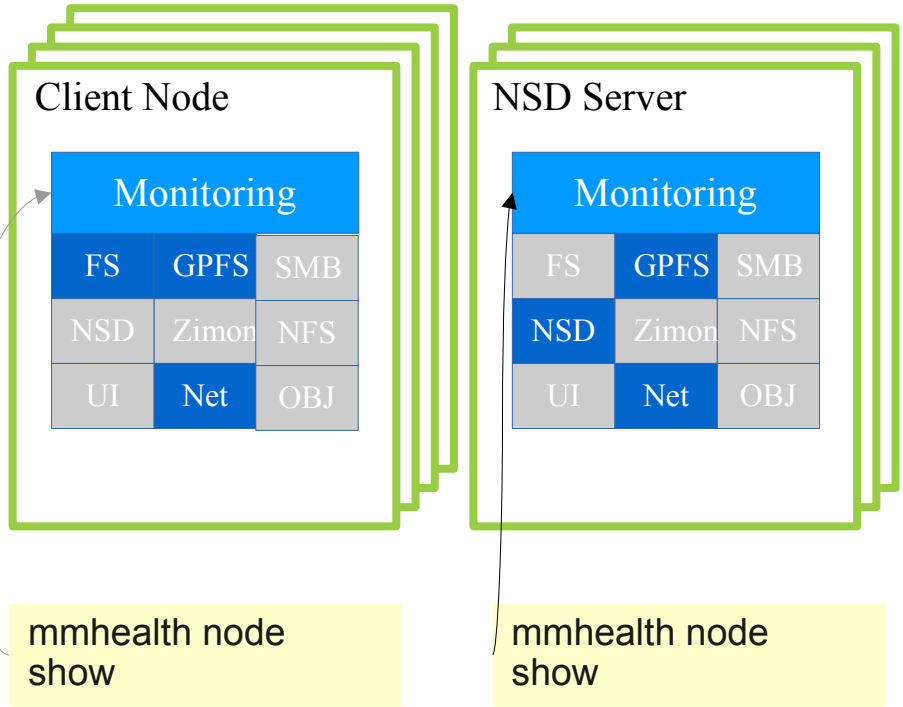
System Health, Performance Monitoring & Call Home

Mathias Dietz
Spectrum Scale RAS Architect

IBM Research and Development
in Kelsterbach/Frankfurt, Germany
mdietz@de.ibm.com

IBM

What is System Health Monitoring?



Health Monitoring Daemon

Backend of the **mmhealth** command

- Introduced with 4.2.1 (gpfs-base)
- Runs on Linux and AIX
- Requires python 2.7.x
- Requires CCR

Decentralized monitoring

- Runs on any cluster node
- Monitors local components only
- Each node knows its own state only
- Avoid communication between nodes to achieve good scalability

Health monitoring

- 27 components covered!
- Better monitoring with each release!

NEW IN 4.2.3

NEW IN 5.0.0

		Zookeeper and Kafka MSGQUEUE	File Audit Logging Consumers FILEAUDITLOG
OBJECT	object authentication AUTH_OBJ	Hadoop connector HADOOPCON	transparent cloud tiering (TCT) CLOUDGW
NFS	file authentication AUTH	FILESYSTEM	DISK
SMB	CES-relevant networks CESNETWORK	performance monitor PERFMON	THRESHOLD
block level storage BLOCK	GPFS-relevant networks NETWORK	REST API monitoring SCALEMGMT	common events GPFS
GNR enclosure ENCLOSURE	GNR physical disk PHYSICALDISK	GUI	Advanced File Management AFM
GNR array ARRAY	GNR virtual disk VIRTUALDISK	GNR recovery group RECOVERYGROUP	CSM-relevant events CLUSTERSTATE

Covered Services

- Over 650 events! Over 170 checks!

GPFS
- quorumloss, ccr_paxos_12_fail
- gpfs_down, longwaiters_found

PERFMON
- pmsensors_up /-down
- pm_collector_up /-down

DISK
- disk_found /-vanished
- disk_up /-down

FILESYSTEM
- fserrinodenummismatch, stale_mount
- fserrallocblock, fserrbaddirblock

CLOUDGW
- cloudgateway_down
- cloudgw_restart

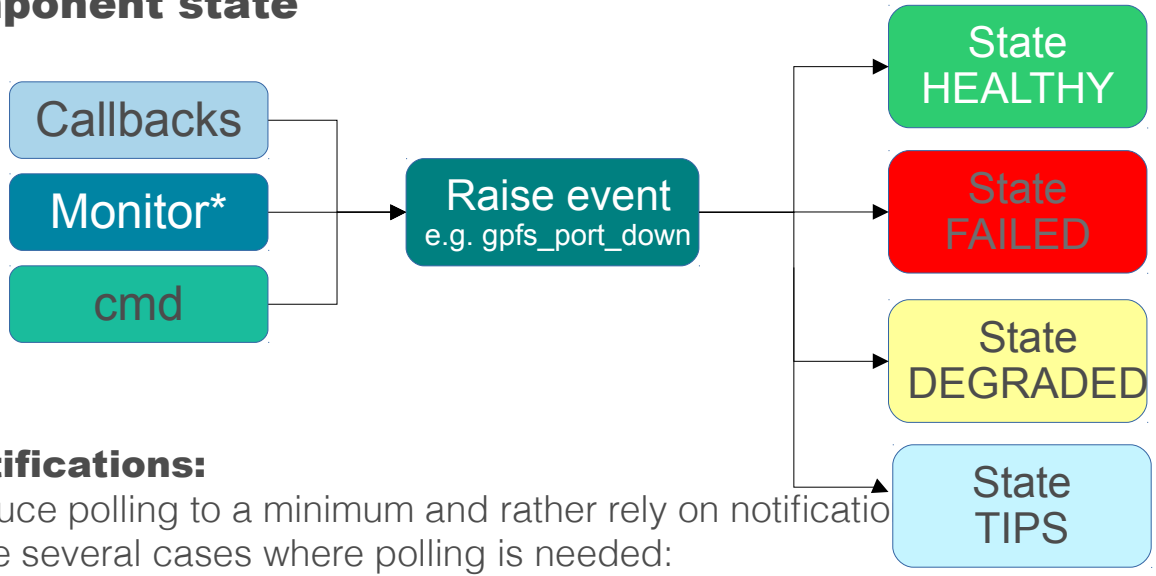
HADOOPCON
- hadoop_namenode_up /-down
- hadoop_datanode_up /-down

NETWORK
- network_down, network_ips_down
- bond_down, network_link_down

GUI
- gui_up /-down
- gui_warn

Health Monitoring

Detect a problem → raise well defined event → update component state



***Polling vs notifications:**

The goal is to reduce polling to a minimum and rather rely on notificatio
However there are several cases where polling is needed:

1) polling needed to build an initial state
2) when the component does not provide notifications (e.g. TCT)
3) when the notifications are incomplete (e.g. disk_failed)
4) To re-sync state when notifications might have been dropped.

System Health Monitoring Interval

System Health framework default intervals

15 sec interval -> CES services
30 sec interval -> GPFS, Network, Filesystem,etc
300 sec interval -> ESS Enclosures
1 hour interval -> Tips
1 day interval -> ESS firmware checks

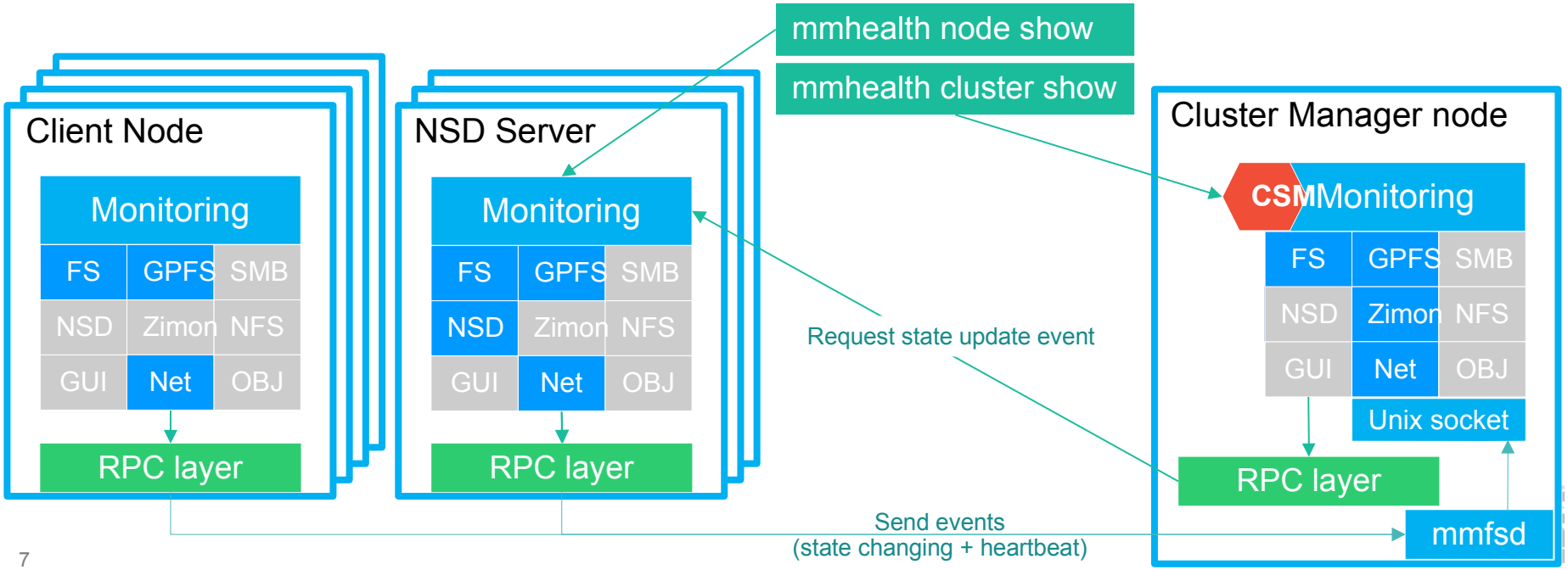
Spectrum Scale >=4.2.3 allows some control on health monitoring frequency

- Monitoring interval High, Medium, Low
- Trade-off: Failure detection time vs resource consumption
- Examples:
  - Low = run monitors rarely → less overhead but longer failure detection time
  - High = run monitors very often → higher overhead, quick failure detection time

Command:
mmhealth config interval off | low | medium | high

Cluster State Overview

- mmhealth cluster show command shows a consolidated cluster wide view
- View node/component state summary for all cluster nodes (eventually consistent)



mmhealth Drilldown

- Is my cluster working fine?

```
[root@ch-41 ~]# mmhealth cluster show

Component            Total         Failed       Degraded        Healthy          Other
-------------------------------------------------------------------------------------
NODE                     5              0              1              4              0
GPFS                     5              0              0              5              0
NETWORK                  5              0              0              5              0
FILESYSTEM               1              0              0              1              0
DISK                     2              0              0              2              0
CES                      2              1              0              1              0
PERFMON                  3              0              0              3              0
```

mmhealth Drilldown

- Where/what is the problem?

```
[root@ch-41 ~]# mmhealth cluster show ces

Component      Node                      Status            Reasons
-------------------------------------------------------------------------------------
CES            ch-41.localnet.com        HEALTHY           -
CES            ch-42.localnet.com        FAILED            ces_network_ips_down,
nfs_in_grace, nfsd_down
```

mmhealth Drilldown

- When did it happen?

```
[root@ch-41 ~]# mmhealth node show -N ch-42

Node name:      ch-42.localnet.com
Node status:    DEGRADED
Status Change:  10 min. ago

Component       Status        Status Change     Reasons
------------------------------------------------------------------------------------
GPFS            HEALTHY       7 days ago        -
NETWORK         HEALTHY       8 days ago        -
FILESYSTEM      HEALTHY       7 days ago        -
DISK            HEALTHY       7 days ago        -
CES             FAILED        9 min. ago        nfsd_down, ces_network_ips_down, nfs_in_grace
PERFMON         HEALTHY       8 days ago        -
```

mmhealth Drilldown

- What is nfsd_down and how do I fix it?

```
[root@ch-41 ~]# mmhealth event show nfsd_down
Event Name:              nfsd_down
Event ID:                999167
Description:             Checks for a NFS service process
Cause:                   The NFS server process was not detected
User Action:             Check the health state of the NFS server and restart, if
necessary. The process might hang or is in a defunct state. Make sure kernel NFS server is
not running
Severity:                ERROR
State:                   FAILED
```

mmhealth eventlog

- mmhealth command allows a lot:
  - View current events and their details
  - Browse event history

```
[root@g5001-21 ~]# mmhealth node eventlog --day
Node name:      g5001-21d.localnet.com
Timestamp                       Event Name              Severity   Details
2018-02-26 16:34:36.405739 CET  callhome_enabled        INFO       Callhome is installed, configured and enabled.
2018-02-26 16:35:36.654185 CET  callhome_not_enabled    TIP        Callhome is not installed, configured or enabled.
2018-02-26 16:36:06.507416 CET  callhome_enabled        INFO       Callhome is installed, configured and enabled.
2018-02-27 08:12:07.508632 CET  gpfs_down               ERROR      The Spectrum Scale service process not running on this node.
Normal operation cannot be done
2018-02-27 08:12:37.642610 CET  gpfs_up                 INFO       The Spectrum Scale service process is running
```

Capabilities

- mmhealth command allows a lot:
  - View current events and their details
  - Browse event history
  - Get tips for a better configuration

NEW IN 4.2.3

```
[root@g5001-21 ~]# mmhealth node show

Node name:      g5001-21d.localnet.com
Node status:    TIPS
Status Change:  7 hours ago

Component      Status        Status Change     Reasons
-------------------------------------------------------------
GPFS           TIPS          7 hours ago       gpfs_pagepool_small
NETWORK        HEALTHY       8 days ago        -
FILESYSTEM     CHECKING      7 hours ago       -
DISK           HEALTHY       7 hours ago       -
PERFMON        HEALTHY       8 days ago        -
THRESHOLD      HEALTHY       8 days ago        -
```

Capabilities

- mmhealth command allows a lot:
  - View current events and their details
  - Browse event history
  - Get tips for a better configuration NEW IN 4.2.3
  - Manage performance thresholds NEW IN 4.2.3

```
[root@g5001-21 ~]# mmhealth thresholds list

### 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
MetaDataPool_capUtil_custom  MetaDataPool_capUtil  90     None    high                 gpfs_fs_name,gpfs_diskpool_name                    300
DataCapUtil_Rule             DataPool_capUtil      90.0   80.0    high                 gpfs_cluster_name,gpfs_fs_name,gpfs_diskpool_name  300
MemFree_Rule                 mem_memfree           50000  100000  low                  node                                               300
MetaDataCapUtil_Rule         MetaDataPool_capUtil  90.0   80.0    high                 gpfs_cluster_name,gpfs_fs_name,gpfs_diskpool_name  300
```

Capabilities

- mmhealth command allows a lot:
  - View current events and their details
  - Browse event history
  - Get tips for a better configuration NEW IN 4.2.3
  - Manage performance thresholds NEW IN 4.2.3
  - Easy scripting with -Y

```
[root@g5001-21 ~]# mmhealth node show gpfs -Y
mmhealth:Event:HEADER:version:reserved:reserved:node:component:entityname:entitytype:event:arguments:activesince:identifier:ishidden:
mmhealth:State:HEADER:version:reserved:reserved:node:component:entityname:entitytype:status:laststatuschange:
mmhealth:State:0:1:::g5001-21d.localnet.com:NODE:g5001-21d.localnet.com:NODE:TIPS:2018-02-27 08%3A12%3A37.713836 CET:
mmhealth:State:0:1:::g5001-21d.localnet.com:GPFS:g5001-21d.localnet.com:NODE:TIPS:2018-02-27 08%3A12%3A37.708570 CET:
mmhealth:Event:0:1:::g5001-21d.localnet.com:GPFS:g5001-21d.localnet.com:NODE:gpfs_pagepool_small::2018-02-19 15%3A25%3A43.147270
CET::no:
```

New Features

- Further new features in 5.0.0:
  - Warn customers about filled /var/mmfs
  - Detect and display when the GPFS daemon is unresponsive
  - Kafka & audit consumer monitor
  - Multiple TCT entities support
  - OS jitter reduction (intervals + clock align)
  - SELinux: mmhealth log rotation support
  - AIX: filesystem monitoring
  - Increased monitoring resilience

New Features

- Further new features in 5.0.0:
  - Event script callout
    - User defined script may be defined as `/var/mmfs/etc/eventsCallback` (must be owned by root + executable)
    - Executed for each event
    - Gets CLI parameters:
      `1 2018-02-28 14:03:46.592724 CET quorum_down gpfs  E D The node is not able to form a quorum with the other available nodes.`
      (<version> <timestamp> <eventName> <component> <identifier> <severity> <state> <msg> <args>)
    - Callout script must be lightweight and return within seconds to not block monitoring thread (Timeout of 30 seconds before it gets killed)

Performance Monitoring

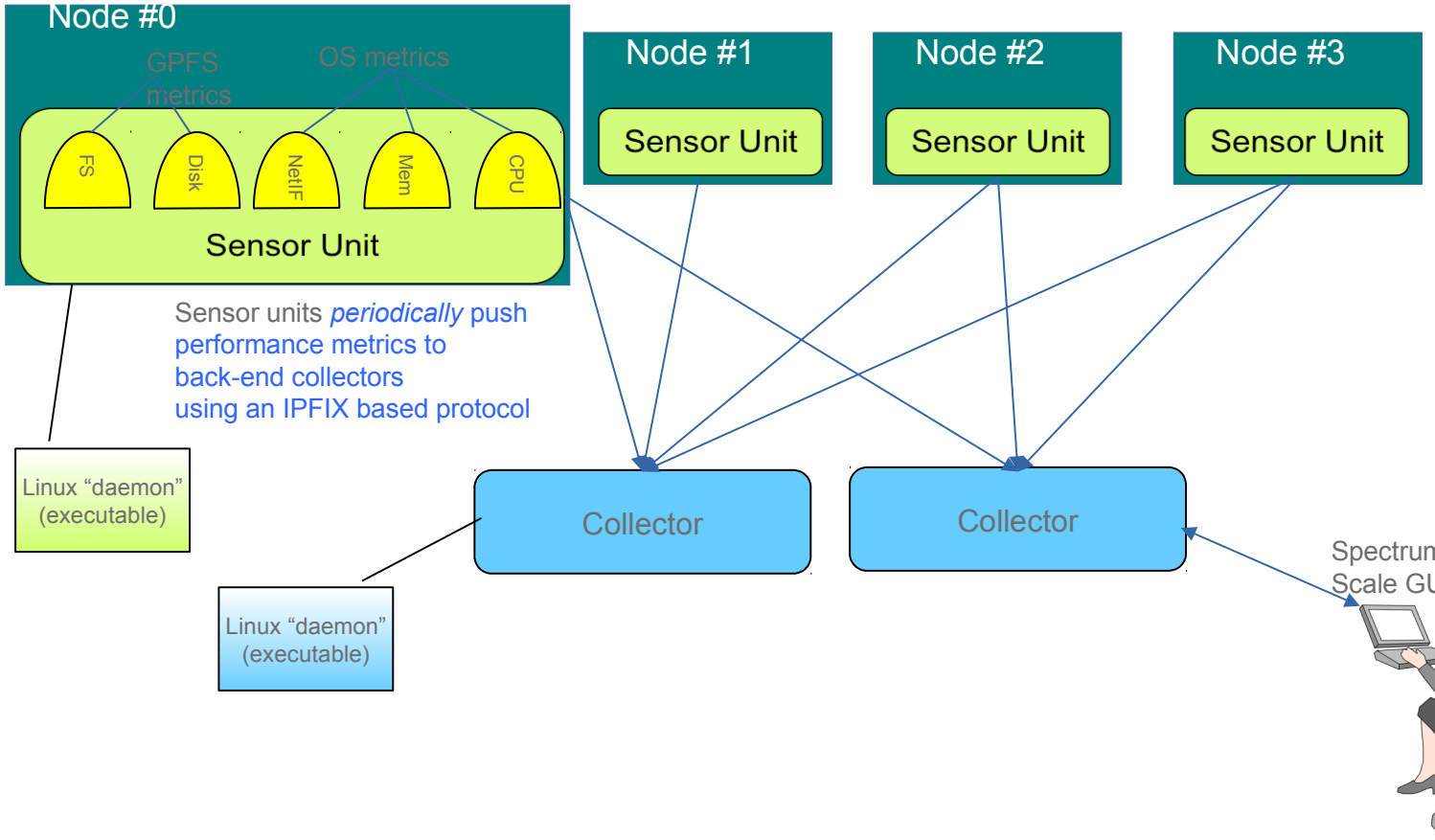
Performance Monitoring for Spectrum Scale / ESS

Spectrum Scale comes with a scalable performance monitoring solution (Zimon)

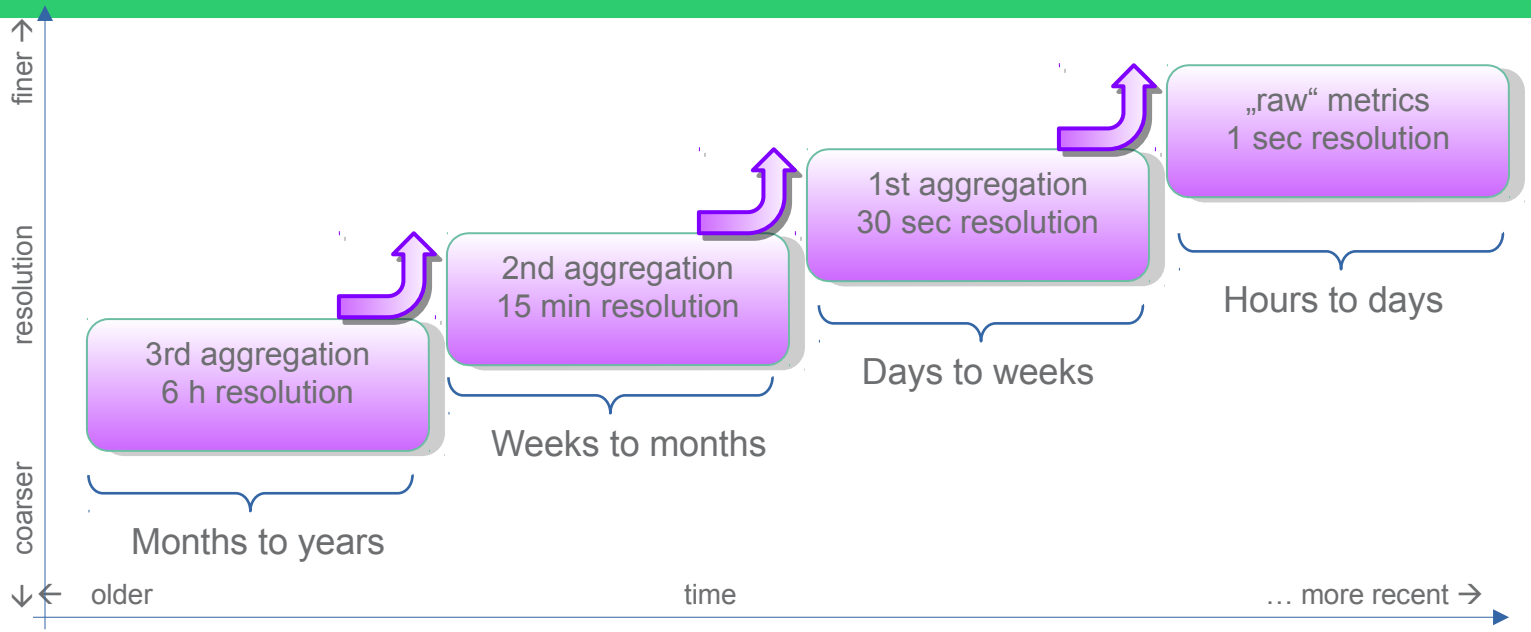
Performance Monitoring Features

- Collect performance data from all nodes in the cluster
  - scales up to a large number of nodes
- Collects various performance metrics from different components
  - >50 Performance sensors (GPFS IO, AFM, SMB, Object,NFS ...) and >1000 Metrics
  - Collect capacity/usage information
- Scalable architecture and extensible design
  - Handles a very large number of performance metrics
  - Easy to add new performance sensors/metrics
- Stores performance data in a time series database
- Time-based aggregation of performance data
- Optional component of Spectrum Scale → but strongly recommended
- Provides multiple interfaces to get access to the data

ZIMon Performance Monitoring Overview



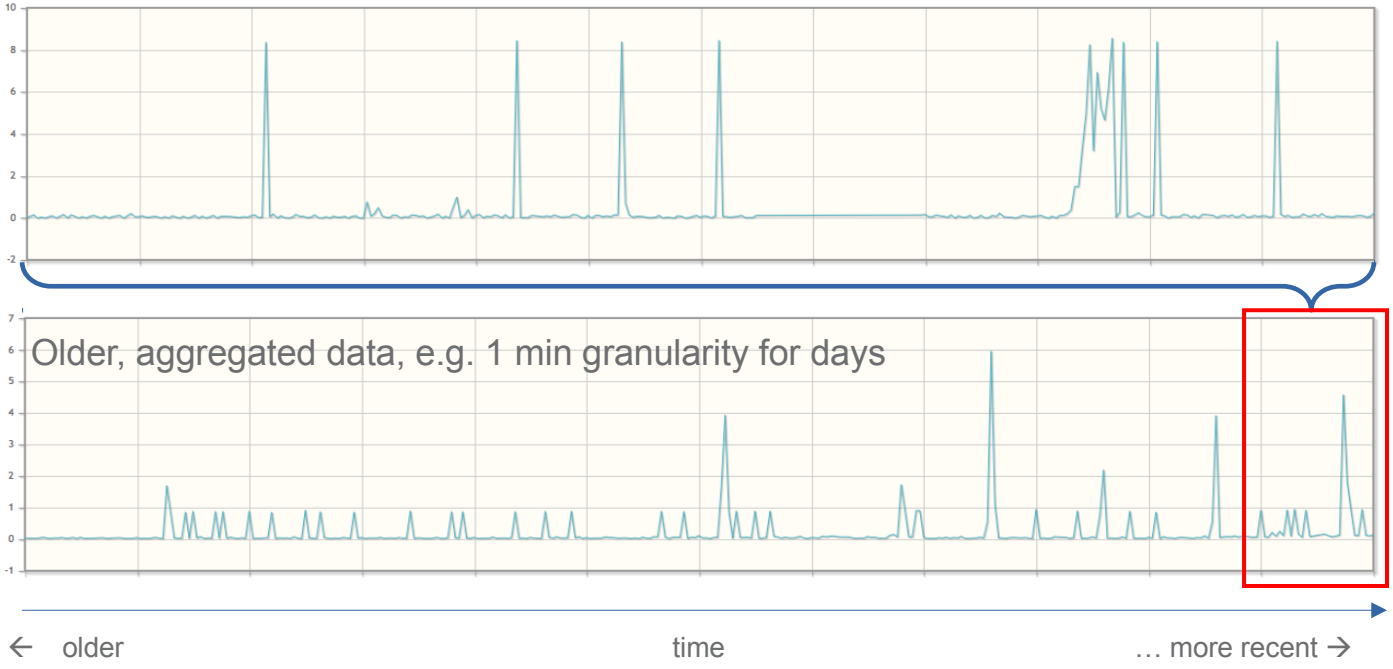
Time-Based Aggregation (Example)



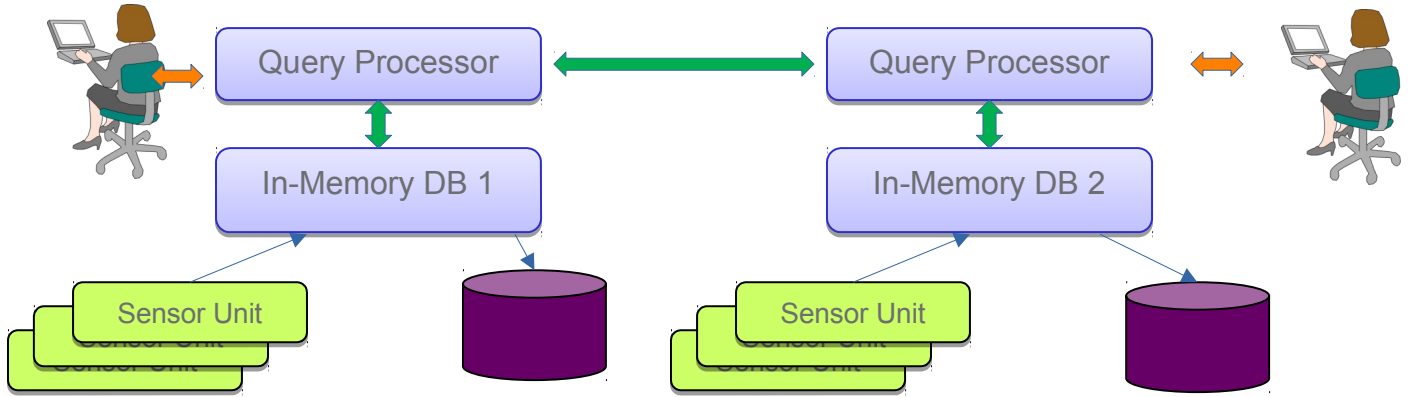
- Each storage domain is limited in the amount of memory
- Older metrics are aggregated and pushed to next aggregation level
- Eventually, metrics data is “forgotten”

Data is retrieved at the best possible time resolution

High resolution for a short time, e.g. 1 sec granularity for hours



Federation of ZIMon Collectors



- Data storage and query processing are distributed among multiple collectors.
- Multiple ZIMon collectors are used to store distinct metrics.
- Queries are divided into sub-queries and distributed to the collector owning the data.

Install TIPs

- If in doubt use the installer
- Having more than one collector distributes the load and allows for full HA
- Use GPFS hostnames (daemon address)

For full HA set colredundancy=2

```
mmperfmon config show --config-file /tmp/pc
vim /tmp/pc
mmperfmon config update --config-file /tmp/pc
```

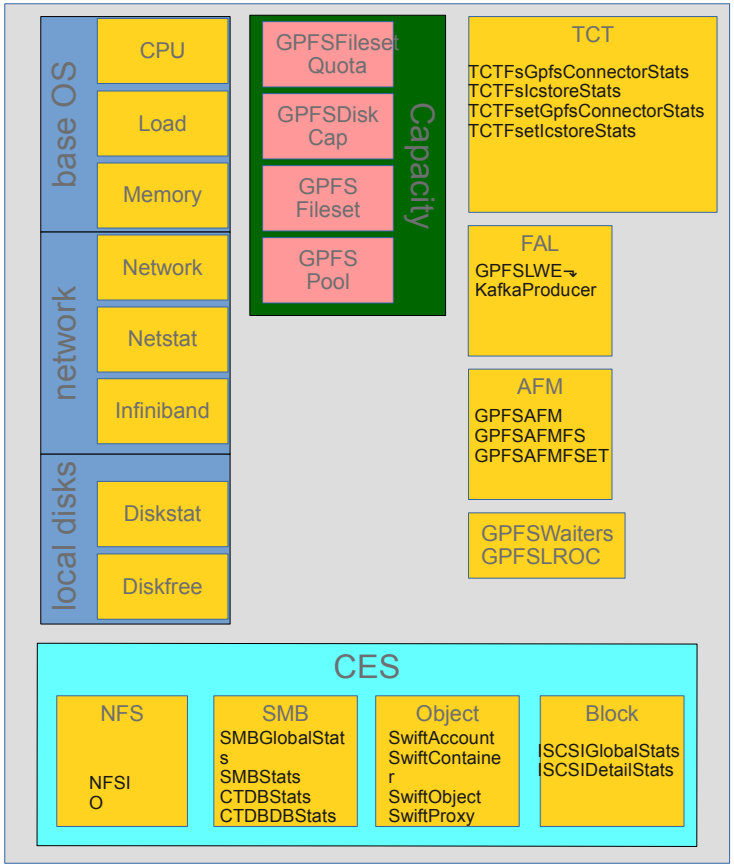
For federation set peer configuration on all collector nodes

```
vim /opt/IBM/zimon/ZIMonCollector.cfg
```

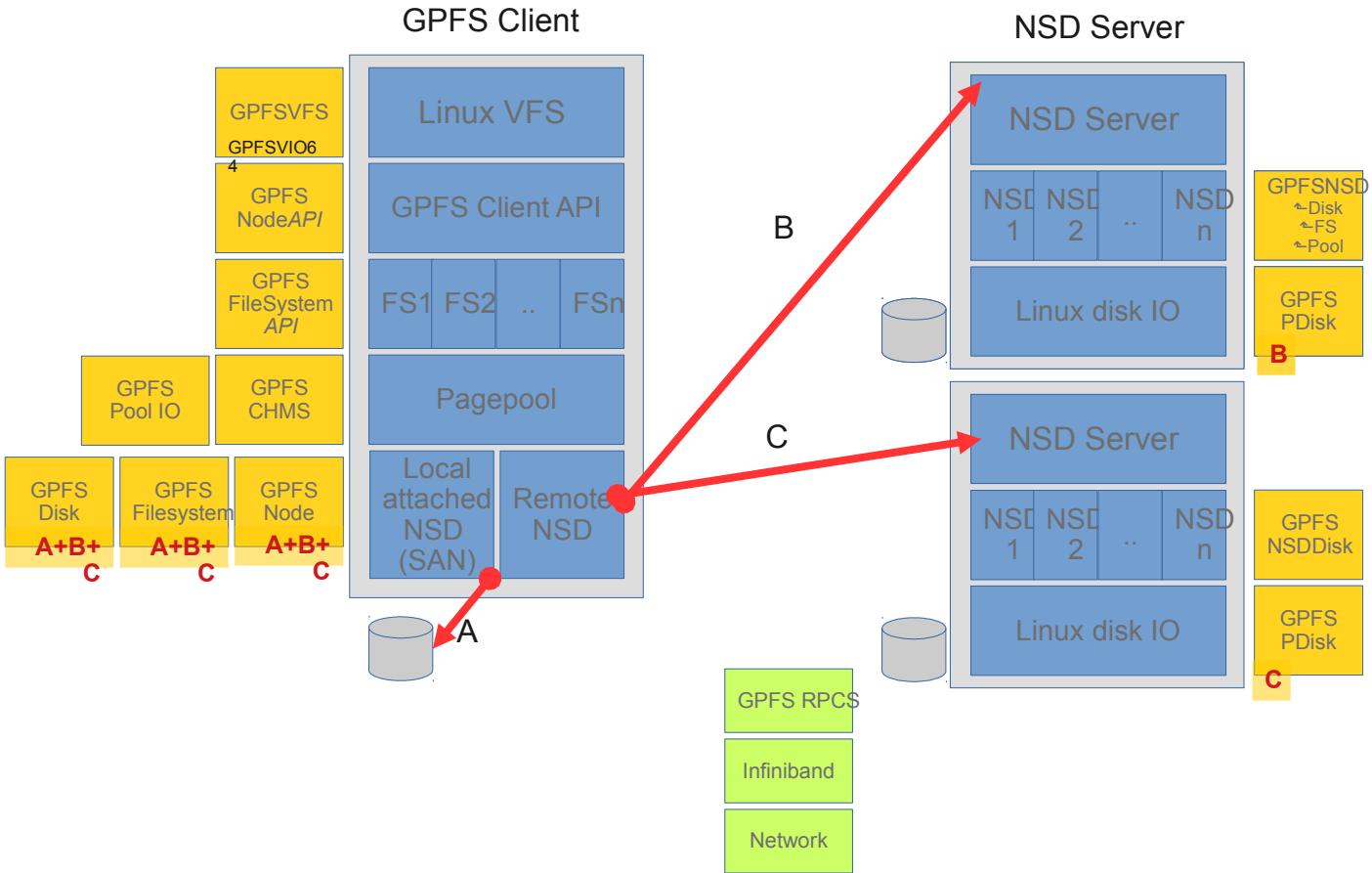
Sensor overview & recommendations

Sensor TIPs

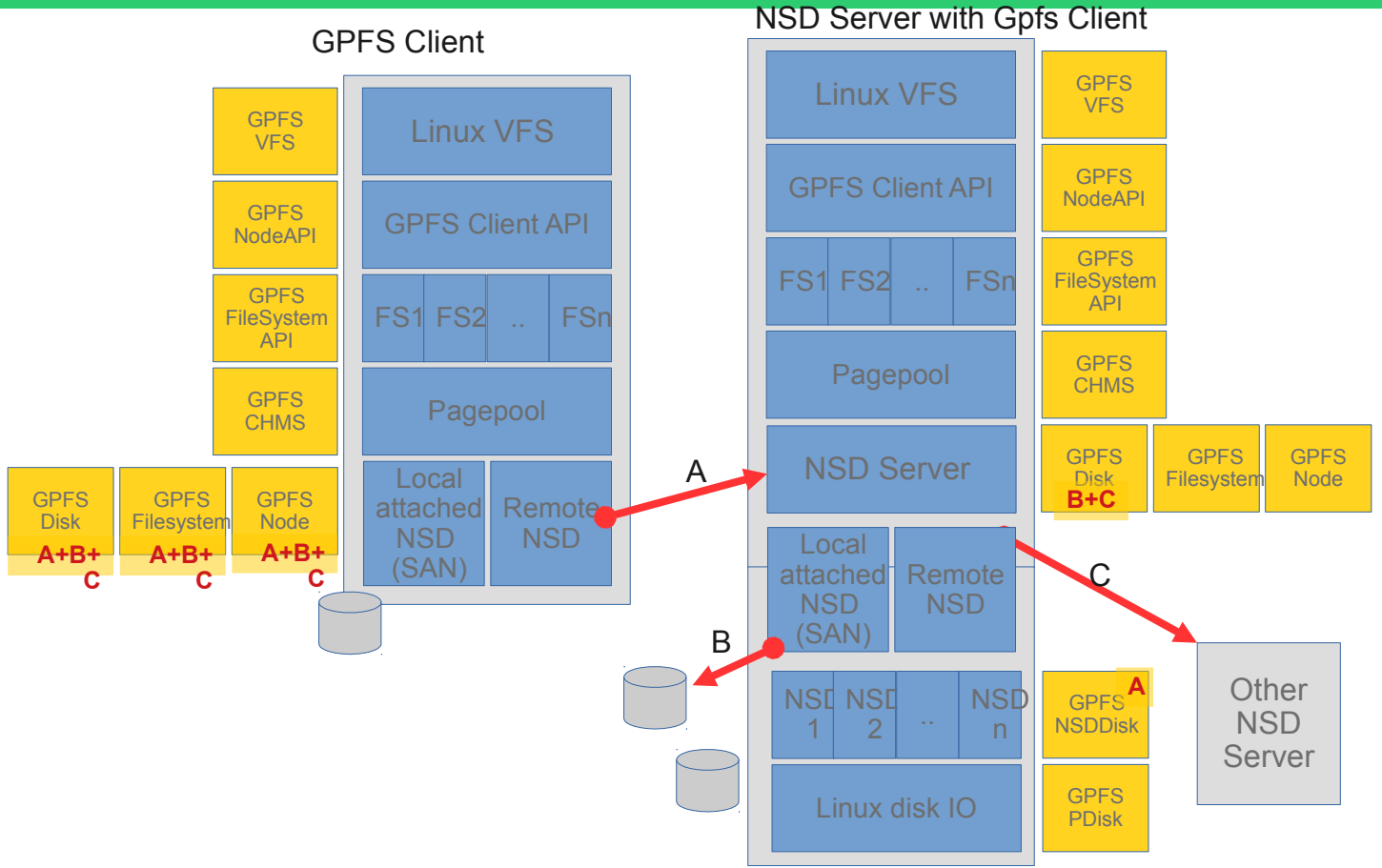
- Capacity sensor (must) be restricted to one node. Data is cluster-wide no need to get it more than once
- Restrict clause has to be the fully qualified node name
- DiskCap (=mmdf) run just once a day (period=86400)
- GpfsFilesetQuota (=mmrepquota) is also heavy, just sample every few hours
- GpfsPool and Fileset are sampled every 5 min to prevent “out of space” conditions
- NFSIO reports data per NFS share, if you have many this will create a lot of data
- GPFSDisk reports IO between nodes and NSDs, if you have numerous of both this will use lots of memory in the collector



Sensor data explained



Sensor data explained



Performance Monitoring for Spectrum Scale / ESS

Interfaces to the performance monitoring data

- Well integrated with Spectrum Scale/ ESS GUI
- Java, C and Python API (internal use only)
- CLI to query performance data
  - mmperfmon query
- Grafana Bridge (OpenSource Monitoring Dashboard)
- REST API to query performance data (planned for Spectrum Scale 5.0)

Performance Monitoring for Spectrum Scale / ESS

Use *mmperfmon query* command to query performance data

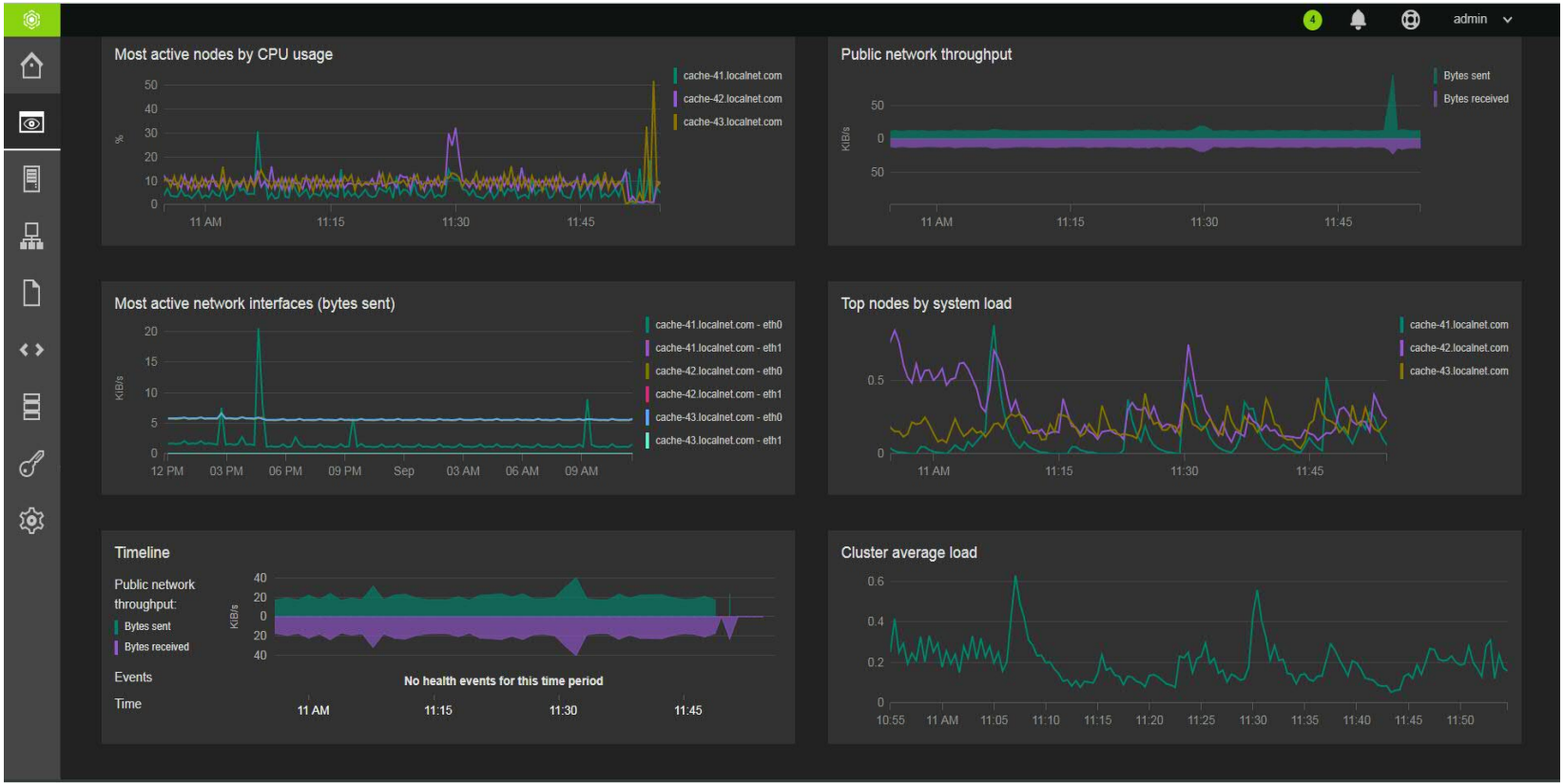
mmperfmon query Metric[,Metric...] | Key[,Key...] | NamedQuery [StartTime EndTime | Duration] [Options]

- Get CPU usage for local node
  - mmperfmon query cpu_user
- Get free memory for another node
  - mmperfmon query mem_memfree –N node2
- Get free memory from all nodes
  - mmperfmon query compareNodes mem_memfree
- Get a list of metrics that can be queries
  - mmperfmon query –list metrics
- Get all metrics for a particular sensor (undocumented feature)
  - mmperfmon query ".* group Netstat"

  - -b bucket-size option , to specify the granularity
  - -n number , to specify how many values are returned
  - --csv for comma separated output

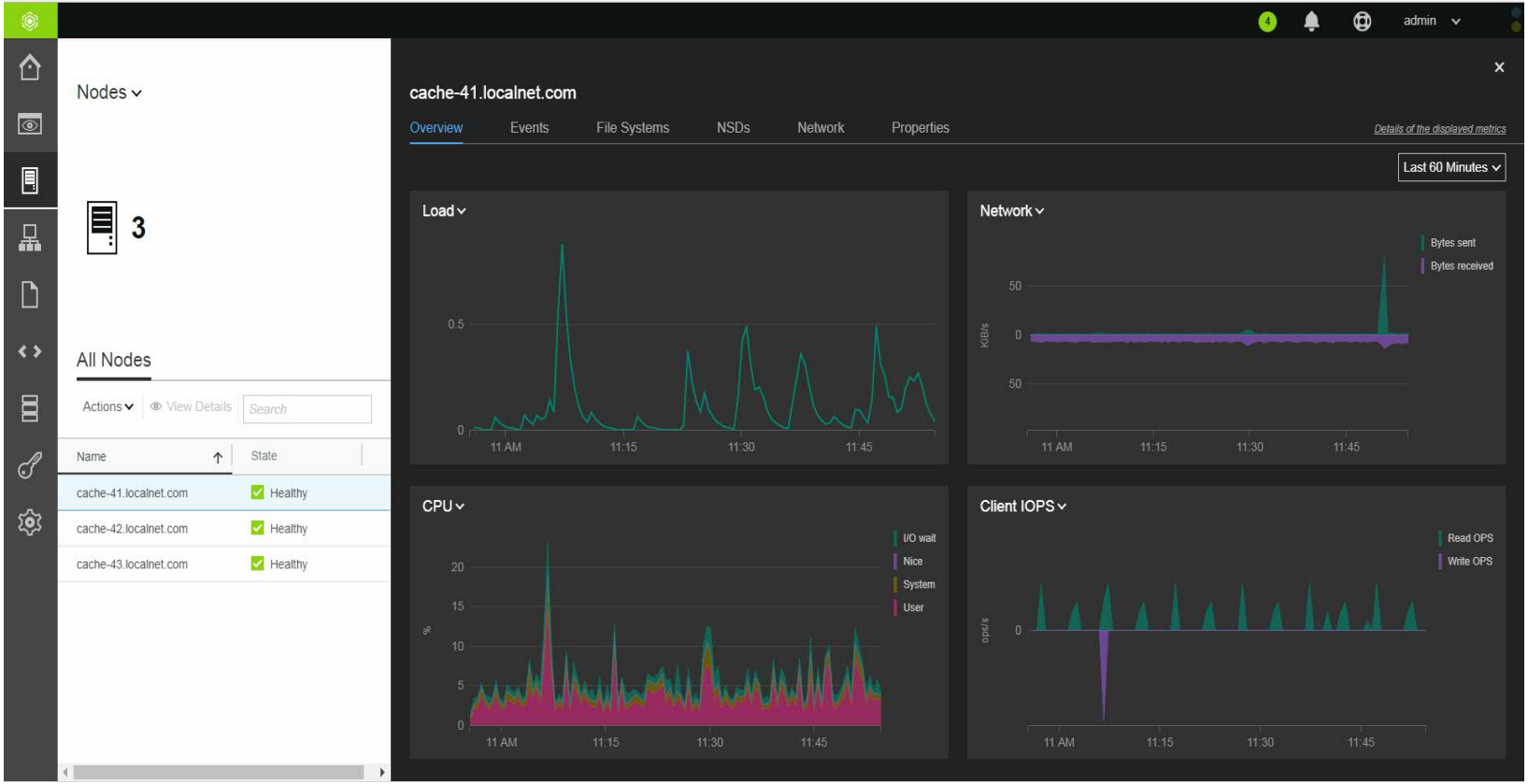
GUI Performance Charts

Spectrum Scale GUI - Fully customizable dashboard



GUI Performance Charts

Spectrum Scale GUI – Hard wired performance charts in detail views:



Performance Monitoring Bridge for Grafana

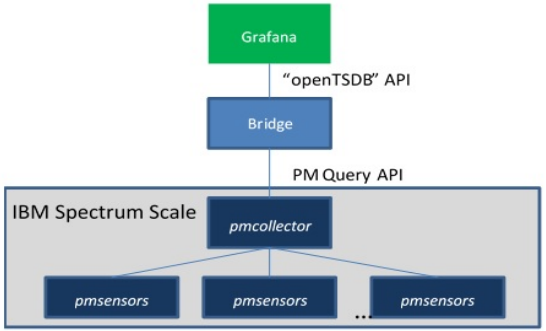


Figure 1. IBM Spectrum Scale integration framework for Grafana

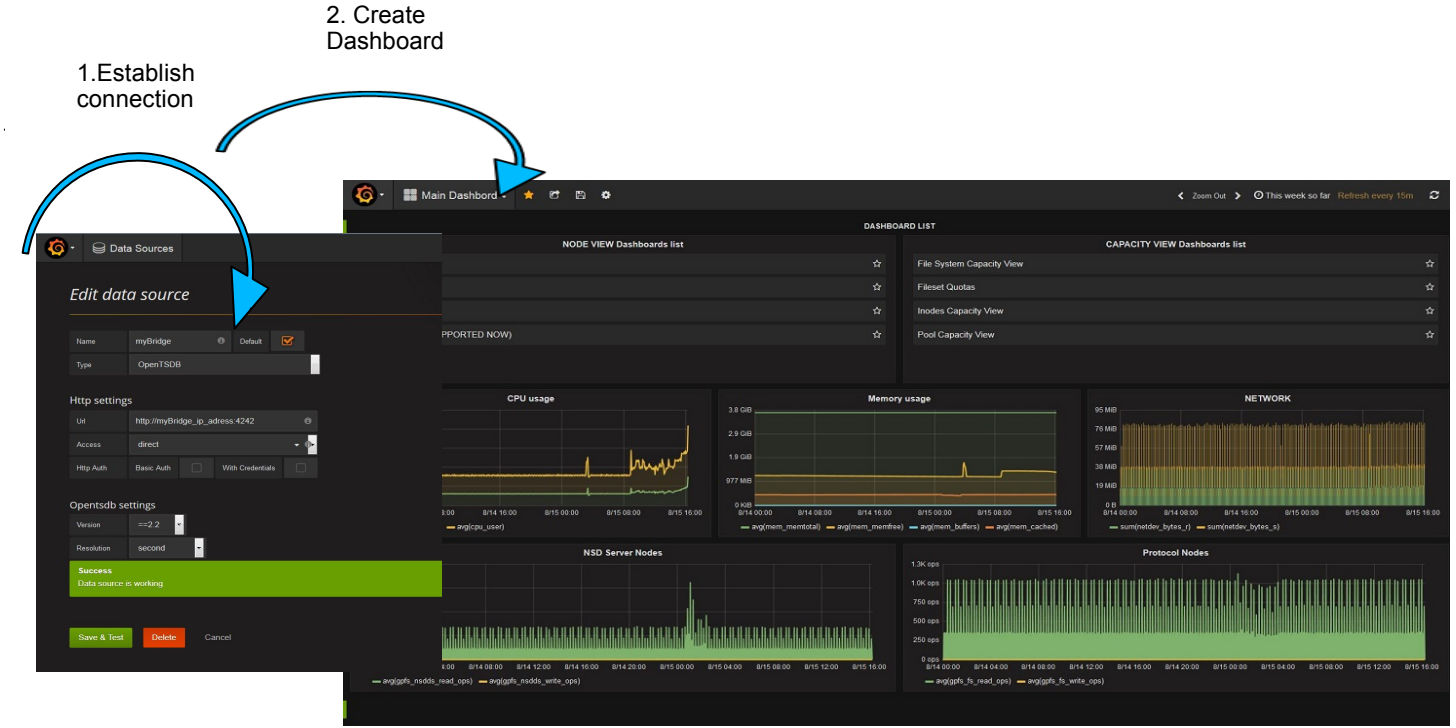
Grafana

- an open source performance data graphical visualizer
- provides a powerful and elegant way to create, explore, and share dashboards and data with your team and the world.

IBM Spectrum Scale Performance Monitoring Bridge

- a Python application
- provides IBM Spectrum Scale performance data to Grafana in „openTSDB“ data exchange format
- communicates with active pmcollector via port 4242

Performance Monitoring Bridge for Grafana



Predefined filesystems capacity thresholds

Spectrum Scale 4.2.2 introduced **predefined filesystem capacity/inode thresholds.**
The capacity metrics will be frequently compared with the rules boundaries by internal monitor process. As soon as one of the metric values exceeds their threshold limit the system health daemon will receive an event notification from monitor process and generate log event and update filesystem status

Predefined filesystem capacity thresholds for

- Fileset-inode spaces
- Data pool capacity
- Metatadata pool capacity
- Memory free <50MB leads to error event, <100MB to warning event

```
[root@gpfsgui-11 ~]# mmhealth thresholds list
### 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          mem_memfree           50000  100000  low                  node                                               300
MetaDataCapUtil_Rule  MetaDataPool_capUtil  90.0   80.0    high                 gpfs_cluster_name,gpfs_fs_name,gpfs_diskpool_name  300
[root@gpfsgui-11 ~]#
```

The violation of any single rule will trigger a health event !

Spectrum Scale 4.2.3 **the ability to set** customer defined thresholds on any given performance metric

Thresholds

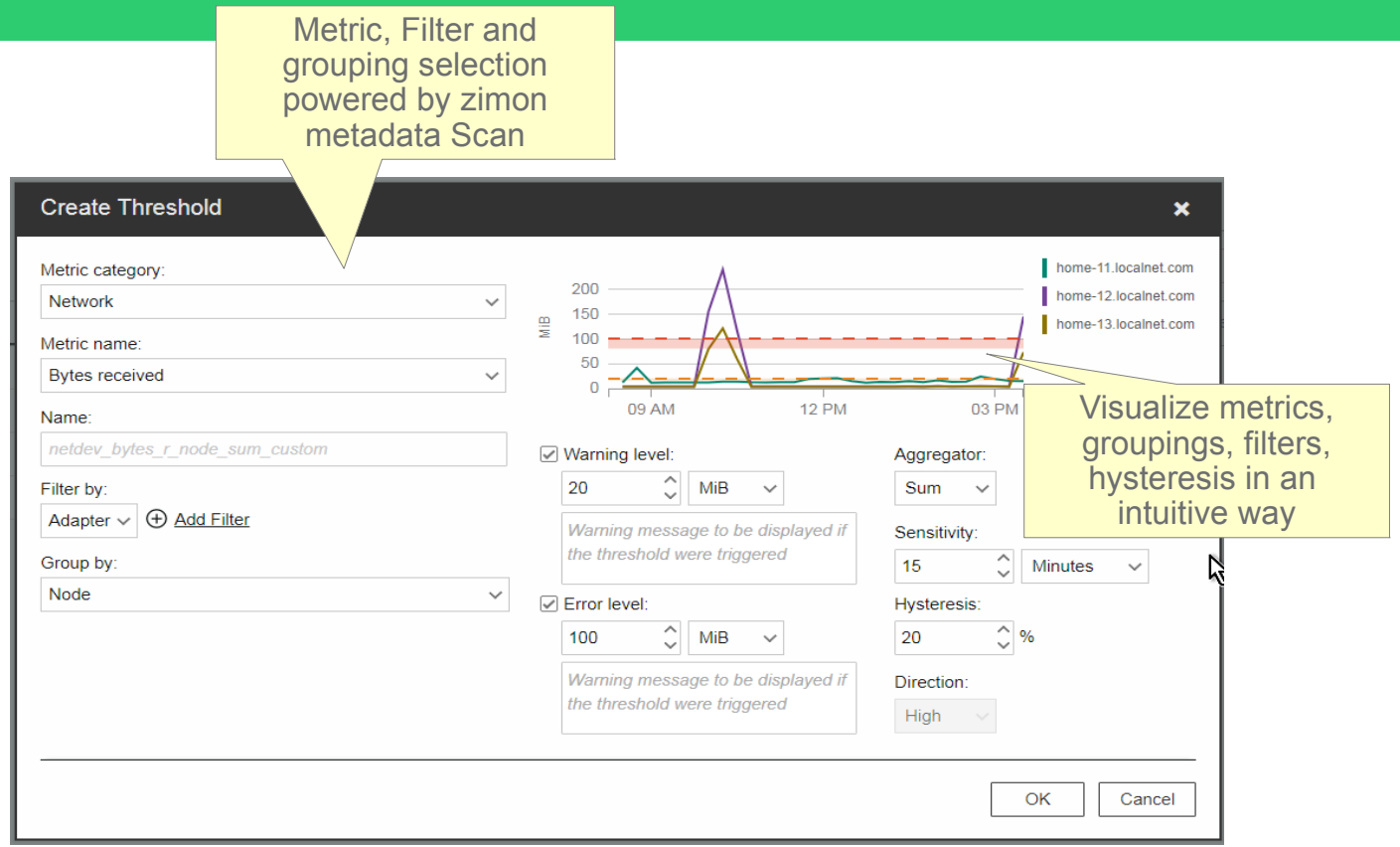
```
mmhealth thresholds add { metric[:sum|avg|min|max|rate]|measurement

    [--errorlevel{threshold error limit}
    [--warnlevel {threshold warn limit}] | --direction {high|low}}

    [--sensitivity {bucketsize}] [--hysteresis {percentage}]

    [--filterBy] [--groupBy ] [--name {ruleName}]

    [--errormsg {user defined action description}]

    [--warnmsg  {user defined action description}]
```

Creates a new thresholds rule for the specified metric or measurement and activates monitoring process for this rule.

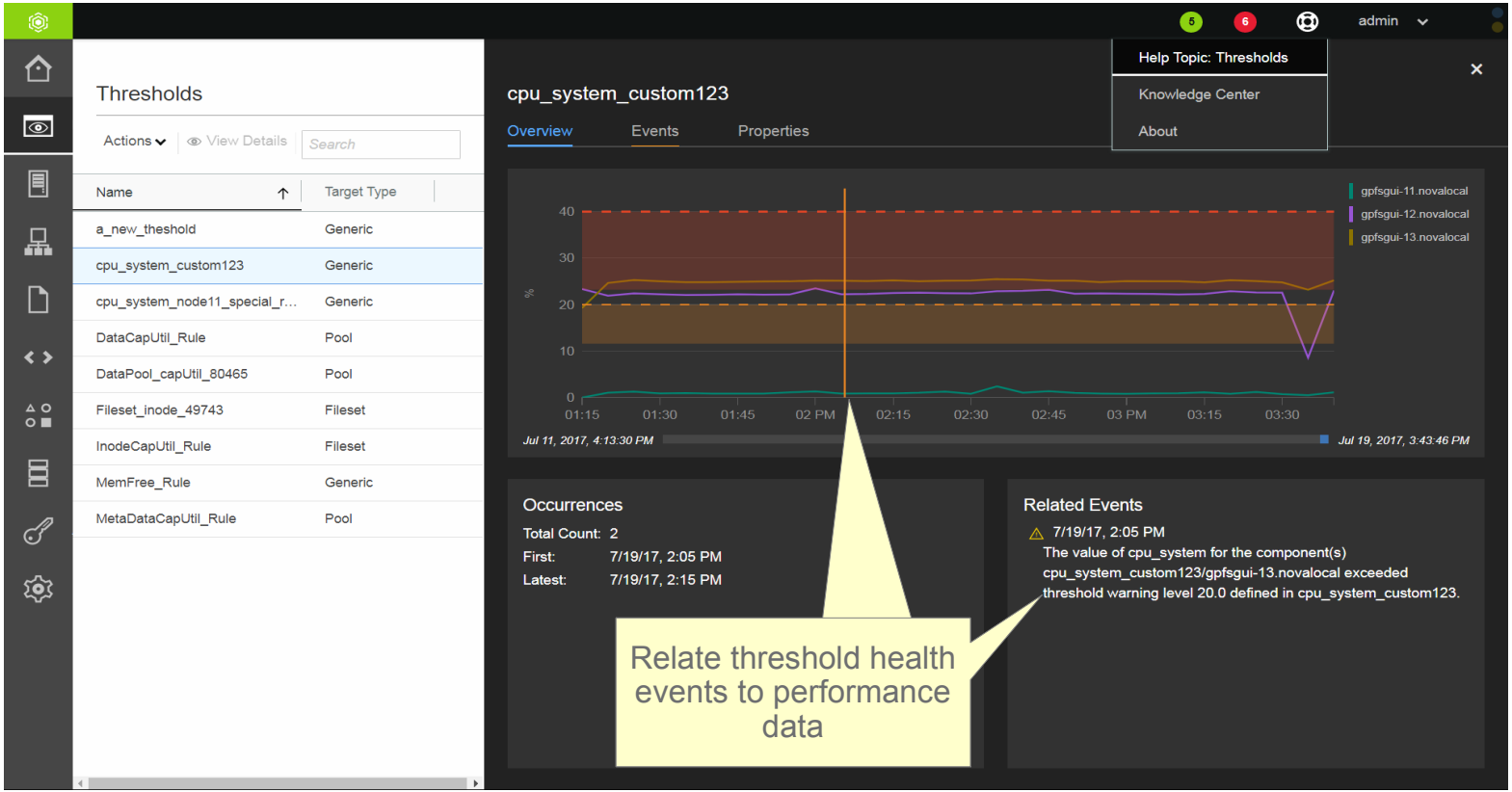
```
mmhealth thresholds add cpu_idle:avg --errorlevel 60 --direction high --name cpu_avg_bynode --groupby node'

mmhealth thresholds add MetaDataPool_capUtil --errorlevel 90 --direction high --groupby gpfs_fs_name --name
myRule'
```

Results of thresholds which are not specific to a node are only shown on the 'active' collector node! (if in doubt check all collector nodes)

GUI Threshold Management



GUI Threshold Monitoring



Call Home / Proactive Service

Overview Spectrum Scale Software Call Home

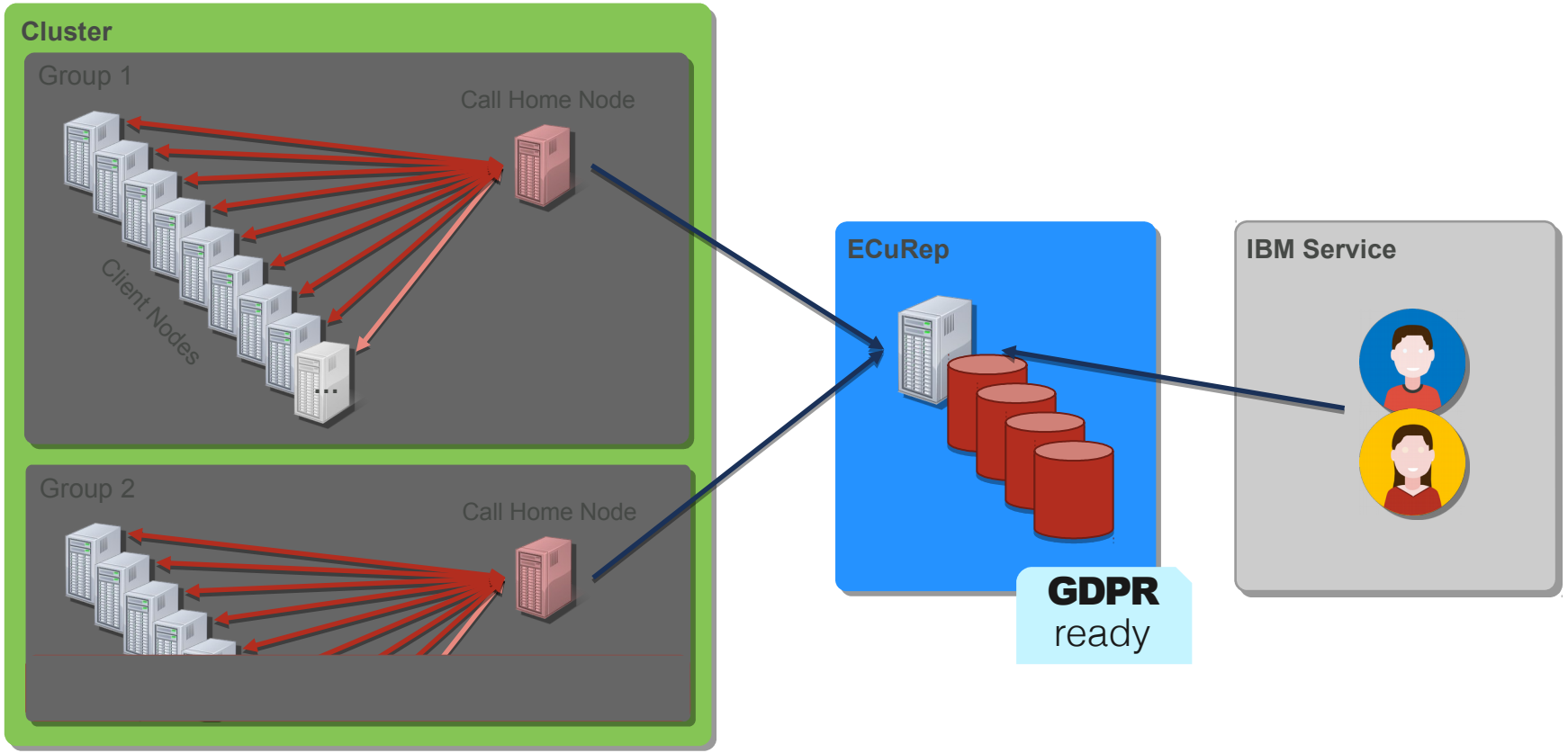
Call Home for Spectrum Scale has been introduced with 4.2.0
Meanwhile it is supported on RHEL7 (x86,ppc,s390), SLES and Ubuntu

It has two main functions:
1) Upload of daily/weekly data collection
2) Upload and data collections triggered by events

Call Home is disabled by default and can be configured using the *mmcallhome* command. Since 5.0 the install toolkit is asking for Call Home setup (opt-out)

Please note: ESS Systems have a dedicated Hardware Call Home feature which is independent from Spectrum Scale Software Call Home.
Recommendation: Enable both !

Software Call Home Architecture



Call Home collects data on a daily and weekly basis

- Spectrum scale configuration
  - mmls* commands, Linux commands like uname, sysctl, CCR content
- Log files (last 10000 lines) and status commands
  - mmfslog, /var/log/messages, mmdiag, mmhealth
- Statistical data
  - object statistics

- Commands/files to collect are defined here:
  - /usr/lpp/mmfs/data/callhome/gather.d/daily.conf|weekly.conf
  - /usr/lpp/mmfs/data/callhome/gather.d/DefaultDaily.ess.conf (on ESS)
- Weekly is meant to collect more heavy weight data

No customer data (e.g. file system content) is collected, but the log files might contain sensitive information

Examples of on the cluster collected datasets

- mmauth_show_Y.txt
- mmces_address_list_Y.txt
- mmdiag_afm_Y.txt
- mmdiag_commands_Y.txt
- mmdiag_deadlock_Y.txt
- mmdiag_health_Y.txt
- mmdiag_lroc_Y.txt
- mmdiag_memory_Y.txt
- mmdiag_network_Y.txt
- mmdiag_nsd_Y.txt
- mmdiag_rpc_24h_Y.txt
- mmdiag_stats_Y.txt
- mmdiag_tokenmgr_Y.txt
- mmdiag_waiters_Y.txt
- mmhealth_cluster_show_verbose_Y.txt
- mmhealth_config_interval_Y.txt
- mmhealth_event_list_hidden_Y.txt
- mmhealth_node_eventlog_day_Y.txt
- mmhealth_node_show_v_Y.txt
- mmhealth_thresholds_list_Y.txt
- mmlscallback_Y.txt
- mmlscluster_ces_Y.txt
- mmlscluster_Y.txt
- mmlsmgr_Y.txt
- mmlsnode_a.txt
- mmlsnsd_L_Y.txt

mmdiag config -Y

mmcallhome:collection:HEADER:version:reserved:reserved:index:schedule:logName:machineType:node:command:timeout:helperOpt
ions:
mmcallhome:collectionRC:HEADER:version:reserved:reserved:rc:starttime:exectime:
mmcallhome:collection:0:1:::COGP001:weekly:mmdiag_config_Y.txt:all:all:/usr/lpp/mmfs/bin/mmdiag --config -Y:10::
mmcallhome:collectionRC:0:1:::0:1504681727:0.12:
mmdiag:config:HEADER:version:reserved:reserved:name:value:changed
mmdiag:config:0:1:::aclHashSpaceSize:2000::
mmdiag:config:0:1:::afmHashVersion:2::
mmdiag:config:0:1:::afmMaxWorkerThreads:1024::
mmdiag:config:0:1:::aioWorkerThreads:256::
mmdiag:config:0:1:::allowDeleteAclOnChmod:1::
mmdiag:config:0:1:::assertOnStructureError:0::
mmdiag:config:0:1:::atimeDeferredSeconds:86400::
mmdiag:config:0:1:::ccrEnabled:1:static:
mmdiag:config:0:1:::cipherList:AUTHONLY:static:
mmdiag:config:0:1:::clusterId:1445794570045570885 9:static:

mmlsnsd -L -Y

mmcallhome:collection:HEADER:version:reserved:reserved:index:schedule:logName:machineType:node:command:timeout:helperOpt
ions:
mmcallhome:collectionRC:HEADER:version:reserved:reserved:rc:starttime:exectime:
mmcallhome:collection:0:1:::COGP011:daily:mmlsnsd_L_Y.txt:all:chnode:/usr/lpp/mmfs/bin/mmlsnsd -L -Y:30::
mmcallhome:collectionRC:0:1:::0:1505738553:0.56:
mmlsnsd:nsd:HEADER:version:reserved:reserved:fileSystem:diskName:volumeId:serverList:
mmlsnsd:nsd:0:1:::gpfs0:disk1:0A006451598C7976:gpfs-81.localnet.com,gpfs-82.localnet.com,gpfs-83.localnet.com:
mmlsnsd:nsd:0:1:::gpfs0:disk6:0A006451598C7992:gpfs-81.localnet.com,gpfs-82.localnet.com,gpfs-83.localnet.com:
mmlsnsd:nsd:0:1:::gpfs1:disk2:0A006451598C797C:gpfs-81.localnet.com,gpfs-82.localnet.com,gpfs-83.localnet.com:
mmlsnsd:nsd:0:1::::disk3:0A006451598C7981:gpfs-81.localnet.com,gpfs-82.localnet.com,gpfs-83.localnet.com:
mmlsnsd:nsd:0:1::::disk4:0A006451598C7987:gpfs-81.localnet.com,gpfs-82.localnet.com,gpfs-83.localnet.com:
mmlsnsd:nsd:0:1::::disk5:0A006451598C798C:gpfs-81.localnet.com,gpfs-82.localnet.com,gpfs-83.localnet.com:

Examples of on the cluster collected datasets

mmlscluster -Y
mmcallhome:collection:HEADER:version:reserved:reserved:index:schedule:logName:machineType:node:command:timeout:helperOptions:
mmcallhome:collectionRC:HEADER:version:reserved:reserved:rc:starttime:exectime:
mmcallhome:collection:0:1:::COGP006:daily:mmlscluster_Y.txt:all:chnode:/usr/lpp/mmfs/bin/mmlscluster -Y:30::
mmcallhome:collectionRC:0:1:::0:1505738549:0.79:
mmlscluster:clusterSummary:HEADER:version:reserved:reserved:clusterName:clusterId:uidDomain:rshPath:rshSudoWrapper:rcpPath:rcpSudoWrapper:repositoryType:primaryServer:secondaryServer:
mmlscluster:clusterNode:HEADER:version:reserved:reserved:nodeNumber:daemonNodeName:ipAddress:adminNodeName:designation:otherNodeRoles:adminLoginName:otherNodeRolesAlias:
mmlscluster:cnfsSummary:HEADER:version:reserved:reserved:cnfsSharedRoot:cnfsMoundPort:cnfsNFSDprocs:cnfsReboot:cnfsMonitorEnabled:cnfsGanesha:
mmlscluster:cnfsNode:HEADER:version:reserved:reserved:nodeNumber:daemonNodeName:ipAddress:cnfsState:cnfsGroupId:cnfsIplist:
mmlscluster:cesSummary:HEADER:version:reserved:reserved:cesSharedRoot:EnabledServices:logLevel:addressPolicy:
mmlscluster:cesNode:HEADER:version:reserved:reserved:nodeNumber:daemonNodeName:ipAddress:cesGroup:cesState:cesIpList:
mmlscluster:cloudGatewayNode:HEADER:version:reserved:reserved:nodeNumber:daemonNodeName:
mmlscluster:clusterSummary:0:1:::gpfs-cluster-8.localnet.com:14457945700455708859:localnet.com:/usr/bin/ssh:no:/usr/bin/scp:no:CCR:gpfs-81.localnet.com::
mmlscluster:clusterNode:0:1:::1:gpfs-81.localnet.com:10.0.100.81:gpfs-81.localnet.com:quorum:X,Z::perfmon,ces:
mmlscluster:clusterNode:0:1:::2:gpfs-82.localnet.com:10.0.100.82:gpfs-82.localnet.com:quorum:X,Z::perfmon,ces:
mmlscluster:clusterNode:0:1:::3:gpfs-83.localnet.com:10.0.100.83:gpfs-83.localnet.com::X,Z::perfmon,ces:
mmlscluster:cesSummary:0:1:::/mnt/gpfs0/ces:OBJ:0:even-coverage:mmlscluster:cesNode:0:1:::1:gpfs-81.localnet.com:10.0.100.81::e:192.168.122.81:
mmlscluster:cesNode:0:1:::2:gpfs-82.localnet.com:10.0.100.82::e:192.168.122.83:
mmlscluster:cesNode:0:1:::3:gpfs-83.localnet.com:10.0.100.83::e:192.168.122.82:

Examples of on the cluster collected datasets

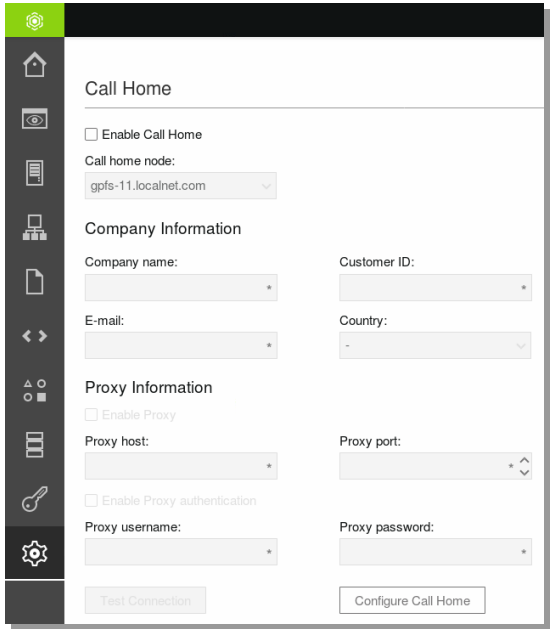
Performance data

```
Row,Timestamp,cpu_system,cpu_user,cpu_contexts
1,1512136018,0.500000,0.500000,1066
2,1512136019,2.530000,1.010000,1504
3,1512136020,0.000000,0.500000,941
4,1512136021,0.000000,0.500000,972
5,1512136022,0.500000,0.000000,992
6,1512136023,0.500000,0.500000,1073
7,1512136024,2.540000,9.140000,1439
8,1512136025,1.010000,6.570000,1114
9,1512136026,0.000000,0.500000,938
10,1512136027,1.510000,0.500000,1156
```

Call Home

What‘s new in 4.2.3.7 and 5.0.0?

- Many improvements and fixes (4.2.3.7 & 5.0.0)
- Install, enable and configure Call Home with the installation toolkit (5.0.0)
- Enable and configure Call Home with the GUI (5.0.0)
- Added -Y option for machine readable output (5.0.0)
- Reworked the collection list, what gets collected and how (5.0.0)
- Added --pmr option to the `mmcallhome run SendFile` command
  - allows to upload data to existing PMRs, also available in the GUI
- Added support for zLinux and Ubuntu



Grow Call Home to become a Proactive-Service

Store anonymized callhome data from all customers in a central database within IBM.

- This includes the full cluster configuration, system events and performance data.

Generate reports from the callhome data to better understand how customers use our system.

Use simple rules as well as advanced analytics to proactively detect issues and give recommendations to the customer and/or support team.

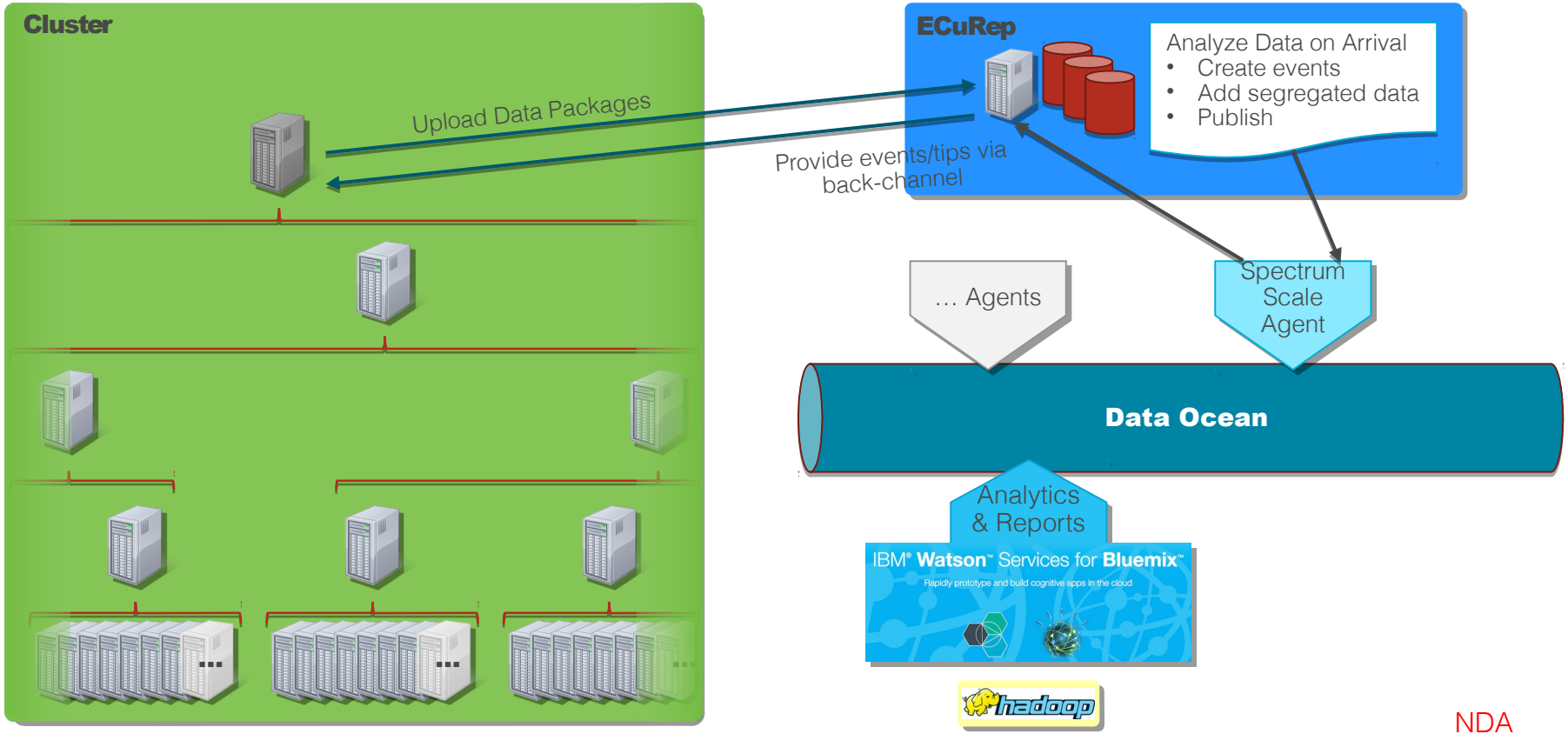
Receive real-time tuning suggestions tailored to customers specific workload and configuration

- Tuning tips based on the measured system performance and configuration
- Tuning tips based on predefined rules to detect common mis-configuration

Receive real-time alerts and other notifications relevant to customers specific configuration

- Get informed about potential issues / mis-configuration based on predefined rules
- Get informed about potential issues based on findings on other systems/customers (advanced analytics)

Grow Call Home to become a Proactive-Service



NDA

Why enable Call Home

- Improve service response time
  - IBM Service team can start with problem analysis without delay
    - Call Home collects the full cluster config, health events and some debug data
  - Future release: Automatically open PMRs for selected cases

- Improve customer support experience (consumability/ease of use)
  - Ability to upload gpfs snaps directly from the system
  - Future release: Easier to open PMRs, through GUI

- Help IBM to improve development and testing
  - See which functionalities / components are used
  - Understand the cluster setup, configuration, workloads, etc.
  - With that: Better guide our test labs on which features/setups/workloads should be tested

Running 4.2.3.7, 5.0.0 or higher? Please enable Call Home!

How to enable Call Home

Decide which mode to use

- ***Automatic Call Home group creation***
  - *Creates one or more Call Home groups*
    - *depends on the number of nodes with internet access*
    - *Internet access nodes can be specified or discovered automatically*
  - *Assigns each cluster node to one of the groups*
  - *Recommended because of simplicity*
- ***Manual Call Home group creation***
  - *Full flexibility to define groups and assign nodes*
  - *Allows to configure Call Home on a subset of nodes only*
  - *Recommended on large clusters (>100 nodes)*

How to enable Call Home

1. Configure basic call home settings

- Set up the customer information, using the command:
  mmcallhome info change --customer-name CustomerName --customer-id CustomerId --email Email --country-code CountryCode
- Set up the scheduled data collection, if needed, using the following commands:
  mmcallhome schedule add --task DAILY
  mmcallhome schedule add --task WEEKLY
- If you are using proxy, configure the proxy settings.
  - Set the proxy location and authentication settings, using the command:
    mmcallhome proxy change --proxy-location ProxyLocation –proxy-port ProxyPort [--proxy-username ProxyUsername –proxy-password ProxyPassword]
  - Enable the proxy, using the command:
    mmcallhome proxy enable [--with-proxy-auth]
- Enable the callhome capability, using the following command:
  mmcallhome capability enable

How to enable Call Home

2. Automatic group creation

mmcallhome group auto [--server servername[,servername]]

- *When omitting the --server option, the nodes with internet access will be discovered automatically*
- *When specifying the --server option:*
  - *the number of servers defines the number of groups*
  - *Make sure the servers have internet access*

```
[root@md-11 ~]# mmcallhome group auto --server md-11
…
[root@md-11 ~]# mmcallhome group list
callHomeGroup   callHomeNode   callHomeChildNodes
--------------- -------------- --------------------
autoGroup_1     md-11          md-11,md-13,md-12
```

How to enable Call Home

2. Manual group creation

mmcallhome group add GroupName server [--node {all | ChildNode[,ChildNode...]}]

- *Don‘t put more than 32 nodes into a group*
- *Make sure that the specified server has internet access*

```
[root@md-11 ~]# mmcallhome group auto --server md-11
…
[root@md-11 ~]# mmcallhome group list
callHomeGroup   callHomeNode   callHomeChildNodes
--------------- -------------- --------------------
autoGroup_1     md-11          md-11,md-13,md-12
```

Questions ?

Legal notices

Copyright © 2018 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 intellectually 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 inquires, 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.