

# An Integrated Approach to Workload and Cluster Management: The HP CMU PBS Professional Connector

---

*Scott Suchyta*

*Altair Engineering Inc., 1820 Big Beaver Road, Troy, MI 48083, USA*

## Contents

1	Abstract.....	3
2	Introduction to HP Insight CMU.....	3
3	Introduction to PBS Professional .....	4
4	Introduction to CMU PBS Professional Connector .....	5
5	Installation of the CMU PBS Professional Connector .....	7
6	Configuring PBS Professional from CMU.....	8
7	Managing and Monitoring PBS Server and Scheduler .....	9
7.1	Configuring PBS Server (qmgr).....	9
7.2	Configuring PBS custom resources (resourcedef).....	10
7.3	Configuring PBS Scheduler (sched_config).....	11
7.4	Restart / Stop / Start PBS Server and Scheduler .....	12
7.5	Show Jobs and Show Reservations .....	12
7.6	Update logical groups (AOEs) / Dynamic Provisioning .....	12
7.7	Update network topologies (Placement Sets) .....	13
8	Managing and Monitoring PBS Professional Compute Nodes.....	14
8.1	Offline / Drain / Online node.....	15
8.2	Put node into maintenance / production mode .....	16
8.3	Enable / Disable PBS OS provisioning.....	18
8.4	Register / Unregister node with PBS .....	19
8.5	Restart / Stop / Start PBS MOM .....	20
8.6	PBS Node State .....	21
8.7	PBS Daemon Status.....	22
8.8	Number of jobs on node .....	22
9	Managing and Monitoring Jobs .....	22
9.1	Delete jobs.....	23
9.2	Suspend / Resume jobs.....	23
9.3	Checkpoint / Restart jobs .....	25
9.4	Show running / suspended / checkpointed jobs .....	27
9.5	User Groups .....	27
10	Online Resources.....	28
10.1	PBS Professional Documentation .....	28
10.2	PBS Professional User Forum.....	29
10.3	PBS Professional Customer Support .....	29
11	Conclusion.....	29
12	References.....	30

## 1 Abstract

Setting up a cluster can be challenging for a system administrator. The burden of installing a complete software stack across hundreds or thousands of compute nodes, combined with the pressure to have the system ready for production quickly (sometimes in days or hours), results in a highly complex project that can be daunting to even the most seasoned administrators. This paper will describe the integration between Altair's PBS Professional and HP's Insight Cluster Management Utility (CMU), explain how the concept of a "connector" unites these two tools to simplify cluster setup and job execution, and provide instructions for performing key PBS Professional tasks within Insight CMU.

## 2 Introduction to HP Insight CMU

HP Insight Cluster Management Utility (CMU) is an efficient and robust hyperscale cluster lifecycle management framework and suite of tools for large Linux clusters such as those found in High Performance Computing (HPC) environments. A simple graphical interface enables an 'at-a-glance' view of the entire cluster across multiple metrics, provides frictionless scalable remote management and analysis, and allows rapid provisioning of software to all the nodes of the system. Insight CMU makes the management of a cluster more user friendly, efficient, and error free than if it were being managed by scripts, or on a node-by-node basis. Insight CMU offers full support for iLO2, iLO3, iLO4 and LO100i adaptors on all the ProLiant HP servers in the cluster.

Insight CMU is highly flexible and customizable, offers both GUI and CLI interfaces, and is being used to deploy a range of software environments, from simple compute farms to highly customized, application-specific configurations. Insight CMU is available for HP ProLiant and HP BladeSystems, with Red Hat Enterprise Linux and Novell SUSE Linux operating systems, including Red Hat Enterprise Linux, Suse Linux Enterprise, CentOS, and Ubuntu. Insight CMU also includes options for monitoring GPUs and for installing GPU drivers and software.

HP Insight CMU offers the following features:

- One-to-N user interface -- the industry's leading cluster GUI and command-line-interface
  - A system admin can click-select a set of servers and groups for cloning, managing or monitoring thousands of servers at once
- Cloning
  - Choose one server and replicate its image over N servers in the cluster. CMU can clone a thousand servers in less than half an hour
- Remote management
  - Choose one set of commands or scripts and execute them on N servers with a single mouse-click

- Monitoring
  - Choose, with one click per sensor, to display real-time metrics for N servers in a pie-chart format or in a 3D tubular display designed for optimal display of data-collections
  - Configure alerts to cause threshold-triggered actions
- Cluster diagnostics
  - Verify that the components, connections and interactions of the cluster infrastructure are all working correctly and at full performance
- Extensible menu of common cluster operations
  - A default drop-down menu of commands includes power on/off, firmware installation, parallel distributed shell and unattended Linux installation
  - Simple file-edits can be used to add operations to the menu
- Scalable performance
  - Hierarchical ultra-lightweight cloning and monitoring operations are automatically done in parallel

### 3 Introduction to PBS Professional

Today's enterprises have grown extremely complex from a computing perspective. IT administrators often need to support hundreds of users running thousands of jobs across many different types of machines in diverse geographical locations. In this distributed heterogeneous environment, it can be extremely difficult for administrators to collect detailed, accurate usage data or to set system-wide resource priorities. As a result, many computing resources are under-utilized while others are over-utilized. At the same time, users are confronted with an ever-expanding array of operating systems and platforms. As a result, each year scientists, engineers, designers and analysts must waste countless hours learning the nuances of different computing environments, rather than focusing on their primary goals.

PBS Professional is the professional version of the Portable Batch System (PBS), a flexible workload management solution originally developed to manage aerospace computing resources at NASA. For more than 20 years, PBS Professional has managed high-performance computing systems at thousands of customer sites; the product is the acknowledged leader in supercomputer workload management and the de facto standard on Linux clusters.

PBS Professional consists of two major component types: **commands** and **daemons / services**. A brief description of each is given below to help readers understand how the components operate together:

- **Commands** – PBS Professional supplies both a graphical interface and a set of POSIX 1003.2d-conforming command-line programs. These are used to submit, monitor, modify and delete jobs. These client commands can be installed on any system type

supported by PBS Professional and do not require local presence of any of the other PBS Professional components.

There are three command classifications in PBS Professional:

1. Commands available to any authorized user
  2. Commands requiring special PBS Professional privileges
  3. Administrator commands which require root access or the equivalent to use
- **Daemons/Services** – Fundamentally, PBS Professional consists of three daemons: the server is the primary interface for the administrator and user, the scheduler orchestrates the workload, and the MOM manages and monitors job execution.

The roles and responsibilities of the three primary daemons/services in PBS Professional are detailed below:

1. **Server** – The server daemon/service, `pbs_server`, is the central component of PBS Professional. All commands and other daemons/services communicate with the `pbs_server` via an Internet Protocol (IP) network. The `pbs_server`'s main function is to provide basic batch services such as receiving/creating a batch job, modifying the job, and passing the job to the execution node. One `pbs_server` manages the machines and jobs in a PBS Professional complex; a secondary server may be configured to handle failover.
2. **Execution Node Manager (MOM)** – A MOM is a daemon/service on the node that actually places a job into execution. The `pbs_mom` daemon is informally called MOM, as it is the mother of all processes for jobs on a particular node. (MOM is a reverse-engineered acronym that stands for Machine Oriented Mini-server.) One MOM runs on each execution node. A node's MOM places a job into execution when it receives a copy of the job from the `pbs_server`. MOM creates a new session for each job and gathers information about the resource usage of that job. MOM also has the responsibility for communicating with other MOMs assigned to the job and returning the job's output to the user when directed to do so by the Server.
3. **Scheduler** – The job scheduler daemon/service, `pbs_sched`, implements the site's scheduling policy, controlling when each job is run and on which resources. The Scheduler may communicate with the various MOMs to query the state of system resources, and with the Server for availability of jobs to execute.

## 4 Introduction to CMU PBS Professional Connector

The CMU PBS Professional Connector provides seamless integration between PBS Professional and HP's Insight CMU, simplifying cluster deployment and management for users and administrators. Following HP's best practices for Insight CMU integration, the

CMU PBS Professional Connector automates the most commonly used tasks for managing a PBS Professional cluster via Insight CMU.

Insight CMU provides a framework to add custom menus to its GUI. The CMU PBS Professional Connector utilizes this framework to automate many of the tasks an administrator executes while managing a cluster. With the CMU PBS Professional Connector, these tasks are made accessible from the menus so administrators can easily access and implement them.

In addition, Insight CMU provides a scalable engine with a default set of sensors that capture metrics such as CPU load, memory usage, I/O performance, and network performance. The CMU PBS Professional Connector takes advantage of this by gathering PBS-specific metrics about nodes and jobs executing on nodes, and exposing these metrics to the administrator and user.

#### Integration highlights:

- Connector approach automates the majority of common admin tasks using HP's preferred method for integration
- Dynamic creation of node groups in the Insight CMU GUI when jobs are executed, so admins and users can easily see utilization and performance metrics related to all nodes where a certain job is running
- Dynamic OS provisioning based on workload in queues, with automatic GUI updates, so PBS Professional can comprehend all existing OS images that can be provisioned on nodes
- Integrated PBS Professional menus in Insight CMU's GUI for registering nodes, onlining/offlining (or "draining") nodes, enabling OS provisioning, deleting jobs and much more
- Easy access to data on jobs and reservations, including which jobs are running, suspended or checkpointed on nodes
- Automatic network topology configuration so PBS Professional can optimally place jobs on appropriate network entities
- Maintenance mode support allowing admins to move 'bad' nodes into maintenance mode for troubleshooting without competing with users' jobs
- One-click access to online resources including PBS Professional documentation, user forums and support

#### Benefits to admin/users:

- Improved visibility into utilization and performance metrics
- Improved access to job data (e.g. jobs on cluster, reservations on cluster...)
- Simplified management (admins can easily perform numerous tasks without needing to open a terminal window)

- Optimized job placement via automatic topology configuration
- Very easy to install
- Easily extensible for site-specific customizations
- One-click access to support resources

## 5 Installation of the CMU PBS Professional Connector

*Before installing the CMU PBS Professional Connector, the PBS Server and Scheduler must be installed and licensed on the Insight CMU management node. Please refer to the [PBS Professional Installation Guide](#)<sup>1</sup> on how to install the PBS Server and Scheduler and license the product.*

The administrator can download the CMU PBS Professional Connector from the PBS Works website: [http://www.pbsworks.com/cmu\\_pbsconnector](http://www.pbsworks.com/cmu_pbsconnector)

The installation of the CMU PBS Professional Connector package is fairly straightforward, but does require some manual steps:

1. Copy the `cmu_pbspro_connector` tarball to `/opt/cmu/contrib`

```
cp cmu_pbspro_connector*.tar.gz /opt/cmu/contrib
```

2. Change directory to `/opt/cmu/contrib`

```
cd /opt/cmu/contrib
```

3. Unpackage the `cmu_pbspro_connector` tarball

```
tar zxvf cmu_pbspro_connector*.tar.gz
```

4. Update `/opt/cmu/etc/ActionAndAlertsFile.txt` with the content provided in the `cmu_pbspro_connector/etc/PBS_Specific_ActionAndAlerts.txt` file.

Administrators will need to manually add the PBS-specific content to the designated areas of the `/opt/cmu/etc/ActionAndAlerts.txt` file. The `PBS_Specific_ActionAndAlerts.txt` file contains a default set of PBS Professional specific metrics that CMU will begin monitoring. Administrators can extend the configuration and capabilities as their cluster and/or requirements grow.

5. Update `/opt/cmu/etc/cmu_custom_menu` with the content provided in the `cmu_pbspro_connector/etc/PBS_Specific_cmu_custom_menu` file. The `PBS_Specific_cmu_custom_menu` file contains the PBS-specific menu options that will be

<sup>1</sup> <http://www.pbsworks.com/documentation/support/PBSProInstallGuide12.pdf>

seen in the CMU GUI. The configuration and capabilities can be extended as the cluster and/or requirements grow.

```
mv ../etc/cmu_custom_menu ../etc/cmu_custom_menu.orig
cat ../etc/cmu_custom_menu.orig
cmu_pbspro_connector/etc/PBS_Specific_cmu_custom_menu >
../etc/cmu_custom_menu
```

## 6. Restart CMU for configuration to take effect

```
/etc/init.d/cmu restart
```

## 6 Configuring PBS Professional from CMU

After deploying the 'golden image' to all the compute nodes within the cluster, with a few clicks of CMU's PBS-specific menu options, administrators can add compute nodes to PBS Professional and begin collecting metrics on job and node status. Details on all menu options and advanced configurations are explained later in the documentation.

To finalize configuration of PBS Professional within the cluster and begin accepting jobs:

1. Add all of the compute nodes to the PBS Server's inventory by selecting them in the frame on the left side of the CMU GUI.
2. Right-click to display the menu, and navigate to *CMU PBS Professional Connector > Register node with PBS*.

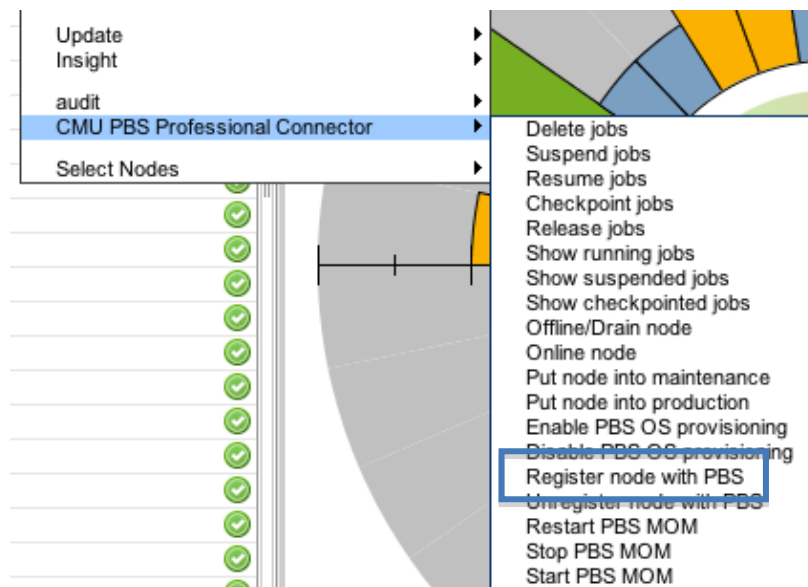


Figure 1: Register node with PBS Professional



Upon completion of the task, a CMU dialog box will appear with the success or failure output. If there are no errors, the administrator and users can now begin submitting jobs to PBS Professional. Please refer to the [PBS Professional User's Guide](#)<sup>2</sup> for details on how to submit jobs.

- For details on how to create logical groups, backup the 'golden image' to a logical group, and clone the compute nodes from within the CMU GUI, please refer to the HP Insight Cluster Management Utility manuals by going to <http://www.hp.com/go/cmu> and clicking on the Technical Documentation link.
- For details on how to install the PBS Professional MOM on the compute node, please refer to the [PBS Professional Installation Guide](#)<sup>3</sup>.

## 7 Managing and Monitoring PBS Server and Scheduler

*CMU relies on an X Windows client to be installed and configured correctly on the system where the CMU client will run. If X Windows is not configured correctly, some of the capabilities described below will not function.*

The configuration of the PBS Server and Scheduler has several important steps, and for a new administrator these tasks can be easily overlooked. This section covers the following menu options exposed in the CMU PBS Professional Connector and explain the benefit of each:

- [Configuring PBS Server \(qmgr\)](#)
- [Configuring PBS Scheduler \(sched config\)](#)
- [Configuring PBS Custom Resources \(resourcedef\)](#)
- [Show jobs](#)
- [Show reservations](#)
- [Update logical groups \(AOEs\) / Dynamic Provisioning](#)
- [Update network topologies \(Placement Sets\)](#)
- [Restart / Stop / Start PBS Server and Scheduler](#)

### 7.1 Configuring PBS Server (qmgr)

Administrators can configure the server by setting server attributes via the qmgr command. The qmgr command is used to create or delete queues, vnodes, and hooks; to set or change vnode, queue, hook, server, or scheduler attributes, including resources; and to view information about hooks, queues, vnodes, the server, and the scheduler. Changes made in qmgr take effect immediately, so there is no need to restart the PBS Server daemon.

<sup>2</sup> <http://www.pbsworks.com/documentation/support/PBSProUserGuide12.pdf>

<sup>3</sup> <http://www.pbsworks.com/documentation/support/PBSProInstallGuide12.pdf>

The *Configure PBS Server (qmgr)* menu option will open a terminal window on the PBS Server with the qmgr prompt ready.

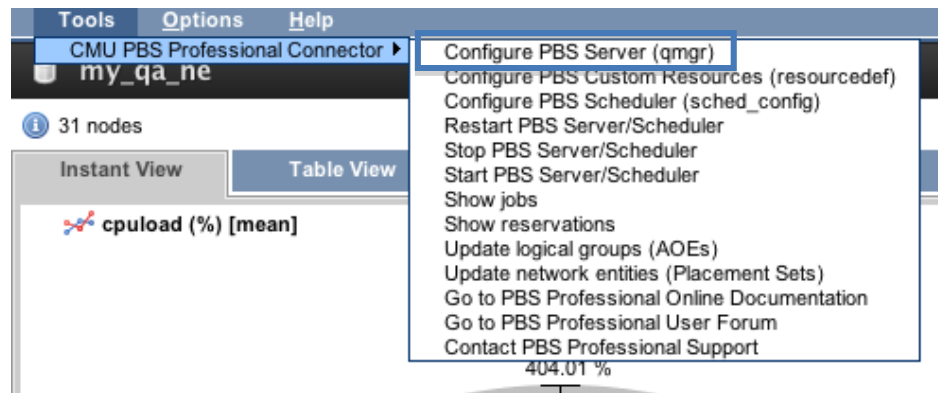


Figure 2: Configure PBS Server (qmgr)

More details about the attributes that can be configured for the PBS Server can be found in the [PBS Professional Administrator's Guide](#)<sup>4</sup> and [PBS Professional Reference Guide](#)<sup>5</sup>.

## 7.2 Configuring PBS custom resources (resourcedef)

PBS Professional has built-in resources such as ncpus, mem, and walltime, which are used in the scheduling of jobs. However, not all sites are created equal. PBS Professional allows for administrators to create custom resources that can be used in scheduling policies and/or meta-data to be used for accounting purposes. For instance, if an administrator has a limited number of application licenses, he/she may want to have PBS Professional ensure that the license is available before dispatching the job to the execution node. Otherwise, the job could fail because the license was not available. Then the user will be upset because they had been waiting in the queue only to find the job failed because of something unrelated to their setup.

Creating a custom resource requires some moving parts, but with the *Configuring PBS Custom Resources (resourcedef)* menu option these are automated for the administrator.

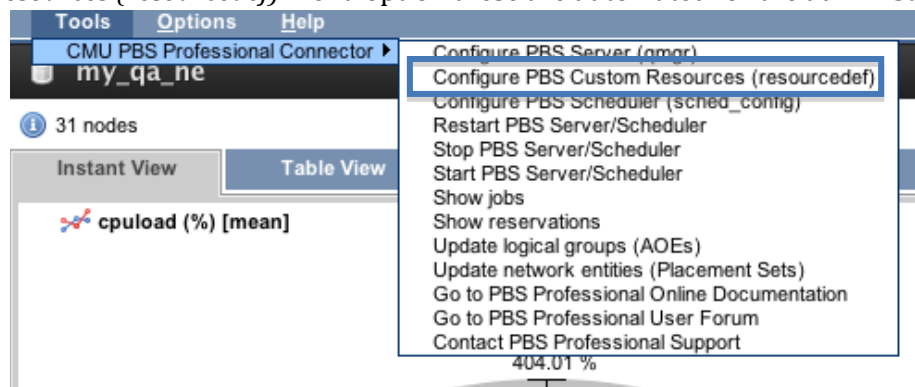


Figure 3: Configure PBS Custom Resources (resourcedef)

<sup>4</sup> <http://www.pbsworks.com/documentation/support/PBSProAdminGuide12.pdf>

<sup>5</sup> <http://www.pbsworks.com/documentation/support/PBSProRefGuide12.pdf>

To configure PBS custom resources, select this menu option -- a terminal window will open with the `$PBS_HOME/server_priv/resourcedef` file open and ready to edit. Upon saving the file, the PBS Server will automatically be restarted. If the new resource needs to be schedulable, modify the `sched_config`. See *Configuring PBS Scheduler (sched\_config)* menu option, and update the resources: scheduler attribute

### 7.3 Configuring PBS Scheduler (sched\_config)

The PBS Scheduler implements scheduling policy (e.g., backfill, preemption, fairshare, dedicated time, job or node sorting). The behavior of the scheduler is primarily controlled by the `$PBS_HOME/sched_priv/sched_config` file, but will also rely on supplemental files for decisions.

The `PBS_HOME/sched_priv/sched_config` contains scheduler configuration options, also called scheduler parameters (e.g. backfill, job\_sort\_key). The configuration file is read on startup and SIGHUP.

The *Configure PBS Scheduler (sched\_config)* menu option will open a terminal window on the PBS Scheduler with the scheduler configuration file (`sched_config`) open and ready to be modified. Upon saving the `sched_config`, the PBS Scheduler will be sent a SIGHUP to ensure the changes are automatically re-read.

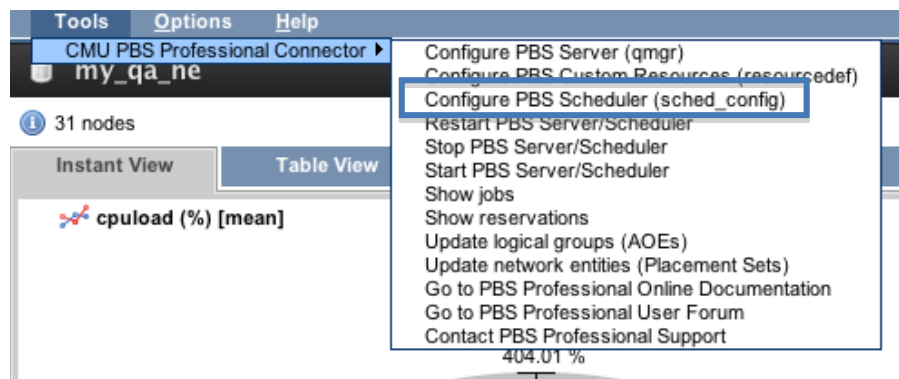


Figure 4: Configure PBS Scheduler (sched\_config)

It is important to note that each entry in the `sched_config` must be a single, unbroken line. Entries must be double-quoted if they contain whitespace.

## 7.4 Restart / Stop / Start PBS Server and Scheduler

Sometimes it is necessary to restart, stop or start the PBS Server and Scheduler. Each of these actions are available in the CMU menu bar, *Tools > CMU PBS Professional Connector*.

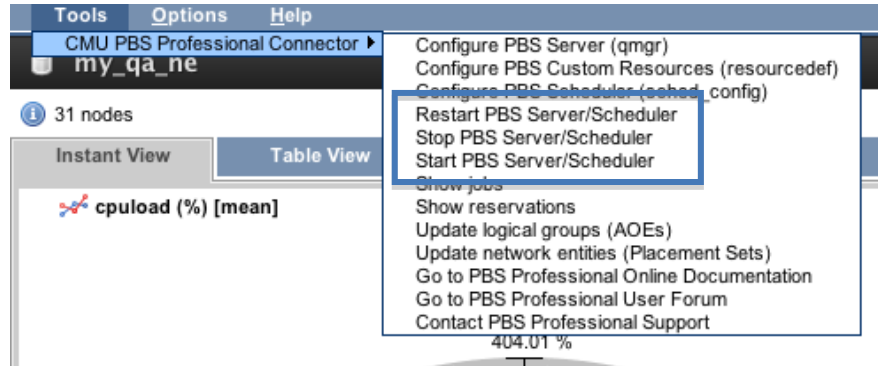


Figure 5: Restart / Stop / Start PBS Server and Scheduler

## 7.5 Show Jobs and Show Reservations

The *Show Jobs* and *Show Reservations* menu options are commonly used to see the status of jobs within the cluster. Instead of opening a terminal on the cluster to view status, these menu options will execute the appropriate commands and provide the output in a CMU dialog box.

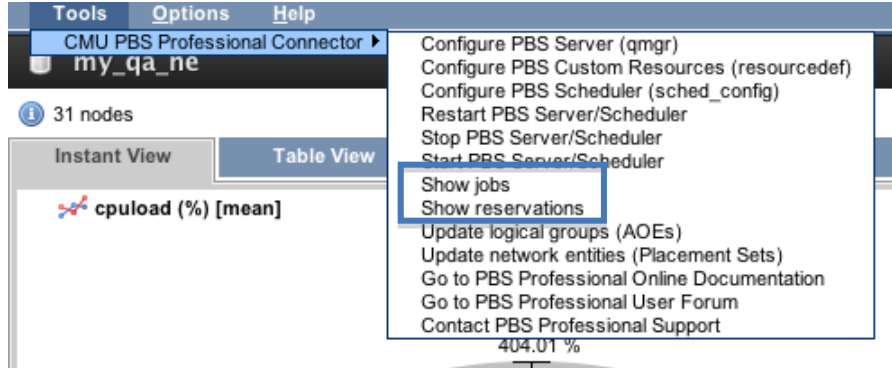


Figure 6: Show jobs and Show reservations

## 7.6 Update logical groups (AOEs) / Dynamic Provisioning

Allowing users to request specific operating systems during submission is typically an advanced configuration. The benefit of dynamically provisioning a cluster is that it enables the efficient utilization of all cluster compute nodes; thus, administrators will not need to dedicate compute nodes to running a specific operating system resulting in under-utilization of the node. PBS Professional refers to logical groups as application operating environment (AOEs).

If users need the ability to request specific operating systems during job submission, the *Update logical groups (AOEs)* menu option will greatly simplify this task by automating a

number of steps -- including associating all known logical groups to a node and configuring the current logical group on the node. PBS Professional will use this information to determine which compute nodes to provision, and when to provision them, by utilizing CMU's cloning capabilities. Please refer to [PBS Professional Administrator's Guide](#)<sup>6</sup> for more details on Dynamic Provisioning.

To update PBS Professional with all known logical groups, navigate to the CMU menu bar, *Tools > CMU PBS Professional Connector > Update logical groups (AOEs)*.

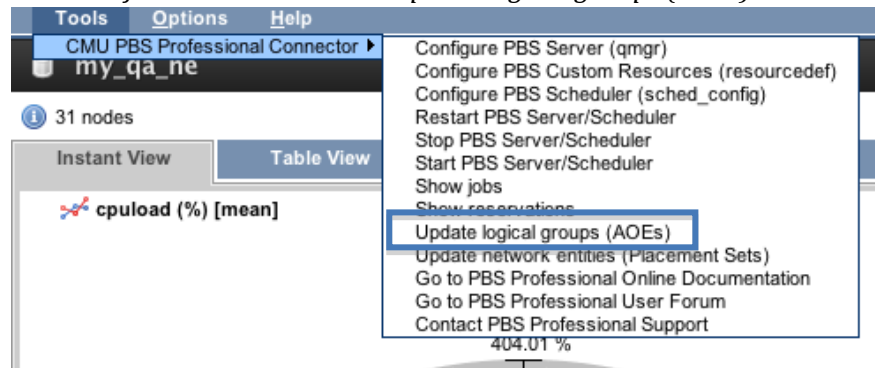


Figure 7: Update logical groups (AOEs)

Upon completion of the task, a CMU dialog box will appear with the success or failure output. If there are no errors, conclude the configuration by selecting the compute nodes eligible for dynamic provisioning and select *Enable PBS OS provisioning* from the custom menu. See "[Enable / Disable PBS OS provisioning](#)" in this document for more information.

Once PBS OS provisioning is enabled on certain nodes, the administrator or users can now begin submitting jobs to PBS Professional requesting specific images. Please refer to the [PBS Professional User's Guide](#)<sup>7</sup> to learn more about how to request a job to provision the compute nodes.

## 7.7 Update network topologies (Placement Sets)

Clusters can have a simple network infrastructure or they can have a complex infrastructure with multiple network switches and/or interconnect technologies being used. Please refer to the HP Insight Cluster Management Utility User Guide for specific details on how to create network entities within the CMU GUI.

If the cluster has multiple network switches or different types of interconnect, the administrator may want to consider leveraging PBS Professional to optimize the placement of the user's job within a group of nodes that have similar characteristics. The placement of the user's job can be transparent to the user submitting the job, meaning that PBS Professional can make the placement decision. If the administrator has advanced users or benchmark users, they may benefit from being able to explicitly express the type of network

<sup>6</sup> <http://www.pbsworks.com/documentation/support/PBSProAdminGuide12.pdf>

<sup>7</sup> <http://www.pbsworks.com/documentation/support/PBSProUserGuide12.pdf>

or interconnect the job will run on. Please refer to the [PBS Professional Administrator's Guide](#)<sup>8</sup> for more details on topology and placement sets.

To configure PBS Professional to be aware of these network topologies, navigate to the CMU menu bar, *Tools > CMU PBS Professional Connector > Update network topologies (Placement Sets)*.

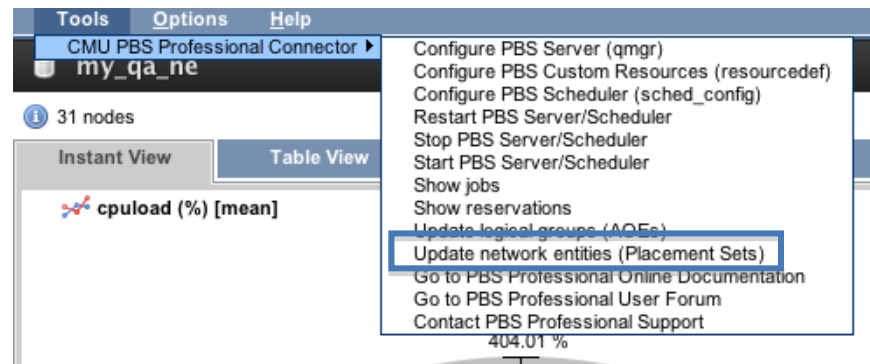


Figure 8: Update network entities (Placement Sets)

Upon completion of this task, a CMU dialog box will appear with the success or failure output. If there are no errors, PBS Professional will now begin considering the network topology within its scheduling decisions.

## 8 Managing and Monitoring PBS Professional Compute Nodes

The cluster administrator may not necessarily be an expert user of the workload manager / job scheduler. However, he/she is typically experienced with administering compute nodes and ensuring performance at expected service level agreements. This section covers how the administrator can manage and monitor the nodes without needing to be an experienced PBS Professional user.

Below are the management-related menu options available to the administrator to act on the selected compute nodes:

- [Offline/Drain / Online node](#)
- [Put node into maintenance / production mode](#)
- [Enable / Disable PBS OS provisioning](#)
- [Register / Unregister node with PBS](#)
- [Restart / Stop / Start PBS MOM](#)

<sup>8</sup> <http://www.pbsworks.com/documentation/support/PBSProAdminGuide12.pdf>



From a monitoring perspective, the administrator can see at-a-glance whether PBS Professional is functioning properly and how many jobs are on the node through the CMU GUI. The administrator can see these metrics within the Table View. Defining new metrics is very flexible and extendible, so as the administrator's cluster knowledge grows he/she will be able to create new metrics to accommodate his organization's changing requirements.

Below are the custom metrics for monitoring that are provided with the CMU PBS Professional Connector:

- [PBS Node State](#)
- [PBS Daemon Status](#)
- [Number of jobs on node](#)

## 8.1 Offline / Drain / Online node

It is possible for a compute node to begin experiencing 'issues' where the cluster administrator may need to investigate the node. If a serious issue has been identified, which may require rebooting the node or installing new software packages, the administrator may want to put the node into maintenance mode. See "[Put node into maintenance / production mode](#)" for more details on what this menu option provides.

Offlining (a.k.a. draining) the node tells the scheduler that the node is no longer eligible to accept new work. However, if a job is running when the node is marked offline, PBS Professional will allow the job to finish, unless the administrator manually intervenes by terminating or re-queuing the job. See [Managing and Monitoring the Jobs](#) for more details on terminating and re-queuing jobs.

Administrators can choose compute nodes to mark offline by selecting the node(s), right-clicking, and navigating to *CMU PBS Professional Connector > Offline/Drain node*.

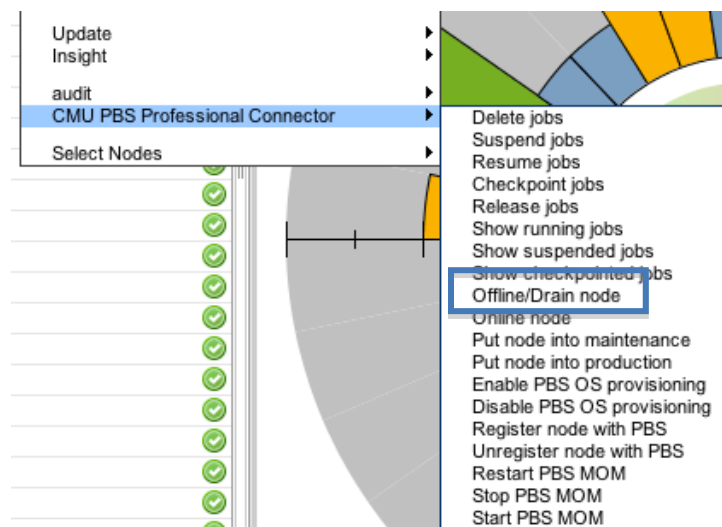


Figure 9: Offline/Drain node

This task will mark the node offline, insert a comment within the node's attribute, and add the node to CMU's custom user group called 'nodes\_offline'. The node's comment will note that the administrator has put the node offline with a date stamp:

```
CMU Administrator put node offline 12/10/2012
```

Once the administrator has finished this investigation, if the node does not require maintenance, he/she can safely online the node. This action will remove the node's comment, remove the node from the CMU custom group (nodes\_offline), and finally mark the node online. The node will then be considered as available in the next scheduling cycle. The administrator can select which compute nodes need to be marked online by selecting the node(s), right-clicking, and navigating to *CMU PBS Professional Connector > Online node*.

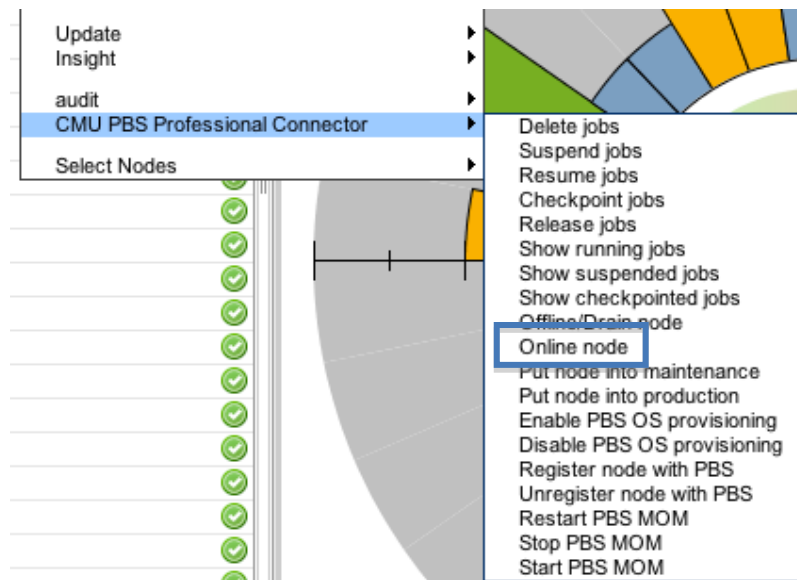


Figure 10: Online node

## 8.2 Put node into maintenance / production mode

It is inevitable that nodes will sometimes require maintenance. The menu option to put a node into maintenance mode will automate multiple tasks, allowing the administrator to safely work on the node without disrupting future workload within the cluster. The node will be assigned to a special queue, called node\_maint that has an Access Control List (ACL) defined. This special queue will enable the administrator (i.e., root) to submit test jobs to the node, verifying that it is operating correctly before releasing the node back into production.



The administrator can select which compute nodes to put into maintenance by selecting the node(s), right-clicking, and navigating to *CMU PBS Professional Connector > Put node into maintenance mode*.

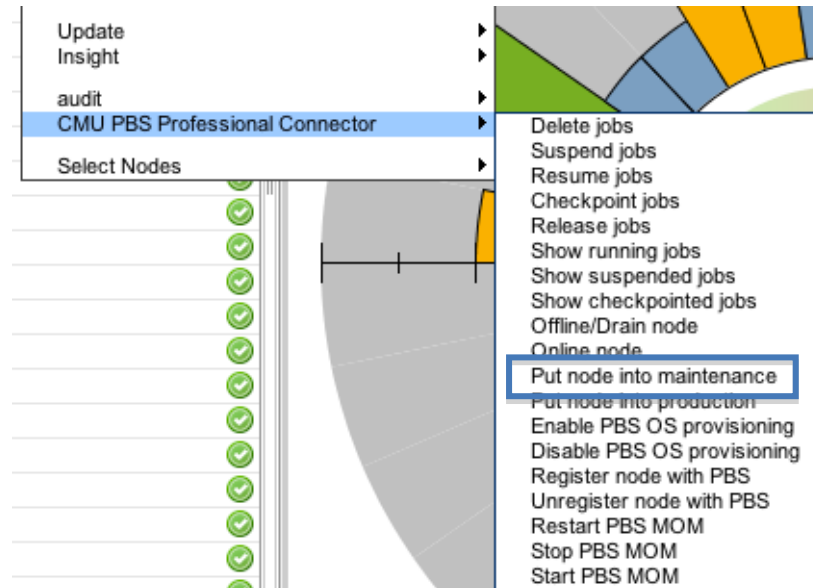


Figure 11: Put node into maintenance

This task will mark the node offline, insert a comment within the node's attribute, assign the node to a nodes\_maint queue, and add the node to the CMU custom user groups (nodes\_offline and nodes\_maint). The node's comment will note that the administrator has put the node into maintenance mode with a date stamp:

```
CMU Administrator put node into maintenance mode 12/10/2012
```

If a job was running on the node when it was selected, PBS Professional will allow the job to run to completion. It is important to verify that no [jobs are running](#) on the node(s) before performing maintenance. If this is not verified before rebooting the node, PBS Professional has a capability called node\_fail\_requeue, which will trigger and requeue the user's job. The potential drawback to not letting the job run to completion is that the job will likely be restarted from the beginning of execution.

Once maintenance is complete on the node, it can be put back into production. The administrator can select which compute nodes need to be put into maintenance by selecting the node(s), right-clicking, and navigating to *CMU PBS Professional Connector > Put node into production mode*.

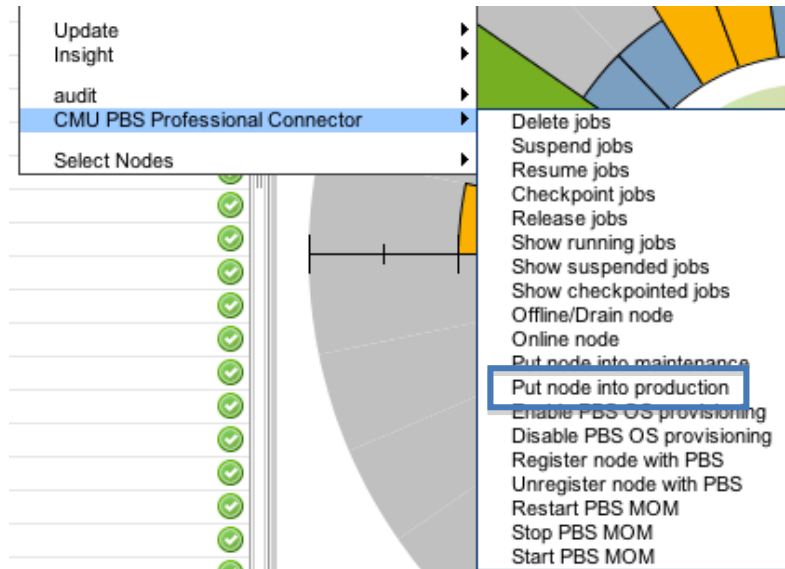


Figure 12: Put node into production

This action will mark the node offline, update the node's comment, remove the node from special queue (nodes\_maint), and remove node from the CMU custom group (nodes\_maint). The node's comment will be updated to say:

```
CMU Administrator put node into production, but node is offline
12/11/2012
```

Note that the node is not automatically marked online yet, because it is prudent for the administrator to perform 'one last check' to make sure the node is functioning correctly. Once the node is verified to be functioning correctly, the administrator can online the node. Refer to the section "Offline / Drain / Online node" for the details involved in onlining the node.

### 8.3 Enable / Disable PBS OS provisioning

The Enable and Disable OS provisioning menu option does exactly what it says. Earlier in this document, we discussed how to configure PBS Professional to recognize available CMU logical groups (see "Update logical groups (AOEs)") to handle dynamically provisioning the operating system on the compute nodes based on the user's job request.

The administrator can select which compute nodes should be eligible for dynamic OS provisioning by selecting the node(s), right-clicking, and navigating to *CMU PBS Professional Connector > Enable or Disable OS provisioning*.

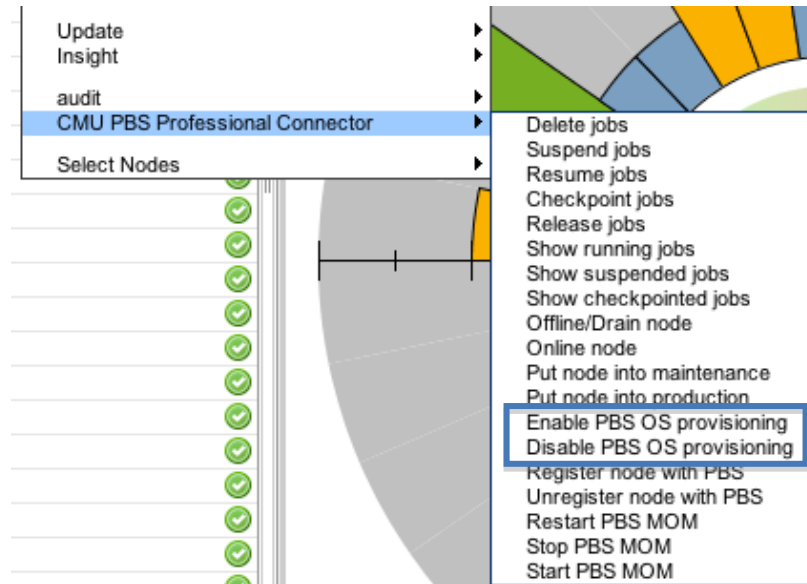


Figure 13: Enable or Disable PBS OS Provisioning

## 8.4 Register / Unregister node with PBS

During the configuration of PBS Professional via CMU, the administrator was instructed to register the nodes with PBS Professional, which entails the node being created within PBS Professional. The node will then be added to the pool of available resources that the scheduler can consider within the scheduling policies.

By unregistering the node from PBS Professional, this means the node is deleted from the pool of resources available to the scheduler. Note that unregistering the node is different than offlining the node or putting the node into maintenance, and should only be executed if the node is to be physically removed from the cluster.

The administrator can select which compute nodes are to be registered or unregistered with PBS Professional by selecting the node(s), right-clicking, and navigating to *CMU PBS Professional Connector > Register or Unregister node with PBS*.

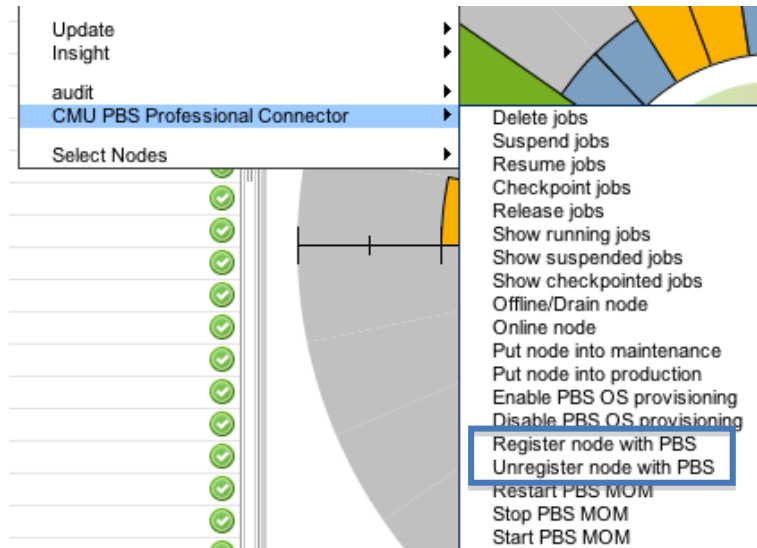


Figure 14: Register or Unregister node with PBS

## 8.5 Restart / Stop / Start PBS MOM

Similar to the PBS Server and Scheduler, sometimes it is necessary to restart, stop or start the PBS MOM. Each of these actions are available after selecting *node > CMU PBS Professional Connector*.

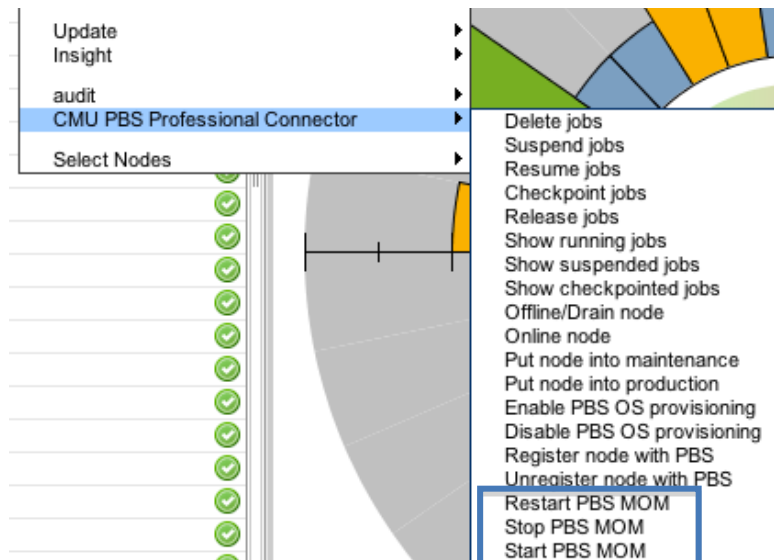


Figure 15: Restart / Stop / Start PBS MOM

## 8.6 PBS Node State

Compute nodes can have different PBS states (e.g., offline, job-busy, down), and in many cases the administrator needs to easily see which nodes are not running jobs, especially the nodes that are considered down. If a node should happen to be Offline or Down, CMU will recognize this state and change the node's badge icon from "OK" to "Warning" (yellow triangle).

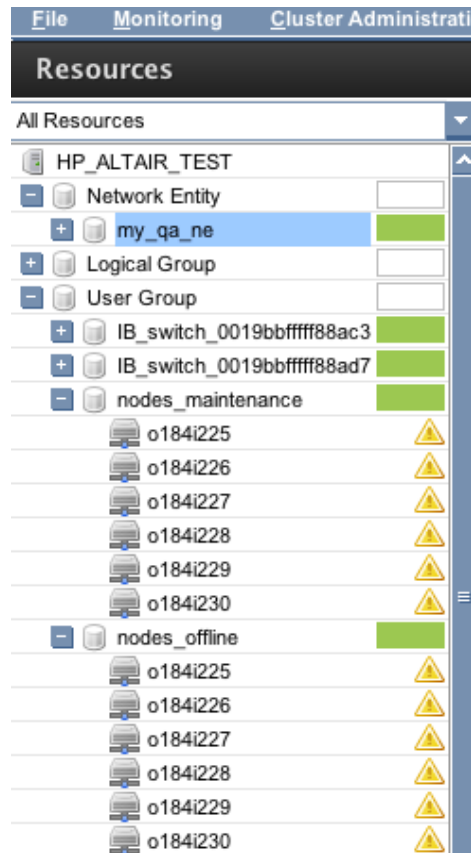


Figure 16: PBS Node State via Tree Panel

The administrator can also see the PBS Node State (node\_state column) along with other metrics being monitored in the CMU Table View.

S...	current_image	pbs_daem...	node_state	job_count
	pbs_rhel6u2_...	Up	job-busy	2
	pbs_rhel6u2_...	Up	job-busy	2
	pbs_rhel6u2_...	Up	job-busy	4
	pbs_rhel6u2_...	Up	job-busy	4
	pbs_rhel6u2_...	Up	job-busy	4

Figure 17: PBS Node State via CMU Table View

## 8.7 PBS Daemon Status

The PBS Daemon Status metric monitors whether the pbs\_mom daemon is running (Up) or not (Down). If the PBS Daemon Status is Down on a node, the administrator will be able to see “Down” in the pbs\_daemon\_status column of the CMU Table View.

S...	current_image	pbs_daem...	node_state	job_count
	pbs_rhel6u2_...	Up	job-busy	2
	pbs_rhel6u2_...	Up	job-busy	2
	pbs_rhel6u2_...	Up	job-busy	4
	pbs_rhel6u2_...	Up	job-busy	4
	pbs_rhel6u2_...	Up	job-busy	4

Figure 18: PBS Daemon Status via CMU Table View

## 8.8 Number of jobs on node

The Number of jobs on node metrics displays exactly what the metric name suggests. When accessing this metrics in the CMU Table View, the administrator can see how many jobs are running on a node in the job\_count column. Some sites configure their systems such that the job is exclusive to the node(s). Other sites may have compute nodes that have the capacity to run multiple jobs on a single node. Refer to the [PBS Professional Administrator's Guide](#)<sup>9</sup> for more details on how to configure nodes to be exclusive or share resources with a job.

S...	current_image	pbs_daem...	node_state	job_count
	pbs_rhel6u2_...	Up	job-busy	2
	pbs_rhel6u2_...	Up	job-busy	2
	pbs_rhel6u2_...	Up	job-busy	4
	pbs_rhel6u2_...	Up	job-busy	4
	pbs_rhel6u2_...	Up	job-busy	4

Figure 19: Number of jobs on node via CMU Table View

## 9 Managing and Monitoring Jobs

Administrators can perform certain actions on jobs. For example, the administrator may want to see all jobs that are either running, suspended or checkpointed on a selection of nodes before [putting the node into maintenance mode](#) or [offlining the node](#). Likewise, if the administrator has detected issues with a node that is running jobs, he/she will have options to delete, suspend/resume, or checkpoint/restart the jobs on the node.

This section covers how the administrator can manage and monitor the jobs on compute nodes without needing to interact directly with PBS Professional. Below are the menu options available to the administrator to act on selected compute nodes:

<sup>9</sup> <http://www.pbsworks.com/documentation/support/PBSProAdminGuide12.pdf>

- [Delete jobs](#)
- [Suspend/Resume jobs](#)
- [Checkpoint/Restart jobs](#)
- [Show running jobs](#)
- [Show suspended jobs](#)
- [Show checkpointed jobs](#)
- [User Groups](#)

## 9.1 Delete jobs

If the administrator needs to delete (a.k.a. terminate) the jobs that are executing on select nodes, this can be accomplished by selecting the node(s), right-clicking, and navigating to *CMU PBS Professional Connector > Delete jobs*.

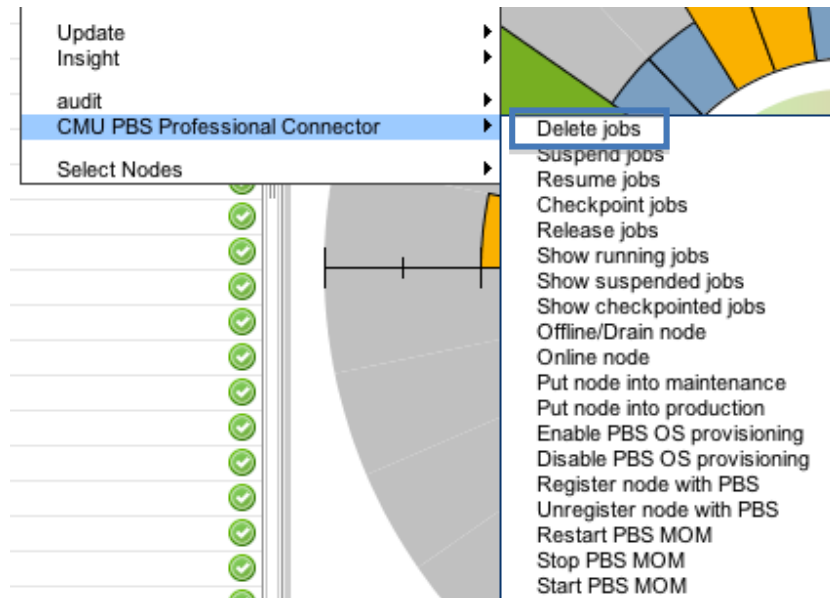


Figure 20: Delete jobs

This menu option will log a message with each job and issue the PBS Professional command to terminate the jobs. The user will see a similar message in the job's STDOUT file:

```
12/10/2012 14:01 CMU Administrator has terminated the job
```

This will help the user understand that the job was terminated by the administrator and did not fail because of something the job script or application did wrong.

## 9.2 Suspend / Resume jobs

Suspending and resuming jobs manually is not frequently done, but the option is provided to the administrator in the event that this becomes necessary. If the site is not using PBS Professional preemption capabilities, the administrator can suspend the jobs on the node(s) to allow a higher priority job to begin executing. When suspending the jobs, the memory

being consumed remains occupied. Therefore, the administrator will need to take care not to over-subscribe the compute resources; otherwise the higher priority job may begin swapping as it competes with the suspended jobs on the node(s).

If the site desires preemption capabilities, it is recommended to configure the PBS Scheduler to handle the site's specific requirements. Refer to the [PBS Professional Administrator's Guide](#)<sup>10</sup>, "Using Preemption" for more details on the capabilities and how to configure this feature.

Suspending jobs on node(s) is accomplished by selecting the node(s), right-clicking, and navigating to *CMU PBS Professional Connector > Suspend jobs*.

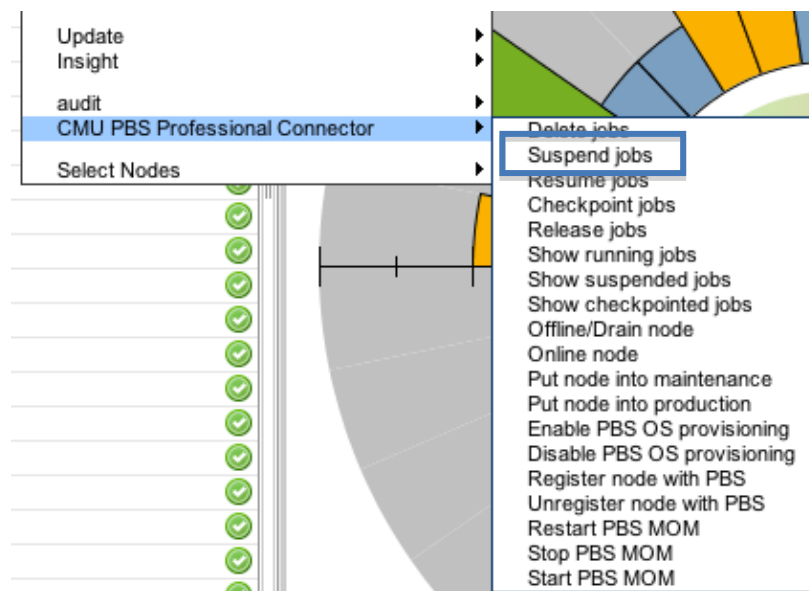


Figure 21: Suspend jobs

Similar to *Delete jobs*, this menu option will log a message with each job and issue the appropriate PBS Professional command. The user will see a similar message in the job's STDOUT file:

```
12/10/2012 14:22 CMU Administrator has suspended the job
```

This will help the user understand that the administrator suspended the job, and when it was suspended.

<sup>10</sup> <http://www.pbsworks.com/documentation/support/PBSProAdminGuide12.pdf>



To resume jobs suspended on the node(s), the administrator will select the node(s), right-click, and navigate to *CMU PBS Professional Connector > Resume jobs*.

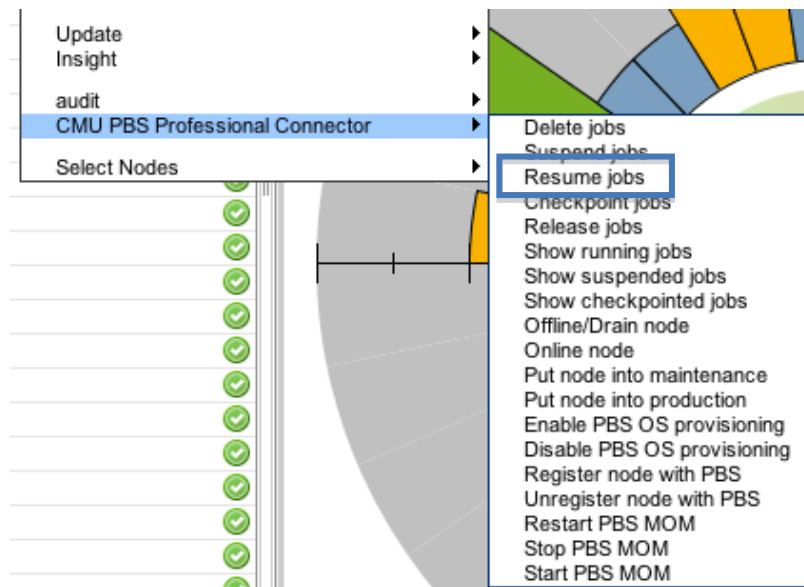


Figure 22: Resume jobs

The user will be informed that the administrator resumed the job, and the PBS Professional command to resume the job will be executed. The user will see a message in the job's STDOUT file:

```
12/10/2012 14:22 CMU Administrator has suspended the job
```

### 9.3 Checkpoint / Restart jobs

Similar to *Suspend/Resume jobs*, the *Checkpoint/Restart jobs* menu options are not used frequently, but when necessary the options are available. Unlike suspending a job, when a job is checkpointed the job releases the memory after it is written to disk.

If the site desires preemption capabilities, it is recommended to configure the PBS Scheduler to handle the site's specific requirements. Refer to the [PBS Professional Administrator's Guide](http://www.pbsworks.com/documentation/support/PBSProAdminGuide12.pdf)<sup>11</sup>, "Using Preemption" for more details on the capabilities and how to configure this feature.

<sup>11</sup> <http://www.pbsworks.com/documentation/support/PBSProAdminGuide12.pdf>

Checkpointing jobs on node(s) is accomplished by selecting the node(s), right-clicking, and navigating to *CMU PBS Professional Connector > Checkpoint jobs*.

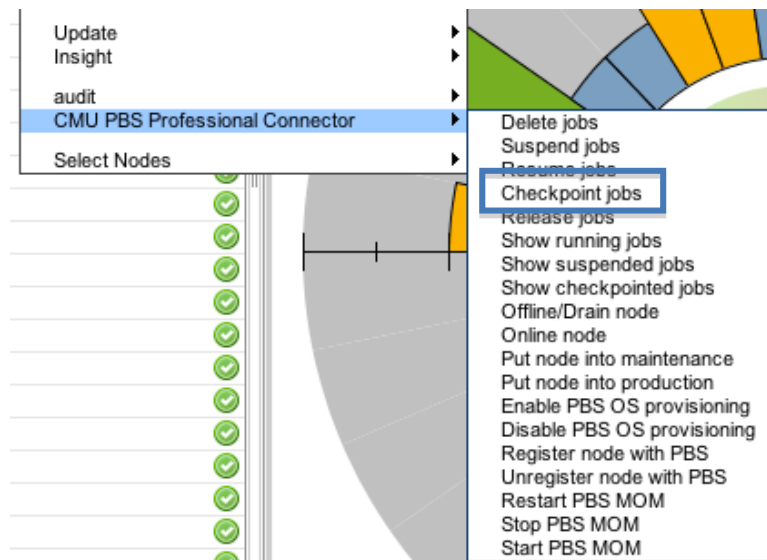


Figure 23: Checkpoint jobs

This menu option will log a message with the job and issue the appropriate PBS Professional command. The user will see a similar message in the job's STDOUT file:

```
12/10/2012 14:53 CMU Administrator has checkpointed the job
```

This will help the user understand that the administrator checkpointed the job, and when it was checkpointed. To release jobs checkpointed on the node(s), the administrator will select the node(s), right-click, and navigating to *CMU PBS Professional Connector > Release jobs*.

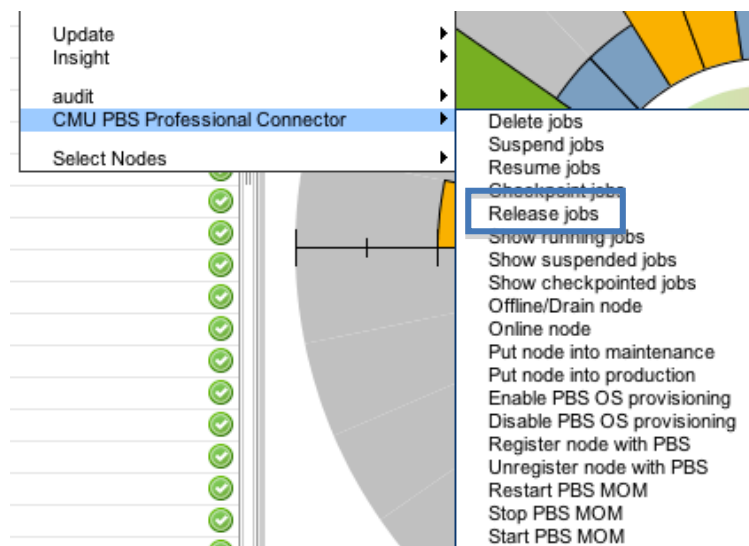


Figure 24: Release jobs

The user will be informed the job is released, and the PBS Professional command to release the job will be executed. The user will see a message in the job's STDOUT file:

12/10/2012 14:22 CMU Administrator has released the job

## 9.4 Show running / suspended / checkpointed jobs

An administrator can get a list of all jobs either running, suspended, or checkpointed on a certain node or set of nodes. Simply select the node(s), right-click, and navigate to *CMU PBS Professional Connector* > *Offline/Show running jobs*. As previously mentioned, an administrator may want to know what jobs are on nodes before [putting the node into maintenance](#) or [offlineing the node](#).

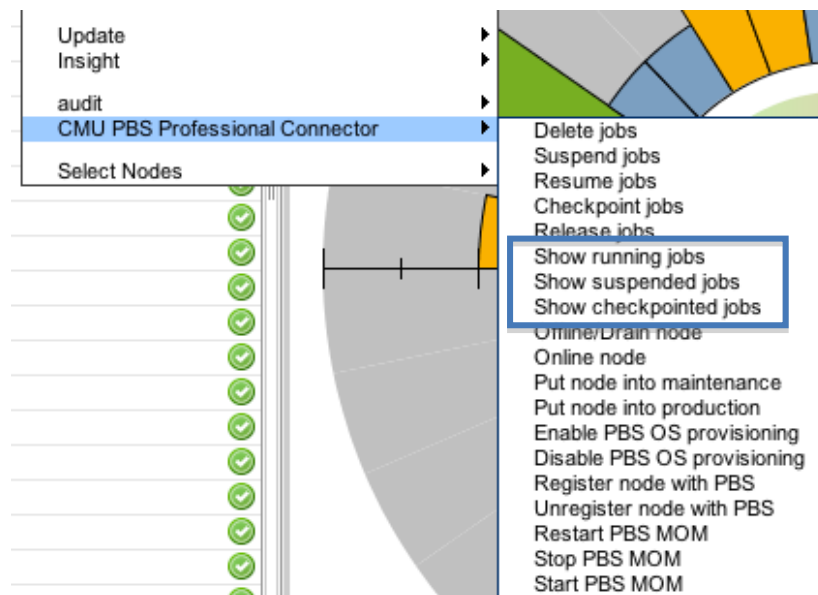


Figure 25: Show running / suspended / checkpointed jobs

## 9.5 User Groups

The CMU PBS Professional Connector takes advantage of CMU's user groups for grouping nodes into objects that make it easier for the administrator or user to visualize what is happening in the system (for instance, all nodes associated to a running job). This is extremely beneficial for the user, who can now monitor the performance of the job and make sure there are no abnormalities in the utilization of the compute node resources.

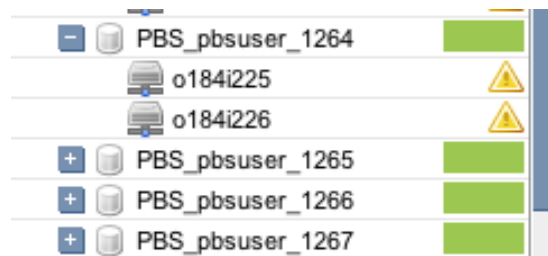


Figure 26: Dynamic User Groups for PBS Professional jobs

The administrator can execute any of the menu options described in [Managing and Monitoring PBS Professional Compute Nodes](#) or [Managing and Monitoring the Jobs](#).

## 10 Online Resources

The CMU PBS Professional Connector exposes all of the PBS Professional online resources via the CMU GUI. This eliminates having to bookmark these resources in a web browser. The following online resources can be obtained by navigating to the CMU menu bar, *Tools* > *CMU PBS Professional Connector*:

- [PBS Professional Documentation](#)
- [PBS Professional User Forum](#)
- [PBS Professional Customer Support](#)

After clicking on one of these menu options, a web browser will open to the appropriate page.

### 10.1 PBS Professional Documentation

The CMU administrator now has easy access to all of the PBS Professional documentation including the Administrator's Guide, User's Guide, Reference Guide, Programmer's Guide, Quick Install, and Release Notes.

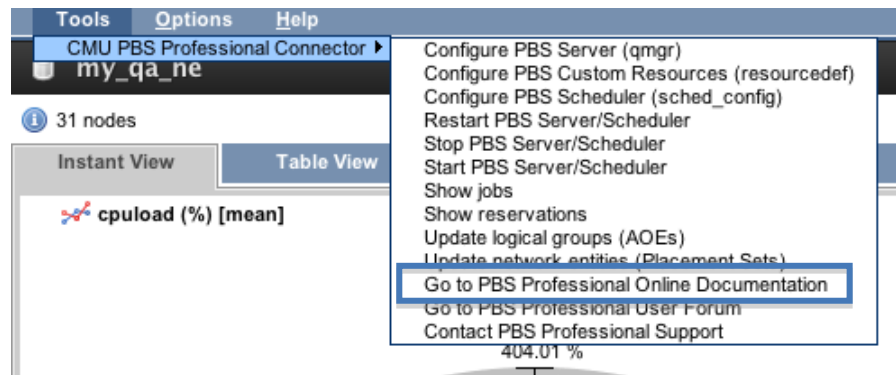


Figure 27: Go to PBS Professional Online Documentation

## 10.2 PBS Professional User Forum

The PBS Professional User Forum allows users to ask fellow PBS Professional users how to setup and configure PBS Professional to achieve desired goals. Users and administrators can also exchange scripts and other tools, if necessary. None of us want to reinvent the wheel, so take advantage of the forum to bounce ideas off each other and spread the knowledge.

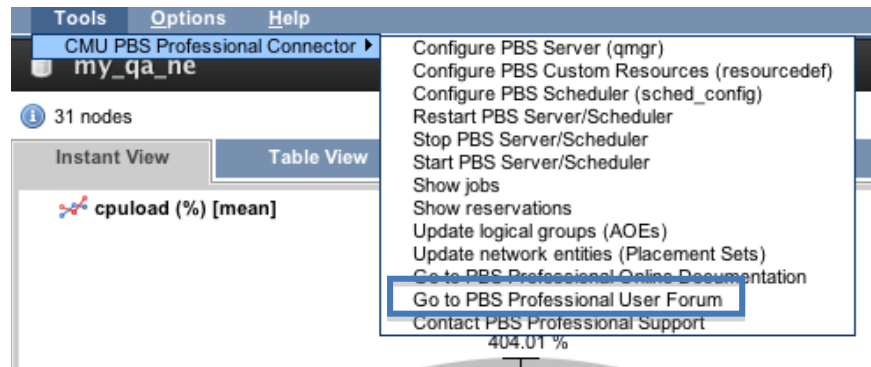


Figure 28: Go to PBS Professional User Forum

## 10.3 PBS Professional Customer Support

If a user or administrator experiences an issue with PBS Professional or would like to request a feature enhancement, this menu option will open the online PBS Professional Customer Support Form. Simply fill out the form and submit. One of Altair's PBS Works Support Team will follow up.

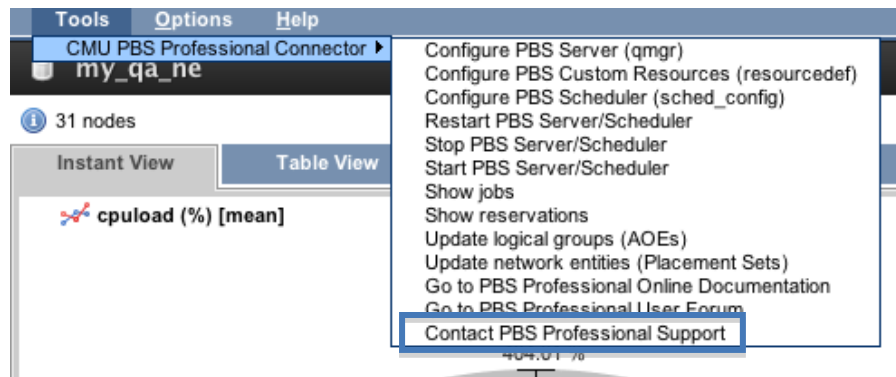


Figure 29: Contact PBS Professional Support

## 11 Conclusion

The Altair CMU PBS Professional Connector provides seamless integration between PBS Professional and HP's Insight CMU, simplifying cluster deployment and management for users and administrators. The Connector enables the administrator to focus on the tasks he or she needs to accomplish, while supporting best practices and without exposing the administrator to a long list of unnecessary details.

## 12 References

- PBS Professional 12.0 Installation Guide:  
<http://www.pbsworks.com/documentation/support/PBSProInstallGuide12.pdf>
- PBS Professional 12.0 Administrator's Guide:  
<http://www.pbsworks.com/documentation/support/PBSProAdminGuide12.pdf>
- PBS Professional 12.0 User's Guide:  
<http://www.pbsworks.com/documentation/support/PBSProUserGuide12.pdf>
- CMU PBS Professional Connector: <http://www.pbsworks.com/cmupbsconnector>
- HP Insight Cluster Management Utility Manual: <http://www.hp.com/go/cmu>

t +1 248.614.2400 • 1820 E. Big Beaver Rd. • Troy, MI 48083-2031 USA • [www.pbsworks.com](http://www.pbsworks.com)

Copyright © 2013 Altair Engineering, Inc. All rights reserved. PBS™, PBS Works™, PBS Professional®, PBS Analytics™, PBS Catalyst™, Copute Manager™, e-BioChem™, e-Compute™, and e-Render™ are trademarks of Altair Engineering, Inc. and are protected under U.S. and international law and treaties. All other marks are the property of their respective owners. This paper is for informational purposes only, and may contain errors; the content is provided as is, without express or implied warranties of any kind.