Monitoring & Serviceability Enhancements in IBM Spectrum Scale

Spectrum Scale German User Meeting 2022
Cologne, Germany – October 19-20, 2022

Pavel Safre (IBM)

Disclaimer

- This information is provided on an "AS IS" basis without warranty of any kind, express or implied, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. Some jurisdictions do not allow disclaimers of express or implied warranties in certain transactions; therefore, this statement may not apply to you.

- 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.

IBM Offering Management

Agenda

- Benefits of monitoring and proactive services

- Improvements for Call Home & Proactive Services

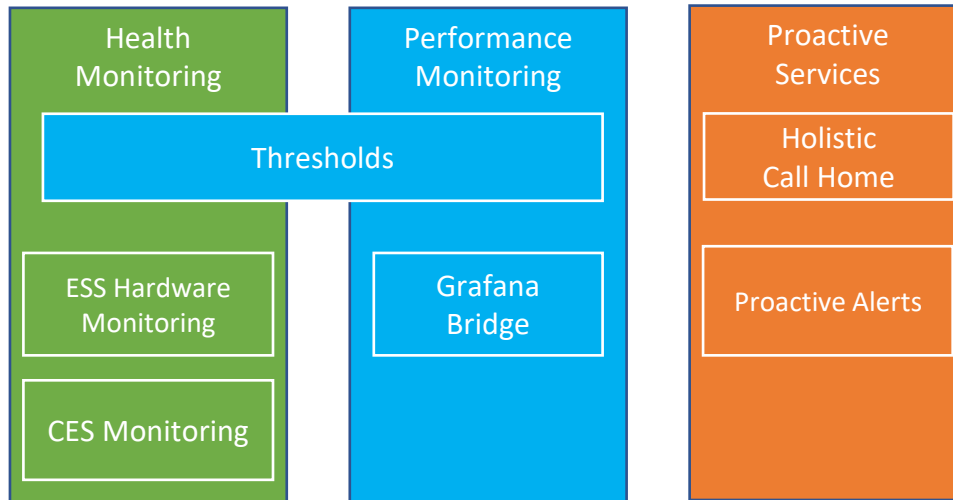- Monitoring Improvements in latest releases

- 5.1.6

Benefits of monitoring and proactive services [1/4]

- Many things can go wrong in a storage cluster:
  - Network issues, disks get broken, settings are wrong, etc.
- There are 3 pillars of monitoring assistance for you in Scale:

1. System Health monitoring
2. Performance monitoring
3. Proactive Services

| Health Monitoring	Performance Monitoring	Proactive Services
Thresholds	Thresholds	Holistic Call Home
ESS Hardware Monitoring	Grafana Bridge	Proactive Alerts
CES Monitoring		

1. **System Health monitoring**

   – Hundreds of checks for dozens of components

   – Informs you about any problems found and what to do to fix it

     ▪ On CLI via `mmhealth cluster show` / `mmhealth node show`

     ▪ In your browser via Scale GUI

     ▪ In your own monitoring tools via Webhooks, Script Callouts, etc.
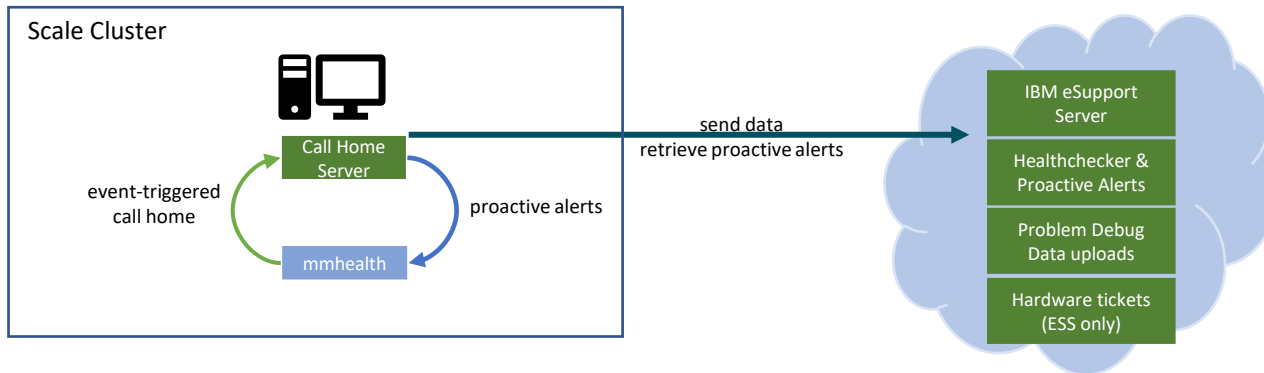
2. **Performance Monitoring**

    – Overview the cluster performance for dozens of metrics

    – Set thresholds to generate health events for cases, relevant for you

    – Scale GUI shows you the details

    – Grafana Bridge allows more customization and integrations into your Grafana instances
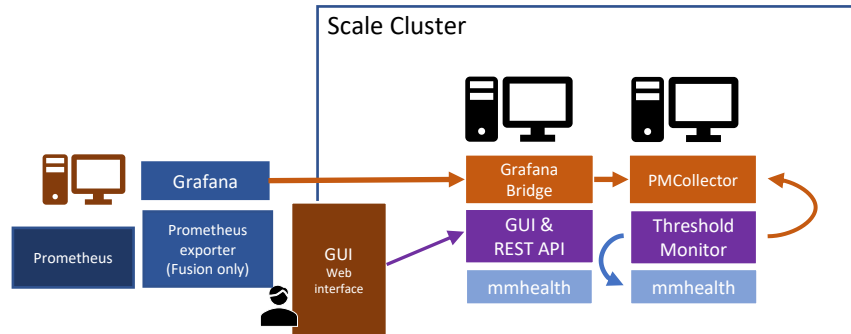
3. **Proactive Services**

    – Extend monitoring every day, especially for urgent topics (e.g. FLASHes)

    – Detect known problems, misconfigurations and best practice violations

    – Provide immediate benefit without a need to wait for an update

    – Synergize the data from call home with the IBM expertise off premise

Scale Cluster

Call Home Server

event-triggered call home

proactive alerts

mmhealth

send data
retrieve proactive alerts

IBM eSupport Server

Healthchecker & Proactive Alerts

Problem Debug Data uploads

Hardware tickets (ESS only)

Proactive Alerts: Healthchecker

Scale Health Checker is an IBM internal Cloud application which is used to
detect best practice violations, common misconfigurations and known problems
by running <u>rules against incoming call home data</u> and sending proactive alerts <u>back to your system (>=5.1.2)</u>

Get informed proactively if your system…

- is impacted by a new issue

   - e.g. when we learn about a new problem from customer A, we can identify
     which other customers are impacted by the same problem and send them a proactive alert

- is impacted by a Spectrum Scale Flash

- is not following "best practices" and could be tuned to improve performance, etc. (Tip)

- is misconfigured which might lead to problems in the future

- Fully integrated with the Spectrum Scale monitoring infrastructure (GUI, mmhealth, event notification, webhook,..)

Benefit from continuous enhancements and new proactive alerts without having to update to a new Scale release !

Healthchecker rules

Call Home Data → Known Issues / Best Practice Violations / Flashs/APARs → Proactive Alerts

Proactive Alerts: mmhealth GUI & CLI

GUI shows proactive alerts like all other events or tips

New **HEALTHCHECK** mmhealth component on Call Home server node

Proactive Alerts Flow

On-prem customer System (ESS, Scale, Fusion)

IBM Cloud
ECuRep

Data Preprocessing

AoATool,
EventGenerator

raw command
output & logs

inventory (JSON)

Health Monitoring

on-prem +
dynamic
alerts

Call Home

data gathering

- GUI Alerts
- mmhealth
- Event Notification
- Fusion Event View

Analysis & Alerting

Healthchecker

Rules

Reporting

Watson Studio

Spectrum Scale Development & Test

Healthcheck Monitor
retrieves Alert Events
(Rest API)

Proactive Alerts

Support Mesh

slack

IBM Support
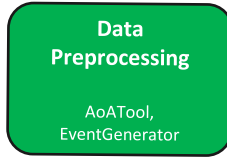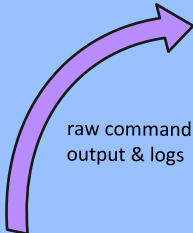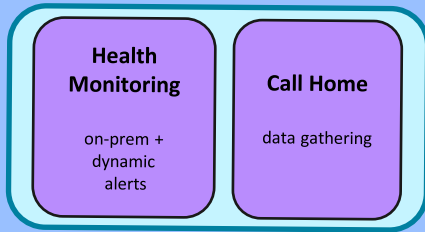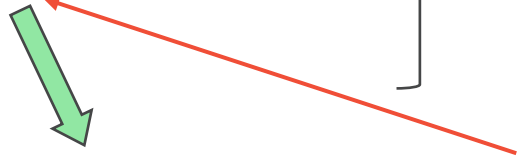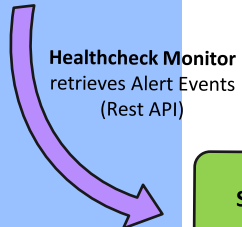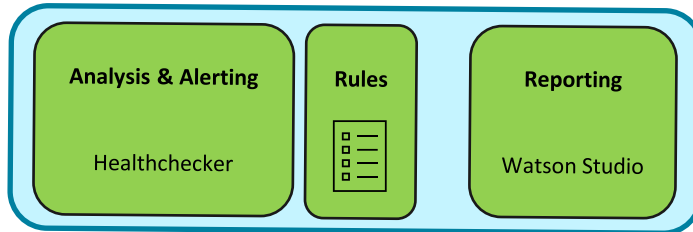
ESS Unified Call Home

ESS used to have 2 separate call home stacks:

– Hardware Call Home (ESA) -> automatically open tickets for hardware failures

– Software Call Home

<u>Disadvantages:</u> more complex installation and maintenance, inconsistent configurations, low usability

ESS 3500 comes with a single "Call Home" stack

Simplified configuration, no inconsistencies , higher usability

– Software Call Home now has the capability to open hardware tickets in Salesforce

– Single management command (mmcallhome)

– ESA will still be used for older ESS systems (e.g. ESS 5000, ESS 3000, etc.)

Health monitoring

↓ Hardware failure event

Ticket data collection

↓ Send to Call Home node

Unified Call Home

Open hardware ticket | Upload data to ECuRep
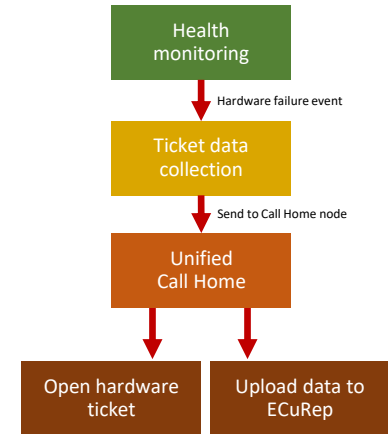
Reminder: Please Enable Call Home

Enable Call Home on your Spectrum Scale clusters !

- Get the benefits now! and help our service team

Call Home Client Value:

– **Proactively detect issues (proactive alerts)**

– Improved service response time and quicker resolution due to

  - Data immediately available to analyse a ticket (Call Home, FTDC)
  - Tools and Salesforce linkage to give support team a quick overview on customer system
  - Automatically open PMRs for hardware failures

– Better customer support experience (consumability/ease of use)

  - Easy and fast sharing of diagnostics data with support
  - Easy way to see history of system configuration changes (config diff)

– Provide customer insights reports for support and development

IBM eSupport / EcuRep

HTTPS Upload

Call Home Server

Spectrum Scale Cluster
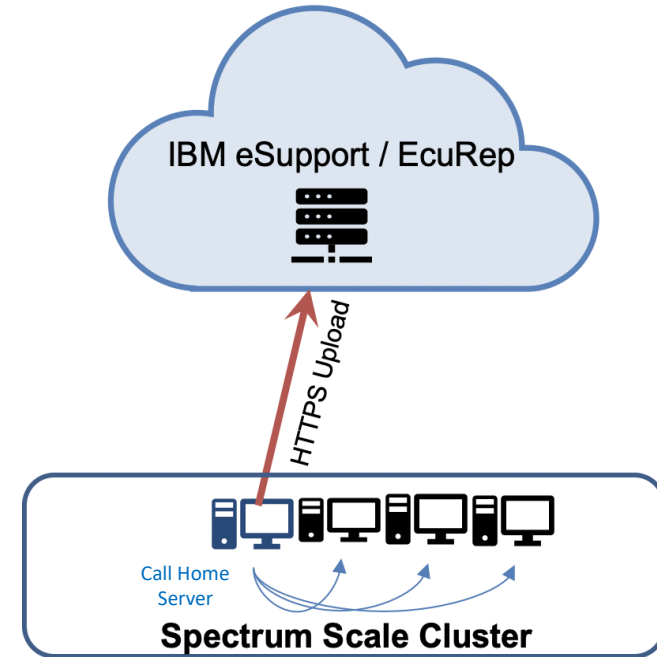
AFM memory queue alerts in mmhealth

- **Task:**

  - Monitor the AFM in queue memory usage on Gateway nodes and raise events upon reaching or exceeding the specified limits.

- **Solution:**

  - New default threshold rule **AFMInQueue_Rule**

```
Warning level        = 80% of afmHardMemThreshold (default 5G)
Error level          = 90% of afmHardMemThreshold (default 5G)
```

  - Based on the perfmon metric **gpfs_afm_used_q_memory**

  - Current limitation: the same memory config is expected for all Gateway nodes.

Improved NSD Disk Monitoring [1/2]

- **Task:**

  – mmhealth should report a specific event for different kinds of **diskDown, diskFailure and diskIOErr** callbacks

- **Solution:**

  – Added many events for different cases:

  - disk_down_io
  - disk_down_unexpected
  - disk_down_unknown
  - disk_down_change

  - disk_down_del
  - disk_down_rpl
  - disk_failed_cb
  - disk_io_err_cb

```
disk_down_io          ERROR     Disk nsd1 is reported as down because of an I/O issue. FS=gpfs0,..
disk_down_unexpected  ERROR     Disk nsd1 is reported as unexpected down. FS=gpfs0
disk_down_change      ERROR     Disk nsd1 is reported as down because the configuration changed. FS=gpfs0
```

Improved NSD Disk Monitoring [2/2]

- **Task:**

  - mmhealth should report if there are not enough healthy descriptor disks for a filesystem

- **Solution:**

  - Added a quorum check per Scale filesystem: desc_disk_quorum_fail / desc_disk_quorum_ok

  - **The new features for this story apply not only to SNC but to all configurations that use replication.**

```
[root@g5140-21 ~]# mmhealth event show desc_disk_quorum_fail
Event Name:             desc_disk_quorum_fail
Event ID:               999577
Description:            Sufficient healthy descriptor disks are not found.
Cause:                  Sufficient healthy descriptor disks are not found. No quorum is found for
the file system.
User Action:            Check the health state of disks, which are declared as descriptor disks
for the file system. A missing quorum might lead to data access loss. For more information, see the
'Disk issues' section in the 'Problem Determination Guide'.
Severity:               WARNING
State:                  DEGRADED
```

Stretch Cluster Monitoring Improvements [1/2]

- **Task:**

  - Extend the stretch cluster monitoring over the 5.1.3 version of it, improving the functionality of the site aware health monitoring

  - Add new checks and events for checking of the descriptor disks per site

- **Solution:**

  - Added checks if it is detected that a filesystem does not have enough desc disks for quorum

    - site_fs_desc_warn / site_fs_desc_fail / site_fs_quorum_fail

  - Added checks to detect if a descriptor disk is not configured in a failure group:

    - disk_fs_desk_ok / disk_fs_desc_missing

  - Added a mmsysmonitor.conf setting **stretch. percent_failed_raise_degraded** (default: 50)
    to allow to change the percentage of failed nodes, at which a DEGRADED event for a site is raised

Stretch Cluster Monitoring Improvements [2/2]

```
[root@gimli-master-1 mmsysmon]# mmhealth cluster show stretchcluster -v
Component           Node                            Status       Reasons
----------------------------------------------------------------------------------------------

STRETCHCLUSTER      gimli-master-1.fyre.ibm.com     DEGRADED     site_fs_quorum_fail
  Frankfurt                                         DEGRADED     site_fs_quorum_fail
  Ehningen                                          HEALTHY      -
  TIEBREAKER                                        HEALTHY      -


[root@gimli-master-1 mmsysmon]# mmhealth event show site_fs_quorum_fail
Event Name:             site_fs_quorum_fail
Event ID:               998421
Description:            Not enough healthy descriptor disk(s) found.
Cause:                  Not enough healthy descriptor disk(s) for quorum for the filesystem at the site.
User Action:            Check the health state of disks which are declared as descriptor disks for the
filesystem. This may lead to data loss. Refer to the 'Disk issues' section in the 'Problem Determination
Guide'
Severity:               ERROR
State:                  FAILED
```

Detection of SCSI passthrough hang issues

- **Task:**

  - mmhealth should report a specific event for disk IO hangs due to SCSI passthrough issues

- **Solution:**

  - Added a new event **passthrough_query_hang** for this case

  - This event is based on the callback "diskIOHang"

    - Just like the general disk IO hang event **kernel_io_hang_detected.**

    - In 5.1.4 the callback was extended to include a reason code, which is analyzed by mmsysmon.

Write cache enabled monitoring on ESS [1/2]

- If write cache is enabled, a data loss is highly likely for spinning disks after a power loss.

- Added monitoring to warn about the potential data loss if write cache is enabled for spinning disks:

  – Write cache enabled –> WARNING event **gnr_pdisk_wcache_enabled** for the DISK

  – Write cache disabled -> HEALTHY event **gnr_pdisk_wcache_disabled**

- To disable the write cache, the following command can be used:

```
sdparm –set=WCE=0 [–S] <device>
```

- The current state of the write cache can be inquired by:

```
sdparm –get=WCE=1 <device>
```

Write cache enabled monitoring on ESS [2/2]

- The WCE monitoring is enabled by default.
  You can disable it with the following command:

```
mmchconfig mmhealthphysicaldiskwcemonitorenabled=no
```

- To check if the WCE monitoring is enabled, issue the following command:

```
mmlsconfig mmhealthphysicaldiskwcemonitorenabled
```

5.1.6

- **Health Monitoring:**
- Enhanced key server problems monitoring
- Integration of mmhealth status into OCP

- **Proactive Services:**
- [CONTINUOUS] Implement rules for new Flashes, best practice violations, etc.
- Product-specific proactive alerts

Questions ?

Thank you for using
IBM Spectrum Scale!