

This command is run from the EMS node.

13. On PPC64BE systems, run phy check and ensure that the phy mapping is OK:

```
gssinstallcheck -N CurrentIoServer --phy-mapping
```

14. Start IBM Spectrum Scale on the I/O server node:

```
mmstartup -N CurrentIoServer-hs
```

Once the IBM Spectrum Scale daemon is successfully started, move back the recovery group that was moved to the peer I/O server node of the same building block in Step 3c above. Move back the cluster manager and the file system manager if required that was moved to the other nodes in step 2.

15. Wait until the I/O server can be seen active from the management server node, using the following command:

```
mmgetstate
```

The management server must be already running for issuing this command.

16. Run **gssinstallcheck** from the management server node:

```
        gssinstallcheck -N IO_NODE
```
17. Repeat preceding steps for the peer I/O server node of the same building block.
18. Repeat all steps in this section for each additional building block.
19. Update the configuration of the nodeclass of I/O server nodes gss_ppc64 and the management server nodeclass ems.

    ```
    /opt/ibm/gss/tools/samples/gssupg530.sh -s gss_ppc64
    /opt/ibm/gss/tools/samples/gssupg530.sh -c ems
    ```

Update the enclosure and drive firmware

1. To update the storage enclosure firmware, run one of the following commands from one I/O Server node of each building block.

   - When upgrade is being performed concurrently:

     ```
     CurrentIoServer$ mmchfirmware --type storage-enclosure
     ```

   - When upgrade is being performed non-concurrently, all attached enclosures can be upgraded in parallel.

     ```
     mmchfirmware --type storage-enclosure -N gss_ppc64
     ```

     Note: The IBM Spectrum Scale daemon must be down on all nodes of the node class gss_ppc64 for parallel upgrade.

2. To update the drive firmware, run the following command from **each** I/O Server node of each building block:

   ```
   CurrentIoServer$ mmchfirmware --type drive
   ```

   The drive update can take some time to complete. You can update the drives more quickly by taking the system offline (shutting down IBM Spectrum Scale) and using the **--fast-offline** option.

Check the installed software and system health

1. Perform a health check:

   ```
   gnrhealthcheck
   /usr/lpp/mmfs/bin/mmhealth node show -N all --verbose
   ```

2. Check for any hardware serviceable events and address them as needed. To view the serviceable events, issue the following command:

   ```
   gssinstallcheck -N ems1,gss_ppc64 --srv-events
   ```

   If any serviceable events are displayed, you can obtain more information by using the --platform-events EVENTLIST flag.

   Note: During the initial deployment of the nodes, SRC BA15D001 might be logged as a serviceable event by Partition Firmware. This is normal and should be cleared after the initial deployment. For more information, see Appendix B, "Known issues," on page 47.

Note: Some of these steps might fail if they are already implemented in previous versions of ESS. If you see any failures indicating **mmperfmon** has already been configured, ignore these failure messages and continue with the remaining steps.

Upgrading GUI

Perform the following steps to upgrade the GUI:

Note: Some of these steps might fail if the GUI is already set up. However, it is important to rerun the upgrade steps using the latest changes.

1. Generate performance collector on the management server node by running the following command. The management server node must be part of the ESS cluster and the node name must be the node name used in the cluster (e.g., ems1-hs).

   ```
   mmperfmon config generate --collectors ems1-hs
   ```
2. Set up the nodes in the *ems nodeclass* and *gss_ppc64 nodeclass* for performance monitoring by running the following command.
   ```
   mmchnode --perfmon -N ems,gss_ppc64
   ```
3. Start the performance monitoring sensors by running the following command.
   ```
   xdsh ems1,gss_ppc64 "systemctl start pmsensors"
   ```
4. Capacity and fileset quota monitoring is not enabled in the GUI by default. You must correctly update the values and restrict collection to the management server node only.
   a. To modify the GPFS Disk Capacity collection interval, run the following command:
      ```
      mmperfmon config update GPFSDiskCap.restrict=EMSNodeName
              GPFSDiskCap.period=PeriodInSeconds
      ```
      The recommended period is 86400 so that the collection is done once per day.
   b. To restrict GPFS Fileset Quota to run on the management server node only, run the following command:
      ```
      mmperfmon config update GPFSFilesetQuota.period=600 GPFSFilesetQuota.restrict=EMSNodeName
      ```
      Here the *EMSNodeName* must be the name shown in the **mmlscluster** output.

      Note: To enable quota, the filesystem quota checking must be enabled. Refer **mmchfs -Q** and **mmcheckquota** commands in the *IBM Spectrum Scale: Command and Programming Reference*.
5. Verify that the values are set correctly in the performance monitoring configuration by running the **mmperfmon config show** command on the management server node. Make sure that GPFSDiskCap.period is properly set, and GPFSFilesetQuota and GPFSDiskCap are both restricted to the management server node only.

   Note: If you are moving from manual configuration to auto configuration then all sensors are set to default. Make the necessary changes using the **mmperfmon** command to customize your environment accordingly. For information on how to configure various sensors using **mmperfmon**, see Manually installing IBM Spectrum Scale GUI.
6. Start the performance collector on the management server node:
   ```
   systemctl start pmcollector
   ```
7. Enable and start gpfsgui:
   ```
   systemctl enable gpfsgui.service
   systemctl start gpfsgui
   ```
8. To launch the ESS GUI in a browser, go to: https://EssGuiNode where ESSGuiNode is the hostname or IP address of the management server node for GUI access. To log in, type admin in the User Name field and your password in the Password field on the login page. The default password for admin is admin001. Walk through each panel and complete the GUI Setup Wizard.

After the GUI is up and running, do the following:

1. Enable the subscription manager by issuing the following commands on the upgraded nodes:
   ```
   subscription-manager config --rhsm.manage_repos=1
   yum clean all
   ```

This completes the upgrade task of the ESS system. For information on applying optimized configuration settings to a set of client nodes or a node class, see "Adding IBM Spectrum Scale nodes to an ESS cluster" on page 65.

Appendix A. Upgrading a cluster containing ESS and protocol nodes

The procedure for upgrading a cluster containing ESS and protocol nodes comprises several phases. Although the protocol node upgrade procedure is detailed here, the same procedure can be tweaked and used for client and NSD nodes as well.

1. "Planning upgrade in a cluster containing ESS and protocol nodes"
2. "Performing upgrade prechecks" on page 41
3. "Upgrading protocol nodes by using the installation toolkit" on page 44
4. "Upgrading OFED, OS, and kernel errata on protocol nodes" on page 44

   This phase comprises the following steps.

   a. Uninstalling OFED
   b. Upgrading OS and rebooting the system
   c. Upgrade kernel and rebooting the system
   d. Upgrading firmware
   e. Build the GPFS portability layer
   f. Installing OFED and rebooting the system

5. "Upgrading ESS" on page 46
6. "Upgrading HMC in PPC64BE systems" on page 46

Planning upgrade in a cluster containing ESS and protocol nodes

Before scheduling an upgrade of a cluster containing ESS and protocol nodes, planning discussions must take place to know the current cluster configuration and to understand which functions might face an outage.

The planning phase comprises the following steps.

1. Note all products and functions currently installed in the cluster that is being upgraded.

   Important: Any function that is actively accessing files on a specific node that is undergoing upgrade might prevent a file system from properly unmounting and thus prevent IBM Spectrum Scale from unloading kernel modules, which is required for RPM updates. If this occurs, the upgrade can be resumed after doing the following steps:

   a. Reboot the node that could not unload kernel modules properly.
   b. Verify that there are no mixed versions of `gpfs.base` or `gpfs.ext` on the node. Resolve this issue by manually upgrading the RPMs that are not at the latest level.
   c. Verify that SMB on this node is not at a level differing from SMB on other nodes within the cluster. Resolve this issue by either upgrading or downgrading the SMB version on this node to match other nodes within the cluster.
   d. Bring the node online and resume the upgrade.

   The following list contains the functions and products that must be considered for upgrade depending on your environment. Upgrade considerations for some of these functions and products are also listed.

   • SMB: Requires quiescing all I/O for the duration of the upgrade. Due to the SMB clustering functionality, differing SMB levels cannot co-exist within a cluster at the same time. This requires a full outage of SMB during the upgrade.

- NFS: Recommended to quiesce all I/O for the duration of the upgrade. NFS experiences I/O pauses, and depending upon the client, mounts and disconnects during the upgrade.
- Object: Recommended to quiesce all I/O for the duration of the upgrade. Object service will be down or interrupted at multiple times during the upgrade process. Clients might experience errors or they might be unable to connect during this time. They should retry as appropriate.
- CES Groups: Follow SMB, NFS, and Object advice for quiescing I/O for the duration of the upgrade.
- TCT: Requires the cloud gateway service to be stopped on all TCT nodes prior to the upgrade by using the following command: `mmcloudgateway service stop -N` *Node | NodeClass*.
- AFM: Active AFM file transfers might hold open the file system.
- ILM: Recommended to quiesce or pause all ILM policies that might be set to trigger during an upgrade window.
- Restripe: Recommended to stop any `mmrestripefs` process prior to an upgrade. For a list of commands that might perform file system maintenance tasks, see "Prepare the system for upgrade" on page 31.
- Snapshot creation or deletion: Recommended to stop or pause any policies that might create or delete snapshots during an upgrade window. For a list of commands that might perform file system maintenance tasks, see "Prepare the system for upgrade" on page 31.
- IBM Spectrum Protect™ : All `mmbackup` operations must be quiesced prior to the upgrade.
- DMAPI flag for file systems: If the cesSharedRoot file system is DMAPI enabled, all HSM services must be stopped prior to upgrade by using `dsmmigfs stop` and `systemctl stop hsm`.

  Furthermore, the installation toolkit upgrade process might fail while attempting to remount cesSharedRoot. This is because HSM processes must be restarted for the file system to mount. Perform this manually if the installation toolkit fails:
  a. Start the HSM service: `systemctl start hsm.service`
  b. Start the HSM daemons: `dsmmigfs start`
  c. Mount the cesSharedRoot file system: `mmmount cesSharedRoot -N` *cesNodes*
  d. Restart the installation toolkit upgrade process.
- IBM Spectrum Archive™ EE: If IBM Spectrum Archive is enabled, do the following steps to upgrade.

  Note: For latest information about IBM Spectrum Archive EE commands, refer to *IBM Spectrum Archive EE documentation on IBM Knowledge Center*.
  a. Stop IBM Spectrum Archive (LTFS) by issuing the following command on all IBM Spectrum Archive EE nodes.

     `ltfsee stop`
  b. Unmount the media by issuing the following command on all IBM Spectrum Archive EE nodes.

     `umount /ltfs`
  c. Deactivate failover operations by issuing the following command on all IBM Spectrum Archive EE nodes.

     `dsmmigfs disablefailover`
  d. Stop the HSM daemons by issuing the following command on all IBM Spectrum Archive EE nodes.

     `dsmmigfs stop`
  e. Stop the HSM service by issuing the following command on all IBM Spectrum Archive EE nodes.

     `systemctl stop hsm.service`
  f. Upgrade using the installation toolkit. For more information, see "Upgrading protocol nodes by using the installation toolkit" on page 44.
  g. Upgrade IBM Spectrum Archive EE, if needed.

h. Start the HSM service by issuing the following command on all IBM Spectrum Archive EE nodes.

    `systemctl start hsm.service`

i. Start the HSM daemons by issuing the following command on all IBM Spectrum Archive EE nodes.

    `dsmmigfs start`

j. Activate failover operations by issuing the following command on all IBM Spectrum Archive EE nodes.

    `dsmmigfs enablefailover`

k. Mount the media by issuing the following command on all IBM Spectrum Archive EE nodes.

    `ltfs -o devname=DEVICE /ltfs`

l. Start IBM Spectrum Archive (LTFS) by issuing the following command on all IBM Spectrum Archive EE nodes.

    `ltfsee start`

For information on how to manage a DMAPI enabled cesSharedRoot file system, see the entry *DMAPI flag for file systems* in this list.

- Encryption
- cNFS
- GUI - How many and which nodes?
- Performance monitoring collectors - How many and where are they located?
- Performance monitoring sensors - Are they installed on all nodes?
- Which nodes run more than one of these functions?

2. Understand the source version and the number of hops needed to move to the target code version across all nodes and functions.
   - For information on IBM Spectrum Scale upgrade paths, see Supported online upgrade paths.
   - For information on ESS upgrade paths, see "Supported upgrade paths" on page 5.

3. Understand if the IBM Spectrum Scale installation toolkit can be used on the protocol nodes and also understand how the installation toolkit performs the upgrade. For information about installation toolkit limitations, see Limitations of the installation toolkit.

   Note: This instruction set assumes that the installation toolkit is being used for protocol nodes.

4. Set expectations for functional currency and outages. For more information, see IBM Spectrum Scale FAQ.

5. Obtain the necessary packages.

6. Decide the upgrade sequence.

7. Decide whether operating system, driver, or firmware updates are needed on protocol nodes. This includes OFED, Power® firmware, x86 firmware. When making this decision, be aware that tools normally used within ESS might not be available to assist with automating these efforts outside of the ESS nodes.

Performing upgrade prechecks

The precheck phase assists with the planning phase and it can be done on a cluster without any harm. It might be useful to run through the precheck steps the day before an upgrade is scheduled, or earlier, to guard against any unexpected situations that might lead to an upgrade failure.

1. Identify the protocol nodes to be upgraded.

2. Verify that a base OS repository exists, and if it does not exist, configure one on all protocol nodes. The repository must reflect the current OS or kernel on the nodes and not the version being upgraded to. The OS upgrade will be done after the IBM Spectrum Scale upgrade using the installation toolkit.

a. Check the current OS and kernel level.

```
uname -a
cat /etc/*release*
```

b. Check the existing repositories.

```
yum repolist
```

c. Pick a file to install for testing the repository and use **yum install** to try the installation. Enter no on the confirmation prompt. If the repository for your current base OS does not exist, create one as follows.

1) Mount the installation image by issuing one of the following command depending on the media type.

- If using an ISO, issue this command:

```
mount -o loop RHEL7.3.iso /mnt
```

- If using DVD media, issue this command:

```
mount -o loop /dev/sr0  /mnt
```

2) Copy the `media.repo` file from the /mnt directory to /etc/yum.repos.d and change its permissions.

```
cp /mnt/media.repo /etc/yum.repos.d/rhel73dvd.repo
chmod 644 /etc/yum.repos.d/rhel73dvd.repo
```

3) Open the file in an editor.

```
vi /etc/yum.repos.d/rhel73dvd.repo
```

4) Configure the following settings in the file.

```
gpgcheck=1
enabled=1
baseurl=file:///mnt/
gpgkey=file:///etc/pki/rpm-gpg/RPM-GPG-KEY-redhat-release
```

5) Clear the caches.

```
yum clean all
subscription-manager clean
```

6) Verify that you can obtain the packages list from the repository.

```
yum list --noplugins
```

Note: When creating a repository, make sure to give it a new name that is different than any existing repositories. This reduces the possibility of yum caching errors.

Note: These steps will vary for SLES 12 nodes and zypper repositories.

3. Verify and, if needed, configure passwordless SSH on all protocol nodes and ESS nodes (EMS and I/O server nodes).

Important: Verify the following combinations.

- All nodes must be able to SSH to themselves and all other nodes using the IP
- All nodes must be able to SSH to themselves and all other nodes using the host name
- All nodes must be able to SSH to themselves and all other nodes using the FQDN

If your IBM Spectrum Scale cluster version is 4.2.2 or later, you can use the **mmnetverify** command to check for passwordless SSH access among nodes.

4. Verify that the contents of the /etc/hosts file on each protocol nodes are in the following format:

```
<IP> <FQDN> <alias>
```

5. Verify that firewall ports required for the necessary functions are open. For more information, see Securing the IBM Spectrum Scale system using firewall.

Important: It is recommended to upgrade the non-ESS nodes such as protocol nodes and NSD nodes to IBM Spectrum Scale 5.0.0.2. The following steps show an upgrade to version 5.0.0.2.

6. Download the new IBM Spectrum Scale self-extracting package using the sub-steps and then place it on the protocol node that you plan to designate as the installer node.

   a. Go to the IBM Spectrum Scale page on Fix Central, select the new **spectrumscale** package and then click **Continue**.

   b. Choose the download option **Download using Download Director** to download the new **spectrumscale** package and place it in the wanted location on the install node.

7. Extract the new IBM Spectrum Scale self-extracting package by using the package name (for example, /tmp/Spectrum_Scale_Protocols_Standard-5.0.0.2_x86_64-Linux_install).

   This creates a new directory structure (/usr/lpp/mmfs/5.0.0.2/).

8. In the installation toolkit, enter the configuration to mirror the current cluster configuration.

   - Do not input the EMS or I/O nodes from the ESS system.

   - If the installation toolkit was previously used, the old clusterdefinition.txt file can be copied to the new code location as follows.

     ```
     cp –p /usr/lpp/mmfs/4.2.3.6/installer/configuration/clusterdefinition.txt \
     /usr/lpp/mmfs/5.0.0.2/installer/configuration
     ```

   - When specifying the installer node, specify the setup type ess.

     ```
     /usr/lpp/mmfs/5.0.0.2/installer/spectrumscale setup -s 192.168.10.1 -st ess
     ```

   - Specify the EMS node.

     ```
     /usr/lpp/mmfs/5.0.0.2/installer/spectrumscale node add -e EMSNode
     ```

   - Input any protocol or non-protocol nodes on which you plan to use the installation toolkit:

     ```
     /usr/lpp/mmfs/5.0.0.2/installer/spectrumscale node add
     ```

   - Input the existing CES shared root file system into the installation toolkit:

     ```
     /usr/lpp/mmfs/5.0.0.2/installer/spectrumscale config protocols -f cesshared -m /ibm/cesshared
     ```

   - Input the existing CES IPs (**mmces address list**) into the installation toolkit:

     ```
     /usr/lpp/mmfs/5.0.0.2/installer/spectrumscale config protocols -e CESIP1,CESIP2,CESIP3
     ```

   - It is not required to input NSD or file system information.

   - Enable performance monitoring reconfiguration to ensure that sensors are also upgraded during the upgrade.

     ```
     ./spectrumscale config perfmon -r on
     ```

   - If the installation toolkit must be setup from scratch, you can refer to this example:

```
./spectrumscale setup -s 192.168.10.1 -st ess          ## IP of installer node that all other nodes can get to
                                                       ## with setup type as ESS
./spectrumscale node add -e EMSNode                    ## Adds the EMS node
./spectrumscale node add node1.gpfs.net –a -p          ## designates this node as the node
                                                       ## that runs mm commands for the installer.
                                                       ## Also designates it as a protocol node
./spectrumscale node add node2.gpfs.net -p
./spectrumscale node add node3.gpfs.net –p
./spectrumscale node add node4.gpfs.net                ## example of a client node
./spectrumscale enable smb                             ## if SMB is active
./spectrumscale enable nfs                             ## if NFS is active
./spectrumscale enable object                          ## if Object is active
./spectrumscale config protocols -e CESIP1,CESIP2,CESIP3       ## CES-IPs gathered from mmces address list
./spectrumscale config protocols -f cesshared -m /ibm/cesshared   ## FS name and mount point for CES shared root
./spectrumscale config perfmon –r on                   ## turn on perfmon recofig only for upgrade to 5.0.0.2
./spectrumscale node list                              ## list out the node config afterwards
./spectrumscale config protocols                       ## shows the protocol config
```

   Note: The installation toolkit can be used for tasks other than upgrade, such as adding new protocols and protocol nodes. If you are planning to do this in the future, you will need to expand the preceding example to input configuration details necessary for each future action. For more information, see Protocols Quick Overview Guide.

9. Run the installation toolkit upgrade precheck.

```
./spectrumscale upgrade --precheck
```

A successful precheck implies that you are ready for using the installation toolkit to perform the upgrade.

10. Double check networking and bonding modes that are in use and save this information in case it is needed later.

11. Check the level of OFED drivers and place the latest OFED package on the nodes, if using Infiniband adapters.

12. Make all possible attempts to quiesce all I/O for Object, SMB, and NFS prior to the upgrade. For information on upgrade considerations for these and other functions, see "Planning upgrade in a cluster containing ESS and protocol nodes" on page 39.

Upgrading protocol nodes by using the installation toolkit

Use these steps to upgrade protocol nodes by using the installation toolkit.

Before proceeding with the upgrade using the installation toolkit, ensure that the toolkit is set up. For more information, see this step of the precheck task.

This phase of the upgrading a cluster containing ESS and protocol nodes procedure is dependent on the successful completion of the planning and precheck phases.

1. Run the installation toolkit upgrade precheck.

   ```
   ./spectrumscale upgrade --precheck
   ```

   If the precheck is successful, proceed to the next step.

   Attention: Make all possible attempts to quiesce all I/O for Object, SMB, and NFS prior to the upgrade. For information on upgrade considerations for these and other functions, see "Planning upgrade in a cluster containing ESS and protocol nodes" on page 39.

2. Run the installation toolkit upgrade procedure.

   ```
   ./spectrumscale upgrade
   ```

   When this procedure is done, components including base GPFS and protocols will have been upgraded on all protocol nodes that were specified to the installation toolkit. This step does not need to be repeated on each node unless only a subset of nodes were specified to the installation toolkit.

   If performance monitoring was not configured correctly on non-ESS nodes before upgrade, then the upgrade does not automatically fix this. In this case, it is advised to rerun the installation.

3. If performance monitoring was not configured correctly on non-ESS nodes before upgrade, rerun the installation toolkit installation procedure.

   ```
   ./spectrumscale install
   ```

   This step sets all non-ESS nodes to performance monitoring nodes, enables protocol sensors, and sets values.

4. If you are using the object protocol, check that object sensors are properly configured.

   ```
   mmperfmon config show
   ```

Upgrading OFED, OS, and kernel errata on protocol nodes

Use these steps to upgrade OFED, OS, and kernel errata on protocol nodes as part of upgrading a cluster containing ESS and protocol nodes.

This phase is not required but it is advisable to match OFED, OS, and kernel errata across all nodes within a cluster to help with performance and to ease debugging. As a part of this procedure, ensure the following:

• Always upgrade IBM Spectrum Scale on protocol nodes prior to OFED, OS, and kernel errata.

- If kernel errata is for a new OS (RHEL7.3 vs RHEL7.2), always update the OS before the kernel errata.
- When taking nodes offline to update OFED, OS, and kernel errata, ensure the following:
  - Quorum does not break
  - Enough NSD nodes remain up to access NSDs
  - The remaining nodes can handle the desired workload

Repeat the following steps on each node.
1. Uninstall the OFED drivers as follows.
   a. Obtain and extract OFED drivers.
   b. Suspend CES on the node being upgraded.
      ```
      mmces node suspend -N NodeBeingUpgraded
      ```
   c. Shut down GPFS on the node being upgraded.
      ```
      mmshutdown -N NodeBeingUpgraded
      ```
   d. Find the uninstallation script within the OFED driver package and execute it on the node being upgraded..
      ```
      mount —o loop mellanox_iso_name /media
      cd /media
      ./uninstall.sh
      ```
2. Create a local repository for the OS upgrade. For information on creating a base RHEL repository, see these steps in the precheck task.

   A repository must be created so that the OS can be upgraded. This repository can be DVD or ISO based. Make sure that you remove any repositories pointing to old OS versions.
3. Upgrade the OS.
   ```
   yum upgrade
   ```
   Review the `yum upgrade` output for any errors that might need to be resolved prior to rebooting and ensure that a clean yum upgrade operation was completed and that it was successful. Reboot the node after OS upgrade.
   ```
   shutdown -r now
   ```
4. Update the kernel errata. Reboot the node after kernel errata update.
   ```
   shutdown -r now
   ```
5. Update the Power8 and x86 firmware. For information on updating Power8 firmware, see Appendix G, "Updating the system firmware," on page 67 and ESS Installation and Deployment Blog.

   x86 firmware update is dependent on the manufacturer, model, and type.
6. Build the GPFS portability layer using the **mmbuildgpl** command. For more information, see Building the GPFS portability layer.
7. Install the latest OFED drivers.

   Note: Do this step only after the OS and kernel are at the latest levels. The OFED level is tied to the kernel so if the kernel changes afterwards, this step might need to be repeated.
   a. Create an updated ISO file for the currently active kernel.
      ```
      mount —o loop mellanox_iso_name /media
      /media/mlnx_add_kernel_support.sh —m /media --make-iso —y --distro rhel7.3 --kmp
      ```
      Ensure that the Linux distribution matches exactly.
   b. Install the OFED drivers from the newly created ISO.
      ```
      umount /media
      mount —o loop newlybuilt_iso_name /media
      cd /media
      ./mlnxofedinstall —q --force
      ```

   Reboot the node after the driver update.
   ```
   shutdown -r now
   ```

8. Verify that GPFS is active on the node and then resume CES.

```
mmgetstate -a
mmces node resume -N NodeBeingUpgraded
mmces node list
mmces service list -a
mmces address list
```

9. Repeat the preceding steps on all non-ESS nodes that the EMS does not upgrade.

Upgrading ESS

While upgrading a cluster containing ESS and protocol nodes, an upgrade of the ESS system itself might occur either before or after the upgrade of protocol nodes. If not yet done, proceed with an upgrade of the ESS system.

For detailed information on the ESS upgrade procedure, see Chapter 4, "Upgrading Elastic Storage Server," on page 31.

Upgrading HMC in PPC64BE systems

While upgrading a cluster containing ESS and protocol nodes, if you had not upgraded HMC before applying the Power8 system firmware, you can proceed with upgrading HMC after the ESS upgrade completes.

For more information, see Appendix H, "Upgrading the Hardware Management Console (HMC)," on page 69.

Appendix B. Known issues

This topic describes known issues for ESS.

ESS 5.3.x issues

The following table describes known issues in ESS 5.3.x and how to resolve these issues. Depending on which fix level you are installing, these might or might not apply to you.

Table 2. Known issues in ESS 5.3.x

Issue	Environment affected	Description	Resolution or action

The **gssgennetworks** script requires high-speed host names to be derived from I/O server (xCAT) host names using suffix, prefix, or both.	High-speed network generation Type: Install Version: All Arch: All Affected nodes: I/O server and EMS nodes	**gssgennetworks** requires that the target host name provided in -N or -G option are reachable to create the high-speed network on the target node. If the xCAT node name does not contain the same base name as the high-speed name you might be affected by this issue. A typical deployment scenario is: gssio1 // xCAT name gssio1-hs // high-speed An Issue scenario is: gssio1 // xCAT name foo1abc-hs // high-speed name	Create entries in the /etc/hosts with node names that are reachable over the management network such that the high-speed host names can be derived from it using some combination of suffix and/or prefix. For example, if the high-speed host names are foo1abc-hs, goo1abc-hs:
			1. Add foo1 and goo1 to the /etc/hosts using management network address (reachable) in the EMS node only.
			2. Use: **gssgennetworks -N foo1,goo1 – suffix abc-hs --create-bond**
			3. Remove the entries foo1 and goo1 from the /etc/hosts file on the EMS node once the high-speed networks are created.
			
			Example of how to fix (/etc/hosts):
			// Before
			`<IP><Long Name><Short Name> 192.168.40.21`
			`gssio1.gpfs.net gssio1 192.168.40.22`
			`gssio2.gpfs.net gssio2 X.X.X.X`
			`foo1abc-hs.gpfs.net foo1abc-hs X.X.X.Y`
			`goo1abc-hs.gpfs.net goo1abc-hs`
			// Fix
			`192.168.40.21 gssio1.gpfs.net gssio1 foo1`
			`192.168.40.22 gssio2.gpfs.net gssio2 goo1`
			`X.X.X.X foo1abc-hs.gpfs.net foo1abc-hs`
			`X.X.X.Y goo1abc-hs.gpfs.net goo1abc-hs`
			`gssgennetworks -N foo1, goo1 --suffix=abc-hs --create-bond`

Running **gssutils** over PuTTY might shows horizontal lines as "qqq" and vertical lines as "xxx".	ESS Install and Deployment Toolkit Type: Install or Upgrade Version: All Arch: All Affected Nodes: EMS and I/O server nodes	PuTTY translation default Remote Character set UTF-8 might not translate horizontal line and vertical character sets correctly.	1. On the PuTTY terminal Window > Translation, change Remote character set from UTF-8 to ISO-8859-1:1998 (Latin-1, West Europe) (this should be the first option after UTF-8). 2. Open session.

Table 2. Known issues in ESS 5.3.x (continued)

Issue	Environment affected	Description	Resolution or action
`gssinstallcheck` might flag an error regarding page pool size in multi-building block situations if the physical memory sizes differ.	Software Validation Type: Install or Upgrade Arch: Big Endian or Little Endian Version: All Affected nodes: I/O server nodes	`gssinstallcheck` is a tool introduced in ESS 3.5, that helps validate software, firmware, and configuration settings. If adding (or installing) building blocks of a different memory footprint installcheck will flag this as an error. Best practice states that your I/O servers must all have the same memory footprint, thus pagepool value. Page pool is currently set at ~60% of physical memory per I/O server node. Example from **gssinstallcheck**: [ERROR] pagepool: found 142807662592 expected range 147028338278 - 179529339371	1. Confirm each I/O server node's individual memory footprint. From the EMS, run the following command against your I/O xCAT group: **xdsh gss_ppc64 "cat/ proc/meminfo | grep MemTotal"** **Note:** This value is in KB.If the physical memory varies between servers and/or building blocks, consider adding memory and re-calculating pagepool to ensure consistency. 2. Validate the pagepool settings in IBM Spectrum Scale: **mmlsconfig | grep -A 1 pagepool** **Note:** This value is in MB. If the pagepool value setting is not roughly ~60% of physical memory, then you must consider recalculating and setting an updated value. For information about how to update the pagepool value, see IBM Spectrum Scale documentation on IBM Knowledge Center.
The GUI might display the long-waiters warning: `Spectrum Scale long-waiters monitoring returned unknown result`	GUI Type: Upgrade Arch: Big Endian Version: All Affected nodes: All	Upon new installs (or upgrades) to ESS 5.3.x, the GUI might show an error due to a bad return code from `mmhealth` in its querying of long-waiters information. `/usr/lpp/mmfs/bin/ mmdiag --deadlock Failed to connect to file system daemon: No such process RC=50`	There is no current workaround but it is advised to verify on the command line that no long-waiters exist. If the system is free from this symptom, mark the event as read on the GUI by clicking under the Action column. Doing so will clear the event.

Table 2. Known issues in ESS 5.3.x (continued)

Issue	Environment affected	Description	Resolution or action
Creating small file systems in the GUI (below 16G) will result in incorrect sizes	GUI Type: Install or Upgrade Arch: Big Endian or Little Endian Version: All Affected nodes: All	When creating file systems in the GUI smaller than 16GB (usually done to create CES_ROOT for protocol nodes) the size will come out larger than expected.	There is currently no resolution. The smallest size you might be able to create is 16GB. Experienced users might consider creating a customer `vdisk.stanza` file for specific sizes you require. You can try one of the following workarounds: • Use three-way replication on the GUI when creating small file systems. • Use **gssgenvdisks** which supports the creation of small file systems especially for CES_ROOT purposes (Refer to the `--crcesfs` flag).
Creating file systems in the GUI might immediately result in lack of capacity data	GUI Type: Install or Upgrade Arch: Big Endian or Little Endian Version: All Affected nodes: All	When creating file systems in the GUI you might not immediately see the capacity data show up.	You may wait up to 24 hours for the capacity data to display or simply use the command line which should accurately show the file system size.

The GUI might show 'unknown' hardware states for storage enclosures and Power 8 servers in the ESS building block.
Part info and firmware levels under the **Hardware Details** panel might also be missing.
Upon adding ESS PPC64LE building-blocks to an existing PPC64BE environment, you might encounter this same issue.
	GUI
	Type: Upgrade
	Arch: Big Endian
	Version: All
	Affected nodes: All
		The ESS GUI (running on the EMS) might show 'unknown' under the Hardware panel for the ESS building block members.
		
		The ESS GUI might also be missing information under Part Info and Firmware version within the **Hardware Details** panel.
			The workaround for this issue is the following:
			1. Login to the EMS
			2. Run the following in order:
			`/usr/lpp/mmfs/gui/cli/runtask RECOVERY_GROUP`
			`/usr/lpp/mmfs/gui/cli/runtask DISK_ENCLOSURES`
			`/usr/lpp/mmfs/gui/cli/runtask ENCLOSURE_FW`
			`/usr/lpp/mmfs/gui/cli/runtask CHECK_FIRMWARE`
			
			After running, the GUI should refresh with the issues resolved.
			Note: If this issue is encountered on adding ESS PPC64LE building-blocks to an existing PPC64BE environment, there is no current workaround as the GUI does not support multiple xCATs.

Canceling disk replacement through GUI leaves original disk in unusable state	GUI Type: Install or Upgrade Arch: Big Endian or Little Endian Version: All Affected nodes: I/O server nodes	Canceling a disk replacement can lead to an unstable system state and must not be performed. However, if you did this operation, use the provided workaround.	Do not cancel disk replacement from the GUI. However, if you did, then use the following command to recover the disk took state: `mmchpdisk <RG> --pdisk <pdisk> --resume`

Table 2. Known issues in ESS 5.3.x (continued)

Issue	Environment affected	Description	Resolution or action
Under **Monitoring** > **Hardware details**, you might see enclosures missing location information.	GUI Type: Install or Upgrade Arch: Big Endian or Little Endian Version: All Affected nodes: N/A	After install or upgrade to ESS 5.3.x, you might see missing location information for the enclosures in your system. This does not reflect the true frame U location which can be observed in the **Monitoring** > **Hardware details** panel.	The current workaround is to wait up to 24 hours for the GUI services to refresh. After this period you will see the enclosure location information fill in.
The GUI wizard might start again after completing the initial setup.	GUI Type: Install Arch: Big Endian Version: All Affected nodes: N/A	After completing the GUI wizard setup on ESS 5.3.x PPC64BE, you might see the start screen again.	If you see the GUI wizard start screen a second time, type the address of the EMS into the browser and press enter. `https://<ip of EMS over management network>` You will then be taken to the GUI home screen.
Upon upgrades to ESS 5.3.x, you might notice missing pools and users in the **Monitoring** > **Capacity** GUI panel	GUI Type: Upgrade Arch: All Version: All Affected nodes: N/A	You might notice one or more missing pools or users after upgrading to ESS 5.3.x in the **Monitoring** > **Capacity** GUI panel. You may also see missing capacity and throughput data under the **Monitoring** > **Nodes** panel.	There is currently no resolution or workaround.Try waiting 24 hours for the GUI to refresh. You can also try clicking Refresh.
Upon upgrades to ESS 5.3.x, you might see several Mellanox OFED `weak-updates` and `unknown symbols` messages on the console during `gss_updatenode`.	OFED Type: Upgrade Arch: Big Endian and Little Endian Version: All Affected nodes: N/A	When building the new OFED driver against the 44.1 kernel, you might see many messages such as `weak-updates` and `unknown symbols`.	There is currently no resolution or workaround. These messages can be ignored.
During firmware upgrades on PPC64LE, `update_flash` might show the following warning: `Unit kexec.service could not be found.`	Firmware Type: Installation or Upgrade Arch: Little Endian Version: All Affected nodes: N/A		This warning can be ignored.

Table 2. Known issues in ESS 5.3.x (continued)

Issue	Environment affected	Description	Resolution or action

Setting target node names within **gssutils** might not persist for all panels. The default host names, such as ems1 might still show.	Deployment
	
	Type: Install or Upgrade
	
	Arch: Big Endian or Little Endian
	
	Version: All
	
	Affected nodes: All
		gssutils allows users to conveniently deploy, upgrade, or manage systems within a GUI-like interface. If you run **gssutils –N** *NODE*, it must store that node name and use it throughout the menu system. There is a bug that might prevent this from working as designed.	Use on one of the following resolutions:
			• Change the JSON file directly as follows.
			  1. Issue this command.
			     `/opt/ibm/gss/tools/bin/gssutils \`
			     `--customize --ems-node-name ems2 \`
			     `--config /tmp/LE.json`
			  2. Issue this command.
			     `/opt/ibm/gss/tools/bin/gssutils \`
			     `--config /tmp/LE.json`
			• For any given command within **gssutils**, you can press "c" to customize. At that point, you can change the node name(s) accordingly.

The GUI wizard might fail due to an error when issuing `mmaddcomp`.	GUI Type: Install Arch: Big Endian or Little Endian Version: All Affected nodes: N/A	During the GUI wizard setup users might hit an error similar to the following. `ERROR: column name is not unique`	Run the final wizard setup step again. After doing this, the error does not occur and you can proceed to the GUI login.

Appendix C. Troubleshooting for ESS on PPC64LE

Note: Most issues on ESS (PPC64LE) are not applicable if the Fusion mode is used.

Here are some tips on how to avoid common issues on ESS (PPC64LE).

- Always use /24 for the FSP network. It is advised to use `10.0.0.0/24`.
- If possible use /24 for the xCAT network. It is advised to use `192.168.X.X/24`.
- Do not overlap subnets on the EMS node. For instance, do not use `192.168.X.X` on both networks.
- Do not use non-traditional subnets such as /26.
- Always verify that all nodes are visible on the FSP network by using **gssdeploy –f**.
- If you get a timeout (8 min/16min) during Genesis discovery look into the following:
  - Is the DHCP server started and without issue (**systemctl status dhcpd**)?
  - Is your subnetting correct?
  - Is your /etc/hosts file correct?
  - Look into using a genesis IP range.

    ```
    Genesis IP range example:
    Add the following to gssdeploy.cfg
    EMS_GENESIS_IP_RANGE=192.168.202.13-192.168.202.14
    ```

    In this case, 202.13 and 202.14 are the nodes that are being tried for deployment. There cannot be any nodes up with IPs in the given range. After setting the range, use **gssdeploy –x** or **gssdeploy –o** again. If all else fails, you can power on the node and boot into petitboot to obtain the deployment MAC address. Once obtained, you can add into the node xCAT definition and complete the steps manually to start deployment.

    Each node ships with an extra cable in HMC port2 intended to be used to troubleshoot issues and access the FSP. It is advised you plug each cable into the FSP VLAN post deployment and set a static IP on the same subnet. Once this is done you can access ASMI remotely from the EMS. Alternatively, you can use this cable to hook a laptop to in the lab to access each node via the default manufacturing static IP.

Another workaround to the Genesis timeout issue is to manually retrieve and insert the MAC addresses into the xCAT node definitions. Use the following steps if the Genesis discovery fails (**gssdeploy -x** times out):

1. Exit **gssdeploy**.
2. Use **rpower** to power off the node(s).

   ```
   rpower NodeName off
   ```
3. Power on the node.

   ```
   rsetboot NodeName hd ; rpower NodeName on
   ```
4. Bring up a console immediately.

   ```
   rcons NodeName -f
   ```
5. When Petitboot comes up on the node, select **Exit to shell**.
6. Use the Linux **ifconfig** command to determine the interface that is holding the IP address.
7. Copy the MAC address and return to the EMS node.
8. Insert the MAC address.

   ```
   chdef NodeName mac=MacAddress
   nodeset NodeName osimage=install-gss_osimage_you_are_deploying
   makedhcp -n ; makedns -n
   ```

At this point you can skip **gssdeploy -x** and move on to the next step in the Quick Deployment Guide.

- Ensure that the storage enclosures are powered off or SAS cables are disconnected before running the **gssdeploy -x** command. If you are unable to power off the storage enclosures or remove the SAS connections before running **gssdeploy -x**, genesis discovery might fail. In that case, exit **gssdeploy** and log in to the I/O server nodes by using the temporary dynamic IP address.
- In most cases, the node IP is different from the one in the /etc/hosts file. You can find it from the **dhcp status** or the **systemctl status dhcpd** commands, or from the journal or from /var/log/messages. It is also displayed in the **rcons** output. Log in and remove the mpt3sas driver (**modprobe -r mpt3sas**) and the nodes finish discovering. Confirm with the command **nodediscoverls** from the EMS node.

Appendix D. Networking requirements for ESS

This topic describes the networking requirements for installing ESS.

Note: The references to HMC are not applicable for the PPC64LE platform.

Networking requirements

The following networks are required:

* **Service network**

  This network connects the flexible service processor (FSP) on the management server and I/O server nodes (with or without the HMC, depending on the platform) as shown in blue in Figure 1 and 2 on the following pages.

* **Management and provisioning network**

  This network connects the management server to the I/O server nodes (and HMCs, if available) as shown in yellow in in Figure 1 and 2 on the following pages. The management server runs DHCP on the management and provisioning network. If a management server is not included in the solution order, a customer-supplied management server is used.

* **Clustering network**

  This high-speed network is used for clustering and client node access. It can be a 10 Gigabit Ethernet (GbE), 40 GbE, or InfiniBand network. It might not be included in the solution order.

* **External and campus management network**

  This public network is used for external and campus management of the management server, the HMC (if available), or both.

Figure 1, Network Topology, is a high-level logical view of the management and provisioning network and the service network for an ESS building block (on **PPC64BE**).
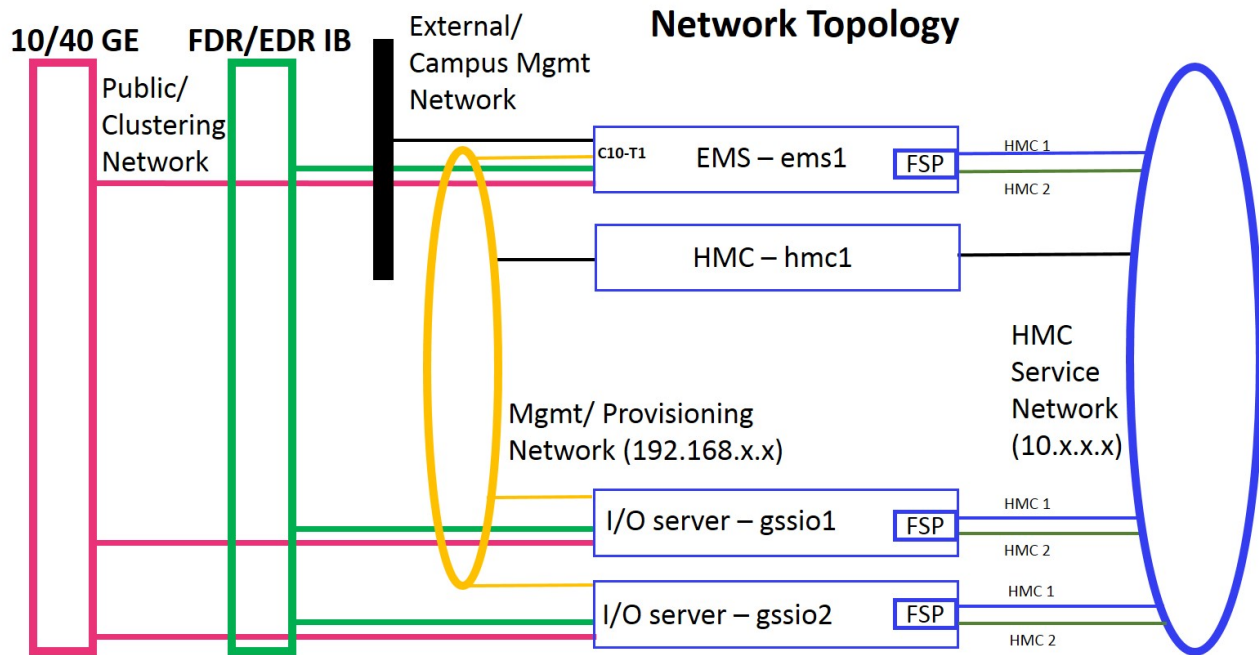
*Figure 1. The management and provisioning network and the service network: a logical view (on **PPC64BE***)*

Figure 2, Network Topology, is a high-level logical view of the management and provisioning network and the service network for an ESS building block (on **PPC64LE**).
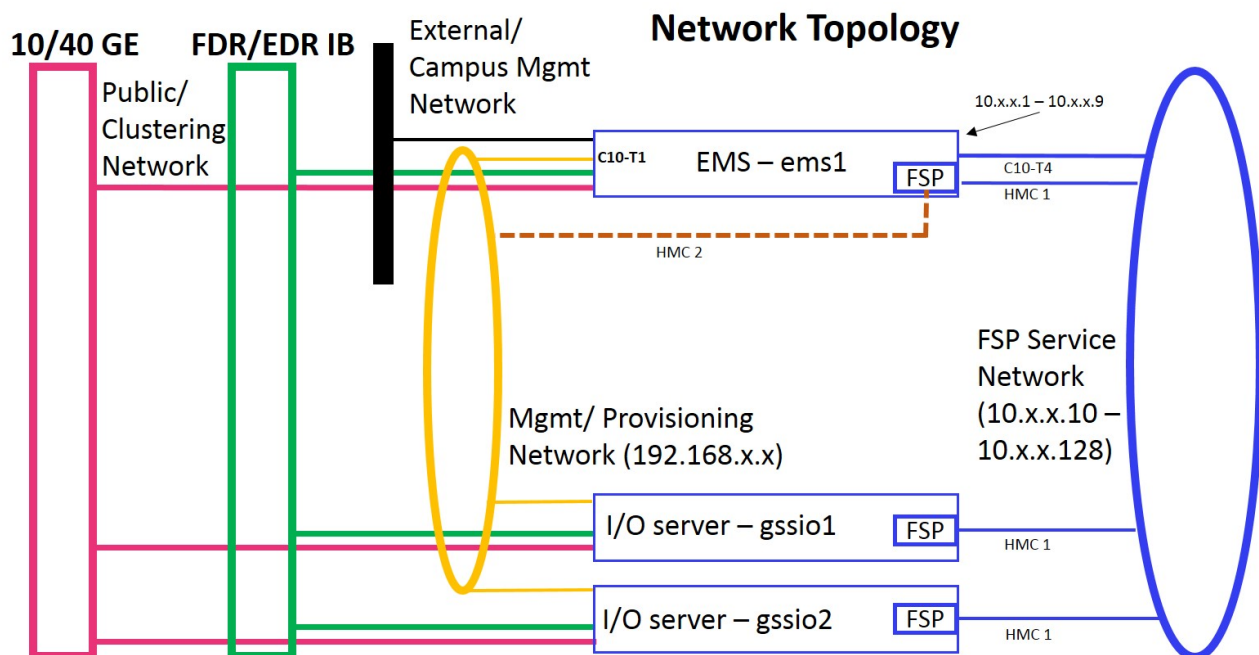

*Figure 2. The management and provisioning network and the service network: a logical view (on **PPC64LE***)*

The management and provisioning network and the service network must run as two non-overlapping networks implemented as two separate physical networks or two separate virtual local-area networks (VLANs).

Tip: HMC 2 is an optional third cable on the management server node that can be connected either to the management network or any other external network provided by the customer. This connection can be added in case the ability to service or control the management server node remotely is required.

The HMC, the management server, and the switches (1 GbE switches and high-speed switches) might not be included in a solution order in which an existing or customer-supplied HMC or management server is used. Perform any advance planning tasks that might be needed to access and use these solution components.

Appendix E. Pre-installation tasks for ESS

This topic provides the pre-installation tasks required for ESS.

Note: The references to HMC are not applicable for the PPC64LE platform.

Table 3. Pre-installation tasks

ESS component	Description	Required actions	System settings
1. Service network **Note:** This network varies depending on the platform (PPC64BE or PPC64LE).	**HMC service network:** This private network connects the HMC with the management server's FSP and the I/O server nodes. The service network must not be seen by the OS running on the node being managed (that is, the management server or the I/O server node). The HMC uses this network to discover the management server and the I/O server nodes and perform such hardware management tasks as creating and managing logical partitions, allocating resources, controlling power, and rebooting. **Note:** HMC is not applicable for the PPC64LE platform. **FSP service network:** This private network connects the FSP interface on EMS and the I/O server nodes. The service network must be seen by the OS running on the EMS node but not by the I/O server nodes being managed.	Perform any advance planning tasks that might be needed to access and use the HMC if it is not part of the solution order and a customer-supplied HMC will be used. Set up this network if it has not been set up already.	Set the HMC to be the DHCP server for the service network.
2. Management and provisioning network	This network connects the management server node with the HMC (when present) and the I/O server nodes. It typically runs over 1Gb. • This network is visible to the OS that is running on the nodes. • The management server uses this network to communicate with the HMC (when present) and to discover the I/O server nodes. • The management server will be the DHCP server on this network. There cannot be any other DHCP server on this network. • This network is also used to provision the node and therefore deploy and install the OS on the I/O server nodes.	Perform any advance planning tasks that might be needed to access and use the management server if it is not part of the solution order and a customer-supplied management server will be used. Set up this network if it has not been set up already.	
3. Clustering network	This network is for high-performance data access. In most cases, this network is also part of the clustering network. It is typically composed of 10GbE, 40GbE, or InfiniBand networking components.	Set up this network if it has not been set up already.	

Table 3. Pre-installation tasks (continued)

ESS component	Description	Required actions	System settings
4. Management network domain	The management server uses this domain for the proper resolution of hostnames.	Set the domain name using *lowercase* characters. Do *not* use any uppercase characters.	Example: `gpfs.net`
5. HMC node (IP address and hostname) **Note:** HMC is not applicable for the PPC64LE platform.	The IP address of the HMC node on the management network has a console name, which is the hostname and a domain name. • This IP address must be configured and the link to the network interface must be up. • The management server must be able to reach the HMC using this address.	Set the fully-qualified domain name (FQDN) and the hostname using *lowercase* characters. Do *not* use any uppercase characters. Do *not* use a suffix of -en*x*, where *x* is any character. Do *not* use an _ (underscore) in the hostname.	Example: IP address: 192.168.45.9 Hostname: `hmc1` FQDN: `hmc1.gpfs.net`
6. Management server node (IP address)	The IP address of the management server node has an FQDN and a hostname. • This IP address must be configured and the link to the network interface must be up. • The management network must be reachable from this IP address.	Set the FQDN and hostname using *lowercase* characters. Do *not* use any uppercase characters. Do *not* use a suffix of -en*x*, where *x* is any character. Do *not* use an _ (underscore) in the hostname.	Example: IP address: 192.168.45.10 Hostname: `ems1` FQDN: `ems1.gpfs.net`

7. I/O server nodes (IP addresses)	The IP addresses of the I/O server nodes have FQDNs and hostnames.
	• These addresses are assigned to the I/O server nodes during node deployment.
	• The I/O server nodes must be able to reach the management network using this address.
		Set the FQDN and hostname using *lowercase* characters. These names must match the name of the partition created for these nodes using the HMC. Do *not* use any uppercase characters. Do *not* use a suffix of -en*x*, where *x* is any character. Do *not* use an _ (underscore) in the host name.	Example:
			
			I/O server 1:
			IP address:
			192.168.45.11
			Hostname: `gssio1`
			FQDN:
			`gssio1.gpfs.net`
			
			I/O server 2:
			IP address:
			192.168.45.12
			Hostname: `gssio2`
			FQDN:
			`gssio2.gpfs.net`

8. Management server node management network interface (PPC64BE) Management server node FSP network interface (PPC64LE)	The management network interface of the management server node must have the IP address that you set in item 6 assigned to it. This interface must have only one IP address assigned. For the PPC64LE system, one additional interface is assigned to FSP network. This interface must have only one IP address assigned.	To obtain this address, run: `ip addr`	Example: `enP7p128s0f0`

Table 3. Pre-installation tasks (continued)

ESS component	Description	Required actions	System settings
9. HMC (**hscroot** password) **Note:** HMC is not applicable for the PPC64LE platform.		Set the password for the **hscroot** user ID.	Example: `abc123` This is the default password.

10. Kernel	Updating the kernel is required for all ESS nodes and it is verified by using **`gssinstallcheck`**.		Example:
			`RHSA-2018-0399-73-`
			`BE-KERNEL.tar.gz`

11. systemd	Updating the systemd service is required for all ESS nodes and it is verified by using **`gssinstallcheck`**.		Example:
			`systemd-RHBA-2018-`
			`0416-73-BE.tar.gz`

12. Network Manager	Updating the Network Manager service is required for all ESS nodes and it is verified by using **`gssinstallcheck`**.		Example:
			`netmgr-RHBA-2017-`
			`2925-73-BE.tar.gz`

13. Customer Red Hat Network (RHN) license keys	If possible, retrieve the RHN license keys for the customer in advance. This allows you to download the kernel, ISO, systemd, and Network Manager ahead of time. This also allows you to register and connect the newly deployed ESS system to RHN to apply security updates prior to leaving the site.		The keys must be available from the customer order. Contact offering management if help is required. **Note:** The customer must have an EU license.
14. I/O servers (user IDs and passwords)	The user IDs and passwords of the I/O servers are assigned during deployment.		Example: User ID: `root` Password: `cluster` (this is the default password)
15. FSP IPMI password	The IPMI password of the FSP. FSP IPMI of all the nodes assumed to be identical.		Example: `PASSWORD`
16. Clustering network (hostname prefix or suffix)	This high-speed network is implemented on a 10Gb Ethernet, 40Gb Ethernet or InfiniBand network.	Set a hostname for this network. It is customary to use hostnames for the high-speed network that use the prefix and suffix of the actual hostname. Do *not* use a suffix of -en*x*, where *x* is any character.	Examples: Suffixes: `-bond0`, `-ib`, `-10G`, `-40G` Hostnames with a suffix: `gssio1-ib`, `gssio2-ib`

Table 3. Pre-installation tasks (continued)

ESS component	Description	Required actions	System settings

17. High-speed cluster network (IP address)	The IP addresses of the management server nodes and I/O server nodes on the high-speed cluster network have FQDNs and hostnames.
	
	In the example, `172.10.0.11` is the IP address that the GPFS daemon uses for clustering. The corresponding FQDN and hostname are `gssio1-ib` and `gssio1-ib.data.net`, respectively.
		Set the FQDNs and hostnames.
		
		Do *not* make changes in the `/etc/hosts` file for the high-speed network until the deployment is complete. Do *not* create or enable the high-speed network interface until the deployment is complete.
			Example:
			
			Management server:
			    IP address:
			    `172.10.0.10`
			    Hostname: `ems1-ib`
			    FQDN:
			    `ems1-ib.gpfs.net`
			
			I/O server 1:
			    IP address:
			    `172.10.0.11`
			    Hostname:
			    `gssio1-ib`
			    FQDN:
			    `gssio1-ib.data.net`
			
			I/O server 2:
			    IP address:
			    `172.10.0.12`
			    Hostname:
			    `gssio2-ib`
			    FQDN:
			    `gssio2-ib.data.net`

18. Red Hat Enterprise Linux 7.3	The Red Hat Enterprise Linux 7.3 DVD or ISO file is used to create a temporary repository for the xCAT installation. xCAT uses it to create a Red Hat Enterprise Linux repository on the management server node.	Obtain this DVD or ISO file and download. For more information, see the Red Hat Enterprise Linux website: **http:// access.redhat.com/ products/red-hat-enterprise-linux/**	Example: `RHEL-7.3-20170219.1-Server-ppc64-dvd1.iso` **Note:** The Red Hat Enterprise Linux 7.3 ISO name depends on the architecture (PPC64BE or PPC64LE).
19. Management network switch	The switch that implements the management network must allow the Bootstrap Protocol (BOOTP) to go through.	Obtain the IP address and access credentials (user ID and password) of this switch. Some switches generate many Spanning Tree Protocol (STP) messages, which interfere with the network boot process. You need to disable STP to mitigate this.	

Table 3. Pre-installation tasks (continued)

ESS component	Description	Required actions	System settings

20. Target file system	You need to provide information about the target file system that is created using storage in the ESS building blocks. This information includes name, block size, file system size, RAID code, etc. This information you is passed on to **gssgenvdisks** to create the customer file system.	Set the target file system name, the mount point, the block size, the number of data NSDs, and the number of metadata NSDs.	Example:
			
			`Block size = 8M,`
			`#datansd=4,`
			`#metadatansd=2`

Appendix F. Installation: reference

This topic provides information on adding IBM Spectrum Scale nodes to an ESS cluster and node name considerations.

Adding IBM Spectrum Scale nodes to an ESS cluster

IBM Spectrum Scale node configuration is optimized for running IBM Spectrum Scale RAID functions.

1. ESS cluster node configuration is optimized for running IBM Spectrum Scale RAID functions. Protocols, other gateways, or any other non-ESS services must not be run on ESS management server nodes or I/O server nodes. In a cluster with high IO load, avoid using ESS nodes as cluster manager or filesystem manager. For optimal performance the NSD client nodes accessing ESS nodes should be properly configured. ESS ships with **gssClientConfig.sh script** located in `/usr/lpp/mmfs/samples/gss/` directory. This script can be used to configure the client as follows:

   ```
   /usr/lpp/mmfs/samples/gss/gssClientConfig.sh <Comma Separated list of
   client nodes or nodeclass>
   ```

   You can run the following to see configuration parameter settings without setting them:

   ```
   /usr/lpp/mmfs/samples/gss/gssClientConfig.sh -D
   ```

   After running this script, restart GPFS on the affected nodes for the optimized configuration settings to take effect.

   Important: Do not run `gssClientConfig.sh` unless you fully understand the impact of each setting on the customer environment. Make use of the `-D` option to decide if all or some of the settings might be applied. Then, individually update each client node settings as required.

2. When IBM Spectrum Scale nodes deployed with protocols are added to the ESS cluster, quorum, cluster manager, and filesystem manager functions should be moved from the ESS to the protocol nodes after adding protocol nodes to the cluster.

For information about adding an IBM Spectrum Scale protocol node to an ESS cluster, see:

- Overview of the IBM Spectrum Scale installation toolkit
- Preparing a cluster that contains ESS for adding protocols
- Spectrum Scale Protocols Quick Overview

Node name considerations

Carefully select the hostname, suffix, and prefix of the management server and I/O server so that the hostname used in the high-speed network and by the ESS cluster can be generated from the suffix or prefix.

High-speed hostnames

Example 1:

```
a-bcd-edf-1
a-bcd-edf-2
a-bcd-edf-3
a-bcd-edf-4
```

Here, `a-bcd-` is the prefix and `edf-1`, `edf-2`, `edf-3`, and `edf-4` are the xCAT names of the nodes.

Example 2:

```
1-a-bcd-edf
2-b-bcd-edf
3-c-bcd-edf
4-d_bcd_edf
```

Here, `-edf` is the suffix and `1-a-bcd`, `2-a-bcd`, `3-a-bcd`, and `4-a-bcd` are the xCAT names of the nodes.

If possible, avoid using high-speed node names with variations at the beginning and the end, such as:

```
A-a-bcd-edf-1
B-b-bdc-edf-2
C-c-bcd-edf-3
D-d-bcd-edf-4
```

In such cases, use the **-N** option and specify the node list with the **gssgencluster** and **gssgenclusterrgs** commands. The node names must be reachable from the management server node. xCAT requires that the target nodes be part of a node group and a warning might be issued if the hostname is not defined as an xCAT object.

Example:

1. The xCAT hostnames are **gssio1**, **gssio2**, **gssio3**, and **gssio4**.
2. The high-speed hostnames are **A-test1**, **B-test2**, **C-test3**, **D-test4**. These hostnames are reachable from the management server node. They are not defined in xCAT.

Run:

```
gssgencluster -C test01 -N A-test1,B-test2,C-test3,D-test4
```

Appendix G. Updating the system firmware

Use this information to obtain and apply the system firmware updates.

The system firmware packages are available in one of the following directories depending on the architecture of the management server node in newly shipped systems:

- **PPC64BE:** `/opt/ibm/gss/install/rhel7/ppc64/firmware`
- **PPC64LE:** `/opt/ibm/gss/install/rhel7/ppc64le/firmware`
- System firmware update files for PPC64BE for updating using HMC:

  ```
  01SV860_138_056.rpm
  01SV860_138_056.xml
  ```
- System firmware update file for PPC64LE for updating using the command line:

  ```
  01SV860_138_056.img
  ```

You can obtain the firmware update packages as follows:

1. Go to the FixCentral website.
2. Download the POWER8 System Firmware FW860.42 (SV860_138) package.

Depending on your platform, use one of the following sets of steps for updating system firmware.

- Update the system firmware on PPC64LE systems as follows.
    1. Unpack the `*img` file in the `/tmp/fwupdate` directory.

       ```
       cd /opt/ibm/gss/install/firmware/
       rpm -ivh 01SV860_138_056.rpm
       ```
    2. Shutdown IBM Spectrum Scale and stop any ongoing I/O on the node.
    3. Verify the firmware level.

       ```
       update_flash -v -f /tmp/fwupdate/01SV860_138_056.img
       ```
    4. Update the system firmware.

       ```
       update_flash -f /tmp/fwupdate/01SV860_138_056.img
       ```

       After issuing this command, the node reboots and updates the firmware. It could take up to 30 minutes for the node to reboot with the new firmware level. You can then run **gssinstallcheck** on the node to verify if the firmware is successfully updated.

To update system firmware on PPC64BE systems, you must use HMC and you must upgrade HMC to 860 SP2 before updating system firmware. For information about upgrading HMC, see HMC V8 Upgrade Procedure.

- Update the system firmware on PPC64BE systems as follows.
    1. From the HMC navigation area, click **Resources** > **All Systems** > **Server** > **Updates**.
    2. From the **Updates** menu, click **Change Licensed Internal Code** > **for the Current Release...**.
    3. Using SFTP, point to the `/opt/ibm/gss/install/firmware` directory on the EMS node. The following files should be present:

       ```
       01SV860_138_056.rpm
       01SV860_138_056.xml
       ```

       Note: For updating the system firmware using HMC, if SFTP to the EMS node does not work, move the `*rpm` and the `*xml` files to a server which is accessible using FTP or SFTP.
    4. Select the update file and update the system firmware. It could take up to 30 minutes to update the firmware using HMC.

Appendix H. Upgrading the Hardware Management Console (HMC)

For PPC64BE deployments, ensure that HMC is properly configured for the management server node and I/O server nodes and partition names are correctly set.

- To apply the HMC V8 update, use the following resources:
  - HMC V8 upgrade procedure: https://www-01.ibm.com/support/docview.wss?uid=nas8N1020108
  - HMC V8 860 files: ftp://public.dhe.ibm.com/software/server/hmc/network/v8860/
  - HMC V8 860 SP2 ISO: ftp://public.dhe.ibm.com/software/server/hmc/updates/ HMC_Update_V8R860_SP2.iso

  After upgrading, the HMC configuration should be similar to:
  Release: 8.6.0 Base Version: V8R8.6.0 Service Pack: 2

  Note: This is not applicable for the PPC64LE platform.

Appendix I. Obtaining kernel for system upgrades

For new system installation, the kernel is shipped with the system. However, for upgrades, you need to obtain and package the kernel update, and then follow the kernel update installation procedure.

You must have a EUS license to download the kernel from Red Hat Network.

Use the following steps during an upgrade to obtain and package the kernel update.

1. Clear the version locks.

   ```
   yum versionlock clear
   ```

2. Connect the management server node to the Red Hat Network.

   ```
   subscription-manager register --username=<X> --password=<Y>
   subscription-manager list --available // list pools
   subscription-manager attach --pool=<X>
   ```

   Or

   ```
   subscription-manager attach --auto
   ```

3. Create a directory for the kernel update package.

   For PPC64BE, issue:

   ```
   mkdir -p /tmp/kernel/RHSA-2018-0399-73-BE/
   ```

   For PPC64LE, issue:

   ```
   mkdir -p /tmp/kernel/RHSA-2018-0399-73-LE/
   ```

4. List all repositories and enable the repositories that are disabled, as required.

   ```
   yum repolist all
   yum-config-manager --enable rhel*
   ```

   Or

   ```
   subscription-manager config --rhsm.manage_repos=1
   ```

5. Download the kernel update package.

   For PPC64BE, issue:

   ```
   yum update *514.44.1* --downloadonly --downloaddir=/tmp/kernel/RHSA-2018-0399-73-BE
   yum update perf-3.10.0-514.44.1.el7.ppc64.rpm --downloadonly --downloaddir=/tmp/kernel/RHSA-2018-0399-73-BE
   yum update python-perf-3.10.0-514.44.1.el7.ppc64.rpm --downloadonly \
   --downloaddir=/tmp/kernel/RHSA-2018-0399-73-BE
   ```

   For PPC64LE, issue:

   ```
   yum update *514.44.1* --downloadonly --downloaddir=/tmp/kernel/RHSA-2018-0399-73-LE
   yum update perf-3.10.0-514.44.1.el7.ppc64le.rpm --downloadonly --downloaddir=/tmp/kernel/RHSA-2018-0399-73-LE
   yum update python-perf-3.10.0-514.44.1.el7.ppc64le.rpm --downloadonly \
   --downloaddir=/tmp/kernel/RHSA-2018-0399-73-LE
   ```

   The command-line kernel download method might fail if a newer kernel is available. In that case, use these steps.

   a. Use one of the following steps depending on your platform:

      - For PPC64BE, go to the following URL: https://access.redhat.com/search/#/ ?q=kernel*514*35*.1*ppc64.rpm&p=1&srch=any&documentKind=
      - For PPC64LE, go to the following URL: https://access.redhat.com/search/#/ ?q=kernel*514*44*.1*ppc64le.rpm&p=1&srch=any&documentKind=

   b. Search for the required or any additional RPMs listed in "About the ESS Red Hat Linux Errata Kernel Update" on page 72 and download them.

6. Package the directory.

For PPC64BE, issue:

```
cd /tmp/kernel ; tar -zcvf kernel-RHSA-2018-0399-73-BE.tar.gz RHSA-2018-0399-73-BE
```

For PPC64LE, issue:

```
cd /tmp/kernel ; tar -zcvf kernel-RHSA-2018-0399-73-LE.tar.gz RHSA-2018-0399-73-LE
```

Note: Make sure that the RPM files are in the RHSA-2018-0399-73-BE or the RHSA-2018-0399-73-LE folder. Do not create any nested folder inside the RHSA-2018-0399-73-BE or the RHSA-2018-0399-73-LE folder and try to place the RPM file in that nested folder. Doing so will result in failure of the kernel patch installation during the cluster deployment.

7. Disable the Red Hat Network connection on the management server node.

```
subscription-manager config --rhsm.manage_repos=0
yum clean all
```

Continue with the kernel update installation steps for `kernel-RHSA-2018-0399-73-BE.tar.gz` or `kernel-RHSA-2018-0399-73-LE.tar.gz` using **gssdeploy -k**. For example, use one of the following commands depending on the architecture to place the kernel updates in the kernel repository:

For PPC64BE, issue:

```
/var/tmp/gssdeploy -k kernel-RHSA-2018-0399-73-BE.tar.gz --silent
```

This command places the kernel update in `/install/gss/otherpkgs/rhels7/ppc64/kernel`

For PPC64LE, issue:

```
/var/tmp/gssdeploy -k kernel-RHSA-2018-0399-73-LE.tar.gz --silent
```

This command places the kernel update in `/install/gss/otherpkgs/rhels7/ppc64le/kernel`

For more information about the kernel update, see "About the ESS Red Hat Linux Errata Kernel Update."

About the ESS Red Hat Linux Errata Kernel Update

This topic provides information about the Red Hat Linux Errata Kernel Update for ESS.

At the time of shipping from factory, most current recommended kernel errata and associated RPMs are provided in the `/home/deploy` directory. It is highly recommended to limit errata updates applied to the Red Hat Enterprise Linux operating system used in the ESS solution to security errata or errata updates requested by service. For more information visit Red Hat's solution article on applying only security updates: https://access.redhat.com/solutions/10021.

Kernel errata updates can be obtained from Red Hat network (RHN) using the supplied license: https://access.redhat.com/errata/#/.

For information about the kernal update for the current release, see https://access.redhat.com/errata/RHSA-2018:0399.

This example shows errata update (RHSA-2018-0399) provided in the `/home/deploy` directory of the EMS node when shipped from factory.

The following packages are provided in `kernel-RHSA-2018-0399-73-BE.tar.gz`:

```
kernel-3.10.0-514.44.1.el7.ppc64.rpm
```

```
kernel-abi-whitelists-3.10.0-514.44.1.el7.noarch.rpm
```

```
kernel-bootwrapper-3.10.0-514.44.1.el7.ppc64.rpm
```

```
kernel-devel-3.10.0-514.44.1.el7.ppc64.rpm
```

```
kernel-doc-3.10.0-514.44.1.el7.noarch.rpm
kernel-headers-3.10.0-514.44.1.el7.ppc64.rpm
kernel-tools-3.10.0-514.44.1.el7.ppc64.rpm
kernel-tools-libs-3.10.0-514.44.1.el7.ppc64.rpm
perf-3.10.0-514.44.1.el7.ppc64.rpm
python-perf-3.10.0-514.44.1.el7.ppc64.rpm
```

The following packages are provided in `kernel-RHSA-2018-0399-73-LE.tar.gz`:

```
kernel-3.10.0-514.44.1.el7.ppc64le.rpm
kernel-abi-whitelists-3.10.0-514.44.1.el7.noarch.rpm
kernel-bootwrapper-3.10.0-514.44.1.el7.ppc64le.rpm
kernel-debug-3.10.0-514.44.1.el7.ppc64le.rpm
kernel-devel-3.10.0-514.44.1.el7.ppc64le.rpm
kernel-doc-3.10.0-514.44.1.el7.noarch.rpm
kernel-headers-3.10.0-514.44.1.el7.ppc64le.rpm
kernel-tools-3.10.0-514.44.1.el7.ppc64le.rpm
kernel-tools-libs-3.10.0-514.44.1.el7.ppc64le.rpm
perf-3.10.0-514.44.1.el7.ppc64le.rpm
python-perf-3.10.0-514.44.1.el7.ppc64le.rpm
```

Appendix J. Obtaining systemd update for system upgrades

For new system installation, the systemd update is shipped with the system and it is available in the /home/deploy directory.. However, for upgrades, you need to obtain and package the systemd update, and then install the systemd update.

You must have a EUS license to download the systemd update from Red Hat Network.

Use the following steps during an upgrade to obtain and package the systemd update.

1. Clear the version locks.

   ```
   yum versionlock clear
   ```

2. Connect the management server node to the Red Hat Network.

   ```
   subscription-manager register --username=<X> --password=<Y>
   subscription-manager list --available // list pools
   subscription-manager attach --pool=<X>
   ```

   Or

   ```
   subscription-manager attach --auto
   ```

3. Create a directory for the systemd update package.

   For PPC64BE, issue:

   ```
   mkdir -p /tmp/systemd/RHBA-2018-0416-BE/
   ```

   For PPC64LE, issue:

   ```
   mkdir -p /tmp/systemd/RHBA-2018-0416-LE/
   ```

4. List all repositories and enable the repositories that are disabled, as required.

   ```
   yum repolist all
   yum-config-manager --enable rhel*
   ```

   Or

   ```
   subscription-manager config --rhsm.manage_repos=1
   ```

5. Download the systemd update package.

   For PPC64BE, issue:

   ```
   yum update systemd*219-42.el7_4.10* --downloadonly —downloaddir=/tmp/systemd/RHBA-2018-0416-BE
   yum update libgudev1-219-42.el7_4.10.ppc64.rpm --downloadonly —downloaddir=/tmp/systemd/RHBA-2018-0416-BE
   yum update libgudev1-devel-219-42.el7_4.10.ppc64.rpm --downloadonly —downloaddir=/tmp/systemd/RHBA-2018-0416-BE
   yum update dracut*033-502*.ppc64.rpm --downloadonly —downloaddir=/tmp/systemd/RHBA-2018-0416-BE
   ```

   For PPC64LE, issue:

   ```
   yum update systemd*219-42.el7_4.10* --downloadonly —downloaddir=/tmp/systemd/RHBA-2018-0416-LE
   yum update libgudev1-219-42.el7_4.10.ppc64le.rpm --downloadonly —downloaddir=/tmp/systemd/RHBA-2018-0416-LE
   yum update libgudev1-devel-219-42.el7_4.10.ppc64le.rpm --downloadonly —downloaddir=/tmp/systemd/RHBA-2018-0416-LE
   yum update dracut*033-502*.ppc64le.rpm --downloadonly —downloaddir=/tmp/systemd/RHBA-2018-0416-LE
   ```

   The command-line kernel download method might fail if a newer kernel is available. In that case, use these steps.

   a. Use one of the following steps depending on your platform:
      - For PPC64BE, go to the following URL: https://access.redhat.com/search/#/ %3Fq=systemd*219-42.el7_4.10*ppc64.rpm%26p=1%26srch=any%26documentKind=
      - For PPC64LE, go to the following URL: https://access.redhat.com/search/#/ %3Fq=systemd*219-42.el7_4.10*ppc64le.rpm%26p=1%26srch=any%26documentKind=

   b. Search for the required or any additional RPMs listed in "About the ESS Red Hat Linux systemd update" on page 76 and download them.

6. Package the directory.

   For PPC64BE, issue:

   ```
   cd /tmp/systemd ; tar -zcvf systemd-RHBA-2018-0416-BE.tar.gz RHBA-2018-0416-BE
   ```

   For PPC64LE, issue:

   ```
   cd /tmp/systemd ; tar -zcvf systemd-RHBA-2018-0416-LE.tar.gz RHBA-2018-0416-LE
   ```

   Note: Make sure that the RPM files are in the RHBA-2018-0416-BE or the RHBA-2018-0416-LE folder. Do not create any nested folder inside the RHBA-2018-0416-BE or the RHBA-2018-0416-LE folder and try to place the RPM file in that nested folder. Doing so will result in failure of the systemd patch installation during the cluster deployment.

7. Disable the Red Hat Network connection on the management server node.

   ```
   subscription-manager config --rhsm.manage_repos=0
   yum clean all
   ```

Continue with the systemd update installation steps for `systemd-RHBA-2018-0416-BE.tar.gz` or `systemd-RHBA-2018-0416-LE.tar.gz` using **gssdeploy -p**. For example, use one of the following commands depending on the architecture to place the systemd update in the patch repository:

For PPC64BE, issue:

```
/var/tmp/gssdeploy -p RHBA-2018-0416-BE.tar.gz —-silent
```

This command places the systemd updates in `/install/gss/otherpkgs/rhels7/ppc64/patch`

For PPC64LE, issue:

```
/var/tmp/gssdeploy -p systemd-RHBA-2018-0416-LE.tar.gz —-silent
```

This command places the systemd updates in `/install/gss/otherpkgs/rhels7/ppc64le/patch`

For more information, see "About the ESS Red Hat Linux systemd update."

About the ESS Red Hat Linux systemd update

This topic provides information about the Red Hat Linux systemd update for ESS.

It is highly recommended to limit errata updates applied to the Red Hat Enterprise Linux operating system used in the ESS solution to security errata or errata updates requested by service. For more information visit Red Hat's solution article on applying only security updates: https://access.redhat.com/solutions/10021.

This example shows systemd update (RHBA-2018-0416) provided in the `/home/deploy` directory of the EMS node when shipped from factory.

For information about the systemd update for the current release, see https://access.redhat.com/errata/RHBA-2018:0416.

The following packages are provided in `systemd-RHBA-2018-0416-BE.tar.gz`:

   `systemd-219-42.el7_4.10.ppc64.rpm`

   `systemd-devel-219-42.el7_4.10.ppc64.rpm`

   `systemd-journal-gateway-219-42.el7_4.10.ppc64.rpm`

   `systemd-libs-219-42.el7_4.10.ppc64.rpm`

   `systemd-networkd-219-42.el7_4.10.ppc64.rpm`

```
systemd-python-219-42.el7_4.10.ppc64.rpm

systemd-resolved-219-42.el7_4.10.ppc64.rpm

systemd-sysv-219-42.el7_4.10.ppc64.rpm

libgudev1-219-42.el7_4.10.ppc64.rpm

libgudev1-devel-219-42.el7_4.10.ppc64.rpm

dracut-033-502.el7.ppc64.rpm

dracut-network-033-502.el7.ppc64.rpm

dracut-config-rescue-033-502.el7.ppc64.rpm
```

The following packages are provided in the `systemd-RHBA-2018-0416-LE.tar.gz`:

```
systemd-219-42.el7_4.10.ppc64le.rpm

systemd-devel-219-42.el7_4.10.ppc64le.rpm

systemd-journal-gateway-219-42.el7_4.10.ppc64le.rpm

systemd-libs-219-42.el7_4.10.ppc64le.rpm

systemd-networkd-219-42.el7_4.10.ppc64le.rpm

systemd-python-219-42.el7_4.10.ppc64le.rpm

systemd-resolved-219-42.el7_4.10.ppc64le.rpm

systemd-sysv-219-42.el7_4.10.ppc64le.rpm

libgudev1-219-42.el7_4.10.ppc64le.rpm

libgudev1-devel-219-42.el7_4.10.ppc64le.rpm

dracut-033-502.el7.ppc64le.rpm

dracut-network-033-502.el7.ppc64le.rpm

dracut-config-rescue-033-502.el7.ppc64le.rpm
```

Appendix K. Obtaining Network Manager updates for system upgrades

For new system installation, the Network Manager update is shipped with the system and it is available in the /home/deploy directory. However, for upgrades, you need to obtain and package the Network Manager update, and then install the Network Manager update.

You must have a EUS license to download the Network Manager update from Red Hat Network.

Use the following steps during an upgrade to obtain and package the Network Manager update.

1. Clear the version locks.

   ```
   yum versionlock clear
   ```

2. Connect the management server node to the Red Hat Network.

   ```
   subscription-manager register --username=<X> --password=<Y>
   subscription-manager list --available // list pools
   subscription-manager attach --pool=<X>
   ```

   Or

   ```
   subscription-manager attach --auto
   ```

3. Create a directory for the Network Manager update package.

   For PPC64BE, issue:

   ```
   mkdir -p /tmp/netmgr/RHBA-2017-2925-BE
   ```

   For PPC64LE, issue:

   ```
   mkdir -p /tmp/netmgr/RHBA-2017-2925-LE
   ```

4. List all repositories and enable the repositories that are disabled, as required.

   ```
   yum repolist all
   yum-config-manager --enable rhel*
   ```

   Or

   ```
   subscription-manager config --rhsm.manage_repos=1
   ```

5. Download the Network Manager update package.

   For PPC64BE, issue:

   ```
   yum update NetworkManager*1.8* --downloadonly —downloaddir=/tmp/netmgr/RHBA-2017-2925-BE
   yum update glib2*2.50* --downloadonly —downloaddir=/tmp/netmgr/RHBA-2017-2925-BE
   ```

   For PPC64LE, issue:

   ```
   yum update NetworkManager*1.8* --downloadonly —downloaddir=/tmp/netmgr/RHBA-2017-2925-LE
   yum update glib2*2.50* --downloadonly —downloaddir=/tmp/netmgr/RHBA-2017-2925-LE
   ```

   The command-line kernel download method might fail if a newer kernel is available. In that case, use these steps.

   a. Use one of the following steps depending on your platform:

      - For PPC64BE, go to the following URL: https://access.redhat.com/search/#/ %3Fq=networkmanager*1.8.0-11.el7_4*ppc64.rpm%26p=1%26srch=any%26documentKind=
      - For PPC64LE, go to the following URL: https://access.redhat.com/search/#/ %3Fq=networkmanager*1.8.0-11.el7_4*ppc64le.rpm%26p=1%26srch=any%26documentKind=

   b. Search for the required or any additional RPMs listed in "About the ESS Red Hat Linux Network Manager update" on page 80 and download them.

6. Package the directory.

   For PPC64BE, issue:

```
cd /tmp/systemd ; tar -zcvf netmgr-RHBA-2017-2925-BE.tar.gz RHBA-2017-2925-BE
```

For PPC64LE, issue:

```
cd /tmp/systemd ; tar -zcvf netmgr-RHBA-2017-2925-LE.tar.gz RHBA-2017-2925-LE
```

Note: Make sure that the RPM files are in the RHBA-2017-2925-BE or the RHBA-2017-2925-LE folder. Do not create any nested folder inside the RHBA-2017-2925-BE or the RHBA-2017-2925-LE folder and try to place the RPM file in that nested folder. Doing so will result in failure of the network manager patch installation during the cluster deployment.

7. Disable the Red Hat Network connection on the management server node.

```
subscription-manager config --rhsm.manage_repos=0
yum clean all
```

8. Place the Network Manager updates in the patch repository.

For PPC64BE, issue:

```
/var/tmp/gssdeploy -p netmgr-RHBA-2017-2925-BE.tar.gz —-silent
```

This command places the Network Manager updates in `/install/gss/otherpkgs/rhels7/ppc64/patch`

For PPC64LE, issue:

```
/var/tmp/gssdeploy -p netmgr-RHBA-2017-2925-LE.tar.gz —-silent
```

This command places the Network Manager updates in `/install/gss/otherpkgs/rhels7/ppc64le/patch` For more information, see "About the ESS Red Hat Linux Network Manager update."

About the ESS Red Hat Linux Network Manager update

This topic provides information about the Red Hat Linux Network Manager update for ESS.

It is highly recommended to limit errata updates applied to the Red Hat Enterprise Linux operating system used in the ESS solution to security errata or errata updates requested by service. For more information visit Red Hat's solution article on applying only security updates: https://access.redhat.com/solutions/10021.

This example shows Network Manager update (RHBA-2017-2925) provided in the `/home/deploy` directory of the EMS node when shipped from factory.

For information about the Network Manager update for the current release, see https://access.redhat.com/errata/RHBA-2017:2925.

The following packages are provided in `netmgr-RHBA-2017-2925-BE.tar.gz`:

    NetworkManager-1.8.0-11.el7_4.ppc64.rpm

    NetworkManager-adsl-1.8.0-11.el7_4.ppc64.rpm

    NetworkManager-bluetooth-1.8.0-11.el7_4.ppc64.rpm

    NetworkManager-glib-1.8.0-11.el7_4.ppc64.rpm

    NetworkManager-glib-devel-1.8.0-11.el7_4.ppc64.rpm

    NetworkManager-libnm-1.8.0-11.el7_4.ppc64.rpm

    NetworkManager-libnm-devel-1.8.0-11.el7_4.ppc64.rpm

    NetworkManager-ppp-1.8.0-11.el7_4.ppc64.rpm

    NetworkManager-team-1.8.0-11.el7_4.ppc64.rpm

    NetworkManager-tui-1.8.0-11.el7_4.ppc64.rpm

```
NetworkManager-wifi-1.8.0-11.el7_4.ppc64.rpm

NetworkManager-wwan-1.8.0-11.el7_4.ppc64.rpm

glib2-2.50.3-3.el7.ppc64.rpm

glib2-devel-2.50.3-3.el7.ppc64.rpm
```

The following packages are provided in the `netmgr-RHBA-2017-2925-LE.tar.gz`:

```
NetworkManager-1.8.0-11.el7_4.ppc64le.rpm

NetworkManager-adsl-1.8.0-11.el7_4.ppc64le.rpm

NetworkManager-bluetooth-1.8.0-11.el7_4.ppc64le.rpm

NetworkManager-glib-1.8.0-11.el7_4.ppc64le.rpm

NetworkManager-glib-devel-1.8.0-11.el7_4.ppc64le.rpm

NetworkManager-libnm-1.8.0-11.el7_4.ppc64le.rpm

NetworkManager-libnm-devel-1.8.0-11.el7_4.ppc64le.rpm

NetworkManager-ppp-1.8.0-11.el7_4.ppc64le.rpm

NetworkManager-team-1.8.0-11.el7_4.ppc64le.rpm

NetworkManager-tui-1.8.0-11.el7_4.ppc64le.rpm

NetworkManager-wifi-1.8.0-11.el7_4.ppc64le.rpm

NetworkManager-wwan-1.8.0-11.el7_4.ppc64le.rpm

glib2-2.50.3-3.el7.ppc64le.rpm

glib2-devel-2.50.3-3.el7.ppc64le.rpm
```

Appendix L. Running gssinstallcheck in parallel

The **gssinstallcheck** command checks various aspects of the installation on all nodes. This command runs on each node sequentially. It has been enhanced such that you can run the **gssinstallcheck** command on all nodes in parallel.

It is advisable to run **gssinstallcheck** in parallel if the number of nodes in the cluster is more than 40 nodes. This is because running this command sequentially on such a large number of nodes takes a significant amount of time.

Note: Parallel **gssinstallcheck** can only be invoked from the management server node. Invoking **gssinstallcheck** parallelly from I/O server nodes will not work.

You can run **gssinstallcheck** in parallel as follows.

```
# xdsh ems1,gss_ppc64 "/opt/ibm/gss/tools/bin/gssinstallcheck -N localhost" | xcoll -n
```

In this command, **gssinstallcheck** is being run from the management server node and all I/O server nodes are a part of the gss_ppc64 xCAT group. Following is a sample output of this command. The output of all the nodes is grouped together if the **gssinstallcheck** output is same across nodes. In the following example, the output texts from gssio1 and gssio2 nodes are identical thus they have been grouped together in a single output. The ems1 output has been separately printed as the output of **gssinstallcheck** on the ems1 node is different. For more information, see the **xdsh** and **xcoll** command documentation.

```
====================================
gssio1,gssio2
====================================
Start of install check
xCAT objects not found for the nodelist localhost
nodelist:       localhost
...
...


====================================
ems1
====================================
Start of install check
xCAT objects not found for the nodelist localhost
nodelist:       localhost

Getting package information.
...
...
```

Appendix M. Shutting down and powering up ESS

The ESS components and frame may need to be powered off in cases such as data center maintenance, relocation, or emergencies. Use the following information to shut down and power up ESS.

Shutting down ESS

1. Verify that the file systems are not needed by users during the time the system will be unavailable.
2. If you are using a remote cluster to mount the ESS file system, unmount the file system by issuing the **mmumount** command from the remote client nodes.
3. Shut down the nodes using the **mmshutdown -N** command. For example:

   mmshutdown -N ems1,gssio1,gssio2
4. If other nodes are attached and ESS nodes are the only quorum and manager nodes, it is recommended that you use the **mmshutdown -a** command to shut down the entire cluster.
5. Verify that IBM Spectrum Scale is shut down on the I/O nodes by issuing the **mmgetstate -a** command.
6. Power off the EMS and I/O nodes by issuing the **mmshutdown -h now** command on each individual node.

   If you are using the Big Endian (BE) platform:
   a. The EMC LPAR, I/O node1 LPAR, and I/O node 2 LPAR will be shut down after you issue the **shutdown -h now**.
   b. Use the HMC to shut down the physical servers.
   c. Verify that the power light on the front of the frame is blinking after the LPARs are shut down.

   If you are using the Big Endian (BE) platform and the HMC resides within this frame:
   a. Power off the HMC. If the HMC controls servers that are outside of this frame, plan appropriately before shutting down.

   If you are using the Little Endian (LE) platform:
   a. The EMC LPAR, I/O node1 LPAR, and I/O node 2 LPAR will be completely shut down after you issue the **shutdown -h now** command.
   b. Verify that the power light on the front of the frame is blinking.
7. Power off all storage by flipping the power switches to off.
8. Before shutting off power to the frame, verify there are no components within the frame that are relied on by external infrastructure such as IB or Ethernet switches. If any of these exist and hardware outside of the frame needs access, plan appropriately before shutting off power to the frame.

Powering up ESS

1. Verify that power is connected to the frame.
2. Turn on all PDUs within the ESS frame.
3. Power on the components in the following order.

   If you are using the Big Endian (BE) platform:
   a. Power on the HMC.
   b. Power on the storage drawers by flipping the power switches on each storage module to on.
   c. Power on the EMS node, I/O node 1 and I/0 node 2.
   d. Wait for the HMC to come online and log in.
   e. Wait for the EMS node, I/O node 1 and I/0 node 2 to be accessible to the HMC.
   f. Once the EMS sees that node, I/O node 1 and I/0 node 2 are powered on, move to the LPAR view for each and power on the associated LPARs:

EMS LPAR

1/O node 1 LPAR

I/O node 2 LPAR

g. Once all LPARs are powered on, ssh to the EMS node and verify that IBM Spectrum Scale has come online by issuing `mmgetstate -N ems1,gssio1,gssio2`. If IBM Spectrum Scale does not automatically start, start it manually by issuing `mmstartup -N ems1,gssio1,gssio2.`

h. Issue the `gnrhealthcheck` and the `mmhealth cluster show` commands, and check the GUI event logs.

If you are using the Little Endian (LE) platform:

a. Power on the storage drawers by flipping the power switches on each storage module to on.

b. Power on the EMS node, I/O node 1 and I/0 node 2.

c. Once all LPARs are powered on, ssh to the EMS node and verify that IBM Spectrum Scale has come online by issuing `mmgetstate -N ems1,gssio1,gssio2`. If IBM Spectrum Scale does not automatically start, start it manually by issuing `mmstartup -N ems1,gssio1,gssio2.`

d. Issue the `gnrhealthcheck` and the `mmhealth cluster show` commands, and check the GUI event logs.

Notices

This information was developed for products and services offered in the U.S.A.

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 Armonk, NY 10504-1785 U.S.A.

For license inquiries regarding double-byte (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

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:

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 states 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 Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply 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 Corporation
Dept. 30ZA/Building 707
Mail Station P300

87

2455 South Road,
Poughkeepsie, NY 12601-5400
U.S.A.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment or 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.

Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

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.

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 the names and addresses used by an actual business enterprise 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.

If you are viewing this information softcopy, the photographs and color illustrations may not appear.

Trademarks

IBM, the IBM logo, and 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.

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, Windows, and Windows NT 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.

Glossary

This glossary provides terms and definitions for the ESS solution.

The following cross-references are used in this glossary:

- *See* refers you from a non-preferred 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 (opens in new window):

http://www.ibm.com/software/globalization/terminology

B

building block
A pair of servers with shared disk enclosures attached.

BOOTP
See *Bootstrap Protocol (BOOTP)*.

Bootstrap Protocol (BOOTP)
A computer networking protocol thst is used in IP networks to automatically assign an IP address to network devices from a configuration server.

C

CEC See *central processor complex (CPC)*.

central electronic complex (CEC)
See *central processor complex (CPC)*.

central processor complex (CPC)
A physical collection of hardware that consists of channels, timers, main storage, and one or more central processors.

cluster
A loosely-coupled collection of independent systems, or *nodes*, organized into a network for the purpose of sharing resources and communicating with each other. See also *GPFS cluster*.

cluster manager
The node that monitors node status using disk leases, detects failures, drives recovery, and selects file system

managers. The cluster manager is the node with the lowest node number among the quorum nodes that are operating at a particular time.

compute node
A node with a mounted GPFS file system that is used specifically to run a customer job. ESS disks are not directly visible from and are not managed by this type of node.

CPC See *central processor complex (CPC)*.

D

DA See *declustered array (DA)*.

datagram
A basic transfer unit associated with a packet-switched network.

DCM See *drawer control module (DCM)*.

declustered array (DA)
A disjoint subset of the pdisks in a recovery group.

dependent fileset
A fileset that shares the inode space of an existing independent fileset.

DFM See *direct FSP management (DFM)*.

DHCP See *Dynamic Host Configuration Protocol (DHCP)*.

direct FSP management (DFM)
The ability of the xCAT software to communicate directly with the Power Systems server's service processor without the use of the HMC for management.

drawer control module (DCM)
Essentially, a SAS expander on a storage enclosure drawer.

Dynamic Host Configuration Protocol (DHCP)
A standardized network protocol that is used on IP networks to dynamically distribute such network configuration parameters as IP addresses for interfaces and services.

E

Elastic Storage Server (ESS)
A high-performance, GPFS NSD solution

made up of one or more building blocks that runs on IBM Power Systems servers. The ESS software runs on ESS nodes - management server nodes and I/O server nodes.

ESS Management Server (EMS)
An xCAT server is required to discover the I/O server nodes (working with the HMC), provision the operating system (OS) on the I/O server nodes, and deploy the ESS software on the management node and I/O server nodes. One management server is required for each ESS system composed of one or more building blocks.

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 (FEK)*, *master encryption key (MEK)*.

ESS See *Elastic Storage Server (ESS)*.

environmental service module (ESM)
Essentially, a SAS expander that attaches to the storage enclosure drives. In the case of multiple drawers in a storage enclosure, the ESM attaches to drawer control modules.

ESM See *environmental service module (ESM)*.

Extreme Cluster/Cloud Administration Toolkit (xCAT)
Scalable, open-source cluster management software. The management infrastructure of ESS is deployed by xCAT.

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 (FEK)*.

file encryption key (FEK)
A key used to encrypt sectors of an individual file. See also *encryption key*.

file system
The methods and data structures used to control how data is stored and retrieved.

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.

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.

flexible service processor (FSP)
Firmware that provices diagnosis, initialization, configuration, runtime error detection, and correction. Connects to the HMC.

FQDN
See *fully-qualified domain name (FQDN)*.

FSP See *flexible service processor (FSP)*.

fully-qualified domain name (FQDN)
The complete domain name for a specific computer, or host, on the Internet. The FQDN consists of two parts: the hostname and the domain name.

G

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 Storage Server (GSS)
A high-performance, GPFS NSD solution made up of one or more building blocks that runs on System x servers.

GSS See *GPFS Storage Server (GSS)*.

H

Hardware Management Console (HMC)
Standard interface for configuring and operating partitioned (LPAR) and SMP systems.

HMC See *Hardware Management Console (HMC)*.

I

IBM Security Key Lifecycle Manager (ISKLM)
For GPFS encryption, the ISKLM is used as an RKM server to store MEKs.

independent fileset
A fileset that has its own inode space.

indirect block
A block that contains 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.

Internet Protocol (IP)
The primary communication protocol for relaying datagrams across network boundaries. Its routing function enables internetworking and essentially establishes the Internet.

I/O server node
An ESS node that is attached to the ESS storage enclosures. It is the NSD server for the GPFS cluster.

IP See *Internet Protocol (IP)*.

IP over InfiniBand (IPoIB)
Provides an IP network emulation layer on top of InfiniBand RDMA networks, which allows existing applications to run over InfiniBand networks unmodified.

IPoIB See *IP over InfiniBand (IPoIB)*.

ISKLM
See *IBM Security Key Lifecycle Manager (ISKLM)*.

J

JBOD array
The total collection of disks and enclosures over which a recovery group pair is defined.

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.

L

LACP See *Link Aggregation Control Protocol (LACP)*.

Link Aggregation Control Protocol (LACP)
Provides a way to control the bundling of several physical ports together to form a single logical channel.

logical partition (LPAR)
A subset of a server's hardware resources virtualized as a separate computer, each with its own operating system. See also *node*.

LPAR See *logical partition (LPAR)*.

M

management network
A network that is primarily responsible for booting and installing the designated server and compute nodes from the management server.

management server (MS)
An ESS node that hosts the ESS GUI and xCAT and is not connected to storage. It can be part of a GPFS cluster. From a system management perspective, it is the

central coordinator of the cluster. It also serves as a client node in an ESS building block.

master encryption key (MEK)
A key that is used to encrypt other keys. See also *encryption key*.

maximum transmission unit (MTU)
The largest packet or frame, specified in octets (eight-bit bytes), that can be sent in a packet- or frame-based network, such as the Internet. The TCP uses the MTU to determine the maximum size of each packet in any transmission.

MEK See *master encryption key (MEK)*.

metadata
A data structure that contains access information about file data. Such structures include inodes, indirect blocks, and directories. These data structures are not accessible to user applications.

MS See *management server (MS)*.

MTU See *maximum transmission unit (MTU)*.

N

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 hexadecimal 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 can contain one or more nodes. In a Power Systems environment, synonymous with *logical partition*.

node descriptor
A definition that indicates how IBM Spectrum Scale uses a node. Possible functions include: manager node, client node, quorum node, and non-quorum node.

node number
A number that is generated and maintained by IBM Spectrum Scale 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 IBM Spectrum Scale 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.

O

OFED See *OpenFabrics Enterprise Distribution (OFED)*.

OpenFabrics Enterprise Distribution (OFED)
An open-source software stack includes software drivers, core kernel code, middleware, and user-level interfaces.

P

pdisk A physical disk.

PortFast
A Cisco network function that can be configured to resolve any problems that could be caused by the amount of time STP takes to transition ports to the Forwarding state.

R

RAID See *redundant array of independent disks (RAID)*.

RDMA
See *remote direct memory access (RDMA)*.

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.

recovery group (RG)
A collection of disks that is set up by IBM Spectrum Scale RAID, in which each disk is connected physically to two servers: a primary server and a backup server.

remote direct memory access (RDMA)
A direct memory access from the memory of one computer into that of another without involving either one's operating system. This permits high-throughput, low-latency networking, which is especially useful in massively-parallel computer clusters.

RGD See *recovery group data (RGD)*.

remote key management server (RKM server)
A server that is used to store master encryption keys.

RG See *recovery group (RG)*.

recovery group data (RGD)
Data that is associated with a recovery group.

RKM server
See *remote key management server (RKM server)*.

S

SAS See *Serial Attached SCSI (SAS)*.

secure shell (SSH)
A cryptographic (encrypted) network protocol for initiating text-based shell sessions securely on remote computers.

Serial Attached SCSI (SAS)
A point-to-point serial protocol that moves data to and from such computer storage devices as hard drives and tape drives.

service network
A private network that is dedicated to

managing POWER8 servers. Provides Ethernet-based connectivity among the FSP, CPC, HMC, and management server.

SMP See *symmetric multiprocessing (SMP)*.

Spanning Tree Protocol (STP)
A network protocol that ensures a loop-free topology for any bridged Ethernet local-area network. The basic function of STP is to prevent bridge loops and the broadcast radiation that results from them.

SSH See *secure shell (SSH)*.

STP See *Spanning Tree Protocol (STP)*.

symmetric multiprocessing (SMP)
A computer architecture that provides fast performance by making multiple processors available to complete individual processes simultaneously.

T

TCP See *Transmission Control Protocol (TCP)*.

Transmission Control Protocol (TCP)
A core protocol of the Internet Protocol Suite that provides reliable, ordered, and error-checked delivery of a stream of octets between applications running on hosts communicating over an IP network.

V

VCD See *vdisk configuration data (VCD)*.

vdisk A virtual disk.

vdisk configuration data (VCD)
Configuration data that is associated with a virtual disk.

X

xCAT See *Extreme Cluster/Cloud Administration Toolkit*.

IBM ®

Printed in USA