

Hewlett Packard
Enterprise

HPE Performance Cluster Manager

Jan Wender
March 15, 2023

The Age of Insight

The age of insight comes with rising demands as compute-intensive applications and workloads are quickly exceeding the capacity of traditional infrastructure.

This poses a major challenge for enterprises looking to get more performance out of a limited budget.

As these requirements continue to evolve, enterprises must be able to upgrade their infrastructure to capitalize on the latest data and technology trends.

**Massive
Data Growth**

**Larger
Models**

**New
Algorithms**

Performance

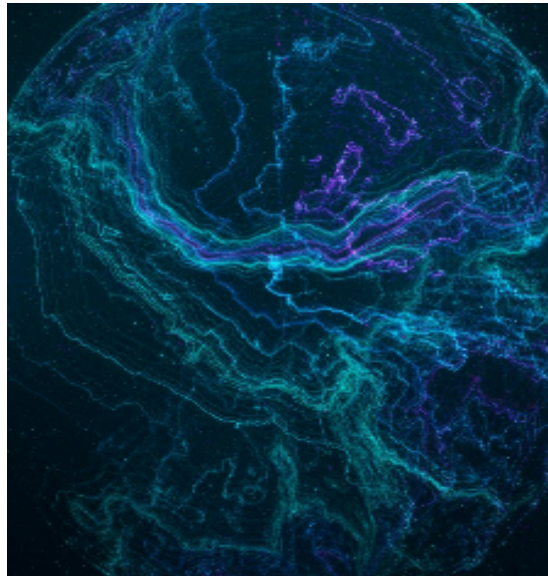
Sustainability

Efficiency

Thriving in the New Age of Insight

New Compute

- Purpose-built density-optimized systems
- Diversity of processors
- Data-intensive applications



New Interconnect

- High bandwidth
- Adaptive routing
- Congestion management
- Ethernet compatibility



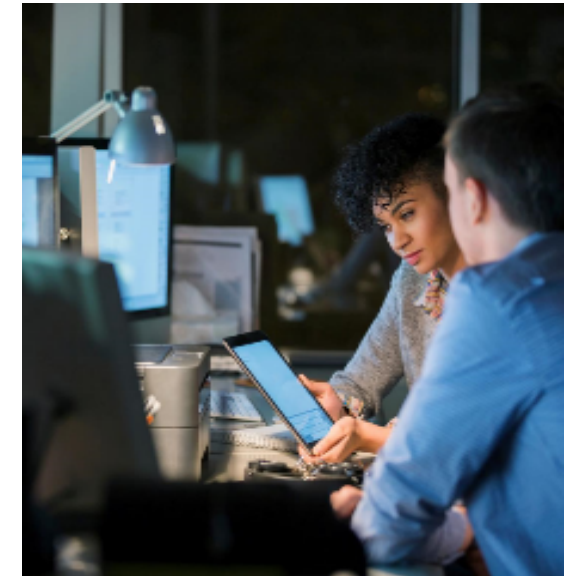
New Storage

- Intelligent data management
- Efficient performance
- Unprecedented scalability



New Software

- Converged workflows
- New standards in manageability
- Cloud experience



Software Requirements for the Era of Insight



Productivity

Turning complex infrastructure into easily manageable and reliable systems capable offering the best ROI.



Performance

Better performance for HPC & converged workloads regardless of underlying architecture.

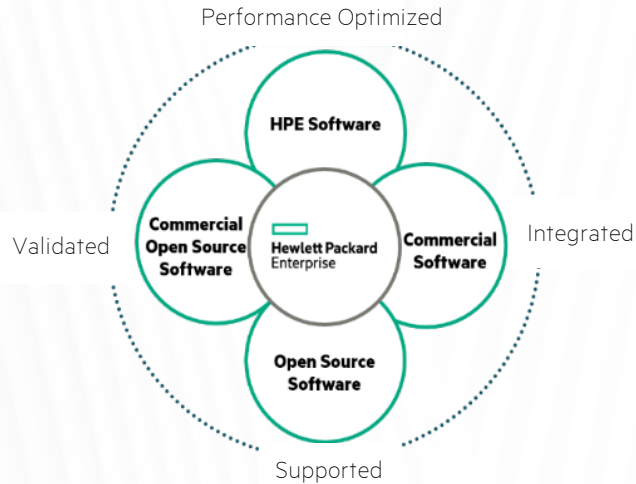


Cloud as an experience

Everything-aaS.
Resources available from everywhere, anytime.



HPE HPC and AI Software Solutions Powering the New Era



From HPC Experts for HPC Experts

HPE leads the industry with experts from Cray, SGI and HPE who have been delivering HPC solutions for over 30 years.

Complete Value Chain

We offer flexible software stacks for all customer needs – from tools for easier system management to suites for faster software development and more – helping to achieve best cost/performance ratio of the systems.

Vision

Software innovations we developed for the first Exascale supercomputers* are available to our HPC customers across the world so they can get results faster.

End-to-End Solutions

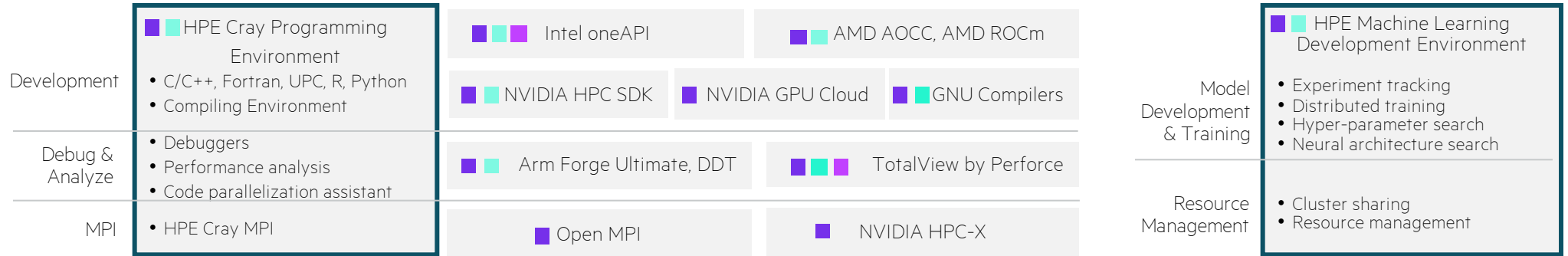
Our experts deliver and support a complete range of solutions and consumption models - supporting customers on their journey towards hybrid models and converged workflows.

* ORNL's Frontier Top 500 Announcement

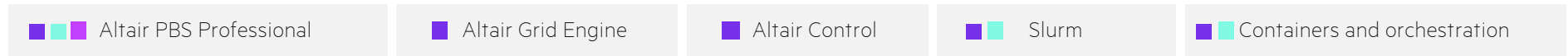
HPE HPC and AI Software Portfolio

Most comprehensive portfolio on the market

Application and Software Development Ecosystem



Workload Management & Orchestration



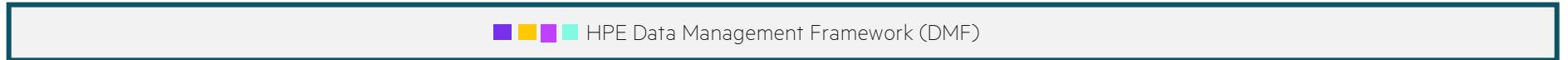
Remote Visualization



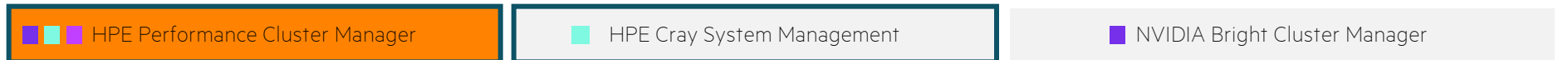
Storage File Systems



Data Management



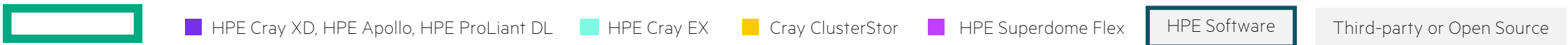
System Management



Fabric Software



Operating System



HPE Performance Cluster Manager

All you need to manage your HPE clusters, keep them healthy and running at peak performance

Flexible, easy to use system management solution offering system administrators all the tools they need to turn even the most complex hardware into easily manageable systems capable to accommodate growing variety of workloads.

Powerful

Comprehensive set of tools you need to manage all aspects of your cluster.

Productive

Designed to maximize productivity of your cluster, automate actions, and optimize running costs.

Flexible

Customize monitoring, alerts, and management actions to best suit your needs.

Scalable

Manage systems from dozens of nodes up to Exascale.

Anywhere

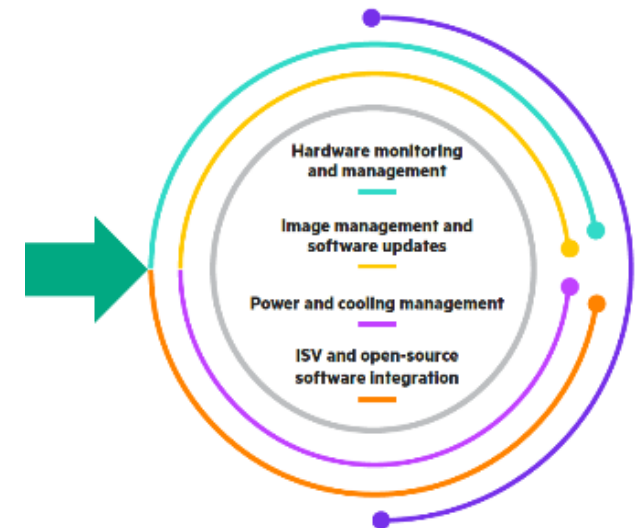
Suitable for both on-premises as well as hybrid HPC deployments.

Proven

Used by hundreds of customers around the globe for over a decade.



SYSTEM SETUP



System Setup



Start Here

The screenshot shows the HPE Performance Cluster Manager web interface. The browser window title is "HPE Performance Cluster Manager v1.3.stout721a.200414.194934.0652c4 - Mozilla Firefox". The address bar shows "https://localhost". The page header includes the Hewlett Packard Enterprise logo and the title "HPE Performance Cluster Manager".

The main content area is divided into two columns:

- Using HPE Performance Cluster Manager:** This column contains a list of links with document icons:
 - [Launch HPE Performance Cluster Manager GUI](#)
 - [Getting Started](#)
 - [Installation Guide](#)
 - [Administration Guide](#)
 - [Power Management Guide](#)
- Cluster status:** This column features a large green circular progress indicator showing "2235 Nodes Total". To the right, a legend indicates the status of the nodes:
 - Up: 2235 Nodes
 - Unknown: 8 Nodes
 - Down: 2 Nodes

The bottom section, titled "Resources", contains links to documentation and support:

- [REST API Documentation](#)
- [Release Notes](#)
- [HPE Performance Cluster Manager Web Site](#) for up-to-date documentation.
- [HPE Support Center](#) for patch announcements and downloads.

A notification at the bottom of the browser window reads: "It looks like you haven't started Firefox in a while. Do you want to clean it up for a fresh, like-new experience? And by the way, welcome back!" with a "Refresh Firefox..." button.

**[CLICK HERE](#)
TO WATCH DEMO**

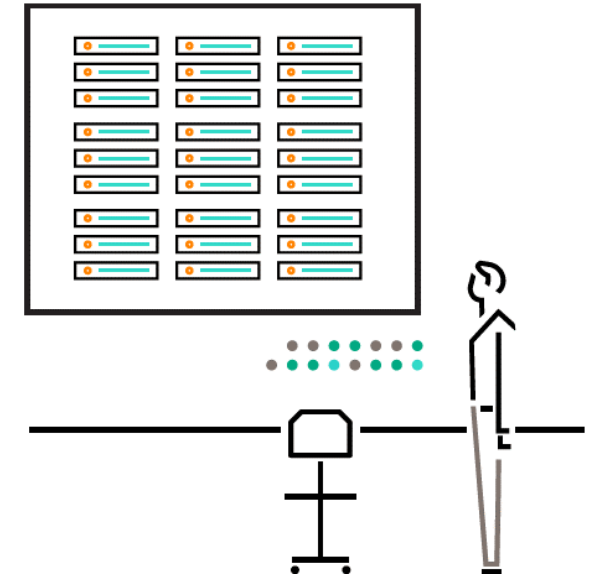
HPE Performance Cluster Manager System Setup

Get into production quickly

- Install all the software and deploy your system over bare metal in minutes rather than hours or days
- Nodes are provisioned in parallel for maximum system performance in one session
- During the setup, hardware elements are automatically discovered and configured

Are you ready to go into production?

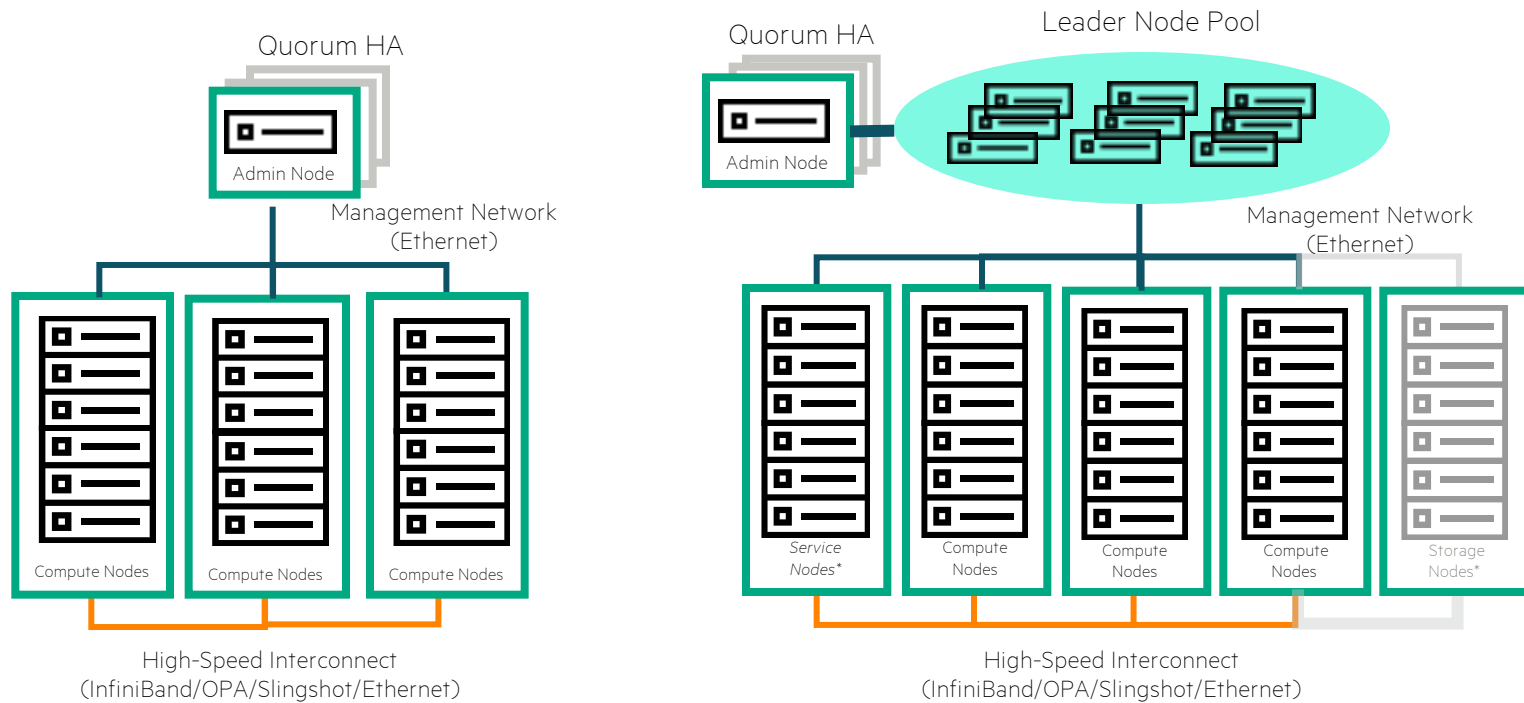
We include the same diagnostic tests in the software suite as we use to test your cluster in the HPE factory before shipping for compliance check.



Grow your cluster without disrupting service

New nodes can be added to the system without requiring the system shutdown with new hardware automatically discovered and provisioned fast.

HPE Performance Cluster Manager System Configurations



Small Scale—Flat Cluster Configurations

- For HPE Apollo/HPE ProLiant clusters
- Manage small HPC clusters with <1K nodes
- Easily scale and add more nodes as needed without disrupting service

Large Scale Cluster Configurations

- For HPE Apollo/HPE ProLiant clusters and HPE Cray Ex systems
- Achieve high resiliency with **Leader Node Pool** which consists of shared leader nodes that manage groups of compute nodes, active-active (RHEL or SLES HA subscription required)
- 3 leader nodes per every 1.5k compute nodes

Avoid failover with Quorum HA

Pool requires a minimum of 3 admin nodes to manage high availability

- Eliminates hardware shared storage requirement
- Uses gluster in sharding mode to host the virtual machine image
- Pacemaker to start and position the virtual machine as needed

NEW

Improved System Configuration

- Simplifies node referencing by enabling creation of many aliases for hosts
- Ability to configure as few or as many networks as required
- Can now add/remove or modify networks through CLI

Monitoring and Management

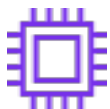


HPE Performance Cluster Manager Monitoring and Management

HPE Performance Cluster Manager offers fine-grained centralized monitoring and management of your cluster to keep it performing at its best:



CPUs



GPUs



**Memory
and discs**



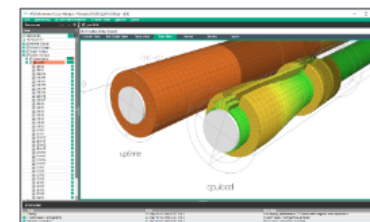
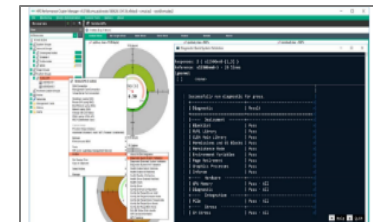
Networking



**Power &
cooling**



Software



- View metrics and alerts via GUI, CLI, Prometheus, Kibana, or Grafana
- Customize system telemetry and alerts to best suit your needs
- Set up automatic reactions to events to prevent failures

HEALTH CHECKS

Help you run applications reliably and at peak performance:



- Introducing monitoring and health-check telemetry for Slingshot fabric and AMD GPUs
- Check nodes before you launch an application to avoid launching jobs on unhealthy nodes
- Analyze why jobs failed or performed poorly
- Run periodic health checks to make sure your cluster is healthy
- Schedule invasive tests during maintenance windows so the system can perform at its best

Hardware Management



Hardware Setup

- Hardware: admin nodes, leader nodes, compute nodes (CPUs & GPUs, with & w/o disk), I/O nodes, service nodes, login nodes, network and NICs
- Automatic discovery and configuration of hardware (incl. iLO-based)
- Secure and authenticated repository for system setup information and telemetry
- Supports standard protocols—Redfish, IPMI, SNMP
- Leader node pools manage multiple groups of nodes in active-active configuration; Quorum HA to avoid failover
- Cluster verification tool for more consistent cluster deployment
- No disruption to existing system with rolling hardware upgrades



Integration

- NVIDIA DCGM integration for NVIDIA GPU management and control
- AMD ROCm support for AMD GPU management and control
- Hardware inventory output in Ansible playbook format
- Integration with NVIDIA UFM connects IB switch telemetry with cluster node telemetry



Maintenance

- All system-wide alerts are consolidated and managed by the Alerta alerts monitoring system
- Pre-defined and customizable alerts for wide variety of events
- Kafka support enables high performance data movement from system telemetry collection to Elastic Stack repository
- Multiple ways to visualize telemetry: Kibana, Grafana, CLI, GUI
- Firmware flashing for BIOS, BMC, CMC, network adapters, switches on all nodes (iLO or non-iLO enabled)
- FRU Inventory & alerts (incl. detecting new hardware, notifications when HW changes name/is replaced/pulled out) for serviceability
- Thermal monitoring and management
- Comprehensive health checks to keep systems resilient
- Check for differences between nodes
- Broadcast commands to multiple nodes
- * Power ramp up/down sequencing of service nodes for more resiliency

Node Status

The screenshot displays the HPE Performance Cluster Manager v1.0 - steelcoe interface. The left sidebar shows a tree view of resources, with 'login-node' selected. The main panel shows the configuration details for 'login-node' under the 'Monitoring' tab. The details are organized into sections: network, image, platform, management, attributes, and internals. Yellow boxes highlight specific sections: 'Network addresses', 'Software image', 'Server', 'Management Network', 'Attributes', and 'Internals'. The 'Attributes' section includes a 'DHCP_BOOTFILE' label. The 'Information' panel at the bottom shows system messages.

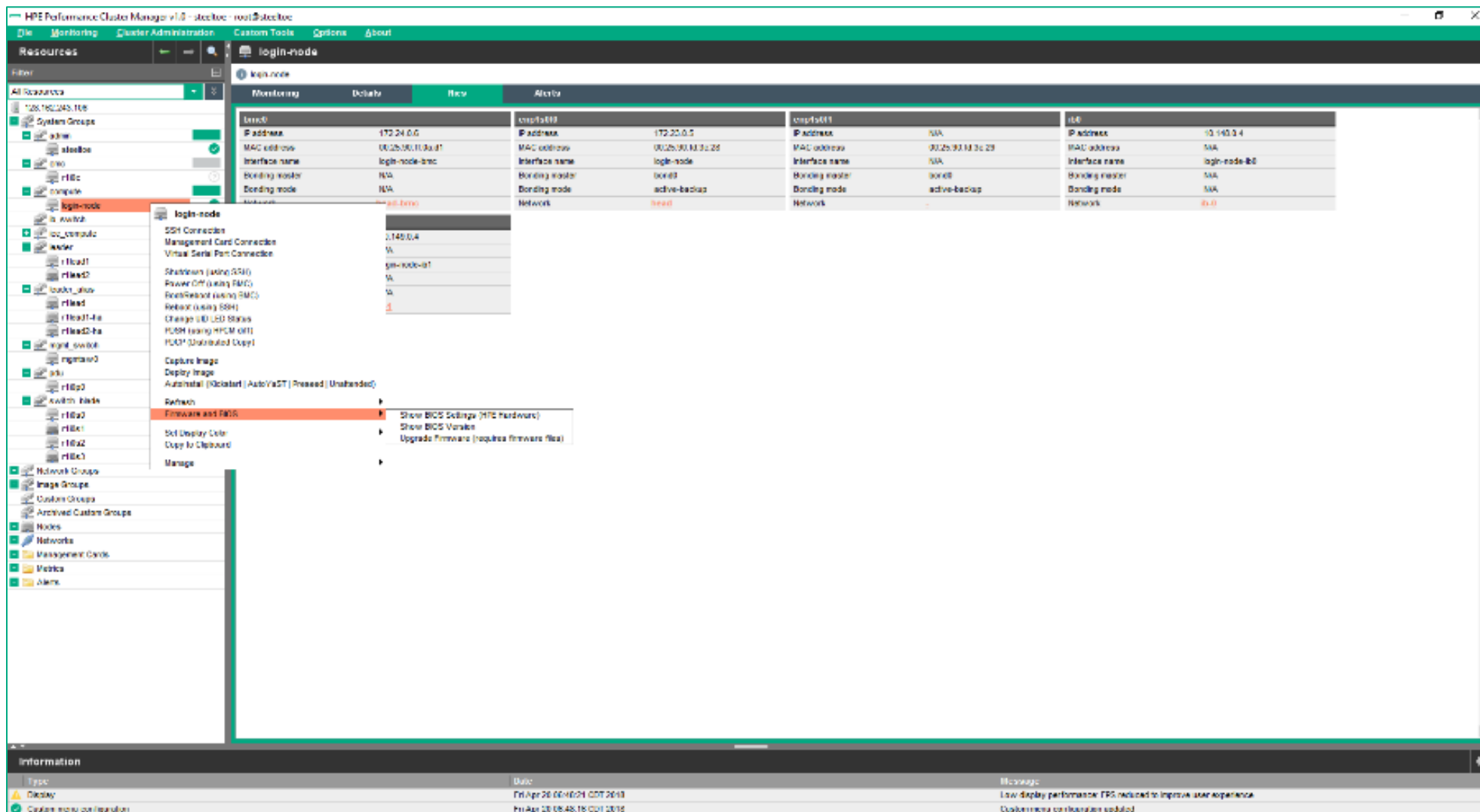
Attribute	Value
ISCSI root	none
network	
Network group	default
Default gateway IP address	172.24.0.6
IP address	00:25:80:14:3c:20
MAC address	255.255.0.0
Subnet mask	default
Management server IP address	default
image	
Image group	slca12a3
Kernel	4.4.120-34.17-default
Cloning block device	default
Cloning date	Fri Apr 20 06:28:14 CDT 2016
platform	
Platform name	generic
Architecture	x86_64
Serial port	lly51
Serial port speed	115200
Vendor arguments	default
management	
Management card type	none
Management card IP address	172.24.0.6
Management card MAC address	00:25:80:14:3c:21
Protocol	default
Username	*****
Password	*****
attributes	
CONSERVER_LOGGING	yes
CONSERVER_ONDEMAND	no
DHCP_BOOTFILE	grub2
DISK_BOOTLOADER	no
MGMT_BONDING	active-backup
PREDICTABLE_NET_NAMES	yes
REDUNDANT_MGMT_NETWORK	yes
SWITCH_MGMT_NETWORK	yes
TEMPO_CPOWER	yes
TPM_BOOT	no
dnx_domain	steelcoe2.america.sgi.com
internals	
ID	17
UUID	fcc3d71e-1b34-4c7a-a7c2-2a574c2d2ef6
Creation time	Fri Apr 20 05:54:28 CDT 2016
Modification time	Fri Apr 20 06:03:01 CDT 2016
Deletion time	

Type	Date	Message
Display	Fri Apr 20 06:40:21 CDT 2016	Low display performance: FPS reduced to improve user experience
Custom menu configuration	Fri Apr 20 06:40:16 CDT 2016	Custom menu configuration updated

Firmware and BIOS Update

Easy firmware and BIOS update from GUI

Updates iLO and other BIOS software



Broadcast Commands to Multiple Nodes

Each window = cluster node

```
[root@r210n0 ~]# hostname
r210n0
[root@r210n0 ~]# crinfo --nodetype
compute
[root@r210n0 ~]# crinfo --rack
2
[root@r210n0 ~]# crinfo | grep NIC
IB_1_NIC="ib1"
GBE_NIC="ens11"
BMC_NIC="bmc0"
IB_0_NIC="ib0"
```

Filesystem	Size	Used	Avail	Use%	Mounted on
system	100	1.5G	19G	18%	/
tmpfs	19G	0	19G	0%	/dev
devtmpfs	19G	0	19G	0%	/dev/shm
tmpfs	19G	25M	19G	1%	/run
tmpfs	19G	0	19G	0%	/sys/fs/cgroup
tmpfs	158M	0	158M	0%	/tmp
tmpfs	3.2G	0	3.2G	0%	/run/user/0

Type command once for 4 nodes

```
[root@r210n11 ~]# hostname
r210n11
[root@r210n11 ~]# crinfo --nodetype
compute
[root@r210n11 ~]# crinfo --rack
2
[root@r210n11 ~]# crinfo | grep NIC
IB_1_NIC="ib1"
GBE_NIC="ens11"
BMC_NIC="bmc8"
IB_0_NIC="ib0"
```

Filesystem	Size	Used	Avail	Use%	Mounted on
tmpfs	19G	1.5G	17G	18%	/
devtmpfs	19G	0	19G	0%	/dev
tmpfs	19G	0	19G	0%	/dev/shm
tmpfs	19G	25M	19G	1%	/run
tmpfs	19G	0	19G	0%	/sys/fs/cgroup
tmpfs	158M	0	158M	0%	/tmp
tmpfs	3.2G	0	3.2G	0%	/run/user/0



Identify Inconsistencies

The screenshot displays the HPE Performance Cluster Manager interface. On the left, a tree view shows resources under 'head-nodes' and 'job-2438'. A list of resources (r2i0n0 to r2i1n3) is shown with status indicators. The main window shows a terminal window with network configuration details for various nodes. A yellow callout box highlights the text 'Immediately find differences'.

Immediately find differences

```
Responses: 10 { r2i0n[0-1,10-17] }
Reference: r2i0n0 - 20 lines
Ignored:
[ ] <none>

kernel_versi
cpuload
memory_us
process_mx
page_cache
buffer_cach
uptime
disk_write
disk_read
net_in
net_out
10% logins
10% swap_aval
10% swap_used
10% cpu_frequer
10% NFS_mount

1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN qlen 1
   link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
   inet 127.0.0.1/8 scope host lo
       valid_lft forever preferred_lft forever
   inet6 ::1/128 scope host
       valid_lft forever preferred_lft forever
3: ens11: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 9000 qdisc mq master br0 stat
   link/ether 44:8a:5b:39:43:45 brd ff:ff:ff:ff:ff:ff (all different)
   link/ether 44:8a:5b:39:43:45 brd ff:ff:ff:ff:ff:ff x 1: r2i0n0
   link/ether 44:8a:5b:39:43:6c brd ff:ff:ff:ff:ff:ff x 1: r2i0n17
   link/ether 44:8a:5b:39:43:53 brd ff:ff:ff:ff:ff:ff x 1: r2i0n15
   link/ether 44:8a:5b:39:c9:5c brd ff:ff:ff:ff:ff:ff x 1: r2i0n16
   link/ether 44:8a:5b:39:43:47 brd ff:ff:ff:ff:ff:ff x 1: r2i0n12
   link/ether 44:8a:5b:39:c9:76 brd ff:ff:ff:ff:ff:ff x 1: r2i0n18
   link/ether 44:8a:5b:39:43:6e brd ff:ff:ff:ff:ff:ff x 1: r2i0n13
   link/ether 44:8a:5b:39:43:29 brd ff:ff:ff:ff:ff:ff x 1: r2i0n1
   link/ether 44:8a:5b:39:43:3b brd ff:ff:ff:ff:ff:ff x 1: r2i0n14
   link/ether 44:8a:5b:39:43:57 brd ff:ff:ff:ff:ff:ff x 1: r2i0n11
   inet6 fe80::468a:5bff:fe39:4345/64 scope link (all different)
       valid_lft forever preferred_lft forever
4: br0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 9000 qdisc noqueue state UP qlen
   link/ether 44:8a:5b:39:43:45 brd ff:ff:ff:ff:ff:ff (all different)
   inet 10.159.4.2/22 brd 10.159.7.255 scope global dynamic br0 (all different)
       valid_lft 27690sec preferred_lft 27690sec (all different)
   inet6 fe80::468a:5bff:fe39:4345/64 scope link (all different)
       valid_lft forever preferred_lft forever
5: ib0: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 65520 qdisc pfifo_fast state >
   link/infiniband 80:00:02:08:fe:80:00:00:00:00:00:00:00:00:7c:fa:90:03:00:b7:ea: (all different)
   inet 10.148.0.91/16 brd 10.148.255.255 scope global ib0 (all different)
```

NVIDIA GPU Management via HPCM

- NVIDIA Datacenter GPU Manager (DCGM) enables GPU diagnostics, health, configs, alerts on a single server
- DCGM features integrated in HPCM:
 - Diagnostic tests
 - Health watches
 - Set/enforce config targets
 - NVLink error counts
 - Device info
 - Topology
- HPCM makes GPU health management cluster-aware by aggregating the GPU health metrics and alerts

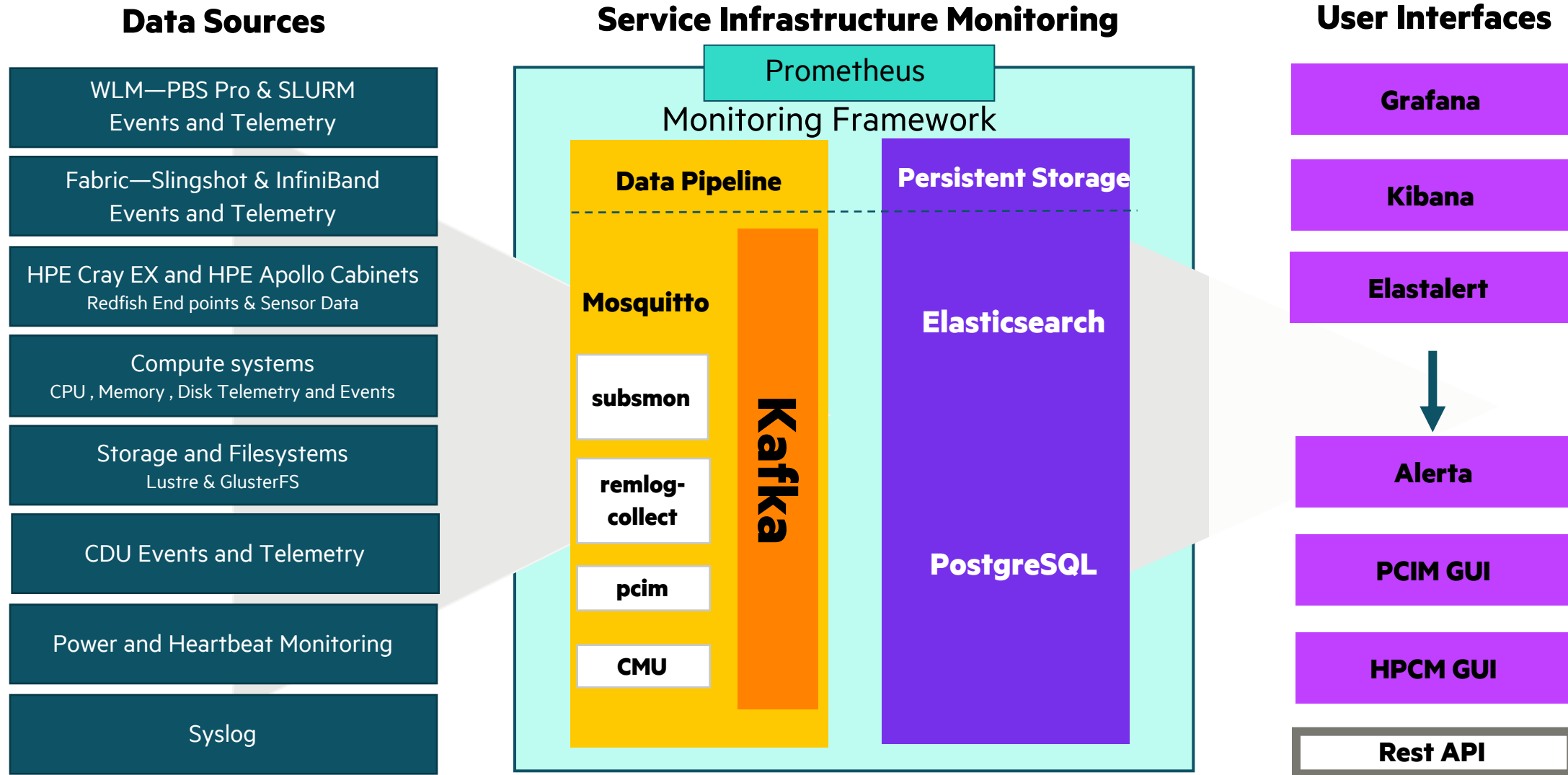
Quick System Validation shows that 2 nodes pass all GPU validation tests

Metrics collected include: GPU load, memory use, power, temperature and more

Diagnostic	Result
----- Deployment -----	
Blacklist	Pass
NVML Library	Pass
CUDA Main Library	Pass
Permissions and OS Blocks	Pass
Persistence Mode	Pass
Environment Variables	Pass
Page Retirement	Pass
Graphics Processes	Pass
Inforom	Pass
----- Hardware -----	
GPU Memory	Pass - All
Diagnostic	Pass - All
----- Integration -----	
PCIe	Pass - All
----- Stress -----	
SM Stress	Pass - All

[CLICK HERE TO WATCH DEMO](#)

HPE Performance Cluster Manager Data Pipeline



Service Infrastructure Monitoring

HPCM enables granular monitoring of services using Prometheus

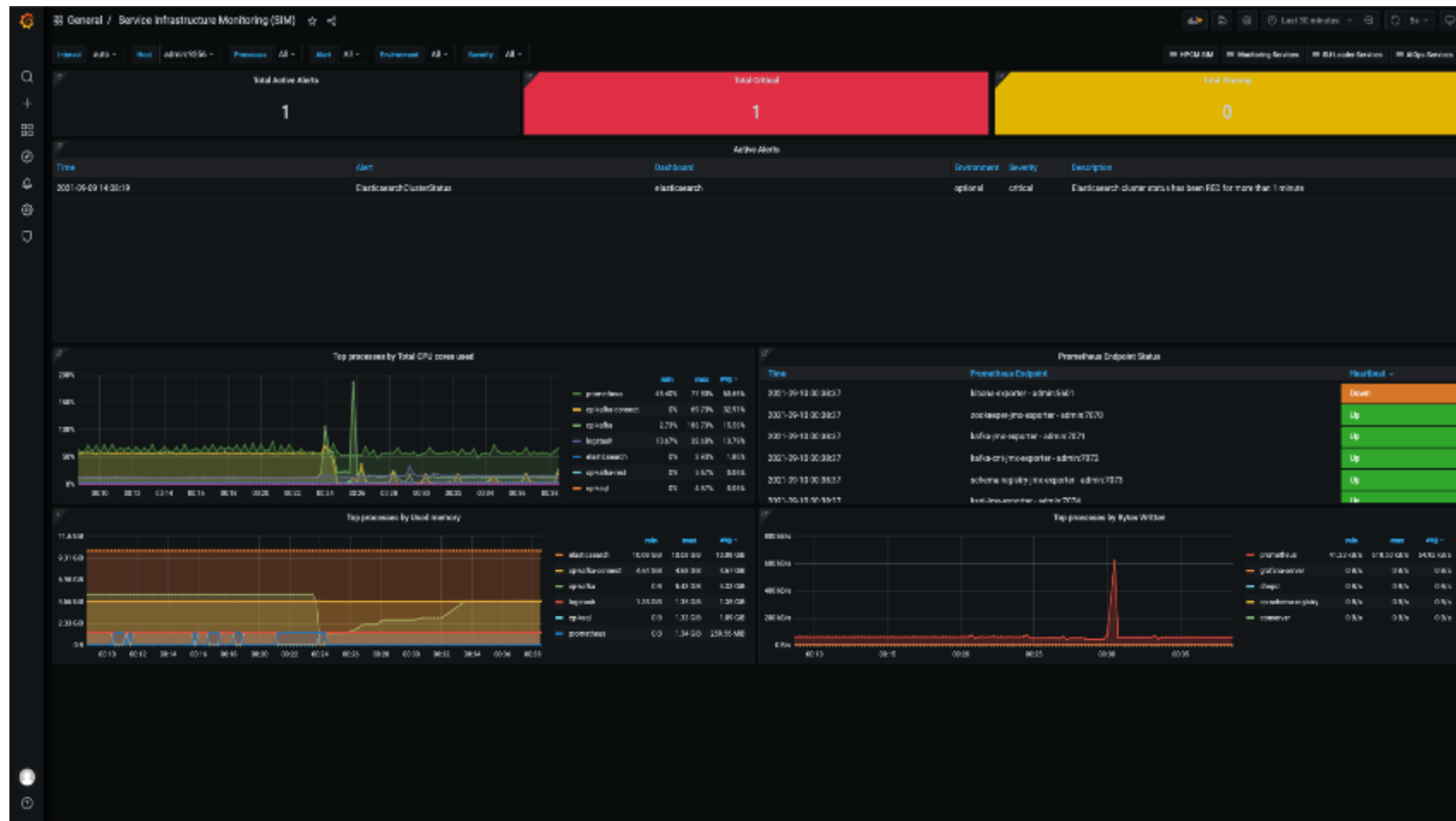
Service Infrastructure Monitoring (SIM)

HPCM enables system administrators to monitor the health and performance of services that run within HPCM.

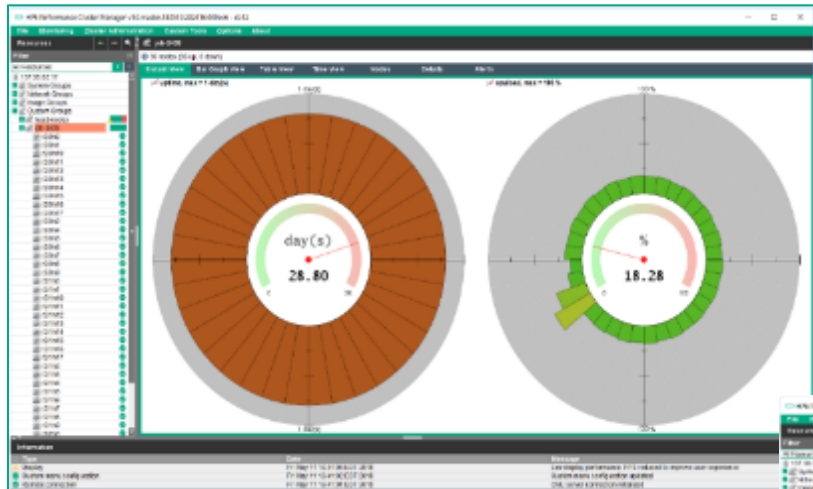
- Kafka
- ElasticSearch
- CTDB
- Kibana
- Gluster
- AIOPs
- Node Status
- CPU usage

Offers comprehensive monitoring and health-check data of the service infrastructure

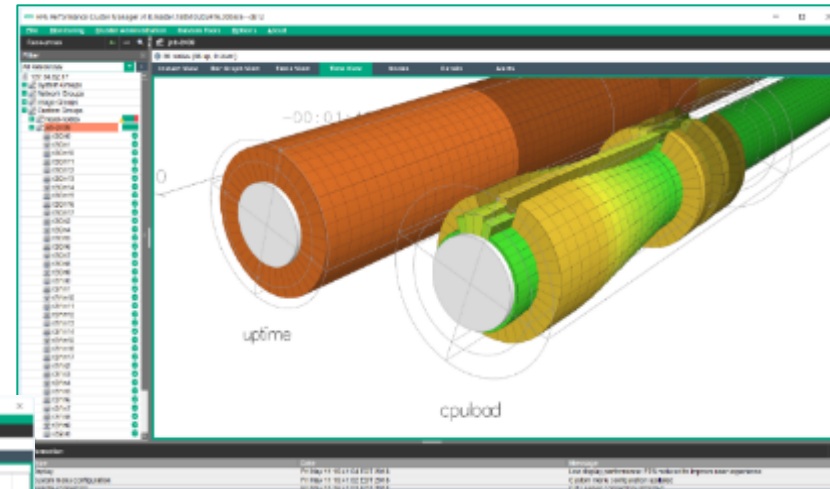
CLICK HERE TO WATCH DEMO



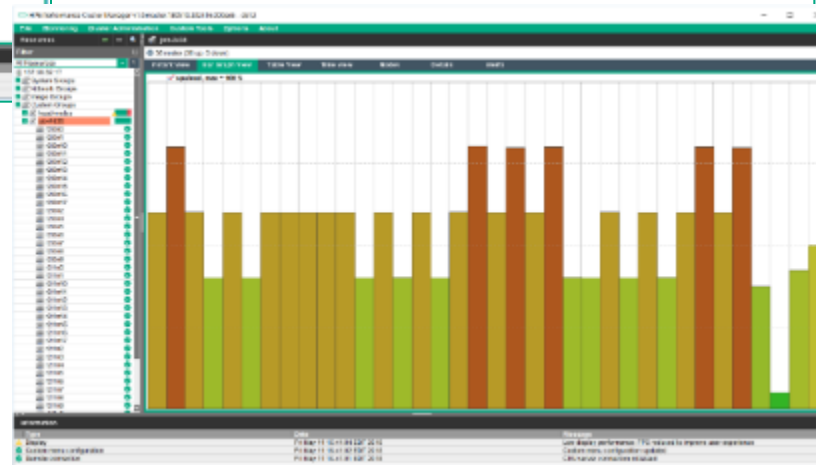
Hardware Monitoring in HPE Performance Cluster Manager GUI



“Instant View” of a set of nodes showing selected hardware metrics (Uptime and CPU load).



3D “Time View” shows the metrics over time - each petal represents a node.



Simplified “Bar Graph” view of the same metric. “Table view” is also available.

HPE Performance Cluster Manager – View and React to Alerts

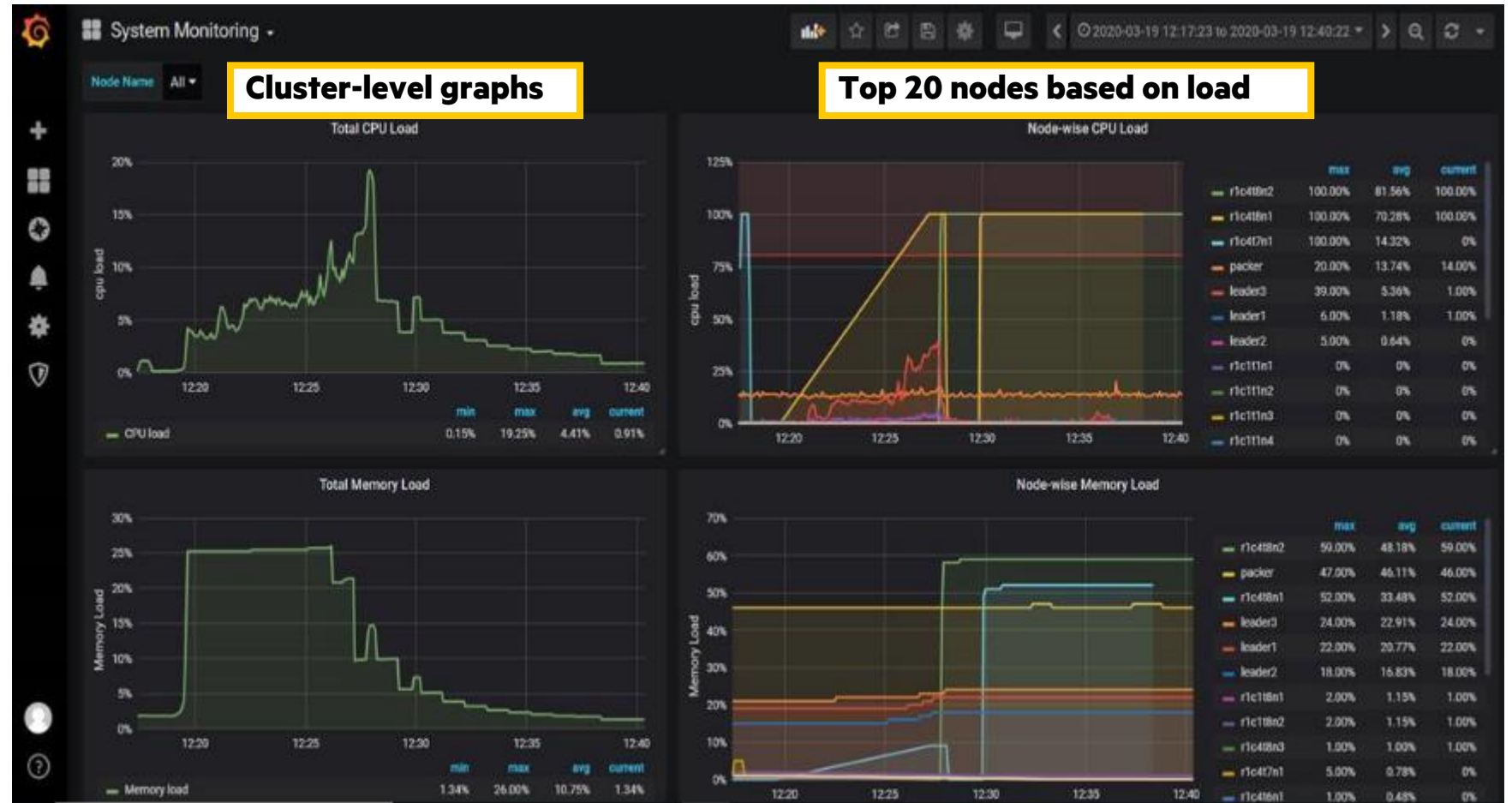
- Best tools available in Open Source, integrated in HPCM for alerts telemetry
- Collection of system-wide alerts enabled through Alerta
- Alerta sends the alerts to Elasticsearch repository and instant visualization through Kibana.
- Real-time alerts management is part of the HPCM cluster health check capability



HPE Performance Cluster Manager – System Monitoring in Grafana

- Preconfigured **System Monitoring Dashboard** in Grafana shows:
 - CPU Load
 - Memory Load
 - CPU Frequency
 - Disk Read
 - Disk Write
 - Process Memory
 - Swap Used
 - Swap Avail
 - Page Cache

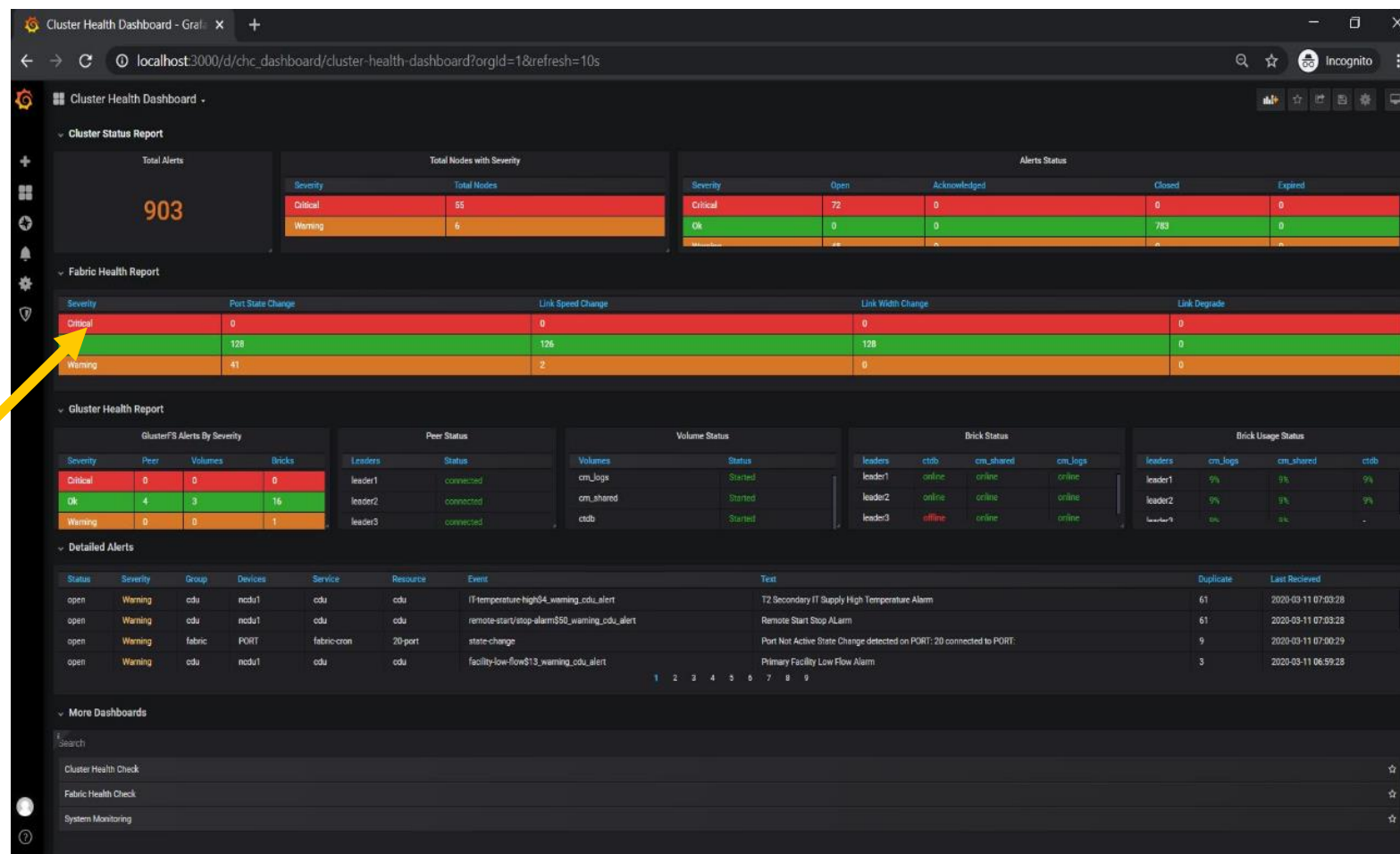
Accessible from
the **Health Management
Dashboard** (see next section).



NEW HPE Slingshot Fabric Monitoring with HPE Performance Cluster Manager

Keeping the clusters productive and resilient

- HPE Slingshot telemetry and alerting
- Monitors fabric at large scale, Exascale
- Runs standard benchmarks for fabric performance checks



HPE Slingshot Fabric Health Report

- Fabric port-state changes
- Link speed changes
- Link width changes
- Link degrade changes
- NUMA consistency

Lustre Filesystem Monitoring and Visualization in Grafana

For easier data management

Lustre Overview Dashboard

Alert numbers as well as information on minimum, maximum and current statistics hardware, including:

- MD CPU Load
- OSS CPU Load
- IME CPU Load
- MDS state
- OSS Bandwidth
- IME Bandwidth
- MDS Load
- OSS Load
- IME Load

Also offered—Lustre Infinite Memory Engine (IME) Dashboard.

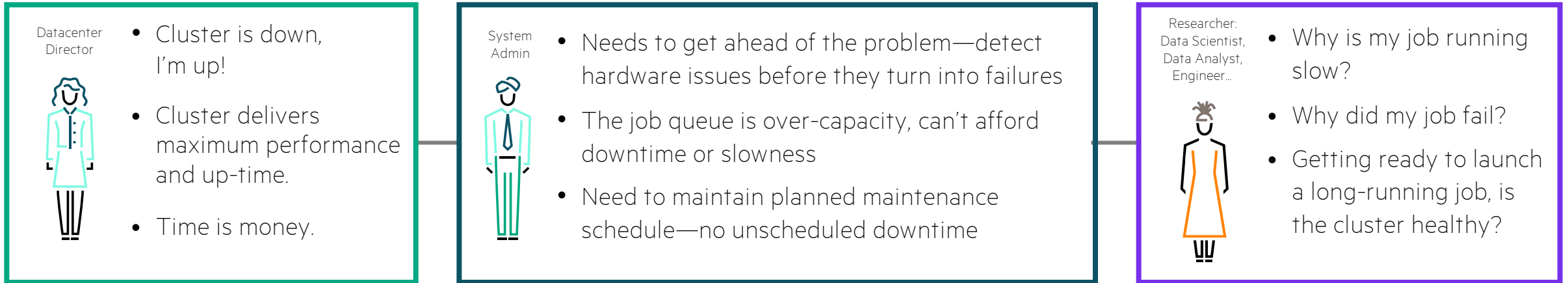


Cluster Health Management

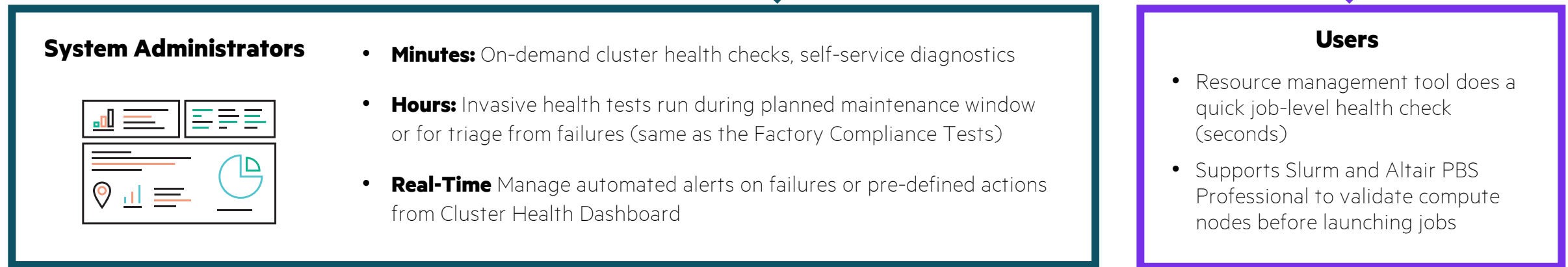


HPE Performance Cluster Manager Cluster Health Management Capabilities

Cluster Health Management addresses data center concerns:



Cluster Health Management capabilities:



Shorter Time to Production with System Diagnostics

New System Setup



HPE Factory

We supply the same system diagnostics used by HPE factory to customers for compliance check



Customer

Test	Description	Test	Description
CPU Uniformity Checks	Finds the outlier nodes with non-uniform configuration	Check the Health of the InfiniBand Fabric	Link integrity checks, Missing links, Down links, link speed and width, Subnet manager availability etc.
CPU Performance Check	Runs High Performance Linpack on each compute nodes and compares the GFLOPS reported on each node by High-Performance LINPACK (HPL).	Check the Health of Gluster Filesystem	Check the Peer node availability, Availability of Volumes, Availability of bricks etc.
Cluster-wide CPU Performance Check	Runs the cluster-wide High Performance Linpack on a cluster.	Disk Health Check	Writes different patterns to disk and read the data back to check the consistency.
Memory Uniformity Checks	Finds the outlier nodes with non-uniform configuration.	I/O Performance Check Tool	Used to benchmark disk performance. Detects the low performing disks on each node.
Memory Performance Check	Runs the STREAM benchmark on each node and reports the TRIAD value. This check reports low performing nodes.	Filesystem Performance Check	Benchmarking and in I/O performance measurement. Detects the slow filesystems and helps to identify the bottlenecks for filesystem finetuning.
Ethernet Uniformity Checks	Finds the outlier nodes with non-uniform fabric configuration.	GPU Health Check Tool	Stress the GPUs to find out the Errors and identify faulty GPUs.
Fabric Checks	Health of InfiniBand HCAs and HPE Slingshot high-speed network (HSN) Cassini cards.	CDU Checks	Reports the health of the cooling distribution units (CDUs) in the cluster.
Fabric Performance Health Check	Conducts a point-to-point MPI bandwidth test to measure InfiniBand fabric performance. Highlights the outlier nodes.	Console Checks	Check the console availability of compute nodes. Report the nodes with inaccessible consoles.

Cluster Health Diagnostics with HPE Performance Cluster Manager

Prevent serious issues - keep your system resilient and in production

Key Hardware

CPU

- Model, Architecture, Online vs Offline Cores, threads
- Industry Standard benchmarks to check CPU Performance

Memory

- Available Memory, Swap space, THP, DIMM Size, DIMM Speed, Balanced DIMM Configurations
- Industry standard benchmarks for Memory Performance Checks

GPU

- GPU Diagnostics
- AMD ROCm support for AMD GPU management and control
- NVIDIA DCGM integration for NVIDIA GPU management and control

Key System Components

- PCI speed, Firmware, Fans, CDU

Disks

- Errors, Bad blocks,
- IOPS, Bandwidth, latency

Ethernet

- Link Speed, Link Detection, Drivers, Firmware, PCI slots, pings

Key Software

Schedulers

- PBS and SLURM
- Node Health and Job Health

GlusterFS

- Node, Volume, Brick Health
- Warns on bricks usage limits

Filesystems

- Bad blocks, Filesystem Errors
- Industry standard Performance Benchmarks

Interconnect

InfiniBand Fabric

- Host Level and Topology Level
- HCA Type, Speed & width, NUMA consistency
- Link Speed, Link width, Missing Links, Link Integrity
- Standard Benchmarks for Fabric Performance Checks

HPE Slingshot Fabric

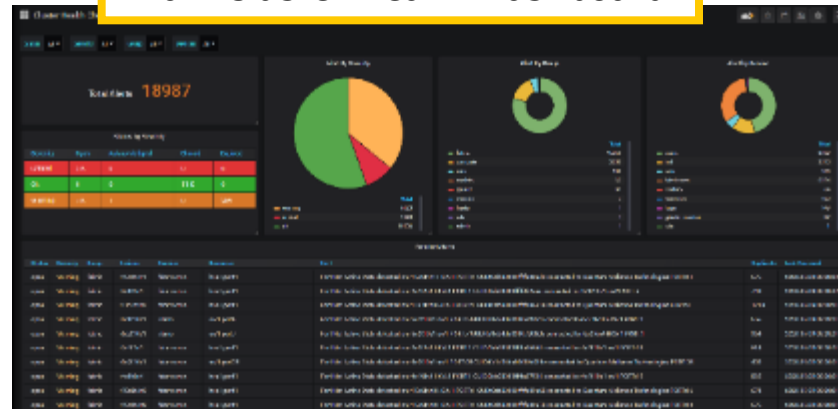
- Host Level and Topology Level
- HCA Type, Speed & width, NUMA consistency
- Link Speed, Link width
- Missing Links, Link Integrity
- Standard Benchmarks for Fabric Performance Checks

Real-Time Cluster Health Monitoring with HPE Performance Cluster Manager

Visualization dashboards

- Cluster Health at-a-Glance—Single Pane view for the complete cluster Health Status
- Live System Monitoring—Dashboards for key metrics like Power, Cooling, CPU, Memory, Disk, Fabric, Gluster, Job Scheduler monitoring metrics
- Scalable—Highly scalable data pipeline at the backend
- Customizable—Create new dashboards easily

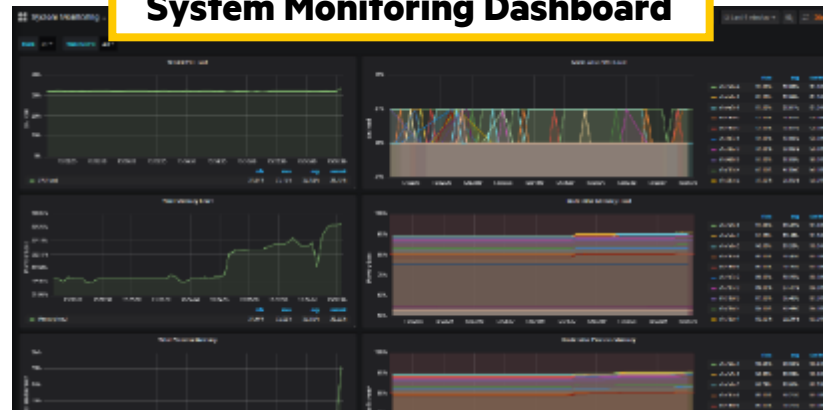
Main Cluster Health Dashboard



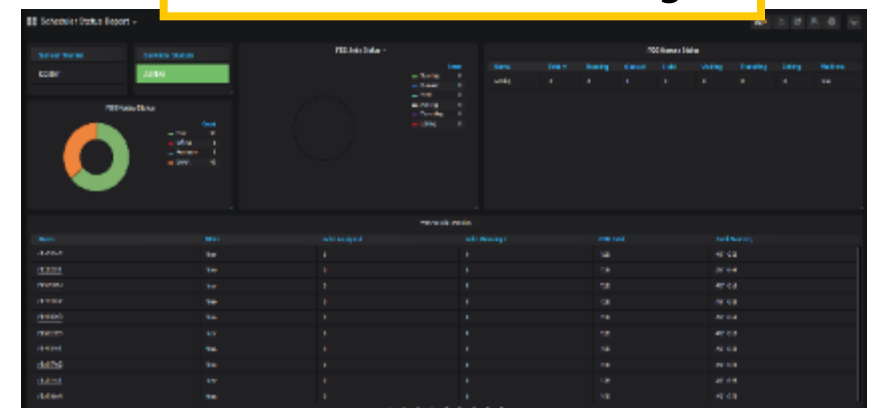
CDU Dashboard



System Monitoring Dashboard



Job Scheduler Monitoring



HPE Performance Cluster Manager Main Cluster Health Dashboard

At-a-glance cluster health overview and links to more detailed reports

Cluster Status Report

Alert numbers as well as information on nodes with severity and alerts status.

Fabric Health Report

Includes fabric port-state changes, link speed changes, and link degrade changes.

Cluster Health Report

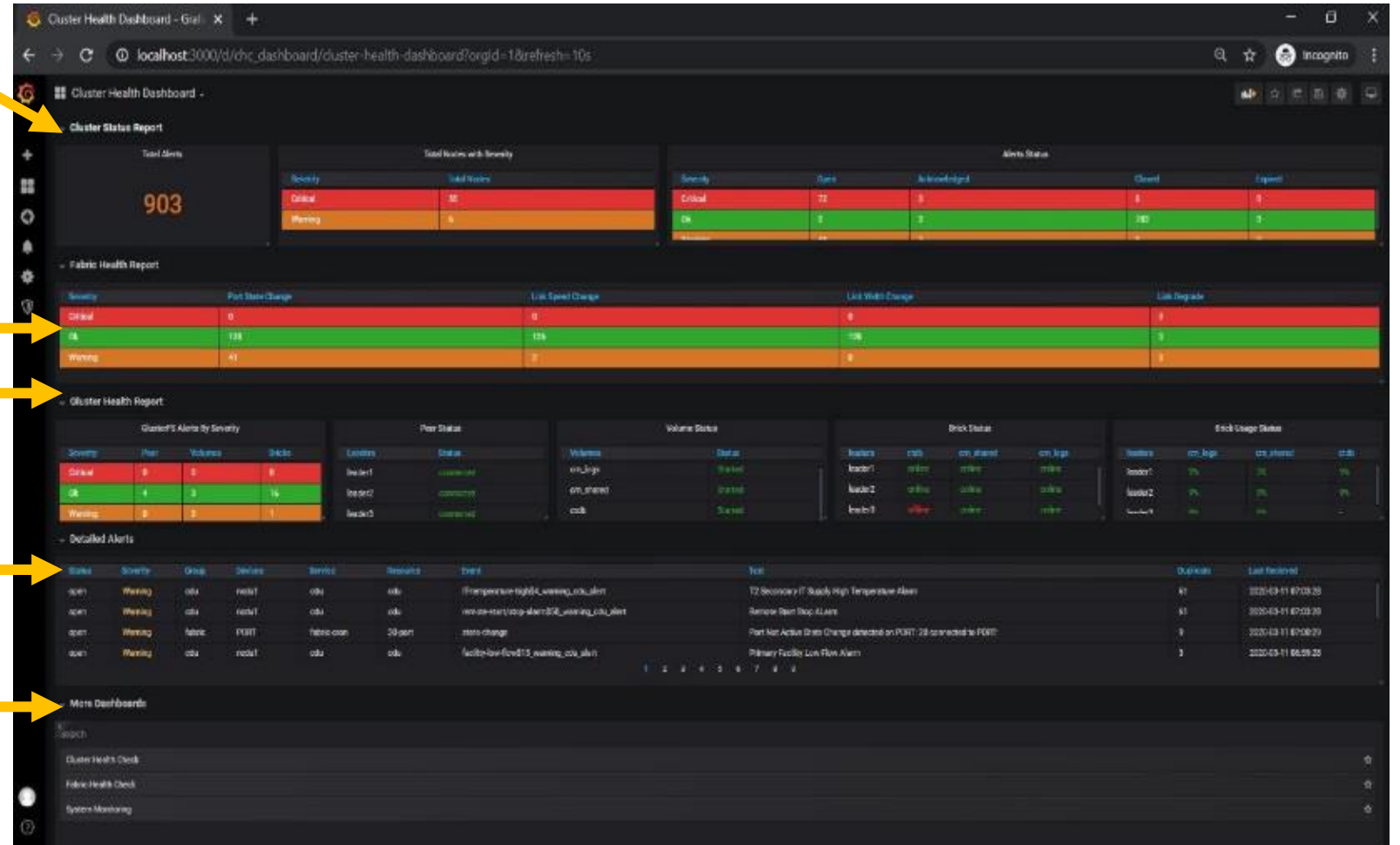
Includes severity levels and includes peer, volume, brick, and brick usage status.

Detailed Alerts

Complete details of the various alerts.

More Dashboards

Link to other dashboards like System Monitoring, Cluster Health Check, Fabric Health Check, CDU monitoring and Scheduler Status Report.



On-Demand Cluster Health Check

- Request health check anytime with data reported in a few seconds
- Hardware level diagnostics for node, chassis, rack, cpus, memory, disks, i/o, Lustre
- Fabric diagnostics
- Firmware/software versions, configuration settings

```
[root@iceadmin ~]# cm health -l
  cpuchk
  cpuperf
  cwcpuperf
  memchk
  memperf
  netchk
  fabricchk
  fabricperf
[root@iceadmin ~]#
```

```
Responses: 3 { r10i4n[0-2] }
Reference: <none>
Ignored:
[ ] <none>
```

Health Check Report

Health check
on selected
nodes

Report:

Analysis report of CPU Health Checks executed on nodes r10i4n[0-2]

1) CPU Configuration Checks

All nodes reported the CPU model name as (Intel(R) Xeon(R) CPU E5-2670 0 @ 2.60GHz) .
All nodes reported the CPU version as Sandy Bridge-E.
All nodes reported the CPU architecture as x86_64.

2) CPU Core Checks

All nodes reported the number of offline cores as 0.
All nodes reported the number of logical cores as 8.
All nodes reported the number of threads as 1.
All nodes reported the number of sockets as 2.

3) CPU Frequency Checks

All nodes reported the CPU frequency driver value as intel_pstate.
All nodes reported the CPU frequency governor value as performance.
Lowest frequency reported for a node: 2999.743
Highest frequency reported for a node: 2999.902
Average frequency reported for a node: 2999.849
All nodes completed the cpu frequency check **successfully**.

Health-aware Job Scheduling with HPE Performance Cluster Manager

Avoid launching jobs on unhealthy nodes

Reliability

- Prevent Job failures due to hardware and software problems

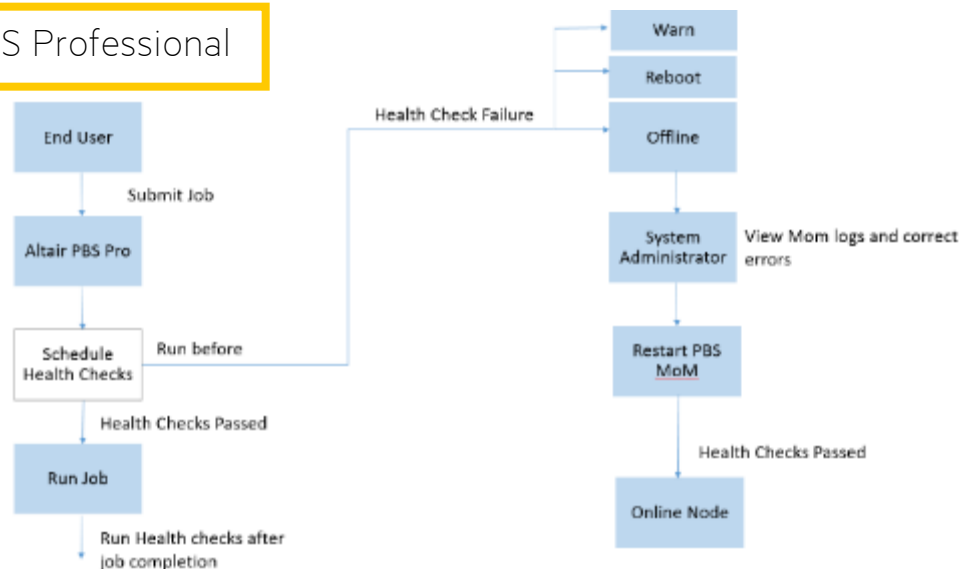
Reduced maintenance time

- Automated detection and notification of unhealthy nodes

Customizable runs

- Periodic health checks
- Prologue & Epilogue health checks

with PBS Professional



With Slurm

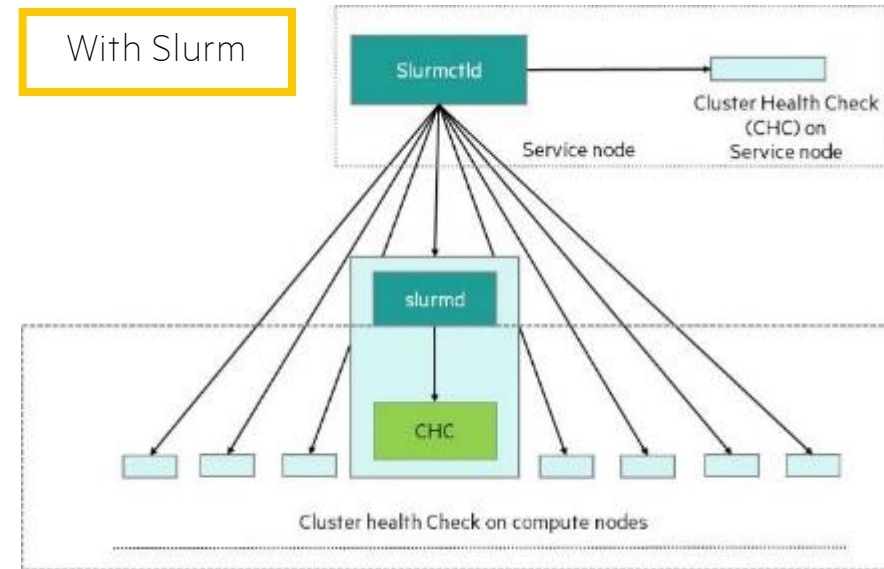


Image Management and Software Updates



Image Management and Software Updates

Run wide range of applications on the same cluster at the same time

- Provision a wide variety of applications on any number of nodes in minutes to keep up with users' demands
- Repurpose nodes on the fly to help minimize waiting times
- Update software and perform other maintenance tasks on idle nodes without interrupting running jobs on the rest of the cluster
- Administrators can query and analyze how well the compute resources are utilized—data, jobs, and users—to make the most of the available resources

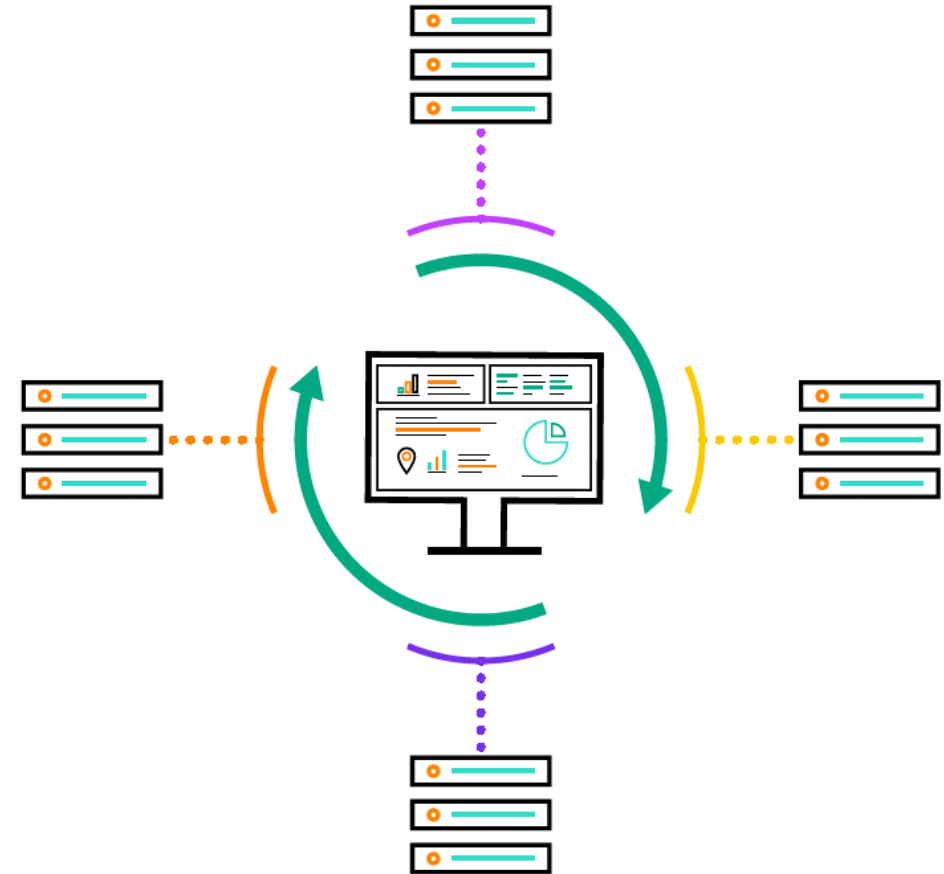


Image Management

Feature	Capability
Image formats	Software image repository supports RPM, ISO, remote repository (URL), Gold image (AutoYaST, Kickstart)
Software repository actions	Capture, backup and redeployment of images
Linux OS distributions	RHEL, Rocky Linux, SLES, CentOS and TOSS multiple versions can be managed and provisioned on all or select nodes
Accountability	Track changes to the archive—who made change and when change was made
Version control	Roll changes forward or backward as required and on-demand
Distributed imaging support	Master images on the admin node are also cached on leader nodes
Security	No secrets are embedded in the image



Secure Provisioning

Feature	Capability
Provisioning	Multi-cast and BitTorrent provisioning enables fast provisioning of software across the cluster in a single provisioning session
Cluster node support	Provisions diskfull and diskless (nfs/tmpfs) nodes
Security	Transfer of secret information (ssh, passwords, etc) occurs before provisioning. No image corruption during provisioning because image is encrypted with key. Compute node root ssh keys kept separate from admin nodes.
Automatic retries	Provisioning is not disrupted by network communication interruptions
System setup	No reboot required after provisioning
Mixed Linux OS distributions	Different OS images and kernels can be provisioned on selected nodes



Software Updates, Backup and Recovery

Feature	Capability
Software updates	Software repositories are used to provide updates Using distro tools (zypper, yum, dnf) to update rpms on images and running nodes
Backup images	Use revision control or clone to save the current images
Backup system	You can use the slot feature to backup the installation
Backup configuration files	Save the current HPCM database in config files that can be used for reinstallations or node rediscovery
Image recovery	Roll back to the previous image version
System recovery	Boot back the system to the previous slot
System reinstallation	Fresh install in another slot Use saved configuration files to quickly bring up the new slot Import the saved images to provision the nodes



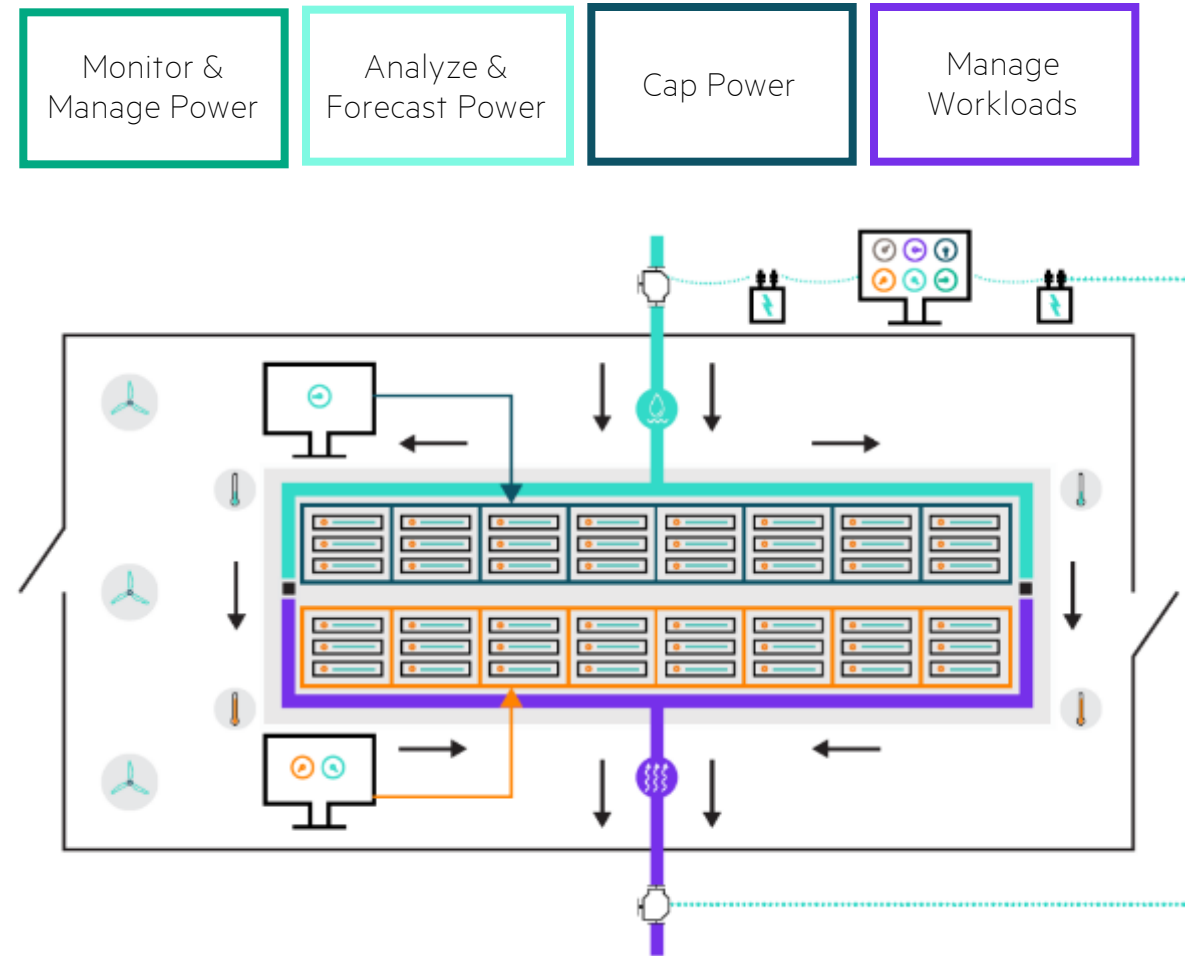
Power and Cooling



HPE Performance Cluster Manager Advanced Power and Cooling Management

More resilient systems and better datacenter economy

- Step-by-step topology and protocol-aware Power On/Off feature enables controlled start and shut-down of the clustered system.
- Management of power and cooling hardware—CPU, GPU, rack, chassis, nodes, rack AC, bulk DC, liquid-cooling infrastructure such as HPE ARCS.
- AIOPs: ML-based real time anomaly detection for cooling hardware
- Power monitoring and analysis—Read and aggregate power and energy at available measurement points for analysis and planning.
- Contain power consumption and apply power per user, per group, and per job (interface with Slurm and Altair PBS Professional) and account for all consumed energy.
- Set limits to trigger a power cap based on environmental failure (power or thermal), data center power capacity, or for other reasons such as workloads, planned brownouts, time of day, and others.



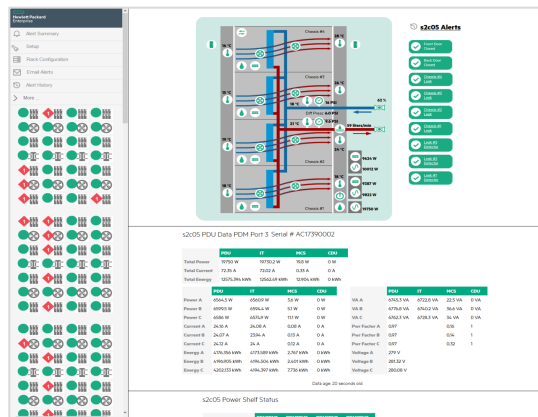
Cooling Device Monitoring in HPCM

The facility manager view

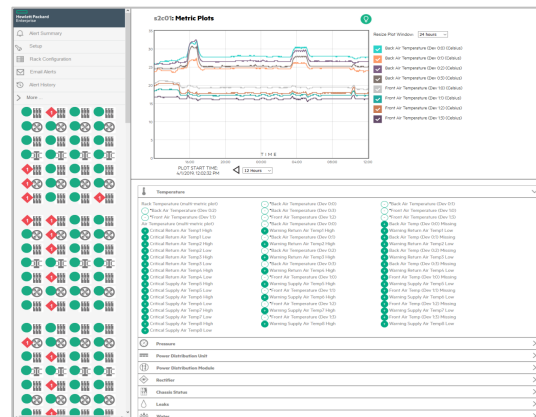
Cooling device & power device monitoring is integrated with HPE Performance Cluster Manager:

- Graphical display for HPE CDUs (i.e. HPE ARCS) with monitoring, alerts & trends. Same for PDUs
- Alerts for pre-configured events
 - Examples: water leakage, temperature changes, lock/unlock doors, power supply failure
 - Alerts can be selectively configured to send email to multiple groups
- Provides plotting for trending and historical analysis

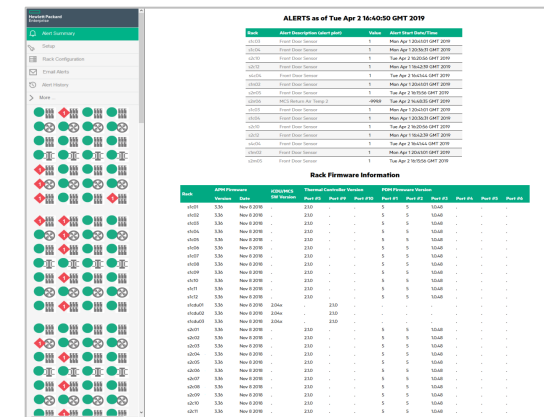
Main Interface



Metric Trends



Alert Summary



CDU Monitoring with HPE Performance Cluster Manager

The system administrator view

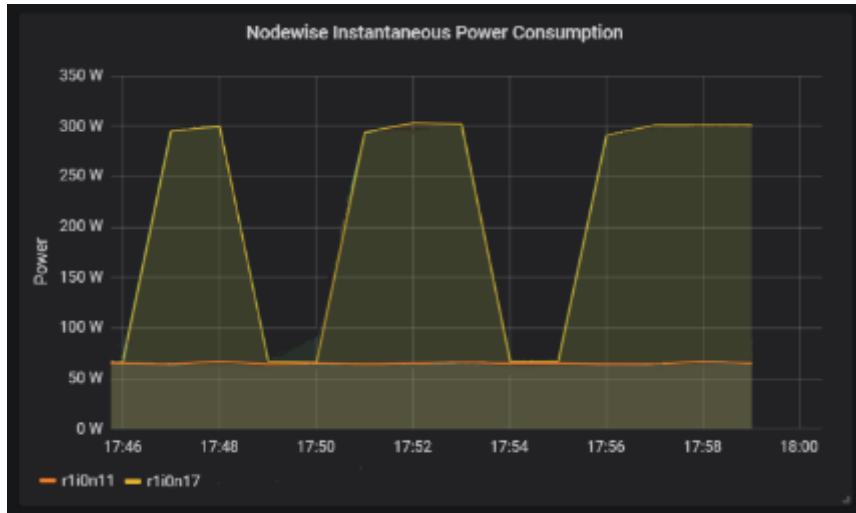
At-a-glance CDU monitoring dashboard is available in Grafana, displaying metrics such as:

- Temperatures of the room and the supplies + CDU
- Water flow I/Os
- Valve controls
- Humidity
- Dew point

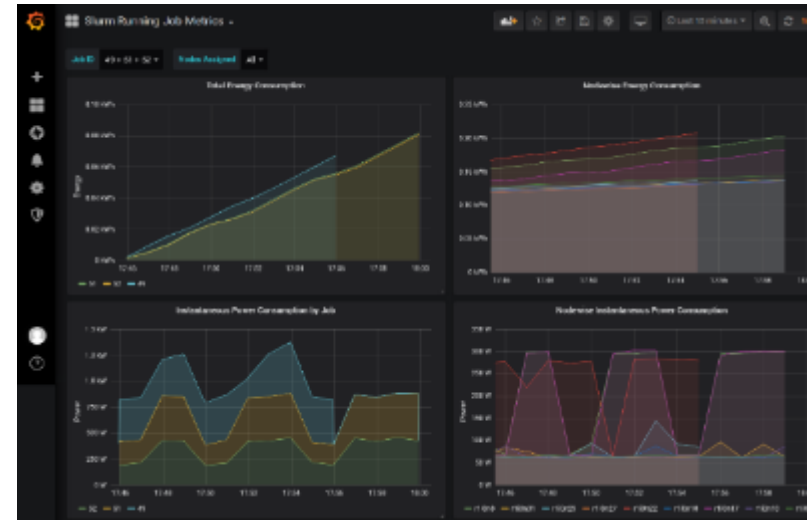
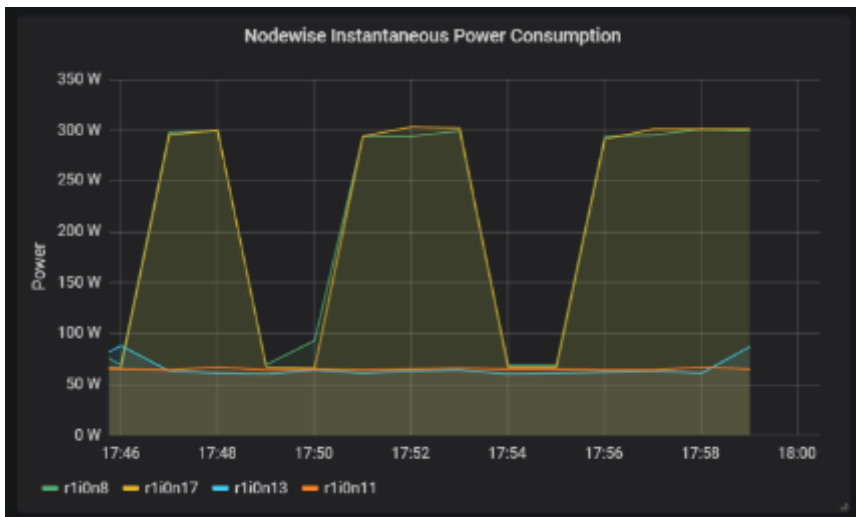


Power Monitoring for Jobs in Grafana – Integration with SLURM

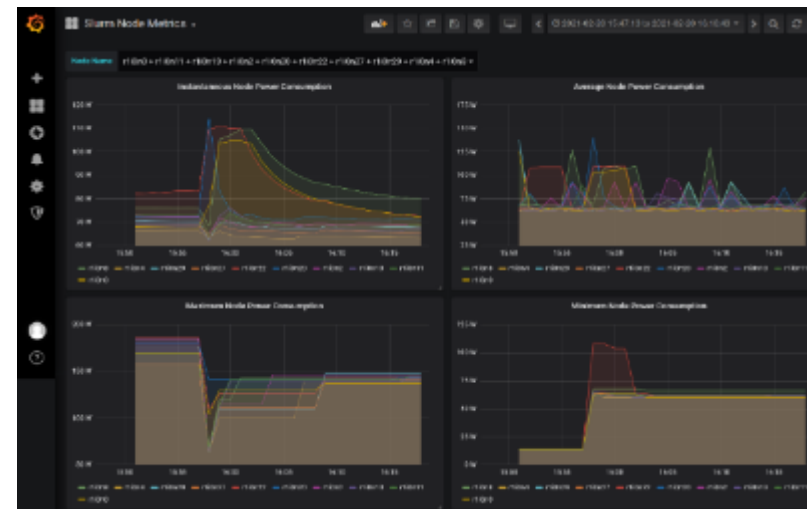
Showing power consumption of 2 nodes used by a single job. One of the nodes is power-capped at 100W.



Showing power consumption of 4 nodes used by two jobs. One of the nodes is power-capped at 100W.



Showing comparison of energy and power consumptions of different jobs selected.



Showing energy and power consumption of Slurm compute nodes.

AIOps in HPE Performance Cluster Manager

HPCM offers real-time anomaly detection on interface hardware such as CDUs and cooling racks using AIOps techniques



Traditional approach to monitoring

- Relying on thresholds to see issues.
- The bigger the system, more data administrators need to sift through and make sense of.
- Too many false alarms mean real issues are often overlooked.



AIOps



- Takes a predictive approach.
- AIOps uses machine learning and deep learning technologies to look for and report trends.

COMPONENTS OF THE HPCM ANOMALY DETECTION FRAMEWORK

AIOps Runtime Framework

Runs anomaly detection models and other streaming processors in real time.

AIOps Pipeline

Injects data from all HPC cluster components into AIOps workspace.

AIOps ML

Implementations of various statistical, ML & DL models for anomaly detection + Supporting tools to train models and analyze results.

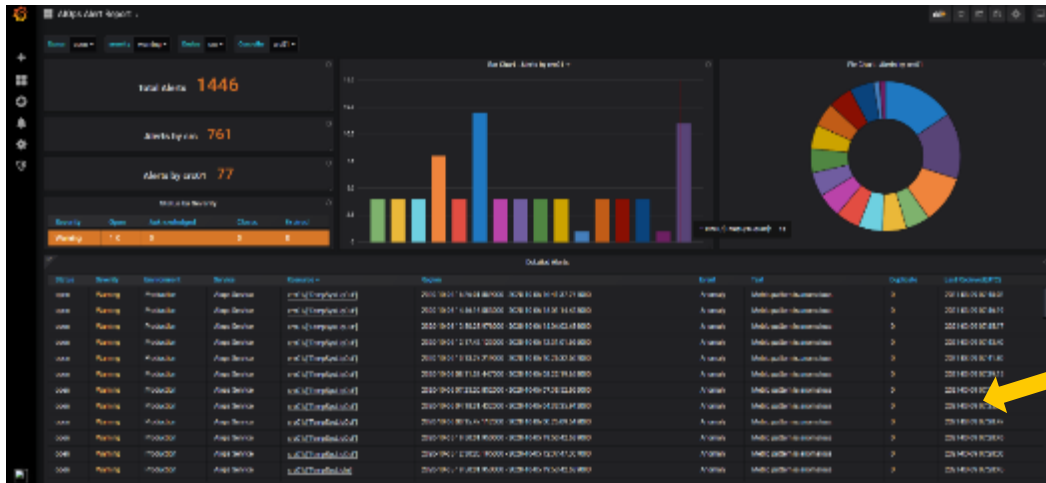
AIOps Visualization

Tools to visualize data and outcomes of machine learning algorithms.

Grafana Alerta Kibana CLI

AIOps in HPE Performance Cluster Manager — Dashboards in Grafana

More insight = better chance to uncover issues before they turn into failures



The **AIOps Alert Report** dashboard displays notifications of anomalies for cooling hardware. This dashboard also displays a pie chart showing where in the system alerts come from.



AIOps Single Metric dashboard contains plots of metric data values (**blue line**), anomaly scores for the monitored metric (**red line**) for a single metric (CDU valve position in this case). An alert is generated (and displayed on the dashboard) when the anomaly score exceeds the anomaly threshold (**yellow line**). The alert expires in the event that additional alerts are not generated during a predefined period of time.



AIOps Multimetric dashboard will display the anomaly a raw metric plot (one that is selected by the user), along with an anomaly score and its threshold for a correlated group of metrics.

ISV and Open-Source Software Integration



ISV and Open-Source Software Integration

	Software	Capability	Use Cases
Workload Management	Altair PBS Professional	<p>Scheduler monitoring in Grafana</p> <p>Track selected nodes' metrics by job</p> <p>Save metrics/job for future resource analyses</p> <p>Access to PBS Professional commands from HPCM menus</p> <p>Altair required software: www.pbsworks.com/hpcm-pbspro-connector</p>	<p>Ongoing monitoring of jobs metrics available via Health check Grafana interface</p> <p>Check on system health before job is scheduled</p> <p>Find CPUload for the job & running time for the job</p> <p>Identify which job is most power efficient</p> <p>Provision OS image of choice on a per-job basis</p> <p>Experiment with adjusting resources</p>
	Slurm	<p>Scheduler monitoring in Grafana</p> <p>Track selected nodes' metrics by job</p> <p>Save metrics/job for future resource analyses</p> <p>Power/energy monitoring capabilities and node level power capping</p> <p>Access to Slurm commands from HPCM menus—Integrated with HPCM via Slurm connector (SPANK)</p>	<p>Ongoing monitoring of jobs metrics available via Health check Grafana interface</p> <p>Check on system health before job is scheduled</p> <p>Find CPUload for the job & running time for the job</p> <p>Identify which job is most power efficient</p> <p>Suspend power for idle nodes</p> <p>Experiment with adjusting resources</p>
Automation	Ansible Playbook	<p>HPCM outputs hardware inventory metrics in Ansible Playbook format</p> <p>Feature is integrated in HPCM</p>	Hardware metrics are ready to be used by Ansible



ISV and Open Source Software Integration

	Software	Capability	Use Cases
Alerts	Alerta	Alerts monitoring system	Collect system-wide alerts
Monitoring	LDMS	Lightweight Distributed Metric Service for run-time resource utilization monitoring	Continuous collection, transport and storage of telemetry data
	Prometheus	Open Source software for hardware monitoring, alerts Integrated in HPCM to monitor for alerts	Checks for services that are running Check for temperature on switches
Repository, Search, Analysis	Elastic Stack	Enables search, analysis and visualization of data in real time	Debug issues by reviewing log data in one UI
	TimescaleDB	Time series database for telemetry data	Efficient storage of telemetry data for analysis and debugging
Network	NVIDIA UFM	View InfiniBand switch metrics with the associated cluster node metrics Use UFM commands from HPCM	Provide more detailed data when reporting InfiniBand issues to HPE Pointnext: "The problem with the switch is on node 15"
GPU	NVIDIA DCGM	NVIDIA GPU monitoring and management Integrated in HPCM	Multi-GPU health management across the cluster



Integration with PBS Professional

- Altair PBS Professional integration into HPCM enables job efficient scheduling
- Access PBS Professional commands from within HPCM
- Track selected nodes' metrics by jobs
- Key functionalities:
 - Enable provisioning of OS image of choice on a per-job basis
 - Delete jobs
 - Suspend jobs
 - Resume jobs
 - Manage suspended jobs
 - Identify job power usage

The screenshot displays the HPCM Performance Cluster Manager interface. The left sidebar shows a tree view of resources, including 'Unassigned nodes' and a list of nodes (n0, n1, n2, n3). A context menu is open over node 'n0', listing various actions such as 'Shutdown (using SSH)', 'Power Off (using BMC)', 'Boot/Reboot (using BMC)', 'Reboot (using SSH)', 'Change UID LED Status', 'PDSH (using HPCM dmtf)', 'PDCP (Distributed Copy)', 'Capture Image', 'Provision Image (Deploy)', 'Autoinstall (Kickstart | AutoYaST | Preseed | Unattended)', 'Refresh', 'Firmware and BIOS', 'SLURM', 'HPCM PBS Professional Connector', 'Set Display Color', 'Copy to Clipboard', and 'Manage'. The 'HPCM PBS Professional Connector' option is highlighted in orange. A yellow box labeled 'PBS Pro commands' points to this menu. The main window shows a 'Monitoring' tab for node 'n0' with a table of metrics:

Metric	n0	Unit
kernel_version	3.10.0-893.21.1.el7.x86_64	release
cpu_load	0	%
memory_used	1	%
	1	%
	3	%
	0	%
	80	day(s)
	No reference value	image
	Inactive action	node_state
	0	job(s)
	Up	Status
	0	MB/s
	0	MB/s
	0	MB/s
	0	MB/s
	0	login(s)
	2047	MB
	0	%
	3200	MHz
	0	mount(s)

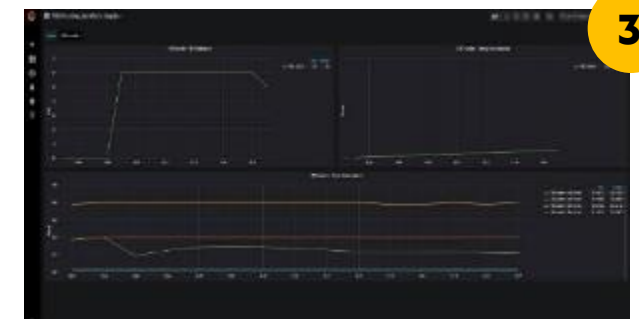
A yellow box at the bottom right of the screenshot contains the text: 'Same integration available for SLURM'.

Scheduler Monitoring in Grafana

(1) Scheduler Status Report Dashboard contains the below items:

- **Server Name:** on which server, PBS is running.
- **Service Status:** shows the service state (active/inactive).
- **PBS Nodes Status:** shows node availability details.
- **PBS Jobs Status:** shows jobs status count.
- **PBS Queue status:** queue details.
- **PBS Vnode Metrics:** node specific metrics report.
 - **(2)** Click on a node name to get node specific jobs history table, CPU free count and free memory graphs.
- **PBS Running Jobs Metrics:** Job level details.
 - **(3)** Click on a job id, it navigates to a new tab which shows job level CPU utilization, energy and power consumption metrics graphs.

• Same dashboards available for SLURM



1

2

3

HPE Performance Cluster Manager

Features At-a-Glance



System Setup

- Provision thousands of cluster nodes in less than an hour, scale cluster without disruption
- Hardware is automatically discovered and configured
- Configuration information stored in a secure repository
- Easily reference nodes by creating aliases for hosts
- Simplified Quorum HA setup

Image Management & Software Update

- Secure software image repository supports Linux OS, middleware and applications
- Image formats: RPM, ISO, gold image
- Any software image can be provisioned on all or select cluster nodes
- Accountability built-in to track changes

Platforms supported:

- HPE Apollo systems
- HPE ProLiant DL
- HPE Superdome Flex
- HPE Cray EX systems

Hardware Monitoring and Management

- Granular Monitoring of Service Infrastructure by Prometheus
- Realtime job-level monitoring of HPE Slingshot fabric
- Slingshot fabric critical alerts reporting
- Redfish event collection from switch controllers
- Collect and storage of telemetry data in a secure repository
- View metrics and alerts via GUI, CLI, Ganglia, Prometheus, Kibana, Grafana
- Setup pre-defined and customizable alerts for events
- Broadcast commands to selected nodes
- FRU inventory for serviceability
- Integrated firmware flashing for BIOS, BMC/iLO, CMC, network adapter/switch
- Central management is protected with secure admin node

Power Monitoring and Management

- Collect and analyze power metrics (for nodes incl. GPU enabled, chassis, rack, PDU, CDU)
- AIOps—multi-metric real-time anomaly detection for cooling units, CPU, GPU, DIMMS and Node Power
- Topology-aware Power On/Off
- Power capping
- Assign/account power resources for jobs
- Job-level power management API

Cluster Health Management

- Validate system diagnostics run in HPE factory on new system after installation at customer site
- Real-time management of alerts data across cluster
- Invasive tests to be used during maintenance window or to diagnose root cause of failures
- Custom cluster health dashboards & alerts in Grafana
- Workload scheduler integration—system health check before running job and real-time monitoring of jobs
- Realtime health-check of HPE Slingshot fabric
- Cluster Configuration Verification Tool (CVT)
- GOUC diagnostics for AMD GPUs and Intel Exascale compute blade

ISV & Open-Source Software Integration

- Altair PBS Professional, Slurm: Launch tool, manage resources, use selected tool commands from HPCM
- NVIDIA UFM: View IB switch metrics with associated cluster node metrics
- Ganglia, Nagios core, Elastic Stack, Alerta Integrated for system monitoring, alerts and analytics
- DCGM connector: Integrated management of NVIDIA GPUs
- Ansible: HPCM stores hardware metrics in Ansible Playbook format

Thank you

Jan Wender

jan.wender@hpe.com

+49 152 0880 1771



a00047447enw