Elastic Storage Server
Version 5.3.6.2

Quick Deployment Guide

SC28-3151-02

Note

Before using this information and the product it supports, read the information in “Notices” on page 147.

This edition applies to version 5.3.6 of the Elastic Storage Server (ESS) for Power®, to version 5 release 0 modification 5 of the following product editions, and to all subsequent releases and modifications until otherwise indicated in new editions:

- IBM Spectrum® Scale Data Management Edition for IBM® ESS (product number 5765-DME)
- IBM Spectrum Scale Data Access Edition for IBM ESS (product number 5765-DAE)

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

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

Contents

Tables

About this information

This information guides you in installing, or upgrading to, version 5.3.6.x of 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 5.3.6.x library consists of following information units. You can access these publications on IBM Knowledge Center or IBM Publications Center.

- *Elastic Storage Server: Quick Deployment Guide*, SC28-3151
- *Elastic Storage Server: Protocol Nodes Quick Deployment Guide*, SC28-3152
- *Elastic Storage Server: Problem Determination Guide*, GC28-3153
- *Elastic Storage Server: Command Reference*, SC28-3154
- *IBM Spectrum Scale RAID: Administration*, SC28-3142
- *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
- *IBM ESS Expansion: Hardware Installation and Maintenance Guide (Model 106)*, SC27-9211
- *IBM ESS Expansion: Overview of CMA and Rail Kit Hardware Fasteners (Model 106)*, SC27-9296
- *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 information about the DCS3700 storage enclosure, see:
  - *System Storage® DCS3700 Quick Start Guide*, GA32-0960-04:

    https://www-01.ibm.com/support/docview.wss?uid=ssg1S7005178
  - *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
- For information about 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

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/gnrfaq.html

Other related information

For information about:

- IBM Spectrum Scale, see:

http://www.ibm.com/support/knowledgecenter/STXKQY/ibmspectrumscale_welcome.html

- IBM Spectrum Scale call home, see Understanding call home.
- Installing IBM Spectrum Scale and CES protocols with the installation toolkit, see Installing IBM Spectrum Scale on Linux® nodes with the installation toolkit.
- IBM POWER8® servers, see IBM Knowledge Center:

  http://www.ibm.com/support/knowledgecenter/POWER8/p8hdx/POWER8welcome.htm
- Extreme Cluster/Cloud Administration Toolkit (xCAT), go to the xCAT website :

  http://xcat.org/

  – xCAT 2.15.1 Release Notes®
- Mellanox OFED Release Notes (4.9), go to https://docs.mellanox.com/display/OFEDv490170/Release%20Notes
- IBM Electronic Service Agent (ESA) documentation, go to https://www-01.ibm.com/support/esa/.
- Drive call home, go to Drive call home in 5146 and 5148 systems.

Conventions used in this information

Table 1 on page viii 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.

Table 1. Conventions (continued)

Convention	Usage
<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	SC28-3151-02	Initial version for ESS 5.3.6.2
2	SC28-3151-01	Initial version for ESS 5.3.6.1
1	SC28-3151-00	Initial version for ESS 5.3.6

Chapter 2. 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).

New features and enhancements

Release	Changes
ESS 5.3.6.2	• Support for IBM Spectrum Scale 5.0.5 PTF4 • New Kernel (3.10.0-1160.6.1) • New Systemd (219-78.el7_9.2) • New Network manager (1.18.8-2.el7_9)
ESS 5.3.6.1	• Support for IBM Spectrum Scale 5.0.5 PTF3 • New Kernel (3.10.0-1127.19.1.el7) • New Systemd (219-67.el7_7.10) • mpt3sas version (31.100.01.00) • New GNR firmware RPM
ESS 5.3.6	• Support for IBM Spectrum Scale 5.0.5 PTF1 efix3 • New Kernel (3.10.0-1062.26.1.el7) • New Systemd (219-67.el7_7.6) • New OFED (MLNX_OFED_LINUX-4.9-0.1.7.3) • New IPR (19512900) • New Power 8 System Firmware (FW860.81 (SV860_215))

Release	Changes
	• New ESA (4.5.5-0) • New xCAT (2.15.1) • Deployment fixes and improvements
ESS 5.3.5.2	• Support for IBM Spectrum Scale 5.0.4 PTF3 + efix 2 • New Kernel (3.10.0-1062.12.1) • Updated systemd (219-67.el7_7.4) • Updated network manager (1.18.0-5.el7_7.2) • Deployment bug fixes
ESS 5.3.5.1	• Support for IBM Spectrum Scale 5.0.4 PTF2 + efix 5 • Support for Red Hat Enterprise Linux 7.7 • New Kernel on top of RHEL 7.7 (3.10.0-1062.9.1) • Support for MLNX_OFED_LINUX-4.7-3.2.9.1 • Support for IPR Firmware 19512600 • Support for new Power 8 Firmware SV860_212 (FW860.80) • Support for new HMC version V9R1M940_SP0 (PPC64BE only) • Support for ESA agent 4.5.3-2
ESS 5.3.5	• Support for IBM Spectrum Scale 5.0.4 PTF1 + efix • Updated kernel, systemd, network manager • Support for ESA agent 4.5.2-1 • Support for IPR Firmware 19512300 • Support for new Power 8 Firmware SV860_205 (FW860.70) • Support for new HMC version V9R1M920_SP0 (PPC64BE only) • Support for IBM Elastic Storage® System 3000 (ESS 3000) • Support for GL3C and GL3S • Support for MES GL2S -> GL3S -> GL4S • Support for Power 8 (PPC64LE) node call home • Enhanced **essutils** (formerly **gssutils**) support for SSRs
ESS 5.3.4.2	• Support for IBM Spectrum Scale 5.0.3 PTF3 • Support for Mellanox OFED 4.6.3 • Support for Red Hat Enterprise Linux 7.6 • Updated kernel, systemd, network manager

Release	Changes
	• Support for xCAT 2.14.6 • Support for ESA agent 4.5.1-1 • Support for IPR Firmware 19512200 • Support for new Power 8 Firmware SV860_205 / FW860.70
ESS 5.3.4.1	• New Mellanox OFED (4.6-3) which reinstates support on Connect-IB adapters • New drive and enclosure firmware updates, which resolves certain stability and enclosure fan speed issues • New patch, which resolves PPC64LE serviceable events from disappearing from the OPAL log • Updated ESA RPM, which fixes a few Call Home issues • Updated IBM Spectrum Scale (5.0.3 PTF2 + efix4) • Updated Kernel version • Minor ESS deployment bug fixes and improvements
ESS 5.3.4	• Support for new security features – Sudo – SELinux – Firewall – Admin mode central – Security certificate support in call home setup • Support for Red Hat Enterprise Linux 7.6 • Support for IBM Spectrum Scale 5.0.3 PTF1 • Support for new xCAT version (2.14.6) • Support for new IPR version (19512200) • Support for new systemd version (219-62.el7_6.6) • Support for new Network Manager version (1.12.0-10.el7_6) • Support for new MOFED version (4.6-1.0.1.2) • Support for new kernel version (3.10.0-957.12.2) • Support for new ESA version (4.5.1-0) • mmvdisk changed to default deployment method • Deployment support for Broadcom adapters • Support for new MES options (GLxC Support) – GL1C to GL2C – GL2C to GL4C – GL4C to GL5C

Release	Changes
	• Deployment improvements and bug fixes

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: 5.0.5.4
- xCAT: 2.15.1
- HMC: V9R1M941 (V9R1M940_SP0)
- System firmware: FW860.81 (SV860_215)
- Red Hat Enterprise Linux: 7.7 (PPC64BE and PPC64LE)
- Kernel: 3.10.0-1160.6.1
- Systemd: 219-78.el7_9.2
- Network Manager: 1.18.8-2.el7_9
- mpt3sas: 31.100.01.00
- IPR: 19512900
- SAS adapter driver: 15.00.00.00
- Support RPMs: gpfs.gnr.support-essbase-1.0.0-2.noarch.rpm
- ESA: 4.5.5-0
- Enclosure firmware:
  - PPC64LE:
    - 2U24 = 4230
    - 5U84 = 4087
    - 4U106 = 5266
  - PPC64BE
    - 2U24 = 65SG
    - DCS3700 = 039D
- OFED: MLNX_OFED_LINUX-4.9-0.1.7.3

  OFED firmware levels:
  - MT27500 = 10.16.1200
  - MT4099 = 2.42.5000
  - MT26448 = 2.9.1326
  - MT4103 = 2.42.5000
  - MT4113 = 10.16.1200
  - MT4115 = 12.27.2008
  - MT4117 = 14.27.2008
  - MT4119 = 16.27.2008
  - MT4120 = 16.27.2008
  - MT4121 = 16.27.2008

– MT4122 = 16.27.2008

Supported editions on each architecture

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

PPC64BE

- Data Access Edition
- Advanced Edition
- Data Management Edition

PPC64LE

- Data Access Edition
- Data Management Edition

ESS best practices and support statements

- `versionlocks` are not enabled in ESS. In the past, `versionlocks` were used to protect against unwarranted kernel and OFED updates. Although `versionlocks` no longer exist, the same rules apply regarding ESS packages verified by **gssinstallcheck** or **gssinstall**. You must get specific approval and guidance from the L2 Service to make any changes to an ESS configuration. The only exceptions are RHEL packages in addition to kernel, systemd, and network manager packages. RHEL packages other than the mentioned packages are customer responsibility and they might be updated for security purposes.
- If you are upgrading to ESS 5.3.6.2, you must convert the environment to **mmvdisk** after the upgrade is completed.
- It is advised that you set autoload to on to enable GPFS to recover automatically in case of a daemon problem. Deployment automatically enables this on new installations but you should disable `autoload` for an upgrade and re-enable it after an upgrade.

  To disable, issue the following command:

  ```
  mmchconfig autoload=no
  ```

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

  ```
  mmchconfig autoload=yes
  ```

- Do not mount the file system on the ESS I/O server nodes.
- It is advised that you disable automount for file systems when performing an upgrade to ESS 5.3.1 or later.

  ```
  mmchfs Device -A no
  ```

  Device is the device name of the file system.

  Automount should automatically be disabled when creating new file systems with **gssgenvdisks**.

  Remember: Mount the file system only on the EMS node where the GUI and the PM collector run.

  .
- Do not configure more than 5 failure groups in a single file system.
- Consider moving all supported Infiniband devices to the Datagram mode (CONNECTED_MODE=no). For more information, see "ESS networking considerations" on page 85.
- Running any additional service or protocols on any ESS node is not supported. This includes installing any additional RPMs, running any protocols (or any other type of service), or mounting the file system on any ESS I/O server node. This also applies to the EMS node, although you must mount the file system to support the IBM Spectrum Scale GUI.

- RoCE (RDMA over Ethernet) is not supported in ESS.
- Consider moving quorum, cluster, and file system management responsibilities from the ESS nodes to other server license nodes within the cluster.
- It is not required, though highly recommended, that the code levels match during a building block addition. Be mindful of changing the release and file system format in mixed IBM Spectrum Scale environments.
- 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.
- It is acceptable for LBS or customers to update any security errata available from Red Hat Network (RHN). Only components checked and protected by ESS (for example, kernel, network manager, systemd) must not be modified unless advised by the IBM service. For more information on applying security erratas see https://access.redhat.com/solutions/10021
- Client node deployment is not supported from the ESS management node.
- You must deploy or add building blocks from an EMS with the same architecture. There must be a dedicated EMS for each architecture (PPC64BE or PPC64LE).
- If running in a mixed architecture environment, the GUI and PM collector are recommended to run on the PPC64LE EMS node.
- Modifying any ESS nodes as a proxy server is not supported.
- PPC64LE to PPC64BE conversions and vice versa are not supported.
- Multiple building blocks are ideal as ESS now by default uses file system level metadata replication. If a single building block is used, by default **gssgenvdisks** uses one failure group and only IBM Spectrum Scale RAID level metadata replication.
- It is recommended to use the highest available block size when creating vdisks or NSDs. The default block size is 16M (current maximum). If the customer primarily generates many tiny files (metadata heavy), consider splitting metadata and data NSDs and using smaller block sizes.
- It is recommended that all nodes in a cluster run the same version of Mellanox OFED.
- Automatic EMS failover is not supported. For help in setting up a redundant, standby EMS, contact the L2 service.
- 4K MTU (InfiniBand) and 9000 MTU (Ethernet) are recommended. Changing to these MTU values requires associated switch-side changes.
- Stretch clusters are supported in various configurations. Contact development or service for guidance.
- If using a PPC64LE building block (8247), note that HMC is not used in that configuration. HMC is applicable for PPC64BE only.
- ConnectX2 (ConnectX-EN) adapters are still supported by ESS.

Obtaining the required Red Hat Enterprise Linux and ESS code

Note: Contact IBM to obtain access to an online folder containing the required items for deployment.

The required Red Hat components and SHA256 are:

- Red Hat Enterprise Linux 7.7 ISO

```
756df1ed5dbe38d5d41d342ce54f72d1f372d3f9bf4f88c3f9bc945ec5841366  RHEL-7.7-20190723.1-Server-
ppc64le-dvd1.iso
aa7d9bce6576eeaf34d13a442d4c668a954b906e54c24bdb4a7dc4181ff9dff5  RHEL-7.7-20190723.1-Server-
ppc64-dvd1.iso
```

- Network manager version : 1.18.8-2.el7_9

```
11d47b72a39a52eef8b8d8b8e90a6853b13c29031674ed66994e00559ae69834  netmgr_5362_LE.tgz
7ca1d5e36356bf87a725f1fbbacc97cfda6a5e826f64f5bdbc0ff718d3037985  netmgr_5362_BE.tgz
```

- Systemd version: 219-78.el7_9.2

```
6d9a703cc0ffa3c1036e953f26d1f3358211bed921402b3c639ca22a5ef7744e  systemd_5362_LE.tgz
16dc8ed912a7e0e83f496af38e0fab75945670633752213282d26369adeccfef  systemd_5362_BE.tgz
```

- Kernel version: 3.10.0-1160.6.1

```
e7ea7a835994a6658cca1335cc94ce46a8b9f9f73567ccd621259ee2ea664bb3  kernel_5362_LE.tgz
3c7ea505ab0953b0638c9206787397d6befa2337af7997940198fca65f5e12fd  kernel_5362_BE.tgz
```

- Power 8 OPAL patch

```
ea9c602234f446f009009eaba9634f40750da4071523d0e2c59dc646a35a2766  opal-patch-le.tar.gz
198cfb1230f6f14a98346f7d923c6e3a2498613a0f6e4229d51f375554c2b252  opal-patch-be.tar.gz
```

On ESS 5.3.6.2 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:

- "Obtaining kernel for system upgrades" on page 113
- "Obtaining systemd update for system upgrades" on page 117
- "Obtaining Network Manager updates for system upgrades" on page 121

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

Available PPC64BE packages:

```
ESS_DA_BASEIMAGE-5.3.6.2-ppc64-Linux.tgz
ESS_ADV_BASEIMAGE-5.3.6.2-ppc64-Linux.tgz
ESS_DM_BASEIMAGE-5.3.6.2-ppc64-Linux.tgz
```

Available PPC64LE packages:

```
ESS_DA_BASEIMAGE-5.3.6.2-ppc64le-Linux.tgz
ESS_DM_BASEIMAGE-5.3.6.2-ppc64le-Linux.tgz
```

ESS 5.3.6.2 can be downloaded from IBM FixCentral.

Once downloaded and placed in `/home/deploy`, untar and uncompress the package to view the contents. For example, for the Data Access edition PPC64LE package, use the following command:

```
tar -zxvf ESS_DA_BASEIMAGE-5.3.6.2-ppc64le-Linux.tgz
```

For example, from the BASEIMAGE tar file, files such as the following get extracted with the preceding command:

- `ESS_5.3.6.2_ppc64le_Release_note_Data_Access.txt`: This file contains the release notes for the latest code.
- `gss_install-5.3.6.2_ppc64le_dataaccess_20201116T223856Z.tgz`: This `.tgz` file contains the ESS code.
- `gss_install-5.3.6.2_ppc64le_dataaccess_20201116T223856Z.sha256`: This `.sha256` file to check the integrity of the tgz file.

Support for signed RPMs

ESS or IBM Spectrum Scale RPMs are signed by IBM.

The PGP key is located in `/opt/ibm/gss/tools/conf`

```
-rw-r-xr-x 1 root root 907 Dec 1 07:45 SpectrumScale_public_key.pgp
```

You can check if an ESS or IBM Spectrum Scale RPM is signed by IBM as follows.

1. Import the PGP key.

   ```
   rpm --import  /opt/ibm/gss/tools/conf/SpectrumScale_public_key.pgp
   ```

2. Verify the RPM.

   ```
   rpm -K RPMFile
   ```

Supported upgrade paths

Note:

A rule of thumb is that you can hop one OS level at a time. For example:

- RHEL 7.5 -> RHEL 7.6 upgrade can be done in one hop

It is recommended that if you are doing an offline upgrade, it is safe to hop two OS releases (RHEL 7.4 -> RHEL 7.6). If you are doing an online upgrade, it is advised to do only one OS hop at a time (RHEL 7.5 -> RHEL 7.6). Review the GNR FAQ to see which ESS releases support the various OS levels.

Mixed environment recommendations

Running ESS building blocks of mixed levels is not recommended. If you choose to do so, following recommendations apply:

- Nodes within a building block must be at the same levels.
- Nodes between building blocks should not be greater than N-2 (OFED 4.4 and OFED 4.6, for example).
- Same rules apply to PPC64BE and PPC64LE mixture.

ESS 3000 considerations (POWER8 EMS)

- If your system came racked with EMS but no ESS3000 (Any supported legacy node and an EMS node are in the order):

  Use the installation flow in this document (**gsschenv**)
- If any other configuration but no ESS 3000:

  Refer to the "Legacy deployment instructions" on page 137.
- If your system comes in any configuration with ESS 3000, refer to the IBM ESS 3000 Version 6.0.1.x Knowledge Center.
- Your EMS node must be at version 5.3.5 or later, with podman installed and C10-T2 connection, to support ESS 3000
- You do not have to upgrade to ESS 5.3.6 to support ESS 3000.
- If EMS + ESS 3000:

  The minimum configuration is EMS at version 5.3.5 with podman + C10-T2 connection + IBM Spectrum Scale 5.0.5.1 (Updated from container).
- If EMS + ESS 3000 + ESS:

  It is advised to upgrade the EMS and ESS first to version 5.3.6.2. This is not a hard requirement, but the version must be 5.3.5 or later.

Security law changes

- New systems and switches shipped from manufacturing now have either an expired password or one set to the serial number of the component.
- You must take input from the customer before deployment starts and change the desired passwords.
- The default root password for the OS is `ibmesscluster`. You are required to change it upon first login. This password must be set the same on each node.

- The default ASMI passwords (login, IPMI, HMC, etc.) are set to the serial number of the server. IPMI must be the same on each node.
- If the 1Gb Cumulus switch is shipped racked, the default password is the serial number (S11 number - label found on the back of the switch). If the switch is shipped unracked, you are required to set the password upon first login. The default password is `CumulusLinux!` but you will be prompted to change the password upon first login. If you have any issues logging in or you need help in setting up a VLAN with the switch, consult this documentation link.
- You must set all required passwords before the deployment begins.

Support for hardware call home

	PPC64BE	PPC64LE
Call home when disk needs to be replaced	Supported	Supported
Enclosure call home	Unsupported	Unsupported
Server call home	Supported through HMC	Supported

For more information, see:

- "Call home configuration" on page 75
- Drive call home in 5146 and 5148 systems

Pre-installation or pre-upgrade checklist

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

	Obtain the kernel, systemd, network manager, RHEL ISO, Power8 OPAL patch (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). To obtain these items from the internal folder, you must be an IBM employee. Business partners or customers cannot access this data directly.
	Ensure that you read all the information in the ESS Quick Deployment Guide. Make sure that you have the latest copy from the IBM Knowledge Center and the version matches accordingly. You should also refer to the related ESS 5.3.6.2 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 HMC v9 (V9R1M940_SP0) if doing a PPC64BE installation. This can be done concurrently. The SSR or the customer might be able to do this ahead of time.
	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 87.
	Review *Elastic Storage Server: Command Reference*.

	Review ESS FAQ and ESS best practices.
	Review the ESS 5.3.6.2 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 completed. **Note:** You would only use **gssdeploy -x** if the legacy installation sequence is used.
	If adding an PPC64LE building block to an existing PPC64BE building block, carefully review "Considerations for adding PPC64LE building blocks to ESS PPC64BE building blocks" on page 143.
	If installing or upgrading protocol nodes, carefully review *Elastic Storage Server: Protocol Nodes Quick Deployment Guide*.
	Carefully study the network diagram for the architecture used. For more information, see "ESS networking considerations" on page 85 and *5148-22L protocol node diagrams* in *Elastic Storage Server: Protocol Nodes Quick Deployment Guide*.
	It is recommended to use a larger block size with IBM Spectrum Scale 5.0.0 or later, even for small I/O tasks. Consult the documentation carefully.
	Ensure that the correct edition of ESS is to be deployed. For example, do not install the Data Management Edition if Data Access Edition is on the order. This must be verified even before Plug-N-Play is attempted.
	Determine the supported high-speed switch bonding mode and, if Infiniband, determine which MTU will be used. The default MTU is now 2048 but it can be changed to 4092.
	Consult the local network team to see if a Fabric diagnostic has taken place. The use of **ibdiagnet** is one way to debug an unhealthy network environment.
	Develop a plan to tune the client nodes. Deployment offers a template but depending on the workload or application type, you might need to make many adjustments. Consult the IBM Spectrum Scale tuning guide.
	Discuss with the customer about password changes required. You must be prepared before starting to set the desired passwords for the customer for the various components: • Operating system (root password should be same on each node) • FSP IPMI password (Needs to be same on each node) • Other ASMI passwords • Management and high-speed switch passwords

Post-installation or post-upgrade checklist

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

	Hardware and software call home have been set up and tested. If applicable, consider postponing the call home setup until the protocol nodes are deployed. • For call home configuration information, see "Call home configuration" on page 75. • For more information, see Drive call home in 5146 and 5148 systems. • 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 21. • For software call home information, see Software call home.
	GUI has been set up and demonstrated to the customer. If applicable, consider postponing the GUI setup until the protocol nodes are deployed.
	GUI SNMP and SMTP 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.
	Client nodes are properly tuned. For more information, see "Adding IBM Spectrum Scale nodes to an ESS cluster" on page 107.
	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**). Do not update any kernel, systemd, or network manager erratas.
	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 *Elastic Storage Server: Protocol Nodes Quick Deployment Guide*.
	Ensure (if possible) that all network switches have had the firmware updated.
	IBM Spectrum Scale release level and file system format have been updated, if applicable.
	If there are more than one building blocks, make sure multiple failure groups are used in a file system and that metadata replication is turned on (-m 2 and -M 2). Do not exceed 5 failure groups or more than 2 metadata replicas.
	Upgrade file system format and move release level to LATEST depending on the IBM Spectrum Scale client node compatibility. Move the Recovery Group format to LATEST.
	Consider offering your clients the following trainings to better enable the proper administration of an IBM Spectrum Scale cluster and ESS storage. • IBM Spectrum Scale Redbook • IBM Spectrum Scale basic training classes • IBM Spectrum Scale advanced training classes
	Take a CCR backup and save it to a secure location in case you need to restore the cluster. Consider implementing the **mmsdrbackup** user exit. For more information, see mmsdrbackup user exit documentation.
	New deployments of ESS 5.3.6.2 are under **mmvdisk** management by default. After upgrading to ESS 5.3.6.2, you must convert the recovery groups to **mmvdisk** as well. Convert to **mmvdisk**, if currently in the legacy mode.

	1. Check if there are any **mmvdisk** node classes. `mmvdisk nodeclass list` 2. If there are none, convert the recovery groups to **mmvdisk** by issuing this command. `gssgenclusterrgs -G gss_ppc64 --suffix=-hs --convert` **Note:** Wait for 5 minutes for daemons to recycle. The file system remains up.
	You must ensure that the Mellanox performance script is run before the system placed into production (post deployment or upgrade). This script is run automatically on new deployments. For upgrades, after the OFED is installed, do the following steps. 1. Run the following command: `/xcatpost/mlnx_params.sh -s 5 -p` 2. Reboot the node. 3. After reboot, verify by using the following command. `/xcatpost/mlnx_params.sh`

Other topics

For help with the following topics, and many others that are unlisted, contact L2 Service.

- Restoring a management server
- Part upgrades or replacements
- VLAN reconfiguration on the 1Gb switch
- Extending the 1Gb Cumulus management switch
- Stretch cluster considerations

Sample installation and upgrade flow

New installations go through manufacturing CSC. The system is fully installed with ESS 5.3.6.2, tested, malfunctioning parts replaced, and required RHEL pieces shipped in `/home/deploy`.

Installation

- SSR checkout complete
- LBS arrival on site
- Plug-n-Play mode demonstrated
- Decisions made on file system names and sizes, block size, host names, IP addresses, and so on
- Check high-speed switch settings and firmware
- Deploy EMS and building block
- Network bonds created
- Cluster created
- Recovery groups, NSDs, file system created
- Stress test performed
- Final checks performed

Proceed to install the protocol nodes, if applicable.

- GUI setup (w/SNMP alerts if desired)
- Call home setup

- Nodes attached to RHN and security updates applied

Upgrade

- Check high speed switch settings and 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)
- Move the release level and the file system format, if applicable. Move the Recovery Group format to LATEST.
- Final checks are performed
- If applicable, upgrade the ESS protocol nodes
- Ensure that call home and GUI are still working as expected
- Use `yum` to upgrade any security related errata (**`yum -y security`**). Do not update any kernel, systemd, or network manager erratas.

Chapter 3. Installing Elastic Storage Server

Note:

- Any versions mentioned in the following steps are just examples and they might vary from the actual product versions.
- These instructions assume a new system from manufacturing (EMS + one or more building-blocks and, optionally, POWER8 protocol nodes). If the EMS is not in the order, follow "Legacy deployment instructions" on page 137.

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

Note: Though manufacturing supplies the relevant pieces needed to deploy the system, it is always a good practice to consult the latest release notes and *Elastic Storage Server: Quick Deployment Guide* for any changes before starting 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
- ESS tarball
- Power 8 OPAL patch

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. Manufacturing has kept intact `/etc/hosts` on each node and a copy of `gssdeploy.cfg` in `/var/tmp` on the EMS node.

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 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 "Optional: Plug-N-Play" on page 34.

Note: If Plug-N-Play was used, clean up the system before proceeding with full installation. For more information, see "Post Plug-N-Play cleanup steps " on page 35.

Before proceeding further, ensure that you have reviewed and completed the pre-installation and pre-upgrade checklist. For more information, see "Pre-installation or pre-upgrade checklist" on page 21.

ESS installation flow

Note: Check if xCAT is installed before proceeding with the following installation flow. It is assumed that this is for a new system that has an EMS in the same order. If this is not the case or xCAT is not installed, use the legacy installation instructions. For more information, see "Legacy deployment instructions" on page 137.

To determine if xCAT is installed, use this command:

```
rpm -qa | grep -i xcat
lsxcatd -v
```

If xCAT is installed but the database is empty, check if there is an xCAT database backup in the `/home/deploy` directory. If it exists, restore the database by using the following command:

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

If xCAT is not installed, then follow the legacy installation instructions. For more information, see “Legacy deployment instructions” on page 137.

Extract the Elastic Storage Server software

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

1. Determine the current ESS version shipped from manufacturing.

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

   Save this information because it is needed later in the installation flow.
2. Unpack the ESS software archive (This is contained in the ESS_DA_BASEIMAGE-5.3.6.2-ppc64le-Linux.tgz file).

   ```
   tar -zxvf gss_install-5.3.6.2_ppc64le_dataaccess_20201116T223856Z.tgz
   ```

3. Check the SHA 256 checksum:

   ```
   shasum -a 256 -c gss_install-5.3.6.2_ppc64le_dataaccess_20201116T223856Z.sha256
   ```

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

   ```
   /bin/sh gss_install-5.3.6.2_ppc64le_dataaccess_20201116T223856Z --remove
   ```

   Depending on the architecture, replace *<ARCH>* with ppc64 or ppc64le.
5. 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-5.3.6.2_ppc64le_dataaccess_20201116T223856Z --text-only
   ```

Deploy the ESS

Before proceeding with the following steps, ensure that you have completed all the steps in “Extract the Elastic Storage Server software” on page 28.

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 xCAT commands 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 not yet executable:

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

2. 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
   ```

3. 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**.

```
# cat gsschenv.cfg
# Modify the following
# HOSTNAME_ORIG =  Original hostnames in your xCAT ESS environment
# IP_ADDR_ORIG = Original IPs in your xCAT ESS environment  want (1 to 1 with
HOSTNAME_ORIG)
# 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)
IP_ADDR_ORIG=(192.168.45.20 192.168.45.21 192.168.45.22)
HOSTNAME_NEW=(modems1 modgssio1 modgssio2)
IP_ADDR_NEW=(192.168.45.40 192.168.45.41 192.168.45.42)
NETMASK="255.255.255.0"
DOMAIN="gpfs.net"
```

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

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

5. Run **systemctl reboot** to reboot the management server node.
6. 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 /etc/hosts
```

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

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

Note: First verify that `/var/tmp/gssdeploy.cfg` exists and that it has the default entries from manufacturing. If it does not exist, you need to create it.

8. Compare the installed ESS version and levels.

Compare that the ESS version installed on each node matches the level downloaded and used from FixCentral. Also, compare that each core component (non-IBM Spectrum Scale configuration) is up-to-date.

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

Compare against the version checked at the beginning of the installation flow.

- **If the version matches with the version that was saved in step 1 of "Extract the Elastic Storage Server software" on page 28**, do the following steps.

  a. Verify the installation by running `gssinstallcheck:`

  ```
  gssinstallcheck -N EMSNode,IONode1,IoNode2
  ```

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

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

  For more information, see "Running gssinstallcheck in parallel" on page 133.

  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

b. If there are any issues reported in **gssinstallcheck**, perform Step 5 to Step 10 of "Update the I/O server nodes" on page 42.

c. If there are no issues reported in **gssinstallcheck**, skip to "Check the system hardware" on page 30 and proceed with the remaining installation flow.

- **If the version does not match with the version that was saved in step 1 of "Extract the Elastic Storage Server software" on page 28, run the following command sequence in the specified order:**

  a. Step 6 to step 10 of "Update the management server node" on page 39.

  b. Step 5 to Step 10 of "Update the I/O server nodes" on page 42.

  c. Proceed with "Check the system hardware" on page 30 and the remaining installation flow.

  Note: Replace *CurrentIoServer* with *gss_ppc64* in these steps.

Check the system hardware

Ensure that this step is fully completed before beginning this procedure.

Power up 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, syslog from all I/O server nodes are directed to the EMS in `/var/log/messages`.
- 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` on the EMS.

1. Run `gssstoragequickcheck`:

```
gssstoragequickcheck -N IONode1,IONode2
```

2. Run `gssfindmissingdisks`:

```
gssfindmissingdisks -N IONode1,IONode2
```

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
```

3. Run `gsscheckdisks`:

```
xdsh IONode GSSENV=INSTALL gsscheckdisks --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.

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

   ```
   gssinstallcheck -N EMSNode,IONode1,IONode2 --srv-events
   ```

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

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. Before creating network bonds, carefully read "ESS networking considerations" on page 85.

- 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. The default MTU is 2048 (2K) and the **gssgennetworks** command supports 2048 (2K) and 4092 (4K) MTU. 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
  ```

Note: For information on Infiniband issue with multiple fabrics, see *Infiniband with multiple fabric* in "Customer networking considerations" on page 87.

Create the cluster, recovery groups, and file system

1. Create the GPFS cluster:

   ```
   gssgencluster -C test01 -G ems1,gss_ppc64 --suffix=-hs --accept-license --add-ems-in-cluster
   ```

   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-filesystem
   ```

   Note: gssgenvdisks, by default, creates vdisk containing both data and metadata with 8+2p RAID code and 16 MB block size.

Check the installed software and system health

1. Run **gssinstallcheck** in parallel 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 "Running gssinstallcheck in parallel" on page 133.

   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. Run the Mellanox performance script.

      ```
      /xcatpost/mlnx_params.sh -s 5 -p
      ```

   b. Shut down GPFS all nodes:

      ```
      mmshutdown -a
      ```

   c. Reboot all server nodes:

      ```
      xdsh gss_ppc64 "systemctl reboot"
      ```

   d. Reboot the management server node:

      ```
      systemctl reboot
      ```

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

   ```
   gssinstallcheck -N IONode1,IONode2 --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 file system 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 EMSNode,IONode1,IONode2 --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 EMSNode,IONode1,IONode2 --srv-events
   ```

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

   Note:

   - On PPC64BE systems, investigate, manage, and close serviceable events from HMC.
   - On PPC64LE systems, ASMI can be used to investigate issues.

8. Verify that NTP is set up and enabled. For more information, see “NTP setup” on page 135.

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.
- It is recommended to delay the GUI setup, if protocol nodes will be immediately deployed. Once the ESS and protocol nodes are deployed, run the wizard to properly discover and slot the nodes into the rack.

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 file system 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
   ```

   Note: If your ESS system came with 5148-22L protocol nodes, wait until the protocol nodes are installed before setting up the GUI.

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. For information on action items to be done after installation, see “Post-installation or post-upgrade checklist” on page 22.

Optional: Plug-N-Play

Using Plug-N-Play with **essutils** is an option. For more information, see *essutils command* in *Elastic Storage Server: Command Reference*. The goal of Plug-N-Play 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 **essutils** to ensure correct disk placement, cabling, networking, and server health.
- Access to EMS is available over SSH for LBS.

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.
- The high-speed network may or may not be ready at the time of Plug-N-Play. For the purposes of this demonstration, the cluster is set up over the 1Gb network.

Note: The code level shipped from manufacturing might be lower than the code level currently available on FixCentral. This can be ignored in the context of Plug-N-Play. Use the version that is available on the system.

Overview

The primary objective 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, 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.
- Basic hardware checking:
  - **gssstoragequickcheck**
  - **gssfindmissingdisks**
  - **gsscheckdisks**
- Basic file system creation (use of entire space, single pool, 16M block size, 8+2p RAID code)
- GUI and performance monitoring setup

Work flow

1. System delivered to customer site. SSR arrives and performs basic unpacking and cabling. All nodes are powered on to the operating system. SSR does a full hardware check using **essutils**. They replace any bad components prior to LBS arrival.
2. LBS logs in to EMS through SSH.
3. Run the **gssprecheck** script in the full install mode and address any errors.

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

   Note: First verify that `/var/tmp/gssdeploy.cfg` exists and it has the default entries from manufacturing. If it does not exist, you need to create it.
4. Proceed to running the standard set of ESS verification checks.
   - **gssstoragequickcheck**
   - **gssfindmissingdisks**
   - **gsscheckdisks**

   For more information, see the usage in "Check the system hardware" on page 30.
5. Create your cluster using **gssgencluster**.

   ```
   gssgencluster -C test01 -G ems1,gss_ppc64 --suffix=-hs --accept-license --add-ems-in-cluster
   ```
6. Create your recovery groups using **gssgencluserrgs**.

   ```
   gssgenclusterrgs -G gss_ppc64
   ```
7. Create your file system using **gssgenvdisks**.

   ```
   gssgenvdisks --create-vdisk --create-filesystem
   ```
8. Set up the performance monitoring collector and sensors. For more information, see this section.

   Note: Ensure that you use the low-speed (1Gb) host names.
9. Start the GUI on the EMS and walk through the setup wizard.

   ```
   systemctl start gpfsgui
   ```

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. This mode shows how fast an ESS system can be brought up and used at a customer site.

Post Plug-N-Play cleanup steps

Use these steps to clean up an ESS system if Plug-N-Play was used before doing full installation.

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 system.

   ```
   mmumount FSDeviceName -a
   ```
4. Delete the file system.

   ```
   mmdelfs FSDeviceName
   ```

Where *FSDeviceName* is the file system device name.

You can also use the **mmvdisk** command to delete the file system along with the vdisk set and the recovery group.

```
mmvdisk filesystem delete --file-system FSDeviceName
```

5. List the vdisk sets.

   ```
   mmvdisk vdiskset list
   ```

6. Delete the vdisk set for the deleted file system.

   ```
   mmvdisk vdiskset delete --vdisk-set VdiskSet
   ```

   This command deletes NSDs, and data and metadata vdisks.

7. Undefine the vdisks sets.

   ```
   mmvdisk vdiskset undefine --vdisk-set VdiskSet
   ```

8. List the recovery groups.

   ```
   mmvdisk recoverygroup list
   ```

9. Delete the recovery groups.

   ```
   mmvdisk recoverygroup delete --recovery-group RecoveryGroup
   ```

10. List the mmvdisk servers.

    ```
    mmvdisk server list
    ```

11. Unconfigure the servers.

    ```
    mmvdisk server unconfigure --node-class NodeClass
    ```

12. Delete the mmvdisk node class.

    ```
    mmvdisk nodeclass delete --node-class NodeClass
    ```

13. Shut down GPFS.

    ```
    mmshutdown -a
    ```

14. Delete the cluster.

    ```
    mmdelnode -a
    ```

15. Break the network bonds on each node.

    ```
    cd /etc/sysconfig/network-scripts ; rm -f *bond*
    nmcli c reload
    ```

Chapter 4. Upgrading Elastic Storage Server

The following steps are required to complete a new ESS upgrade.

Note:

- Protocol nodes must be upgraded after the upgrade on ESS nodes is completed.
- Any security features must be disabled while upgrading the cluster. Security features must be enabled after the upgrade cycle is completed.
- Offline upgrades from any prior ESS version are supported but use the guidance to general best practices (typically, maximum N-2 Red Hat Enterprise Linux version hop). An offline upgrade involves taking GPFS down cluster wide before doing the upgrade. Carefully follow the documented steps and look for actions specific to offline upgrades.
- Offline upgrades can be done to multiple building blocks at once.

CAUTION: Make sure that quorum rules for the whole cluster are considered before starting the offline upgrade.

- Power 8 firmware is not automatically updated. In the following procedure, the best time to upgrade the Power 8 firmware is advised. For more information, see “Updating the system firmware” on page 109.
- The HMC (PPC64BE) can be upgraded at any time without disrupting a production system.
- Package names mentioned in the following steps are examples and they might not match with the package names that you are upgrading to. This is also applicable to the architecture mentioned in the example package names. The default architecture is PPC64LE.
- If possible, perform the upgrade while the system is offline. This is faster and it is less prone to issues.

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 *Upgrading a cluster containing ESS and protocol nodes* in *Elastic Storage Server: Protocol Nodes Quick Deployment Guide*.

Prerequisites

Before you begin the upgrade procedure, do the following:

- Obtain the ESS tarball, kernel, systemd, RHEL ISO, Power 8 OPAL patch, and network manager packages for the architecture being used.
- Archive the current contents of `/home/deploy` and move the 5.3.6.2 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
```

- All health checks must be clean before attempting to upgrade a system. This applies to offline and online upgrades.
- Ensure that the HMC is upgraded to the recommended level. For more information, see “Upgrading the Hardware Management Console (HMC)” on page 111.
- 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.

- Ensure that you have reviewed and completed the pre-installation and pre-upgrade checklist. For more information, see "Pre-installation or pre-upgrade checklist" on page 21.
- If you do not have an active cluster, ignore any commands related to starting and stopping IBM Spectrum Scale (**mmstartup** or **mmshutdown**). If you are preforming an online upgrade, it is assumed that you have an active cluster.
- Manually disable any RHN plugin in the file `/etc/yum/pluginconf.d/rhnplugin.conf` and disable or remove any repos in `/etc/yum.repos.d` before starting the upgrade.
- On each node, consider cleaning up old kernels to make space in `/boot`. You may need to do this on each node prior to upgrade.

```
package-cleanup --oldkernels --count=1
```

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`
4. Stop the creation and deletion of snapshots using `mmcrsnapshot` and `mmdelsnapshot` during the upgrade window.
5. Run the following checks.
   a. Run **gssinstallcheck** in parallel to verify that the current system is up to date.

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

   b. Run the serviceable events check and rule out any bad hardware components.

      ```
      gssinstallcheck -N EMSNode,IONode1,IONode2 --srv-events
      ```

   c. Check for any open GUI events or tips.

Upgrading the Elastic Storage Server

1. Unpack the ESS software archive (This is contained in the ESS_DA_BASEIMAGE-5.3.6.2-ppc64-Linux.tgz file.

   ```
   tar -zxvf gss_install-5.3.6.2_ppc64le_dataaccess_20201116T223856Z.tgz
   ```
2. Check the SHA256 checksum:

```
shasum -a 256 -c gss_install-5.3.6.2_ppc64le_dataaccess_20201116T223856Z.sha256
```

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

```
/bin/sh gss_install-5.3.6.2_ppc64le_dataaccess_20201116T223856Z --remove
```

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

   Note: If you are upgrading to 5.3.6.2 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-5.3.6.2_ppc64le_dataaccess_20201116T223856Z –-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-5.3.6.2_ppc64le_dataaccess_20201116T223856Z --text-only
```

5. Make the gssdeploy script executable:

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

6. Perform cleanup and save a backup copy of the xCAT database:

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

7. 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
```

   Note: Before copying gssdeploy.cfg.default to gssdeploy.cfg, it is advised that you save any current copy of gssdeploy.cfg. You can then use that as a reference.

8. 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.

9. Customize the gssdeploy.cfg configuration file according to your environment.

Update the management server node

1. On the management server node, stop GUI services:

```
systemctl stop gpfsgui
```

2. Copy the RHEL 7.7 ISO file to the directory specified in the gssdeploy.cfg file. The default is /opt/ibm/gss/iso.

3. 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 GPFS using one of the following steps depending on whether online or offline upgrade is being performed.

- **Online upgrade only**

  a. Review the quorum information prior to shutting down GPFS:

  ```
  mmgetstate -s
  ```

  b. Shut down GPFS.

  ```
  mmshutdown
  ```

- **Offline upgrade only**

  Shut down GPFS on the entire cluster.

  Warning: Ensure that this is the step that you want to perform.

  ```
  mmshutdown -a
  ```

6. Set up the kernel, systemd, and network manager errata, and OPAL patch repositories. For example, use the following command on PPC64LE systems:

```
/var/tmp/gssdeploy -k /home/deploy/kernel_ESS_5362_LE.tgz -p \
/home/deploy/systemd_5362_LE.tgz,/home/deploy/netmgr_5362_LE.tgz,\
/home/deploy/opal-patch-le.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 root root 47647044 Nov 17 01:49 kernel-3.10.0-1160.6.1.el7.ppc64le.rpm
```

```
/install/gss/otherpkgs/rhels7/<arch>/patch
-rw-r--r-- 1 root root 5460520 Nov 17 01:49 systemd-219-78.el7_9.2.ppc64le.rpm
-rw-r--r-- 1 root root 2001044 Nov 17 21:54 NetworkManager-1.18.8-2.el7_9.ppc64le.rpm
```

Wrong file permission leads 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.

   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.

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. Update the system firmware.

    Before starting up GPFS, it is a good time to upgrade the system firmware. For more information, see “Updating the system firmware” on page 109.

    - **Online upgrade only**

      Update the EMS node only.

    - **Offline upgrade only**

      System firmware of all nodes can be updated at this time in case of an offline upgrade.

    Note: A system firmware update requires a node reboot.

11. If the firmware update is not needed, use **systemctl reboot** to reboot the management server node.
12. Perform the following steps to upgrade IBM Spectrum Scale RAID configuration parameters.

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

13. Start up GPFS.

    - **Online upgrade only**

      Run this command on the management server node.

```
mmstartup
```

    - **Offline upgrade only**

      Skip this step

14. Verify that GPFS is in the active state before upgrading the I/O server nodes.

    - **Online upgrade only**

      Run this command.

```
mmgetstate
```

    - **Offline upgrade only**

Skip this step

Do not proceed if the system is not active.

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

```
gssinstallcheck -N ems1
```

Note: Offline upgrade only Some checks cannot be performed at this time because the cluster is offline. It is good to verify that the base upgrade has completed on this node (OFED, IPR, Kernel, etc.).

Update the I/O server nodes

Note: This procedure is for online upgrade thus many of the steps can be skipped or done in parallel in case of offline upgrades. Look for **Offline upgrade only** tags in each step for guidance.

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.

- **Online upgrade only**

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.

- **Offline upgrade only**

Skip this step

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

- **Online upgrade only (Legacy)**

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
```

- **Online upgrade only (mmvdisk environment)**

a. To list the recovery groups, run:

```
mmvdisk recoverygroup list
```

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

```
mmvdisk recoverygroup list --recovery-group rg_gssio1-hs --server
```

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:

```
mmvdisk recoverygroup change --recovery-group rg_gssio1-hs --active gssio2-hs
mmvdisk recoverygroup change --recovery-group rg_gssio2-hs --active gssio2-hs
```

- **Offline upgrade only**

Skip this step

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 GPFS on the current I/O server node while maintaining quorum.

- **Online upgrade only**

```
mmumount all -N CurrentIoServer-hs
```

```
mmshutdown -N CurrentIoServer-hs
```

- **Offline upgrade only**

Skip this step

5. Run `updatenode`.

- **Online upgrade only**

```
updatenode CurrentIoServer -P gss_updatenode
```

- **Offline upgrade only**

Run in parallel.

Instead of *CurrentIoServer* use the xCAT group gss_ppc64. For example:

```
updatenode gss_ppc64 -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.

- **Online upgrade only**

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

- **Offline upgrade only**

Run in parallel.

Instead of *CurrentIoServer* use the xCAT group gss_ppc64. For example:

```
xdsh gss_ppc64 "systemctl reboot"
```

7. Run updatenode again (if instructed to do so).
   - **Online upgrade only**

     ```
     updatenode CurrentIoServer -P gss_updatenode
     ```

   - **Offline upgrade only**

     Run in parallel.

     Instead of *CurrentIoServer* use the xCAT group gss_ppc64. For example:

     ```
     updatenode gss_ppc64 -P gss_updatenode
     ```

8. Update OFED.
   - **Online upgrade only**

     ```
     updatenode CurrentIoServer -P gss_ofed
     ```

   - **Offline upgrade only**

     Run in parallel.

     Instead of *CurrentIoServer* use the xCAT group gss_ppc64. For example:

     ```
     updatenode gss_ppc64 -P gss_ofed
     ```

9. Update IP RAID FW in the I/O Server node that is being upgraded.
   - **Online upgrade only**

     ```
     updatenode CurrentIoServer -P gss_ipraid
     ```

   - **Offline upgrade only**

     Run in parallel.

     Instead of *CurrentIoServer* use the xCAT group gss_ppc64. For example:

     ```
     updatenode gss_ppc64 -P gss_ipraid
     ```

10. **Online upgrade only** Perform the system firmware update. For more information, see “Updating the system firmware” on page 109.

    Note: A system firmware update requires a node reboot.
11. If the firmware update is not needed, reboot the I/O server node.
    - **Online upgrade only**

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

    - **Offline upgrade only**

      Run in parallel.

      Instead of *CurrentIoServer* use the xCAT group gss_ppc64. For example:

      ```
      xdsh gss_ppc64 "systemctl reboot"
      ```

12. Update the SAS host adapter firmware on *CurrentIoServer*.

- **Online upgrade only**

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

  Here `CurrentIoServer` is an I/O server node and the command is run on the I/O server node.
- **Offline upgrade only**

  Run this command on both I/O server nodes.

13. Update the node configuration.
    - **Online upgrade only**

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

      This command is run from the EMS node.
    - **Offline upgrade only**

      Run the following command.

```
/opt/ibm/gss/tools/samples/gssupgrade.sh -s
```

14. On PPC64BE systems, run phy check and ensure that the phy mapping is OK.
    - **Online upgrade only**

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

    - **Offline upgrade only**

      Run the following command.

```
gssinstallcheck -N IONode1,IONode2 –phy-mapping
```

15. Start GPFS on the I/O server node.
    - **Online upgrade only**

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

      Once the GPFS 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.
    - **Offline upgrade only**

      Run the following command to start GPFS cluster wide.

```
mmstartup -a
```

      Moving recovery groups, or cluster or file system manager is not needed at this point.

    Important: If this is the last node in the building-block that is being updated, ensure that you balance the active recovery group ownership before finishing up. For more information, see this step. Each recovery group must have a unique active server (rg_gssio1-hs owned by gssio1-hs and rg_gssio2-hs owned by gssio2-hs).
16. Wait until the I/O server can be seen active from the management server node, using the following command.
    - **Online upgrade only**

```
mmgetstate -N CurrentIoServer-hs
```

      The management server must be already running for issuing this command.
    - **Offline upgrade only**

Run the following command to check cluster state in parallel.

```
mmgetstate -a
```

17. Convert to **mmvdisk**, if currently in the legacy mode.
    a. Check if there are any **mmvdisk** node classes.

    ```
    mmvdisk nodeclass list
    ```

    b. If there are none, convert the recovery groups to **mmvdisk** by issuing this command.

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

    Note: Wait for 5 minutes for daemons to recycle. The file system remains up.
18. Update **mmvdisk** node class best-practice settings.

    ```
    /usr/lpp/mmfs/bin/mmvdisk server configure --update --node-class {{ item }} --recycle one
    ```

    This command ensures that the **mmvdisk** node class(es) are in sync with the best-practices settings.
19. Run `gssinstallcheck`.
    - **Online upgrade only**

      Run this command from the management server node.

      ```
      gssinstallcheck -N IONode
      ```

    - **Offline upgrade only**

      Run the following command to perform final installation check in parallel.

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

20. **Online upgrade only** Repeat preceding steps for the peer I/O server node of the same building block.

    Offline upgrade only Skip this step.
21. **Online upgrade only** Repeat all steps in this section for each additional building block.

    Offline upgrade only Skip this step.

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 EMSNode,IONode1,IONode2 --srv-events
```

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

   Note:

   - On PPC64BE systems, investigate, manage, and close serviceable events from HMC.
   - On PPC64LE systems, ASMI can be used to investigate issues.
   - 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 “Known issues” on page 51.

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

Warning: Apart from restarting the GUI, the sensors, and the collector, these steps might not be necessary. Carefully consider if updating the **mmperfmon** configuration is required. It depends on how old is the release that you are upgrading from. Run **mmperfmon config show** before determining the actions to take next.

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
```

2. Manually enable any RHN plugin in the file `/etc/yum/pluginconf.d/rhnplugin.conf` and enable any repos in `/etc/yum.repos.d` that were disabled before starting the upgrade.

This completes the upgrade task of the ESS system. For information on action items to be done after installation, see “Post-installation or post-upgrade checklist” on page 22.

Chapter 5. Adding building blocks to an existing ESS cluster

A building block is made up of a pair of Power 8 servers and one or more storage enclosures. When adding building blocks in an ESS cluster, perform the following steps. You can add more than one building blocks at a time.

Attention: Be careful when running this procedure on a production system. Failure to follow the listed procedure might result in re-deployment of a running I/O server node, loss of quorum, and/or data loss. Be especially carefully before running **gssdeploy -d** or **gsscheckdisks**. It is always a good practice to back up the xCAT database prior to starting this procedure and for the customer to back up their data, if possible.

1. Power on and boot the new Power 8 servers within the building block(s). Do not power up the associated storage enclosures.

   For example, a GL3C building block is made up of two Power 8 (PPC64LE) servers and three 4U106 enclosures. At this step, power on only the servers but do not apply power to the enclosures.
2. Add the new building block to the `/etc/hosts` file.
3. Find the new building block serial numbers. The subnet and mask are typically 10.0.0.1/24 by default. However, check the customer FSP network.

   ```
   /var/tmp/gssdeploy -f subnet/mask
   ```
4. Find rack positions that are used for GUI.

   ```
   /var/tmp/gssdeploy -i
   ```
5. Update `gssdeploy.cfg` to make the following changes.
   - Change `DEPLOYMENT_TYPE` to `ADD_BB`.
   - Change `GSS_GROUP` to something other than `gss_ppc64` or `ces_ppc64`. For example, `new_bb`.
   - Add the new serial numbers and node names to the `gssdeploy.cfg` file.
6. Run **gssdeploy -o** to add the new building block to xCAT.

   ```
   /var/tmp/gssdeploy -o
   ```
7. Run **gssprecheck** on the new node group and address any issues.

   ```
   /opt/ibm/gss/tools/samples/gssprecheck -G new_bb --install --file /var/tmp/gssdeploy.cfg
   ```
8. Run **gssdeploy -d** to deploy the new building block.

   ```
   /var/tmp/gssdeploy -d
   ```
9. After about 30 minutes, run health checks on the new building block and address any issues.
   a. **gssstoragquickcheck -N** *NewIONode1,NewIONode2*
   b. **gssfindmissingdisks -N** *NewIONode1,NewIONode2*
   c. **xdsh** *NewIONode* **GSSENV=INSTALL gsscheckdisks --encl all --iotest a --write-enable**
10. Perform the steps in this procedure: “Set up the high-speed network” on page 31.
11. Run **gssinstallcheck** on the new group. Ignore any GPFS related settings for now.

    ```
    xdsh new_bb "/opt/ibm/gss/tools/bin/gssinstallcheck -N localhost" | xcoll -n
    ```
12. Add the new nodes to the cluster.

```
gssaddnode -N NewIONode1,NewIONode2 --suffix=-hs --accept-license --nodetype gss --cluster-
node low_speed_name_of_cluster_node
```

13. Create recovery groups.

```
mmvdisk server configure --nc new_node_class --recycle one ; mmvdisk rg create --rg
newrg1>,<newrg2 --nc new_node_class
```

14. Create vdisks and NSDs. Make sure that you match the block size and the RAID code.

```
mmvdisk vs define --vs new_vdisk_set --rg newrg1,newrg2 --code 8+2p --bs 8M --ss 100% --nsd-
usage dataAndMetadata --sp system
```

    This is an example command. For best practices, see **mmvdisk documentation** in *IBM Spectrum Scale RAID: Administration*.

15. Reboot both new I/O server nodes.

```
xdsh new_bb "systemctl reboot"
```

    When both nodes are back up, you can start GPFS and move on to the next step.

16. Add NSDs to an existing file system. For example:

```
mmvdisk fs add --file-system existing_filesystem --vdisk-set new_vdisk_set
```

17. Run restripe if needed.

```
mmrestripefs -b
```

18. Update the performance monitoring list by using the **mmchnode** command. For more information, see mmchnode command.
19. Put the new nodes back in the gss_ppc64 node group and delete the temporary group, and then comment out the GSS_GROUP line in gssdeploy.cfg.

    For example, add nodes to gss_ppc64 as follows.

```
chdef -t group gss_ppc64 -o gssio3,gssio4 -p
```

20. Update call home for new building blocks.

```
gsscallhomeconf -E ems1 -N EMSNode,IONode1,IONode2 --suffix=-hs --register=all
```

Note: Now that you have additional building blocks, ensure that file system metadata replication is enabled (-m 2). The new pair of NSDs should be in a new failure group. The maximum number of failure groups is 5.

Known issues

This topic describes known issues for ESS.

ESS 5.3.6.2 issues

The following table describes known issues in ESS 5.3.6.2 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 5.3.6.2

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: Both Affected nodes: All	**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`

Table 2. Known issues in ESS 5.3.6.2 (continued)

Issue	Environment affected	Description	Resolution or action
Running **essutils** over PuTTY might show horizontal lines as “qqq” and vertical lines as “xxx”.	ESS Install and Deployment Toolkit Type: Install or Upgrade Version: All Arch: Both Affected nodes: All	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.
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: Both 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.

Table 2. Known issues in ESS 5.3.6.2 (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: Both 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).
Canceling disk replacement through GUI leaves original disk in unusable state	GUI Type: Install or Upgrade Arch: Both 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**
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 5.3.6.2 (continued)

Issue	Environment affected	Description	Resolution or action
Infiniband with multiple fabric is not supported with **gssgennetworks**	Type: Install and Upgrade Arch: Both Version: All	In a multiple fabric network, the Infiniband Fabric ID might not be properly appended in the `verbsPorts` configuration statement during the cluster creation. Incorrect `verbsPort` setting might cause the outage of the IB network.	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. During cluster creation, the `verbsPorts` setting is applied and there is a probability that the IB network becomes unreachable, if multiple fabric are set up during the cluster deployment. 3. Ensure that the GPFS daemon is running and then run the **mmfsadm test verbs config | grep verbsPorts** command. `# mmfsadm test verbs config | grep verbsPorts` `mmfs verbsPorts: mlx5_0/1/4 mlx5_1/1/7` In this example, the adapter `mlx5_0, port 1` is connected to fabric 4 and the adapter `mlx5_1 port 1` is connected to fabric 7. Run the following command and ensure that `verbsPorts` is correctly configured to the GPFS cluster. `# mmlsconfig | grep verbsPorts` `verbsPorts mlx5_0/1 mlx5_1/1` Here, it can be seen that the fabric is not configured even though IB was configured with multiple fabric. This is a known issue. Now, modify the `verbsPorts` setting for each node or node class to take the subnet into account. `# verbsPorts="$(echo $(mmfsadm test verbs config` `| grep verbsPorts | awk '{ $1=""; $2=""; $3=""; print $0} '))"` `# echo $verbsPorts mlx5_0/1/4 mlx5_1/1/7` `# mmchconfig verbsPorts="$verbsPorts" -N gssio1` `mmchconfig: Command successfully completed` `mmchconfig: Propagating the cluster configuration data to` `all affected nodes.` `This is an asynchronous process.` Here, the node can be any GPFS node or node class. Thereafter, verify that the new, correct `verbsPorts` setting is listed in the output. `# mmlsconfig | grep verbsPorts` `verbsPorts mlx5_0/1/4 mlx5_1/1/7`

Table 2. Known issues in ESS 5.3.6.2 (continued)

Issue	Environment affected	Description	Resolution or action
A failed disk's state wouldn't be changed to drained or replace in new enclosures that are added by MES procedure and never be used for any file system.	Type: IBM Spectrum Scale RAID Arch: Little Endian Version: ESS 5.3.6.2 Affected Nodes: N/A	If a user runs **mmvdisk pdisk change --simulate-failing** to fail two pdisks in the new enclosure(s) that are added by using the MES procedure and never be used for any file system, the state of the second pdisk stays at `simulate-failing`. Then, GUI cannot detect that the second failed disk is replaceable, same as command line which fails because of the state of the disk.	Run the **mmchpdisk --diagnose** on the failing disk. Or, run **mmshutdown** or **mmstartup** on the I/O server node that serves the recovery group that the simulate-failing pdisk belongs to. **Important:** This workaround causes a failover.
Cable pulls might result in I/O hang and application failure.	Type: IBM Spectrum Scale RAID Arch: Both Version: ESS 5.3.6.2 Affected Nodes: I/O server nodes	If SAS cables are pulled, I/O might hang for an extended period of time during RG or path recovery which could lead to application failures.	Change `nsdRAIDEventLogShortTermDelay` to 30ms (The default is 3000ms): 1. Run **mmchconfig nsdRAIDEventLogShortTermDelay=30**. 2. Restart GPFS.
gssinstall_<arch> and **gssinstallcheck** report NOT_INST for a few GPFS group RPMs.	Type: Deployment Arch: Little Endian Version: ESS 5.3.6.2 Affected Nodes: ALL	By default, deployment does not install RPMs for file audit logging support.	This is the expected behavior. If the file audit logging feature is required, you can manually install these packages from the EMS GPFS repository. • `gpfs.kafka` • `gpfs.libkafka`

Table 2. Known issues in ESS 5.3.6.2 (continued)

Issue	Environment affected	Description	Resolution or action
When doing a disk revive operation, you might see it fail with a return code 5.	Type: RAS Arch: Both Version: ESS 5.3.6.2 Affected Nodes: N/A	If a pdisk fails (or simulated to fail), a user would perform a revive operation using **mmchpdisk**. This command fails with a return code 5. `Command: err 5: tschpdisk --recovery-group rg_essio52-ce --pdisk e2s024 --state ok 2019-02-25_10:11: 29.742-0600: Input/output error`	No workaround is required. Though the command shows an error and bad return code, the operation will work (pdisk revived) after a few minutes.
When running **gssinstallcheck**, the system profile setting is found `activated`, instead of the expected setting `scale`	Type: RAS Arch: Both Version: ESS 5.3.6.2 Affected Nodes: N/A	Users run **gssinstallcheck** to verify various settings on an ESS 5.3.6.2 deploy or upgrade. There is a rare situation where the output of the system profile check is incorrect. This is due to the tuned service running on the node hitting a problem. To verify that tuned has failed, run: `systemctl status tuned`	The workaround is to restart the tuned service on the failed node and re-running **gssinstallcheck** (or manually check using **tuned-adm verify**). To restart the service, run: `systemctl restart tuned`

Table 2. Known issues in ESS 5.3.6.2 (continued)

Issue	Environment affected	Description	Resolution or action
During enclosure firmware updates, you might hit an issue where GPFS might crash and the file system shows stale file handle.	Type: Upgrade Arch: Both Version: ESS 5.3.6.2 Affected Nodes: I/O	During Enclosure Firmware test (upgrade), you may see a GPFS crash with signal 11 (visible in `/var/adm/ras/mmfs.log.latest`).	The current workaround is to restart GPFS on the affected nodes.
Deleting loghome vdisk and recreating it without deleting the corresponding recovery group might lead to loss access or data.	Type: Recovery usage Arch: Both Version: All Affected Nodes: I/O	A loghome vdisk is created when a recovery group is created using the **gssgenclusterrgs** script. If you delete the loghome vdisk, the corresponding recovery group must be also deleted. If a loghome vdisk is recreated without deleting the corresponding recovery group, a rare race condition could arise which might not properly flush key metadata to the disk associated with the recovery group of the loghome. Thus, it might lead to internal metadata inconsistencies resulting in loss of access or loss of data.	Ensure that if a loghome vdisk is deleted, the corresponding recovery group is also deleted. When the recovery group is deleted, you can recreate the recovery group using **gssgenclusterrgs** and proceed as normal usage.
The **Pools** page of the ESS GUI might indicate `No capacity available.` for the given system.	Type: GUI Versions: All Arch: Both Affected nodes: All	In certain scenarios during a deployment or an upgrade, the ESS GUI might show that capacity is not available in the **Pools** page.	It is advised that you wait up to 24 hrs. for all GUI refresh tasks to update. If you still see the problem, you can obtain the correct capacity by using the command line options.

Table 2. Known issues in ESS 5.3.6.2 (continued)

Issue	Environment affected	Description	Resolution or action
gssgenclusterrgs might fail because **mmvdisk** node class already exists.	Type: Install Versions: Both Arch: All Affected nodes: I/O	When performing a new installation, the creation of the recovery groups might fail due to a timing conflict with the **mmvdisk** node class generation. This is when using the wrapper scripts which now, by default, use **mmvdisk**.	**New installations only (not for adding building blocks)** Check if the mmvdisk node class exists: `mmvdisk nc list` If so, do the following: 1. Unconfigure servers.`mmvdisk server unconfigure --node-class <class>` 2. Delete the node class.`mmvdisk nodeclass delete --node-class <class>` 3. Try **gssgenclusterrgs** again.
gssgenclusterrgs might fail due to longer than expected recovery group names.	Type: Install/Add Versions: All Arch: Both Affected nodes: I/O	The **gssgenclusterrgs** command might fail because **mmvdisk** does not support names over a certain length.	Use **mmvdisk** commands directly and create recovery groups with names of a shorter character length.
gssgenvdisks does not work properly if using **--vdisk-size** or **--use-da**	Type: Install/Add Versions: All Arch: Both Affected nodes: I/O	The **gssgenvdisks** command might not return the expected result in situations where the usage of a single pool or hybrid environment requires exact data sizes. For example, `--vdisk-size` when used with `--use-da` might not give the intended result.	Use **mmvdisk** commands directly or modify the vdisk stanza when precise vdisk sizes are required.
Duplicate PMRs might be generated for node call home	Versions: ESS 5.3.6.2 Arch: Little Endian Affected nodes: All	Duplicate PMRs might be generated unless a step is taken to avoid this behavior.	When doing the call home setup by using the **gsscallhomeconf** command, use the `--stop-auto-event-report` flag. Using this flag resolves this issue.

Table 2. Known issues in ESS 5.3.6.2 (continued)

Issue	Environment affected	Description	Resolution or action
MTM does not match the failing unit information reference in the PMR log file	Versions: ESS 5.3.6.2 Arch: Little Endian Affected nodes: All	The MTM does not match the failing unit reference in the PMR log file in cases such as a power cord pull and PS pull. For more information, see this example.	Note that the failing unit might be reported as the EMS node but the PMR details reference the failing node/part correctly..
During a rolling upgrade, **mmhealth** might show the error `local_exported_fs_unavail` even though the file system is still mounted.	Area: RAS Type: Upgrade Arch: Both Version: ESS 5.3.6.2	During an online rolling upgrade (Updating of one ESS I/O node at a time but maintaining quorum), **mmhealth** might display an error indicating that the local exported file system is unavailable. This message is erroneous. For more information, see this example.	Restart **mmsysmon** on each node called out by **mmhealth**.
Hardware details – EMS and I/O server nodes are not showing **Asset number**.	Type: Install / Add Version: All Arch: All Affected Nodes: EMS node.	Hardware details – EMS and I/O server nodes are not showing **Asset number**	Use the **mmlscomp** command to view this information.
xCAT related commands do not work on EMS node if security is enabled using the **gss_security** tool.	Type: Install / upgrade Version: 5.3.6.2 Arch: All Affected Nodes: EMS node	After running the **gss_security** tool to enable security, as part of the security enablement, the `httpd` daemon is shut down. Due to this, xCAT commands fail. For more information, see “Enabling security in ESS” on page 97.	Disable security on the EMS node by using **gss_security -d** command and retry xCAT commands.

Table 2. Known issues in ESS 5.3.6.2 (continued)

Issue	Environment affected	Description	Resolution or action
Failed to start IBM.ESAGENT subsystem due to wrong Java installed	Type: Upgrade Version: 5.3.x.x Arch: All Affected Nodes: All	Upon upgrade, the JAVA pointer might be incorrect and might cause issues when starting ESA.	To fix this, issue, run the following command: `yum reinstall java-1.8.0-openjdk-headless-1.8.0.222.b03-1.el7.ppc64le` To verify that the issue is fixed, run the following command: `which java; java -version` The expected output is: `/bin/java` ` openjdk version "1.8.0_222-ea"` ` OpenJDK Runtime Environ-ment (build 1.8.0_222-ea-b03)` ` OpenJDK 64-Bit Server VM (build 25.222-b03, mixed mode)` After these pets, restart ESA as follows: `systemctl restart esactl` Verify that the service has started as follows: `systemctl status esactl`
Email notification by using ESA does not work	Type: Install/ Upgrade Version: 5.3.x.x Arch: All Affected Nodes: All	An ESA patch is required for SNMP notifications to work as designed.	A patch is being tested at this time to check if it addresses the issue. Contact support to check if a fix is available.
POWER8 protocol nodes are not supported with Fiber Channel (or other storage adapters).	Type: Install Version: 5.3.x.x Arch: All Affected Nodes: All	ESS deployment currently does not support POWER8 protocol nodes that have storage adapters inserted. You might still use these nodes but ESS deployment installation and upgrade will not be supported.	

Table 2. Known issues in ESS 5.3.6.2 (continued)

Issue	Environment affected	Description	Resolution or action
Call home setup has a requirement that all nodes of PPC64LE architecture must be registered in ESA.	Type: Install Version: 5.3.x.x Arch: All Affected Nodes: All	There is currently a limitation wherein for call home to work correctly, all nodes in the cluster that are PPC64LE must be registered with ESA. If all nodes of that architecture are not registered, call home might not work as designed. This is usually the case when PPC64LE client or protocol nodes are in the same cluster as ESS nodes but they are not registered with ESA.	
Race condition in `opal-elog` that can hit a kernel panic in function `elog_work_fn`. This is experienced when the GUI is running HW_INVENTORY commands to POWER servers.	Type: Kernel panic Version: All Arch: PPC64LE Affected Nodes: EMS, I/O, and protocol nodes	This issue was found with RHEL 7 kernel (Bugzilla 1873189) while `opal-elog` is handling an excessive amount of OPAL error log events. The GUI runs **`ipmi fru print`** commands as part of its HW_INVENTORY checks. The bug might be hit during these intervals due to the excessive amount of OPAL events are being generated.	A fix is being worked on by Red Hat to provide a new kernel to address this race condition. There is a known issue with OPAL on Power nodes wherein too many OPAL requests might cause a system hang. This issue does not affect ESS 3000 nodes. In response, consider disabling the **HW_INVENTORY GUI** task to reduce requests to the FSP. `/usr/lpp/mmfs/gui/cli/chtask HW_INVENTORY --inactive`

ESS 5.3.6.1C issues

The following table describes known issues in ESS 5.3.6.1C and how to resolve these issues. Depending on which fix level you are installing, these might or might not apply to you.

Table 3. Known issues in ESS 5.3.6.1C

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: Both 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 **essutils** over PuTTY might show horizontal lines as “qqq” and vertical lines as “xxx”.	ESS Install and Deployment Toolkit Type: Install or Upgrade Version: All Arch: Both 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 3. Known issues in ESS 5.3.6.1C (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: Both 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.
Creating small file systems in the GUI (below 16G) will result in incorrect sizes	GUI Type: Install or Upgrade Arch: Both 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).

Table 3. Known issues in ESS 5.3.6.1C (continued)

Issue	Environment affected	Description	Resolution or action
Canceling disk replacement through GUI leaves original disk in unusable state	GUI Type: Install or Upgrade Arch: Both 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**
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 3. Known issues in ESS 5.3.6.1C (continued)

Issue	Environment affected	Description	Resolution or action
Infiniband with multiple fabric is not supported with **gssgennetworks**	Type: Install and Upgrade Arch: Both Version: All	In a multiple fabric network, the Infiniband Fabric ID might not be properly appended in the `verbsPorts` configuration statement during the cluster creation. Incorrect `verbsPort` setting might cause the outage of the IB network.	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. During cluster creation, the `verbsPorts` setting is applied and there is a probability that the IB network becomes unreachable, if multiple fabric are set up during the cluster deployment. 3. Ensure that the GPFS daemon is running and then run the **mmfsadm test verbs config | grep verbsPorts** command. `# mmfsadm test verbs config | grep verbsPorts` `mmfs verbsPorts: mlx5_0/1/4 mlx5_1/1/7` In this example, the adapter `mlx5_0, port 1` is connected to fabric 4 and the adapter `mlx5_1 port 1` is connected to fabric 7. Run the following command and ensure that `verbsPorts` is correctly configured to the GPFS cluster. `# mmlsconfig | grep verbsPorts` `verbsPorts mlx5_0/1 mlx5_1/1` Here, it can be seen that the fabric is not configured even though IB was configured with multiple fabric. This is a known issue. Now, modify the `verbsPorts` setting for each node or node class to take the subnet into account. `# verbsPorts="$(echo $(mmfsadm test verbs config` `| grep verbsPorts | awk '{ $1=""; $2=""; $3=""; print $0} '))"` `# echo $verbsPorts mlx5_0/1/4 mlx5_1/1/7` `# mmchconfig verbsPorts="$verbsPorts" -N gssio1` `mmchconfig: Command successfully completed` `mmchconfig: Propagating the cluster configuration data to` `all affected nodes.` `This is an asynchronous process.` Here, the node can be any GPFS node or node class. Thereafter, verify that the new, correct `verbsPorts` setting is listed in the output. `# mmlsconfig | grep verbsPorts` `verbsPorts mlx5_0/1/4 mlx5_1/1/7`

Table 3. Known issues in ESS 5.3.6.1C (continued)

Issue	Environment affected	Description	Resolution or action
A failed disk's state wouldn't be changed to drained or replace in new enclosures that are added by MES procedure and never be used for any file system.	Type: IBM Spectrum Scale RAID Arch: Little Endian Version: ESS 5.3.6.1C Affected Nodes: N/A	If a user runs **mmvdisk pdisk change --simulate-failing** to fail two pdisks in the new enclosure(s) that are added by using the MES procedure and never be used for any file system, the state of the second pdisk stays at `simulate-failing`. Then, GUI cannot detect that the second failed disk is replaceable, same as command line which fails because of the state of the disk.	Run the **mmchpdisk --diagnose** on the failing disk. Or, run **mmshutdown** or **mmstartup** on the I/O server node that serves the recovery group that the simulate-failing pdisk belongs to. **Important:** This workaround causes a failover.
Cable pulls might result in I/O hang and application failure.	Type: IBM Spectrum Scale RAID Arch: Both Version: ESS5.3.6.1C Affected Nodes: I/O server nodes	If SAS cables are pulled, I/O might hang for an extended period of time during RG or path recovery which could lead to application failures.	Change `nsdRAIDEventLogShortTermDelay` to 30ms (The default is 3000ms): 1. Run **mmchconfig nsdRAIDEventLogShortTermDelay=30**. 2. Restart GPFS.
gssinstall_<arch> and **gssinstallcheck** report NOT_INST for a few GPFS group RPMs.	Type: Deployment Arch: Little Endian Version: ESS 5.3.6.1C Affected Nodes: ALL	By default, deployment does not install RPMs for file audit logging support.	This is the expected behavior. If the file audit logging feature is required, you can manually install these packages from the EMS GPFS repository. • `gpfs.kafka` • `gpfs.libkafka`

Table 3. Known issues in ESS 5.3.6.1C (continued)

Issue	Environment affected	Description	Resolution or action
When doing a disk revive operation, you might see it fail with a return code 5.	Type: RAS Arch: Both Version: ESS 5.3.6.1C Affected Nodes: N/A	If a pdisk fails (or simulated to fail), a user would perform a revive operation using **mmchpdisk**. This command fails with a return code 5. `Command: err 5: tschpdisk --recovery-group rg_essio52-ce --pdisk e2s024 --state ok 2019-02-25_10:11: 29.742-0600: Input/output error`	No workaround is required. Though the command shows an error and bad return code, the operation will work (pdisk revived) after a few minutes.
When running **gssinstallcheck**, the system profile setting is found `activated`, instead of the expected setting `scale`	Type: RAS Arch: Both Version: ESS 5.3.6.1C Affected Nodes: N/A	Users run **gssinstallcheck** to verify various settings on an ESS 5.3.6.1C deploy or upgrade. There is a rare situation where the output of the system profile check is incorrect. This is due to the tuned service running on the node hitting a problem. To verify that tuned has failed, run: `systemctl status tuned`	The workaround is to restart the tuned service on the failed node and re-running **gssinstallcheck** (or manually check using **tuned-adm verify**). To restart the service, run: `systemctl restart tuned`

Table 3. Known issues in ESS 5.3.6.1C (continued)

Issue	Environment affected	Description	Resolution or action
During enclosure firmware updates, you might hit an issue where GPFS might crash and the file system shows stale file handle.	Type: Upgrade Arch: Both Version: ESS 5.3.6.1C Affected Nodes: I/O	During Enclosure Firmware test (upgrade), you may see a GPFS crash with signal 11 (visible in `/var/adm/ras/mmfs.log.latest`).	The current workaround is to restart GPFS on the affected nodes.
Deleting loghome vdisk and recreating it without deleting the corresponding recovery group might lead to loss access or data.	Type: Recovery usage Arch: Both Version: All Affected Nodes: I/O	A loghome vdisk is created when a recovery group is created using the **gssgenclusterrgs** script. If you delete the loghome vdisk, the corresponding recovery group must be also deleted. If a loghome vdisk is recreated without deleting the corresponding recovery group, a rare race condition could arise which might not properly flush key metadata to the disk associated with the recovery group of the loghome. Thus, it might lead to internal metadata inconsistencies resulting in loss of access or loss of data.	Ensure that if a loghome vdisk is deleted, the corresponding recovery group is also deleted. When the recovery group is deleted, you can recreate the recovery group using **gssgenclusterrgs** and proceed as normal usage.

Table 3. Known issues in ESS 5.3.6.1C (continued)

Issue	Environment affected	Description	Resolution or action
The `lvm2-monitor` service scans all external disks. Thus, it can cause long delays upon I/O server node boot.	Type: RAS Versions: All Arch: All Affected nodes: I/O	New to RHEL 7.6, the `lvm2-monitor` service runs upon boot and scans all disks including those within the externally attached enclosures. This results in significantly increased I/O node boot times.	Disable the device scan and monitoring by doing the following: 1. Add `global_filter = [ "r|.*/|" ]` in the `devices` section of the `/etc/lvm/lvm.conf` file. 2. Reboot the system. 3. Run **`systemctl start lvm2-monitor`**. The reboot time will be as normally expected.
The **Pools** page of the ESS GUI might indicate `No capacity available.` for the given system.	Type: GUI Versions: All Arch: All Affected nodes: N/A	In certain scenarios during a deployment or an upgrade, the ESS GUI might show that capacity is not available in the **Pools** page.	It is advised that you wait up to 24 hrs. for all GUI refresh tasks to update. If you still see the problem, you can obtain the correct capacity by using the command line options.
During a rolling upgrade, **`mmhealth`** might show the error `local_exported_fs_unavail` even though the file system is still mounted.	Area: RAS Type: Upgrade Arch: Both Version: ESS 5.3.6.1C	During an online rolling upgrade (Updating of one ESS I/O node at a time but maintaining quorum), **`mmhealth`** might display an error indicating that the local exported file system is unavailable. This message is erroneous. For more information, see this example.	Restart **`mmsysmon`** on each node called out by **`mmhealth`**.
`mmnetmetrics` problem log shows message: `cannot run`	Type: SVT GPFS mmfs.log.latest Version: ESS 5.3.6.1C	During the start of GPFS, the `mmfs.log.latest` file shows the error message: `/opt/IBM/gpfs/bin/mmnetmetrics cannot run with error 2: No such file or directory.`	This message does not affect any function or workloads. No workaround is required.

Table 3. Known issues in ESS 5.3.6.1C (continued)

Issue	Environment affected	Description	Resolution or action
systemd issue when just doing reboot	Type: SVT GPFS systemd Version: ESS 5.3.6.1C	After system reboot, GPFS might not be able to be started. **mmstartup** could give the error message: `Failed to start gpfs.service: Connection timed out.` This error is due to systemd not function properly.	This issue does not affect any function or workloads. No workaround is required. However, you can do the following: From the EMS node, power down and power up the system with **rpower off** and **rpower on** commands.
`ibm-crassd` validation sensors do not log power cord or any other monitoring	Type: SVT ibm-crassd Version: ESS 5.3.6.1C	The following issue was found in `ibm-crassd`: If the BMC system clock gets wrongly configured that is if it has an old date then the monitoring does not work. There are some bugs with the ToolTip output. If BMC clock gets scrambled, `ibm-crassd` validation sensors do not log power cord or any other monitoring. The `/opt/ibm/ras/lib/crassd.jar` file needs to be patched in this case.	Log in to BMC GUI and change the time zone to match your environment.

Example of MTM not matching failing unit information in PMR log

```
**************  FAILING UNIT INFORMATION **************
Service Agent Date, Time:  2019-11-14 17:19:46 UTC

UTC:                         2019-11-14 17:19:46 GMT
Machine Type/Feat:       5148       <------------- This data continues to be from the ems node
XXX
Model:                   21L        <------------- This data continues to be from the ems node
XXX
Serial:                  005789A    <------------- This data continues to be from the ems node
XXX
Unit Name:               essio41-fo
Sys Feature/Fnc 20:
```

```
Bundled Problem Report:
Indicator Mode (LP/GL):
Sys Attn/Info Act (Y/N):
```

Example output of mmhealth command

```
Component       Status        Status Change     Reasons
-------------------------------------------------------------------------------------------
GPFS            HEALTHY       6 min. ago        -
NETWORK         HEALTHY       20 min. ago       -
FILESYSTEM      DEGRADED      18 min. ago       local_exported_fs_unavail(gpfs1, gpfs0)
DISK            HEALTHY       6 min. ago        -
NATIVE_RAID     HEALTHY       6 min. ago        -
PERFMON         HEALTHY       19 min. ago       -
THRESHOLD       HEALTHY       20 min. ago       -
```

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.

Call home configuration

ESS 4.5 introduced ESS Management Server and I/O Server HW call home capability in ESS 5146 systems, where hardware events are monitored by the HMC managing these servers.

When a serviceable event occurs on one of the monitored servers, the Hardware Management Console (HMC) generates a call home event. This feature is only available in the 5146 systems as the 5146 systems are managed by the HMC. This feature is not available in 5148 systems as the 5148 systems are not managed by the HMC.

ESS 5.X provides additional call home capabilities for the drives in the attached enclosures of ESS 5146 and ESS 5148 systems. The call home for drive events does not require HMC, and uses the Electronic Service Agent (ESA) running on the EMS node.

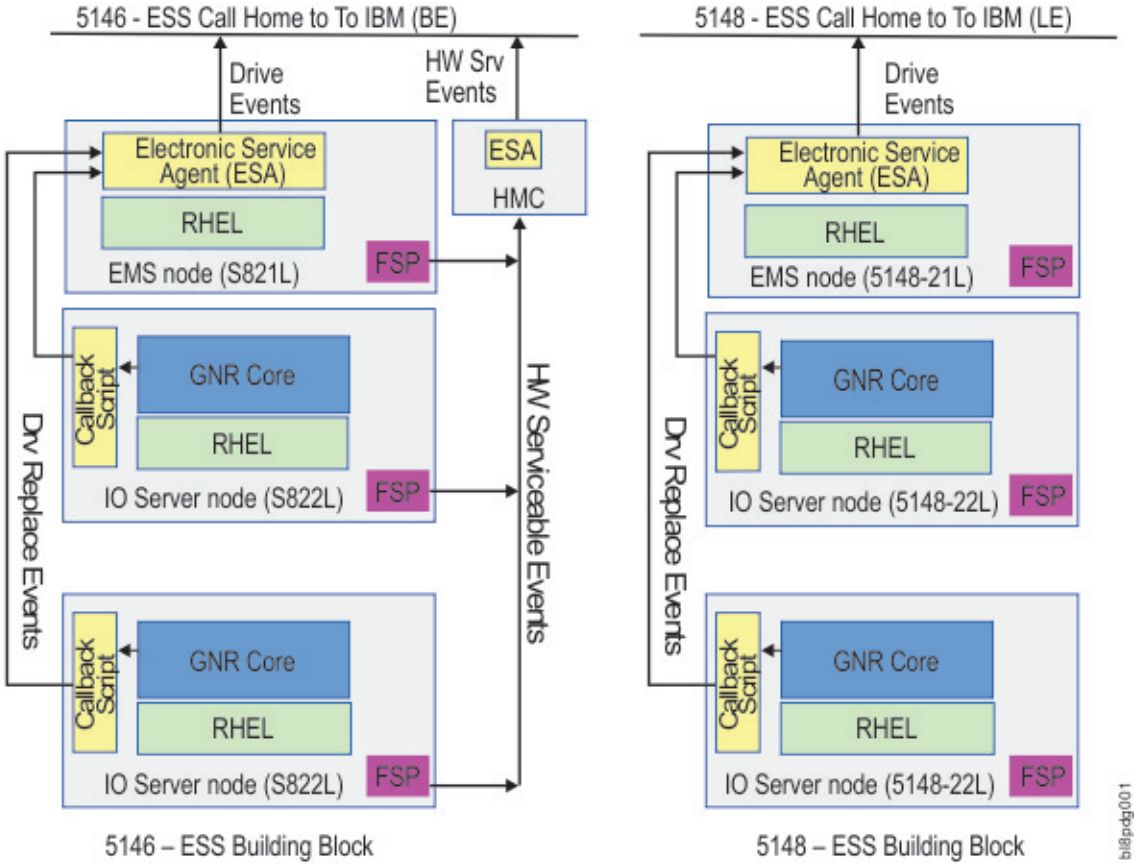

Figure 1. ESS Call Home block diagram

In ESS 5146 the HMC obtains the health status from the Flexible Service Process (FSP) of each server. When there is a serviceable event detected by the FSP, it is sent to the HMC, which initiates a call home event if needed. This function is not available in ESS 5148 systems.

ESS call home detects any issues with enclosures disks, issues with POWER systems such as DIMM failure, etc., and IBM Spectrum Scale software call home issues.

The callback script communicates with the ESA over a REST API. The ESA is installed in the ESS Management Server (EMS), and initiates a call home task. The ESA is responsible for automatically opening a Service Request (PMR) with IBM support, and managing the end-to-end life cycle of the problem.

IBM Electronic Service Agent (ESA) for PowerLinux version 4.1 and later can monitor the ESS systems. It is installed in the ESS Management Server (EMS) during the installation of ESS version 5.X, or when upgrading to ESS 5.X.

The IBM Electronic Service Agent is installed when the **gssinstall** command is run. The **gssinstall** command can be used in one of the following ways depending on the system:

- For 5146 system:

```
gssinstall_ppc64 -u
```

- For 5148 system:

```
gssinstall_ppc64le -u
```

The rpm files for the esagent are located in the `/install/gss/otherpkgs/rhels7/<arch>/gss` directory.

Issue the following command to verify that the rpm for the esagent is installed:

```
rpm -qa | grep esagent
```

This gives an output similar to the following:

```
esagent.pLinux-4.5.5-0.noarch
```

If ESA is not installed, issue the following command:

```
cd /install/gss/otherpkgs/rhels7/<arch>/gss
rpm -ihv --nodeps esagent.pLinux-4.5.5-0.noarch.rpm
```

After the ESA is installed, the ESA portal can be reached by going to the following link.

```
https://<EMS or ip>:5024/esa
```

For example:

```
https://192.168.45.20:5024/esa
```

The ESA uses port 5024 by default. It can be changed by using the ESA CLI if needed. For more information on ESA, see IBM Electronic Service Agent. On the Welcome page, log in to the IBM Electronic Service Agent GUI. If an untrusted site certificate warning is received, accept the certificate or click **Yes** to proceed to the IBM Electronic Service Agent GUI. You can get the context sensitive help by selecting the **Help** option located in the upper right corner.

After you have logged in, go to the **Main Activate ESA**, to run the activation wizard. The activation wizard requires valid contact, location and connectivity information.

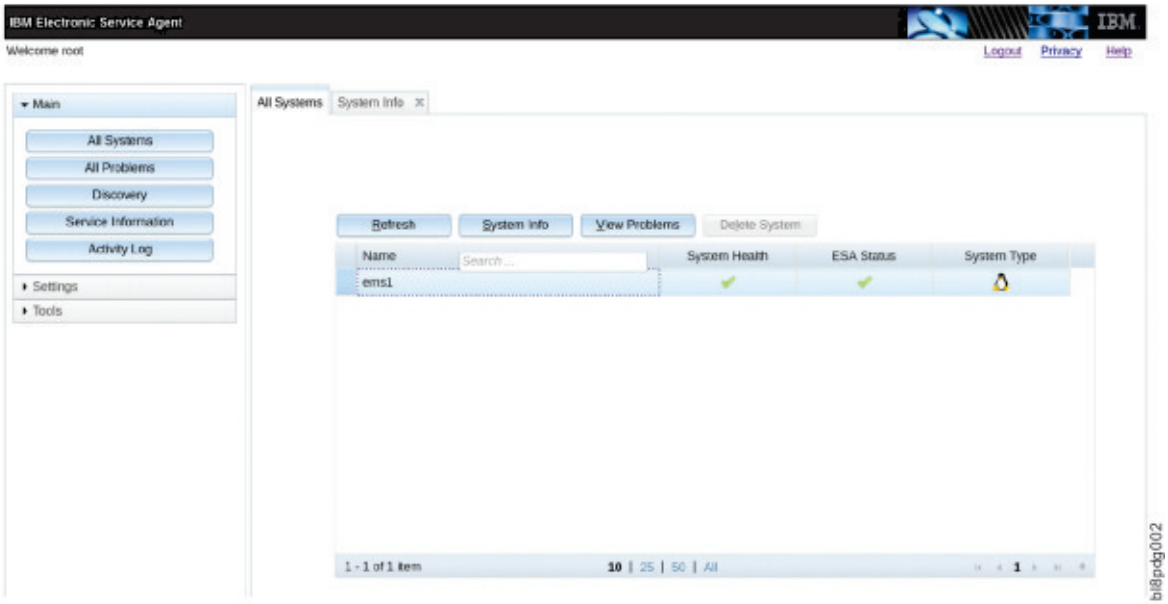

Figure 2. ESA portal after login

The `All Systems` menu option shows the node where ESA is installed. For example, ems1. The node where ESA is installed is shown as `PrimarySystem` in the **System Info**. The ESA Status is shown as **Online** only on the `PrimarySystem` node in the **System Info** tab.

Note: The ESA is not activated by default. In case it is not activated, you will get a message similar to the following:

```
[root@ems1 tmp]# gsscallhomeconf -E ems1 --show
IBM Electronic Service Agent (ESA) is not activated.
Activated ESA using /opt/ibm/esa/bin/activator -C and retry.
```

Entities or systems that can generate events are called endpoints. The EMS, I/O Server Canisters, and attached enclosures can be endpoints in ESS. Only enclosure endpoints can generate events, and the only event generated for call home is the disk replacement event. In the ESS 5146 systems, HMC can generate call home for certain node-related events.

In ESS, the ESA is only installed on the EMS, and automatically discovers the EMS as `PrimarySystem`. The EMS and I/O Server Canisters have to be registered to ESA as endpoints. The **gsscallhomeconf** command is used to perform the registration task. The command also registers enclosures attached to the I/O servers by default.

The software call home is registered based on the customer information given while configuring the ESA agent. A software call home group `auto` is configured by default and the EMS node acts as the software call home server. The weekly and daily software call home data collection configuration is also activated by default.

The software call home uses the ESA network connection settings to upload the data to IBM. The ESA agent network setup must be complete and working for the software call home to work.

Note: You cannot configure the software call home without configuring the ESA. For more information, see *Software call home* in *Elastic Storage Server: Problem Determination Guide*

```
# gsscallhomeconf -h
usage: gsscallhomeconf [-h] ([-N NODE-LIST] [--show] [--prefix PREFIX]
                       [--suffix SUFFIX] -E ESA-AGENT [--register {node,all}]
                       [--no-swcallhome] [--icn ICN] [--crvpd]
                       [--serial SOLN-SERIAL] [--model SOLN-MODEL] [--verbose]
                       [--esa-hostname-fqdn ESA_HOSTNAME_FQDN]
                       [--stop-auto-event-report]

optional arguments:
  -h, --help            show this help message and exit
```

```
  -N NODE-LIST          Provide a list of nodes to configure.
  --show                Show call home configuration details.
  --prefix PREFIX       Provide hostname prefix. Use = between --prefix and
                        value if the value starts with -.
  --suffix SUFFIX       Provide hostname suffix. Use = between --suffix and
                        value if the value starts with -.
  -E ESA-AGENT          Provide nodename for esa agent node
  --register {node,all}
                        Register endpoints(nodes, enclosure or all) with ESA.
  --no-swcallhome       Do not configure software callhome while configuring
                        hardware callhome
  --icn ICN             Provide IBM Customer Number for Software callhome.
  --crvpd               Create vpd file.
  --serial SOLN-SERIAL  Provide ESS solution serial number.
  --model SOLN-MODEL    Provide ESS model. Applicable only for BE (ppc64)
                        models.
  --verbose             Provide verbose output
  --esa-hostname-fqdn ESA_HOSTNAME_FQDN
                        Fully qualified domain name of ESA server for
                        certificate validation.
  --stop-auto-event-report
                        Stop report of automatic event to ESA in case of any
                        hardware call home event reported to system.

This program can be used to configure ESS for call home using ESA agent.
Example: esscallhomeconf -E ems1 -N ems1,gss_ppc64
```

A sample output is as follows:

```
# gsscallhomeconf -E ems1 -N ems1,gss_ppc64 --suffix=-ib –icn 123456

2017-02-07T21:46:27.952187 Generating node list...
2017-02-07T21:46:29.108213 nodelist:    ems1 essio11 essio12
2017-02-07T21:46:29.108243 suffix used for endpoint hostname: -ib
End point ems1-ib registered successfully with systemid 802cd01aa0d3fc5137f006b7c9d95c26
End point essio11-ib registered successfully with systemid c7dba51e109c92857dda7540c94830d3
End point essio12-ib registered successfully with systemid 898fb33e04f5ea12f2f5c7ec0f8516d4
End point enclosure G5CT018 registered successfully with systemid
c14e80c240d92d51b8daae1d41e90f57
End point enclosure G5CT016 registered successfully with systemid
524e48d68ad875ffbeeec5f3c07e1acf
ESA configuration for ESS Callhome is complete.

Started configuring software callhome
Checking for ESA is activated or not before continuing.
Fetching customer detail from ESA.
Customer detail has been successfully fetched from ESA.
Setting software callhome customer detail.
Successfully set the customer detail for software callhome.
Enabled daily schedule for software callhome.
Enabled weekly schedule for software callhome.
Direct connection will be used for software calhome.
Successfully set the direct connection settings for software callhome.
Enabled software callhome capability.
Creating callhome automatic group
Created auto group for software call home and enabled it.
Software callhome configuration completed.
```

The **gsscallhomeconf** command logs the progress and error messages in the `/var/log/messages` file. There is a **--verbose** option that provides more details of the progress, as well error messages. The following example displays the type of information sent to the `/var/log/messages` file in the EMS by the **gsscallhomeconf** command.

```
[root@ems1 vpd]# grep ems1 /var/log/messages | grep gsscallhomeconf

Feb 8 01:37:39 ems1 gsscallhomeconf: [I] End point ems1-ib registered successfully with
 systemid 802cd01aa0d3fc5137f006b7c9d95c26
Feb 8 01:37:40 ems1 gsscallhomeconf: [I] End point essio11-ib registered successfully
 with systemid c7dba51e109c92857dda7540c94830d3
Feb 8 01:37:41 ems1 gsscallhomeconf: [I] End point essio12-ib registered successfully
 with systemid 898fb33e04f5ea12f2f5c7ec0f8516d4
Feb 8 01:43:04 ems1 gsscallhomeconf: [I] ESA configuration for ESS Callhome is complete.
```

Attention: The **gsscallhomeconf** command also configures the IBM Spectrum Scale call home setup. The IBM Spectrum Scale call home feature collects files, logs, traces, and details of certain system health events from the I/O and EMS nodes and services running on those nodes. These details are shared with the IBM support center for monitoring and problem determination. For

more information on IBM Spectrum Scale call home, see the *Understanding call home* section in the IBM Knowledge Center.

The endpoints are visible in the ESA portal after registration, as shown in the following figure:

Figure 3. ESA portal after node registration

Detailed information about the node can be obtained by selecting **System Information**. Here is an example of the system information:

System Information

Property	Value
Name	essio12.isst.gpfs.ibm.net
Machine Type	8247
Machine Model	22L
Serial Number	2145B3A
Manufacturer	IBM
Operating System	Linux
OS Type	Linux
OS Version	3.10.0-327.36.3.el7.ppc64
OS Additional Version	
IP Address	192.168.1.103 192.168.2.103
Firmware	
PM Enabled	No
ESA Status	Offline
System ID	898fb33e04f5ea12f2f5c7ec0f8516d4

bi8pdg005

Figure 4. System information details

When an endpoint is successfully registered, the ESA assigns a unique system identification (system id) to the endpoint. The system id can be viewed using the `--show` option.
For example:

```
[root@ems1 vpd]# gsscallhomeconf -E ems1 --show
System id and system name from ESA agent

{"c14e80c240d92d51b8daae1d41e90f57": "G5CT018"

"c7dba51e109c92857dda7540c94830d3": "essio11-ib",
"898fb33e04f5ea12f2f5c7ec0f8516d4": "essio12-ib",
"802cd01aa0d3fc5137f006b7c9d95c26": "ems1-ib",
"524e48d68ad875ffbeeec5f3c07e1acf": "G5CT016"
}
```

When an event is generated by an endpoint, the node associated with the endpoint must provide the system id of the endpoint as part of the event. The ESA then assigns a unique event id for the event. The system id of the endpoints are stored in a file called `esaepinfo01.json` in the `/vpd` directory of the EMS and I/O servers that are registered. The following example displays a typical `esaepinfo01.json` file:

```
[root@ems1 vpd]# cat esaepinfo01.json
{
"encl": {
"G5CT016": "524e48d68ad875ffbeeec5f3c07e1acf",
"G5CT018": "c14e80c240d92d51b8daae1d41e90f57"},
"esaagent": "ems1", "node": {
"ems1-ib": "802cd01aa0d3fc5137f006b7c9d95c26",
"essio11-ib": "c7dba51e109c92857dda7540c94830d3",
 "essio12-ib": "898fb33e04f5ea12f2f5c7ec0f8516d4"
}
```

In the ESS 5146, the **gsscallhomeconf** command requires the ESS solution vpd file that contains the IBM Machine Type and Model (MTM) and serial number information to be present. The vpd file is used by the ESA in the call home event. If the vpd file is absent, the **gsscallhomeconf** command fails, and displays an error message that the vpd file is missing. In this case, you can rerun the command with the `--crvpd` option, and provide the serial number and model number using the `--serial` and `--model` options. In ESS 5148, the vpd file is auto generated if not present.

The system vpd information is stored in the `essvpd01.json` file in the EMS `/vpd` directory. Here is an example of a vpd file:

```
[root@ems1 vpd]# cat essvpd01.json
{
"groupname": "ESSHMC", "model": "GS2",
 "serial": "219G17G", "system": "ESS", "type": "5146"
}
[root@ems1 vpd]# cat essvpd01.json
{
"groupname": "ESSHMC", "model": "GS2",
 "serial": "219G17G", "system": "ESS", "type": "5146"
}
```

To check if the ESA rpms are installed, run the following command:

```
rpm -qa | grep esagent
```

To check if the ESA is configured and activated, run the following command:

```
gsscallhomeconf -E ems1 --show
```

For more information on ESA configuration and activation, see *Login and activation* in *Elastic Storage Server: Problem Determination Guide*. For information on network connectivity and end-to-end setup, see *Test call home* in *Elastic Storage Server: Problem Determination Guide*.

Attention: The EMS node must not have any other version of JDK installed than Open JDK 8. If the EMS node has any other JDK version installed such as Open JDK 11 or some other vendor provided JDK then the ESA daemon startup failure might occur. You must remove all other versions of JDK installed on EMS except `gpfs.java` and reinstall Open JDK 8 to get the ESA 455 working.

A callback is a one-time event. Therefore, it is triggered when the disk state changes to `replace`. If the ESA misses the event , for example if the EMS is down for maintenance, the call home event is not generated by the ESA.

To mitigate this situation, the `callhomemon.sh` script is provided in the `/opt/ibm/gss/tools/samples` directory of the EMS. This script checks for pdisks that are in the `replace` state, and sends an event to the ESA to generate a call home event if there is no open PMR for the corresponding physical drive. This script can be run on a periodic interval. For example, every 30 minutes.

In the EMS, create a cronjob as follows:

1. Open crontab editor using the following command:

   ```
   crontab -e
   ```

2. Setup a periodic cronjob by adding the following line:

   ```
   */30 * * * * /opt/ibm/gss/tools/samples/callhomemon.sh
   ```

3. View the cronjob using the following command:

```
crontab -l
 [root@ems1 deploy]# crontab -l
*/30 * * * * /opt/ibm/gss/tools/samples/callhomemon.sh
```

GSS command logging

All GSS commands are logged at `var/log/gss/5.3.6.x`.

The `var/log/gss/5.3.6.x` directory contains the following log files:

gsscommand.log
This log file contains all the sub-commands which are executed when any GSS command is run, its error code, and output of each individual command within a script.

gssinstallcheck.log
This log file contains complete user output (verbose) of each individual command within a script.

gssinstall.log
This log file contains all the information to debug any exception encountered during execution of a GSS command.

gssprecheck.*timestamp*
This log file contains the command execution debug data for the **gssprecheck** command.

ESS networking considerations

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), 25 GbE, 40 GbE, 100 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.

- **IBM Elastic Storage Server networking with Mellanox adapters**

  Mellanox ConnectX-2 adapter cards improve network performance by increasing available CPU bandwidth, which enhances performance in virtualized server environments. Mellanox ConnectX-2 adapter cards provide:

  – Data Center Bridging (DCB)
  – Fibre Channel over Ethernet (FCoE)
  – SR-IOV

  For information on using Mellanox adapter cards, see: http://www.mellanox.com/page/ethernet_cards_overview

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**).

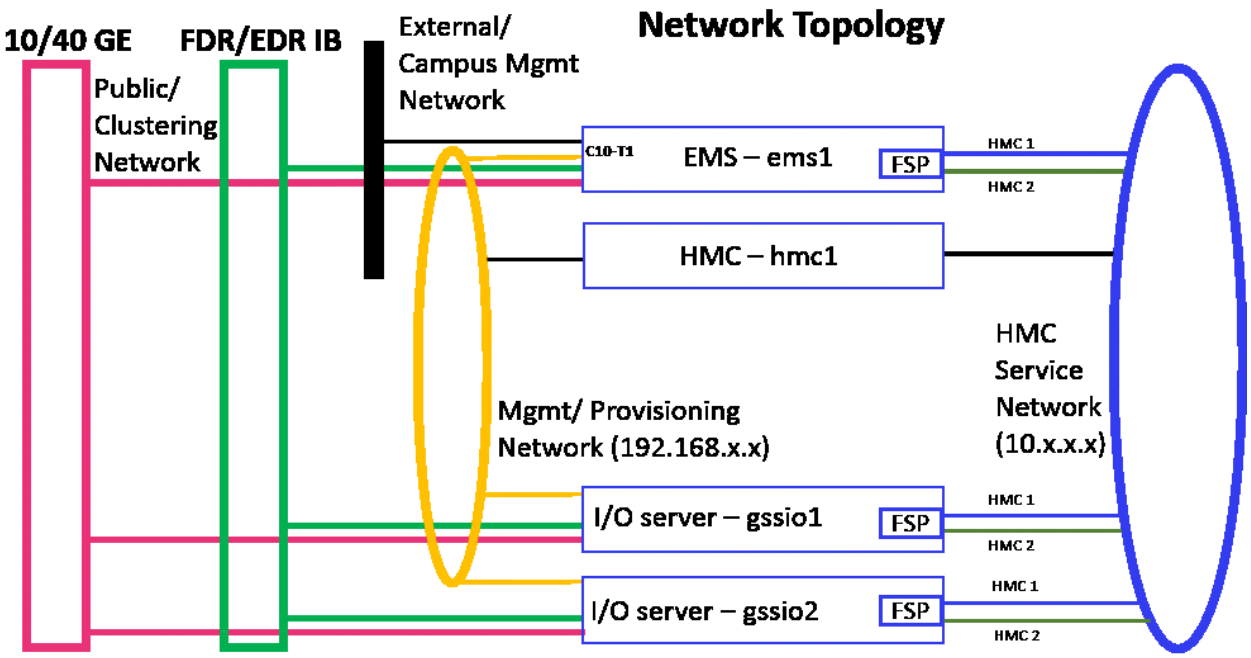

*Figure 5. The management and provisioning network and the service network: a logical view (on **PPC64BE**)*

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**).

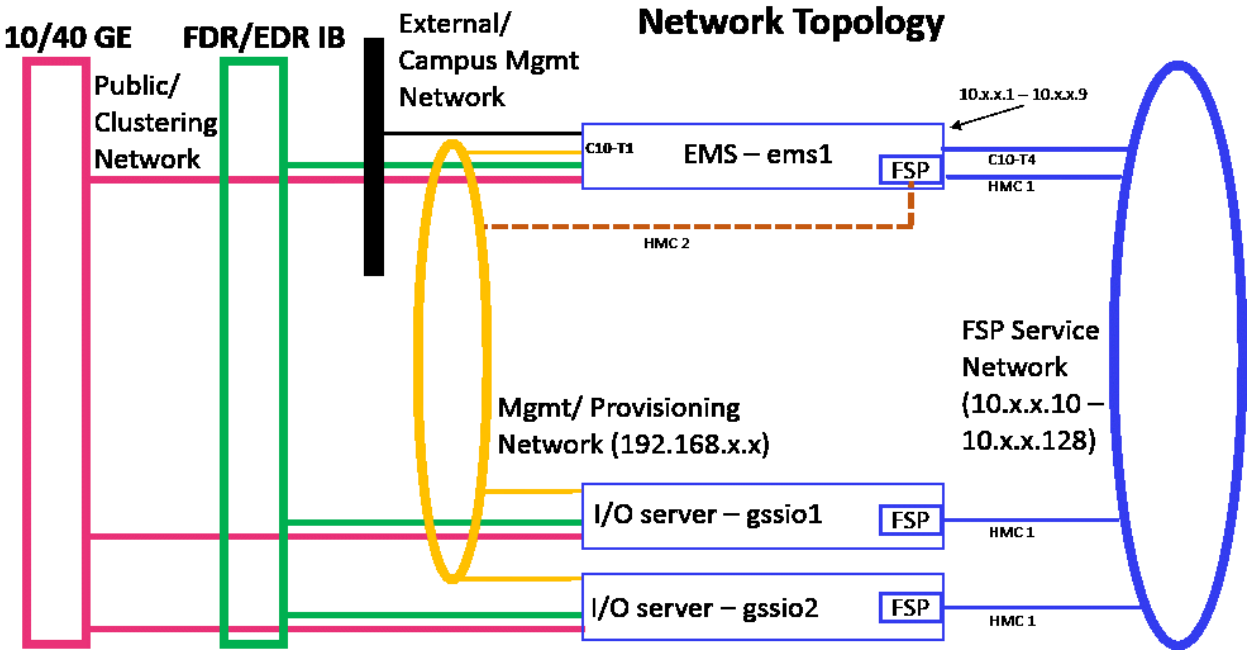

*Figure 6. The management and provisioning network and the service network: a logical view (on **PPC64LE**)*

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.

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.

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. Incorrect `verbsPorts` setting might cause the outage of the IB network. 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. During cluster creation, the `verbsPorts` setting is applied and there is a probability that the IB network becomes unreachable, if multiple fabrics are set up during the cluster deployment.
3. Ensure that the GPFS daemon is running and then run the **mmfsadm test verbs config | grep verbsPorts** command.

These steps show the Fabric ID found for each link.

For example:

```
# mmfsadm test verbs config | grep verbsPorts
mmfs verbsPorts: mlx5_0/1/4 mlx5_1/1/7
```

In this example, the adapter `mlx5_0, port 1` is connected to fabric 4 and the adapter `mlx5_1 port 1` is connected to fabric 7. Now, run the following command and ensure that `verbsPorts` settings are correctly configured to the GPFS cluster.

```
# mmlsconfig | grep verbsPorts
verbsPorts mlx5_0/1 mlx5_1/1
```

Here, it can be seen that the fabric has not been configured even though IB was configured with multiple fabric. This is a known issue.

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

```
[root@gssio1 ~]# verbsPorts="$(echo $(mmfsadm test verbs config | \
grep verbsPorts | awk '{ $1=""; $2=""; $3=""; print $0} '))"
# echo $verbsPorts
mlx5_0/1/4 mlx5_1/1/7
```

```
# mmchconfig verbsPorts="$verbsPorts" -N gssio1
mmchconfig: Command successfully completed
mmchconfig: Propagating the cluster configuration data to all
  affected nodes.  This is an asynchronous process.
```

Here, the node can be any GPFS node or node class. Once the `verbsPorts` setting is changed, make sure that the new, correct `verbsPorts` setting is listed in the output of the **mmlsconfig** command.

```
# mmlsconfig | grep verbsPorts
verbsPorts mlx5_0/1/4 mlx5_1/1/7
```

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.

Table 4. Network switch firmware

Type	IBM MTM	Melannox switch model	Description	Latest validated switch OS (June 2020)
IB - FDR	8828-F36 8828-F37	SX6036	36-port FDR switch	mlnxOS 3.6.8012
IB - EDR	8828-E36 8828-E37	SB7700	36-port EDR switch (`switchIB1`)	mlnxOS 3.9.0300
IB - EDR	8828-G36 8828-G37	SB7800	36-port EDR switch (`switchIB2`)	mlnxOS 3.9.0300
ETH - 1GbE	8831-S52	Edgecore AS4610	48-port 1G + 4-port 10G SFP+	cumulus-3.7.12a
ETH - 40GbE	8831-NF2	SX1710	36-port 40G switch	Onyx 3.6.8012
ETH - 10GbE	8831-S48	SX1410	48-port 10G + 12-port 40G	Onyx 3.6.8012
ETH - 100GbE	8831-00M	SN2700	32-port 40G/100G	Onyx 3.9.0300
ETH - 10/25/40/100	8831-25M	SN2410	48-port 10G/25G + 8-port 40G/100G	Onyx 3.9.0300

Security

The following topics describe how to enable security related settings in ESS.

Enabling SELinux in ESS

Enabling SELinux in an ESS environment is a two-step process and it can be enabled for EMS and I/O server nodes using the **gss_selinux** post script.

By default, any node in an ESS cluster has SELinux disabled. You can run the **gss_selinux** post script using the **updatenode** command. This script can be run after the deployment of EMS node or I/O server nodes is complete.

- Enable SELinux on the EMS node as follows.
  a) Run the **gss_selinux** post script on the EMS node.

```
# updatenode ems1 -V -P "gss_selinux"
```

  Note: Make sure that you reboot the node when the **gss_selinux** script completes.

  b) Reboot the node.

```
# systemctl reboot
```

  The node is rebooted and it comes up with SELinux in Permissive mode.

```
# sestatus
SELinux status:                 enabled
SELinuxfs mount:                /sys/fs/selinux
SELinux root directory:         /etc/selinux
Loaded policy name:             targeted
Current mode:                   permissive
Mode from config file:          permissive
Policy MLS status:              enabled
Policy deny_unknown status:     allowed
Max kernel policy version:      31
```

  c) Rerun the **gss_selinux** script with the `-e` switch to enforce SELinux.

```
# updatenode ems1 -V -P "gss_selinux -e"
```

  No reboot is required in this case.

```
# sestatus
SELinux status:                 enabled
SELinuxfs mount:                /sys/fs/selinux
SELinux root directory:         /etc/selinux
Loaded policy name:             targeted
Current mode:                   enforcing
Mode from config file:          enforcing
Policy MLS status:              enabled
Policy deny_unknown status:     allowed
Max kernel policy version:      31
```

  After SELinux is enabled, kernel logs any activity in the `/var/log/audit/audit.log` file.

- Enable SELinux on I/O server nodes as follows.

a) Run the **gss_selinux** post script on the I/O server nodes.

```
# updatenode gss_ppc64 -V -P "gss_selinux"
```

Note: Make sure that you reboot the node when the **gss_selinux** script completes.

b) Reboot the I/O server nodes.

```
# systemctl reboot
```

The node is rebooted and it comes up with SELinux in Permissive mode.

c) Rerun the **gss_selinux** script with the `-e` switch to enforce SELinux.

```
# updatenode gss_ppc64 -V -P "gss_selinux -e"
```

No reboot is required in this case.

After SELinux is enabled, kernel logs any activity in the `/var/log/audit/audit.log` file.

- Disable SELinux on ESS nodes as follows.
  - To disable SELinux on the EMS node, use the following command.

    ```
    # updatenode ems1 -V -P "gss_selinux -d"
    ```

    Reboot the node after the script completes. When the node comes up after reboots, SELinux is disabled.

    You can check the status as follows.

    ```
    # sestatus
    SELinux status:                 disabled
    ```

  - To disable SELinux on the I/O server nodes, use the following command.

    ```
    # updatenode gss_ppc64 -V -P "gss_selinux -d"
    ```

    Reboot the node after the script completes. When the node comes up after reboots, SELinux is disabled. Any I/O server node name such as IO3 or IO4 can also be used instead of the xCAT group name.

  Protocol node consideration: You can also use this procedure to enable SELinux on protocol nodes. Consult IBM Spectrum Scale NFS, SMB and object documentation for information on SELinux support.

  Additional information: Any mentioned security item is an optional feature and you can enable it on demand for an ESS cluster. Post scripts can be run using xCAT **updatenode** command after deployment of the node is done and before creating the GPFS cluster. In upgrade cases, any such post script must be run after stopping the GPFS cluster. Do not attempt to run any security script while GPFS cluster is up and running.

Working with sudo user in an ESS Environment

Enabling sudo requires a sudo-capable user (`gpfsadmin`) to be added to all nodes which are a part of or which are going to be a part of an ESS cluster. Sudo must be enabled for EMS, I/O server nodes, or protocol nodes using the **gss_sudo** post script.

Note: Sudo user across all GPFS nodes must have the same Linux group ID and user ID.

- “Protocol nodes” on page 93
- “Help text gss_sudo script” on page 93

Enabling sudo on Linux nodes

You can enable sudo configuration on a Linux node using the `-e` switch of the **gss_sudo** script.

```
# updatenode ems1 -V -P "gss_sudo -e"
```

This command creates the `gpfsadmin` Linux user and `gpfs` Linux group at node and performs all necessary sudoers set up. For detailed information, see the `/etc/sudoers.d/ess_sudoers` file.

User can now log in to the node server using the `gpfsadmin` user and they can perform GPFS administration tasks.

Make sure that the **gss_sudo** script is run on all GPFS nodes (EMS node, I/O server nodes, protocol nodes, and any client nodes) as part of the cluster to be completely compliant with the sudo requirement. Change the node name in the **updatenode** command accordingly. Enabling sudo also allows the `gpfsadmin` user to administer xCAT and the GPFS GUI on the EMS node.

The xCAT policy table contains a new entry for the sudo user as shown in this example:

```
# tabdump policy
#priority,name,host,commands,noderange,parameters,time,rule,comments,disable
"1","root",,,,,,"allow",,
...
...
...
"1107","gpfsadmin",,,,,,"allow",,   (A new policy has been added to xCAT policy table)
```

Disabling sudo on Linux nodes

You can disable sudo configuration on a Linux node using the `-d` switch of the **gss_sudo** script.

```
# updatenode ems1 -V -P "gss_sudo -d"
```

Disabling sudo reverts the xCAT policy table to its previous state, deletes `/etc/sudoers.d/ess_sudoers` file, and deletes the `gpfsadmin` user from the Linux node. Make sure that you have disabled sudo user configuration on all GPFS nodes (EMS node, I/O server nodes, protocol nodes, and any client nodes) as part of the cluster to be completely compliant with the sudo requirement. Change the node name in the **updatenode** command accordingly.

The xCAT policy table now does not contain the entry for the sudo user as shown in this example:

```
# tabdump policy
#priority,name,host,commands,noderange,parameters,time,rule,comments,disable
"1","root",,,,,,"allow",,
...
...
...
```

Important: You must not disable sudo user until the GPFS cluster is set to configure not to use sudo wrapper and sudo user. Failing to do so might result in cluster corruption.

Enabling sudo with GPFS cluster

Once the sudo feature is enabled, make sure that you use `--use-sudo-wrapper` and `--sudo-user` option while creating a new GPFS cluster using **gssgencluster**. For more information, see **gssgencluster** command. If there is an existing cluster available, it must be converted to use sudo wrapper and sudo user by using the **gss_sudo -u** option. For more information, see gss_sudo command help text.

For example, consider a cluster which is created earlier and it is not using sudo wrapper and sudo user.

```
# mmlscluster

GPFS cluster information
========================
  GPFS cluster name:         scalecluster.gpfs.net
  GPFS cluster id:           15270568330550226974
  GPFS UID domain:           scalecluster.gpfs.net
  Remote shell command:      /usr/bin/ssh (No SUDO Wrapper used here)
  Remote file copy command:  /usr/bin/scp (No SUDO Wrapper used here)
  Repository type:           CCR

 Node  Daemon node name   IP address     Admin node name      Designation
-------------------------------------------------------------------------------
   1   io3-10g.gpfs.net   198.51.100.14  io3-10g.gpfs.net     quorum-manager
   2   io4-10g.gpfs.net   198.51.100.15  io4-10g.gpfs.net     quorum-manager
   3   ems1-10g.gpfs.net  198.51.100.13  ems2-10g.gpfs.net    quorum
```

You can configure the cluster to use sudo by issuing the following command.

```
# updatenode ems1 -V -P "gss_sudo -u"
```

```
# mmlscluster

GPFS cluster information
========================
  GPFS cluster name:         scalecluster.gpfs.net
  GPFS cluster id:           15270568330550226974
  GPFS UID domain:           scalecluster.gpfs.net
  Remote shell command:      sudo wrapper in use (SUDO wrapper now in use)
  Remote file copy command:  sudo wrapper in use (SUDO wrapper now in use)
  Repository type:           CCR

 Node  Daemon node name   IP address     Admin node name    Designation
-------------------------------------------------------------------------
   1   io3-10g.gpfs.net   198.51.100.14  io3-10g.gpfs.net   quorum-manager
   2   io4-10g.gpfs.net   198.51.100.15  io4-10g.gpfs.net   quorum-manager
   3   ems2-10g.gpfs.net  198.51.100.13  ems2-10g.gpfs.net  quorum
```

In the preceding **mmlscluster** command output, remote shell and remote copy commands are changed to use sudo wrapper (**sshwrap** and **scpwrap**).

The sudoUser **mmlsconfig** parameter is now set to gpfsadmin.

```
# mmlsconfig sudoUser
sudoUser gpfsadmin
```

Important:

- The **gss_sudo** script switches -u, -n and -s must not be used for nodes other than the EMS node.
- The IBM Spectrum Scale GUI services must be restarted by using **systemctl restart gpfsgui** after enabling or disabling sudo in a GPFS cluster.
- The sudo user password must be set to a new password before using it.

Disabling sudo with GPFS cluster

You can unconfigure a sudo-enabled GPFS cluster to not use sudo wrapper by using the -n switch of the **gss_sudo** script.

For example, consider a cluster which is created earlier and it is using sudo wrapper and sudo user.

```
# mmlscluster

GPFS cluster information
========================
  GPFS cluster name:         scalecluster.gpfs.net
  GPFS cluster id:           15270568330550226974
  GPFS UID domain:           scalecluster.gpfs.net
  Remote shell command:      sudo wrapper in use (SUDO wrapper now in use)
  Remote file copy command:  sudo wrapper in use (SUDO wrapper now in use)
  Repository type:           CCR

 Node  Daemon node name    IP address     Admin node name   Designation
```

```
--------------------------------------------------------------------------------
   1   io3-10g.gpfs.net   198.51.100.14  io3-10g.gpfs.net   quorum-manager
   2   io4-10g.gpfs.net   198.51.100.15  io4-10g.gpfs.net   quorum-manager
   3   ems2-10g.gpfs.net  198.51.100.13  ems2-10g.gpfs.net  quorum
```

You can configure the cluster to not to use sudo by issuing the following command.

```
# updatenode ems1 -V -P "gss_sudo -n"
```

```
# mmlscluster

GPFS cluster information
========================
  GPFS cluster name:         scalecluster.gpfs.net
  GPFS cluster id:           15270568330550226974
  GPFS UID domain:           scalecluster.gpfs.net
  Remote shell command:      /usr/bin/ssh (No SUDO Wrapper used here)
  Remote file copy command:  /usr/bin/scp (No SUDO Wrapper used here)
  Repository type:           CCR

 Node  Daemon node name      IP address      Admin node name   Designation
--------------------------------------------------------------------------------
   1   io3-10g.gpfs.net     198.51.100.14   io3-10g.gpfs.net   quorum-manager
   2   io4-10g.gpfs.net     198.51.100.15   io4-10g.gpfs.net   quorum-manager
   3   ems1-10g.gpfs.net    198.51.100.13   ems2-10g.gpfs.net  quorum
```

In the preceding **mmlscluster** command output, remote shell and remote copy commands are changed to use ssh and scp instead of sudo wrapper (**sshwrap** and **scpwrap**).

The sudoUser **mmlsconfig** parameter is now set to undefined.

```
# mmlsconfig sudoUser
sudoUser (undefined)
```

Important:

- The **gss_sudo** script switches -u, -n and -s must not be used for nodes other than the EMS node.
- The IBM Spectrum Scale GUI services must be restarted by using **systemctl restart gpfsgui** after enabling or disabling sudo in a GPFS cluster.

I/O server nodes

I/O server nodes must also have sudo user gpfsadmin configured if the ESS cluster is going to be managed with a sudo user.

```
# updatenode gss_ppc64 -V -P "gss_sudo -e"
```

Important: The **gss_sudo** script switches -u, -n and -s must not be used for nodes other than the EMS node.

Protocol nodes

Protocol nodes must also have sudo user gpfsadmin configured if the ESS cluster is going to be managed with a sudo user.

```
# updatenode ces_ppc64 -V -P "gss_sudo -e"
```

Important: The **gss_sudo** script switches -u, -n and -s must not be used for nodes other than the EMS node.

Help text gss_sudo script

```
# ./gss_sudo --help

usage: gss_sudo [-e] [-d] [-u] [-s SUDO_USER ] [-g SUDO_USER_GROUP] [-n] [-h]
```

```
GSS Enable and Disable SUDO script.

optional arguments:
  -e | --enable
    Configure SUDO user (default: gpfsadmin) on Node and do necessary configuration.
    It will create a user with password 'cluster' but password will be expired
    as soon as user gets created. Please set new password for user before using it.


  -d | --disable
    Disable SUDO user (default: gpfsadmin) and deleting it.
    Use root if you disabled SUDO using -d option.
    GPFS cluster will continue to use Remote shell command as sshwrap and
    Remote file copy command as scpwrap. Use -n option if you get away
    from sshwrap and scpwrap


  -u | --use-sudo-wrapper
    Convert existing GPFS cluster to use GPFS SUDO feature.
    Doing this will move cluster Remote shell command from /usr/bin/ssh to
    sshwrap and Remote file copy command from /usr/bin/scp to scpwrap.

      [gpfsadmin@ems2 ~]$ sudo /usr/lpp/mmfs/bin/mmlscluster

      GPFS cluster information
      ========================
        GPFS cluster name:         scalecluster_sudo.gpfs.net
        GPFS cluster id:           15270568330549627553
        GPFS UID domain:           scalecluster_sudo.gpfs.net
        Remote shell command:      sudo wrapper in use
        Remote file copy command:  sudo wrapper in use
        Repository type:           CCR

       Node  Daemon node name      IP address     Admin node name   Designation
      ------------------------------------------------------------------------------
         1   io3-10g.gpfs.net     198.51.100.14  io3-10g.gpfs.net   quorum-manager
         2   io4-10g.gpfs.net     198.51.100.15  io4-10g.gpfs.net   quorum-manager
         3   ems2-10g.gpfs.net    198.51.100.13  ems2-10g.gpfs.net  quorum

  -s | --sudo-user [SUDO_USER]
    SUDO user name which should be used for GPFS administration. By default
    gpfsadmin user will be configured by gss_sudo user. Use gpfsadmin user
    or use other SUDO user.

  -g | --sudo-user-group [SUDO_USER_GROUP]
    SUDO group name for SUDO user. By default it will be 'gpfs'.

  -n | --nouse-sudo-wrapper
    Convert existing GPFS cluster *NOT* to use GPFS SUDO wrapper.
    Doing this will move cluster from Remote shell command from scpwrap to
    /usr/bin/ssh and Remote file copy command from scpwrap to /usr/bin/scp

      [gpfsadmin@ems2 ~]$ sudo /usr/lpp/mmfs/bin/mmlscluster

      GPFS cluster information
      ========================
        GPFS cluster name:         scalecluster_sudo.gpfs.net
        GPFS cluster id:           15270568330549627553
        GPFS UID domain:           scalecluster_sudo.gpfs.net
        Remote shell command:      /usr/bin/ssh
        Remote file copy command:  /usr/bin/scp
        Repository type:           CCR

       Node  Daemon node name    IP address     Admin node name    Designation
      ------------------------------------------------------------------------------
       1     io3-10g.gpfs.net   198.51.100.14   io3-10g.gpfs.net   quorum-manager
       2     io4-10g.gpfs.net   198.51.100.15   io4-10g.gpfs.net   quorum-manager
       3     ems2-10g.gpfs.net  198.51.100.13   ems2-10g.gpfs.net  quorum


  -h | --help
    Print help usages.
```

Using the central administration mode in an ESS environment

Enabling the central administration mode, by setting `adminMode` attribute to `central`, prevents unwanted passwordless SSH access from any non-admin GPFS nodes to any another GPFS node. In case of ESS, it is assumed that the EMS node is the only node which acts as an admin mode.
For more information, see *adminMode configuration attribute* in *IBM Spectrum Scale: Administration Guide*.

Running the **gss_admincentral** script along with **updatenode -k** configures `adminMode=central` in an ESS cluster. By default, passwordless SSH setup between all nodes is enabled.

Only the EMS node is allowed to do passwordless SSH to all other GPFS nodes. However, other nodes such as the I/O server nodes, protocol nodes, and client nodes cannot do SSH back to the EMS or other GPFS nodes once `adminMode` is set to `central` and the node security context is updated.

- "Enabling the central administration mode" on page 95
- "Disabling the central administration mode" on page 95
- "Help text gss_admincentral script" on page 96

Enabling the central administration mode

Enabling the central administration mode is a two-step procedure.

1. Run the **gss_admincentral** script with -e option.

   ```
   # updatenode ems1 -V -P "gss_admincentral -e"
   ```

   Note: After running this command any future deployment of new nodes only have the `adminMode` attribute set to `central`, by default. For existing nodes in the cluster, you must update the xCAT security context by running the following command.

2. Update the xCAT security context using the **updatenode Node -k** script.

   ```
   # updatenode gss_ppc64,ces_ppc64 -V -k
   ...
   Password: <Type EMS node root Password here>
   ...
   ...
   ```

   Note:

   - If you do not run the **updatenode Node -k** command, the central administration mode gets enabled for any new nodes deployed using the current EMS node. However, existing nodes can still do passwordless SSH between each other.
   - In case of an upgrade, if you want to enable the central administration mode then run the same commands.
   - Make sure that you do not run **updatenode admin_node -V -k** on the EMS node which is the admin node.
   - Running **gss_admincentral** script against non-EMS nodes is not allowed.

The **gss_admincentral** script can be run after the deployment of the EMS node or I/O server nodes or protocol nodes is completed.

Disabling the central administration mode

Disabling the central administration mode is a two-step procedure.

1. Run the **gss_admincentral** script with -d option.

   ```
   # updatenode ems1 -V -P "gss_admincentral -d"
   ```

Note: After running this command any future deployment of new nodes only have the central administration mode disabled. For existing nodes in the cluster, you must update the xCAT security context by running the following command.

2. Update the xCAT security context using the **updatenode Node -k** script.

```
# updatenode gss_ppc64,ces_ppc64 -V -k
...
Password: <Type EMS node root Password here>
...
...
```

Note:

- If you do not run the **updatenode Node -k** command, the central administration mode gets disabled for any new nodes deployed using the current EMS node. However, existing nodes cannot do passwordless SSH between each other.
- In case of an upgrade, if you want to disable the central administration mode then run the same commands.
- Make sure that you do not run **updatenode admin_node -V -k** on the EMS node which is the admin node.
- Running **gss_admincentral** script against non-EMS nodes is not allowed.

Help text gss_admincentral script

```
# ./gss_admincentral -h

  usage: gss_admincentral [-e] [-d] [-h]

  Enable xCAT to setup passwordless SSH between all nodes depending on enabling
  and disabling switch.

  optional arguments:
    -e | --enable
      Disable xCAT both way passwordless for any future node deployment using
      EMS node. Login from newly deployed node to xcat management server will
      be prohibited.

      Always run gss_admincentral against EMS node only. Issuing gss_admincentral
      command from nodes other than EMS node is not allowed.

      In order to enable admin central (i.e disabling  passwordless SSH from
      existing non-admin node to admin node) please run
      "updatenode non_admin_node -V -k" on xCAT management node. Doing so will
      automatically delete private key on non-admin node.

      Make sure you must not run "updatenode admin_node -V -k" against EMS node
      i.e. admin node.

    -d | --disable
      Enable xCAT both way passwordless SSH for any future node deployment using
      EMS node. Login from newly deployed node to xcat management server will
      be allowed.

      Always run gss_admincentral against EMS node only. Issuing gss_admincental
      command from nodes other than EMS node is not allowed.

      In order to disable admin central (i.e re-enabling passwordless SSH from
      existing non-admin node to admin node) please run
      "updatenode non_admin_node -V -k" on xCAT management node. Doing so will
      automatically copy private key from admin node to non-admin node.

      Make sure you must not run "updatenode admin_node -V -k" against EMS node
      i.e. admin node.

    -h | --help
      Print help usages.
```

Enabling firewall in ESS

Enabling firewall in an ESS environment is a one-step process and it can be enabled for EMS, I/O server nodes, and protocol nodes using the **gss_firewall** post script.

By default, any node in an ESS cluster has firewall disabled. You can run the **gss_firewall** post script using the **updatenode** command. This script can be run after the deployment of EMS node or I/O server nodes is complete.

- Enable firewall on the EMS node by running the **gss_firewall** post script.

  ```
  # updatenode ems1 -V -P "gss_firewall -e"
  ```

  You can check the status of the firewall as follows.

  ```
  # firewall-cmd --state
  running
  ```

  You can verify the open firewall ports, by running **gss_firewall** with the -c switch. When the script completes, the required ports in firewall are verified.

  ```
  # updatenode ems1 -V -P "gss_firewall -c"
  ```

- Enable firewall on I/O server nodes by running the **gss_firewall** post script.

  ```
  # updatenode gss_ppc64 -V -P "gss_firewall -e"
  ```

  You can check the status of the firewall as follows.

  ```
  # firewall-cmd --state
  running
  ```

  You can verify the open firewall ports, by running **gss_firewall** with the -c switch. When the script completes, the required ports in firewall are verified.

  ```
  # updatenode gss_ppc64 -V -P "gss_firewall -c"
  ```

- Disable firewall on the EMS node by running the **gss_firewall** post script.

  ```
  # updatenode EMS1 -V -P "gss_firewall -d"
  ```

- Disable firewall on I/O server nodes by running the **gss_firewall** post script.

  ```
  # updatenode gss_ppc64 -V -P "gss_firewall -d"
  ```

  Protocol node consideration: You can also use these steps to enable firewall on protocol nodes.

Enabling security in ESS

Enabling security in an ESS environment is a one-step process and it can be enabled for EMS, I/O server nodes, and protocol nodes by using the **gss_security** script.

On the EMS node, the script is located in the `/opt/ibm/gss/xcat/postscripts/` directory. On the I/O server nodes and the protocol nodes, the script is located in `/xcatpost/`.

By default, any node in an ESS environment has security disabled. This script can be run after the deployment of the EMS node or the I/O server nodes is complete.

Note: If security is enabled on the EMS node, xCAT commands fail. Hence, before installation, configuration, or upgrade, disable security on the EMS node.

By default, any node in an ESS environment has security disabled. When you enable security on the node, the following changes occur:

- OS hardening is enabled by disabling TCP timestamps and ICMP protocol in network packets on the node.
- The HTTPd server is disabled from running on the node.

  Note: All services that are using the HTTPd server, including xCAT, might be affected when HTTPd is disabled.
- Strong ciphers, Macs, and KexAlgorithms are enabled on the node.
- SSH timeout is set to 300 seconds (5 minutes).

Note: To enable, disable or verify security on any node, the script must be run from that node locally.

- Enable security on the node as follows.

  ```
  # ./gss_security -e
  ```

  A sample output is as follows:

  ```
  gss_security [INFO]: Enabling Security...
  gss_security [INFO]: Security is enabled.
  ```

- Disable security on the node as follows.

  ```
  # ./gss_security -d
  ```

  A sample output is as follows:

  ```
  gss_security [INFO]: Disabling Security...
  gss_security [INFO]: Security is disabled.
  ```

- Check the status of the security settings as follows.

  ```
  # ./gss_security -c
  ```

  A sample output is as follows:

  ```
  gss_security [INFO]: SSH timout is enabled
  gss_security [INFO]: ICMP timestamp is disabled
  gss_security [INFO]: TCP timestamp is disabled
  gss_security [INFO]: httpd is disabled
  gss_security [INFO]: Strong Ciphers, MACs and KexAlgorithms are enabled.
  ```

Protocol node consideration: You can also use this procedure to enable security on protocol nodes.

Support for hybrid enclosures

ESS supports hybrid enclosures that comprise four or two enclosures containing only hard disk drives (HDDs) and one or two enclosures of only solid state drives (SSDs).

These are the hybrid enclosure models that are supported. The support for GH14S and GH24S is added in ESS 5.3.1. The support for GH12S is added in ESS 5.3.1.1. The support for GH22S is added in ESS 5.3.3.

GH14S
1 2U24 (SSD) and 4 5U84 (HDD) enclosures

GH14S recovery groups (RGs)

Each of the two GH14S RGs have a 12-disk SSD user declustered array (DA) and a 167-disk HDD user DA.

GH24S
2 2U24 (SSD) and 4 5U84 (HDD) enclosures

GH24S recovery groups

Each of the two GH24S RGs have a 24-disk SSD user DA and a 167-disk HDD user DA.

GH12S
1 2U24 (SSD) and 2 5U84 (HDD) enclosures

GH12S recovery groups

Each of the two GH12S RGs have a 12-disk SSD user DA and an 83-disk HDD user DA.

GH22S
2 2U24 (SSD) and 5 5U84 (HDD) enclosures

GH22S recovery groups

Each of the two GH22S RGs have a 24-disk SSD user DA and an 83-disk HDD user DA.

For information about 2U24 and 5U84 enclosures, see IBM ESS Expansion documentation.

Hybrid enclosure support in gssgenvdisks

The **gssgenvdisks** command can detect hybrid enclosures. In case of hybrid enclosures, the **gssdgenvdisks** command requires two declustered arrays (DAs). One DA comprises HDDs only and the other one comprises SSDs only.

gssgenvdisks provides a default placement policy in case hybrid enclosures are used. According to this policy, any data vdisk is placed on the DA which is composed of HDDs only and any metadata vdisk is placed on the DA which is composed of SSDs only.

In the following example, DA1 consists of HDDs and DA2 consists of SSDs. This example shows that the default data vdisk is placed in DA1 and the metadata vdisk is placed in DA2.

```
# gssgenvdisks --create-vdisk --create-filesystem --verbose \
  --filesystem-name gpfs0 --reserved-space-percent 1
...
...
# mmlsvdisk
                                                                declustered     block size
 vdisk name                   RAID code        recovery group     array           in KiB
remarks
 ------------------           ---------------  ------------------ -----------     ----------
-------
 rg_io41_ce_Data_16M_2p_1     8+2p             rg_io41-ce         DA1               16384
 rg_io41_ce_Data_16M_2p_2     8+2p             rg_io41-ce         DA1               16384
 rg_io41_ce_MetaData_1M_3W_1  3WayReplication  rg_io41-ce         DA2                1024
 rg_io41_ce_loghome           4WayReplication  rg_io41-ce         DA1                2048
log
```

```
 rg_io41_ce_logtip            2WayReplication  rg_io41-ce          NVR                2048
logTip
 rg_io41_ce_logtipbackup      Unreplicated     rg_io41-ce          SSD                2048
logTipBackup
 rg_io42_ce_Data_16M_2p_1     8+2p             rg_io42-ce          DA1               16384
 rg_io42_ce_Data_16M_2p_2     8+2p             rg_io42-ce          DA1               16384
 rg_io42_ce_MetaData_1M_3W_1  3WayReplication  rg_io42-ce          DA2                1024
 rg_io42_ce_loghome           4WayReplication  rg_io42-ce          DA1                2048
log
 rg_io42_ce_logtip            2WayReplication  rg_io42-ce          NVR                2048
logTip
 rg_io42_ce_logtipbackup      Unreplicated     rg_io42-ce          SSD                2048
logTipBackup
```

You can override the default vdisk placement policy used in case of a hybrid enclosure system, by using the `--use-only-da` option. If the `--use-only-da` option is used, only the specified DA is considered for the vdisk creation. If the system is a hybrid enclosure and there are multiple DAs, the DAs available in the recovery group are not considered except for the one specified with the `--use-only-da` option.

You can create vdisks on other DAs using the BM Spectrum Scale RAID command.

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 5. 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.	

Table 5. Pre-installation tasks (continued)

ESS component	Description	Required actions	System settings
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.	
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`

Table 5. Pre-installation tasks (continued)

ESS component	Description	Required actions	System settings
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`
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: `kernel_ESS_5362_LE.tar.gz`
11. systemd	Updating the systemd service is required for all ESS nodes and it is verified by using **gssinstallcheck**.		Example: `systemd_ESS_5362_LE.tgz`
12. Network manager	Updating the network manager service is required for all ESS nodes and it is verified by using **gssinstallcheck**.		Example: `netmanager-ESS_5362_LE.tgz`

Table 5. Pre-installation tasks (continued)

ESS component	Description	Required actions	System settings
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 and EMS (user IDs and passwords)	The user IDs and passwords of the I/O servers and EMS are assigned during deployment.		Example: User ID: `root` Password: `esscluster` **Note:** The default password is set to expire upon first login and it must be changed.
15. FSP IPMI password	The IPMI password of the FSP. FSP IPMI of all the nodes assumed to be identical.		Example: `PASSW0RD`
16. Clustering network (hostname prefix or suffix)	This high-speed network is implemented on a 10Gb Ethernet, 25Gb Ethernet, 40Gb Ethernet, 100Gb 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`, `-25G`, `-40G`, `-100G` Hostnames with a suffix: `gssio1-ib`, `gssio2-ib`

Table 5. 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, `198.51.100.11` 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: `198.51.100.10` Hostname: `ems1-ib` FQDN: `ems1-ib.gpfs.net` I/O server 1: IP address: `198.51.100.11` Hostname: `gssio1-ib` FQDN: `gssio1-ib.data.net` I/O server 2: IP address: `198.51.100.12` Hostname: `gssio2-ib` FQDN: `gssio2-ib.data.net`
18. Red Hat Enterprise Linux 7.7	The Red Hat Enterprise Linux 7.7 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-7.7-server-ppc64le.iso` **Note:** The Red Hat Enterprise Linux 7.7 ISO name depends on the architecture (PPC64BE or PPC64LE).

Table 5. Pre-installation tasks (continued)			
ESS component	**Description**	**Required actions**	**System settings**
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.	
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 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 = 16M, #datansd=4, #metadatansd=2`

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
```

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_215_165.rpm
01SV860_215_165.xml
```

- System firmware update file for PPC64LE for updating using the command line:

```
01SV860_215_165.img
```

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.
  a) Unpack the `*img` file in the `/tmp/fwupdate` directory.

```
cd /opt/ibm/gss/install/rhel7/ppc64le/firmware
rpm -ivh 01SV860_215_165.rpm
```

  b) Shut down IBM Spectrum Scale and stop any ongoing I/O on the node.
  c) Verify the firmware level.

```
update_flash -v -f /tmp/fwupdate/01SV860_215_165.img
```

  d) Update the system firmware.

```
update_flash -f /tmp/fwupdate/01SV860_215_165.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 V9R1M920_SP0 before updating system firmware. For information about upgrading HMC, see HMC V9 Upgrade Procedure.

- Update the system firmware on PPC64BE systems as follows.
  a) From the HMC navigation area, click **Resources** > **All Systems** > **Server** > **Updates**.
  b) From the **Updates** menu, click **Change Licensed Internal Code** > **for the Current Release...**.
  c) Using SFTP, point to the `/opt/ibm/gss/install/rhel7/ppc64/firmware` directory on the EMS node.
  The following files should be present:

```
01SV860_215_165.rpm
01SV860_215_165.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.
  d) Select the update file and update the system firmware.
  It could take up to 30 minutes to update the firmware using HMC.

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 V9 update, use the following resources:
  - HMC V9 upgrade procedure: https://www.ibm.com/support/pages/hmc-v9-network-installation-images-and-installation-instructions
  - HMC V9 files: ftp://public.dhe.ibm.com/software/server/hmc/recovery_images/HMC_Recovery_V9R1M910_1_x86.iso
  - HMC V9 update: ftp://public.dhe.ibm.com/software/server/hmc/updates/HMC_Update_V9R1M941_x86.iso

  After upgrading, the HMC configuration should be similar to: `V9R1M940_SP0`

  Note: This is not applicable for the PPC64LE platform.

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-2020-5023-BE/
```

   For PPC64LE, issue:

```
mkdir -p /tmp/kernel/RHSA-2020-5023-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 *1160.6.1* --downloadonly --downloaddir=/tmp/kernel/RHSA-2020-5023-BE
yum update bpftool-3.10.0-1160.6.1.el7.ppc64.rpm --downloadonly --downloaddir=/tmp/kernel/
RHSA-2020-5023-BE
yum update perf-3.10.0-1160.6.1.el7.ppc64.rpm --downloadonly --downloaddir=/tmp/kernel/
RHSA-2020-5023-BE
yum update python-perf-3.10.0-1160.6.1.el7.ppc64.rpm --downloadonly \
--downloaddir=/tmp/kernel/RHSA-2020-5023-BE
```

   For PPC64LE, issue:

```
yum update *1160.6.1* --downloadonly --downloaddir=/tmp/kernel/RHSA-2020-5023-LE
yum update bpftool-3.10.0-1160.6.1.el7.ppc64le.rpm --downloadonly --downloaddir=/tmp/kernel/
RHSA-2020-5023-LE
yum update perf-3.10.0-1160.6.1.el7.ppc64le.rpm --downloadonly --downloaddir=/tmp/kernel/
RHSA-2020-5023-LE
yum update python-perf-3.10.0-1160.6.1.el7.ppc64le.rpm --downloadonly \
--downloaddir=/tmp/kernel/RHSA-2020-5023-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: Kernel 1160.6.1 packages - PPC64BE
- For PPC64LE, go to the following URL: Kernel 1160.6.1 packages - PPC64LE

b. Search for the required or any additional RPMs listed in “About the ESS Red Hat Linux Errata Kernel Update ” on page 114 and download them.

6. Package the directory.

   For PPC64BE, issue:

   ```
   cd /tmp/kernel ; tar -zcvf kernel_ESS_5362_BE.tar.gz RHSA-2020-5023-BE
   ```

   For PPC64LE, issue:

   ```
   cd /tmp/kernel ; tar -zcvf kernel_ESS_5362_LE.tar.gz RHSA-2020-5023-LE
   ```

   Note: Make sure that the RPM files are in the `RHSA-2020-5023-BE` or the `RHSA-2020-5023-LE` folder. Do not create any nested folder inside the `RHSA-2020-5023-BE` or the `RHSA-2020-5023-LE` folder and try to place the RPM file in that nested folder. Doing so results 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_ESS_5362_BE.tar.gz` or `kernel_ESS_5362_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_ESS_5362_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_ESS_5362_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 ” on page 114.

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.

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

For information about the kernel update for the current release, see RHBA-2020:3528.

This example shows errata update (RHSA-2020-5023) provided in the `/home/deploy` directory of the EMS node when shipped from factory.

The following packages are provided in `kernel_ESS_5362_BE.tar.gz`:

```
kernel-3.10.0-1160.6.1.el7.ppc64.rpm
kernel-abi-whitelists-3.10.0-1160.6.1.el7.noarch.rpm
kernel-bootwrapper-3.10.0-1160.6.1.el7.ppc64.rpm
```

kernel-devel-3.10.0-1160.6.1.el7.ppc64.rpm
kernel-doc-3.10.0-1160.6.1.el7.noarch.rpm
kernel-headers-3.10.0-1160.6.1.el7.ppc64.rpm
kernel-tools-3.10.0-1160.6.1.el7.ppc64.rpm
kernel-tools-libs-3.10.0-1160.6.1.el7.ppc64.rpm
kernel-tools-libs-devel-3.10.0-1160.6.1.el7.ppc64.rpm
bpftool-3.10.0-1160.6.1.el7.ppc64.rpm
perf-3.10.0-1160.6.1.el7.ppc64.rpm
python-perf-3.10.0-1160.6.1.el7.ppc64.rpm

The following packages are provided in kernel_ESS_5362_LE.tar.gz:

kernel-3.10.0-1160.6.1.el7.ppc64le.rpm
kernel-abi-whitelists-3.10.0-1160.6.1.el7.noarch.rpm
kernel-bootwrapper-3.10.0-1160.6.1.el7.ppc64le.rpm
kernel-devel-3.10.0-1160.6.1.el7.ppc64le.rpm
kernel-doc-3.10.0-1160.6.1.el7.noarch.rpm
kernel-headers-3.10.0-1160.6.1.el7.ppc64le.rpm
kernel-tools-3.10.0-1160.6.1.el7.ppc64le.rpm
kernel-tools-libs-3.10.0-1160.6.1.el7.ppc64le.rpm
kernel-tools-libs-devel-3.10.0-1160.6.1.el7.ppc64le.rpm
bpftool-3.10.0-1160.6.1.el7.ppc64le.rpm
perf-3.10.0-1160.6.1.el7.ppc64le.rpm
python-perf-3.10.0-1160.6.1.el7.ppc64le.rpm

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-2020-5007-BE/
   ```

   For PPC64LE, issue:

   ```
   mkdir -p /tmp/systemd/RHBA-2020-5007-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*219-78.el7_9.2* --downloadonly –downloaddir=/tmp/systemd/RHBA-2020-5007-BE
   yum update libgudev1-219-78.el7_9.2.ppc64.rpm --downloadonly –downloaddir=/tmp/systemd/
   RHBA-2020-5007-BE
   yum update libgudev1-devel-219-78.el7_9.2.ppc64.rpm --downloadonly –downloaddir=/tmp/systemd/
   RHBA-2020-5007-BE
   ```

   For PPC64LE, issue:

   ```
   yum update systemd*219-78.el7_9.2* --downloadonly –downloaddir=/tmp/systemd/RHBA-2020-5007-LE
   yum update libgudev1-219-78.el7_9.2.ppc64le.rpm --downloadonly –downloaddir=/tmp/systemd/
   RHBA-2020-5007-LE
   yum update libgudev1-devel-219-78.el7_9.2.ppc64le.rpm --downloadonly –downloaddir=/tmp/
   systemd/RHBA-2020-5007-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*78*el7*9*2*ppc64.rpm%26p=1%26sort=relevant%26scoped=false%26language=en
- For PPC64LE, go to the following URL: https://access.redhat.com/search/#/%3Fq=systemd*219*78*el7*9*2*ppc64le.rpm%26p=1%26sort=relevant%26scoped=false%26language=en

b. Search for the required or any additional RPMs listed in "About the ESS Red Hat Linux systemd update" on page 118 and download them.

6. Package the directory.

   For PPC64BE, issue:

   ```
   cd /tmp/systemd ; tar -zcvf systemd_ESS_5362_BE.tgz RHBA-2020-5007-BE
   ```

   For PPC64LE, issue:

   ```
   cd /tmp/systemd ; tar -zcvf systemd_ESS_5362_LE.tgz RHBA-2020-5007-LE
   ```

   Note: Make sure that the RPM files are in the RHBA-2020-5007-BE or the RHBA-2020-5007-LE folder. Do not create any nested folder inside the RHBA-2020-5007-BE or the RHBA-2020-5007-LE folder and try to place the RPM file in that nested folder. Doing so results 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_ESS_5362_BE.tgz` or `systemd_ESS_5362_LE.tgz` 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 systemd_ESS_5362_BE.tgz.tar.gz --silent
```

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

For PPC64LE, issue:

```
/var/tmp/gssdeploy -p systemd_ESS_5362_LE.tgz --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" on page 118.

About the ESS Red Hat Linux systemd update

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

This example shows systemd update (RHBA-2020-5007) 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-2020:5007.

The following packages are provided in `systemd_ESS_5362_BE.tgz`:

```
systemd-219-78.el7_9.2.ppc64.rpm
systemd-devel-219-78.el7_9.2.ppc64.rpm
systemd-journal-gateway-219-78.el7_9.2.ppc64.rpm
systemd-libs-219-78.el7_9.2.ppc64.rpm
```

systemd-networkd-219-78.el7_9.2.ppc64.rpm
systemd-python-219-78.el7_9.2.ppc64.rpm
systemd-resolved-219-78.el7_9.2.ppc64.rpm
systemd-sysv-219-78.el7_9.2.ppc64.rpm
libgudev1-219-78.el7_9.2.ppc64.rpm
libgudev1-devel-219-78.el7_9.2.ppc64.rpm

The following packages are provided in the systemd_ESS_5362_LE.tgz:

systemd-219-78.el7_9.2.ppc64le.rpm
systemd-devel-219-78.el7_9.2.ppc64le.rpm
systemd-journal-gateway-219-78.el7_9.2.ppc64le.rpm
systemd-libs-219-78.el7_9.2.ppc64le.rpm
systemd-networkd-219-78.el7_9.2.ppc64le.rpm
systemd-python-219-78.el7_9.2.ppc64le.rpm
systemd-resolved-219-78.el7_9.2.ppc64le.rpm
systemd-sysv-219-78.el7_9.2.ppc64le.rpm
libgudev1-219-78.el7_9.2.ppc64le.rpm
libgudev1-devel-219-78.el7_9.2.ppc64le.rpm

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-2020-5025-BE
   ```

   For PPC64LE, issue:

   ```
   mkdir -p /tmp/netmgr/RHBA-2020-5025-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*1.18.8-2.el7_9* --downloadonly –downloaddir=/tmp/netmgr/
   RHBA-2020-5025-BE
   ```

   For PPC64LE, issue:

   ```
   yum update NetworkManager*1.18.8-2.el7_9* --downloadonly –downloaddir=/tmp/netmgr/
   RHBA-2020-5025-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: NetworkManager 1.18.8-2.el7_9 - PPC64BE
      - For PPC64LE, go to the following URL: NetworkManager 1.18.8-2.el7_9 - PPC64LE

   b. Search for the required or any additional RPMs listed in “About the ESS Red Hat Linux Network Manager update” on page 122 and download them.

6. Package the directory.

   For PPC64BE, issue:

   ```
   cd /tmp/netmgr ; tar -zcvf netmgr_5362_BE.tgz RHBA-2020-5025-BE
   ```

   For PPC64LE, issue:

   ```
   cd /tmp/netmgr ; tar -zcvf netmgr_5362_LE.tgz RHBA-2020-5025-LE
   ```

   Note: Make sure that the RPM files are in the RHBA-2020-5025-BE or the RHBA-2020-5025-LE folder. Do not create any nested folder inside the RHBA-2020-5025-BE or the RHBA-2020-5025-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_5362_BE.tgz --silent
   ```

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

   For PPC64LE, issue:

   ```
   /var/tmp/gssdeploy -p netmgr_5362_BE.tgz --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” on page 122.

About the ESS Red Hat Linux Network Manager update

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

This example shows the Network Manager update (RHBA-2020-5025) 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 RHBA-2020:5025.

The following packages are provided in `netmgr_5362_BE.tgz`:

```
NetworkManager-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-adsl-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-bluetooth-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-glib-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-glib-devel-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-libnm-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-libnm-devel-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-ppp-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-ovs-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-team-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-tui-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-wifi-1.18.8-2.el7_9.ppc64.rpm
```

NetworkManager-wwan-1.18.8-2.el7_9.ppc64.rpm
NetworkManager-dispatcher-routing-rules-1.18.8-2.el7_9.noarch.rpm
NetworkManager-config-server-1.18.8-2.el7_9.noarch.rpm

The following packages are provided in the netmgr_5362_LE.tgz:

NetworkManager-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-adsl-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-bluetooth-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-glib-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-glib-devel-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-libnm-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-libnm-devel-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-ppp-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-ovs-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-team-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-tui-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-wifi-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-wwan-1.18.8-2.el7_9.ppc64le.rpm
NetworkManager-config-server-1.18.8-2.el7_9.noarch.rpm
NetworkManager-dispatcher-routing-rules-1.18.8-2.el7_9.noarch.rpm

ESS 5.3.6.1C

- IBM Spectrum Scale for IBM Elastic Storage Server Version 5.0.5.3
- Edition: Data Management for PPC64LE Architecture
- ESS version: 5.3.6.1C

```
gss_install-5.3.6.1C_ppc64le_datamanagement_20200918T022503Z.tgz
```

Summary of changes

1. IBM Spectrum Scale for ESS core: IBM Spectrum Scale 5.0.5.3
2. Support of Red Hat Enterprise Linux Server release 7.6 (Maipo) for PPC64LE (Power 9)
   - Kernel: 4.14 kernel-alt packages
   - Kernel version: 4.14.0-115.26.1.el7a.ppc64le
3. OFED: MLNX_OFED_LINUX-4.9-0.1.7.3
   - Includes firmware manager in ISO
   - Change of driver and adapter firmware
   - ConnectX-5 Ex CA Type: MT4121
   - Firmware: 16.27.2008
4. xCAT version: 2.15.1
5. System firmware
   - SMC BMC FW (2.06)
   - PNOR firmware (20190503)

     Release: open-power-SUPERMICRO-P9DSU-0.55.1-20190503
   - Boot Adapter (Adaptec):

     Firmware: 3.80

     Driver: 1.2.1.50877
   - NVR (LogTip): FW: 3.0.20.1
6. Storage firmware
   - SAS Adapter (0x3180) firmware: 15.00.00.00
   - Mpt3sas driver level: 31.100.01.00
   - Enclosure 5147-106 firmware: 524D
   - Drive ST800FM0183 (800GB SSD) firmware: 4036
   - Drive ST10000NM0226 (10TB HD) firmware: ECGE

MES upgrade flow

For customers who are looking to add storage to an existing building block, this option is now supported. The goal of MES is to expand a customer's available storage, either to existing file system or a new one, without the need to buy an entirely new building block. MES is designed to be an online operation and to not interrupt customer workloads.

Supported paths

- GS1S -> GS2S
- GS2S -> GS4S
- GL1S -> GL2S
- GL2S -> GL4S
- GL2S -> GL3S
- GL2S -> GL4S
- GL4S -> GL6S
- GL1C -> GL2C
- GL2C -> GL4C
- GL4C -> GL5C

You may not hop multiple building block types at a time in the same MES upgrade session. For example:

GL1S -> GL4S is not supported.

You must first go from GL1S -> GL2S and then from GL2S -> GL4S.

Prerequisites

1. All new or existing building blocks must be at ESS 5.3.6 or later. If there are protocol nodes in the setup, they must also be upgraded to the matching ESS version.
2. A system must be completely healthy prior to the **mmvdisk** conversion and the MES migration.
3. The recovery groups must be converted to **mmvdisk** before MES can begin. See the Flow section that follows.
4. Heavy IBM Spectrum Scale and I/O operations must be suspended prior to the **mmvdisk** conversion and the MES operation.
5. Additional enclosures cannot span frames.
6. If space needs to be made, for example for moving of the EMS, this has to be planned for accordingly.
7. Legacy enclosures (PPC64BE) and the new 4U106 (LE) model are not supported.
8. LBS must wear an ESD wrist band when physically working on the hardware (like plugging in SAS cables).

MES upgrade considerations

- Do not try to configure call home before MES is complete, that is until resizing is done.
- You can perform additional MES upgrades while the DA's are rebalancing.
- You can restripe the file system while the DA's are rebalancing.
- Although it is recommended, you do not have to rebalance the file system if NSDs are added during MES.

SAS cable plug-in tips

- Unlatch the cable arm from the I/O server node into which you will be plugging in the SAS cable.
- Remove the blue cap.
- Make sure the location code label from the cable matches the port location code and the port number.
- Remove the cap from the SAS cable connector and plug it into the port.
- You should hear a click when the cable is inserted correctly.

SSR tasks

SSR is are responsible for the following tasks.

1. Code 20 of the new enclosures - replacing parts as needed.
2. Running or labeling the new SAS cable connections.
3. Potentially making space in the frame - Moving the EMS.

SSR is not responsible for checking system health using **essutils** like in a rackless or a rackful solution.

LBS tasks

LBS is responsible for the following tasks.

1. Upgrade of ESS 5.3.6 - prior to MES engagement.
2. **mmvdisk** conversion
3. Pre and post **mmvdisk** conversion and MES health checks.
4. Plugging the SAS cables into the adapters and enclosures.
5. Performing MES software functions such as conversion and resizing.
6. New storage management functions such as adding new space to existing file system and creating a new file system.
7. Restriping the file system.
8. Replacing any bad parts such as disks or cables.
9. Pre and post engagement operations

Flow

TDA process ensures that the customer is prepared for the MES upgrade. Considerations such as if there is enough room in the rack or usage of the file system space are planned out.

LBS

1. LBS performs normal ESS software upgrade. Customer must be at ESS 5.3.6 for MES. This upgrade is treated as a separate engagement than the future MES operation. No **mmvdisk** conversion is done at this time.

=== MES operation begins ===

SSR

1. The SSR arrives at the customer site. If the EMS needs to be moved, the SSR shuts down GPFS and powers down the server to move. For more information, see “Shutting down and powering up ESS” on page 145.
2. The SSR places the new enclosures in the rack and establishes power connection. Based on the lights, the SSR performs a code 20 operation. If lights indicate any problem, they might need to take a service action.
3. The SSR runs the new SAS cable connections and labels in a bundle and hooks them to the frame. Later when LBS comes they simply plug in the connections when required in the flow.

4. The SSR places the EMS (if required) back into the existing frame or a new frame. Network connections and power are restored. Once the server is powered on, the SSR (or customer) can start GPFS to return the EMS back into the cluster.

LBS

1. Power on the new enclosure(s).
   - For GLxS or GSxS, the power cord should be connected. Press the switch to turn on the enclosure(s).
   - For GLxC, the power cord should be disconnected. Plug in the power cord to turn on the enclosure(s).
2. Verify that the system is converted to mmvdisk.
   a. **mmvdisk nodeclass list** - This command shows if the **mmvdisk** node class exists.
3. Upon arrival LBS should first perform the normal upgrade-related system verification steps. Run the following from the EMS:
   a. **gnrhealthcheck** - This command determines if there are any issues in various areas of ESS. Any problems that show up must be addressed before MES starts.
   b. **gssinstallcheck -N** *IONode1,IONode2* - This command checks the system to ensure all components match ESS 5.3.6 levels. If there are protocol nodes in the setup, check them using **gssinstallcheck -N** *ProtocolNode1,...,ProtocolNodeN*
   c. **/opt/ibm/gss/tools/samples/gssprecheck -N ems1 --upgrade --file /var/tmp/gssdeploy.cfg** - This command checks for common issues prior to doing any upgrade.
   d. **mmhealth node show -N all --verbose** - This command shows any system health related issues to address.
   e. Look for any events or tips that are open in the GUI. These also show up when you issue **mmhealth** but it is good to check in the GUI as well.
   f. **gssinstallcheck -N** *EMSNode,IONode1,IONode2* **--srv-events** - This command checks for any serviceable events reported from the Power 8 servers.
4. Convert to **mmvdisk**, if currently in legacy mode:
   a. **gssgenclusterrgs -G gss_ppc64 --suffix=-hs --convert**
   b. Wait for 5 minutes for daemons to recycle; file system stays up.
5. Verification steps:
   a. **mmgetstate -a** - Issue this command to ensure that all daemons are active.
   b. **mmlsmount all -L** - Issue this command to ensure that all mount points are still up. The file system must only be mounted on the EMS and protocol nodes (if applicable).

   After these issues are resolved, MES upgrade can begin.
6. Start by moving both recovery groups to `gssio2-hs`.

   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 and the current master server, run:

   ```
   mmvdisk recoverygroup list
   ```

   b. 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:

   ```
   mmvdisk recoverygroup change --recovery-group rg_gssio1-hs --active gssio2-hs
   ```

Running **mmvdisk recoverygroup list** should show both RGs actively managed by `gssio2-hs`.

7. Plug in the SAS cables for gssio1 on the server and enclosure ends. Shut down GPFS and then reboot the I/O node. Wait for 5 minutes for the node to reboot and paths to be rediscovered. Run the following commands to ensure that gssio1 has discovered the new enclosures.

   a. **gssstoragequickcheck -N gssio1**

   b. **gssfindmissingdisks -N gssio1**

   Both commands should return with no issues and recognize the new enclosure and disk counts. The paths should also be without error. Once this is complete, start IBM Spectrum Scale on the node in question. Once IBM Spectrum Scale is `active` proceed to the next step.

8. Move the recovery group ownership to `gssio1-hs`. Use the same commands as used in this step. Make sure to use the correct RG names.
9. Perform this step for `gssio2`.
10. Rebalance both recovery groups. You must use the following **mmvdisk** commands instead of the legacy **mmchrecoverygroup** command.

    a. **mmvdisk recoverygroup change --recovery-group rg_gssio1-hs --active gssio1-hs**

    b. **mmvdisk recoverygroup change --recovery-group rg_gssio2-hs --active gssio2-hs**

    c. Check that the ownership has changed using the **mmvdisk recoverygroup list** command listed in step 4.

11. Perform step 5 again before starting the resize command. Use the group name to check all servers at once; instead of `-N gssioX` use `-G gss_ppc64`.
12. Update MES enclosure and drive firmware. If there are any issues, you should stop and replace any disks or enclosures that could not be updated for some reason.

    `CurrentIoServer` implies running the command from either of I/O server nodes in the MES building block.

    Note: It might take up to an hour for the firmware upgrade to complete. You might notice that the fan starts to run at high speed. This is a known issue.

    a. `CurrentIoServer$` **mmchfirmware --type storage-enclosure**

    b. `CurrentIoServer$` **mmchfirmware --type drive**

    c. **mmhealth node show -N all --verbose**- This command shows any system health related issues to address.

    d. **gnrhealthcheck** - This command determines if there are any issues in various areas of ESS. Any problems that show up must be addressed before MES starts.

13. Add new storage into recovery groups.

    gssgenclusterrgs -N gssio3,gssio4 --suffix=-te0 --resize

14. Verify that the new storage is available and the DA is rebalancing.

    mmvdisk recoverygroup list --recovery-group *RG* **--all** - Run for both recovery groups. Note that the additional free space available in the DA and the background task for each DA is showing as `rebalancing`.

15. Start up the GUI and use **Edit rack components** to have the GUI discover the new topologies and make changes to the frame accordingly. Changes such as modify ESS model to consume more U space, move EMS, and so on.
16. Reconfigure call home.

    ```
    gsscallhomeconf -E ems1 -N EMSNode,IONode1,IONode2 --suffix=-hs --register=all
    ```

At this point, discussions with the customers need to be had on what to do with the free space (**gssgenvdisks**).

1. Add to the existing file system?
    a. See the add building block flow for tips on creating new NSDs and adding to an existing file system.
    b. Consider file system restripe at the end which might take time. (**mmrestripefs** *FileSystem* **-b**)
2. Create a new file system ?

- See the installation section on how to use **gssgenvdisks** on creating a new file system.

Before LBS operation is complete, it is good to run the health checks in step #1 again besides **gssprecheck**. If everything is clean, the operation is complete.

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.
...
...
```

NTP setup

Ensure that NTP is configured on the management server node to act as an NTP server. The NTP service must be enabled and the `chronyd` service must be disabled.

Run the following from the management node.

1. Stop the NTP server on all ESS nodes.

```
xdsh ems,gss_ppc64 "systemctl stop ntpd"
xdsh ems,gss_ppc64 "systemctl enable ntpd"
```

2. Stop and disable the `chronyd` service.

```
xdsh ems,gss_ppc64 "systemctl stop chronyd"
xdsh ems,gss_ppc64 "systemctl disable chronyd"
```

3. Assign the management server node as the NTP server.

```
makentp
```

   This command assigns the management server node as the NTP server. If there is any other NTP server already running then follow the Red Hat Enterprise Linux documentation to synchronize the time.

4. Edit the `/etc/ntp.conf` file accordingly.

```
# Use the local clock
server 127.127.1.0 prefer
fudge 127.127.1.0 stratum 10

restrict default kod nomodify notrap noquery nopeer
restrict 127.0.0.1

# Modify the line below to match your system
restrict 192.168.202.0 mask 255.255.255.0 nomodify notrap
driftfile /var/lib/ntp/ntp.drift
logfile /var/log/ntp.log
broadcastdelay 0.009
keys /etc/ntp/keys
```

5. Restart NTP on the management server.

```
systemctl restart ntpd
```

6. On the I/O server nodes, modify `/etc/ntp.conf` accordingly.

```
# Modify the line below to match your system
server 192.168.202.20
driftfile /var/lib/ntp/drift
disable auth
restrict 127.0.0.1
```

7. Synchronize time with the management server node.

```
ntpdate ems1
```

8. Start NTP.

```
systemctl start ntpd
```

Legacy deployment instructions

Use these steps for the legacy method of rebuilding the xCAT database from scratch and reinstalling the I/O server nodes. Use these steps only if there are problems with the normal installation flow or if you are deploying a new system that did not come with an EMS node.

Extract the ESS software and set up the local repository

1. Unpack the ESS software archive (This is contained in the ESS_DA_BASEIMAGE-5.3.6.2-ppc64le-Linux.tgz file).

   ```
   tar -zxvf gss_install-5.3.6.2_ppc64le_dataaccess_20201116T223856Z.tgz
   ```

2. Check the SHA 256 checksum:

   ```
   shasum -a 256 -c gss_install-5.3.6.2_ppc64le_dataaccess_20201116T223856Z.sha256
   ```

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

   ```
   /bin/sh gss_install-5.3.6.2_ppc64le_dataaccess_20201116T223856Z --remove
   ```

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

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-5.3.6.2_ppc64le_dataaccess_20201116T223856Z --text-only
   ```

5. Make the gssdeploy script executable, if it is not yet executable:

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

6. 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
   ```

Rediscover and rebuild the xCAT database

1. 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.
- 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 returns the chassis serial numbers and FSP IP addresses of the EMS and I/O server nodes in the building block(s).
- There is a slight hang when **gssdeploy -f** attempts to query the FSP IP address configured on the EMS operating system. This operation eventually times out and fails which is the normal behavior. The only EMS FSP IP that should be discovered is the one assigned to HMC port 1.
- 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`.

The main purpose of **gssdeploy -i** is to properly identify the slot of the ESS components within the IBM Spectrum Scale GUI. This is important for disk and other hardware replacements in the future. If using the default naming conventions, the bottom most server found in a frame is gssio1, then gssio2, and so on.

Note: Upgrading to HMC SP3 might affect support for hardware call home.

2. 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**]
- Type of deployment: DEPLOYMENT_TYPE

  The default type of deployment is ESS. It can also be ADD_BB.

  ESS: Deploys I/O server nodes.

  ADD_BB: Adds new building block of I/O server nodes.

  PPC64LE protocol nodes can be deployed using the CES or ADD_CES deployment types.

  CES: Deploys protocol nodes.

  ADD_CES: Adds new protocol nodes.
- 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
- xCAT Group: GSS_GROUP

  For example, gss_ppc64, ces_ppc64

  Note: Modification is only required when adding a protocol node to existing setup or adding an ESS I/O server node. A temporary group name is used for that operation.

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`.

3. Copy the RHEL 7.7 ISO file to the directory specified in the `gssdeploy.cfg` file.
4. 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: Do the following steps before running **gssdeploy -x**.

- Power down the storage enclosures, or remove the SAS cables.
- Make sure that you update the `/etc/hosts` file with the xCAT node names and IP addresses that match with values defined in `gssdeploy.cfg`.

5. 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.

6. Log out and then log back in to acquire the environment updates.
7. 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
```

8. Set up the kernel, systemd, and network manager errata, and OPAL patch repositories. For example, use the following command on PPC64BE systems:

```
/var/tmp/gssdeploy -k /home/deploy/kernel_ESS_5362_LE.tgz -p \
/home/deploy/systemd_5362_LE.tgz,/home/deploy/netmgr_5362_LE.tgz,\
/home/deploy/opal-patch-le.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: systems and network manager patch). 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 root root 47647044 Nov 17 01:49 kernel-3.10.0-1160.6.1.el7.ppc64le.rpm
```

```
/install/gss/otherpkgs/rhels7/<arch>/patch
-rw-r--r-- 1 root root 5460520 Nov 17 01:49 systemd-219-78.el7_9.2.ppc64le.rpm
-rw-r--r-- 1 root root 2001044 Nov 17 21:54 NetworkManager-1.18.8-2.el7_9.ppc64le.rpm
```

Wrong file permission leads to node deployment failure.

9. Run updatenode.

```
updatenode ems1 -V -P gss_updatenode
```

Note: If the kernel changes, you will be prompted to reboot and re-run updatenode.

10. Run gss_ofed.

```
updatenode ems1 -V -P gss_ofed
```

11. Run gss_iprraid.

```
updatenode ems1 -V -P gss_iprraid
```

12. Reboot the node.

Deploy the I/O server nodes

1. Before initiating the deployment of the I/O server nodes, do the following on the management server node:
   a. Verify that the running kernel level is at the desired level (for example, 1127.19.1) using the **uname -a** command.
   b. Verify that there are no repository errors and all repositories are in place (patch, kernel, etc) 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** in parallel 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 "Running gssinstallcheck in parallel" on page 133.

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: rhels7.7-ppc64le-install-gss
node: rhels7.7-ppc64le-install-gss
```

PPC64BE installations:

```
node: rhels7.7-ppc64-install-gss
node: rhels7.7-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.

You can also use the following command to view the progress of the node installation:

```
rcons NodeName -f
```

To exit an rcons session, press Ctrl+E followed by C and then the period key (.).

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 -N EMSNode,IONode1,IONode2
```

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

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

For more information, see "Running gssinstallcheck in parallel" on page 133.

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 might be flagged by the `gssinstallcheck` script. They will go away after the remaining installation steps are completed.

At this point, you might perform the following tasks by using the provided tools:

- Create network bonds (**gssgennetworks**)
- Create cluster (**gssgencluster**)
- Create recovery groups (**gssgenclusterrgs**)
- Create vdisks, NSDs, and file systems (**gssgenvdisks**)

For recovery groups and file system, it is recommended to use **mmvdisk** directly. For additional post-installation steps such as configuring GUI or call home or doing health checks, etc., see command examples in other sections in this document.

Considerations for adding PPC64LE building blocks to ESS PPC64BE building blocks

When adding PPC64LE nodes to ESS PPC64BE systems, following considerations apply.

- All nodes must be at IBM Spectrum Scale 4.2.3.6 or later, and ESS 5.0.2 or later.
- The tested flow that is recommended for this procedure is available in the Box.
- Each architecture must have its own EMS node.
- PPC64LE must be the primary EMS node where GUI and performance collector services run.
- Shut down the GUI on the PPC64BE system and then change the performance collector to the PPC64LE EMS.
- Add the PPC64LE nodes as sensor nodes.
- You must have a flat network for the 1Gb xCAT. All nodes must be reachable and resolvable. The same consideration applies for the high speed network.
- Update and copy `/etc/hosts` to all nodes. Run **makedns** on both EMS nodes.
- Be mindful of pools. Do not mix SSD and HDD for instance. You might need to also set up policy files.
- Before starting the GUI, update the component database.

```
mmaddcompspec default --replace
```

- After starting the GUI, add the xCAT PPC64LE IP to the hardware monitoring list.
- Run the **Edit Rack Components** wizard in the GUI.

Note:

- In the **Hardware Monitoring** page of the GUI, add the IPs of all EMS' over the xCAT network.
- The collector must run on the same node as the GUI. Make sure to shut down the PPC64BE GUI and move the collector to the PPC64LE EMS before starting the GUI there.
- Before starting the collector on the PPC64LE EMS and GUI, back up and restore the data from the PPC64BE EMS (`/opt/IBM/zimon/data`).

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
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 that 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 must 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*.

Product Number: 5765-DME
5765-DAE

SC28-3151-02