



# Keep your Spectrum Scale cluster HEALTHY with MAPS

Anna Greim and Pavel Safre  
05.03.2020

# Disclaimer



- IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion. Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision. The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.
- IBM reserves the right to change product specifications and offerings at any time without notice. This publication could include technical inaccuracies or typographical errors. References herein to IBM products and services do not imply that IBM intends to make them available in all countries.





What is MAPS?

MAPS' tools to keep my cluster HEALTHY

Use Cases

Improvements in 5.0.4 / 5.0.5

Tips and Tricks: Call Home



# What is MAPS?



- **Monitoring**, **Availability** and **Proactive Services**

## On premise

### Sysmonitor

mmsysmon.py, mmhealth

### CES

mmces, IP balancing/failover

### Call Home

mmcallhome

### Perfmon

mmperfmon, thresholds,  
Grafana Bridge

### Protocol Tracing

mmprotocoltrace

### FTDC

gpfs.snap, \*.snap.py

## IBM site

### Proactive Services



# MAPS' tools to keep my cluster HEALTHY? [1/2]



- **mmhealth node/cluster show**

- Hundreds of automatic checks / approx. 1000 events
- All results at one glance!
- Be informed via user defined scripts

- **mmhealth thresholds**

- Events based on performance data
- Define own rules!

- **mmperfmon / Grafana Bridge**

- Monitor performance of your cluster



# MAPS' tools to keep my cluster HEALTHY? [2/2]



- **mmcallhome**

- Improved service response times
- Proactively detect issues
- Better dev. test coverage for your cluster

- **mmprotocoltrace**

- In case of protocol issues quickly analyse the root cause

- **CES IP management**

- Minimize impact on node / network failures



# Use Case 1: mmhealth Drilldown [1/2]



- **Situation:**

- The customer checks his cluster as everyday using the **mmhealth** command and sees that something is **FAILED** or **DEGRADED**

- **Wish:**

- The cluster should be HEALTHY for every component

- **Solution:**

- Drilling down to find the cause of the problem with **mmhealth** and fix it.



# Use Case 1: mmhealth Drilldown [2/2]



- Everyday cluster monitoring:

```
mmhealth cluster show
```

- If any component is not HEALTHY, see which RAS event is active and on which node:

```
mmhealth cluster show <component>
```

- To see more details for the node:

```
mmhealth node show -N <nodeName>
```

- To see the details for a RAS event (including cause + user action):

```
mmhealth event show <eventName>
```





# Use Case 2: RAM and OOM problem [1/13]



- **Situation:**

- The customer realizes that some applications might use too much RAM

- **Wish to be informed:**

- when there's almost no free RAM left
- after the OOM killer killed any applications
- via e-mail when this occurs

- **Solution:**

- Thresholds + user callback scripts



# Use Case 2: RAM and OOM problem [2/13]



- Was OOM killer triggered?

test-21.localnet.com

Overview **Events** File Systems NSDs

Actions ▾ | Current Issues ▾ | ↻ Last Updated: 08:02 | ⬇ Export |

Severity	Event Time	Active Until	Event Name
⚠ Warning	02.03.20 15:03:25		out_of_memory

4

1

admin ▾

ALERTS

AGE

⚠ GPFS (1) 17 hours



# Use Case 2: RAM and OOM problem [3/13]



## ■ What does this event mean?

Event Details

Event name: out\_of\_memory

Component: GPFS

Entity type: Node

Entity name: test-21.localnet.com

Event time: 02.03.20 15:03:25

Active until:

Message: Detected Out of memory killer conditions in system log

Description: In an out of memory condition the OOM killer will terminate the process with the highest memory utilization score. This may affect Spectrum scale processes and cause subsequent issues.

Cause: The dmesg command returned log entries which are written by the OOM killer.

User action: Check the memory usage on the node. Identify the reason for the out of memory condition and check the system log to find out which processes have been killed by OOM killer. You might need to recover these processes manually or reboot the system to get to a clean state. Run the command 'mmhealth resolve event out\_of\_memory' once you recovered the system to remove this warning event from mmhealth.

Reporting node: test-21.localnet.com

Event type: Active health state of an entity which is monitored by the system.

Previous

Next

Close

### Cause:

The dmesg command returned log entries which are written by the OOM killer.

### User action:

Check the memory usage on the node. Identify the reason for the out of memory condition and check the system log to find out which processes have been killed by OOM killer. You might need to recover these processes manually or reboot the system to get to a clean state. Run the command 'mmhealth resolve event out\_of\_memory' once you recovered the system to remove this warning event from mmhealth.



# Use Case 2: RAM and OOM problem [4/13]



- A threshold **MemFree\_Rule** was triggered, then GUI daemon was killed:









Severity	Event Time	Reporting Node	Event Name	Event T...	Acti...	Entity Name
Info	02.03.20 15:10:09	<a href="#">test-21.localnet.com</a>	thresholds_normal	State		MemFree_Rule
Info	02.03.20 15:03:50	<a href="#">test-21.localnet.com</a>	gui_refresh_task_successful	State		<a href="#">test-21.localnet.com</a>
Info	02.03.20 15:03:35	<a href="#">test-21.localnet.com</a>	gui_up	State		<a href="#">test-21.localnet.com</a>
Warning	02.03.20 15:03:25	<a href="#">test-21.localnet.com</a>	out_of_memory	State		<a href="#">test-21.localnet.com</a>
Info	02.03.20 15:01:58	<a href="#">test-21.localnet.com</a>	ccr_comm_dir_ok	State		<a href="#">test-21.localnet.com</a>
Info	02.03.20 15:01:56	<a href="#">test-23.localnet.com</a>	ccr_quorum_nodes_ok	State		<a href="#">test-23.localnet.com</a>
Error	02.03.20 15:01:56	<a href="#">test-21.localnet.com</a>	gui_down	State		<a href="#">test-21.localnet.com</a>
Warning	02.03.20 15:01:56	<a href="#">test-21.localnet.com</a>	thresholds_warn	State		MemFree_Rule



# Use Case 2: RAM and OOM problem [5/13]



- OOM killer run was detected:

Severity	Event Time	Reporting Node	Event Name	Event T...	Acti...	Entity Name
 Info	02.03.20 15:10:09	<a href="#">test-21.localnet.com</a>	thresholds_normal	State		MemFree_Rule
 Info	02.03.20 15:03:50	<a href="#">test-21.localnet.com</a>	gui_refresh_task_successful	State		<a href="#">test-21.localnet.com</a>
 Info	02.03.20 15:03:35	<a href="#">test-21.localnet.com</a>	gui_up	State		<a href="#">test-21.localnet.com</a>
 Warning	02.03.20 15:03:25	<a href="#">test-21.localnet.com</a>	out_of_memory	State		<a href="#">test-21.localnet.com</a>
 Info	02.03.20 15:01:58	<a href="#">test-21.localnet.com</a>	ccr_comm_dir_ok	State		<a href="#">test-21.localnet.com</a>
 Info	02.03.20 15:01:56	<a href="#">test-23.localnet.com</a>	ccr_quorum_nodes_ok	State		<a href="#">test-23.localnet.com</a>
 Error	02.03.20 15:01:56	<a href="#">test-21.localnet.com</a>	gui_down	State		<a href="#">test-21.localnet.com</a>
 Warning	02.03.20 15:01:56	<a href="#">test-21.localnet.com</a>	thresholds_warn	State		MemFree_Rule



# Use Case 2: RAM and OOM problem [6/13]



- Later GUI went up again, RAM usage threshold was cleared on the next check:

Severity	Event Time	Reporting Node	Event Name	Event T...	Acti...	Entity Name
Info	02.03.20 15:10:09	test-21.localnet.com	thresholds_normal	State		MemFree_Rule
Info	02.03.20 15:03:50	test-21.localnet.com	gui_refresh_task_successful	State		test-21.localnet.com
Info	02.03.20 15:03:35	test-21.localnet.com	gui_up	State		test-21.localnet.com
Warning	02.03.20 15:03:25	test-21.localnet.com	out_of_memory	State		test-21.localnet.com
Info	02.03.20 15:01:58	test-21.localnet.com	ccr_comm_dir_ok	State		test-21.localnet.com
Info	02.03.20 15:01:56	test-23.localnet.com	ccr_quorum_nodes_ok	State		test-23.localnet.com
Error	02.03.20 15:01:56	test-21.localnet.com	gui_down	State		test-21.localnet.com
Warning	02.03.20 15:01:56	test-21.localnet.com	thresholds_warn	State		MemFree_Rule



# Use Case 2: RAM and OOM problem [7/13]



## ▪ Where is **MemFree\_Rule** defined?

Thresholds

<a href="#">+ Create Threshold</a>   <a href="#">Edit</a>   <a href="#">View Details</a>   <a href="#">Actions</a> ▼   <a href="#">Export</a>							
Name	↑	Target Type	Metric	Warning Level	Error Level	Filter	Active Events
DataCapUtil_Rule		Pool	Measurements - DataPool_capUtil	80 %	90 %		0
InodeCapUtil_Rule		Fileset	Measurements - Fileset_inode	80 %	90 %		0
MemFree_Rule		Node	Measurements - MemoryAvailable_percent	5 %			0
MetaDataCapUtil_Rule		Pool	Measurements - MetaDataPool_capUtil	80 %	90 %		0
SMBConnPerNode_Rule		Node	- Current connections		3000 #		0
SMBConnTotal_Rule		Generic	- Current connections		20000 #		0



# Use Case 2: RAM and OOM problem [8/13]



- I get it: GUI is great, but can I have CLI please?

```
[root@test-21 ~]# mmhealth thresholds list
```

```
active_thresholds_monitor: test-21.localnet.com
```

```
### Threshold Rules ###
```

rule_name	metric	error	warn	direction	filterBy	groupBy	sensitivity
InodeCapUtil_Rule	Fileset_inode	90.0	80.0	high		gpfs_cluster_name, gpfs_fs_name, gpfs_fset_name	300
DataCapUtil_Rule	DataPool_capUtil	90.0	80.0	high		gpfs_cluster_name, gpfs_fs_name, gpfs_diskpool_name	300
MemFree_Rule	MemoryAvailable_percent	None	5.0	low		node	300-min
SMBConnPerNode_Rule	current_connections	3000	None	high		node	300
SMBConnTotal_Rule	current_connections	20000	None	high			300
MetaDataCapUtil_Rule	MetaDataPool_capUtil	90.0	80.0	high		gpfs_cluster_name, gpfs_fs_name, gpfs_diskpool_name	300





# Use Case 2: RAM and OOM problem [9/13]



- Default thresholds are not enough?

```
[root@test-21 ~]# mmhealth thresholds add
Missing arguments.
Usage:
  mmhealth thresholds add { metric[:sum|avg|min|max|rate]|measurement
    [--errorlevel{threshold error limit} [--warnlevel{threshold warn limit}][--direction {high|low}}
    [--sensitivity {bucketsize}] [--hysteresis {percentage}]
    [--filterBy] [--groupBy ] [--name {ruleName}]
    [--errormsg {user defined action description}]
```



# Use Case 2: RAM and OOM problem [10/13]



- I want more details on the RAM usage over time:

```
[root@test-21 ~]# mmpperfmon query mem_memfree --bucket-size 60 2020-03-02-14:55:00 2020-03-02-15:05:00
```

Legend:

1: test-21.localnet.com|Memory|mem\_memfree

Row	Timestamp	mem_memfree
1	2020-03-02-14:56:00	954.7 MB
2	2020-03-02-14:57:00	954.9 MB
3	2020-03-02-14:58:00	753.5 MB
4	2020-03-02-14:59:00	85.4 MB
5	2020-03-02-15:00:00	85.4 MB
6	2020-03-02-15:01:00	85.3 MB
7	2020-03-02-15:02:00	992.8 MB
8	2020-03-02-15:03:00	1797.7 MB
9	2020-03-02-15:04:00	1511.2 MB
10	2020-03-02-15:05:00	1427.2 MB



# Use Case 2: RAM and OOM problem [11/13]



## ▪ What caused OOM?

```
[root@test-21 ~]# mmpfmon report top --bucket-size 10 2020-03-02-15:01:30
```

Top values in format:

{process name | PID} {cpu per mil} {memory per mil}

```
Time                test-21.localnet.com
```

```
20-03-02-14:59:30  tail 156 627
```

```
    kswapd0 102 0
```

```
    mmfsd 19 268
```

```
    java 18 9
```

```
    pmsensors 9 0
```

```
    kworker/1:1H 4 0
```

```
[...]
```

→ we used **tail /dev/zero** to trigger the OOM killer



# Use Case 2: RAM and OOM problem [12/13]



- IMPORTANT: **out\_of\_memory** is a resolvable event!
- After resolving the issue (e.g. banning users for **tail /dev/zero** calls) resolve the event:

```
[root@test-21 ~]# mmhealth event resolve out_of_memory  
Successfully resolved event out_of_memory for entity  with event out_of_memory_ok.
```



# Use Case 2: RAM and OOM problem [13/13]



- Push on failure to be informed ASAP:
  - A sample events callback script:

```
[root@g5040-11 ~]# vim /var/mmfs/etc/eventsCallback
#!/bin/bash
# <version> <date> <time> <timezone> <event> <component> <id> <severity> <state> <message> <arguments>
if [[ $# -gt 10 ]]; then
    #echo "Got CLI parameters: $@" >> /root/callback.txt
    if [[ "$5" == 'thresholds_warn' && "${7#*TriggerCallbackRule}" != "$7" ]]; then
        echo -e "Threshold details:\n${@}" | mail -s "Threshold $7 triggered" admin@company.com
    fi
fi
```

- The resulting e-mails then look this way:
- Or use GUI ;)



**Threshold TriggerCallbackRule/g5040-11.localnet.com triggered**

**root** to: Pavel Safr

09.07.2019 17:55

[Show Details](#)

Threshold details:

```
1 2019-07-09 17:55:33.363401 CEST thresholds_warn threshold
TriggerCallbackRule/g5040-11.localnet.com W D The value of mem_memfree
for the component(s) TriggerCallbackRule/g5040-11.localnet.com exceeded
threshold warning level 2000000 defined in TriggerCallbackRule.
2000000,mem%5Fmemfree,TriggerCallbackRule
```



# Use Case 3: No access to a share [1/4]



## ▪ Situation:

- The customer cannot access an SMB share

```
smb: \foo\> ls
NT_STATUS_ACCESS_DENIED listing \foo\*
```

- The customer checks the permissions and they seem to be OK

## ▪ Wish:

- Resolve the issue and get an access to the SMB share

## ▪ Solution:

- Use **mmprotocoltrace** + **mmcallhome**



# Use Case 3: No access to a share [2/4]



- The customer contacts the support and is asked to provide tracing data for the operation
- He starts the trace:

```
[root@ces5050-41 ~]# mmprotocoltrace start network smb winbind -f --client-ips 9.155.106.232
Setting up traces
Trace 'a078911e-1571-497b-803f-974f47bc6750' created successfully for 'network'
Trace '904f871e-35f1-4388-955d-25f0c8ceda47' created successfully for 'smb'
Trace '468fd3f8-770e-4650-b759-22776dce6f27' created successfully for 'winbind'
...
Trace ID:          a078911e-1571-497b-803f-974f47bc6750
State:             ACTIVE
...
```

- Recreates the issue:

```
[root@ces5050-41 ~]# smbclient //localhost/smbexport -U MZDOM\\User%password
Try "help" to get a list of possible commands.
smb: \foo\> ls
NT_STATUS_ACCESS_DENIED listing \foo\*
```



# Use Case 3: No access to a share [3/4]



- Subsequently he stops the trace:

```
[root@ces5050-41 ~]# mmprotocoltrace stop
Stopping traces

...

Trace tar file has been written to '/tmp/mmfs/smb.trace.20200303_060522.tar.gz'
```

- **If he wants**, he can now unpack the archive and go through the contained trace logs
  - A message in the SMB log states that the ID mapping (Windows SID ↔ UNIX ID) does not exist
- Solution: the user manually configures the missing mapping and the share works





# Use Case 3: No access to a share [4/4]



- Subsequently he stops the trace:

```
[root@ces5050-41 ~]# mmprotocoltrace stop
Stopping traces

...

Trace tar file has been written to '/tmp/mmfs/smb.trace.20200303_060522.tar.gz'
```

- **Alternatively**, he shares the trace data with the support team using mmcallhome:

```
[root@ces5050-41 ~]# mmcallhome run SendFile --file /tmp/mmfs/smb.trace.20200303_060522.tar.gz
Running sendFile... (In case of network errors, it may take over 20 minutes for retries.)
Successfully uploaded the given file
Run mmcallhome status list --verbose to see the package name
```

– Root cause found: missing ID mapping (Windows SID ↔ UNIX ID)

- Solution: the user manually configures the missing mapping and the share works



# Use Case 4: Unknown setting changed [1/3]



- **Situation:**

- My cluster is extremely slow since yesterday
- I know, that I might have run a command in a wrong terminal a few days ago

- **Wish:**

- Find out, what exactly was changed and change it back

- **Solution:**

- `mmcallhome status diff`



# Use Case 4: Unknown setting changed [2/3]



- Does **mmhealth** detect any problems?

```
[root@g5050-11 ~]# mmhealth node show
```

```
Node name:      g5050-11d.localnet.com
Node status:    HEALTHY
Status Change:  Now
IP mode:        LEGACY
```

Component	Status	Status Change	Reasons
<hr/>			
GPFS	HEALTHY	Now	-
NETWORK	HEALTHY	1 day ago	-
FILESYSTEM	HEALTHY	Now	-
DISK	HEALTHY	Now	-
CALLHOME	HEALTHY	8 hours ago	-
PERFMON	HEALTHY	1 day ago	-
THRESHOLD	HEALTHY	1 day ago	-



# Use Case 4: Unknown setting changed [3/3]



- Can perhaps call home help me?

```
[root@g5050-11 ~]# mmcallhome run GatherSend --task daily # we want a checkpoint NOW!
One time run completed with success

[root@g5050-11 ~]# mmcallhome status diff --last-days 2

Active Config      (modified)
  Maxfilestocache   : 100010  --> 10

Cluster Data       (modified)
  Maxfilestocache   : 100010  --> 10
```

- Solution: increase **maxFilesToCache** back to the normal + restart **mmfsd**



# Improvements in 5.0.4 / 5.0.5 [1/4]



- Call Home:
  - **Configuration diff**
  - **Test connection for ungrouped nodes**
  - **Add-/delete- group nodes**
  - Mandatory config + TIPs
  - 5.0.4 Increased recommended group size to 128 nodes
  - 5.0.4 Removed package gpfs.callhome
  - 5.0.4 Hardware call home (ESA) trigger for specific events



# Improvements in 5.0.4 / 5.0.5 [2/4]



- mmprotocoltrace:
  - **Increased usability**
  - 5.0.4 Improved backend engine
- CES
  - 5.0.4.2 HDFS integration
  - 5.0.4 CES disable confirmation



# Improvements in 5.0.4 / 5.0.5 [3/4]



## ■ Perfmon:

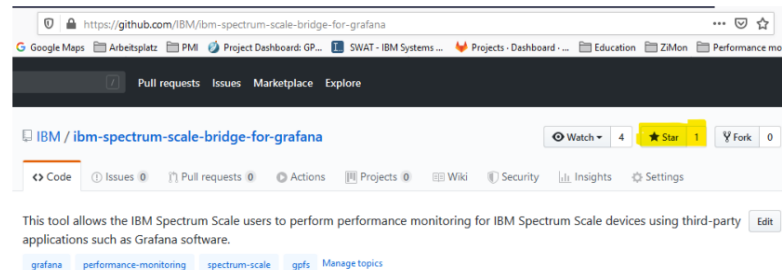
- Default sensors filter to reduce keypace with Docker (volatile naming) (originated from a service ticket)

Workaround: mmperfmon delete --expired-keys

- **Top-k Phase 1: top CPU consumers**

+ also top-log (CPU + MEM) in gpfs.snap

- **5.0.4** Added disk write latency measurements



## ■ Grafana Bridge Migration to GitHub:

<https://github.com/IBM/ibm-spectrum-scale-bridge-for-grafana>



# Improvements in 5.0.4 / 5.0.5 [4/4]



- mmhealth:
  - numactl monitoring
  - OOM monitoring
  - ESS3000 monitoring
  - 5.0.4 DNS issues monitoring (specifically AD)
  - 5.0.4 ExtAuthMonitor: UDP + TCP checks
  - 5.0.4 Verification that required file systems are mounted







- Improved usability for mmprotocoltrace
  - Expected behavior = real behavior
  - Usability shortcuts to streamline the usage
  - Less commands do more
  - More details about failures



# mmprotocoltrace Usability [2/7]: start



- Block start step until all nodes are ACTIVE  
(otherwise customers start their workloads before the tracing is ready)
- At the end of the execution show the current trace state

## BEFORE

```
[root@ces5033-31 ~]# mmprotocoltrace start network
Setting up traces
Trace '3bf4f361-0005-4469-af0a-f0097a95608e' created successfully for 'network'

[root@ces5033-31 ~]# # manually poll "mmprotocoltrace status network"

[root@ces5033-31 ~]# mmprotocoltrace status network
Trace ID:      3bf4f361-0005-4469-af0a-f0097a95608e
State:         ACTIVE
User ID:       root
Protocol:      network
Start Time:    14:23:29 09/10/19
End Time:      14:33:29 09/10/19
Client IPs:
Origin Node:   ces5033-31.localnet.com
Syscall:       False
Syscall Only: False
Nodes:
  Node Name:    ces5033-31.localnet.com
  State:        ACTIVE
  Trace Location: /tmp/mmfs/network.20191009_142329.trc
```

## AFTER

```
[root@ces5040-41 ~]# mmprotocoltrace start network
Setting up traces
Trace '4300981f-59cb-4eb7-bc56-e04e90f0801c' created successfully for 'network'

Waiting for all participating nodes...

Trace ID:      4300981f-59cb-4eb7-bc56-e04e90f0801c
State:         ACTIVE
Protocol:      network
Start Time:    14:22:12 09/10/19
End Time:      14:32:12 09/10/19
Trace Location: /tmp/mmfs/network.20191009_142212.trc
Origin Node:   ces5040-41.localnet.com
Nodes:
  Node Name:    ces5040-41.localnet.com
  State:        ACTIVE
```

# mmprotocoltrace Usability [3/7]: start



- mmprotocoltrace start:
  - If any of the nodes is FAILED, the trace is aborted

## BEFORE

```
[root@ces5033-31 ~]# mmprotocoltrace start network
Setting up traces
Trace 'cce2d03f-b538-4100-b325-7ff35e65277d' created successfully for 'network'

[root@ces5033-31 ~]# # manually poll "mmprotocoltrace status network"

[root@ces5033-31 ~]# mmprotocoltrace status network
Trace ID:      cce2d03f-b538-4100-b325-7ff35e65277d
State:         ACTIVE
User ID:       root
Protocol:      network
Start Time:    14:48:57 09/10/19
End Time:      14:58:57 09/10/19
Client IPs:
Origin Node:   ces5033-31.localnet.com
Syscall:       False
Syscall Only: False
Nodes:
  Node Name:    ces5033-31.localnet.com
  State:        FAILED
  Trace Location: /tmp/mmfs/network.20191009_144857.trc
  Failure Reason: Could not find the tcpdump executable

[root@ces5033-31 ~]# # manually stop + clear the trace
```

## AFTER

```
[root@ces5040-41 ~]# mmprotocoltrace start network
Setting up traces
Trace '0ae4a665-a716-4ff1-83a4-2634767f3072' created successfully for 'network'

Waiting for all participating nodes...
For the protocol 'network' some of the nodes failed to start the trace.
Aborting the trace

Waiting for all participating nodes...

Trace ID:      0ae4a665-a716-4ff1-83a4-2634767f3072
State:         FAILED
Protocol:      network
Start Time:    14:53:42 09/10/19
End Time:      15:03:42 09/10/19
Trace Location: /tmp/mmfs/network.20191009_145342.trc
Origin Node:   ces5040-41.localnet.com
Nodes:
  Node Name:    ces5040-41.localnet.com
  State:        FAILED
  Failure Reason: Could not find the tcpdump executable
```

# mmprotocoltrace Usability [4/7]: start



- mmprotocoltrace start:
  - Auto clear traces on start with -f or with a CLI confirmation (yes/no)

## BEFORE

```
[root@ces5033-31 ~]# mmprotocoltrace start network
Setting up traces
Trace '3bf4f361-0005-4469-af0a-f0097a95608e' for protocol 'network' already
exists
Something failed when creating traces

[root@ces5033-31 ~]# # manually run "mmprotocoltrace clear network"

[root@ces5033-31 ~]# # manually run "mmprotocoltrace start network" again
```

## AFTER

```
[root@ces5040-41 ~]# mmprotocoltrace start network
For the following protocols traces are still running: network

Starting new traces requires that all previous traces for the corresponding
protocols are cleared.
Do you want to clear these traces? (yes/no - default no): yes
...
```



# mmprotocoltrace Usability [5/7]: defaults



- The subcommands "stop", "status" and "clear" now assume "all" if no protocols are specified

## BEFORE

```
[root@ces5033-31 ~]# mmprotocoltrace status
Too few arguments for this command
Spectrum Scale - Protocol Trace Management Tool
Usage:
  mmprotocoltrace start <identifier> [<identifier> ...]
                        [-c <clientIPs>] [-d <duration>] [-l <logFileDir>] [-N <nodes>] [-f]
  mmprotocoltrace stop <identifier> [<identifier> ...]
  mmprotocoltrace status <identifier> [<identifier> ...] [-v]
...
specified the option -f while starting the trace.

[root@ces5033-31 ~]# # run explicitly "mmprotocoltrace status network"
```

## AFTER

```
[root@ces5040-41 ~]# mmprotocoltrace status
Trace ID:      88186b61-a0ac-495b-aba9-8b8f361a6b1b
State:         ACTIVE
Protocol:      network
Start Time:    15:12:12 09/10/19
End Time:      15:22:12 09/10/19
Trace Location: /tmp/mmfs/network.20191009_151212.trc
Origin Node:   ces5040-41.localnet.com
Nodes:
  Node Name:   ces5040-41.localnet.com
  State:       ACTIVE

  Node Name:   ces5040-42.localnet.com
  State:       ACTIVE

  Node Name:   ces5040-43.localnet.com
  State:       ACTIVE

  Node Name:   ces5040-44.localnet.com
  State:       ACTIVE
```



# mmprotocoltrace Usability [6/7]: status



- After a trace is over, shows where to find results (previously: only shown once when stopping the trace)

## BEFORE

```
[root@ces5033-31 ~]# mmprotocoltrace status network
Trace ID:      2343bf34-175f-46ef-9734-dbf9003c6cc
State:         DONE
User ID:       root
Protocol:      network
Start Time:    14:59:13 09/10/19
End Time:      15:09:13 09/10/19
Client IPs:
Origin Node:   ces5033-31.localnet.com
Syscall:       False
Syscall Only: False
Nodes:
  Node Name:    ces5033-33.localnet.com
  State:        DONE
  Trace Location: /tmp/mmfs/network.20191009_145913.trc

  Node Name:    ces5033-31.localnet.com
  State:        DONE
  Trace Location: /tmp/mmfs/network.20191009_145913.trc

  Node Name:    ces5033-32.localnet.com
  State:        DONE
  Trace Location: /tmp/mmfs/network.20191009_145913.trc
```

## AFTER

```
[root@ces5040-41 ~]# mmprotocoltrace status network
Trace ID:      e27e0d87-35d4-434f-bdf3-726dfcf4acd9
State:         DONE
Protocol:      network
Start Time:    14:59:03 09/10/19
End Time:      15:09:03 09/10/19
Trace Location: /tmp/mmfs/network.20191009_145903.trc
Origin Node:    ces5040-41.localnet.com
Trace results file: ces5040-41.localnet.com:/tmp/mmfs/network.trace.20191009_150347.tar.gz
Nodes:
  Node Name:    ces5040-41.localnet.com
  State:        DONE

  Node Name:    ces5040-42.localnet.com
  State:        DONE

  Node Name:    ces5040-43.localnet.com
  State:        DONE
```

# mmprotocoltrace Usability [7/7]: errors



- If an unhandled exception occurs, a specific log location is now referenced (before: "see logs")

## BEFORE

```
[root@ces5033-31 ~]# mmprotocoltrace stop network
Stopping traces
Trace '67f98781-cblf-465e-b7dc-f0f9e7c63632' stopped for network
Waiting for traces to complete
Waiting for node 'ces5033-31.localnet.com'
Waiting for node 'ces5033-32.localnet.com'
...
Finishing trace '67f98781-cblf-465e-b7dc-f0f9e7c63632'
Successfully copied file from 'ces5033-31.localnet.com:/tmp/mmfs/network.20191009_155043.trc'
Successfully copied file from 'ces5033-32.localnet.com:/tmp/mmfs/network.20191009_155043.trc'
There were issues gathering files. Check logs for more info
Trace tar file has been written to '/tmp/mmfs/network.trace.20191009_155235.tar.gz'
```

## AFTER

```
[root@ces5040-41 ~]# mmprotocoltrace stop network
Stopping traces

Trace '653a617d-ff01-4443-9aaa-02f72fee5ee7' stopped for network

Waiting for all participating nodes...

Collecting data from the participating nodes '653a617d-ff01-4443-9aaa-02f72fee5ee7'
Collected 'ces5040-41.localnet.com:/tmp/mmfs/network.20191009_155052.trc'
Collected 'ces5040-42.localnet.com:/tmp/mmfs/network.20191009_155052.trc'
There were issues gathering files. Check /var/adm/ras/mmprotocoltrace.log for more info
Trace tar file has been written to '/tmp/mmfs/network.trace.20191009_155140.tar.gz'
```

# CALLHOME: add-/delete-node



- Previously:
  - Need to remove nodes from a group or add to it?
  - Remove the group and then create a new one
- Now:

```
gpfsadmin@sudo5050-21

[root@sudo5050-21 gpfsadmin]# mmcallhome group list
callHomeGroup  callHomeNode  callHomeChildNodes
-----
autoGroup_1    sudo5050-21    sudo5050-21,sudo5050-22,sudo5050-23,sudo5050-24

[root@sudo5050-21 gpfsadmin]# mmcallhome group change autoGroup_1 --delete-nodes 3,4
Call home group autoGroup_1 has been changed
[root@sudo5050-21 gpfsadmin]# mmcallhome group list
callHomeGroup  callHomeNode  callHomeChildNodes
-----
autoGroup_1    sudo5050-21    sudo5050-21,sudo5050-22 

[root@sudo5050-21 gpfsadmin]# mmcallhome group change autoGroup_1 --add-nodes 10.0.100.23
Call home group autoGroup_1 has been changed
[root@sudo5050-21 gpfsadmin]# mmcallhome group list
callHomeGroup  callHomeNode  callHomeChildNodes
-----
autoGroup_1    sudo5050-21    sudo5050-21,sudo5050-22,
```





# CALLHOME: test outside of groups



- Previously:

- `mmcallhome test connection` only for call home group members
- Other nodes can do this implicitly, e.g. using `mmcallhome group add` or using undocumented internal commands

- Now:

## BEFORE

```
[root@ces5041-41 ~]# mmcallhome group list
No callhome group defined.

[root@ces5041-41 ~]# mmcallhome test connection
Failed to obtain a valid call home node.
A call home group for this node needs to be defined using group command.

[root@ces5041-41 ~]# # implicit workaround:
[root@ces5041-41 ~]# mmcallhome group add g1 ces5041-41
Call home child nodes = ces5041-41.localnet.com
Call home group g1 has been created
[root@ces5041-41 ~]# # now we have to clean up:
[root@ces5041-41 ~]# mmcallhome group delete g1
Successfully deleted
```

## AFTER

```
[root@sudo5050-21 gpfsadmin]# mmcallhome group list
No callhome group defined.

[root@sudo5050-21 gpfsadmin]# mmcallhome test connection
## Starting connectivity test between the current node and IBM
(the current node is not a part of a call home group)

Current proxy status: disabled

Connection: OK
```

# Tips and Tricks: Call Home [1/2]



- Who has Spectrum Scale / ESS Software Call Home enabled?
- Running 4.2.3.8, 5.0.0 or higher?
  - Please enable Call Home
- Why?
  - Faster service response times
  - Proactive issues detection
  - Easier and faster sharing data with support
  - Configuration changes analysis
  - Better test coverage for your configuration
- Automatic service tickets for hardware failures (ESS only)



# Tips and Tricks: Call Home [2/2]



- Who has Spectrum Scale / ESS Software Call Home enabled?
- Running 4.2.3.8, 5.0.0 or higher?
  - Please enable Call Home
- How?
  - Spectrum Scale installer
  - Spectrum Scale GUI
  - mmcallhome command

The image shows a terminal window and the Spectrum Scale GUI. The terminal window displays the output of the `mmcallhome --help` command, showing usage information and a list of subcommands: `group` (Administer the group), `capability` (Enable, disable, or query capabilities), `info` (Administer customer data), `proxy` (Administer proxy configuration), and `schedule` (Administer call home scheduling). The GUI is the 'Call Home' configuration page, which includes sections for enabling Call Home, setting the Call home node (currently `gpfs-11.localnet.com`), Company Information (Company name, E-mail, Customer ID, Country), Proxy Information (Proxy host, Proxy port, Proxy username, Proxy password, and Proxy authentication), and buttons for 'Test Connection' and 'Configure Call Home'.

```
[root@gpfs-11 installer]# ./spectrumscale callhome
usage: spectrumscale callhome [-h]
                               {enable,disable,list}

[root@gpfs-11 ~]# mmcallhome --help
Usage:
  mmcallhome group      Administer the group
  mmcallhome capability Enable, disable, or query capabilities
  mmcallhome info       Administer customer data.
  mmcallhome proxy      Administer proxy configuration.
  mmcallhome schedule   Administer call home scheduling
```



Questions ?



Thank you for your attention !!!

# Legal notices



Copyright © 2019 by International Business Machines Corporation. All rights reserved.

No part of this document may be reproduced or transmitted in any form without written permission from IBM Corporation.

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This document could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or program(s) described herein at any time without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectual property rights, may be used instead.

THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER OR IMPLIED. IBM LY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted, if at all, according to the terms and conditions of the agreements (e.g., IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. 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 in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. IBM makes no representations or warranties, ed or implied, regarding non-IBM products and services.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing  
IBM Corporation  
North Castle Drive  
Armonk, NY 1 0504- 785  
U.S.A.



# Information and trademarks



IBM, the IBM logo, ibm.com, IBM System Storage, IBM Spectrum Storage, IBM Spectrum Control, IBM Spectrum Protect, IBM Spectrum Archive, IBM Spectrum Virtualize, IBM Spectrum Scale, IBM Spectrum Accelerate, Softlayer, and XIV are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at <http://www.ibm.com/legal/copytrade.shtml>

The following are trademarks or registered trademarks of other companies.

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

IT Infrastructure Library is a Registered Trade Mark of AXELOS Limited.

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a registered trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

Cell Broadband Engine is a trademark of Sony Computer Entertainment, Inc. in the United States, other countries, or both and is used under license therefrom.

ITIL is a Registered Trade Mark of AXELOS Limited.

UNIX is a registered trademark of The Open Group in the United States and other countries.

\* All other products may be trademarks or registered trademarks of their respective companies.

## Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice.

Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the 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.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

This presentation and the claims outlined in it were reviewed for compliance with US law. Adaptations of these claims for use in other geographies must be reviewed by the local country counsel for compliance with local laws.



# Special notices



This document was developed for IBM offerings in the United States as of the date of publication. IBM may not make these offerings available in other countries, and the information is subject to change without notice. Consult your local IBM business contact for information on the IBM offerings available in your area.

Information in this document concerning non-IBM products was obtained from the suppliers of these products or other public sources. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. Send license inquiries, in writing, to IBM Director of Licensing, IBM Corporation, New Castle Drive, Armonk, NY 10504-1785 USA.

All statements regarding IBM future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

The information contained in this document has not been submitted to any formal IBM test and is provided "AS IS" with no warranties or guarantees either expressed or implied.

All examples cited or described in this document are presented as illustrations of the manner in which some IBM products can be used and the results that may be achieved. Actual environmental costs and performance characteristics will vary depending on individual client configurations and conditions.

IBM Global Financing offerings are provided through IBM Credit Corporation in the United States and other IBM subsidiaries and divisions worldwide to qualified commercial and government clients. Rates are based on a client's credit rating, financing terms, offering type, equipment type and options, and may vary by country. Other restrictions may apply. Rates and offerings are subject to change, extension or withdrawal without notice.

IBM is not responsible for printing errors in this document that result in pricing or information inaccuracies.

All prices shown are IBM's United States suggested list prices and are subject to change without notice; reseller prices may vary.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

Any performance data contained in this document was determined in a controlled environment. Actual results may vary significantly and are dependent on many factors including system hardware configuration and software design and configuration. Some measurements quoted in this document may have been made on development-level systems. There is no guarantee these measurements will be the same on generally-available systems. Some measurements quoted in this document may have been estimated through extrapolation. Users of this document should verify the applicable data for their specific environment.

