Lenovo Distributed Storage Solution for IBM® Spectrum Scale™ (DSS-G)

Spectrum Scale Update Bundle for DSS-G

`dss-g-update-standard-5.1.2-2.9`

DSS-G Spectrum Scale Update Procedure

April 27, 2022

hpcstorage@lenovo.com

April 2022

© Copyright Lenovo 2018–2022
LIMITED AND RESTRICTED RIGHTS NOTICE: If data or software is delivered pursuant to a General Services Administration "GSA" contract, use, reproduction, or disclosure is subject to restrictions set forth in Contract No. GS- 35F-05925

Contents

© 2017–2022 Lenovo. All rights reserved.

List of Acronyms

CCR	Clustered Configuration Repository
DSS	Distributed Storage Solution
DSS-G	DSS for IBM Spectrum Scale
ECE	Erasure Code Edition
ESM	Enclosure Services Module
FAQ	Frequently Asked Questions
FS	File System
GPFS	General Parallel File System
GSS	GPFS Storage Server
GUI	Graphical User Interface
I/O	Input/Output
JBOD	Just a Bunch of Disks
MTM	Machine Type Model
NSD	Network Shared Disk
NVMe	Non-Volatile Memory express
OS	Operating System
RAID	Redundant Array of Independent Disks
RG	Recovery Group
RHEL	Red Hat Enterprise Linux
RPM	Red Hat Package Manager
SUB	Spectrum Scale Update Bundle
xCAT	Extreme Cloud Administration Toolkit

© 2017–2022 Lenovo. All rights reserved.

Spectrum Scale Update Procedure

1. Introduction

This document is intended for experienced Spectrum Scale users. Lenovo Professional Services are available if needed.

The intent of this document is to provide a procedure for updating the Spectrum Scale software on one or more Lenovo DSS-G (Distributed Storage Solution for IBM Spectrum Scale) building block(s). Only the Spectrum Scale software stack is updated; the hardware, firmware levels, and the OS (Operating System) including device drivers remain the same. **This procedure only supports DSS-G configurations running either DSS-G release 2.6a and above for the DSS-G v2 stream, or DSS-G release 3.0a and above.**

Please review this document carefully before performing a Spectrum Scale update on DSS-G building block(s). In case that a customer does not feel comfortable to perform the update on their own, Lenovo recommends to engage the Lenovo Professional Services team to perform the Spectrum Scale update. Reach out to your Lenovo account team and/or send email to hpcstorage@lenovo.com in order to establish the contact to the Lenovo Professional Services organization.

1.1 Style guide

The following style is used throughout this document:

- Note box:

  ! Important information and items to make note of.

- Warning box:

  STOP Crucial information that could result in potential loss of data if not carefully accounted for.

- Shell command (gray) and console output (blue) boxes:

```
# Linux commands that one can copy and paste into a linux shell
dssg="Distributed Storage Solution for IBM Spectrum Scale"
echo Lenovo $dssg "(DSS-G)"
```

```
Lenovo Distributed Storage Solution for IBM Spectrum Scale (DSS-G)
```

Various commands also appear `inline` within the text. The following syntax coloring is used for the `shell` commands, `xCAT` commands, `Spectrum Scale` commands, and `DSS-G` commands.

© 2017–2022 Lenovo. All rights reserved.

1.2 Purpose

This document describes the procedure to update the Spectrum Scale software components of the DSS-G solution without going through a complete re-installation of the DSS-G software stack.

1.2.1 Supported DSS-G hardware

The supported DSS-G building block hardware is:

- **DSS-G2xy configuration:** 2× identical DSS (Distributed Storage Solution) servers and a variable number of external JBOD (Just a Bunch of Disks) storage enclosures:
  - Lenovo ThinkSystem SR650 V2 servers with MTM (Machine Type Model) 7Z73-CTOLWW (DSS-G gen3), or
  - Lenovo ThinkSystem SR650 servers with MTM 7X06-CTOLWW or 7X06-CTO1WW (DSS-G gen2), and
  - Lenovo Storage D3284 with MTM 6413-HC1 or 6413-LC1, and/or
  - Lenovo Storage D1224 with MTM 4587-HC2 or 4587-LC2 external storage enclosures.

  ! New "L" MTMs for the SR650 servers and for the external storage enclosures have been introduced in 2020. For compatibility with existing configurations, especially regarding building block expansion, the enclosure "L" models are changed back to "H" models before Spectrum Scale can use them; see **section 2.3.2** of the **DSS-G Installation & Integration Guide** for details.

- **DSS-G100 ECE (Erasure Code Edition) configuration:** 6× to 32× identical DSS servers equipped with either four or eight NVMe (Non-Volatile Memory express) storage drives each:
  - Lenovo ThinkSystem SR630 V2 servers with MTM 7Z71-CTOLWW (DSS-G gen3) or
  - Lenovo ThinkSystem SR630 servers with MTM 7X02-CTOLWW or 7X02-CTO1WW (DSS-G gen2).

Refer to **section 2.1.1** of the **DSS-G Concepts & Administration Guide** for details on the supported DSS-G configurations.

1.2.2 Spectrum Scale Update Bundle (SUB) for DSS-G

This update procedure uses a dedicated package, the **SUB (Spectrum Scale Update Bundle) for DSS-G**, that installs a new level of Spectrum Scale on top of an existing DSS-G installation. The SUB package provides a set of RPMs (Red Hat Package Manager) that are specific to a given installation of the DSS-G software stack as well as other resources that are required to properly handle the DSS-G hardware. Each SUB package specifically targets a given DSS-G edition, `standard` **or** `advanced`, that is already installed on the DSS-G hardware. The filename of the SUB package (for instance, `dss-g-update-standard-5.1.2-2.9.tgz`) reflects both the target DSS-G edition and the level of the Spectrum Scale software it provides.

© 2017–2022 Lenovo. All rights reserved.

Make sure to select the correct edition of the SUB package that matches the edition of the installed DSS-G software stack.

The versioning of the Spectrum Scale software follows the `V.R.M-P.e` nomenclature, where:

`V` = version
`R` = release
`M` = mod
`P` = PTF
`e` = efix

The `V.R.M.P` nomenclature may also be used when no efix level is listed.

The Spectrum Scale support model allows for updating between adjacent releases only, *i.e.,* where either only `R` is changing by at most 1 (for instance, 4.1 → 4.2) or `V` is changing by at most 1 from the lastest available `R` level (for instance, 4.2 → 5.0 is allowed, but 4.1 → 5.0 is not).

1.2.3 Applicability of the SUB

This SUB package applies to installations of DSS-G only and according to the DSS-G life cycle management policy (`https://www.lenovo.com/dssg#lcm`). **This SUB package does not support upgrading Spectrum Scale on DSS servers running DSS-G release 2.5c and older. In this case the storage servers must be upgraded to DSS-G release 2.6a or 3.0a at a minimum (preferably DSS-G release 2.9b or 4.1b) prior to applying this SUB package.**

Installing the SUB for DSS-G is **not** supported on GSS (GPFS Storage Server) hardware. SUBs for GSS are also available; refer to the GSS documentation.

This SUB package may be used in two scenarios where the update package is an adjacent release to the installed level of Spectrum Scale (see section 1.2.2):

1. If `V` is not changed: the update is allowed.
2. If `V` is changed: the target system must be running DSS-G 2.9b for the update to be allowed.

Any other scenario, especially updates between non-adjacent releases, is not allowed.

This document provides update instructions using the `dss-g-update-standard-5.1.2-2.9.tgz` update package.

1.3 Planning the Spectrum Scale update

1.3.1 Overall infrastructure impact

The update of the Spectrum Scale software on the DSS-G building block(s) has an impact on the rest of the Spectrum Scale cluster. While the Spectrum Scale software provides interoperability with older releases, the Spectrum Scale level running on the nodes mounting the DSS-G file

© 2017–2022 Lenovo. All rights reserved.

system(s) should be aligned with that of the DSS-G building block(s) being updated. The purpose of section 2 is to provide guidance on how to update Spectrum Scale and associated file system version(s) on the non-storage nodes in the Spectrum Scale cluster(s).

1.3.2 Update overview

Below is a summary of the main steps that will be performed in order to update the Spectrum Scale software on one or multiple DSS-G building blocks:

1. Update the whole Spectrum Scale infrastucture; see section 1.3.1.
2. Perform the **online** update of Spectrum Scale: for each DSS-G build block, carry out a rolling update by re-installing the Spectrum Scale software on each DSS server, one after another. The file systems served by the updated DSS-G building block are still available via manual failover to the peer node (for DSS-G2xy configurations) or by temporarily "suspending" the node being upgraded (for DSS-G100 ECE configurations).

   ! As much as 50% performance degradation can be expected while updating one of the two DSS servers of a DSS-G2xy building block.

   ! For DSS-G100 ECE configurations, the **online** DSS-G upgrade may be possible only when the erasure code level and file system block sizes have been initially selected according to the recommendations described in the following resource:

   `https://www.ibm.com/support/knowledgecenter/STXKQY_ECE_5.0.5/com.ibm.spectrum.scale.ece.v5r05.doc/b1lece_plan_recommendations.htm`

   Otherwise, the DSS-G100 ECE building block must be upgraded **offline**, *i.e.,* with all file systems unmounted.
3. Finalize the Spectrum Scale update for the whole DSS-G server cluster, including updates to the recovery group version, cluster's release level, and file system versions.

© 2017–2022 Lenovo. All rights reserved.

2. Overall Spectrum Scale update

Since the DSS-G servers are rarely the only Spectrum Scale nodes in a customer environment, updating Spectrum Scale on the DSS-G building block(s) is part of an overall Spectrum Scale update process. The latter means that all Spectrum Scale nodes leveraging the DSS-G storage, either from the local storage cluster (for instance, quorum nodes or protocol servers) or from remote clusters (for instance, compute nodes) should run the same level of the Spectrum Scale software. After a few general considerations regarding Spectrum Scale (multi-)clusters and rolling updates in section 2.1, several recommendations regarding the update of remote Spectrum Scale clusters as well as non-DSS-G nodes in the DSS-G server cluster are discussed in section 2.2.

2.1 General considerations

This section briefly discusses Spectum Scale multi-clusters and rolling updates.

2.1.1 DSS-G server cluster and Spectrum Scale remote cluster(s)

The Spectrum Scale environment includes one or more DSS-G building blocks as well as other Spectrum Scale nodes, in one or more Spectrum Scale clusters. It is the recommended best practice that the DSS-G servers (plus any additional quorum node) form their own Spectrum Scale cluster, the *DSS-G server cluster* (also referred to as *storage cluster*). File systems created on the DSS-G building blocks are then accessed from Spectrum Scale nodes in other Spectrum Scale cluster(s), the *Spectrum Scale remote cluster(s)*, through the Spectrum Scale multi-cluster feature. However, this is not a mandatory requirement, and DSS-G servers can also be in the same Spectrum Scale cluster as the Spectrum Scale "client" nodes that access the file systems.

Section 2.2 provides several recommendations on the high-level migration steps for the overall Spectrum Scale environment, including the Spectrum Scale remote cluster(s) and non-DSS-G nodes in the DSS-G server cluster. For details on updating non-DSS-G nodes, refer to the IBM Spectrum Scale product documentation. Specific instructions are however provided in section 2.2.4 for updating Spectrum Scale on any GUI (Graphical User Interface) server of the storage cluster. The remaining sections of this document then provide the detailed update steps for the DSS-G building block(s) of the storage cluster.

2.1.2 Spectrum Scale rolling updates

Spectrum Scale supports "N-1" backward compatibility between two adjacent Spectrum Scale releases, for example between Spectrum Scale v5.0 and Spectrum Scale v5.1. This compatibility enables "rolling updates" of a Spectrum Scale cluster to a new level, one node at a time, without shutting Spectrum Scale down on other nodes. It also allows individual clusters in a Spectrum Scale multi-cluster environment to be updated at their own schedules. Access to the file system data can be preserved even though some Spectrum Scale cluster(s) in a multi-cluster environment may still be running Spectrum Scale v5.0, while others are running Spectrum Scale v5.1.

This functionality is part of the base Spectrum Scale software, it is not specific to DSS-G. The update planning for the overall Spectrum Scale environment should be performed with reference to the IBM Spectrum Scale product documentation. For general Spectrum Scale update

© 2017–2022 Lenovo. All rights reserved.

and coexistence considerations, please refer to the following IBM Spectrum Scale "Migration, coexistence and compatibility" document:

https://www.ibm.com/support/knowledgecenter/en/STXKQY_5.0.4/com.ibm.spectrum.scale.v5r04.doc/bl1ins_migratl.htm

Please also check the "IBM Spectrum Scale FAQ (Frequently Asked Questions)" and the "IBM Spectrum Scale RAID (Redundant Array of Independent Disks) FAQ" for the most up-to-date information on IBM Spectrum Scale (see **appendix D – Further information** of the **DSS-G Concepts & Administration Guide** for web links).

2.2 Update the remote clusters and non-DSS-G nodes

This section describes how to perform Spectrum Scale updates on remote clusters (if any) mounting the file system(s) served by DSS-G, as well as any non-DSS-G server that may belong to the DSS-G server cluster (for instance, quorum and/or protocol nodes).

2.2.1 Determine the DSS-G server cluster's release level

Run the following commands to obtain the Spectrum Scale release level from the DSS-G server cluster.

1. From a DSS server (`$DSS_A` per best practice), list all Spectrum Scale nodes in the DSS-G server cluster, determine if Spectrum Scale is `active` on all nodes in this cluster, and obtain the minimum release level of this Spectrum Scale cluster:

```
ssh root@$DSS_A
mmlscluster
mmgetstate -a
mmlsconfig minReleaseLevel
exit
```

2. Determine the Spectrum Scale daemon version that is running on the DSS-G servers. Assuming that an xCAT (Extreme Cloud Administration Toolkit) group `dssg` is defined, run the following command from the xCAT server:

```
xdsh dssg "grep Version: /var/adm/ras/mmfs.log.latest"
```

   In case that there are other (non-DSS-G) nodes in the Spectrum Scale cluster, determine the Spectrum Scale version of those Spectrum Scale nodes as well.

2.2.2 Update Spectrum Scale remote clusters

Before updating the DSS-G server cluster, verify the Spectrum Scale multi-cluster setup and update all Spectrum Scale remote clusters (if any).

The following commands provide a view of the Spectrum Scale remote clusters that mount file system(s) from the DSS-G server cluster. From a DSS server (`$DSS_A` per best practice), list the local file system(s) and their version(s), list remote cluster relationships, and determine which of the file systems of the DSS-G server cluster are mounted for each remote cluster(s):

© 2017–2022 Lenovo. All rights reserved.

```
ssh root@$DSS_A
mmlsfs all_local -V --create-time
mmauth show all
localcluster=$(mmlscluster -Y | grep -v HEADER | grep clusterSummary | cut -d: -f7)
echo "LOCAL: $localcluster"
if mmauth show all -Y > /dev/null 2>&1 ; then  # option -Y is supported
  remoteclusters=( $(mmauth show all -Y | grep -Ev "$localcluster|HEADER" | cut -d: -f7) )
else
  remoteclusters=( $(mmauth show all | grep name: | awk '{print $3}' | sort -u | grep -v "$localcluster") )
fi
for c in "${remoteclusters[@]}" ; do mmlsmount all_local -C "$c" -L ; done
```

The `$localcluster` variable contains the name of the (local) DSS-G storage cluster. The names of the remote Spectrum Scale clusters (if any) are stored in the `$remoteclusters` array. The Spectrum Scale software must be updated on these remote clusters before updating the DSS-G server cluster. If otherwise there isn't any remote cluster mounting the file system(s) served by this Spectrum Scale storage cluster, the output of the last command is empty.

The examples below illustrate the following situation:

- The `DSS.cluster` storage cluster contains a single DSS-G building block with the `dss17,dss18` server pair.
- The building block serves a `fs_16m` Spectrum Scale file system.
- This file system is mounted by multiple client nodes from a Spectrum Scale `x36n36-ib0.cluster` remote cluster.

The commands listed above are repeated on the `dss17` server:

```
# list the local file system(s) and selected attributes
[root@dss07 ~]# mmlsfs all_local -V --create-time

File system attributes for /dev/fs_16m:
=======================================
flag                value                    description
------------------- ------------------------ -----------------------------------
 --create-time      Mon Nov 18 17:57:46 2019 File system creation time
 -V                 22.00 (5.0.4.0)          File system version
```

```
# list the Spectrum Scale local and remote clusters and their relationships
[root@dss17 ~]# mmauth show all

Cluster name:        x36n36-ib0.cluster
Cipher list:         AUTHONLY
SHA digest:          8ef451ffc3ae7d7e032c53f31c3e5118756937108cd19cd02a4aa4c5192bb65e
File system access:  fs_16m    (rw, root allowed)

Cluster name:        DSS.cluster (this cluster)
Cipher list:         AUTHONLY
SHA digest:          16d89395a3eee5c5053ac8943f31972810a32e781b45a4bf0dc69f7cb868c709
File system access:  (all rw)
```

© 2017–2022 Lenovo. All rights reserved.

```
# store and show the name of the local cluster
[root@dss17 ~]# localcluster=$(mmlscluster -Y | grep -v HEADER | grep clusterSummary | cut -d: -f7)
[root@dss17 ~]# echo "LOCAL: $localcluster"
```

```
LOCAL: DSS.cluster
```

```
# store the name of the remote cluster(s) and show which file system(s) are mounted; using -Y variant
[root@dss17 ~]# remoteclusters=( $(mmauth show all -Y | grep -Ev "$localcluster|HEADER" | cut -d: -f7) )
[root@dss17 ~]# for c in "${remoteclusters[@]}" ; do mmlsmount all_local -C "$c" -L ; done
```

```
File system fs_16m is mounted on 16 nodes in cluster x36n36-ib0.cluster:
  172.17.0.46     x36n46-ib0
  172.17.0.43     x36n43-ib0
  172.17.0.41     x36n41-ib0
  172.17.0.55     x36n55-ib0
  172.17.0.44     x36n44-ib0
  172.17.0.53     x36n53-ib0
  172.17.0.54     x36n54-ib0
  172.17.0.42     x36n42-ib0
  172.17.0.50     x36n50-ib0
  172.17.0.49     x36n49-ib0
  172.17.0.45     x36n45-ib0
  172.17.0.51     x36n51-ib0
  172.17.0.36     x36n36-ib0
  172.17.0.47     x36n47-ib0
  172.17.0.48     x36n48-ib0
  172.17.0.52     x36n52-ib0
```

Here, the level of the Spectrum Scale software running on the 16 client nodes of the `x36n36-ib0.cluster` remote cluster must be updated to match the level that will be installed on the DSS nodes of the `DSS.cluster` storage cluster.

After Spectrum Scale has been updated on the remote cluster(s), the release level (`minReleaseLevel` Spectrum Scale configuration parameter) of the remote cluster(s) may or may not be updated to `LATEST`, as appropriate for the overall Spectrum Scale environment. Use `mmchconfig release=LATEST` as described in section 4.3 if updating the remote cluster's release level. This is not relevant to the DSS-G server cluster (unless a remote cluster is exporting Spectrum Scale file systems back to the DSS-G server cluster).

2.2.3 Update non-DSS-G nodes in the DSS-G server cluster

Before updating the DSS-G servers in the DSS-G server cluster, it is advisable to update all non-DSS-G nodes in the cluster (if any). These may comprise Spectrum Scale protocol servers, management or quorum nodes, DSS-G GUI nodes, and the xCAT management server when it is used as the third quorum node of a single DSS-G2xy building block. Section 2.2.4 specifically details the procedure to update GUI node(s) in the DSS-G server cluster. Other non-DSS-G nodes in this cluster may follow a similar procedure by skipping the RPMs specific to Spectrum Scale RAID and the GUI, and applying other Spectrum Scale resources as needed.

After Spectrum Scale has been updated on the non-DSS-G nodes of the DSS-G server cluster, the `minReleaseLevel` of the DSS-G server cluster or the file system version of the Spectrum Scale file system(s) in the DSS-G server cluster should **not** be changed at this time, because the

© 2017–2022 Lenovo. All rights reserved.

DSS-G servers still need to be updated. The release level of the DSS-G server cluster and the versions of the file system(s) will be updated in section 4.3 after all DSS-G building blocks have been updated.

2.2.4 Update GUI node(s) in the DSS-G server cluster

This section details the procedure to update Spectrum Scale on a GUI node that manages a DSS-G server (storage) cluster by leveraging this SUB package for DSS-G. As detailed in **the PDF document** of the **DSS-G Graphical User Interface**, the GUI server(s) in the storage cluster require the same set of Spectrum Scale RPMs that are installed on the DSS-G servers. In the following, the GUI node is referenced with `$GUI`.

1. Copy the `dss-g-update-standard-5.1.2-2.9.tgz` update tarball to the GUI server, unpack it on the GUI server, and go to the newly-created directory for this update package:

```
# assume dss-g-update-standard-5.1.2-2.9.tgz is copied into /install on the xCAT server
xdcp $GUI /install/dss-g-update-standard-5.1.2-2.9.tgz
ssh root@$GUI
tar xvzf dss-g-update-standard-5.1.2-2.9.tgz
cd dss-g-update-standard-5.1.2-2.9
```

   From now on, all commands run from the GUI node.

2. Stop the GUI service:

```
systemctl stop gpfsgui
```

3. Shut down Spectrum Scale on the GUI node. If it has a quorum role in the Spectrum Scale storage cluster, verify that there is enough quorum first. The following command reports the number of quorum nodes in the storage cluster:

```
mmlscluster -Y | grep -i "clusternode.*quorum" | wc -l
```

   Then shut down Spectrum Scale on the GUI server:

```
mmshutdown
```

4. Verify the list of installed Spectrum Scale RPMs. This list should match the RPMs installed on the DSS-G storage servers complemented with GUI-specific RPMs (including `gpfs.gui` and `gpfs.gss.pmcollector`):

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

   In case this list reports the `gpfs.kafka` and `gpfs.librdkafka` RPMs, remove them from the GUI server:

```
yum remove -y gpfs.kafka gpfs.librdkafka
```

© 2017–2022 Lenovo. All rights reserved.

5. Create `shell` variables (arrays) that list generic and GUI-specific RPMs from the SUB package depending on the version of RHEL (Red Hat Enterprise Linux) running on the GUI server:

```
osverbase=$(cat /etc/os-release | grep VERSION_ID= | cut -d\" -f2 | cut -d. -f1)
rpms=( $(find -name \*rpm | grep '\.rpm$') )
generic=( $(printf "%s\n" "${rpms[@]}" | grep -v "\.el") )
specific=( $(printf "%s\n" "${rpms[@]}" | grep "\.el${osverbase}") )
```

   Then update the installed Spectrum Scale RPMs:

```
yum update -y "${generic[@]}" "${specific[@]}"
```

6. Once the update of all RPMs is complete, rebuild the GPFS (General Parallel File System) portability layer and restart Spectrum Scale:

```
mmbuildgpl
mmstartup
```

7. Verify that Spectrum Scale is now running the updated level with `mmdiag` and by listing again the installed RPMs:

```
mmdiag --version
rpm -qa | grep gpfs | sort
```

8. Finally, restart the `pmsensors` and `pmcollector` services followed by the GUI service:

```
systemctl restart pmsensors
systemctl restart pmcollector
systemctl start gpfsgui
```

Refresh the GUI in the web browser by clicking the browser reload button. The login page of the DSS-G GUI should appear again within a minute.

2.2.5 Client settings

When the Spectrum Scale software is updated outside of the DSS-G servers (*i.e.*, compute clients mounting the DSS-G file system(s), Spectrum Scale management or quorum nodes), be sure to align the Spectrum Scale configuration of these nodes by running the most current version of the `dssClientConfig.sh` script. The latter is found under the `/opt/lenovo/dss/bin/` on the updated DSS-G servers. Refer to **section 6 – Configuring Spectrum Scale client nodes** of the **DSS-G Concepts & Administration Guide** for details.

! Spectrum Scale protocol servers usually require a specific set of tuning parameters and should not run the `dssClientConfig.sh` script. GUI nodes managing the DSS-G servers do not need running that script either.

© 2017–2022 Lenovo. All rights reserved.

3. Update the DSS-G building block(s)

In this section, we cover the steps needed to update the Spectrum Scale software on the DSS-G building block(s).

3.1 Prerequisites

Before updating Spectrum Scale on a DSS-G building block, verify that the DSS-G building block is healthy. If there are any issues with the building block, correct them before proceeding. If the issues can not be corrected, check with Lenovo support regarding the best solution for your issue(s).

3.1.1 Perform a DSS-G health check

Before starting the update, check if the DSS-G building block is healthy and **fix any issues that are found.** If the update is started while there are errors such as missing disk paths to the `$DSS_B` peer server of a DSS-G2xy configuration, this may cause loss of access to the Spectrum Scale file systems when the Spectrum Scale daemon is shut down during the update process. Run the following commands to check the health of the DSS-G building block:

- For DSS-G2xy configurations:

```
ssh root@$DSS_A
DSS_A=dss01 ; DSS_B=dss02
RG_A=$DSS_A ; RG_B=$DSS_B
mmgetstate -a

# mmvdisk management commands
mmvdisk rg list --rg $RG_A --all      # check that all disks have 2 active paths
mmvdisk rg list --rg $RG_B --all      # check that all disks have 2 active paths
mmvdisk pdisk list --rg all --not-ok  # "mmvdisk: All pdisks are ok." is expected

# legacy management commands
mmlsrecoverygroup $RG_A -L --pdisk  # check that all disks have 2 active paths
mmlsrecoverygroup $RG_B -L --pdisk  # check that all disks have 2 active paths
mmlspdisk all --not-ok              # "no disks were found" is expected response

mmlsenclosure all --not-ok  # no output is expected response
gnrhealthcheck --local
```

- For DSS-G100 ECE configurations:

```
ssh root@$DSS_A
rg=$(mmvdisk rg list -Y | grep -v HEADER | cut -d: -f7)  # only one RG exists
nc=$(mmvdisk nc list -Y | grep -v HEADER | cut -d: -f7)  # node class name
mmvdisk rg list --rg "$rg" --all      # check that all pdisks have 1 active path
mmvdisk pdisk list --rg all --not-ok  # "mmvdisk: All pdisks are ok." is expected

# both commands do not come clean with DSS-G 3.0, 4.0 and 4.1; see note below
mmlsenclosure all -N "$nc" --not-ok -L
gnrhealthcheck --local
```

© 2017–2022 Lenovo. All rights reserved.

When updating DSS-G 3.0 with SR630 servers and DSS-G 4.0 or 4.1 with SR630 V2 servers, the `mmlsenclosure` command above reports that each DSS-G100 ECE node "needs service". The extended command output provided by the `-L` option should flag that the issue comes from the "esm" component of the servers' internal enclosure that holds the NVMe storage. Using a 6-node DSS-G100 ECE configuration as an example:

```
[root@ece01 ~]# mmlsenclosure all -N nc_ece01 --not-ok -L
```

```
                 needs  nodes
serial number    service
-------------    ------- ------
J100MM5X         yes     ece04-ib0.cluster
J100MM5Z         yes     ece06-ib0.cluster
J100MM61         yes     ece05-ib0.cluster
J100MM63         yes     ece01-ib0.cluster
J100MM64         yes     ece02-ib0.cluster
J100MM67         yes     ece03-ib0.cluster

component type  serial number     component id    failed value   unit   properties  fru     location
--------------  -------------     ------------    ------ -----   ----   ----------  ---     --------
esm             J100MM5X          0               no                     NO_BUS_FAILURE NOT_FAILING NOT_CRITICAL
esm             J100MM5Z          0               no                     NO_BUS_FAILURE NOT_FAILING NOT_CRITICAL
esm             J100MM61          0               no                     NO_BUS_FAILURE NOT_FAILING NOT_CRITICAL
esm             J100MM63          0               no                     NO_BUS_FAILURE NOT_FAILING NOT_CRITICAL
esm             J100MM64          0               no                     NO_BUS_FAILURE NOT_FAILING NOT_CRITICAL
esm             J100MM67          0               no                     NO_BUS_FAILURE NOT_FAILING NOT_CRITICAL
```

As a consequence the consecutive `gnrhealthcheck` command also reports enclosure problems for the same reason.

In this case, the DSS-G100 ECE servers are actually fine and do not need be fixed. The error is caused by a minor mistake in the definition of the ESM (Enclosure Services Module) component for the SR630 server in DSS-G 3.0 and for the SR630 V2 server in DSS-G 4.0 and 4.1 that does not cause any functional issue. This problem is fixed in this SUB package.

In a DSS-G cluster where all DSS-G servers are at the same DSS-G code level, the `gnrhealthcheck` script can also be run without the `--local` option. But if there are different DSS-G code levels on different servers in the cluster, it is best to only invoke it as shown above. That script is not designed to run in mixed environments, and may produce misleading output when run globally on a cluster with a mix of different DSS-G releases.

3.1.2 Review state of nodes

Check the overall health of the building block nodes using the `mmhealth` command:

```
mmhealth node show -N all
```

Below is an example output for a DSS-G2xy configuration, showing the first node only:

```
mmhealth node show -N all
```

```
Node name:      dss13.cluster
Node status:    HEALTHY
Status Change:  2 days ago
```

© 2017–2022 Lenovo. All rights reserved.

```
Component       Status        Status Change     Reasons
------------------------------------------------------------------------------
GPFS            TIPS          2 days ago        callhome_not_enabled
NETWORK         HEALTHY       2 days ago        -
FILESYSTEM      HEALTHY       2 days ago        -
DISK            HEALTHY       2 days ago        -
NATIVE_RAID     HEALTHY       2 days ago        -

[...]
```

Specifically review `FILESYSTEM`, `DISK`, and `NATIVE_RAID` for any problems and correct issues found.

3.1.3 Review state of file system(s)

Beyond the status of `FILESYSTEM` from the previous section, it is important to perform an explicit check to review the state of the file system before updating:

1. Use `mmlsfs` to list all of the Spectrum Scale file system device names:

   ```
   mmlsfs all
   ```

   This command gives a detailed list of all available file system devices and their associated properties.

   In contrast, the following commands only collect and show the actual device names that are required for the next step:

   ```
   scaledevices=( $(mmlsfs all -Yd | awk -F: '$3==0 {print $7}') )
   printf "%s\n" "${scaledevices[@]}"
   ```

   ```
   dsshome
   ```

   Here only one file system device is listed; it is stored in the `scaledevices` array variable.

2. For each filesystem device name, check the status of the file system with `mmlsdisk`:

   ```
   for d in "${scaledevices[@]}"; do mmlsdisk "$d" -e 2>&1 | sed "s/^/${d}: /"; done
   ```

   ```
   dsshome: All disks up and ready
   ```

   This command loops over all file system devices stored in the `scaledevices` array, runs `mmlsdisk <device> -e` for each device, and prepends the device name to each output line. In the example above, all "disks" (actually Spectrum Scale NSD (Network Shared Disk) devices) are "up and ready" for the `dsshome` file system, which is the expected state.

   > **!** While file systems look healthy at the disk (NSD) abstraction level, it does not necessarily means that all other abstraction layers underneath (for instance, pdisks) are healthy. The more detailed checks performed in section 3.1.1 should expose any such potential issues.

© 2017–2022 Lenovo. All rights reserved.

If all file systems report “All disks up and ready” the next step can be skipped.

3. In case the file systems are not nominal, the `mmlsdisk` command loop above may report something like:

```
dsshome: disk         driver   sector     failure holds    holds                            storage
dsshome: name         type       size       group metadata data  status        availability pool
dsshome: ------------ -------- ------ ----------- -------- ----- ------------- ------------ ------------
dsshome: dss13_Data   nsd         512          30 no       yes   to be emptied up           data
```

Here, for the `dss13_Data` disk (NSD) the availability is “up” as expected but the status is “to be emptied” (instead of “ready”). This situation means that the NSD has been suspended. See `man mmlsdisk` for a description of the status and availability fields.

The status and/or availability of those non-nominal NSD must be resolved before updating, using the `mmchdisk` command to “resume” or “stop” then “start” the NSD. See the command syntax and options with `mmchdisk --help` and check `man mmchdisk` for a complete description of the command.

> **!** The `mmchdisk` commands may require recycling Spectrum Scale on all nodes.

3.2 Spectrum Scale update of the DSS-G servers

The DSS servers of a DSS-G building block are updated in sequence, one server after another:

- For **DSS-G2xy configurations**, Spectrum Scale on the `$DSS_A` server is updated first while the `$DSS_B` server continues to serve both the `RG_A` and the `RG_B` recovery groups of that building block. The full sequence of steps in sections 3.2.1 to 3.2.11 must be performed when updating the `$DSS_A` server, except steps sections 3.2.4 and 3.2.10 which are relevant to DSS-G100 ECE configurations only. Next, Spectrum Scale on the `$DSS_B` server is updated after both recovery groups have been moved to `$DSS_A`. The same sequence of steps needs to be repeated to update the `$DSS_B` server, interchanging server and recovery group names as explained in section 3.2.12.
- For **DSS-G100 ECE configurations**, Spectrum Scale is updated on each DSS server while the single scale-out recovery group is handled by all other DSS nodes of the building block. The same sequence of steps in sections 3.2.1 to 3.2.11 must be performed except those to fail over and shut down Spectrum Scale (sections 3.2.2 and 3.2.3) then startup Spectrum Scale and fail back the active RG (Recovery Group) (sections 3.2.8 and 3.2.9) which are applicable to DSS-G2xy configurations only and are replaced by the steps in sections 3.2.4 and 3.2.10. The `$DSS_A` server refers to the DSS-G100 ECE node being upgraded.

In all cases, **only one DSS server per DSS-G building block must be updated at a time.**

If more than one DSS-G building block needs to be updated in the DSS-G server cluster, repeat these steps for each DSS-G building block, one building block after another.

© 2017–2022 Lenovo. All rights reserved.

3.2.1 Move Spectrum Scale management functions away from $DSS_A

Before shutting down Spectrum Scale on the $DSS_A server to perform the Spectrum Scale update, it is important to check for Spectrum Scale cluster functions that may involve this server and move them away from $DSS_A to the $DSS_B server of the same building block.

1. Login to the node: `ssh root@$DSS_A`
2. Check the **node quorum** with `mmlscluster | grep quorum; mmgetstate -a`. If node quorum would be at risk when shutting down server $DSS_A, assign another quorum node to ensure that node quorum is maintained during the update:

```
mmchnode --quorum -N $NewNode  # only if node quorum is at risk
```

3. Check the **cluster configuration servers** with `mmlscluster | grep "server:"`. If the output is blank, skip this step. Otherwise, re-assign another cluster configuration server if $DSS_A should currently have this role. This only applies when using traditional server-based configuration repository; it is not applicable using CCR (Clustered Configuration Repository). The role that $DSS_A is providing must be moved to a different node:

```
mmchcluster -p $DSS_B  # example if $DSS_A is primary
```

4. Check the **cluster manager node** with `mmlsmgr`, and if $DSS_A is a manager node, assign a different manager node:

```
mmchmgr -c $DSS_B      # only if $DSS_A is cluster manager
```

5. Check the **file system manager node** assignments with `mmlsmgr`. Change those roles that reside on $DSS_A to assign a different manager node:

```
mmchmgr $FS $DSS_B     # only if $DSS_A is file system manager for $FS
```

> **!** There may be multiple $FS file systems managed by $DSS_A; all of them need to be moved to the alternate node.

3.2.2 Fail over the $DSS_A recovery group (RG_A) to $DSS_B (DSS-G2xy only)

In preparation to shut down the $DSS_A server, verify that the recovery group RG_B is served by its default server $DSS_B, and fail over the RG_A recovery group from server $DSS_A to server $DSS_B. Perform the following steps on the $DSS_A server:

```
ssh root@$DSS_A
DSS_A=dss01 ; DSS_B=dss02
RG_A=$DSS_A ; RG_B=$DSS_B

# mmvdisk management commands
mmvdisk rg list --rg $RG_B --server     # $DSS_B should be serving $RG_B
mmvdisk rg list --rg $RG_A --server     # $DSS_A should still be serving $RG_A
```

© 2017–2022 Lenovo. All rights reserved.

```
mmvdisk rg change --rg $RG_A --active $DSS_B
mmvdisk rg list --rg $RG_A --server     # $DSS_B should now be serving $RG_A

# legacy management commands
mmlsrecoverygroup $RG_B -L | tail -n 4  # active RG server should be $DSS_B
mmlsrecoverygroup $RG_A -L | tail -n 4  # active RG server should still be $DSS_A
mmchrecoverygroup $RG_A --active $DSS_B
mmlsrecoverygroup $RG_A -L | tail -n 4  # active RG server should now be $DSS_B
```

3.2.3 Shut down Spectrum Scale on `$DSS_A` (DSS-G2xy only)

Now Spectrum Scale can be shut down on server `$DSS_A`:

```
ssh root@$DSS_A
mmgetstate
mmshutdown
mmgetstate
```

> **!** Only the Spectrum Scale daemon need be shut down at this time. Do not shut down or power off the node.

The node is now ready for Spectrum Scale update.

3.2.4 Suspend the `$DSS_A` server (DSS-G100 ECE only)

DSS-G100 ECE configurations use a single scale-out RG for the building block. Updating Spectrum Scale on a DSS server relies on suspending this server and all the pdisks it owns from the RG. Suspending a server is normally used temporarily for quick maintenance, as data from the suspended pdisks is rebuilt onto the other nodes after five minutes. For details refer to:

https://www.ibm.com/support/knowledgecenter/SSYSP8_5.3.5/com.ibm.spectrum.scale.raid.v5r04.adm.doc/bl8adm_mmvdisk_recoverygroup.htm

The suspend procedure is also used to verify that the file system configuration on the DSS-G100 ECE building block allows an **online** DSS-G upgrade, *i.e.,* with all file systems mounted. If the latter were created with different erasure codes and file system block sizes than those recommended, then suspending the node may not be allowed. In this case an **offline** Spectrum Scale update must be performed. Refer to **section 5.1.5** of the **DSS-G Concepts & Administration Guide** and the following resource for details:

https://www.ibm.com/support/knowledgecenter/STXKQY_ECE_5.0.5/com.ibm.spectrum.scale.ece.v5r05.doc/b1lece_plan_recommendations.htm

The procedure to suspend any `$DSS_A` server from a DSS-G100 configuration is as follows. Here, the `ece01` node is used as an example; the name of the recovery group is also `ece01` since it has

© 2017–2022 Lenovo. All rights reserved.

been derived from the first node when creating the RG. A DSS-G100 ECE configuration with 6× DSS servers is used to illustrate the procedure below.

1. By contrast to paired DSS-G2xy configuration with multiple recovery groups, the scale-out RG uses multiple log groups (including a `root` log group for RG metadata and event logs) to balance the I/O (Input/Output) workload. See the following resource for details:

   https://www.ibm.com/support/knowledgecenter/en/SSYSP8_5.3.5/com.ibm.spectrum.scale.raid.v5r04.adm.doc/bl1adv_loggroups.htm

   Verify the assignement of the log groups (and associated vdisks) to the DSS nodes with the following command:

   ```
   [root@ece01 ~]# mmvdisk rg list --rg ece01 --server
   ```

   ```
    node
   number  server                          active   remarks
   ------  ------------------------------  -------  -------
        1  ece01-ib0.cluster               yes      serving ece01: root, LG006, LG012
        2  ece02-ib0.cluster               yes      serving ece01: LG005, LG011
        3  ece03-ib0.cluster               yes      serving ece01: LG004, LG010
        4  ece04-ib0.cluster               yes      serving ece01: LG003, LG009
        5  ece05-ib0.cluster               yes      serving ece01: LG001, LG007
        6  ece06-ib0.cluster               yes      serving ece01: LG002, LG008
   ```

2. Suspend the `ece01-ib0` node:

   ```
   [root@ece01 ~]# mmvdisk rg change --rg ece01 --suspend -N ece01-ib0
   ```

   An **online** Spectrum Scale update is allowed when the command is successful and the Spectrum Scale daemon and all pdisks on this node automatically shut down:

   ```
   mmvdisk: Suspended all pdisks from node 'ece01-ib0.cluster'.
   mmvdisk: Stopped serving log groups from node 'ece01-ib0.cluster'.
   mmvdisk: Shutting down GPFS on node 'ece01-ib0.cluster'.
   mmvdisk: Suspended node 'ece01-ib0.cluster' in 'ece01'.
   mmvdisk: Non-critical rebuilds deferred for 20 minutes in recovery group 'ece01'.
   mmvdisk: After node maintenance is finished, run the command:
   mmvdisk:      mmvdisk recoverygroup change --resume --recovery-group ece01 -N ece01-ib0.cluster
   ```

   If otherwise any file system served by the scale-out RG does not comply to the recommended erasure codes and file system block sizes, the command may fail with an error similar to the following:

   ```
   mmvdisk: Cannot suspend node 'ece01-ib0.cluster' in 'ece01'.
   mmvdisk: Vdisk 'RG001LG001VS002' of recovery group 'ece01'
   mmvdisk: has only '1 node' fault tolerance.
   mmvdisk: This is less than the minimum '1 node + 1 pdisk' fault tolerance
   mmvdisk: required in order to suspend a node. To see the fault tolerances
   mmvdisk: in the recovery group, use the command:
   mmvdisk:      mmvdisk recoverygroup list --recovery-group ece01 --fault-tolerance
   mmvdisk: Cannot suspend node 'ece01-ib0.cluster' in 'ece01'.
   mmvdisk: Command failed. Examine previous error messages to determine cause.
   ```

© 2017–2022 Lenovo. All rights reserved.

In this case, perform an **offline** Spectrum Scale update. All file systems must be unmounted and Spectrum Scale shut down on all nodes of the DSS-G100 ECE configuration:

```
mmumount all -a  # unmount all file systems from all nodes
mmshutdown -a    # shut down Spectrum Scale on all DSS-G100 nodes
```

Since the Spectrum Scale daemon is shut down on all storage servers of the DSS-G100 ECE configuration, all of these DSS nodes may also be updated at once in section 3.2.5.

3. Verify the `mmfsd` daemon is down:

```
[root@ece01 ~]# mmgetstate -a

 Node number  Node name        GPFS state
------------------------------------------
       1      ece01-ib0        down
       2      ece02-ib0        active
       3      ece03-ib0        active
       4      ece04-ib0        active
       5      ece05-ib0        active
       6      ece06-ib0        active
```

4. Check the new balancing of the log groups over the remaining DSS nodes:

```
[root@ece01 ~]# mmvdisk rg list --rg ece01 --server

 node
number  server                          active   remarks
------  ------------------------------  -------  -------
     1  ece01-ib0.cluster               no       configured, suspended
     2  ece02-ib0.cluster               yes      serving ece01: LG005, LG011
     3  ece03-ib0.cluster               yes      serving ece01: LG004, LG010
     4  ece04-ib0.cluster               yes      serving ece01: LG002, LG009, LG012
     5  ece05-ib0.cluster               yes      serving ece01: root, LG001, LG007
     6  ece06-ib0.cluster               yes      serving ece01: LG003, LG006, LG008
```

The node is now ready for the Spectrum Scale update.

3.2.5 Update Spectrum Scale on the `$DSS_A` server

This step performs a re-installation of the Spectrum Scale software on the `$DSS_A` server, removing the existing level and installing the new level provided by the present update package.

Except if performing an offline update of Spectrum Scale, do not perform the Spectrum Scale update on multiple nodes of the DSS-G building block at the same time.

© 2017–2022 Lenovo. All rights reserved.

Copy the `dss-g-update-standard-5.1.2-2.9.tgz` update tarball to the `$DSS_A` server, unpack it on the DSS-G server, and go to the newly-created directory for this update package:

```
# assume dss-g-update-standard-5.1.2-2.9.tgz is copied into /install on the xCAT server
xdcp $DSS_A /install/dss-g-update-standard-5.1.2-2.9.tgz
ssh root@$DSS_A
tar xvzf dss-g-update-standard-5.1.2-2.9.tgz
cd dss-g-update-standard-5.1.2-2.9
```

Then verify that Spectrum Scale is down on the node:

```
# should return: "Failed to connect to file system daemon: No such process"
mmdiag --version
```

The actual Spectrum Scale update can now be performed by running the `update.sh` script:

```
./update.sh
```

The script detects the version and edition of Spectrum Scale that is already installed and that of the update, and asks for confirmation before updating when no mismatch exists.

> **!** The installation of the new RPMs uses the `yum` package manager; it is expected that the message `Warning: RPMDB altered outside of yum.` will appear before updating the Spectrum Scale RPMs. This warning is harmless and can be ignored.

3.2.6 Update the DSS-G server settings on `$DSS_A`

For optimal operation after Spectrum Scale has been updated to a newer level, the configuration of the `$DSS_A` server must be updated by applying a large number of `mmchconfig` parameters specific to DSS-G. In the context of new DSS-G installations, this is done when creating a Spectrum Scale cluster (see **section 4.1 – Create a production Spectrum Scale storage cluster** of the **DSS-G Concepts & Administration Guide** for details). By contrast, in the present update procedure the `$DSS_A` server already belongs to an existing Spectrum Scale cluster and the DSS-G settings must be applied manually. This is the task of the `dssServerConfig.sh` script. See **section 7.4.16 – `dssServerConfig.sh`** of the **DSS-G Concepts & Administration Guide** for details.

> **!** Running `dssServerConfig.sh` is crucial especially when changing the Spectrum Scale release, for instance, updating Spectrum Scale from version 4.2 to 5.0.

The `dssServerConfig.sh` script runs from the `$DSS_A` server, applying settings to itself (the script argument is a Spectrum Scale nodeclass):

```
ssh root@$DSS_A dssServerConfig.sh $DSS_A
```

The updated settings will become effective when restarting Spectrum Scale in section 3.2.8. Both `$DSS_A` and `$DSS_B` servers shall have the same settings applied at the end of the update procedure of the DSS-G building block. Alternatively, for instance in the context of a

© 2017–2022 Lenovo. All rights reserved.

cluster-wide maintenance window, the settings can be applied offline to all DSS-G servers at once (see section 4.1).

3.2.7 Verify the DSS-G topology

As a basic sanity check, verify that the DSS-G topology of the building block is correct:

```
ssh root@$DSS_A
dssgcktopology <nodelist>
```

where `<nodelist>` is a comma-separated list of DSS servers of the building block that is being updated.

Refer to **section 7.4.3 – `dssgcktopology` of the DSS-G Concepts & Administration Guide** for explanations of the output.

3.2.8 Start Spectrum Scale on `$DSS_A` (DSS-G2xy only)

When the DSS-G topology is as expected, Spectrum Scale can be started again on the `$DSS_A` server:

```
ssh root@$DSS_A
mmgetstate
mmstartup
mmgetstate

mmdiag --version  # waits for mmfsd to start
mmgetstate
```

The final `mmgetstate` invocation must return that the node is `active`.

3.2.9 Fail back the `$DSS_A` recovery group (`RG_A`) to `$DSS_A` (DSS-G2xy only)

Verify that the recovery groups and pdisks are in the expected state using the following commands:

```
ssh root@$DSS_A
DSS_A=dss01 ; DSS_B=dss02
RG_A=$DSS_A ; RG_B=$DSS_B

# mmvdisk management commands
mmvdisk rg list --rg $RG_A --server
mmvdisk rg list --rg $RG_B --server
mmvdisk pdisk list --rg $RG_A
mmvdisk pdisk list --rg $RG_B

# legacy management commands
mmlsrecoverygroup $RG_A -L | tail -n 4
mmlsrecoverygroup $RG_B -L | tail -n 4
mmlsrecoverygroup $RG_A -L --pdisk
mmlsrecoverygroup $RG_B -L --pdisk
```

© 2017–2022 Lenovo. All rights reserved.

In particular, make sure that:

- Each storage pdisk is reported with 2 active paths and state `ok`;
- Both pdisks associated with the `NVR` declustered array are present with 1 active path and state `ok`.

In case a `NVR` pdisk is missing, it is crucial to recover this pdisk as detailed in **section 4.2.3 – Repairing the NVR device** of the **DSS-G Problem Determination & Remediation Guide** before failing back the RGs.

Then, manually fail back the recovery groups to the expected server:

```
# mmvdisk management commands
mmvdisk rg change --rg $RG_A --active $DSS_A
mmvdisk rg change --rg $RG_B --active $DSS_B

# legacy management commands
mmchrecoverygroup $RG_A --active $DSS_A
mmchrecoverygroup $RG_B --active $DSS_B
```

3.2.10 Resume `$DSS_A` (DSS-G100 ECE only)

The next step is to add back the node in the scale-out recovery group. The following procedure uses `ece01` for the `$DSS_A` node from a 6-node DSS-G100 ECE configuration as an example.

1. Verify the Spectrum Scale daemon is still down:

```
[root@ece01 ~]# mmgetstate
```

```
 Node number  Node name        GPFS state
------------------------------------------
       1      ece01-ib0        down
```

2. If the node has been successfully suspended in section 3.2.4, resume the node:

```
[root@ece01 ~]# mmvdisk recoverygroup change --resume --recovery-group ece01 -N ece01-ib0
```

```
mmvdisk: Starting GPFS on node 'ece01-ib0'.
mmvdisk: Waiting up to 5 minutes for the GPFS daemon on node 'ece01-ib0' to join the cluster.
mmvdisk: Serving log groups is already enabled on node 'ece01-ib0'.
mmvdisk: Resumed all pdisks from node 'ece01-ib0'.
mmvdisk: Rebuilds are no longer deferred in recovery group 'ece01'.
mmvdisk: Resumed node 'ece01-ib0' in 'ece01'.
```

   Otherwise (offline DSS-G upgrade) restart the Spectrum Scale daemon using `mmstartup` on the node.

3. Verify the Spectrum Scale daemon is now up and the level is as expected, and check that the log groups are rebalanced on the upgraded server:

```
[root@ece01 ~]# mmgetstate -a     # all nodes must be active
[root@ece01 ~]# mmdiag --version  # build "5.1.2.2 efix9 (1149565)" is expected
[root@ece01 ~]# mmvdisk rg list --rg ece01 --server
```

© 2017–2022 Lenovo. All rights reserved.

```
 node
number  server                          active   remarks
------  --------------------------------  -------  -------
     1  ece01-ib0.cluster                 yes      serving ece01: LG008, LG009
     2  ece02-ib0.cluster                 yes      serving ece01: LG005, LG011
     3  ece03-ib0.cluster                 yes      serving ece01: LG004, LG010
     4  ece04-ib0.cluster                 yes      serving ece01: LG002, LG012
     5  ece05-ib0.cluster                 yes      serving ece01: root, LG001, LG007
     6  ece06-ib0.cluster                 yes      serving ece01: LG003, LG006
```

3.2.11 Move Spectrum Scale management functions back to `$DSS_A`

If any Spectrum Scale management functions have been moved away in section 3.2.1, they may be moved back to `$DSS_A` at this time:

```
ssh root@$DSS_A

mmlscluster | grep quorum
mmgetstate -a
mmlscluster
mmlsmgr
mmchmgr -c $DSS_A
mmchmgr $FS $DSS_A  # for each $FS file system
```

3.2.12 Repeat on the next server of the DSS-G building block

The same sequence of steps described in sections 3.1.1 to 3.2.11 now needs to be repeated for the next server of the DSS-G building block. For DSS-G2xy configurations, replace references to `$DSS_A` and `$RG_A` with `$DSS_B` and `$RG_B`, respectively. For DSS-G100 ECE configurations, loop over all the remaining DSS nodes, replacing references to `$DSS_A` with the current node name.

© 2017–2022 Lenovo. All rights reserved.

4. Finalize the update of the DSS-G server cluster

After all non-DSS-G nodes and all DSS-G servers have been updated, the DSS-G server cluster update can be finalized.

4.1 DSS-G server settings

The configuration of the DSS server must be updated for the latest level of Spectrum Scale installed. This is normally performed **online** while updating each DSS server of a DSS-G building block in sequence, as detailed in section 3.2.

Alternatively, when a maintenance window is scheduled for an **offline** update of the whole Spectrum Scale cluster (including *e.g.,* client nodes, protocol servers) it is possible to update the configuration of all DSS-G servers at once before they are put back in production. This is done by running the `dssServerConfig.sh` script from a DSS-G server (here `$DSS_A`) that belongs to the `dssg` nodeclass for which the settings will be applied:

```
# the Spectrum Scale daemon is down on all DSS-G servers
ssh root@$DSS_A dssServerConfig.sh dssg
```

Optimal settings will then apply consistently across the storage cluster once the Spectrum Scale daemon is restarted on all DSS-G nodes. See section 3.2.6 for details and recommendations in using the `dssServerConfig.sh` script.

4.2 Update the recovery group version

It is recommended to let the system run at the old recovery group format level for a while (one day to a week) before performing the RG format update, because **this is not a reversible step.** Therefore, ensure that DSS-G is running reliably before performing this step.

The RG format version should then be migrated to the `LATEST` level for the scale-out recovery group (DSS-G100 ECE configurations) or for both RGs (DSS-G2xy configurations). Before changing the RG format version, also check that the RG is currently served by its default RG server. For DSS-G2xy configurations, the active RG server for `$RG_A` should be `$DSS_A`, and the active RG server for `$RG_B` should be `$DSS_B`.

Run the following commands to update the format for `$RG_A`:

```
ssh root@$DSS_A

# mmvdisk management commands
mmvdisk rg list   --rg $RG_A --version
mmvdisk rg change --rg $RG_A --version LATEST
mmvdisk rg list   --rg $RG_A --version

# legacy management commands (not applicable to DSS-G100)
mmlsrecoverygroup $RG_A -L | grep version -A2
mmchrecoverygroup $RG_A --version LATEST
mmlsrecoverygroup $RG_A -L | grep version -A2

exit
```

© 2017–2022 Lenovo. All rights reserved.

For DSS-G2xy configurations, perform the same for `$RG_B`:

```
ssh root@$DSS_B

# mmvdisk management commands
mmvdisk rg list   --rg $RG_B --version
mmvdisk rg change --rg $RG_B --version LATEST
mmvdisk rg list   --rg $RG_B --version

# legacy management commands (not applicable to DSS-G100)
mmlsrecoverygroup $RG_B -L | grep version -A2
mmchrecoverygroup $RG_B --version LATEST
mmlsrecoverygroup $RG_B -L | grep version -A2

exit
```

After this sequence of steps, the RG version should have been updated from *e.g.,* "4.2.2.0" to "5.0.0.0". In the `mmfs.log`, entries similar to the following indicate the successful update to the RG format version:

```
tail /var/adm/ras/mmfs.log.latest
```

```
Mon Feb 20 22:32:09.232 2019:
  [I] Command: tschrecgroup --recovery-group dss02 --version LATEST
Mon Feb 20 22:32:09.251 2019:
  [C] Command: err 0: tschrecgroup --recovery-group dss02 --version LATEST.
```

4.3 Update the DSS-G cluster's release level and file system versions

The remaining cluster-wide settings that can and should be updated are the cluster's release level, and the file system versions.

- **DSS-G cluster's release level.** As a best practice, keep the DSS-G server cluster at the previous level for several days to ensure that everything is stable, because **those changes are not reversible.** Then run the following on one Spectrum Scale node in the DSS-G server cluster (`$DSS_A` per best practice):

  ```
  ssh root@$DSS_A
  mmlsconfig minReleaseLevel  # check current release level
  mmchconfig release=LATEST
  mmlsconfig minReleaseLevel  # validate the new level
  ```

- **File system versions.** For each Spectrum Scale file system `$FS` that is served by the DSS-G server cluster, the file system version can be upgraded to `full` to enable the latest file system features. Here, running from `$DSS_A` per best practice:

  ```
  ssh root@$DSS_A
  mmlsfs $FS --create-time -V  # check the creation time and version of $FS
  mmchfs $FS -V full
  mmlsfs $FS --create-time -V  # validate the new version
  ```

© 2017–2022 Lenovo. All rights reserved.

> **!** A more limited `compat` version change is also possible. This should not be required here since, at this time of the Spectrum Scale update procedure, all remote Spectrum Scale clusters and all nodes in the DSS-G server cluster should have been updated to at least the same Spectrum Scale version. See sections 2.2.2 and 2.2.3.

4.4 Verify file system mount(s)

Upgrading Spectrum Scale may have remounted the file system(s) internally to the current DSS-G building block. For instance:

```
[root@dss23 ~]# mmlsmount all -L | grep -E "^File system|internal"
```

```
File system fs_4m is mounted on 7 nodes:
  172.17.1.24     dss24-ib0                 (internal mount)
  172.17.1.23     dss23-ib0                 (internal mount)
```

Internal mounts generally mean that the file system is not visibly mounted on the corresponding storage nodes, and system utilities such as `ls` will report empty content for these nodes. If the FS (File System) is properly mounted on other nodes (for instance clients) they will be able to access it normally. The following command can be used to remount all FSs that are reported as internal mounts:

```
# create an array fs[name]="nodes" for internal mounts
eval declare -A fs=( $(mmlsmount all -Y | grep :INT: | awk -F: '{printf "[%s]+=\"%s \"\n",$8,$12}') )
# mount the FSs on the corresponding nodes
for f in "${!fs[@]}"; do
  mmmount "$f" -N "${fs[$f]// /,}"
done
```

Then verify that the FS mounts are as expected by re-running the `mmlsmount` command above.

© 2017–2022 Lenovo. All rights reserved.