IBM Elastic Storage System 3000
Version 6.0.1

Service Guide

SC28-3158-02

Note

Before using this information and the product it supports, read the information in "Notices" on page 61.

This edition applies to version 6 release 0 modification 1 of the following product 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)

IBM welcomes your comments; see the topic "How to submit your comments" on page x. 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

Figures

Tables

About this information

Who should read this information

This information is intended for administrators of IBM Elastic Storage® System (ESS) that includes IBM Spectrum Scale RAID.

IBM Elastic Storage System information units

IBM Elastic Storage System (ESS) 3000 documentation consists of the following information units.

Information unit	Type of information	Intended users
Hardware Planning and Installation Guide	This unit provides ESS 3000 information including technical overview, planning, installing, troubleshooting, and cabling.	System administrators and IBM support team
Quick Deployment Guide	This unit provides ESS 3000 information including the software stack, deploying, upgrading, and best practices.	System administrators, analysts, installers, planners, and programmers of IBM Spectrum Scale clusters who are very experienced with the operating systems on which each IBM Spectrum Scale cluster is based
Service Guide	This unit provides ESS 3000 information including events, servicing, and parts listings.	System administrators and IBM support team
Problem Determination Guide	This unit provides ESS 3000 information including setting up call home, replacing servers, issues, maintenance procedures, and troubleshooting.	System administrators and IBM support team
Command Reference	This unit provides information about ESS commands and scripts.	System administrators and IBM support team
IBM Spectrum Scale RAID: Administration	This unit provides IBM Spectrum Scale RAID information including administering, monitoring, commands, and scripts.	• System administrators of IBM Spectrum Scale systems • Application programmers who are experienced with IBM Spectrum Scale systems and familiar with the terminology and concepts in the XDSM standard

Related information

Related information

For information about:

- IBM Spectrum Scale, see:

  http://www.ibm.com/support/knowledgecenter/STXKQY/ibmspectrumscale_welcome.html
- mmvdisk command, see mmvdisk documentation.

- Mellanox OFED (MLNX_OFED v4.9-0.1.7.0) Release Notes, go to https://docs.mellanox.com/display/OFEDv490170/Release%20Notes

Conventions used in this information

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

Table 1. Conventions

Convention	Usage
bold	Bold words or characters represent system elements that you must use literally, such as commands, flags, values, and selected menu options. Depending on the context, **bold** typeface sometimes represents path names, directories, or file names.
bold underlined	bold underlined keywords are defaults. These take effect if you do not specify a different keyword.
constant width	Examples and information that the system displays appear in constant-width typeface. Depending on the context, constant-width typeface sometimes represents path names, directories, or file names.
italic	*Italic* words or characters represent variable values that you must supply. *Italics* are also used for information unit titles, for the first use of a glossary term, and for general emphasis in text.
<key>	Angle brackets (less-than and greater-than) enclose the name of a key on the keyboard. For example, <Enter> refers to the key on your terminal or workstation that is labeled with the word *Enter*.

****	In command examples, a backslash indicates that the command or coding example continues on the next line. For example:
	`mkcondition -r IBM.FileSystem -e "PercentTotUsed > 90" \`
	`-E "PercentTotUsed < 85" -m p "FileSystem space used"`

{item}	Braces enclose a list from which you must choose an item in format and syntax descriptions.
[item]	Brackets enclose optional items in format and syntax descriptions.
<Ctrl-*x*>	The notation <Ctrl-*x*> indicates a control character sequence. For example, <Ctrl-c> means that you hold down the control key while pressing <c>.
item...	Ellipses indicate that you can repeat the preceding item one or more times.
|	In *synopsis* statements, vertical lines separate a list of choices. In other words, a vertical line means *Or*. In the left margin of the document, vertical lines indicate technical changes to the information.

How to submit your comments

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

scale@us.ibm.com

Chapter 1. Events

The recorded events are stored in the local database on each node. The user can get a list of recorded events by using the **mmhealth node eventlog** command. Users can use the **mmhealth node show** or **mmhealth cluster show** commands to display the active events in the node and cluster respectively.

The recorded events can also be displayed through the GUI.

When you upgrade to IBM Spectrum Scale 5.0.5.3 or a later version, the nodes where no `sqlite3` package is installed have their RAS event logs converted to a new database format to prevent known issues. The old RAS event log is emptied automatically. You can verify that the event log is emptied either by using the **mmhealth node eventlog** command or in the IBM Spectrum Scale GUI.

Note: The event logs are updated only the first time IBM Spectrum Scale is upgraded to version 5.0.5.3 or higher.

The following sections list the RAS events that are applicable to various components of the IBM Spectrum Scale system:

Array events

The following table lists the events that are created for the *Array* component.

Table 2. Events for the Array component

Event	Event Type	Severity	Message	Description	Cause	User Action
gnr_array_found	INFO_ADD_ENTITY	INFO	GNR declustered array {0} was found.	A GNR declustered array listed in the IBM Spectrum Scale configuration was detected.		N/A
gnr_array_needsservice	STATE_CHANGE	WARNING	GNR declustered array {0} needs service.	The declustered array state needs service.	N/A	N/A
gnr_array_ok	STATE_CHANGE	INFO	GNR declustered array {0} is ok.	The declustered array state is ok.	N/A	N/A
gnr_array_unknown	STATE_CHANGE	WARNING	GNR declustered array {0} is in unknown state.	The declustered array state is unknown.	N/A	N/A
gnr_array_vanished	INFO_DELETE_ENTITY	INFO	GNR declustered array {0} has vanished.	A GNR declustered array listed in the IBM Spectrum Scale configuration was not detected.	A GNR declustered array, listed in the IBM Spectrum Scaleconfiguration as mounted before, is not found. This could be a valid situation	Run the **mmlsrecoverygroup** command to verify that all the expected GNR declustered arrays exist.

Canister events

The following table lists the events that are created for the *Canister* component.

Table 3. Events for the Canister component

Event	Event Type	Severity	Message	Description	Cause	User Action
bootdrive_installed	STATE_CHANGE	INFO	The bootdrive attached to port {0} is available.	The bootdrive is available.	The **tsplatformstat -a** command returns the bootdrives as expected.	N/A

Table 3. Events for the Canister component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
bootdrive_mirror_degraded	STATE_CHANGE	WARNING	The bootdrive's mirroring is degraded.	The bootdrive's mirroring is degraded.	The **tsplatformstat -a** command returns a DEGRADED value for at least one partition.	N/A
bootdrive_mirror_failed	STATE_CHANGE	ERROR	The bootdrive's mirroring is failed.	The bootdrive's mirroring is failed.	The **tsplatformstat -a** command returns a FAILED value for at least one partition.	N/A
bootdrive_mirror_ok	STATE_CHANGE	INFO	The bootdrive's mirroring is OK.	The bootdrive's mirroring is OK.	The **tsplatformstat -a** command returns optimal for all partitions.	N/A
bootdrive_mirror _unconfigured	STATE_CHANGE	WARNING	The bootdrive's mirroring is unconfigured.	The bootdrive's mirroring is unconfigured.	The **tsplatformstat -a** command returns unconfigured for mirroring.	N/A
bootdrive_missing	STATE_CHANGE	ERROR	The bootdrive on port {0} is missing or dead.	One bootdrive is missing or dead. Redundancy is not given anymore.	The **tsplatformstat -a** command returns only one instead of two bootdrives. Two drives are expected to ensure redundancy.	Inspect that the drive is correctly installed on the referenced port. Else insert or replace the drive.
bootdrive_smart_failed	STATE_CHANGE	ERROR	The smart assessment of bootdrive {0} attached to port {1} does not return OK.	The bootdrive's smart assessment does not return OK.	The **tsplatformstat -a** command does not return a PASSED value in the **selfAssessment** field for the bootdrive.	Verify the smart status of the bootdrive using **tsplatformstat** command or `smartctl`.
bootdrive_smart_ok	STATE_CHANGE	INFO	The smart assessment of bootdrive {0} attached to port {1} returns OK.	The bootdrive's smart assessment returns OK.	The **tsplatformstat -a** command returns a PASSED in the **selfAssessment** field for the bootdrive.	N/A
can_fan_failed	STATE_CHANGE	WARNING	Fan {0} is failed.	The fan state is failed.	The **mmlsenclosure** command reports the fan as failed.	Check the fan status by using the **mmlsenclosure** command. Replace the fan module in the canister.
can_fan_ok	STATE_CHANGE	INFO	Fan {0} is OK.	The fan state is OK.	The **mmlsenclosure** command reports the fan as working.	N/A
can_temp_bus_failed	STATE_CHANGE	WARNING	Temperature sensor {0} I2C bus is failed.	The temperature sensor I2C bus failed.	The **mmlsenclosure** command reports the temperature sensor with a failure.	Check the temperature status by using the **mmlsenclosure** command.
can_temp_high_critical	STATE_CHANGE	WARNING	Temperature sensor {0} measured a high temperature value.	The temperature exceeded the actual high critical threshold value for at least one sensor.	The **mmlsenclosure** command reports the temperature sensor with a failure.	Check the temperature status by using the **mmlsenclosure** command.
can_temp_high_warn	STATE_CHANGE	WARNING	Temperature sensor {0} measured a high temperature value.	The temperature exceeded the actual high warning threshold value for at least one sensor.	The **mmlsenclosure** command reports the temperature sensor with a failure.	Check the temperature status by using the **mmlsenclosure** command.

Table 3. Events for the Canister component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
can_temp_low_critical	STATE_CHANGE	WARNING	Temperature sensor {0} measured a low temperature value.	The temperature has fallen below the actual low critical threshold value for at least one sensor.	The **mmlsenclosure** command reports the temperature sensor with a failure.	Check the temperature status by using the **mmlsenclosure** command.
can_temp_low_warn	STATE_CHANGE	WARNING	Temperature sensor {0} measured a low temperature value.	The temperature has fallen below the actual low warning threshold value for at least one sensor.	The **mmlsenclosure** command reports the temperature sensor with a failure.	Check the temperature status by using the **mmlsenclosure** command.
can_temp_sensor_failed	STATE_CHANGE	WARNING	Temperature sensor {0} is failed.	The temperature sensor state is failed.	The **mmlsenclosure** command reports the temperature sensor with a failure.	Check the temperature status by using the **mmlsenclosure** command. Replace the canister.
can_temp_sensor_ok	STATE_CHANGE	INFO	Temperature sensor {0} is OK.	The temperature sensor state is OK.	N/A	N/A
canister_failed	STATE_CHANGE	ERROR	Canister {0} is failed.	The canister is reporting a failed hardware state. This might be caused by a failure of an underlying component. For example, the fan.	The **mmlsenclosure** command reports the canister as failed.	Check for detailed error events of canister components by using the **mmhealth** command. Inspect the output of **mmlsenclosure all -L** command for the referenced canister.
canister_ok	STATE_CHANGE	INFO	Canister {0} is OK.	The canister state is OK.	The **mmlsenclosure** command reports the canister as failed.	N/A
cpu_inspection_failed	STATE_CHANGE	ERROR	The inspection of the CPU slots found a mismatch	Number of populated CPU slots, number of enabled CPUs, number of CPU cores, number of CPU threads or CPU speed is not as expected.	The **/opt/ibm/gss/tools/bin/ess3kplt** command returned an `InspectionPassed unequal to True` value.	Check for specific events related to CPUs by using the **mmhealth** command. Inspect the output of the **ess3kplt** command for details.
cpu_inspection_passed	STATE_CHANGE	INFO	The CPUs of the canister are OK.	The CPU speed and number of populated CPU slots is as expected.	The **/opt/ibm/gss/tools/bin/ess3kplt** command returned an `InspectionPassed equal to True` value.	N/A
cpu_speed_ok	STATE_CHANGE	INFO	The CPU speed is OK.	The speed of all CPUs is as expected.	The **/opt/ibm/gss/tools/bin/ess3kplt** command returned no speed errors.	N/A
cpu_speed_wrong	STATE_CHANGE	ERROR	One or more CPUs have an unsupported speed.	The speed of one or more CPUs is not as expected. This configuration is not supported.	The **/opt/ibm/gss/tools/bin/ess3kplt** command returned one or more speed errors.	Inspect the output of the **ess3kplt** command to see which CPUs have an unsupported speed.
dimm_inspection_failed	STATE_CHANGE	ERROR	The inspection of the memory dimm slots found a failure.	The capacity, speed, or number of populated dimm slots is not as expected.	The **/opt/ibm/gss/tools/bin/ess3kplt** command returned an `InspectionPassed unequal to True` value.	Check for specific events related to dimms by using the **mmhealth** command. Inspect the output of the **ess3kplt** command for details.
dimm_inspection_passed	STATE_CHANGE	INFO	The memory dimms of the canister is OK.	The capacity, speed, and number of populated dimm slots is as expected.	The **/opt/ibm/gss/tools/bin/ess3kplt** command returned an `InspectionPassed equal to True` value.	N/A

Table 3. Events for the Canister component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
dimm_size_ok	STATE_CHANGE	INFO	All installed memory dimms have the expected capacity.	The capacity of all populated memory dimm slots is as expected.	The **/opt/ibm/gss/tools/bin/ess3kplt** command returned no capacity errors.	N/A
dimm_size_wrong	STATE_CHANGE	ERROR	One or more memory dimm modules have an unsupported capacity.	The capacity of one or more memory dimm slots is not as expected. This configuration is not supported.	The **/opt/ibm/gss/tools/bin/ess3kplt** command returned some capacity errors.	Inspect the output of the **ess3kplt** command to see which memory dimm slots have an unsupported capacity and replace those dimm modules.
dimm_speed_ok	STATE_CHANGE	INFO	All installed memory dimms have the expected speed.	The speed of all populated memory dimm slots is as expected.	The **/opt/ibm/gss/tools/bin/ess3kplt** command returned no speed errors.	N/A
dimm_speed_wrong	STATE_CHANGE	ERROR	One or more memory dimm modules have an unsupported speed.	The speed of one or more memory dimm slots is not as expected. This configuration is not supported.	The **/opt/ibm/gss/tools/bin/ess3kplt** command returned some speed errors.	Inspect the output of the **ess3kplt** command to see which memory dimm slots have an unsupported speed and replace those dimm modules.
pair_canister_missing	STATE_CHANGE	WARNING	Pair canister {0} is missing or dead.	Could not get the state of the pair canister. It might be missing or dead.	The **mmlsenclosure** command reports only one canister instead of two.	Check for detailed error events of the referenced canister node by using the **mmhealth** command. Inspect the output of the **mmlsenclosure all -L** command for the referenced canister.
pair_canister_visible	STATE_CHANGE	INFO	Pair canister {0} is visible.	Successfully get the state of the pair canister.	The **mmlsenclosure** command reports both canisters.	N/A

Enclosure events

The following table lists the events that are created for the *Enclosure* component.

Table 4. Events for the Enclosure component

Event	Event Type	Severity	Message	Description	Cause	User Action
adapter_bios_notavail	STATE_CHANGE	WARNING	The bios level of adapter {0} is not available.	The bios level of the adapter is not available.	N/A	Check the installed BIOS level using the **mmlsfirmware** command.
adapter_bios_ok	STATE_CHANGE	INFO	The BIOS level of adapter {0} is correct.	The BIOS level of the adapter is correct.	N/A	N/A
adapter_bios_wrong	STATE_CHANGE	WARNING	The bios level of adapter {0} is wrong.	The bios level of the adapter is wrong.	N/A	Check the installed BIOS level using the **mmlsfirmware** command.
adapter_firmware_notavail	STATE_CHANGE	WARNING	The firmware level of adapter {0} is not available.	The firmware level of the adapter is not available.	N/A	Check the installed BIOS level using the **mmlsfirmware** command.
adapter_firmware_ok	STATE_CHANGE	INFO	The firmware level of adapter {0} is correct.	The firmware level of the adapter is correct.	N/A	N/A
adapter_firmware_wrong	STATE_CHANGE	WARNING	The firmware level of adapter {0} is wrong.	The firmware level of the adapter is wrong.	N/A	Check the installed BIOS level using the **mmlsfirmware** command.

Table 4. Events for the Enclosure component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
current_failed	STATE_CHANGE	ERROR	currentSensor {0} failed.	The currentSensor state is failed.	N/A	N/A
current_ok	STATE_CHANGE	INFO	currentSensor {0} is ok.	The currentSensor state is ok.	N/A	N/A
current_warn	STATE_CHANGE	WARNING	currentSensor {0} is degraded.	The currentSensor state is degraded.	N/A	N/A
dcm_drawer_open	STATE_CHANGE	WARNING	DCM {0} drawer is open.	The DCM drawer is open.	N/A	N/A
dcm_failed	STATE_CHANGE	WARNING	DCM {0} is failed.	The DCM state is failed.	N/A	N/A
dcm_not_available	STATE_CHANGE	WARNING	DCM {0} is not available.	The DCM is not installed or not responding.	N/A	N/A
dcm_ok	STATE_CHANGE	INFO	DCM {id[1]} is ok.	The DCM state is ok.	N/A	N/A
drawer_failed	STATE_CHANGE	ERROR	drawer {0} is failed.	The drawer state is failed.	N/A	N/A
drawer_ok	STATE_CHANGE	INFO	drawer {0} is ok.	The drawer state is ok.	N/A	N/A
drive_firmware_notavail	STATE_CHANGE	WARNING	The firmware level of drive {0} is not available.	The firmware level of the drive is not available.	N/A	Check the installed firmware level using the **mmlsfirmware** command.
drive_firmware_ok	STATE_CHANGE	INFO	The firmware level of drive {0} is correct.	The firmware level of the drive is correct.	N/A	N/A
drive_firmware_wrong	STATE_CHANGE	WARNING	The firmware level of drive {0} is wrong.	The firmware level of the drive is wrong.	N/A	Check the installed firmware level using the **mmlsfirmware** command.
enclosure_data	STATE_CHANGE	INFO	Enclosure data found.	Successfully queried the enclosure details.	The **mmlsenclosure all -L -Y** command reports enclosure data.	N/A
enclosure_firmware_notavail	STATE_CHANGE	WARNING	The firmware level of enclosure {0} is not available.	The firmware level of the enclosure is not available.	N/A	Check the installed firmware level using the **mmlsfirmware** command.
enclosure_firmware_ok	STATE_CHANGE	INFO	The firmware level of enclosure {0} is correct.	The firmware level of the enclosure is correct.	N/A	N/A

Table 4. Events for the Enclosure component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
enclosure_firmware_unknown	STATE_CHANGE	WARNING	The firmware level of enclosure {0} is unknown.	The SAS card is unable to read enclosure firmware.	The SAS card does not report the enclosure firmware.	Check the SAS connectivity from node to enclosure. Use the **mmlsrecoverygroup rg_name -L --pdisk** command to verify if all the paths to pdisk are available. Check the SAS connectivity using a combination of the **mmgetpdisktopology** and the **topsummary** command. If there is an issue with the SAS HBA or SAS Cable, reboot the node to see if this resolves the issue. If not contact your IBM representative.
enclosure_firmware_wrong	STATE_CHANGE	WARNING	The firmware level of enclosure {0} is wrong.	The firmware level of the enclosure is wrong.	N/A	Check the installed firmware level using **mmlsfirmware** command.
enclosure_found	INFO_ADD_ENTITY	INFO	Enclosure {0} was found.	A GNR enclosure listed in the IBM Spectrum Scale configuration was detected.	N/A	N/A
enclosure_needsservice	STATE_CHANGE	WARNING	Enclosure {0} needs service.	The enclosure needs service.	N/A	N/A
enclosure_ok	STATE_CHANGE	INFO	Enclosure {0} is ok.	The enclosure state is ok.	N/A	N/A
enclosure_unknown	STATE_CHANGE	WARNING	Enclosure state {0} is unknown.	The enclosure state is unknown.	N/A	N/A
enclosure_vanished	INFO_DELETE_ENTITY	INFO	Enclosure {0} has vanished.	A GNR enclosure listed in the IBM Spectrum Scale configuration was not detected.	A GNR enclosure, listed in the IBM Spectrum Scale configuration as mounted before, is not found. This could be a valid situation.	Run the **mmlsenclosure** command to verify that all expected enclosures exist.
esm_absent	STATE_CHANGE	WARNING	ESM {0} is absent.	The ESM state is not installed .	N/A	N/A
esm_failed	STATE_CHANGE	WARNING	ESM {0} is failed.	The ESM state is failed.	N/A	N/A
esm_ok	STATE_CHANGE	INFO	ESM {0} is ok.	The ESM state is ok.	N/A	N/A
expander_absent	STATE_CHANGE	WARNING	expander {0} is absent.	The expander is absent.	N/A	N/A
expander_failed	STATE_CHANGE	ERROR	expander {0} is failed.	The expander state is failed.	N/A	N/A
expander_ok	STATE_CHANGE	INFO	expander {0} is ok.	The expander state is ok.	N/A	N/A
fan_failed	STATE_CHANGE	WARNING	Fan {0} is failed.	The fan state is failed.	N/A	N/A
fan_ok	STATE_CHANGE	INFO	Fan {0} is ok.	The fan state is ok.	N/A	N/A
fan_speed_high	STATE_CHANGE	WARNING	Fan {0} speed is too high	The fan speed is out of the tolerance range	N/A	Check the enclosure cooling module LEDs for fan faults.

Table 4. Events for the Enclosure component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
fan_speed_low	STATE_CHANGE	WARNING	Fan {0} speed is too low	The fan speed is out of the tolerance range	N/A	Check the enclosure cooling module LEDs for fan faults.
no_enclosure_data	STATE_CHANGE	WARNING	Enclosure data and state information cannot be queried.	Cannot query the enclosure details. State reporting for all enclosures and canisters will be incorrect.	The **mmlsenclosure all -L -Y** command fails to report any enclosure data.	Run the **mmlsenclosure** command to check for errors. Use the **lsmod** command to verify that the pemsmod is loaded.
power_high_current	STATE_CHANGE	WARNING	Power supply {0} reports high current.	The DC power supply current is greater than the threshold.	N/A	N/A
power_high_voltage	STATE_CHANGE	WARNING	Power supply {0} reports high voltage.	The DC power supply voltage is greater than the threshold.	N/A	N/A
power_no_power	STATE_CHANGE	WARNING	Power supply {0} has no power.	Power supply has no input AC power. The power supply may be turned off or disconnected from the AC supply.	N/A	N/A
power_supply_absent	STATE_CHANGE	WARNING	Power supply {0} is missing.	The power supply is missing	N/A	N/A
power_supply_failed	STATE_CHANGE	WARNING	Power supply {0} is failed.	The power supply state is failed.	N/A	N/A
power_supply_off	STATE_CHANGE	WARNING	Power supply {0} is off.	The power supply is not providing power.	N/A	N/A
power_supply_ok	STATE_CHANGE	INFO	Power supply {0} is ok.	The power supply state is ok.	N/A	N/A
power_switched_off	STATE_CHANGE	WARNING	Power supply {0} is switched off.	The requested on bit is off, indicating that the power supply has not been manually turned on or been requested to turn on by setting the requested on bit.	N/A	N/A
sideplane_failed	STATE_CHANGE	ERROR	sideplane {0} failed.	The sideplane state is failed.	N/A	N/A
sideplane_ok	STATE_CHANGE	INFO	sideplane {0} is ok.	The sideplane state is ok.	N/A	N/A
temp_bus_failed	STATE_CHANGE	WARNING	Temperature sensor {0} I2C bus is failed.	The temperature sensor I2C bus has failed.	N/A	N/A
temp_high_critical	STATE_CHANGE	WARNING	Temperature sensor {0} measured a high temperature value.	The temperature has exceeded the actual high critical threshold value for at least one sensor.	N/A	N/A
temp_high_warn	STATE_CHANGE	WARNING	Temperature sensor {0} measured a high temperature value.	The temperature has exceeded the actual high warning threshold value for at least one sensor.	N/A	N/A
temp_low_critical	STATE_CHANGE	WARNING	Temperature sensor {0} measured a low temperature value.	The temperature has fallen below the actual low critical threshold value for at least one sensor.	N/A	N/A

Table 4. Events for the Enclosure component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
temp_low_warn	STATE_CHANGE	WARNING	Temperature sensor {0} measured a low temperature value.	The temperature has fallen below the actual low warning threshold value for at least one sensor.	N/A	N/A
temp_sensor_failed	STATE_CHANGE	WARNING	Temperature sensor {0} is failed.	The temperature sensor state is failed.	N/A	N/A
temp_sensor_ok	STATE_CHANGE	INFO	Temperature sensor {0} is ok.	The temperature sensor state is ok.	N/A	N/A
voltage_bus_failed	STATE_CHANGE	WARNING	Voltage sensor {0} I2C bus is failed.	The voltage sensor I2C bus has failed.	N/A	N/A
voltage_high_critical	STATE_CHANGE	WARNING	Voltage sensor {0} measured a high voltage value.	The voltage has exceeded the actual high critical threshold value for at least one sensor.	N/A	N/A
voltage_high_warn	STATE_CHANGE	WARNING	Voltage sensor {0} measured a high voltage value.	The voltage has exceeded the actual high warning threshold value for at least one sensor.	N/A	N/A
voltage_low_critical	STATE_CHANGE	WARNING	Voltage sensor {0} measured a low voltage value.	The voltage has fallen below the actual low critical threshold value for at least one sensor.	N/A	N/A
voltage_low_warn	STATE_CHANGE	WARNING	Voltage sensor {0} measured a low voltage value.	The voltage has fallen below the actual low warning threshold value for at least one sensor.	N/A	N/A
voltage_sensor_failed	STATE_CHANGE	WARNING	Voltage sensor {0} is failed.	The voltage sensor state is failed.	N/A	N/A
voltage_sensor_ok	STATE_CHANGE	INFO	Voltage sensor {0} is ok.	The voltage sensor state is ok.	N/A	N/A

Physical disk events

The following table lists the events that are created for the *Physical disk* component.

Table 5. Events for the physical disk component

Event	Event Type	Severity	Message	Description	Cause	User Action
gnr_nvram_degraded	STATE_CHANGE	WARNING	The NVDIMM of the pdisk {0} is degraded.	The NVRAM drive of the disk is in degraded state.	The **tslsnvramstatus** command shows degraded state for the NVRAM drive of the disk.	N/A
gnr_nvram_disarmed	STATE_CHANGE	ERROR	The NVDIMM of the pdisk {0} is disarmed.	NVDIMM is unable to preserve future content.	The **tslsnvramstatus** command reports disarmed failure condition for the NVRAM drive of the disk.	Identify the NVDIMM cards or BPM, which encountered the errors from FSP log or call home data, and replace the faulty NVDIMM cards, BPM or both as soon as possible.

Table 5. Events for the physical disk component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
gnr_nvram_erased	STATE_CHANGE	ERROR	The NVDIMM of the pdisk {0} reports erased image.	Image erased. The NVDIMM contents not persisted.	The **tslsnvramstatus** command reports `erased` failure condition for the NVRAM drive of the disk.	Verify that any NVDIMM cards, BPM encountered any errors from FSP log or call home data. If any errors are found then replace the faulty NVDIMM cards, BPM or both as soon as possible. If no errors are found then try to add the drive back to RG.
gnr_nvram_error	STATE_CHANGE	ERROR	The NVDIMM of the pdisk {0} is failed.	The NVRAM drive of the disk is in `error` state.	The **tslsnvramstatus** command shows `failed` state for the NVRAM drive of the disk.	N/A
gnr_nvram_ok	STATE_CHANGE	INFO	The NVDIMM of the pdisk {0} is normal.	The NVRAM drive of the disk is in `normal` state.	N/A	N/A
gnr_nvram_persist_error	STATE_CHANGE	ERROR	The NVDIMM of the pdisk {0} could not persist.	NVDIMM failed to save or restore the memory contents.	The **tslsnvramstatus** command reports `FailToPersist` failure condition for the NVRAM drive of the disk.	Identify the NVDIMM cards or BPM, which encountered the errors from FSP log or call home data, and replace the faulty NVDIMM cards, BPM or both as soon as possible.
gnr_nvram_unhealthy	STATE_CHANGE	WARNING	The NVDIMM of the pdisk {0} is unhealthy.	Error is detected but save or restore might still work for the NVRAM drive of the disk.	The **tslsnvramstatus** command reports `unhealthy` failure condition for the NVRAM drive of the disk.	Identify the NVDIMM cards or BPM, which encountered the errors from FSP log or call home data, and replace the faulty NVDIMM cards, BPM or both.
gnr_pdisk_degraded	WARNING	WARNING	GNR pdisk {0} is degraded.	The pdisk state is degraded.	N/A	N/A
gnr_pdisk_diagnosing	INFO	WARNING	GNR pdisk {0} is diagnosing.	The pdisk state is diagnosing.	N/A	N/A
gnr_pdisk_draining	STATE_CHANGE	ERROR	GNR pdisk {0} is draining.	The pdisk state is draining.	N/A	N/A
gnr_pdisk_disks	STATE_CHANGE	INFO	Pdisks found on this node.	Pdisks found		N/A
gnr_pdisk_found	INFO_ADD_ENTITY	INFO	GNR pdisk {0} was found.	A GNR pdisk listed in the IBM Spectrum Scale configuration was detected.	N/A	N/A
gnr_pdisk_maintenance	STATE_CHANGE	WARNING	GNR pdisk {0} is in maintenance.	The GNR pdisk is in maintenance because the state is either `suspended`, `serviceDrain`, `pathMaintenance` or `deleting`. This might be caused by some administration commands like **mmdeldisk**.	The **mmlspdisk** command displays `maintenance` user condition for the disk.	Complete the maintenance action. Contact IBM support if you are not sure how to solve this problem.
gnr_pdisk_missing	STATE_CHANGE	WARNING	GNR pdisk {0} is missing.	The pdisk state is missing.	N/A	N/A

Table 5. Events for the physical disk component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
gnr_pdisk_needanalysis	STATE_CHANGE	ERROR	GNR pdisk {0} needs analysis.	The GNR pdisk has a problem that has to be analyzed and solved by an expert.	The **mmlspdisk** command displays `attention` user condition for the disk.	Contact IBM support if you are not sure how to solve this problem.
gnr_pdisk_nodisks	STATE_CHANGE	INFO	No pdisks found on this node.	No pdisks found, but some pdisks are expected on recovery group nodes.	The **mmvdisk pdisk list** command returned no pdisks.	Run the **mmvdisk pdisk list** command to verify if this is correct.
gnr_pdisk_ok	STATE_CHANGE	INFO	GNR pdisk {0} is ok.	The pdisk state is ok.	N/A	N/A
gnr_pdisk_replaceable	STATE_CHANGE	ERROR	GNR pdisk {0} is replaceable.	The pdisk state is replaceable.	N/A	N/A
gnr_pdisk_sedlocked	STATE_CHANGE	ERROR	GNR pdisk {0} is locked (Self-encrypting drive).	A self-encrypting drive which has encryption enabled is locked. GNR does not have access to any data on the drive.	The **mmlspdisk** command shows that the pdisk state contains `sedLocked`.	The drive must be unlocked to be used by GNR.
gnr_pdisk_unknown	STATE_CHANGE	WARNING	GNR pdisks are in unknown state.	The pdisk state is unknown.	N/A	N/A
gnr_pdisk_vanished	INFO_DELETE_ENTITY	INFO	GNR pdisk {0} has vanished.	A GNR pdisk listed in the IBM Spectrum Scale configuration was not detected.	A GNR pdisk, listed in the IBM Spectrum Scale configuration as mounted before, is not found. This could be a valid situation.	Run the **mmlspdisk** command to verify that all expected GNR pdisk exist.
gnr_pdisk_vwce	STATE_CHANGE	ERROR	GNR pdisk {0} has volatile write cache enabled.	Volatile write cache is enabled on the drive. Already committed writes could be lost in case of power loss. GNR will read-only from this disk.	The **mmlspdisk** command shows that the pdisk state contains VWCE.	Check why the volatile write cache is enabled (e.g. new drive added with wrong default, wrong UDEV rules) and fix the modes using the **sg_wr_modes** command.
ssd_endurance_ok	STATE_CHANGE	INFO	The `ssdEndurancePercentage` of GNR pdisk {0} is ok.	The `ssdEndurancePercentage` value is ok.	N/A	N/A
ssd_endurance_warn	STATE_CHANGE	WARNING	The `ssdEndurancePercentage` of GNR pdisk {0} is on a warning value.	The `ssdEndurancePercentage` value is warning.	The `ssdEndurancePercentage` value of the pdisk is between 95 and 100.	SSDs have a finite lifetime based on the number of drive writes per day. The ssd-endurance-percentage values actually reported will be a number between 0 and 255. This value indicates the percentage of life that is used by the drive. The value 0 indicates that full life remains, and 100 indicates that the drive is at or past its end of life. The drive must be replaced when the value exceeds 100", "state":"DEGRADED " }.

Recovery group events

The following table lists the events that are created for the *Recovery group* component.

Table 6. Events for the Recovery group component

Event	Event Type	Severity	Message	Description	Cause	User Action
gnr_rg_failed	STATE_CHANGE	ERROR	GNR recoverygroup {0} is not active.	The recovery group is not active.	N/A	N/A
gnr_rg_found	INFO_ADD_ENTITY	INFO	GNR recovery group {0} was found.	A GNR recovery group listed in the IBM Spectrum Scale configuration was detected.	N/A	N/A
gnr_rg_ok	STATE_CHANGE	INFO	GNR recoverygroup {0} is ok.	The recovery group is ok.	N/A	N/A
gnr_rg_vanished	INFO_DELETE_ENTITY	INFO	GNR recovery group {0} has vanished.	A GNR recovery group listed in the IBM Spectrum Scale configuration was not detected.	A GNR recovery group, listed in the IBM Spectrum Scale configuration as mounted before, is not found. This could be a valid situation.	Run the **mmlsrecoverygroup** command to verify that all expected GNR recovery group exist.

Server events

The following table lists the events that are created for the *Server* component.

Server events

Table 7. Server events

Event	Event Type	Severity	Message	Description	Cause	User Action
cpu_peci_ok	STATE_CHANGE	INFO	PECI state of CPU {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
cpu_peci_failed	STATE_CHANGE	ERROR	PECI state of CPU {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
cpu_qpi_link_ok	STATE_CHANGE	INFO	QPI Link of CPU {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.

Table 7. Server events (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
cpu_qpi_link_failed	STATE_CHANGE	ERROR	QPI Link of CPU {0} is failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
cpu_temperature_ok	STATE_CHANGE	ERROR	QPI Link of CPU {0} is failed.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
cpu_temperature_ok	STATE_CHANGE	INFO	CPU {0} temperature is normal ({1}).	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
cpu_temperature_failed	STATE_CHANGE			The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_power_supply_ temp_ok	STATE_CHANGE	INFO	Temperature of Power Supply {0} is ok. ({1})	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_power_supply_ temp_failed	STATE_CHANGE	ERROR	Temperature of Power Supply {0} is too high. ({1})	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_power_supply_oc_line_ 12V_ok	STATE_CHANGE	INFO	OC Line 12V of Power Supply {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_power_supply_oc_line_ 12V_failed	STATE_CHANGE	ERROR	OC Line 12V of Power Supply {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_power_supply_ov_line_ 12V_ok	STATE_CHANGE	INFO	OV Line 12V of Power Supply {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_power_supply_ov_line_ 12V_failed	STATE_CHANGE	ERROR	OV Line 12V of Power Supply {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_power_supply_uv_line_ 12V_ok	STATE_CHANGE	INFO	UV Line 12V of Power Supply {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_power_supply_uv_line_ 12V_failed	STATE_CHANGE	ERROR	UV Line 12V of Power Supply {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_power_supply_aux_line_ 12V_ok	STATE_CHANGE	INFO	AUX Line 12V of Power Supply {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.

Table 7. Server events (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
server_power_supply_aux_line_ 12V_failed	STATE_CHANGE	ERROR	AUX Line 12V of Power Supply {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_power_supply_ fan_ok	STATE_CHANGE	INFO	Fan of Power Supply {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_power_supply_ fan_failed	STATE_CHANGE	ERROR	Fan of Power Supply {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_power_supply_ voltage_ok	STATE_CHANGE	INFO	Voltage of Power Supply {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_power_supply_ voltage_failed	STATE_CHANGE	ERROR	Voltage of Power Supply {0} is not ok.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_power_ supply_ok	STATE_CHANGE	INFO	Power Supply {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_power_ supply_failed	STATE_CHANGE	ERROR	Power Supply {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
pci_riser_temp_ok	STATE_CHANGE	INFO	The temperature of PCI Riser {0} is ok. ({1})	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
pci_riser_temp_failed	STATE_CHANGE	ERROR	The temperature of PCI Riser {0} is too high. ({1})	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_fan_ok	STATE_CHANGE	INFO	Fan {0} is ok. ({1})	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_fan_failed	STATE_CHANGE	ERROR	Fan {0} failed. ({1})	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
dimm_ok	STATE_CHANGE	INFO	DIMM {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
dimm_failed	STATE_CHANGE	ERROR	DIMM {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.

Table 7. Server events (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
pci_ok	STATE_CHANGE	INFO	PCI {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
pci_failed	STATE_CHANGE	ERROR	PCI {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
fan_zone_ok	STATE_CHANGE	INFO	Fan Zone {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
fan_zone_failed	STATE_CHANGE	ERROR	Fan Zone {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
drive_ok	STATE_CHANGE	INFO	Drive {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
drive_failed	STATE_CHANGE	ERROR	Drive {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
dasd_backplane_ok	STATE_CHANGE	INFO	DASD Backplane {0} is ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
dasd_backplane_failed	STATE_CHANGE	ERROR	DASD Backplane {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_cpu_ok	STATE_CHANGE	INFO	All CPUs of server {0} are fully available.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_cpu_failed	STATE_CHANGE	ERROR	At least one CPU of server {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_dimm_ok	STATE_CHANGE	INFO	All DIMMs of server {0} are fully available.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_dimm_failed	STATE_CHANGE	ERROR	At least one DIMM of server {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_pci_ok	STATE_CHANGE	INFO	All PCIs of server {0} are fully available.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.

Table 7. Server events (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
server_pci_failed	STATE_CHANGE	ERROR	At least one PCI of server {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_ps_conf_ok	STATE_CHANGE	INFO	All Power Supply Configurations of server {0} are ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_ps_conf_failed	STATE_CHANGE	ERROR	At least one Power Supply Configuration of server {0} is not ok.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_ps_heavyload _ok	STATE_CHANGE	INFO	No Power Supplies of server {0} are under heavy load.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_ps_heavyload _failed	STATE_CHANGE	ERROR	At least one Power Supply of server {0} is under heavy load.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_ps_resource _ok	STATE_CHANGE	INFO	Power Supply resources of server {0} are ok.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_ps_resource _failed	STATE_CHANGE	ERROR	At least one Power Supply of server {0} has insufficient resources.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_ps_unit_ok	STATE_CHANGE	INFO	All Power Supply units of server {0} are fully available.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_ps_unit_failed	STATE_CHANGE	ERROR	At least one Power Supply unit of server {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_ps_ambient_ok	STATE_CHANGE	INFO	Power Supply ambient of server {0} is ok.	The GUI checks the hardware state using xCAT.	he hardware part is ok.	None.
server_ps_ambient _failed	STATE_CHANGE	ERROR	At least one Power Supply ambient of server {0} is not okay.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_boot_status_ok	STATE_CHANGE	INFO	The boot status of server {0} is normal.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_boot_status _failed	STATE_CHANGE	ERROR	System Boot failed on server {0}.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.

Table 7. Server events (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
server_planar_ok	STATE_CHANGE	INFO	Planar state of server {0} is healthy, the voltage is normal ({1}).	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_planar_failed	STATE_CHANGE	ERROR	Planar state of server {0} is unhealthy, the voltage is too low or too high ({1}).	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_sys_board_ok	STATE_CHANGE	INFO	The system board of server {0} is healthy.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_sys_board _failed	STATE_CHANGE	ERROR	The system board of server {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_system_event _log_ok	STATE_CHANGE	INFO	The system event log of server {0} operates normally.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_system_event _log_full	STATE_CHANGE	ERROR	The system event log of server {0} is full.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
server_ok	STATE_CHANGE	INFO	The server {0} is healthy.	The GUI checks the hardware state using xCAT.	The hardware part is ok.	None.
server_failed	STATE_CHANGE	ERROR	The server {0} failed.	The GUI checks the hardware state using xCAT.	The hardware part failed.	None.
hmc_event	STATE_CHANGE	INFO	HMC Event: {1}	The GUI collects events raised by the HMC.	An event from the HMC arrived.	None.

Virtual disk events

The following table lists the events that are created for the *Virtual disk* component.

Table 8. Events for the virtual disk component

Event	Event Type	Severity	Message	Description	Cause	User Action
gnr_vdisk_critical	STATE_CHANGE	ERROR	GNR vdisk {0} is critical degraded.	The vdisk state is critical degraded.	N/A	N/A
gnr_vdisk_degraded	STATE_CHANGE	WARNING	GNR vdisk {0} is degraded.	The vdisk state is degraded.	N/A	N/A
gnr_vdisk_found	INFO_ADD_ENTITY	INFO	GNR vdisk {0} was found.	A GNR vdisk listed in the IBM Spectrum Scale configuration was detected.	N/A	N/A
gnr_vdisk_offline	STATE_CHANGE	ERROR	GNR vdisk {0} is offline.	The vdisk state is offline.	N/A	N/A
gnr_vdisk_ok	STATE_CHANGE	INFO	GNR vdisk {0} is ok.	The vdisk state is ok.	N/A	N/A

Table 8. Events for the virtual disk component (continued)

Event	Event Type	Severity	Message	Description	Cause	User Action
gnr_vdisk_unknown	STATE_CHANGE	WARNING	GNR vdisk {0} is unknown.	The vdisk state is unknown.	N/A	N/A
gnr_vdisk_vanished	INFO_DELETE_ENTITY	INFO	GNR vdisk {0} has vanished.	A GNR vdisk listed in the IBM Spectrum Scale configuration was not detected.	A GNR vdisk, listed in the IBM Spectrum Scale configuration as mounted before, is not found. This could be a valid situation.	Run the **mmlsvdisk** command to verify that all expected GNR vdisk exist.

Chapter 2. Servicing

Service information is intended for IBM authorized service personnel only. Consult the terms of your warranty to determine the extent to which you can attempt to accomplish any IBM ESS 3000 system maintenance.

IBM service support representatives and lab based services personnel can access service information through the following link.

Servicing (service personnel only)

Note: An IBM intranet connection is required.

Removing and replacing a drive

Use the following procedures to remove a faulty drive from a node canister and replace it with a new one from stock.

Attention:

- 
When you replace this part, you must follow recommended procedures for handling electrostatic discharge (ESD)-sensitive devices.
- No tools are required to complete this task. Do not remove or loosen any screws.
- This procedure requires access to either management GUI or CLI command as a root user. IBM service personnel need to coordinate with the customer to work on this procedure.
- You can identify the failed drive by the amber fault LED on the drive carrier. If the fault LED is lit on a drive, it is safe to replace the drive.

  You can also locate unhealthy drives in the management GUI, either from the **Storage** > **Physical Disks** page or from the list of events that are available under the **Monitoring** > **Events** page. You can also select the **Display unhealthy devices** option in the **Monitoring** > **Hardware Details** page to see all the unhealthy devices including the faulty disks.
- Every drive slot of an operational control enclosure must contain either a drive or a blank filler, and must not be left empty for more than 10 minutes during servicing. Ensure that you have read and understood all these instructions and have the replacement drive available and unpacked before you remove the existing drive.
- During the replacement process, the replacement drive is automatically checked to confirm if it is a valid FRU and has the same capacity as all other drives in this system. The drive firmware will be automatically upgraded, if required.

Safety preparation

Read the safety precautions in the *IBM Systems Safety Notices*. These guidelines help you safely work with the system.

Removing the drive

In ESS 3000 system, a faulty drive has amber fault LED ON (non-flashing) on the driver carrier. You can remove a drive using either the management GUI or using the CLI commands.

- **Option 1. Removing a drive using management GUI**

  To replace disks using the management GUI, follow the direct maintenance procedure (DMP) that is available in the *Replace disks* topic in the *Elastic Storage Server 5.3.5: Problem Determination Guide*.
- **Option 2. Removing a drive using CLI commands**

If one or more pdisks in the recovery group is marked for replacement, the following command reports it with a yes in the “needs service” column:

```
mmvdisk recoverygroup list
```

In the following example, the BB01L recovery group needs service:

```
# mmvdisk recoverygroup list

                                                          needs    user
recovery group  active   current or master server        service  vdisks  remarks
--------------  -------  -------------------------------  -------  ------  -------
BB01L           yes      server01.gpfs.net                yes           3
BB01R           yes      server02.gpfs.net                no            3
```

This happens when the number of failed pdisks in one of the recovery group's declustered arrays reaches or exceeds the replacement threshold for the declustered array.

Pdisks that have reached the threshold for replacement are listed with the following command:

```
mmvdisk pdisk list --replace
```

```
# mmvdisk pdisk list --recovery-group all --replace

recovery group  pdisk         priority  FRU (type)       location
--------------  ------------  --------  ---------------  --------
BB01L           e2s11             1.15  00W1240          Enclosure 2 Drive 11
BB01L           e3s01             1.15  00W1240          Enclosure 3 Drive 1

mmvdisk: A lower priority value means a higher need for replacement.
```

Preparing disks for replacement

1. Prepare each of the pdisk name entries for replacement with the following command:

   ```
   mmvdisk pdisk replace –-prepare –-recovery-group <rg name> --pdisk <pdisk name>
   ```

   To prepare pdisk e2s11 of recovery group BB01L for replacement, run the following command:

   ```
   # mmvdisk pdisk replace --prepare --recovery-group BB01L --pdisk e2s11
   mmvdisk: Suspending pdisk e2s11 of RG BB01L in location SX32901810-11.
   mmvdisk: Location SX32901810-11 is Enclosure 2 Drive 11.
   mmvdisk: Carrier released.
   mmvdisk:
   mmvdisk:   - Remove carrier.
   mmvdisk:   - Replace disk in location SX32901810-11 with type '00W1240'.
   mmvdisk:   - Reinsert carrier.
   mmvdisk:   - Issue the following command:
   mmvdisk:
   mmvdisk:   mmvdisk pdisk replace --recovery-group BB01L --pdisk 'e2s11'
   ```

   The drive associated with the pdisk name in the previous command should now have flashing amber fault LED to indicate it is safe to remove this drive.

Removing the disk physically

1. Press the blue touchpoint to unlock the latching handle, as shown in this figure.

Figure 1. Unlocking the drive and release latch

2. Lower the handle and slide the drive out of the enclosure, as shown in this figure.

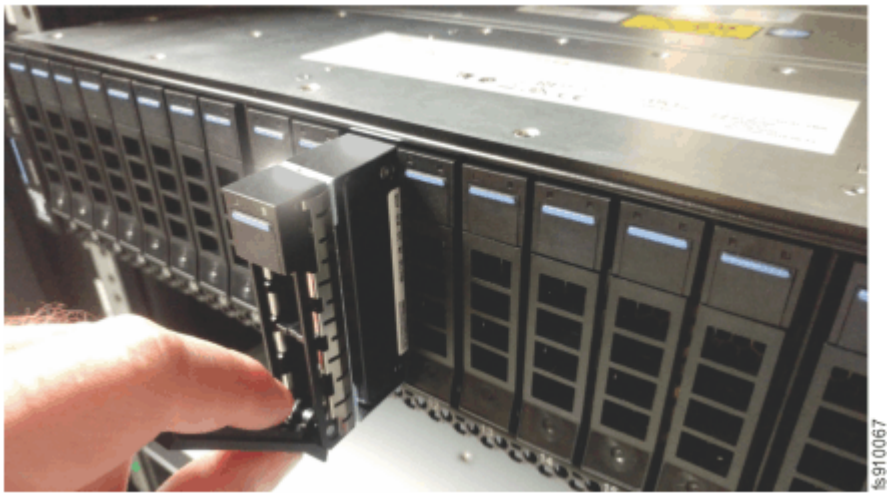

Figure 2. Removing the drive

Replacing the drive

1. Ensure that the LED indicators are at the top of the drive.
2. Press the blue touchpoint to unlock the latching handle on the new drive.
3. Slide the new drive into the node canister, as shown in the Figure 3 on page 22. Press on the drive label near the bottom of the drive to ensure that the drive is fully inserted into the slot.

Figure 3. Inserting the new drive

4. Finish inserting the new drive by closing the handle until the latch clicks into place.

Figure 4. Completing the drive installation

5. After the drive is replaced run the following command:

```
mmvdisk pdisk replace –-rg <rg name> --pdisk <pdisk name>
```

For example, finish replacing pdisk e2s11 with the new physical disk by running the following command:

```
# mmvdisk pdisk replace --recovery-group BB01L --pdisk e2s11
mmvdisk:
mmvdisk: Preparing a new pdisk for use may take many minutes.
mmvdisk:
mmvdisk: The following pdisks will be formatted on node ess01io1:
mmvdisk:       /dev/sdrk
mmvdisk:
mmvdisk: Location SX32901810-11 is Enclosure 2 Drive 11.

mmvdisk: Pdisk e2s11 of RG BB01L successfully replaced.
mmvdisk: Carrier resumed.
```

6. Repeat the steps listed in the Preparing disks for replacement, Removing the disk physically and Replacing the drive sections for each pdisk that needs to be replaced as marked in the output of the **mmvdisk pdisk list --replace** command.

Removing and replacing a drive blank

Use the following procedures to remove a faulty drive slot filler and replace it with a new one from stock. Drive slot fillers are passive components that regulate airflow through the control enclosure.

Notes:

- Every drive slot of an operational control enclosure must contain either a drive or a drive slot filler. A drive slot must not be left empty for more than **10 minutes** during servicing. Ensure that you have read and understood the removal and replacement instructions, and have the replacement part unpacked before you remove the existing drive slot filler.
- No tools are required to complete this task. Do not remove or loosen any screws.

1. Unpack the replacement drive slot filler from its packaging.

Removing the drive slot filler

2. Use your thumb and fore finger to pinch the latch of the faulty drive blank.

Removing a drive blank assembly

3. Gently slide the release latch up to unlock the handle.
4. Pull the faulty drive slot filler from the drive slot.

Replacing a drive blank assembly

5. Hold the drive blank the correct way up, as shown in Figure 5 on page 23.

Figure 5. Correct drive blank orientation

6. Slide the replacement drive blank into the empty drive slot.

Removing and replacing a power supply unit

You can remove and replace either of the two hot-swap redundant power supply units (PSUs) in an Elastic Storage System 3000 control enclosure. These redundant power supplies operate in parallel, one continuing to power the enclosure if the other fails.

Notes:

- This procedure requires access to either management GUI or CLI command as a root user. IBM service personnel need to coordinate with the customer to work on this procedure.

- Do not insert a PSU if the PSU slot does not contain a power interposer.
- Do not operate the enclosure without a power interposer and PSU in a PSU slot for longer than **5 minutes**. Operating for longer than this period might cause the enclosure to shut down due to overheating.
- In case a PSU fails, you should expect the following LED indication:
  - Enclosure LED amber fault is ON (non-flashing)
  - Amber fault LED on the faulty PSU is also ON
- In the management GUI, you can access the details of the power supply units from the **Monitoring** > **Hardware Details** page as shown in the Figure 6 on page 24:

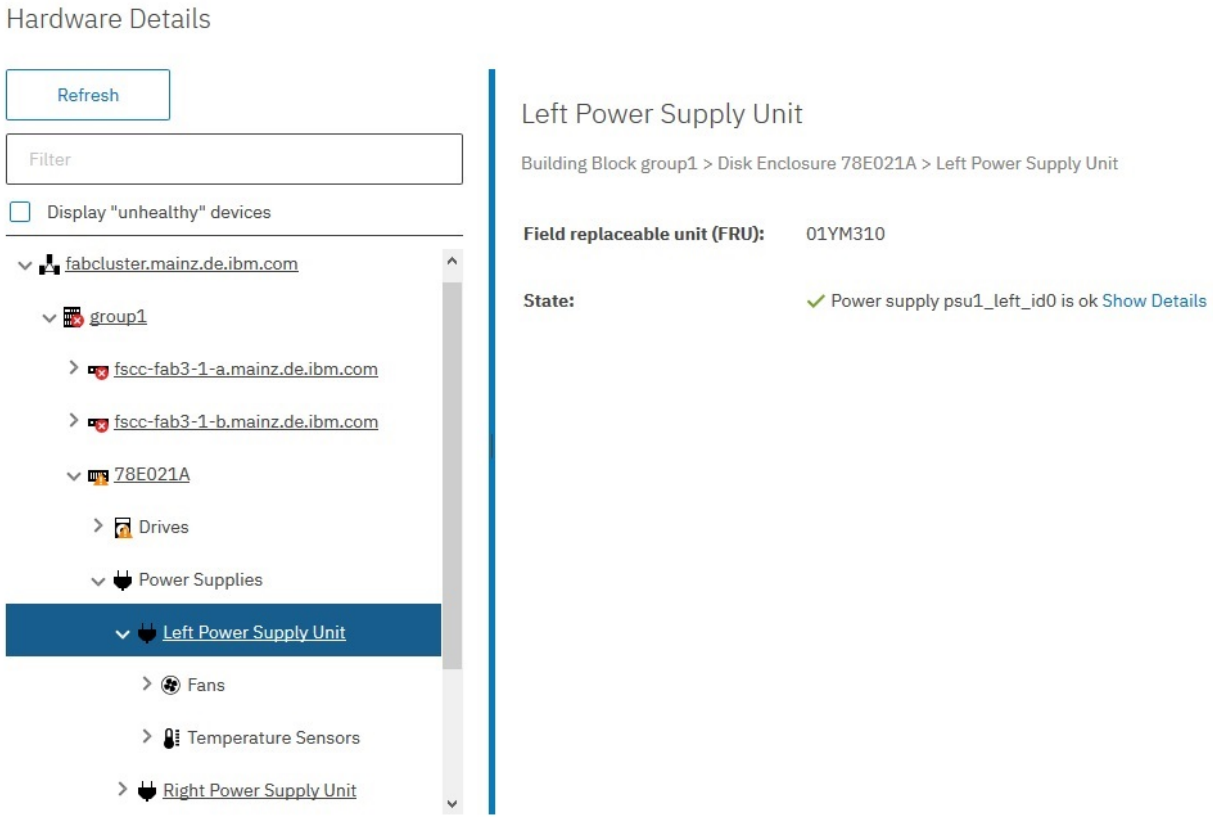

Figure 6. Details of Power Supply Units in the management GUI

Two sets of power supply units are available for each enclosure.

- Remove the replacement PSU from its packaging and have it available before carrying out this procedure.
- No tools are required to complete this task. Do not remove or loosen any screws.
- Although many components are hot-swappable, they are intended to be used only when your system is not active (no I/O operations).
- Be careful when you are replacing the hardware components that are located in the back of the system. Do not inadvertently disturb or remove any cables that you are not instructed to remove.
- 

When you replace this part, you must follow recommended procedures for handling electrostatic discharge (ESD)-sensitive devices.

1. In the management GUI, you can identify the faulty PSU from the **Monitoring** > **Hardware** page.

You can also run the **mmhealth node show enclosure** command on the canister of the affected enclosure.

To identify the affected enclosure, run the **mmhealth cluster show enclosure** command. The faulty enclosure will be in an DEGRADED or FAILED state.

A faulty power supply would be indicated by the power_supply_failed, power_supply_absent, power_high_power, power_high_current or power_no_power events. In the given example, the event message shows the psu1_left_id0 power supply needs to be replaced:

```
mmhealth node show enclosure

Node name:      fab3-1-b.example.com

Component     Status        Status Change     Reasons
-------------------------------------------------------------------------------------------
-----------------------------------------------
ENCLOSURE     DEGRADED      1 day ago         power_supply_failed(78E021A)
  78E021A     DEGRADED      1 day ago         power_supply_failed(78E021A)


Event                       Parameter     Severity    Active Since      Event Message
-------------------------------------------------------------------------------------------
------------------------------
power_supply_failed         78E021A       WARNING     Now               Power supply
psu1_left_id0 is failed
```

Removing the PSU

2. Release the cable retention clip and disconnect the power cord from the power supply unit that you are replacing.

   This figure shows the location of the cable retention clip 1, LED indicator 2, PSU 1 release tab 3, PSU handle 4, and the power interposer release tab 5.

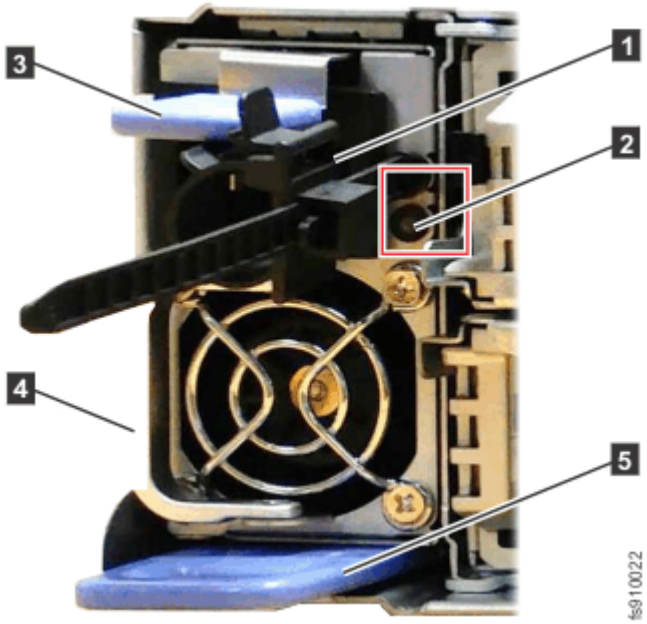

Figure 7. Features of a power supply unit

3. Fold out the handle so that it extends perpendicular to the PSU.
4. Holding the PSU handle, press and hold the PSU release tab and steadily pull the handle horizontally to slide the PSU from the enclosure, as shown in this figure.

   Support the PSU with your other hand as it is released from the enclosure.

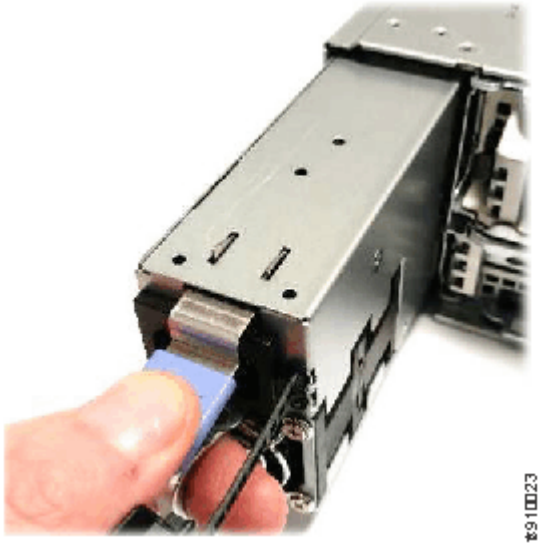

Figure 8. Removing the power supply unit

Important: Insert the replacement PSU within 5 minutes; otherwise, the system may overheat and shut down.

Replacing the PSU

5. Fold out the handle so that it extends perpendicular to the PSU.
6. While keeping the PSU handle extended and supporting the PSU, slide the power supply into the enclosure until the release tab engages with a "click".
7. Connect the power cord to the power supply and to a properly grounded electrical outlet. Secure the cable with the cable retention clip on the rear of the power supply unit.

   Note: After the power cord is connected to the electrical outlet, make sure that the LED indicator, shown in Figure 7 on page 25, is lit.

Removing and replacing a power interposer

A faulty power interposer can be removed and replaced with an identical or equivalent replacement part from FRU stock.

A power interposer forms part of each power supply unit (PSU) slot and fills the space between the PSU and the midplane. It can only be removed after its PSU is removed from the rear of the enclosure. Before you remove or replace a power interposer, review the following guidelines for this procedure:

- Ensure that you identify the correct PSU and power interposer for removal. If a PSU or a power interposer has a fault that prevents it from powering the enclosure and the functional power supply is removed, the node canisters in the control enclosure will shut down.

  In case a PSU or Power Interposer fails, you will see the following LED indication:

  – Enclosure LED amber fault is ON (non-flashing)
  – Amber fault LED on the faulty PSU is also ON
- Replace a power interposer only after the PSU replacement has failed to eliminate the fault condition.
- Do not insert a PSU into the slot while the power interposer is removed.

  Warning: Do not operate the enclosure without a power interposer and PSU in a PSU slot for longer than **5 minutes**. Operating for longer than this period might cause the control enclosure to shut down due to overheating.
- No tools are required to complete this task. Do not remove or loosen any screws.

1. Remove the power supply unit, as described in “Removing and replacing a power supply unit” on page 23.

Removing the power interposer

2. Remove the power interposer by pulling on the blue handle that is located beneath the PSU slot. Figure 9 on page 27 shows an example.

Figure 9. Sliding out the power interposer

3. Slide the power interposer out until it is clear of the enclosure rear, as shown in Figure 10 on page 27.

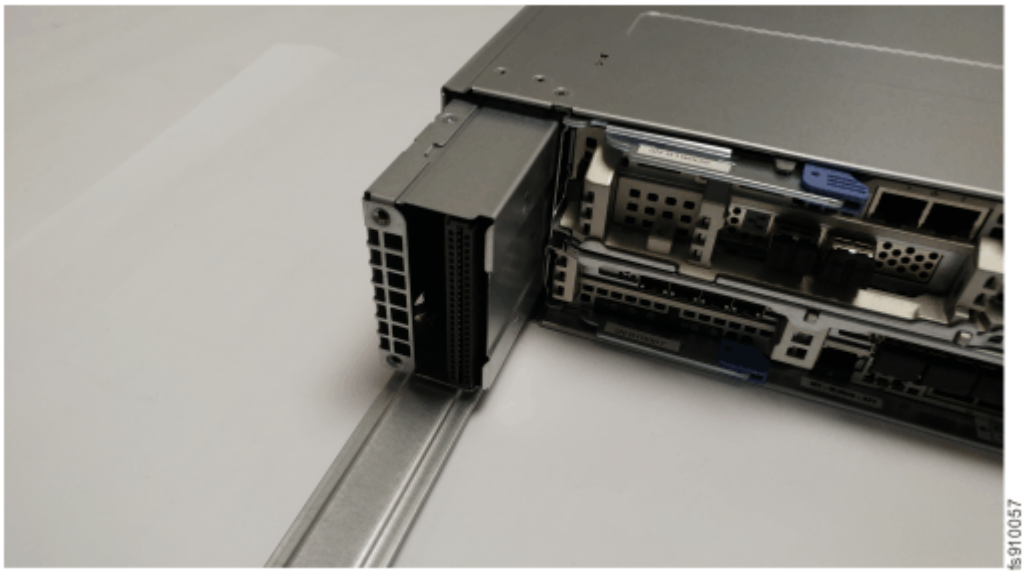

Figure 10. Removing a power interposer

Replacing the power interposer

4. Identify the correct empty power slot where the power interposer is to be installed.
5. Hold the power interposer so that the connectors are near the PSU slot, as shown in Figure 11 on page 28 (1).
6. Slide the power interposer into the PSU slot until the handle is the only part that is exposed at the rear of the enclosure, as shown in Figure 11 on page 28 (2).

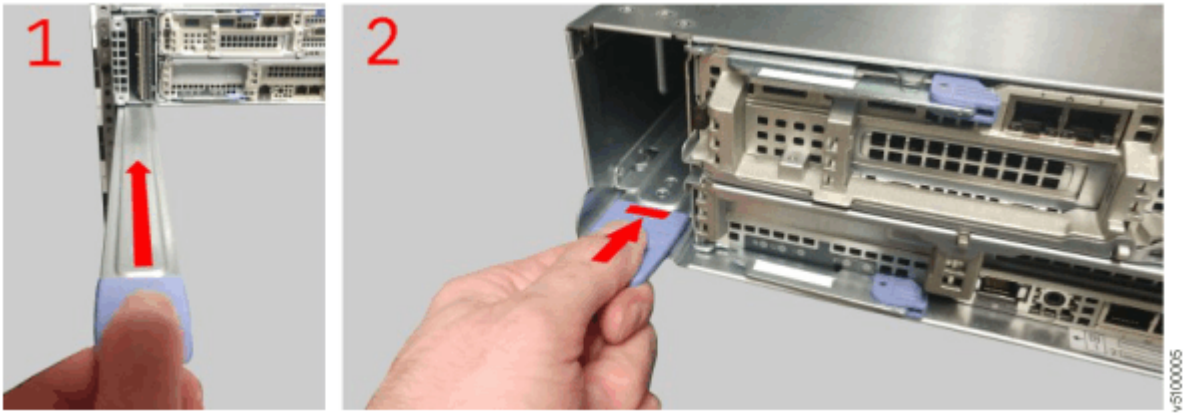

Figure 11. Inserting the new power interposer

7. Replace the PSU that was removed in step “1” on page 27. Follow the procedure that is described in “Removing and replacing a power supply unit” on page 23.
8. Reconnect all cables.

Miscellaneous equipment specification (MES) instructions

Any ESS 3000 hardware change, which can be an addition, improvement, removal, or any combination of three can be done by using the MES instructions.

ESS 3000 storage drives MES upgrade

An offline IBM Elastic Storage System 3000 (ESS 3000) MES upgrade is supported for customers who want to upgrade a 12-drive ESS 3000 to a 24-drive ESS 3000.

To upgrade the system, the NVMe drives with the same size as the existing 12 drives must be used. This MES upgrade makes the available storage double in the existing ESS 3000. For this offline procedure, a planned downtime is required. Existing customer recovery group and file system data are preserved before the MES upgrade.

Supported path: 12 NVMe ESS 3000 -> 24 NVMe ESS 3000

Prerequisites

- All new or existing building blocks must be at the ESS 5.3.5.1 or ESS 3000 6.0.0.1 level. If the setup has any protocol nodes, these nodes must also be upgraded to ESS 5.3.5.1 levels (underlying code + IBM Spectrum Scale 5.0.4.2 verified by using the **gssinstallcheck** command).
- The system must be healthy before the ESS 3000 storage MES upgrade.
- The existing ESS 3000 must be a properly installed 12 NVMe system with 12 NVMe drives correctly located in slots 1-6 and 13-18.
- All file systems that use the ESS 3000 that is being upgraded must be unmounted.
- All cluster functions, for example, quorum or manager, stop on the two canister servers. If the canister servers need to be up to maintain cluster quorum, then either other quorum nodes must be defined and made available or GPFS on the whole cluster must be shut down.
- Unlike a regular ESS MES upgrade, you need not to ensure that all recovery groups in the cluster are converted to the **mmvdisk** management because the **mmvdisk** management is already required for ESS 3000.
- LBS must wear an ESD wrist band when they work on the hardware, for example, inserting NVMe drives.

MES upgrade considerations

- Do not try to configure call home before MES is complete, that is, until the resizing is done.

- When the resizing is done and the upgraded ESS 3000 is back online, you can perform other ESS and GPFS operations.

Note: GPFS uses preferentially the new network shared disks (NSDs) to store data of a new file system. GPFS has four new NSDs that are the same as the four original NSDs, the workload per server is the same as it was before. The new file system data goes to the four new NSDs, like before the resizing, the original file system data goes to the four original NSDs.

Consider the necessity of restriping and the current demands on the system. New data that is added to the file system is correctly striped. Restriping a large file system requires many insert operations and delete operations and might affect system performance. Plan to perform this task when system demand is low.

MES upgrade steps

1. Ensure that the technical delivery assessment (TDA) process is complete before you start the MES upgrade.
2. Ensure that the system is at the ESS 3000 6.0.0.1 level for the storage MES.
3. Ensure that neither canister server of the ESS 3000 building block participates as a quorum node.
4. Ensure that I/O to the affected file systems is stopped, and the file systems that use the ESS 3000 are unmounted.
5. Ensure that GPFS is not set to autoload.
6. Ensure that all client nodes unmount the ESS 3000 file system.
7. Ensure that the automount is disabled on the file systems and the remote clusters.
8. Issue the **mmshutdown** command on the ESS 3000 canister servers.
9. Power off the ESS 3000 by removing the cables that are at the back of the storage enclosure.
10. Ensure that the 12 new NVMe drives are of the same size as the original 12 NVMe drives.
11. Insert the 12 new NVMe drives into slots 7 - 12 and 19 - 24.
12. Do not move the original 12 NVMe drives to different slots!
13. Power on the ESS 3000 and restart the OS on the canister servers. GPFS must not start because the autostart is turned off.
14. Verify the new 24 NVMe disk topology on both canisters by issuing the following command:

    ```
    # mmvdisk server list --disk-topology --node-class this ESS 3000 node class
    ```

    - Both canisters must show a 24 NVMe disk topology.
    - If any errors such as new disk is bad or does not show up are reported, the errors must be fixed before proceeding.
15. When GPFS is down on both canisters, issue the following command on one of the canisters to check whether the new drives are not at the latest level from the expected level:

    ```
    # mmlsfirmware --type drive
    ```

16. If the drive firmware is not at the latest level, issue the following command from one of the canisters to update:

    ```
    # mmchfirmware --type drive
    ```

    a. After the **mmchfirmware** command completes, verify that the drive firmware levels are correct by issuing the following command again:

    ```
    # mmchfirmware --type drive
    ```

    b. If some drives failed to update to the current level, rerun the **mmchfirmware** command to complete the process.

17. Start GPFS on the two canisters (or the whole cluster if needed for quorum).
18. Issue the following command on the affected recovery group. This command adds the new capacity, and rebalancing data begins onto the new disks.

    ```
    # mmvdisk recoverygroup resize --recovery-group ess3k recoverygroup
    ```

19. Update the node class server configuration for 24 drives by issuing the following command:

    ```
    # mmvdisk server configure --update --recycle 1 --node-class this ESS 3000 node class
    ```

20. If the GPFS autostart was enabled and had to be disabled, enable autostart again. Also, if the automount was disabled, you can enable the automount again. You can resume normal cluster operations after the GPFS autostart and the automount.

The customer can use the new space by creating new vdisksets from the available space. When the vdisksets are added to the existing file system (if required), the restripe operation can be run. For more information, see *IBM Spectrum Scale: Administration Guide*.

ESS 3000 adapter MES upgrade: EC64 and EC67

Overview

This procedure is intended for the concurrent (online) adapter MES installation to add the third adapter to each of the server canisters of ESS 3000. The adapter options to choose from include EC64 (InfiniBand) and EC67 (Ethernet).

Because the adapter MES is designed to be a concurrent procedure, it is recommended to perform this procedure during a service window of low activity.

Objectives

- Install the new adapter pair, one in each server node canister, by using an online process.
- Incorporate the new interfaces into an existing single master bond.

High-level MES upgrade steps

1. Prepare canister A (bottom canister slot) for adapter MES.
2. Unmount file systems that are mounted on canister A.
3. Shut down GPFS only in canister A.
4. Remove canister A from the enclosure.
5. Install the hardware – add an adapter to the available middle slot in canister A.
6. Insert canister A into the enclosure again and do basic checks.
7. Incorporate new interfaces to the existing single master bond in canister A.
8. Start GPFS in canister A.
9. Mount file systems on canister A again.
10. Prepare canister B (top canister slot) for an adapter MES.
11. Unmount file systems that are mounted on canister B.
12. Shut down GPFS only in canister B.
13. Remove canister B from the enclosure.
14. Install the hardware – add an adapter to the available middle slot in canister B.
15. Reinsert canister B into the enclosure and do basic checks.
16. Incorporate new interfaces to the existing single master bond in canister B.
17. Start GPFS in canister B.

18. Mount file systems on canister B.

- Steps 1–3 are expected to be customer task
- Steps 4–6 are SSR task
- Steps 7–12 are expected to be customer task
- Steps 13–15 are SSR task
- Steps 16–18 are expected to be customer task

Summary

The goal of this procedure is primarily to add a third high-speed adapter into each ESS 3000 canister. Customer can add supported InfiniBand or Ethernet adapters into the third PCI slot.

- The PCI address is af:00.1
- The adapter type is ConnectX-5 [ConnectX-5 Ex]

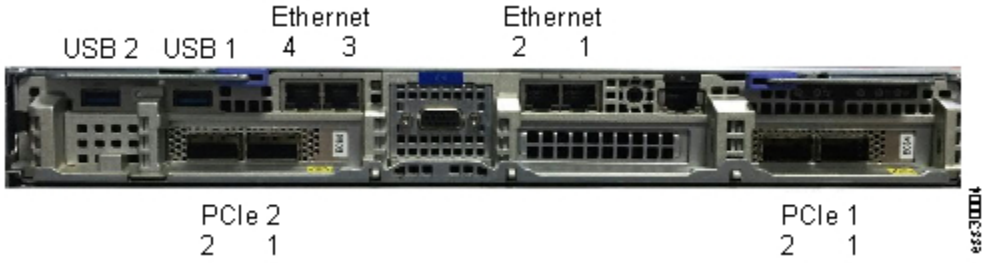

Figure 12. Ethernet ports on canister 1 (upper canister)

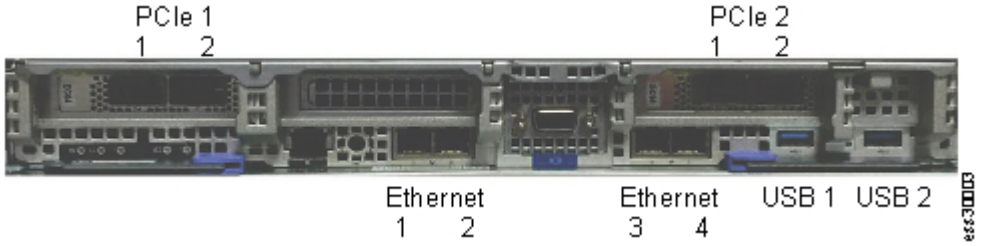

Figure 13. Ethernet ports on canister 2 (lower canister)

Note: These images show the PCIe ports for two adapters for each canister. The MES upgrade installs the third adapter with two more ports in the PCIe slot between the original two adapters.

Online adapter MES procedure

1. Prepare canister A (bottom canister slot) for an adapter MES. (Customer task)
   a. Determine the node class of the ESS 3000 enclosure where you want to perform MES.

      ```
      # mmvdisk nodeclass list
      ```

   b. Display the state of the `nodeclass` associated with the ESS 3000.

      ```
      # mmgetstate -N the ESS 3000 nodeclass
      ```

      Example:

      ```
      [root@ess3k5a ~]# mmvdisk nodeclass list

      node class                  recovery groups
      --------------------------  --------------------------
      ece_nc_1                    ece_rg_1
      ess_x86_64_mmvdisk_78E016N  ess3k_78E016N
      ess_x86_64_mmvdisk_78E05N1  ess3k_78E05N1
      ```

```
gssio1_ibgssio2_ib          rg_gssio1-ib, rg_gssio2-ib

[root@ess3k5a ~]# mmgetstate -N ess_x86_64_mmvdisk_78E05N1

 Node number  Node name        GPFS state
------------------------------------------
      21      ess3k5a-ib       active
      22      ess3k5b-ib       active
```

c. If canister A is currently designated as a quorum node, move the quorum to a different node.

1) To determine which nodes are quorum nodes, issue the following command:

```
# mmlscluster
```

2) To evaluate quorum status, issue the following command:

```
# mmgetstate -s
```

3) If necessary, to designate a different node to be a quorum node, issue the following command:

```
# mmchnode –quorum -N node not in this enclosure's nodeclass
```

4) If necessary, to remove a quorum node from canister A, issue the following command:

```
# mmchnode –noquorum -N canister A
```

2. Unmount file systems that are mounted on canister A. (Customer task)

   Unmount GPFS file systems from canister A.

```
# mmumount fs1 -N canister A
```

   or you can unmount all file systems at a time from canister A:

```
# mmumount all -N canister A
```

   Example:

```
[root@ess3k5a]# mmumount fs3k70 -N ess3k5a
```

3. Shut down GPFS only in canister A. (Customer task)

```
# mmshutdown -N canister A
```

   Example:

```
[root@ess3k5a]# mmshutdown -N ess3k5a
```

4. Remove canister A from the enclosure. (SSR task)

   Record the high speed and management network cabling to canister A. After adapter MES is done, the network cabling needs to be done as recorded here.

   Disconnect high speed and management network cables from canister A ONLY and remove the canister from the bottom slot of the enclosure.

   Note: Do not unplug power cables to either PSU.

5. Install the hardware – add an adapter to the available middle slot in canister A. (SSR task)

   Remove top cover of canister A and install one adapter in the 3rd PCIe slot that is located between the two existing adapters. For more information about these steps, see the *IBM Elastic Storage System 3000: Service Guide*.

6. Insert canister A into the enclosure again and do basic checks. (SSR task)

   a. Put back the top cover of canister A and insert it back to its original canister slot (bottom).

   b. Reconnect original high speed and management network cables.

c. Connect the new high speed cables to the new adapter if provided by the customer.

d. Perform basic checks via the technician port in canister A by using the “essserv1” service ID.

e. Perform the following checks in the **essutils** menu:

1) Option **2. Quick storage configuration check**

2) Option **3. Check enclosure cabling and paths to disks**

Both options must be successful. Option **2.** must show six network adapters.

Example:

```
localhost: Valid Network Adapter Configuration. Number of Adapter(s) found: 6
These checks are run per canister.
```

7. Incorporate new interfaces to the existing single master bond in canister A. (Customer task)

a. Verify two new ports, which are also called devices or interfaces, are recognized.

```
# ip a
```

```
# nmcli d
```

Example:

```
[root@ess3k5b ~]# nmcli d
DEVICE TYPE STATE CONNECTION bond0 bond connected bond-bond0 enp29s0f0 ethernet
connected enp29s0f0
ib0 infiniband connected bond-slave-ib0
ib1 infiniband connected bond-slave-ib1
ib2 infiniband connected bond-slave-ib2
ib3 infiniband connected bond-slave-ib3
ib4 infiniband connected bond-slave-ib4
ib5 infiniband connected bond-slave-ib5
enp1s0 ethernet unavailable -- enp29s0f1 ethernet unavailable -- enp29s0f2 ethernet
unavailable -- enp29s0f3 ethernet unavailable -- lo loopback unmanaged --
```

b. If this MES upgrade for EC64 (InfiniBand), issue the following commands:

```
# nmcli c add type infiniband ifname new interface name #1 master bond name
```

```
# nmcli c add type infiniband ifname new interface name #2 master bond name
```

Example:

```
[root@ess3k5b ~]# nmcli c add type infiniband ifname ib4 master bond0
Connection 'bond-slave-ib4' (30d93d60-0846-46c6-8baf-53e1093e67eb) successfully added.
[root@ess3k5b ~]# nmcli c add type infiniband ifname ib5 master bond0
Connection 'bond-slave-ib5' (a1c3bdfc-1b49-49a5-8966-19d6e91c4c56) successfully added.
```

c. If this MES upgrade for EC67 (Ethernet), issue the following commands:

```
# nmcli c add type ethernet ifname new interface name #1 master bond name
```

```
# nmcli c add type ethernet ifname new interface name #2 master bond name
```

d. To validate the newly added interfaces that are part of the bond, issue the following command:

```
# nmcli c
```

Example:

```
[root@ess3k5a ~]# nmcli c
NAME             UUID                                   TYPE        DEVICE
enp29s0f0        ada957e2-feba-d40d-5e1d-5639f49aa2d4   ethernet    enp29s0f0
bond-bond0       e26c9f46-7df6-45f8-a4d7-d322838d34da   bond        bond0
bond-slave-ib0   4c318f6d-d18c-45a6-b76b-9ec5daafd971   infiniband  ib0
bond-slave-ib1   4bd92a1e-278f-4282-bcb2-1d7b85ab5510   infiniband  ib1
bond-slave-ib2   f737ea73-27bb-4a41-b3e7-4862bac41f15   infiniband  ib2
```

```
bond-slave-ib3   c55b57b7-d5c7-4705-a4b1-909a6ea53ec2  infiniband  ib3
bond-slave-ib4   6e0e78f1-1425-4546-be66-efedaae6aa90  infiniband  ib4
bond-slave-ib5   ace1cc7d-ff25-4503-a41b-1ec5806162d0  infiniband  ib5
enp1s0           157075f3-08e7-4d87-91f4-ff90141f8e84  ethernet    --
enp29s0f1        e2cc75e6-a2d8-45f4-8010-56f92c50158b  ethernet    --
enp29s0f2        ce01d41e-24d1-41fe-b038-0298e7565123  ethernet    --
enp29s0f3        e6f8b4e8-93d5-4008-9598-9878d8def28c  ethernet    --
ethernet-enp1s0  8d5550ae-52e4-4a2e-8037-e09c2df60dbc  ethernet    --
servport         28be0740-4dfd-4646-8f06-331de13f2c7f  ethernet    --
```

8. Start GPFS in canister A. (Customer task)

```
# mmstartup -N canister A
```

   a. Verify that GPFS is active on both A and B canisters.

```
# mmgetstate -N this ESS 3000 nodeclass
```

   b. Do the following extra steps on canister A, if MES is for EC64 (InfiniBand):

      1) Update the `verbsPorts` list, first start GPFS manually.

      2) Ensure that correct entries are listed in `verbsPorts` for the target node class by issuing the following command:

```
/opt/ibm/ess/tools/samples/essServerConfig.sh node class name
```

```
# mmlsconfig -Y | grep -i verbsPort
```

         Example:

```
[root@ess3k5a ~]# mmlsconfig -Y | grep -i verbsPort
mmlsconfig::0:1:::verbsPorts:mlx5_1/1::
mmlsconfig::0:1:::verbsPorts:mlx5_0/1 mlx5_1/1:bodhi1-ib,gssio2-ib,bodhi_nc1,ems:
mmlsconfig::0:1:::verbsPorts:mlx5_0/1 mlx5_1/1 mlx5_2/1 mlx5_3/1:ess_x86_64:
mmlsconfig::0:1:::verbsPorts:mlx5_2/1 mlx5_3/1 mlx5_4/1
mlx5_5/1:ess_x86_64_mmvdisk_5:
mmlsconfig::0:1:::verbsPorts:mlx5_2/1 mlx5_3/1:ess_x86_64_mmvdisk_6:
mmlsconfig::0:1:::verbsPorts:mlx5_0/1:gss_ppc64:
```

9. Mount file systems to canister A again. (Customer task)

   a. Check mounted file systems and mount to canister A again, if necessary.

```
# mmlsmount fs1 -L
```

   or

```
# mmlsmount all -L
```

   b. Move the quorum node back to the original state, if necessary.

10. Prepare canister B (top canister slot) for an adapter MES. (Customer task)

   a. Determine the node class of the ESS 3000 enclosure where you want to perform MES upgrade.

```
# mmvdisk nodeclass list
```

   b. Display the state of the `nodeclass` associated with the ESS 3000.

```
# mmgetstate -N the ESS 3000 nodeclass
```

      Example:

```
[root@ess3k5a ~]# mmvdisk nodeclass list

node class                 recovery groups
-------------------------  -------------------------
ece_nc_1                   ece_rg_1
ess_x86_64_mmvdisk_78E016N  ess3k_78E016N
ess_x86_64_mmvdisk_78E05N1  ess3k_78E05N1
gssio1_ibgssio2_ib         rg_gssio1-ib, rg_gssio2-ib
```

```
[root@ess3k5a ~]# mmgetstate -N ess_x86_64_mmvdisk_78E05N1

 Node number  Node name        GPFS state
------------------------------------------
      21      ess3k5a-ib       active
      22      ess3k5b-ib       active
```

c. If canister B is currently designated as a quorum node, move the quorum to a different node.

1) To determine which nodes are quorum nodes, issue the following command:

```
# mmlscluster
```

2) To evaluate quorum status, issue the following command:

```
# mmgetstate -s
```

3) If necessary, to designate a different node to be a quorum node, issue the following command:

```
# mmchnode –quorum -N node not in this enclosure's nodeclass
```

4) If necessary, to remove a quorum node from canister B, issue the following command:

```
# mmchnode –quorum -N canister B
```

11. Unmount file systems that are mounted on canister B. (Customer task)

Unmount GPFS file systems from canister B:

```
# mmumount fs1 -N canister B
```

or you can unmount all file systems at a time from canister B.

```
# mmumount all -N canister B
```

Example:

```
[root@ess3k5a]# mmumount fs3k70 -N ess3k5
```

12. Shut down GPFS only in canister B. (Customer task)

```
# mmshutdown -N canister B
```

Example:

```
[root@ess3k5b]# mmshutdown -N ess3k5b
```

13. Remove canister B from the enclosure. (SSR task)

Record the high speed and management network cabling to canister B. After adapter MES is done, the network cabling needs to be done as recorded here.

Disconnect high speed and management network cables from canister B ONLY and remove the canister from the top slot of the enclosure.

Note: Do not unplug power cables to either PSU.

14. Install the hardware – add an adapter to the available middle slot in canister B. (SSR task)

Remove top cover of canister B and install an adapter in the 3rd PCIe slot that is located between two existing adapters. For more information about these steps, see *IBM Elastic Storage System 3000: Service Guide*.

15. Insert canister B into the enclosure again and do basic checks. (SSR task)

a. Put back the top cover of canister B and insert it back to its original canister slot (top).

b. Reconnect original high speed and management network cables.

c. Connect the new high speed cables to the new adapter if provided by the customer.

d. Perform basic checks via the technician port in canister B using the “essserv1” service ID.

e. Perform the following checks in the **essutils** menu:

1) Option **2. Quick storage configuration check**

2) Option **3. Check enclosure cabling and paths to disks**

Both options must be successful. Option **2.** must show six network adapters.

Example:

```
localhost: Valid Network Adapter Configuration. Number of Adapter(s) found: 6
These checks are run per canister.
```

16. Incorporate new interfaces to the existing single master bond in canister B. (Customer task)

a. Verify two new ports, which are also called devices or interfaces, are recognized.

```
# ip a
```

```
# nmcli d
```

b. If this MES upgrade for EC64 (InfiniBand), issue the following commands:

```
# nmcli c add type infiniband ifname new interface name #1 master bond name
```

```
# nmcli c add type infiniband ifname new interface name #2 master bond name
```

Example:

```
[root@ess3k5b ~]# nmcli c add type infiniband ifname ib4 master bond0
Connection 'bond-slave-ib4' (30d93d60-0846-46c6-8baf-53e1093e67eb) successfully added.
[root@ess3k5b ~]# nmcli c add type infiniband ifname ib5 master bond0
Connection 'bond-slave-ib5' (a1c3bdfc-1b49-49a5-8966-19d6e91c4c56) successfully added.
```

c. If this MES upgrade for EC67 (Ethernet), issue the following commands:

```
# nmcli c add type ethernet ifname new interface name #1 master bond name
```

```
# nmcli c add type ethernet ifname new interface name #2 master bond name
```

d. To validate the newly added interfaces that are part of the bond, issue the following command:

```
# nmcli c
```

Example:

```
[root@ess3k5a ~]# nmcli c
NAME             UUID                                   TYPE        DEVICE
enp29s0f0        ada957e2-feba-d40d-5e1d-5639f49aa2d4   ethernet    enp29s0f0
bond-bond0       e26c9f46-7df6-45f8-a4d7-d322838d34da   bond        bond0
bond-slave-ib0   4c318f6d-d18c-45a6-b76b-9ec5daafd971   infiniband  ib0
bond-slave-ib1   4bd92a1e-278f-4282-bcb2-1d7b85ab5510   infiniband  ib1
bond-slave-ib2   f737ea73-27bb-4a41-b3e7-4862bac41f15   infiniband  ib2
bond-slave-ib3   c55b57b7-d5c7-4705-a4b1-909a6ea53ec2   infiniband  ib3
bond-slave-ib4   6e0e78f1-1425-4546-be66-efedaae6aa90   infiniband  ib4
bond-slave-ib5   ace1cc7d-ff25-4503-a41b-1ec5806162d0   infiniband  ib5
enp1s0           157075f3-08e7-4d87-91f4-ff90141f8e84   ethernet    --
enp29s0f1        e2cc75e6-a2d8-45f4-8010-56f92c50158b   ethernet    --
enp29s0f2        ce01d41e-24d1-41fe-b038-0298e7565123   ethernet    --
enp29s0f3        e6f8b4e8-93d5-4008-9598-9878d8def28c   ethernet    --
ethernet-enp1s0  8d5550ae-52e4-4a2e-8037-e09c2df60dbc   ethernet    --
servport         28be0740-4dfd-4646-8f06-331de13f2c7f   ethernet    --
```

17. Start GPFS in canister B. (Customer task)

```
# mmstartup -N canister B
```

a. Verify that GPFS is active on A and B canisters.

```
# mmgetstate -N this ESS 3000 nodeclass
```

b. Do the following extra steps on canister B, if MES is for EC64 (InfiniBand):

1) Update the `verbsPorts` list, first start GPFS manually.

2) Ensure that correct entries are listed in `verbsPorts` for the target node class by issuing the following command:

```
/opt/ibm/ess/tools/samples/essServerConfig.sh node class name
```

```
# mmlsconfig -Y | grep -i verbsPort
```

Example:

```
[root@ess3k5a ~]# mmlsconfig -Y | grep -i verbsPort
mmlsconfig::0:1:::verbsPorts:mlx5_1/1::
mmlsconfig::0:1:::verbsPorts:mlx5_0/1 mlx5_1/1:bodhi1-ib,gssio2-ib,bodhi_nc1,ems:
mmlsconfig::0:1:::verbsPorts:mlx5_0/1 mlx5_1/1 mlx5_2/1 mlx5_3/1:ess_x86_64:
mmlsconfig::0:1:::verbsPorts:mlx5_2/1 mlx5_3/1 mlx5_4/1 mlx5_5/1:ess_x86_64_mmvdisk_5:
mmlsconfig::0:1:::verbsPorts:mlx5_2/1 mlx5_3/1:ess_x86_64_mmvdisk_6:
mmlsconfig::0:1:::verbsPorts:mlx5_0/1:gss_ppc64:
```

18. Mount file systems on canister B again. (Customer task)

a. Check mounted file systems and mount to canister B again, if necessary.

```
# mmlsmount fs1 -L
```

or

```
# mmlsmount all -L
```

b. Move the quorum node back to the original state, if necessary.

ESS 3000 adapter MES upgrade: AJP1

This procedure is intended for the concurrent (online) adapter MES installation to add the third adapter to each of the server canisters of ESS 3000. The adapter option is AJP1 (VPI).

While the adapter MES is designed to be a concurrent procedure, it is recommended to perform this procedure during a service window of low activity.

Prerequisites

- ESS 3000 6.0.1.1
- Existing adapters that are installed in the system are VPI

Objectives

- Install the new adapter pair, one in each server node canister, by using an online process.
- Incorporate the new interfaces into an existing single master bond.

High-level MES upgrade steps

1. Prepare canister A (bottom canister slot) for adapter MES.
2. Unmount file systems mounted on canister A.
3. Shut down GPFS only in canister A.
4. Remove canister A from the enclosure.
5. Install the hardware - add one adapter to the available middle slot in canister A.
6. Reinsert canister A into enclosure and do basic checks.

7. Incorporate new interfaces to the existing single master bond in canister A.
8. Start GPFS in canister A.
9. Remount file systems to canister A.
10. Prepare canister B (top canister slot) for adapter MES.
11. Unmount file systems mounted on canister B.
12. Shut down GPFS only in canister B.
13. Remove canister B from the enclosure.
14. Install the hardware - add one adapter to the available middle slot in canister B.
15. Reinsert canister B into enclosure and do basic checks.
16. Incorporate new interfaces to the existing single master bond in canister B.
17. Start GPFS in canister B.
18. Remount file systems to canister B.

- Steps 1 – 3 are to be customer tasks
- Steps 4 – 6 are SSR tasks
- Steps 7 – 12 are expected to be customer tasks
- Steps 13 – 15 are SSR tasks
- Steps 16 – 18 are expected to be customer tasks

Summary

The goal of this procedure is primarily to add a third high-speed adapter into each ESS 3000 canister. Customer can add supported InfiniBand or Ethernet adapters into the third PCI slot.

- The PCI address is af:00.1
- The adapter type is ConnectX-5 [ConnectX-5 VPI]

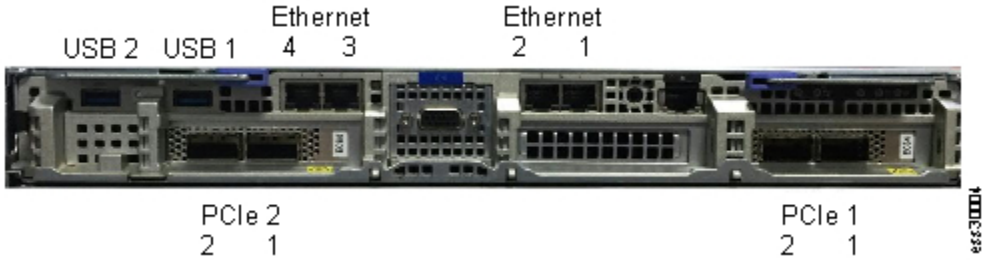

Figure 14. Ethernet ports on canister 1 (upper canister)

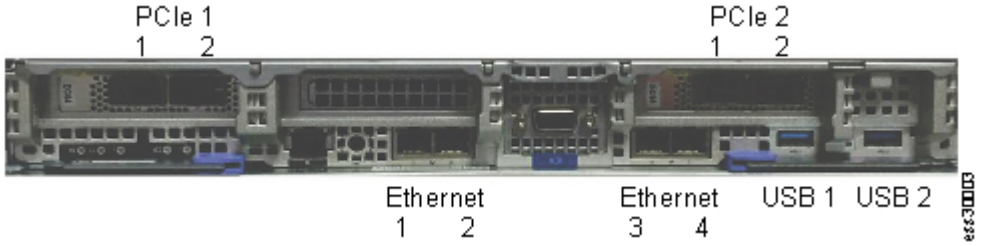

Figure 15. Ethernet ports on canister 2 (lower canister)

Note: These images show the PCIe ports for two adapters for each canister. The MES upgrade installs the third adapter with two more ports in the PCIe slot between the original two adapters.

Online adapter MES procedure

1. Prepare canister A (bottom canister slot) for an adapter MES. (Customer task)
   a. Determine the node class of the ESS 3000 enclosure where you want to perform MES.

      ```
      # mmvdisk nodeclass list
      ```

   b. Display the state of the nodeclass associated with the ESS 3000.

      ```
      # mmgetstate -N the ESS 3000 nodeclass
      ```

      Example:

      ```
      [root@ess3k5a ~]# mmvdisk nodeclass list

      node class                  recovery groups
      --------------------------  -------------------------
      ece_nc_1                    ece_rg_1
      ess_x86_64_mmvdisk_78E016N  ess3k_78E016N
      ess_x86_64_mmvdisk_78E05N1  ess3k_78E05N1
      gssio1_ibgssio2_ib          rg_gssio1-ib, rg_gssio2-ib

      [root@ess3k5a ~]# mmgetstate -N ess_x86_64_mmvdisk_78E05N1

       Node number  Node name        GPFS state
      ------------------------------------------
            21      ess3k5a-ib       active
            22      ess3k5b-ib       active
      ```

   c. If canister A is currently designated as a quorum node, move the quorum to a different node.
      1) To determine which nodes are quorum nodes, issue the following command:

         ```
         # mmlscluster
         ```

      2) To evaluate quorum status, issue the following command:

         ```
         # mmgetstate -s
         ```

      3) If necessary, to designate a different node to be a quorum node, issue the following command:

         ```
         # mmchnode –quorum -N the node which is not in the enclosure's nodeclass
         ```

      4) If necessary, to remove a quorum node from canister A, issue the following command:

         ```
         # mmchnode –noquorum -N canister A
         ```

2. Unmount file systems that are mounted on canister A. (Customer task)

   Unmount GPFS file systems from canister A.

   ```
   # mmumount fs1 -N canister A
   ```

   or you can unmount all file systems at a time from canister A:

   ```
   # mmumount all -N canister A
   ```

   Example:

   ```
   [root@ess3k5a]# mmumount fs3k70 -N ess3k5a
   ```

3. Shut down GPFS only in canister A. (Customer task)

   ```
   # mmshutdown -N canister A
   ```

   Example:

   ```
   [root@ess3k5a]# mmshutdown -N ess3k5a
   ```

4. Remove canister A from the enclosure. (SSR task)

   Record the high speed and management network cabling to canister A. After adapter MES is done, the network cabling needs to be done as recorded here.

   Disconnect high speed and management network cables from canister A ONLY and remove the canister from the bottom slot of the enclosure.

   Note: Do not unplug power cables to either PSU.

5. Install the hardware – add an adapter to the available middle slot in canister A. (SSR task)

   Remove top cover of canister A and install one adapter in the 3rd PCIe slot that is located between the two existing adapters. For more information about these steps, see the *IBM Elastic Storage System 3000: Service Guide*.

6. Insert canister A into the enclosure again and do basic checks. (SSR task)

   a. Put back the top cover of canister A and insert it back to its original canister slot (bottom).

   b. Reconnect original high speed and management network cables.

   c. Connect the new high-speed cables to the new adapter if provided by the customer.

   d. Perform basic checks via the technician port in canister A by using the “essserv1” service ID.

   e. Perform the following checks in the **essutils** menu:

      1) Option **2. Quick storage configuration check**

      2) Option **3. Check enclosure cabling and paths to disks**

      Both options must be successful. Option **2** must show six network adapters.

      Example:

      ```
      localhost: Valid Network Adapter Configuration. Number of Adapter(s) found: 6
      These checks are run per canister.
      ```

7. Incorporate new interfaces to the existing single master bond in canister A. (Customer task)

   To configure each of the two new ports to be either InfiniBand or Ethernet, see the *ConnectX-5 VPI support on ESS 3000* topic in the *IBM Elastic Storage System 3000: Quick Deployment Guide*. Contact technical support, if required.

8. Start GPFS in canister A. (Customer task)

   ```
   # mmstartup -N canister A
   ```

   a. Verify that GPFS is active on both A and B canisters.

      ```
      # mmgetstate -N ESS 3000 nodeclass
      ```

   b. Do the following extra steps on canister A, if MES is for InfiniBand:

      1) Update the `verbsPorts` list, first start GPFS manually.

      2) Ensure that correct entries are listed in `verbsPorts` for the target node class by issuing the following command:

         ```
         /opt/ibm/ess/tools/samples/essServerConfig.sh node class name
         ```

         ```
         # mmlsconfig -Y | grep -i verbsPort
         ```

         Example:

         ```
         [root@ess3k5a ~]# mmlsconfig -Y | grep -i verbsPort
         mmlsconfig::0:1:::verbsPorts:mlx5_1/1::
         mmlsconfig::0:1:::verbsPorts:mlx5_0/1 mlx5_1/1:bodhi1-ib,gssio2-ib,bodhi_nc1,ems:
         mmlsconfig::0:1:::verbsPorts:mlx5_0/1 mlx5_1/1 mlx5_2/1 mlx5_3/1:ess_x86_64:
         mmlsconfig::0:1:::verbsPorts:mlx5_2/1 mlx5_3/1 mlx5_4/1
         mlx5_5/1:ess_x86_64_mmvdisk_5:
         ```

```
mmlsconfig::0:1:::verbsPorts:mlx5_2/1 mlx5_3/1:ess_x86_64_mmvdisk_6:
mmlsconfig::0:1:::verbsPorts:mlx5_0/1:gss_ppc64:
```

9. Mount file systems to canister A again. (Customer task)
   a. Check mounted file systems and mount to canister A again, if necessary.

   ```
   # mmlsmount fs1 -L
   ```

   or

   ```
   # mmlsmount all -L
   ```

   b. Move the quorum node back to the original state, if necessary.
10. Prepare canister B (top canister slot) for an adapter MES. (Customer task)
    a. Determine the node class of the ESS 3000 enclosure where you want to perform MES upgrade.

    ```
    # mmvdisk nodeclass list
    ```

    b. Display the state of the nodeclass associated with the ESS 3000.

    ```
    # mmgetstate -N the ESS 3000 nodeclass
    ```

    Example:

    ```
    [root@ess3k5a ~]# mmvdisk nodeclass list

    node class                  recovery groups
    --------------------------  --------------------------
    ece_nc_1                    ece_rg_1
    ess_x86_64_mmvdisk_78E016N  ess3k_78E016N
    ess_x86_64_mmvdisk_78E05N1  ess3k_78E05N1
    gssio1_ibgssio2_ib          rg_gssio1-ib, rg_gssio2-ib

    [root@ess3k5a ~]# mmgetstate -N ess_x86_64_mmvdisk_78E05N1

     Node number  Node name        GPFS state
    -----------------------------------------
          21      ess3k5a-ib       active
          22      ess3k5b-ib       active
    ```

    c. If canister B is currently designated as a quorum node, move the quorum to a different node.
       1) To determine which nodes are quorum nodes, issue the following command:

       ```
       # mmlscluster
       ```

       2) To evaluate quorum status, issue the following command:

       ```
       # mmgetstate -s
       ```

       3) If necessary, to designate a different node to be a quorum node, issue the following command:

       ```
       # mmchnode –quorum -N the node which is not in this enclosure's nodeclass
       ```

       4) If necessary, to remove a quorum node from canister B, issue the following command:

       ```
       # mmchnode –quorum -N canister B
       ```

11. Unmount file systems that are mounted on canister B. (Customer task)

    Unmount GPFS file systems from canister B:

    ```
    # mmumount fs1 -N canister B
    ```

    or you can unmount all file systems at a time from canister B.

    ```
    # mmumount all -N canister B
    ```

Example:

```
[root@ess3k5a]# mmumount fs3k70 -N ess3k5
```

12. Shut down GPFS only in canister B. (Customer task)

    ```
    # mmshutdown -N canister B
    ```

    Example:

    ```
    [root@ess3k5b]# mmshutdown -N ess3k5b
    ```

13. Remove canister B from the enclosure. (SSR task)

    Record the high speed and management network cabling to canister B. After adapter MES is done, the network cabling needs to be done as recorded here.

    Disconnect high speed and management network cables from canister B ONLY and remove the canister from the top slot of the enclosure.

    Note: Do not unplug power cables to either PSU.

14. Install the hardware – add an adapter to the available middle slot in canister B. (SSR task)

    Remove top cover of canister B and install an adapter in the 3rd PCIe slot that is located between two existing adapters. For more information about these steps, see *IBM Elastic Storage System 3000: Service Guide*.

15. Insert canister B into the enclosure again and do basic checks. (SSR task)

    a. Put back the top cover of canister B and insert it back to its original canister slot (top).

    b. Reconnect original high speed and management network cables.

    c. Connect the new high speed cables to the new adapter if provided by the customer.

    d. Perform basic checks via the technician port in canister B using the “essserv1” service ID.

    e. Perform the following checks in the **essutils** menu:

       1) Option **2. Quick storage configuration check**

       2) Option **3. Check enclosure cabling and paths to disks**

    Both options must be successful. Option **2** must show six network adapters.

    Example:

    ```
    localhost: Valid Network Adapter Configuration. Number of Adapter(s) found: 6
    These checks are run per canister.
    ```

16. Incorporate new interfaces to the existing single master bond in canister B. (Customer task)

    To configure each of the two new ports to be either InfiniBand or Ethernet, see the *ConnectX-5 VPI support on ESS 3000* topic in the *IBM Elastic Storage System 3000: Quick Deployment Guide*. Contact technical support, if required.

17. Start GPFS in canister B. (Customer task)

    ```
    # mmstartup -N canister B
    ```

    a. Verify that GPFS is active on BOTH canisters.

    ```
    # mmgetstate -N ESS 3000 nodeclass
    ```

    b. Do the following extra steps on canister B, if MES is for AJP1 (InfiniBand):

       1) Update the `verbsPorts` list, first start GPFS manually.

2) Ensure that correct entries are listed in `verbsPorts` for the target node class by issuing the following command:

```
/opt/ibm/ess/tools/samples/essServerConfig.sh node class name
```

```
# mmlsconfig -Y | grep -i verbsPort
```

Example:

```
[root@ess3k5a ~]# mmlsconfig -Y | grep -i verbsPort
mmlsconfig::0:1:::verbsPorts:mlx5_1/1::
mmlsconfig::0:1:::verbsPorts:mlx5_0/1 mlx5_1/1:bodhi1-ib,gssio2-ib,bodhi_nc1,ems:
mmlsconfig::0:1:::verbsPorts:mlx5_0/1 mlx5_1/1 mlx5_2/1 mlx5_3/1:ess_x86_64:
mmlsconfig::0:1:::verbsPorts:mlx5_2/1 mlx5_3/1 mlx5_4/1 mlx5_5/1:ess_x86_64_mmvdisk_5:
mmlsconfig::0:1:::verbsPorts:mlx5_2/1 mlx5_3/1:ess_x86_64_mmvdisk_6:
mmlsconfig::0:1:::verbsPorts:mlx5_0/1:gss_ppc64:
```

18. Mount file systems on canister B again. (Customer task)
    a. Check mounted file systems and mount to canister B again, if necessary.

```
# mmlsmount fs1 -L
```

or

```
# mmlsmount all -L
```

    b. Move the quorum node back to the original state, if necessary.

ESS 3000 memory MES upgrade

Overview

This procedure is intended for the non-concurrent (offline) memory MES installation of feature code ACGB to upgrade cache memory with additional 768 GB (24 x 32 GB) for each system.

Objectives

- Install all memory DIMMs to both server canisters.
- Configure GPFS pagepool size to 60% (of cache memory size).

High-level MES upgrade steps

1. Stop all I/O to one or more file systems that are mounted on the MES target ESS 3000 system.
2. Unmount one or more file systems and prepare GPFS to stay in the offline mode.
3. Shut down GPFS.
4. Power off ESS 3000.
5. Install the memory.
6. Power on ESS 3000 and do basic checks.
7. Configure GPFS page pool size to the 60% target.
8. Restore GPFS normal operational mode.
9. Mount one or more file systems that are designated to the target ESS 3000.
10. Resume I/O.

- Customer tasks are steps 1 - 3.
- SSR tasks are steps 4 - 6.
- Customer tasks are steps 7 - 10.

Summary

The goal of this procedure is to add additional memory into each ESS 3000 canister. When the physical memory is installed, the customer can complete the operation by increasing the GPFS page pool. For more information, see the Planning for hardware chapter of the *IBM Elastic Storage System 3000: Hardware Planning and Installation Guide*.

Offline memory MES procedure

1. Stop all I/O workload that is running on one or more file systems that are mounted on the target ESS 3000 for MES (customer task).

   Ensure that all I/O workload is stopped, and do the next step.
2. Turn off the GPFS autoload and automount settings on both server canisters during the MES process (customer task).
   a. Log in as root to each canister and issue the following commands:
      1) To get the information about a GPFS cluster, issue the following command:

         ```
         # mmlscluster
         ```

         Identify and copy and paste the server node names that are associated with the target ESS 3000 into a note.
      2) To get file systems information, issue the following command:

         ```
         # mmvdisk fs list
         ```

         Identify and copy and paste one or more file system names that are associated with the target ESS 3000 into a note.
      3) To unmount one or more file systems, issue the following command:

         ```
         # mmumount <filesystem name> -a
         ```
      4) Turn off the GPFS automount by issuing the following command:

         ```
         # mmchfs <filesystem> -A no
         ```
      5) Turn off the GPFS autoload by issuing the following command:

         ```
         # mmchconfig autoload=no
         ```
      6) Confirm that one or more target file systems are no longer mounted by issuing either of the following command:

         ```
         # df -h
         ```

         ```
         # mslsmount
         ```

         When you issue one of these commands, the file systems must not be displayed in the output.
      7) Double check and confirm that the GPFS automount is disabled by issuing the following command:

         ```
         # mmlsfs <filesystem name>
         ```
      8) Ensure that `autoload=no` is under the nodeclass entry that is associated with the target ESS 3000 by issuing the following command:

         ```
         # mmlsconfig
         ```
3. Shut down GPFS in both canisters by issuing the following command (customer task):

```
# mmshutdown -N <Node server names separated by a comma>
```

a. Ensure that the node servers associated with the target ESS 3000 are shut down by issuing the following command:

```
# mmgetstate -a
```

4. Pull power cables that are connected to the ESS 3000 system (SSR task).
   a. Confirm with customer to ensure that customer did all required steps, and then disconnect the power cables.
   b. To shut down the storage enclosure, unplug both power cords that are on both the sides of the ESS 3000 system.
5. For MES memory installation, do the following steps (SSR task).
   a. Pull out the ESS 3000 from the frame.
   b. Open each canister.
   c. Insert the additional DIMMS into each canister.

      For more information about these steps, see the *IBM Elastic Storage System 3000: Service Guide* or the Planning for hardware chapter of the *IBM Elastic Storage System 3000: Hardware Planning and Installation Guide*.
6. Power on ESS 3000 and do basic checks (SSR task).
   a. After basic checks completion, place everything back into the frame and reinsert power cables. This step restarts the nodes.

      You can use the procedure in the Installing chapter of the *IBM Elastic Storage System 3000: Hardware Planning and Installation Guide* to do the following steps:

      1) Plug your laptop to point-to-point to each container technician port.
      2) Log in as `essserv1`.
      3) Run the **essutils** command.

   b. In the **essutils** command, use the SSR tools option to run the following option on each canister:

      - Option #6. Check the memory

```
Running: /opt/ibm/ess/tools/bin/ess3kplt -t memory –local

ESS3K Mem Inspection:

InspectionPassed:       True
Total Available Slots:  24     (expected 24)
Total Installed Slots:  24     (expected 12 or 24)
DIMM Capacity Errors:   0     (Number of DIMMs with a size different from 32 GB)
DIMM Speed Errors:      0     (Number of DIMMs with a speed of neither 2400 nor 2666
MT/s)
Inspection DateTime:    2020-02-19 15:50:41.031501

Press Enter to continue...

The key here is the Total Installed Slots should now be ‘24’ and the
InspectionPassed:       True
```

      1) Check the memory of each canister.
      2) If any issues, go back to the main page of essutils -> Advanced Tasks -> Check the memory. This option dumps the complete list of DIMMs in each slot.
7. Configure GPFS page pool size to the 60% target (customer task).

   Find the node class name to use, and list the current pagepool settings by issuing the following commands from either one of the server canisters:

```
# mmvdisk nc list
```

a. Identify the node class name that is associated with the system by going through MES.

Example

```
[root@ess3k5a ~]# mmvdisk nc list

node class            recovery group
--------------------  ---------------
ess_x86_64_mmvdisk    ess3k
ess_x86_64_mmvdisk_5  ess3k5
gssio1_ibgssio2_ib    -
```

b. Gather the current pagepool configuration by issuing the following command:

```
# mmvdisk server list –-nc <node class name> --config
```

Example

```
[root@ess3k5a ~]# mmvdisk server list --nc ess_x86_64_mmvdisk_5 --config

node
number  server                              active   memory    pagepool  nsdRAIDTracks
------  --------------------------------  -------  --------  --------  -------------
    21  ess3k5a-ib.example.net               no        754 GiB    75 GiB         131072
    22  ess3k5b-ib.example.net               no        754 GiB    75 GiB         131072
```

Here you can see that the pagepool is less than 25% of physical memory.

c. To change the pagepool percentage, check that GPFS is running:

d. Restart the GPFS by issuing the following command:

```
# mmstartup -N <node class name>
```

Example

```
[root@ess3k5b ~]# mmstartup -N ess_x86_64_mmvdisk_5
Wed Feb 19 16:37:02 EST 2020: mmstartup: Starting GPFS ...
```

e. Change the pagepool to 60%, which is 460G by issuing the following command:

```
# mmchconfig pagepool=460G -N <node class name>
```

f. Ensure that the 460G pagepool setting is listed for the target node class by issuing the following command:

```
# mmlsconfig -Y | grep -i pagepool
```

8. Restore GPFS normal operational mode and confirm pagepool configuration setting (customer task). Do the following steps on both canisters:

a. Restart the server by issuing the following command:

```
# systemctl reboot
```

b. When the server is up again, do a basic ping test between the canister over the high-speed interface.

c. If the ping is successful, start GPFS again by issuing the following command:

```
# mmstartup -N <node class name>
```

d. Ensure that node servers are active before you do the next step by issuing the following command:

```
# mmgetstate -a
```

You can use the following command also to check the pagepool:

```
# mmvdisk server list --nc <node class name> --config
```

Example

```
[root@ess3k5a ~]# mmvdisk server list --nc ess_x86_64_mmvdisk_5 --config

 node
number  server                              active   memory    pagepool  nsdRAIDTracks
------  ----------------------------------  -------  --------  --------  -------------
    21  ess3k5a-ib.example.net                 yes       754 GiB   459 GiB         131072
    22  ess3k5b-ib.example.net                 yes       754 GiB   459 GiB         131072
```

e. Turn on the GPFS automount by issuing the following command:

```
# mmchfs <filesystem> -A yes
```

f. Turn on the GPFS autoload by issuing the following command:

```
# mmchconfig autoload=yes
```

g. Double check and confirm that the GPFS automount is enabled by issuing the following command:

```
# mmlsfs <filesystem name>
```

h. Ensure that `autoload=yes` is under the node class entry that is associated with the target ESS 3000 by issuing the following command:

```
# mmlsconfig
```

9. Mount the file system (customer task).

a. Mount each file system individually by issuing the following command:

```
# mmmount <filesystem> -a
```

b. If you want to mount all file systems in all nodes, you can issue the following command:

```
# mmmount all -N all
```

c. Ensure that the file systems are up.

d. Confirm that one or more target file systems are mounted by issuing the following command:

```
# mmlsmount <filesystem> -L
```

10. Do health check by issuing the following command, and resume I/O because the MES is complete (customer task):

```
# mmhealth node show
```

ESS 3000 storage drives concurrent MES upgrade

An online IBM Elastic Storage System 3000 (ESS 3000) MES upgrade is supported for customers who want to upgrade a 12-drive ESS 3000 to a 24-drive ESS 3000.

To upgrade the system, the NVMe drives with the same size as the existing 12 drives must be used. This MES upgrade doubles the available storage capacity in the existing ESS 3000. For this concurrent upgrade, it is recommended to perform this operation during a period of low workload stress. Existing customer recovery group and file system data are preserved during the MES upgrade.

Supported upgrade path: 12 NVMe ESS 3000 -> 24 NVMe ESS 3000

Prerequisites

- All new or existing building blocks must be at the ESS 5.3.5.2, ESS 3000 6.0.0.2 or later level. If the setup has any protocol nodes, these nodes must also be upgraded to ESS 5.3.5.2 levels (underlying

code IBM Spectrum Scale 5.0.4.3 must be verified by using the **gssinstallcheck** or **essinstallcheck** command).

- The system must be healthy before the ESS 3000 storage MES upgrade.
- The existing ESS 3000 must be a properly installed 12 NVMe system with 12 NVMe drives correctly located in slots 1 - 6 and 13 - 18.
- If the canister servers are allocated as quorum servers, understand the implications of losing a quorum server on one canister server at a time during this operation. If you do not want to lose the quorum, move the quorum to different servers during this procedure.
- It is recommended to wear an ESD wrist band when you work on the hardware, for example, inserting NVMe drives.

MES upgrade considerations

GPFS uses preferentially the new network shared disks (NSDs) to store data of a new file system. GPFS has four new NSDs that are the same as the four original NSDs and the workload on each server is the same as it was before. The new file system data goes to the four new NSDs, like before the resizing, the original file system data goes to the four original NSDs. Consider the necessity of restriping and the current demands on the system. New data that is added to the file system is correctly striped. Restriping a large file system requires many insert operations and delete operations, which might affect the system performance. Restripe a large file system, when the system demand is low.

Concurrent MES upgrade steps

1. Ensure that the technical delivery assessment (TDA) process is complete before you start the MES upgrade.
2. Ensure that the system is at the ESS 3000 6.0.0.2 or later level for the storage MES.
3. If losing quorum on each canister server during MES upgrade causes an issue with the customer configuration, move the quorum to other supported nodes in the system.
4. Ensure that the 12 new NVMe drive FRUs are of the same capacity as the original 12 NVMe drives.
5. Insert the 12 new NVMe drive FRUs into slots 7 - 12 and 19 - 24.

   Do not move the original 12 NVMe drives to different slots!
6. Verify the new 24 NVMe disk topology on both canisters by issuing the following command:

   ```
   # mmvdisk server list --disk-topology --node-class this ESS 3000 node class
   ```

   Example

   ```
   [ess3ka ~]# mmvdisk server list --disk-topology --node-class ess_x86_64_mmvdisk_5_mySN
    node                                    needs    matching
   number  server                          attention   metric   disk topology
   ------  --------------------------------  ---------  --------  -------------
       21  ess3ka-ib.example.net               no          100/100  ESS3K SN0 24 NVMe
       22  ess3kb-ib.example.net               no          100/100  ESS3K SN0 24 NVMe
   ```

   - Both canisters must show a 24 NVMe disk topology.
   - If any errors such as a new disk is bad or does not show up are reported, the errors must be fixed before proceeding.
7. To check whether all 24 NVMe drives have the latest firmware level, issue the following command from one of the canisters:

   ```
   # mmlsfirmware --type drive
   ```

   Example

```
                    enclosure        firmware  available
type  product id      serial number    level     firmware  location
----  ----------      -------------    --------  --------  --------
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 10
```

```
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 11
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 12
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 13
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 14
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 15
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 16
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 17
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 18
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 19
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 1
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 20
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 21
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 22
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 23
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 24
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 2
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 3
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 4
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 5
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 6
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 7
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 8
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ISN1I SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 9
```

8. If the drive firmware is not at the latest level, issue the following command from one of the canisters to update the drive firmware:

```
# mmchfirmware --type drive
```

   a. After the **mmchfirmware** command completes, verify that the drive firmware levels are correct by issuing the following command again:

```
# mmlsfirmware --type drive
```

   Example

```
[ess3ka ~]# date; mmlsfirmware --type drive
Sat Feb 22 16:42:06 EST 2020

                     enclosure        firmware  available
type  product id     serial number    level     firmware  location
----  ----------     -------------    --------  --------  --------
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E *SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 12
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E *SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 19
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E *SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 20
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E *SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 21
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E *SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 23
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E *SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 24
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E *SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 7
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E *SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 8
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E *SN1ISN1I ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 9
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 10
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 11
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 13
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 14
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 15
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 16
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 17
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 18
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 1
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 22
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 2
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 3
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 4
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 5
drive 3.84TB NVMe G3 Tier-1 Flash mySN SN1ESN1E SN1ISN1I  ess3ka-ib Rack myrack U37-38, Enclosure 5141-AF8-mySN Drive 6
```

   b. If some drives failed to update to the current level, rerun the following command to update the drives to the current level:

```
# mmchfirmware --type drive
```

9. Issue the following command against the affected recovery group. This command adds the new capacity and begins rebalancing data onto the new disks.

```
# mmvdisk recoverygroup resize --recovery-group ESS 3000 recoverygroup
```

   Example

```
[ess3ka ~]#    mmvdisk recoverygroup resize --recovery-group ess3k_mySN
mmvdisk: Obtaining pdisk information for recovery group 'ess3k_mySN'.
```

```
mmvdisk: Analyzing disk topology for node 'ess3ka-ib.example.net'.
mmvdisk: Analyzing disk topology for node 'ess3kb-ib.example.net'.
mmvdisk: Validating existing pdisk locations for recovery group 'ess3k_mySN'.
mmvdisk: The resized server disk topology is 'ESS3K SN0 24 NVMe'.
mmvdisk: Validating declustered arrays for recovery group 'ess3k_mySN'.
mmvdisk: Adding new pdisks to recovery group 'ess3k_mySN'.
mmvdisk: Updating declustered array attributes for recovery group 'ess3k_mySN'.
mmvdisk: Successfully resized recovery group 'ess3k_mySN'
```

Note: Use the **mmvdisk recoverygroup list** command to determine the proper recovery group.

10. Update the node class server configuration for 24 drives by issuing the following command:

```
# mmvdisk server configure --update --node-class this ESS 3000 node class --recycle 1
```

Example

```
[ess3ka ~]# mmvdisk server configure --update --node-class ess_x86_64_mmvdisk_5_mySN --
recycle 1
mmvdisk: Checking resources for specified nodes.
mmvdisk: Node class 'ess_x86_64_mmvdisk_5_mySN' has a shared recovery group disk topology.
mmvdisk: Updating configuration for node class 'ess_x86_64_mmvdisk_5_mySN' (recovery group
'ess3k5_mySN').
mmvdisk: Restarting GPFS on the following nodes:
mmvdisk:     ess3ka-ib.example.net
mmvdisk: Restarting GPFS on the following nodes:
mmvdisk:     ess3kb-ib.example.net
```

Important: This command automatically stops and restarts GPFS on each canister server in a serial fashion by using the `--recycle 1` option. If you do not want to stop and restart GPFS, the server can be configured without the `--recycle 1` option. The customer can restart GPFS manually on each canister server by the process at each step. For more information about manually restarting GPFS, see “Example: Manually restarting GPFS on the ESS 3000 canisters” on page 52.

11. Verify that the newly added space is available to the system.

```
# mmvdisk pdisk list --recovery-group ESS 3000 recovery group
```

Example

```
[ess3ka ~]# mmvdisk pdisk list --rg ess3k_mySN

                        declustered
recovery group  pdisk     array     paths  capacity  free space  FRU (type)       state
--------------  -------  ----------  -----  --------  ----------  ---------------  -----
ess3k_mySN      e1s01     DA1        2      3576 GiB  1610 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s02     DA1        2      3576 GiB  1626 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s03     DA1        2      3576 GiB  1616 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s04     DA1        2      3576 GiB  1612 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s05     DA1        2      3576 GiB  1612 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s06     DA1        2      3576 GiB  1622 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s07     DA1        2      3576 GiB  2902 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s08     DA1        2      3576 GiB  2892 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s09     DA1        2      3576 GiB  2902 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s10     DA1        2      3576 GiB  2896 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s11     DA1        2      3576 GiB  2902 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s12     DA1        2      3576 GiB  2896 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s13     DA1        2      3576 GiB  1602 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s14     DA1        2      3576 GiB  1612 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s15     DA1        2      3576 GiB  1622 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s16     DA1        2      3576 GiB  1632 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s17     DA1        2      3576 GiB  1640 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s18     DA1        2      3576 GiB  1600 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s19     DA1        2      3576 GiB  2892 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s20     DA1        2      3576 GiB  2900 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s21     DA1        2      3576 GiB  2892 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s22     DA1        2      3576 GiB  2902 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s23     DA1        2      3576 GiB  2900 GiB    3.84TB NVMe G3   ok
ess3k_mySN      e1s24     DA1        2      3576 GiB  2894 GiB    3.84TB NVMe G3   ok
```

```
# mmvdisk recoverygroup list --recovery-group this ESS 3000 recovery group --declustered-
array
```

Example

```
[ess3ka ~]# mmvdisk rg list --rg ess3k_mySN --declustered-array

declustered   needs          vdisks      pdisks     replace        capacity
   array     service  type  user log  total spare  threshold  total raw free raw  background task
-----------  -------  ----  ---- ---  ----- -----  ---------  --------- --------  ---------------
DA1          no       NVMe     4   5     24     2          2     76 TiB   45 TiB  rebalance (18%)

mmvdisk: Total capacity is the raw space before any vdisk set definitions.
mmvdisk: Free capacity is what remains for additional vdisk set definitions.
[ess3ka ~]#
```

12. Allocate the new space by using the **mmvdisk** command set.

    An example of adding the new storage to an existing file system is as follows:

    a. Define a vdisk set that is tied to the existing recovery group.

```
[ess3ka ~]#  mmvdisk vdiskset define --vdisk-set vs_ess3k_1 --recovery-group ess3k_mySN --code 8+2p
--block-size 4m --set-size 40%

mmvdisk: Vdisk set 'vs_ess3k_1' has been defined.
mmvdisk: Recovery group 'ess3k_mySN' has been defined in vdisk set 'vs_ess3k_1'.

                     member vdisks
vdisk set       count   size   raw size  created  file system and attributes
--------------  -----  -------- --------  -------  -------------------------
vs_ess3k_1          4   6152 GiB 7820 GiB  no       -, DA1, 8+2p, 4 MiB, dataAndMetadata, system

                declustered                capacity            all vdisk sets defined
recovery group     array     type  total raw  free raw  free%  in the declustered array
--------------  -----------  ----  ---------  --------  -----  ------------------------
ess3k_mySN      DA1          NVMe     76 TiB    15 TiB    20%  vs_ess3k, vs_ess3k_1

                                    vdisk set map memory per server
node class                   available  required  required per vdisk set
---------------------------  ---------  --------  ----------------------
ess_x86_64_mmvdisk_5_mySN    102 GiB  7080 MiB  vs_ess3k (1074 MiB), vs_ess3k_1 (1077 MiB)

[ess3ka ~]#
```

    b. Create the newly defined vdisk set.

```
[ess3ka ~]#  mmvdisk vdiskset create --vdisk-set vs_ess3k_1

mmvdisk: 4 vdisks and 4 NSDs will be created in vdisk set 'vs_ess3k_1'.
mmvdisk: (mmcrvdisk) [I] Processing vdisk RG002LG001VS003
mmvdisk: (mmcrvdisk) [I] Processing vdisk RG002LG002VS003
mmvdisk: (mmcrvdisk) [I] Processing vdisk RG002LG003VS003
mmvdisk: (mmcrvdisk) [I] Processing vdisk RG002LG004VS003
mmvdisk: Created all vdisks in vdisk set 'vs_ess3k_1'.
mmvdisk: (mmcrnsd) Processing disk RG002LG001VS003
mmvdisk: (mmcrnsd) Processing disk RG002LG002VS003
mmvdisk: (mmcrnsd) Processing disk RG002LG003VS003
mmvdisk: (mmcrnsd) Processing disk RG002LG004VS003
mmvdisk: Created all NSDs in vdisk set 'vs_ess3k_1'.
[ess3ka ~]#
```

    c. Add the new vdisk set to the existing file system.

```
[ess3ka ~]#    mmvdisk filesystem add --file-system ess3k --vdisk-set vs_ess3k_1

mmvdisk: The following disks of ess3k will be formatted on node ess3k3a.example.net:
mmvdisk:     RG002LG001VS003: size 6299740 MB
mmvdisk:     RG002LG002VS003: size 6299740 MB
mmvdisk:     RG002LG003VS003: size 6299740 MB
mmvdisk:     RG002LG004VS003: size 6299740 MB
mmvdisk: Extending Allocation Map
mmvdisk: Checking Allocation Map for storage pool system
mmvdisk: Completed adding disks to file system ess3k.

[ess3ka ~]#
```

    d. Verify that the vdisk set is added to the file system.

```
[ess3ka ~]#  mmvdisk vs list --vs vs_ess3k_1

                     member vdisks
vdisk set       count   size   raw size  created  file system and attributes
--------------  -----  -------- --------  -------  -------------------------
vs_ess3k_1          4  6152 GiB 7820 GiB  yes      ess3k, DA1, 8+2p, 4 MiB, dataAndMetadata, system
```

```
                declustered                capacity           all vdisk sets defined
recovery group    array     type  total raw  free raw  free%  in the declustered array
--------------  -----------  ----  ---------  --------  -----  ------------------------
ess3k_mySN      DA1          NVMe     76 TiB    15 TiB    20%  vs_ess3k, vs_ess3k_1

                                       vdisk set map memory per server
node class                    available  required  required per vdisk set
----------------------------  ---------  --------  ----------------------
ess_x86_64_mmvdisk_5_mySN    102 GiB  7080 MiB  vs_ess3k (1074 MiB), vs_ess3k_1 (1077 MiB)

[ess3ka ~]#
```

The customer can use the new space by creating new vdisk sets from the available space. When the vdisk sets are added to the existing file system (if required), the restripe operation can be run. For more information, see *IBM Spectrum Scale: Administration Guide*.

Example: Manually restarting GPFS on the ESS 3000 canisters

You can manually stop and start GPFS to solidify the nodes configuration changes on both canisters.

For configuration changes to take effect, restart GPFS on one canister at a time, and ensure that at least one of the canisters is always in the active state. It is also important to understand the effect of losing quorum on the canister during the GPFS recycling.

1. Check the quorum state by issuing the following command:

   ```
   # mmgetstate -s
   ```

   A sample output is as follows:

   ```
    Node number  Node name        GPFS state
   -------------------------------------------
         21      ess3ka-ib         active


   Summary information
   ---------------------
   Number of nodes defined in the cluster:           10
   Number of local nodes active in the cluster:      10
   Number of remote nodes joined in this cluster:     0
   Number of quorum nodes defined in the cluster:     6
   Number of quorum nodes active in the cluster:      6
   Quorum = 4, Quorum achieved
   ```

2. Shut down GPFS on a canister by issuing the following command:

   ```
   # mmshutdown -N canisterA
   ```

3. Start GPFS on a canister by issuing the following command:

   ```
   # mmstartup -N canisterA
   ```

4. Check the state of GPFS on both canisters by issuing the following command:

   ```
   # mmgetstate -N this ESS 3000 node class
   ```

   A sample output is as follows:

   ```
    Node number  Node name        GPFS state
   -------------------------------------------
         21      ess3ka-ib       arbitrating
         22      ess3kb-ib       active
   ```

5. Issue the following command until GPFS is in the active state on both canisters:

   ```
   # mmgetstate -N this ESS 3000 node class
   ```

   A sample output is as follows:

   ```
    Node number  Node name        GPFS state
   -------------------------------------------
   ```

```
       21      ess3ka-ib        active
       22      ess3kb-ib        active
```

6. Repeat the **mmshutdown** command and the **mmstartup** command on *canisterB*.

Chapter 3. Part Listings

View a complete list of CRU and FRU information.

In addition to a detailed list of CRU and FRU information along with their part numbers, you can also view the part numbers for FRU - Cables in a separate list.

CRU part number list

The CRU part numbers are listed in the table.

CRU Part Numbers List

Table 9. CRU Part Numbers

Description	Part Number
SFF Enclosure bezel, right	00Y2512
2000 W AC PSU	01YM310
Power Interposer	01YM314
Rail kit	01YM317
1.92 TB 2.5" NVMe Flash drive	01LL512
3.84 TB 2.5" NVMe Flash drive	01LL513
7.68 TB 2.5" NVMe Flash drive	01LL514
15.36 TB 2.5" NVMe Flash drive	01LL515
Left Brand Bezel	01LL519

FRU part number list

The FRU part numbers are listed in the table.

FRU Part Number

Table 10. FRU Part Numbers

Description	Part Number
PCIe4 LP 2-Port VPI 100Gb IB-EDR/Ethernet adapter	0000001LL586
PCIe4 LP 2-port 100 Gb EDR IB CAPI adapter	0000000WT176
100 GbE Optical Transceiver QSFP28	0000001FT706
PCIe4 LP 2-port 100 Gb ROCE EN LP adapter	0000001FT742
Altsrc 2P100HP (*Arab Nation orders, China Source)	0000001FT769
2-Port 100G LP (*Arab Nation orders, China Source)	0000001FT812
Node canister: ESS 3000 (5141-AF8)	0000001LL518

Table 10. FRU Part Numbers (continued)	
Description	**Part Number**
END7 Power Cable - Drawer to IBM PDU - C13/C20 (250V/10A) for India	0000001PP687
END5 power cord (9.2 ft), Drawer to IBM PDU - C13/C20 (250V/10A) for India	0000001PP688
Trusted Platform Module (TPM)	0000001YM315
Drive Blank	0000001YM705
DIMM Filler	0000001YM789
PCIe riser card with bracket assembly	0000001YM902
Control Enclosure Midplane Assembly	0000002PX580
BLADE QSFP+ 40	0000049Y7928
32 GB DDR4 DIMM	01LJ207
800GB M.2 SSD	01LL516
Control Enclosure Midplane Assembly	01YM312
Fan module	01YM313

Cable part number list

The cable part numbers are listed in the table.

Cable part number list

Table 11. Cable Part Numbers	
Description	**Part Number**
IB cbl 2 M	0000000RX861
CR2032 coin cell	0000000RY543
1 M EDR IB COPPER CABLE/TRANSCEIVER, QSFP/QSFP	0000000WT049
2 M EDR IB COPPER CABLE/TRANSCEIVER, QSFP/QSFP	0000000WT050
3 M EDR IB OPTICAL CABLE/TRANSCEIVER, QSFP/QSFP	0000000WT051
5 M EDR IB OPTICAL CABLE/TRANSCEIVER, QSFP/QSFP	0000000WT052
10 M EDR IB OPTICAL CABLE/TRANSCEIVER, QSFP/QSFP	0000000WT053
15 M EDR IB SWITCH OPTICAL CABLE, QSFP/QSFP	0000000WT054
1.5M EDR IB COPPER CABLE/TRANSCEIVER, QSFP/QSFP	0000000WT082
20M EDR IB SWITCH OPTICAL CABLE, QSFP/QSFP	0000000WT083

Table 11. Cable Part Numbers (continued)	
Description	**Part Number**
30M EDR IB OPTICAL CABLE/TRANSCEIVER, QSFP/QSFP	0000000WT084
50M EDR IB SWITCH OPTICAL CABLE, QSFP/QSFP	0000000WT085
3M FDR10 MLX	0000000WT098
5M FDR10 MLX	0000000WT099
10M FDR10 MLX	0000000WT100
15M FDR10 MLX	0000000WT101
20M FDR10 MLX	0000000WT102
30M FDR10 MLX	0000000WT103
50M FDR10 MLX	0000000WT104
25M EDR IB SWITCH OPTICAL CABLE, QSFP/QSFP	0000000WT124
3M, Green Ethernet Cat 5E cable	0000001AF037
10M, Green Ethernet Cat 5E cable	0000001AF038
3M, Blue Ethernet Cat 5E cable	0000001AF040
10M, Blue Ethernet Cat 5E cable	0000001AF041
25M, Blue Ethernet Cat 5E cable	0000001AF042
10M, Yellow Ethernet Cat 5E cable	0000001AF043
3M, Yellow Ethernet Cat 5E cable	0000001AF045
0.5M QSFP28 passive copper 100Gb Ethernet cable	0000001FT718
1M QSFP28 passive copper 100Gb Ethernet cable	0000001FT719
1.5M QSFP28 passive copper 100Gb Ethernet cable	0000001FT720
2M QSFP28 passive copper 100Gb Ethernet cable	0000001FT721
3M QSFP28 AOC 100Gb Ethernet cable	0000001FT722
5M QSFP28 AOC 100Gb Ethernet cable	0000001FT723
10M QSFP28 AOC 100Gb Ethernet cable	0000001FT724
20M QSFP28 AOC 100Gb Ethernet cable	0000001FT725
30M QSFP28 AOC 100Gb Ethernet cable	0000001FT726
50M QSFP28 AOC 100Gb Ethernet cable	0000001FT727
100M QSFP28 AOC 100Gb Ethernet cable	0000001FT728
15M QSFP28 AOC 100Gb Ethernet cable	0000001FT730
END3 Power Cable - Drawer to IBM PDU - C13/C14, 200-240V/10A for India	0000001KV679
END0 Power Cord M (6.5 foot), Drawer to IBM PDU - C13/C14 (250V/10A) for India	0000001KV680

Table 11. Cable Part Numbers (continued)	
Description	**Part Number**
END1 Power Cord M (9 foot), Drawer to IBM PDU - C13/C14 (250V/10A) for India	0000001KV681
END2 Power Cord m (14 ft), Drawer to IBM PDU - C13/C14 (250V/10A) for India	0000001KV682
5M, Blue Ethernet Cat 5E cable	0000002CL468
5M, Green Ethernet Cat 5E cable	0000002CL469
5M, Yellow Ethernet Cat 5E cable	0000002CL470
ELC5 Power Cable - Drawer to IBM PDU - C13/C20 (250V/10A)	0000002EA542
6665 Power Cablem (9.2 ft), Drawer to IBM PDU - C13/C20 (250V/10A)	0000039M5392
6672 Power Cord M (6.5 foot), Drawer to IBM PDU - C13/C14 (250V/10A)	0000039M5508
6671 Power Cord M (9 foot), Drawer to IBM PDU - C13/C14 (250V/10A)	0000039M5509
6458 Power Cord.m (14 ft), Drawer to IBM PDU - C13/C14 (250V/10A)	0000039M5510
IB OPTIC 10M	0000041V2458
IB OPTIC 31M	0000045D6369

Accessibility features for the system

Accessibility features help users who have a disability, such as restricted mobility or limited vision, to use information technology products successfully.

Accessibility features

The following list includes the major accessibility features in IBM Spectrum Scale RAID:

- Keyboard-only operation
- Interfaces that are commonly used by screen readers
- Keys that are discernible by touch but do not activate just by touching them
- Industry-standard devices for ports and connectors
- The attachment of alternative input and output devices

IBM Knowledge Center, and its related publications, are accessibility-enabled. The accessibility features are described in IBM Knowledge Center (www.ibm.com/support/knowledgecenter).

Keyboard navigation

This product uses standard Microsoft Windows navigation keys.

IBM and accessibility

See the IBM Human Ability and Accessibility Center (www.ibm.com/able) for more information about the commitment that IBM has to accessibility.

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 IBM Elastic Storage System 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 System (ESS)
A high-performance, GPFS NSD solution made up of one or more building blocks. 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 System (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 provides 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 ESS 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 ESS 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 ESS 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 ESS, 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*.

Index

Index

Product Number: 5765-DME
5765-DAE

SC28-3158-02