Bright Cluster Manager & SLURM

Maximum Productivity, Minimum Complexity

Robert Stober
Systems Engineer

Bright Computing

1. Develops and supports Bright Cluster Manager for HPC systems and server farms
2. Incorporated in USA (HQ in San Jose, California)
3. Backed by ING Bank as shareholder and investor
4. Sells through a rapidly growing network of resellers and OEMs world-wide
5. Customers and resellers in US, Canada, Brazil, Europe, Middle-East, India, Singapore, Japan, China
6. Installations in Academia, Government, Industry, ranging from 4 node to TOP500 systems

Customers

Academia

- The City University of New York
- University of Houston
- Drexel University
- Texas A&M
- 東京工業大学 Tokyo Institute of Technology
- Virginia Tech
- 東京大学 The University of Tokyo
- University of Oxford
- University of Cambridge
- Goethe Universität Frankfurt am Main
- UCL
- EPFL École Polytechnique Fédérale de Lausanne
- Universität Stuttgart
- King's College London
- Technische Universität Dresden
- university of groningen
- technische universität dortmund
- Universität Duisburg Essen

Government

- Fermilab
- NERSC
- Sandia National Laboratories
- TACC
- Proudman Oceanographic Laboratory, Natural Environment Research Council
- Department of Defense, United States of America
- SINTEF
- cea
- TNO
- CNRS
- CCLRC
- INRIA
- UZ Leuven
- ESIEE Amiens
- De Nayer Instituut, Sint-Katelijne-Waver
- Max-Planck-Gesellschaft
- CERN
- CASPUR

Industry

- Boeing
- ارامكو السعودية Saudi Aramco
- Rock Solid Images, a member of the OHM Group
- Schering-Plough
- PACCAR
- ING
- Regeneron
- GL Operating 24/7
- Masdar, Abu Dhabi Future Energy Company, شركة أبوظبي لطاقة المستقبل
- DNV
- CD-adapco
- Genus
- Vattenfall
- dice Program
- Atlas Copco
- DELPHI
- zitron
- vonRoll
- COMPONENTA
- BMT ARGOSS

The Commonly Used "Toolkit" Approach

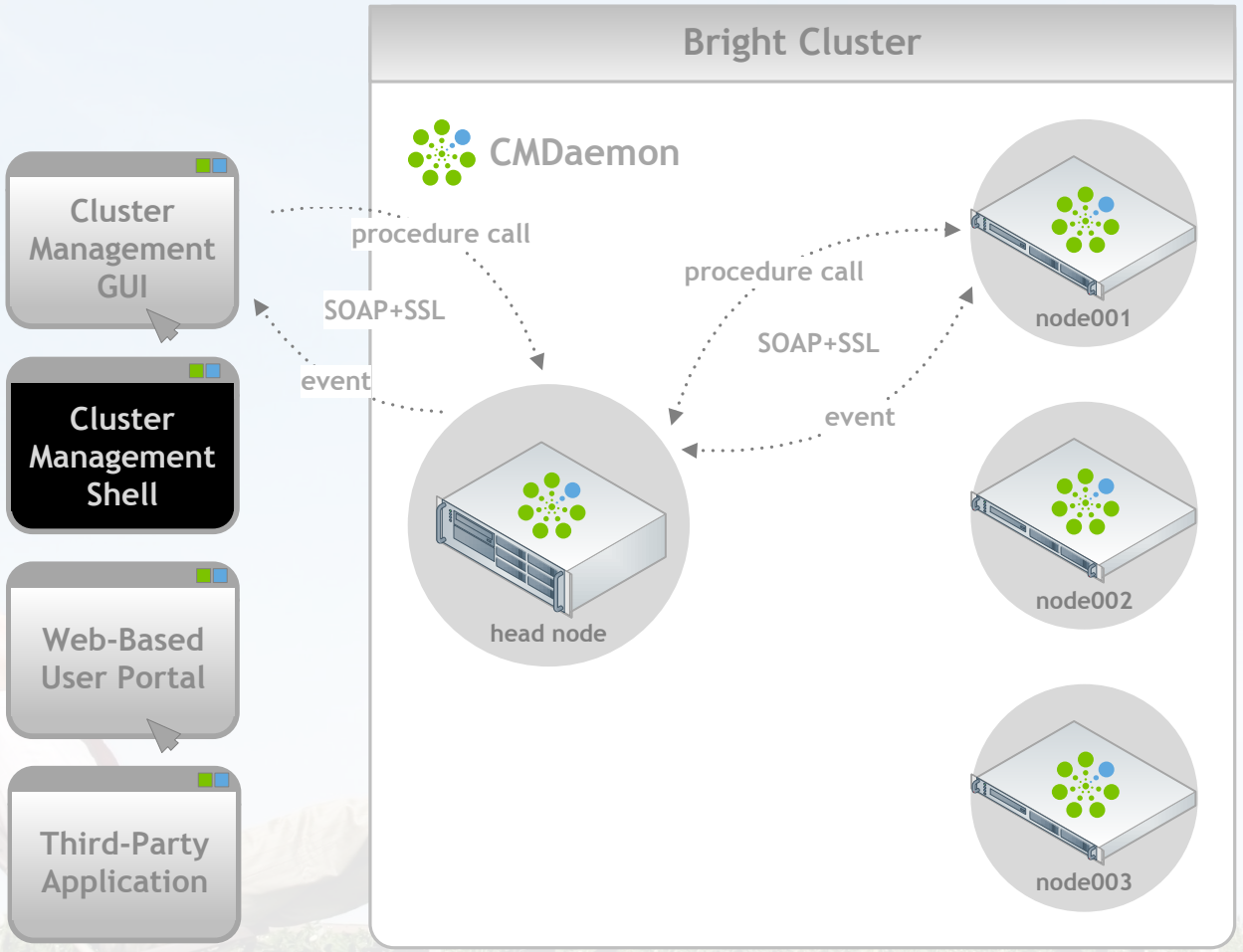
- Most HPC cluster management solutions use the "toolkit" approach (Linux distro + tools)
  - Examples: Rocks, PCM, OSCAR, UniCluster, CMU, etc.
  - Tools typically used: Ganglia, Cacti, Nagios, Cfengine, System Imager, xCAT, Puppet, Cobbler, Hobbit, Big Brother, Zabbix, Groundwork, etc.
- Issues with the "toolkit" approach:
  - Tools rarely designed to work together
  - Tools rarely designed for HPC
  - Tools rarely designed to scale
  - Each tool has its own command line interface and GUI
  - Each tool has its own daemon and database
  - Roadmap dependent on developers of the tools
- Making a collection of unrelated tools work together
  - Requires a lot of expertise and scripting
  - Rarely leads to a really easy-to-use and scalable solution

About Bright Cluster Manager

- Bright Cluster Manager takes a much more fundamental & integrated approach
  - Designed and written from the ground up
  - Single cluster management daemon provides all functionality
  - Single, central database for configuration and monitoring data
  - Single CLI and GUI for ALL cluster management functionality

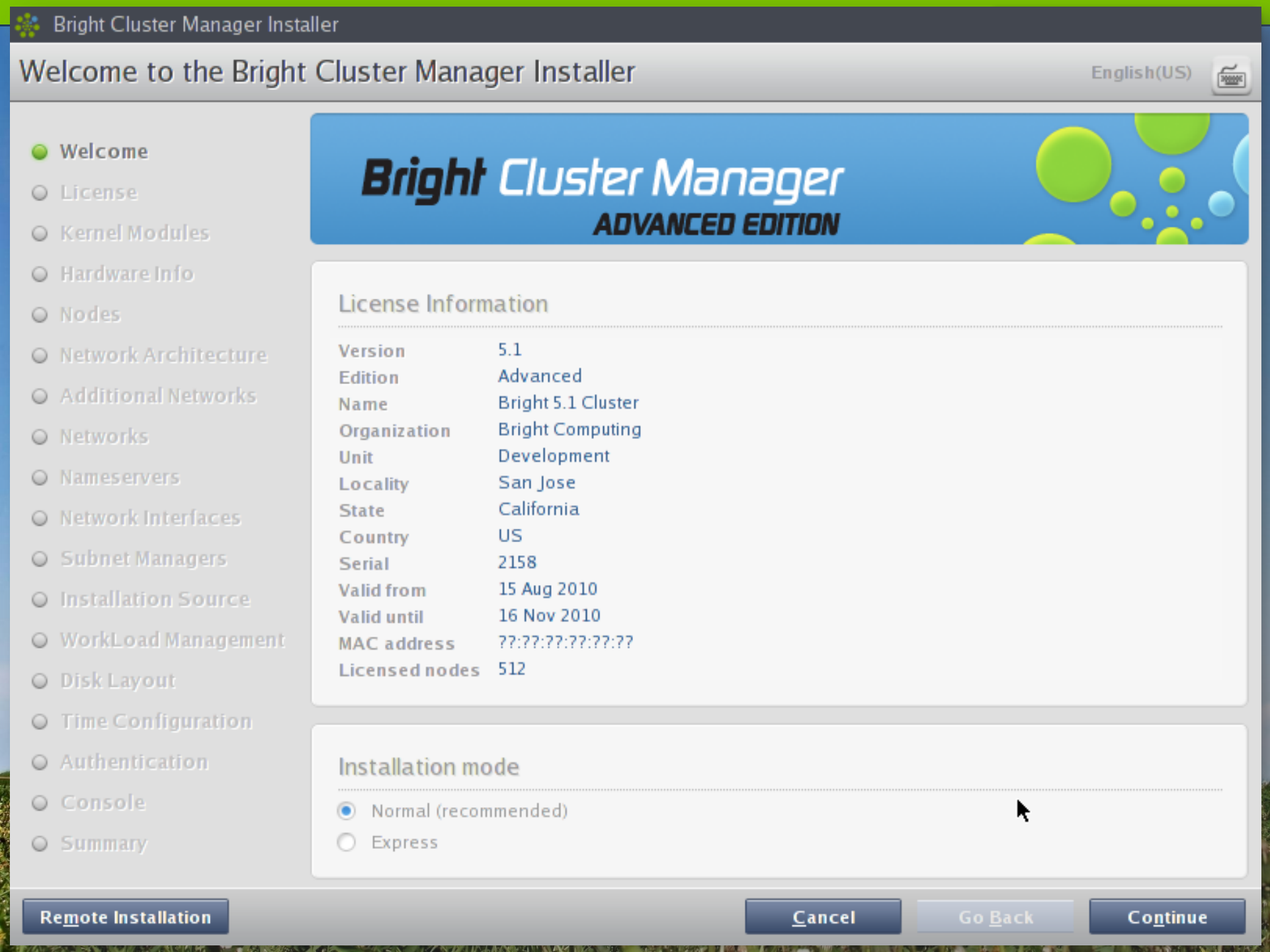
- Which makes Bright Cluster Manager …
  - Extremely easy to use
  - Extremely scalable
  - Secure & reliable
  - Complete
  - Flexible
  - Maintainable

Architecture



Workload Manager Integration

- Bright supports several workload managers including SLURM
- SLURM is now the default
- Automatic installation
- Automatic configuration
- Pre-job health checks
- Sampling and visualization of workload manager metrics
- Job monitoring and control
- Creation and editing of queues
- Integration of Bright and SLURM failover


Bright Cluster Manager Installer
Welcome to the Bright Cluster Manager Installer
English(US)
Welcome
License
Kernel Modules
Hardware Info
Nodes
Network Architecture
Additional Networks
Networks
Nameservers
Network Interfaces
Subnet Managers
Installation Source
WorkLoad Management
Disk Layout
Time Configuration
Authentication
Console
Summary
Bright Cluster Manager
ADVANCED EDITION
License Information
Version 5.1
Edition Advanced
Name Bright 5.1 Cluster
Organization Bright Computing
Unit Development
Locality San Jose
State California
Country US
Serial 2158
Valid from 15 Aug 2010
Valid until 16 Nov 2010
MAC address ??:??:??:??:??:??
Licensed nodes 512
Installation mode
Normal (recommended)
Express
Remote Installation
Cancel
Go Back
Continue

SLURM Installation

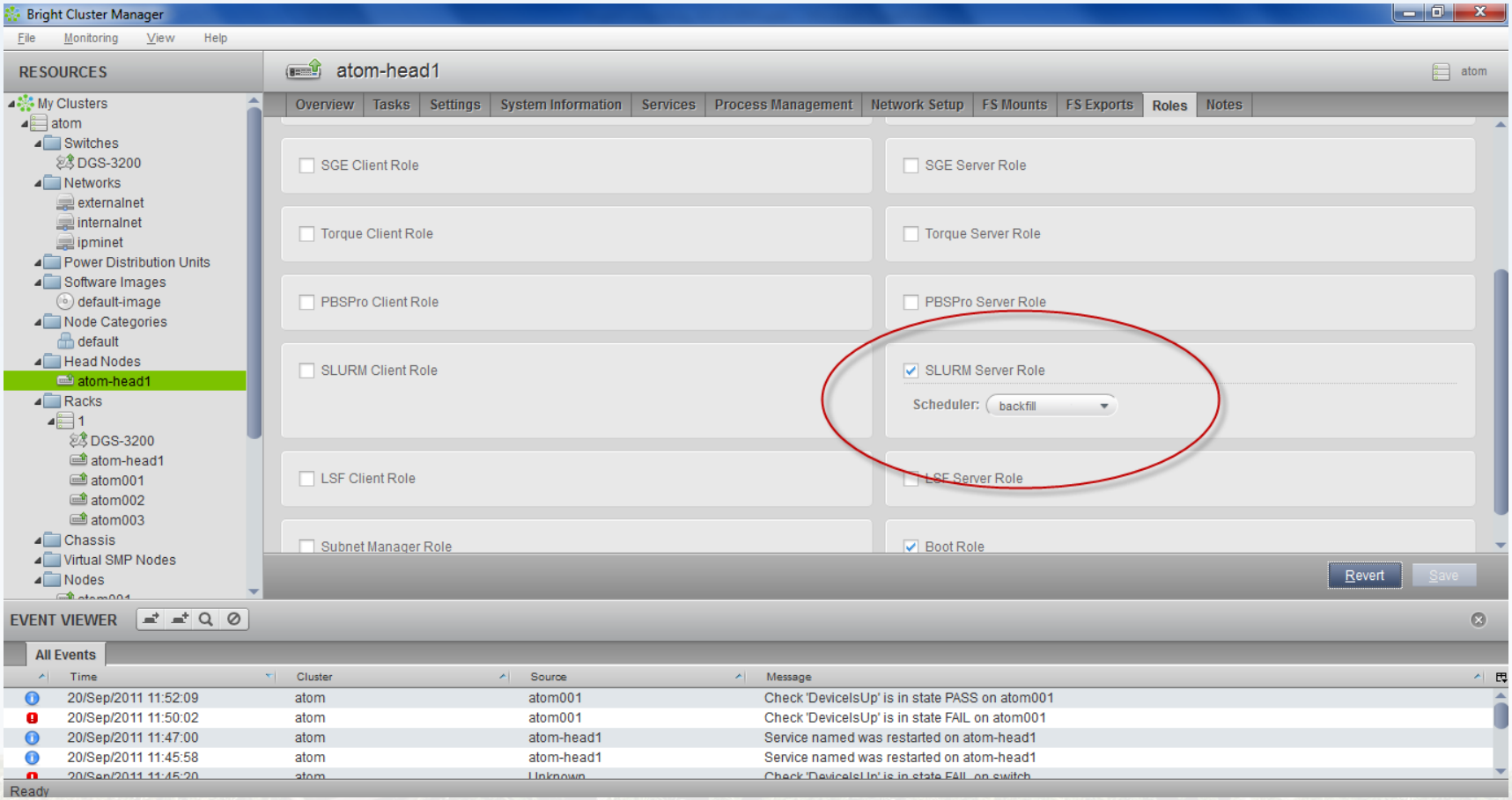
SLURM is installed in a shared directory

- Current version is 2.2.4
- /cm/shared is mounted on the nodes by default

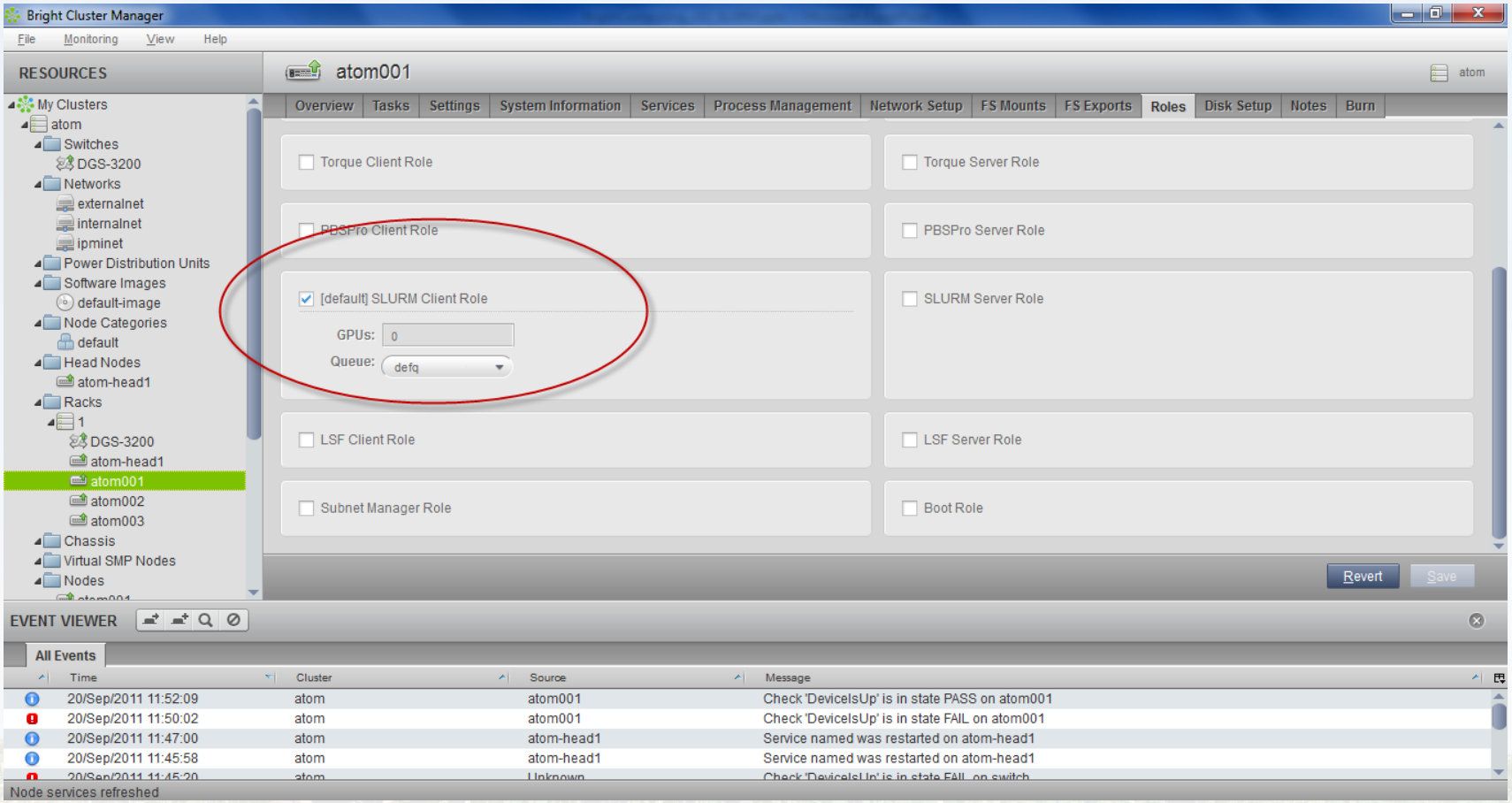
```
[root@atom-head1 apps]# cd /cm/shared/apps/slurm/current

[root@atom-head1 current]# ls -l
total 28
drwxr-xr-x 2 root root 4096 Sep  2 18:32 bin
drwxr-xr-x 4 root root 4096 Sep  3 02:10 cm
drwxr-xr-x 2 root root 4096 Sep  3 02:10 etc
drwxr-xr-x 3 root root 4096 Sep  2 18:32 lib64
drwxr-xr-x 5 root root 4096 Sep  2 18:32 man
drwxr-xr-x 2 root root 4096 Sep  2 18:34 sbin
drwxr-xr-x 4 root root 4096 Sep  2 18:32 share
```

SLURM Server Role



SLURM Client Role



SLURM Installation

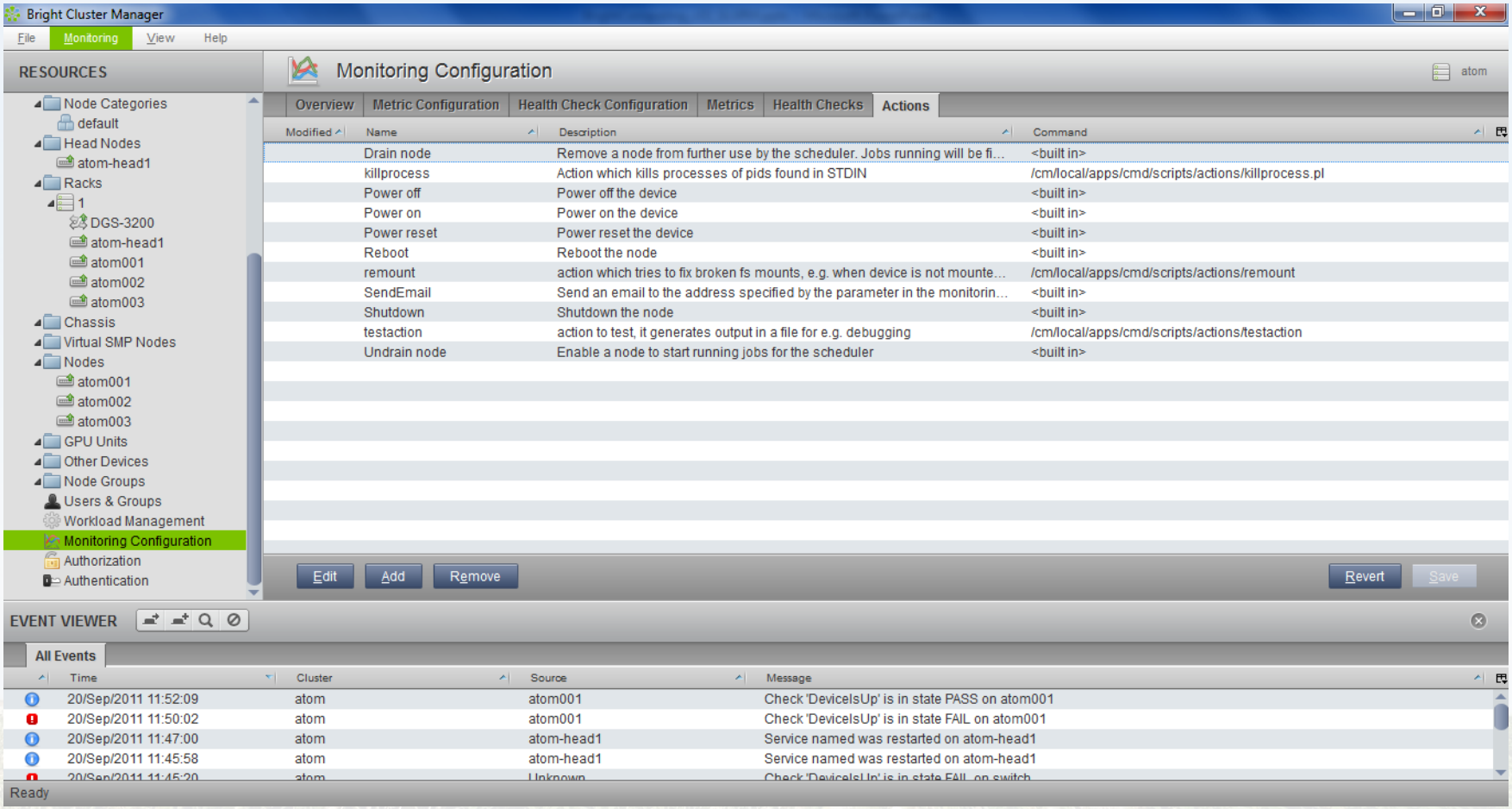
Portions of the SLURM config files are autogenerated

- Applies when a role has been assigned to a node
- Example: slurm.conf file

```
# BEGIN AUTOGENERATED SECTION -- DO NOT REMOVE
# Scheduler
SchedulerType=sched/backfill
# Master nodes
ControlMachine=atom-head1
ControlAddr=atom-head1
# Nodes
NodeName=atom[001-003]
# Partitions
PartitionName=defq Nodes=atom[001-003] Default=YES MinNodes=1
MaxNodes=UNLIMITED MaxTime=UNLIMITED AllowGroups=ALL Priority=1
DisableRootJobs=NO RootOnly=NO Hidden=NO Shared=NO
# END AUTOGENERATED SECTION   -- DO NOT REMOVE
```

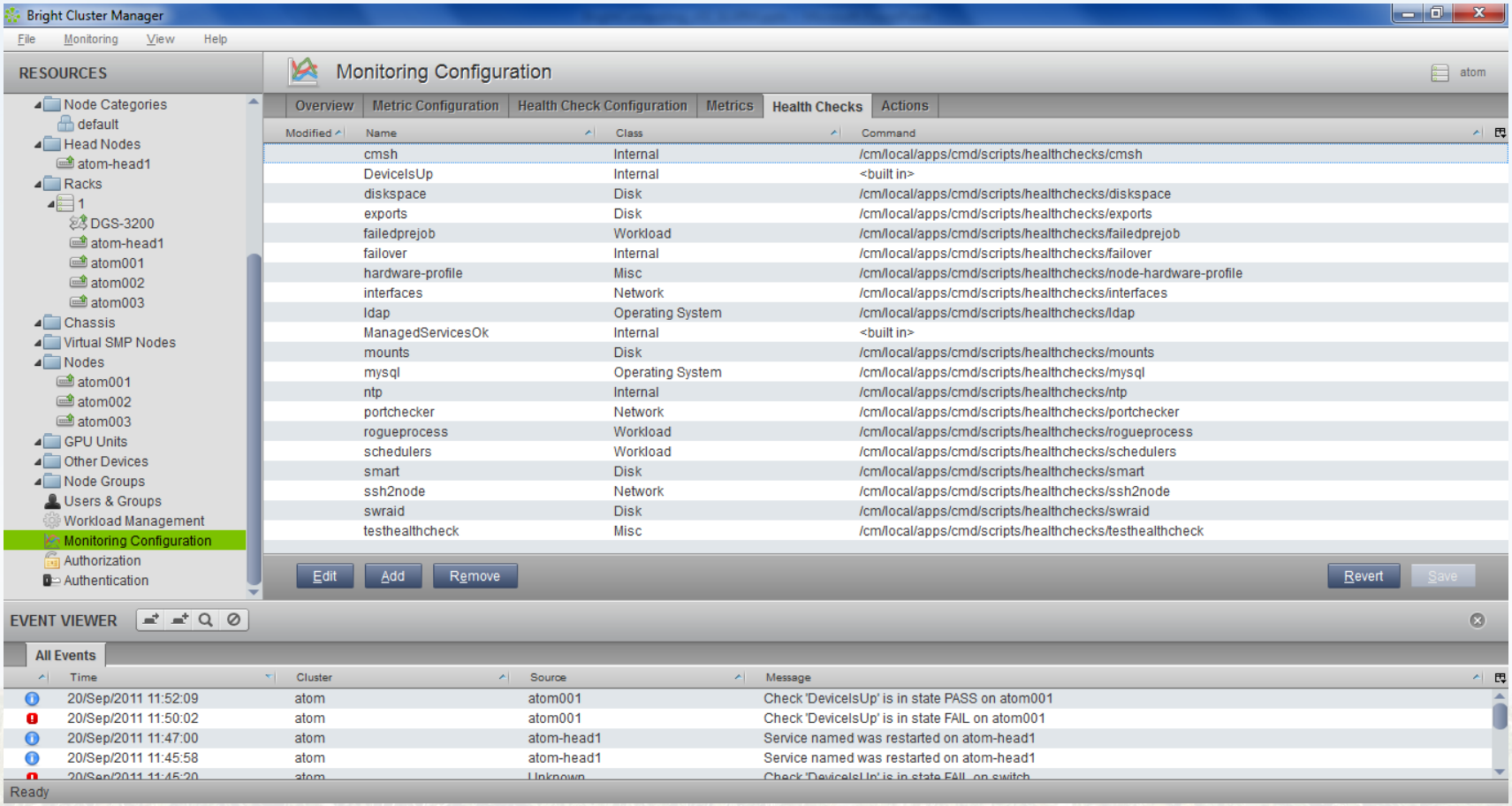
Bright Monitoring Framework

- **Actions**



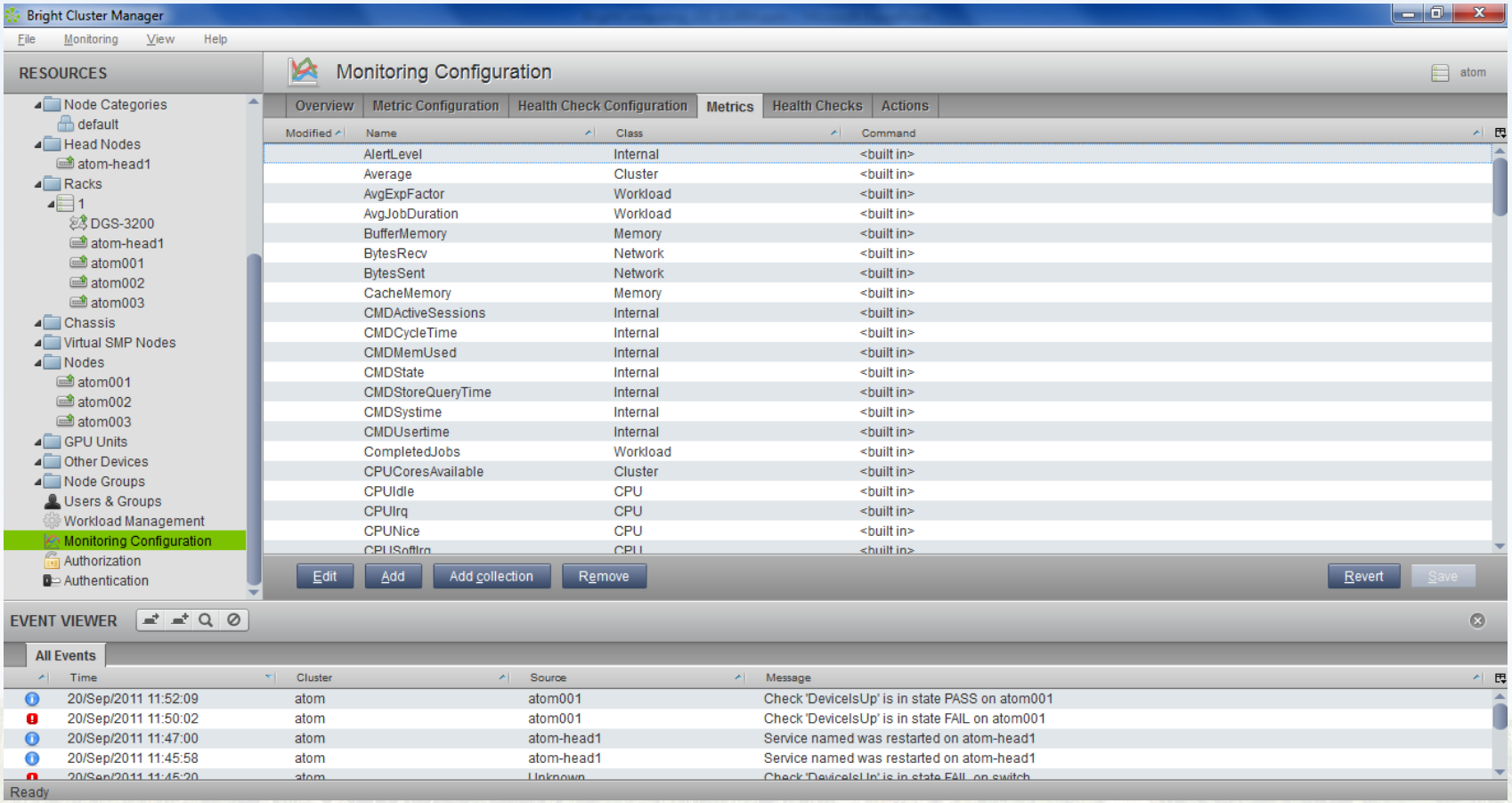
Bright Monitoring Framework

- **Health Checks**



Bright Monitoring Framework

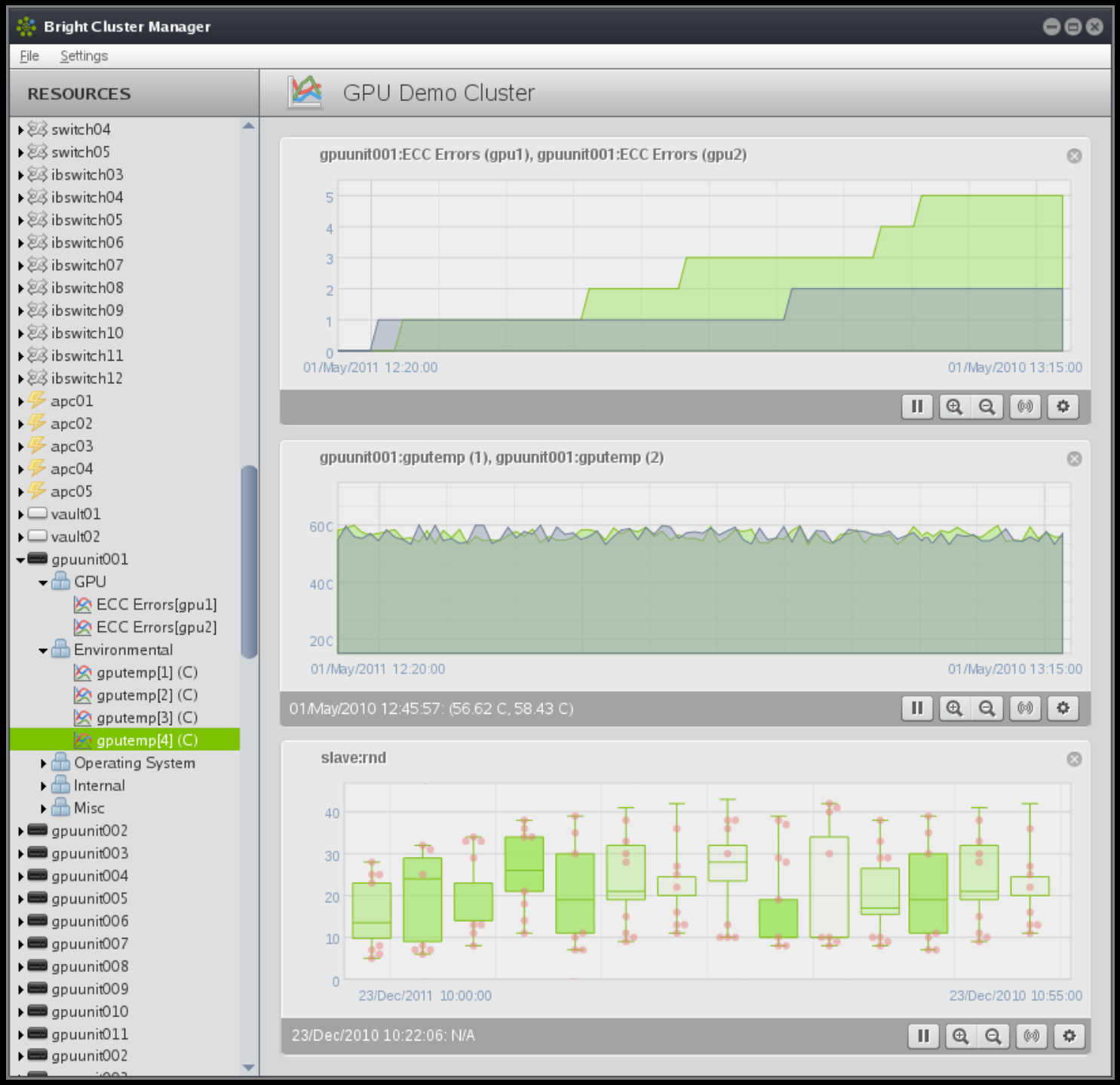
- **Metrics**

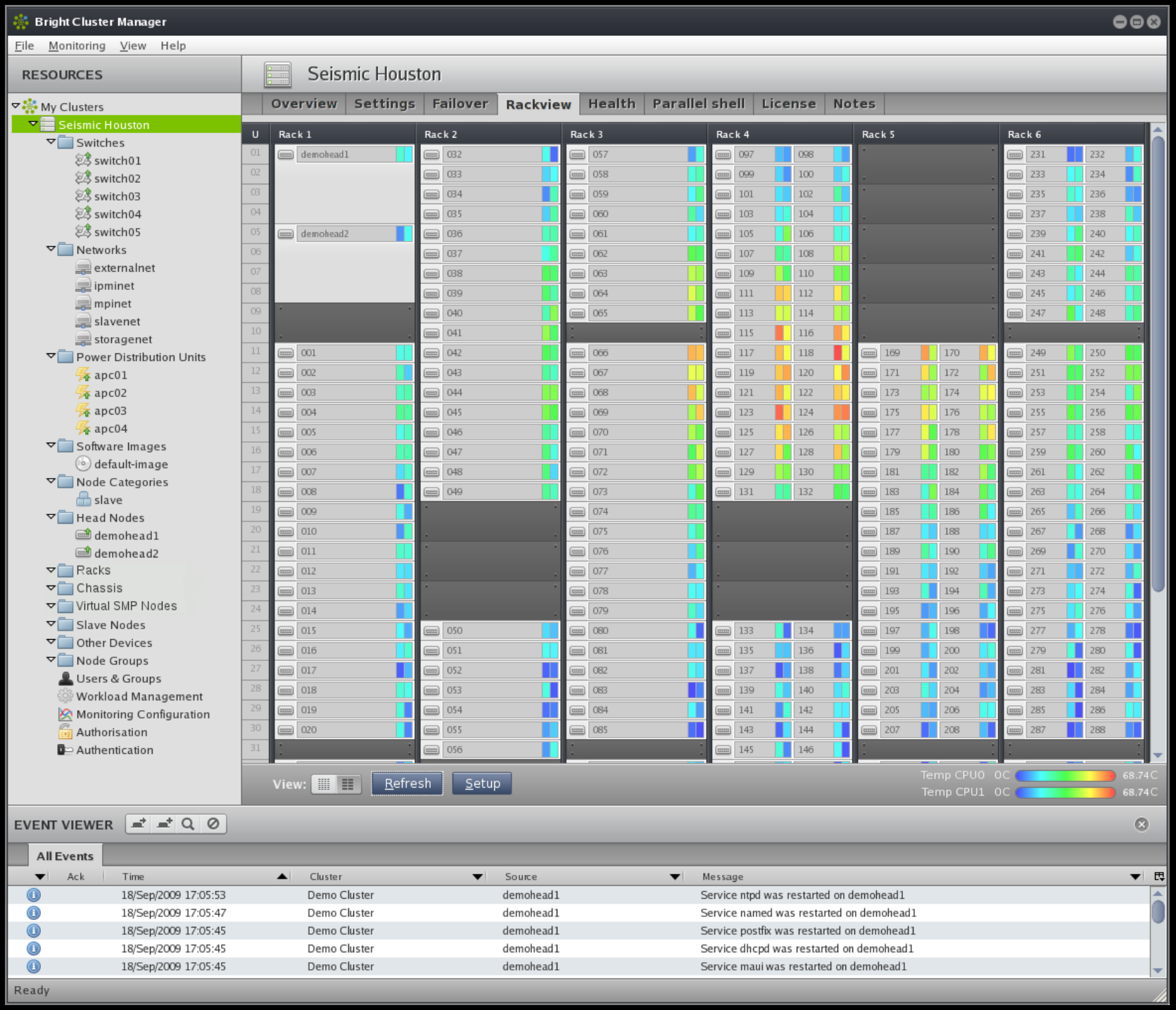


SLURM Metrics

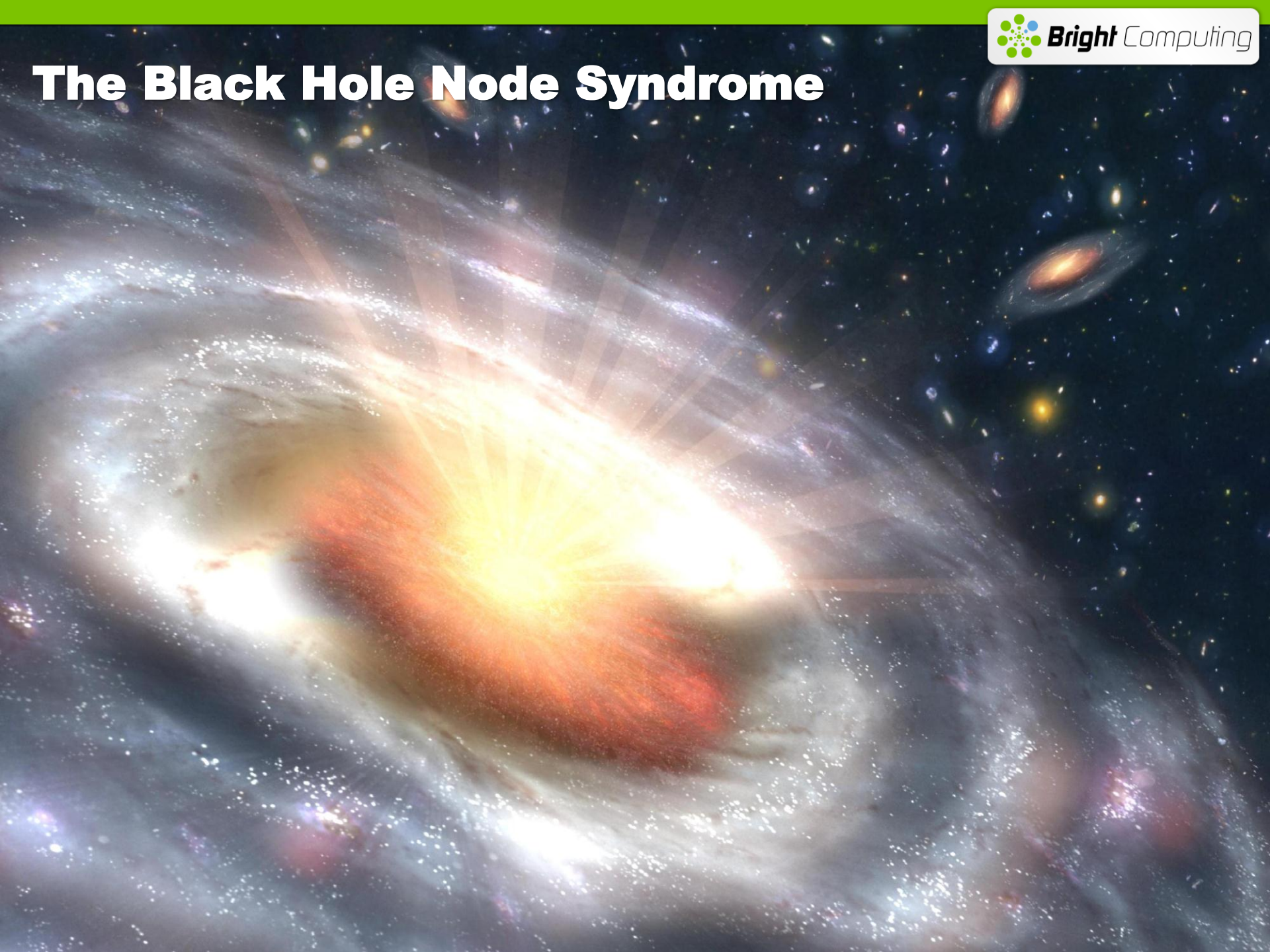
Bright collects the following basic workload manager metrics

- Running jobs
- Failed jobs
- Queued jobs
- Average expansion factor: This is by what factor, on average, jobs took longer to run than expected. The expectation is according to heuristics based on duration in past and current job queues, as well as node availability
- Estimated delay: Estimated Delay to execute jobs
- Average run time (per queue)


Bright Cluster Manager
File
Settings
RESOURCES
GPU Demo Cluster
switch04
switch05
ibswitch03
ibswitch04
ibswitch05
ibswitch06
ibswitch07
ibswitch08
ibswitch09
ibswitch10
ibswitch11
ibswitch12
apc01
apc02
apc03
apc04
apc05
vault01
vault02
gpuunit001
GPU
ECC Errors[gpu1]
ECC Errors[gpu2]
Environmental
gputemp[1] (C)
gputemp[2] (C)
gputemp[3] (C)
gputemp[4] (C)
Operating System
Internal
Misc
gpuunit002
gpuunit003
gpuunit004
gpuunit005
gpuunit006
gpuunit007
gpuunit008
gpuunit009
gpuunit010
gpuunit011
gpuunit002
gpuunit001:ECC Errors (gpu1), gpuunit001:ECC Errors (gpu2)
5
4
3
2
1
0
01/May/2011 12:20:00
01/May/2010 13:15:00
gpuunit001:gputemp (1), gpuunit001:gputemp (2)
60C
40C
20C
01/May/2011 12:20:00
01/May/2010 13:15:00
01/May/2010 12:45:57: (56.62 C, 58.43 C)
slave:rnd
40
30
20
10
0
23/Dec/2011 10:00:00
23/Dec/2010 10:55:00
23/Dec/2010 10:22:06: N/A

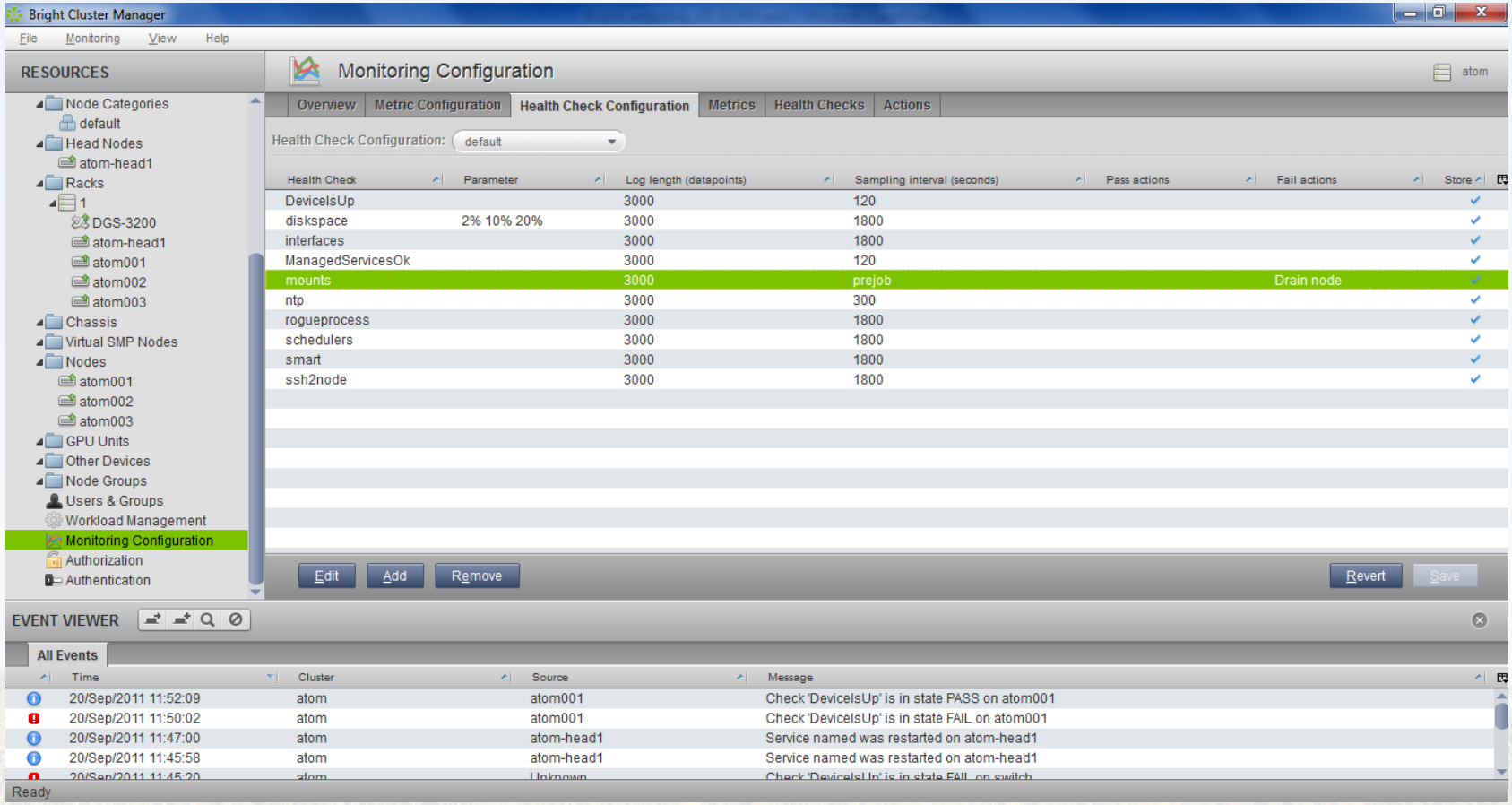

Bright Cluster Manager
File Monitoring View Help
RESOURCES
Seismic Houston
My Clusters
Seismic Houston
Switches
switch01
switch02
switch03
switch04
switch05
Networks
externalnet
ipminet
mpinet
slavenet
storagenet
Power Distribution Units
apc01
apc02
apc03
apc04
Software Images
default-image
Node Categories
slave
Head Nodes
demohead1
demohead2
Racks
Chassis
Virtual SMP Nodes
Slave Nodes
Other Devices
Node Groups
Users & Groups
Workload Management
Monitoring Configuration
Authorisation
Authentication
Overview Settings Failover Rackview Health Parallel shell License Notes
U
Rack 1
Rack 2
Rack 3
Rack 4
Rack 5
Rack 6
demohead1
demohead2
View:
Refresh
Setup
Temp CPU0 0C 68.74C
Temp CPU1 0C 68.74C
EVENT VIEWER
All Events
Ack
Time
Cluster
Source
Message
18/Sep/2009 17:05:53 Demo Cluster demohead1 Service ntpd was restarted on demohead1
18/Sep/2009 17:05:47 Demo Cluster demohead1 Service named was restarted on demohead1
18/Sep/2009 17:05:45 Demo Cluster demohead1 Service postfix was restarted on demohead1
18/Sep/2009 17:05:45 Demo Cluster demohead1 Service dhcpd was restarted on demohead1
18/Sep/2009 17:05:45 Demo Cluster demohead1 Service maui was restarted on demohead1
Ready

The Black Hole Node Syndrome



Pre-Job Health Checks

- **Any Bright health check can be configured as a pre-job health check**

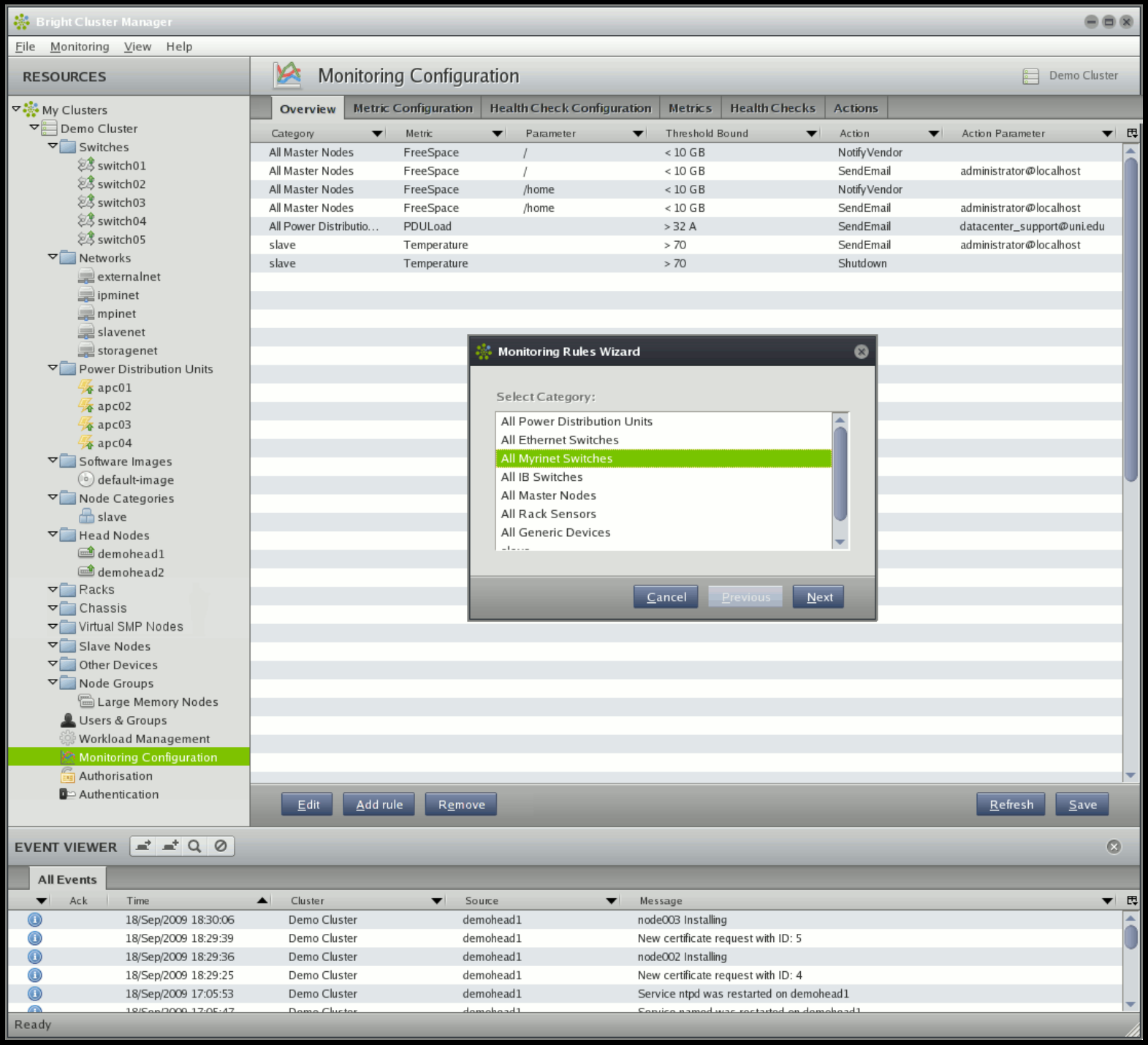


Pre-Job Health Checks

- Bright prolog script (cmprolog) is configured as the SLURM prolog script
- The prolog script calls all the configured pre-job health checks
- It instructs the cmdaemon on the execution node, to execute all pre-job health checks.
- If any of the pre-job health checks fail
  - The cmprolog script exits with code 99
  - The node is drained

```
scontrol state=DRAIN NodeName=node001
```

  - Administrator is notified of the failure.


Bright Cluster Manager
File Monitoring View Help
RESOURCES
Monitoring Configuration
Demo Cluster
My Clusters
Demo Cluster
Switches
switch01
switch02
switch03
switch04
switch05
Networks
externalnet
ipminet
mpinet
slavenet
storagenet
Power Distribution Units
apc01
apc02
apc03
apc04
Software Images
default-image
Node Categories
slave
Head Nodes
demohead1
demohead2
Racks
Chassis
Virtual SMP Nodes
Slave Nodes
Other Devices
Node Groups
Large Memory Nodes
Users & Groups
Workload Management
Monitoring Configuration
Authorisation
Authentication
Overview
Metric Configuration
Health Check Configuration
Metrics
Health Checks
Actions
Category
Metric
Parameter
Threshold Bound
Action
Action Parameter
All Master Nodes FreeSpace / < 10 GB NotifyVendor
All Master Nodes FreeSpace / < 10 GB SendEmail administrator@localhost
All Master Nodes FreeSpace /home < 10 GB NotifyVendor
All Master Nodes FreeSpace /home < 10 GB SendEmail administrator@localhost
All Power Distributio... PDULoad > 32 A SendEmail datacenter_support@uni.edu
slave Temperature > 70 SendEmail administrator@localhost
slave Temperature > 70 Shutdown
Monitoring Rules Wizard
Select Category:
All Power Distribution Units
All Ethernet Switches
All Myrinet Switches
All IB Switches
All Master Nodes
All Rack Sensors
All Generic Devices
Cancel
Previous
Next
Edit
Add rule
Remove
Refresh
Save
EVENT VIEWER
All Events
Ack
Time
Cluster
Source
Message
18/Sep/2009 18:30:06 Demo Cluster demohead1 node003 Installing
18/Sep/2009 18:29:39 Demo Cluster demohead1 New certificate request with ID: 5
18/Sep/2009 18:29:36 Demo Cluster demohead1 node002 Installing
18/Sep/2009 18:29:25 Demo Cluster demohead1 New certificate request with ID: 4
18/Sep/2009 17:05:53 Demo Cluster demohead1 Service ntpd was restarted on demohead1
Ready

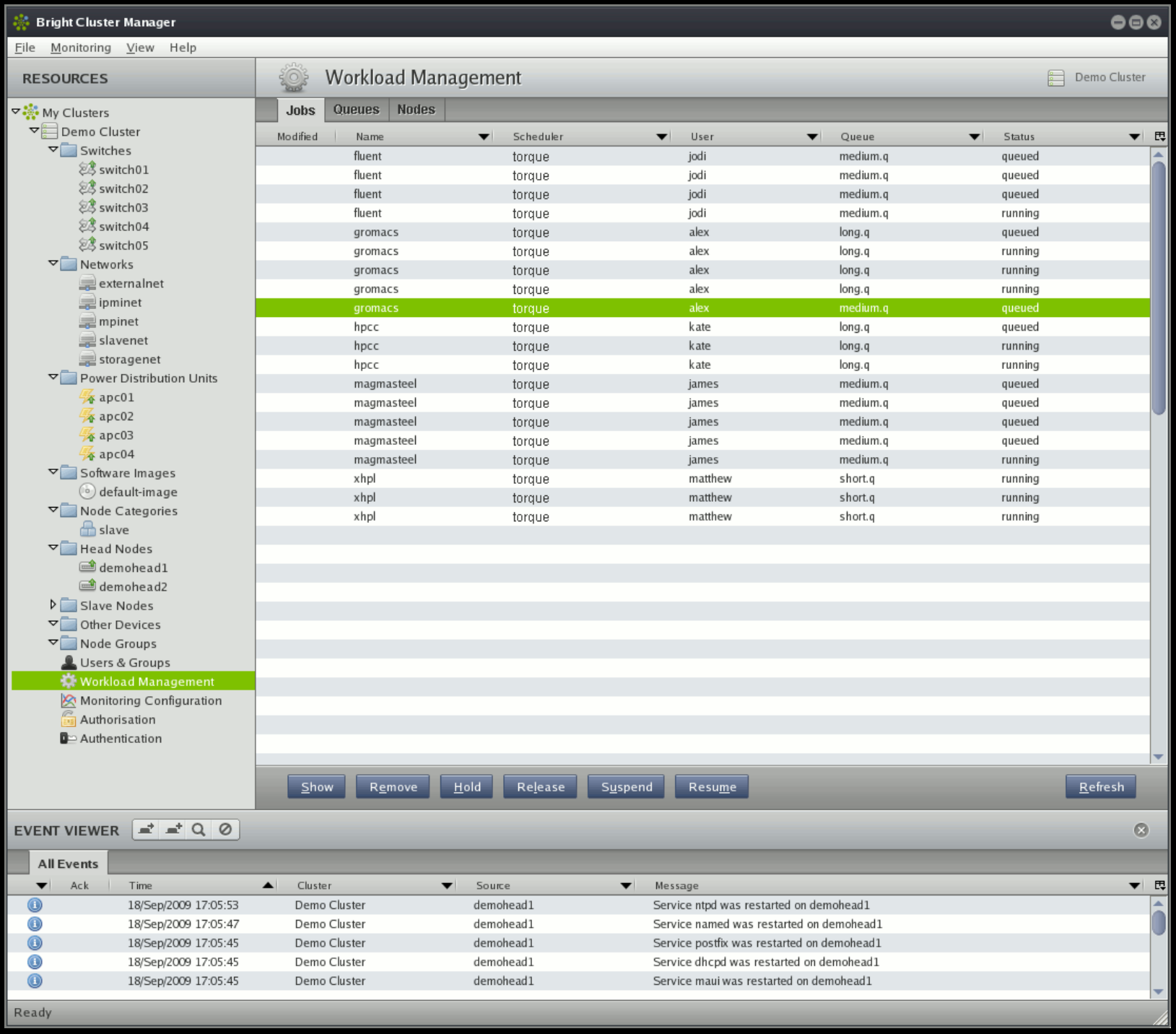
SLURM Configuration

SLURM Failover

- The SLURM failover role becomes available when a Bright failover node is configured
- When the failover occurs the SLURM DbdHost is changed to the secondary head node

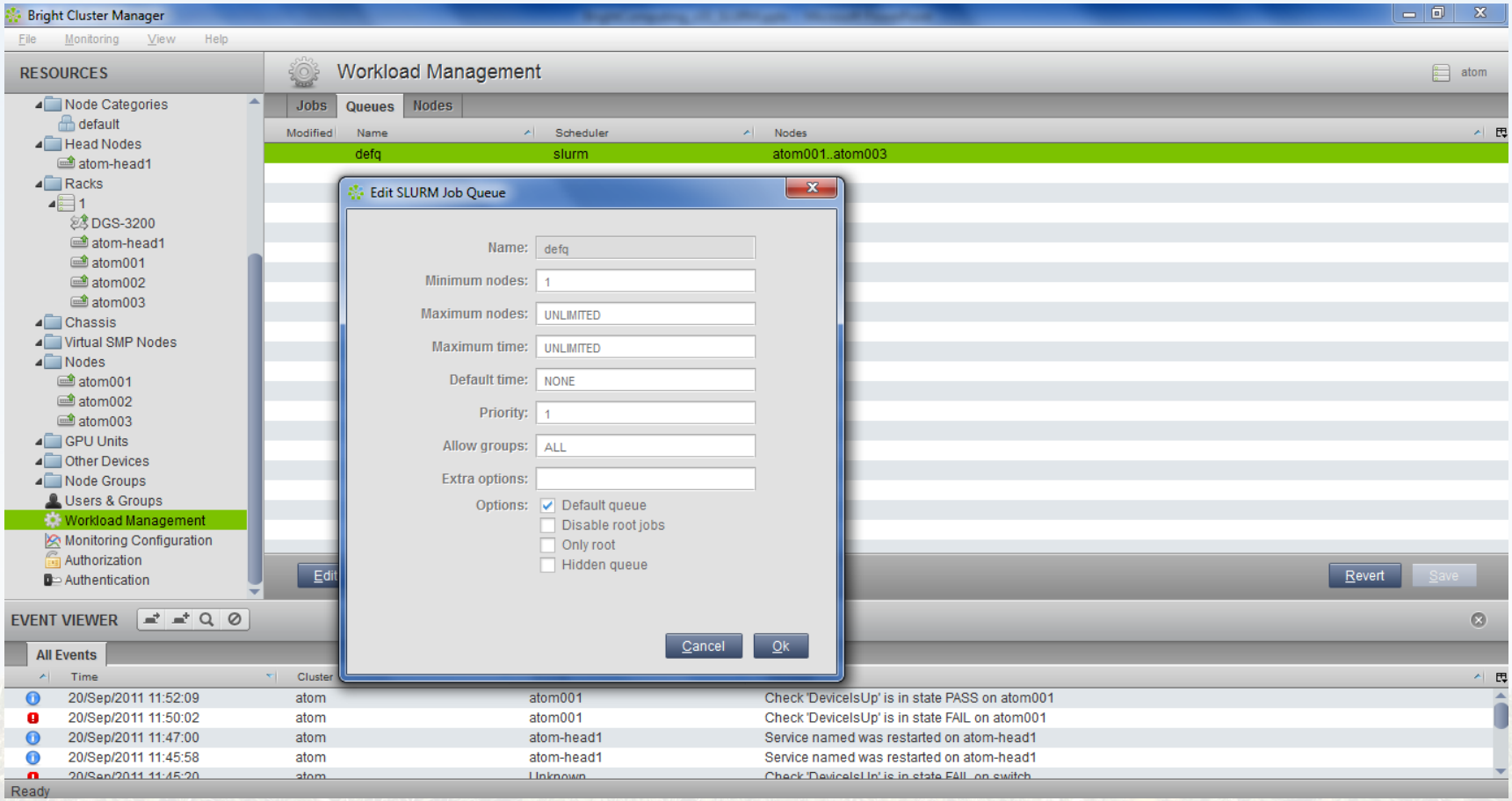
Bright provides the capability to

- Monitor, kill, suspend, resume, hold and release jobs
- Add, remove and edit queues
- View, drain and undrain nodes


Bright Cluster Manager
File Monitoring View Help
RESOURCES
My Clusters
Demo Cluster
Switches
switch01
switch02
switch03
switch04
switch05
Networks
externalnet
ipminet
mpinet
slavenet
storagenet
Power Distribution Units
apc01
apc02
apc03
apc04
Software Images
default-image
Node Categories
slave
Head Nodes
demohead1
demohead2
Slave Nodes
Other Devices
Node Groups
Users & Groups
Workload Management
Monitoring Configuration
Authorisation
Authentication
Workload Management
Demo Cluster
Jobs Queues Nodes
Modified Name Scheduler User Queue Status
fluent torque jodi medium.q queued
fluent torque jodi medium.q queued
fluent torque jodi medium.q queued
fluent torque jodi medium.q running
gromacs torque alex long.q queued
gromacs torque alex long.q running
gromacs torque alex long.q running
gromacs torque alex long.q running
gromacs torque alex medium.q queued
hpcc torque kate long.q queued
hpcc torque kate long.q running
hpcc torque kate long.q running
magmasteel torque james medium.q queued
magmasteel torque james medium.q queued
magmasteel torque james medium.q queued
magmasteel torque james medium.q queued
magmasteel torque james medium.q running
xhpl torque matthew short.q running
xhpl torque matthew short.q running
xhpl torque matthew short.q running
Show Remove Hold Release Suspend Resume Refresh
EVENT VIEWER
All Events
Ack Time Cluster Source Message
18/Sep/2009 17:05:53 Demo Cluster demohead1 Service ntpd was restarted on demohead1
18/Sep/2009 17:05:47 Demo Cluster demohead1 Service named was restarted on demohead1
18/Sep/2009 17:05:45 Demo Cluster demohead1 Service postfix was restarted on demohead1
18/Sep/2009 17:05:45 Demo Cluster demohead1 Service dhcpd was restarted on demohead1
18/Sep/2009 17:05:45 Demo Cluster demohead1 Service maui was restarted on demohead1
Ready

SLURM Configuration

- Example of editing an existing queue



Questions?

Robert Stober
robert.stober@brightcomputing.com
+1 209 986 9298