Spectrum Scale & ESS Upgrade

May 17th, 2018 GPFS User's Group (Boston)

Aaron Palazzolo (*aspalazz@us.ibm.com*)
Muthu Muthiah (*mutmuthi@in.ibm.com*)

IBM Spectrum Scale & ESS Upgrade

What are we working on?

- **Consistency of Upgrade flow and rules among components**

- **Clear categorization of components that support online vs offline upgrade**

- **Tools to assist with checking settings before/after upgrade**

- **Ease of 'shutdown'**

- **End to End documentation for upgrade + cheat sheets**
    - Workflow examples for manually invoked upgrade
    - Workflow examples for Install Toolkit invoked upgrade
    - OS / driver (OFED) upgrade workflow and guidance

- **Install Toolkit Enhancement for toleration of more cluster states and offline upgrade**

- **ESS + Protocols Upgrade workflow**

- **ESS: gssutils menu driven upgrade, also encompassing protocol nodes**

Definitions

Online Upgrade

- **Component Level:** Ability for a component to co-exist at N and N-1 levels simultaneously while functional, allowing staggered node-at-a-time or multiple node upgrades while still maintaining operation.

- **Cluster Level:** Ability for the cluster to remain online and deliver function of all components within while individual or multiple nodes upgrade from an N-1 to N level of code. Components unable to meet the component level definition of online upgrade are brought down for/during upgrade but this does not affect other components nor the cluster as a whole.

- **Inter-cluster Level:** Storage cluster is upgraded in an online/rolling fashion and redundancy with remotely mounted clusters is such that storage remains available for remote clusters during the upgrade

Definitions

Offline Upgrade

- **Component Level:** Component must be brought offline on all nodes before an upgrade of that component can occur. Complete outage of function delivered by component. Once in this state, a component must be able to handle upgrade from any level to any level so long as all components are also upgraded to the same go-to level.

- **Cluster Level:** All nodes of a cluster must be brought offline before an upgrade can occur. A complete cluster outage. The entire cluster can then be upgraded from any level to any new level

- **Inter-cluster Level:** Storage cluster must be brought offline, thus causing remotely mounted clusters to lose connectivity to storage.
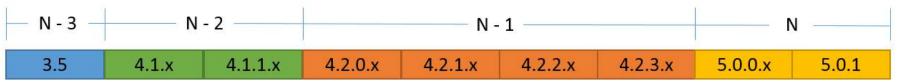
Definitions

Upgrade Completion

An upgrade of a Spectrum Scale cluster is complete only after the following items have occurred:

- All nodes within the cluster are now at the upgraded level

- All Spectrum Scale components and dependencies within every node have been upgraded

- 'mmchconfig release=LATEST' has been run on the cluster, enabling access to new functionality *(this has implications and is a point of no return)*

- 'mmchfs –V full' has been run on the cluster, enabling new functionality that requires different on-disk data structures (this has implications, especially for remote clusters which may be unable to mount the local cluster file systems after this change)
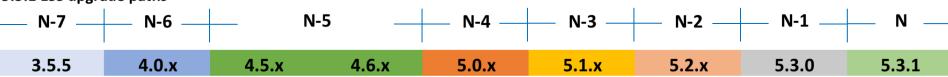
Spectrum Scale and ESS releases

5.0.1 release upgrade paths

N - 3	N - 2		N - 1				N	
3.5	4.1.x	4.1.1.x	4.2.0.x	4.2.1.x	4.2.2.x	4.2.3.x	5.0.0.x	5.0.1

Online upgrade supported from N-1 to N

Offline upgrade supported from any level to N

5.3.1 ESS upgrade paths

N-7	N-6	N-5		N-4	N-3	N-2	N-1	N
3.5.5	4.0.x	4.5.x	4.6.x	5.0.x	5.1.x	5.2.x	5.3.0	5.3.1

Online/Offline upgrade supported from N-2 or N-1 to N

IBM

Spectrum Scale upgrade paths

IBM Spectrum Scale supported upgrade paths

Use this information to understand the supported upgrade paths for IBM Spectrum Scale™.

The following table lists the supported online upgrade paths for IBM Spectrum Scale.

Table 1. IBM Spectrum Scale supported online upgrade paths

Upgrading from:	To: 3.5	To: 4.1.0.x	To: 4.1.1.x	To: 4.2.0.x	To: 4.2.1.x	To: 4.2.2.x	To: 4.2.3.x	To: 5.0.0.x	To: 5.0.1.x
3.5	✓	✓	✓	X	X	X	X	X	X
4.1.0.x	--	✓	✓	✓	✓	✓	✓	X	X
4.1.1.x	--	--	✓	✓	✓	✓	✓	X	X
4.2.0.x	--	--	--	✓	✓	✓	✓	✓	✓
4.2.1.x	--	--	--	--	✓	✓	✓	✓	✓
4.2.2.x	--	--	--	--	--	✓	✓	✓	✓
4.2.3.x	--	--	--	--	--	--	✓	✓	✓
5.0.0.x	--	--	--	--	--	--	--	✓	✓
5.0.1.x	--	--	--	--	--	--	--	--	✓

- ✓: Supported
- X: Not supported
- --: Not applicable

https://www.ibm.com/support/knowledgecenter/en/STXKQY_5.0.1/com.ibm.spectrum.scale.v5r01.doc/bl1ins_migratl.htm

IBM Spectrum Scale & ESS Upgrade

Upgrading From	To 3.5.5 (or earlier)	To 4.0.x	To 4.5/4.6	To 5.0.x	To 5.1.x	To 5.2.0	To 5.3.0	To 5.3.1
3.5.5 (or earlier)	Yes	Yes	Yes	No	No	No	No	NO
4.0.x	-	Yes	Yes	Yes	No	No	No	No
4.5/4.6	-	-	Yes	Yes	Yes	No	No	No
5.0.x	-	-	-	Yes	Yes	Yes	No	No
5.1.x	-	-	-	-	Yes	Yes	Yes	No
5.2.0	-	-	-	-	-	Yes	Yes	Yes
5.3.0	-	-	-	-	-	-	Yes	Yes
5.3.1	-	-	-	-	-	-	-	Yes

IBM Spectrum Scale & ESS Upgrade

ESS version	Spectrum Scale	OS	Kernel errata	OFED	Firmware	IPR	Systemd	Netmgr
5.0.2	4.2.2-3 efix11	RHEL7.2	3.10.0-327.53.1.el7.ppc64	MLNX_OFED_LINUX-3.4-2.0.0.1	FW860.10 (SV860_056)	15511300	219-30.e7_3.8	N/A
5.1.1 (LE+BE)	4.2.3.2	RHEL7.2 (LE+BE)	3.10.0-327.55.3.el7.ppc64 + ppc64le	MLNX_OFED_LINUX-4.0-2.0.0.3	FW860.30 (SV860_103)	15511800	219-30.el7_3.8	N/A
5.2 (LE+BE)	4.2.3-4	RHEL7.3 (LE+BE)	3.10.0-514.26.2.el7.ppc64 + ppc64le	MLNX_OFED_LINUX-4.1-0.1.4.1	FW860.30 (SV860_103)	16519500	219-30.el7_3.9	1.4.0-20.el7_3
5.3 (LE+BE)	5.0.0-1 (GNR efix)	RH7.3 (LE+BE)	3.10.0-514.44.1.ppc64 + ppc64le	MLNX_OFED_LINUX-4.1-4.1.6.1	SV860_138 (FW860.42)	17518300	219-42.el7_4.10	1.8.0-11.el7_4

Sample OS/OFED Upgrade Flow

1) Verify services provided by the node are redundant/hosted on other nodes.... or be prepared for an outage

2) Suspend/Stop all GPFS functions/services on the node to be upgraded

3) Shutdown GPFS on this node
- A shutdown may fail to unload GPFS kernel modules if files are still open. Check lsof if this occurs and either attempt to kill all associated processes the FS, or reboot the node. Set autoload = off before rebooting so that gpfs doesn't come back up on this node.
- Health Monitoring may hold open a file system in rare cases. It may be necessary to stop if shutdown fails.

4) **If using OFED - OFED driver uninstall. T**his can be done from the old OFED driver directory structure or the new one - there is typically a Mellanox uninstall script bundled into these packages.

5) Upgrade the OS via your favorite method *(package manager, manually, etc...)*
- pay attention to any dependency failures. With RedHat, we often see python rpm upgrade failures due to conflicts with CES Object. With the OS itself, we often see failures with a few filesystem related rpms when the OS repo is NFS mounted.
- Make sure you have enough space for your OS to upgrade. SLES snapshots, /boot filling up due to too many old kernels .... are just some of the failure causes we see

6) When upgrade is done, reboot

7) Install the new OFED driver

9) reboot

10) *check any OFED/OS settings and make sure previous settings are restored if overwritten* ****
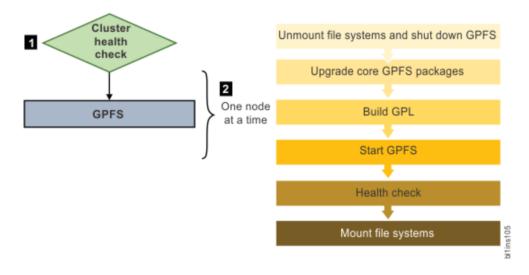
Sample OS/OFED Upgrade Flow

11) Build GPFS portability layer (don't do this until after you've rebooted and applied the latest kernel)

12) Verify no conflicting stuff has been enabled/started with the OS upgrade
- check node for an OS NFS process, for instance. Stop/disable if it exists - this will conflict with CES Ganesha
- check SELinux settings
- check firewall settings

13) Start GPFS, start/resume any functions/services previously stopped.
*** set autoload to yes is needed if you had set autoload to no back on step 3 ****

13) Check health

14) repeat on other node

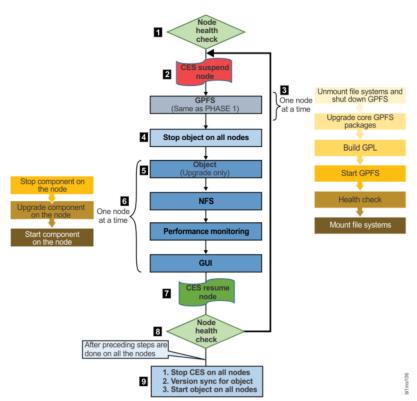## How does the Install Toolkit Upgrade?

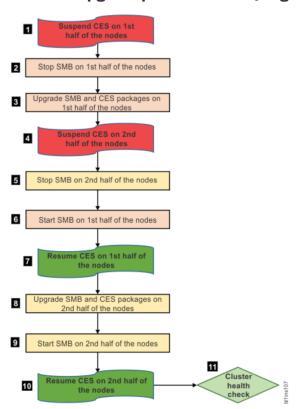### Phase 1 - upgrade non-protocol nodes

How does the Install Toolkit Upgrade?

Phase 2 - upgrade protocol nodes (stage 1)

Phase 2 - upgrade protocol nodes (stage 2)

Sample Install Toolkit Upgrade Steps

1) Download the latest Spectrum Scale Protocols Package and pick a non-ESS node in the cluster to extract it upon

2) Install Toolkit is in the /usr/lpp/mmfs/x.x.x.x/installer directory: ./spectrumscale

3) Setup the Install Toolkit
> **If a Spectrum Scale cluster with no ESS in the same cluster**
> > *./spectrumscale setup -s <IP of installer node> -st ss*
> **If a Spectrum Scale cluster with an ESS in the same cluster**
> > *./spectrumscale setup -s <IP of installer node> -st ess*

4) Populate the Toolkit config (available in 4.2.3, works best at 5.0.0.2 and higher*) - if incompatible with your environment, manually input config*
> **If a Spectrum Scale cluster with no ESS in the same cluster**
> > ./spectrumscale config populate -N <Node in the cluster>
> **If a Spectrum Scale cluster with an ESS in the same cluster**
> > ./spectrumscale config populate -N <EMS node>

5) Check the toolkit config afterwards
> ./spectrumscale node list
> ./spectrumscale filesystem list
> ./spectrumscale nsd list
> ./spectrumscale config gpfs
> ./spectrumscale config protocols
> ./spectrumscale config object
> ./spectrumscale callhome config
> ./spectrumscale fileauditlogging list

Sample Install Toolkit Upgrade Steps

6) If you don't want to upgrade all nodes that Install Toolkit knows about, remove the ones you don't want
     ./spectrumscale node delete <node name>

7) Run the upgrade preceheck
     ./spectrumscale upgrade –precheck

8) Make sure you have base OS repos available on each node the toolkit will upgrade

9) Run the upgrade
     ./spectrumscale upgrade

IBM Spectrum Scale & ESS Upgrade

Sample manual upgrade steps (protocol nodes)

Assumes 4 CES RHEL7 nodes and assumes all 3 protocols. If you are not running SMB, you don't have to take the outage in step 10

1) suspend on node #1 and node #2

2) stop SMB, NFS, OBJ on node #1 and node #2. Also stop OBJ on node #3 and node #4

3) unmount all FSs on node #1 and node #2. If unmount fails, this is your first clue that shutdown won't go well

4) shutdown gpfs on node #1 and node #2
- If shutdown fails to unload kernel - make sure autoload is set to no and reboot. Resume with step 5 after reboot

5) stop pmsensors on node #1 and node #2

6) upgrade rpms in gpfs_rpms, gpfs_rpms/rhel7, smb_rpms/rhel7, object_rpms/rhel7, ganesha_rpms, zimon_rpms/rhel7
- don't do an upgrade gpfs.*. Best to use yum (or apt/zipper) and specify the full path to each rpm you will be upgrading.

7) mmbuildgpl

8) do not bring node #1 and node #2 online

9) suspend node #3 and node #4

10) stop SMB, NFS, and OBJ (OBJ should already be stopped) on node #3 and node #4
Complete outage of SMB. Pause of NFS.

11) mmstartup node #1 and node #2

Sample manual upgrade steps (protocol nodes)

12) start SMB, NFS, and OBJ on node #1 and node #2

13) resume CES on node #1 and node #2
- validate mmhealth cluster show and cluster show CES

14) repeat steps 3->7 on node #3 and #4

15) verify SMB levels on all nodes

16) mmstartup node #3 and #4

17) start SMB and others on node #3 and #4

18) resume CES on node #3 and #4

19) verify all pmsensors are running. Start them if not on all nodes

20) check mmpermon config show -> verify SMB sensors and restrictions/periods exist

21) validate mmhealth cluster show and cluster show CES

Upgrade Links

Spectrum Scale Main Upgrade page

https://www.ibm.com/support/knowledgecenter/en/STXKQY_5.0.1/com.ibm.spectrum.scale.v5r01.doc/bl1ins_migratl.htm

Spectrum Scale Protocols Quick Overview Guide

https://www.ibm.com/developerworks/community/wikis/home?lang=en#!/wiki/General%20Parallel%20File%20System%20(GPFS)/page/Protocols%20Quick%20Overview%20for%20IBM%20Spectrum%20Scale

ESS Upgrade

- Always check the latest ESS Quick Deployment Guide for your code level

https://www.ibm.com/support/knowledgecenter/SSYSP8_5.3.0/com.ibm.ess.v5r3.qdg.doc/bl8qdg_top.htm

ESS + Scale Protocol node upgrade

- See the ESS 5.2, and later, Quick Deployment Guides
- See the Spectrum Scale Protocols Quick Overview Guide

THE END