

SGI® Altix® XE Cluster
Quick Reference Guide

007-5474-002

COPYRIGHT

© 2008 SGI. All rights reserved; provided portions may be copyright in third parties, as indicated elsewhere herein. No permission is granted to copy, distribute, or create derivative works from the contents of this electronic documentation in any manner, in whole or in part, without the prior written permission of SGI.

LIMITED RIGHTS LEGEND

The electronic (software) version of this document was developed at private expense; if acquired under an agreement with the USA government or any contractor thereto, it is acquired as “commercial computer software” subject to the provisions of its applicable license agreement, as specified in (a) 48 CFR 12.212 of the FAR; or, if acquired for Department of Defense units, (b) 48 CFR 227-7202 of the DoD FAR Supplement; or sections succeeding thereto. Contractor/manufacturer is Silicon Graphics, Inc., 1140 East Arques Avenue, Sunnyvale, CA 94085-4602.

TRADEMARKS AND ATTRIBUTIONS

Silicon Graphics, SGI, Altix and the SGI logo are registered trademarks of SGI, in the United States and/or other countries worldwide.

Voltaire is a registered trademark of Voltaire Inc.

Platform Manager is a trademark of Platform Computing Inc.

Scali Manage is a trademark of Platform Computing Inc.

SMC is a registered trademark of SMC Networks Inc.

Linux is a registered trademark of Linus Torvalds.

Unix is a registered trademark of the Open Group.

Windows is a registered trademark of Microsoft Corporation.

InfiniBand is a trademark of the InfiniBand Trade Association.

PBS Professional is a trademark of Altair Grid Technologies, LLC.

All other trademarks mentioned herein are the property of their respective owners.

Record of Revision

Version	Description
-001	March 2008 First publication. Note that substantial content included in this document was originally published in SGI publication 007-4979-00x.
-002	July 2008 Modifications to accommodate the Platform Manager software (formerly Scali Manage) release 5.7 for use with SGI Altix XE clusters plus other miscellaneous updates.

Contents

1. SGI Altix XE Cluster Quick-reference	1
Overview	1
Site Plan Verification	3
Unpacking and Installing a Cluster Rack	3
Booting the XE Cluster	3
SGI Altix XE250 Node Front Controls and Indicators	3
SGI Altix XE240 Head Node Front Controls and Indicators	4
Altix XE310 or XE320 Compute Node Controls and Indicators	6
Cluster Configuration Overview	7
Power Down the Cluster	14
Powering Off Manually	15
Ethernet Network Interface Card (NIC) Guidelines	16
Cluster Management (Head Node) IP Addresses	16
Changing the NIC1 (Customer Domain) IP Address	17
Cluster Compute Node IP Addresses	19
Switch Connect and IP Address	19
Web or Telnet Access to Maintenance Port on the Gigabit Ethernet Switch	20
Web or Telnet Access to the Compute-Traffic Gigabit Ethernet Switch	20
Serial Access to the SMC Switch	21
InfiniBand Switch Connect and IP Address	22
Web or Telnet Access to the InfiniBand Switch	22
Serial Access to the Switch	23
Using the 1U Console Option	24
Installing or Updating Software	25
Accessing BIOS Information	25
Platform Manage Troubleshooting Tips	26
NFS Quick Reference Points	27

Related Publications	28
Third-Party Clustering Documents	30
Customer Service and Removing Parts	31
Contacting the SGI Customer Service Center	32
Cluster Administration Training from SGI	32
2. Administrative Tips and Adding a Node	33
Administrative Tips	34
Start the Platform Manager GUI	36
Head Node Information Screen	37
Adding a Node Starting from the Main GUI Screen.	38
Adding a Cluster Compute Node.	39
Selecting the Server Type	40
Network BMC Configuration	41
Select Preferred Operating System	42
Node Network Configuration Screen	43
DNS and NTP Configuration Screen.	45
NIS Configuration Screen	46
Platform Manager Options Screen	47
Configuration Setup Complete Screen	48
Checking the Log File Entries (Optional)	49
Setting a Node Failure Alarm on Platform Manage	50
3. IPMI Commands Overview	57
User Administration	58
Typical ipmitool Command Line	58
Adding a User to the BMC	58
Configuring a NIC	58
Display a current LAN Configuration	58
Configure a Static IP Address	58
Serial-over-lan Commands	59
Configuring SOL	59
Connecting to Node Console via SOL	60
Deactivating an SOL Connection	60

Sensor commands	60
Displaying all Objects in SDR	60
Displaying all Sensors in the System	60
Displaying an Individual Sensor	60
Chassis Commands	61
Chassis Identify.	61
Controlling System Power	61
Changing System Boot Order	61
SEL Commands.	61

SGI Altix XE Cluster Quick-reference

Overview

Your SGI® Altix® XE cluster system ships with a variety of hardware and software documents in both hard copy and soft copy formats. Hard copy documents are in the packing box and soft copy documents are located on your system hard disk in both

`/opt/sgi/Factory-Install/Docs`

and

`/opt/sgi/Factory-Install/CFG`

Additional third-party documentation may be shipped on removable media (CD/DVD) included with your shipment.

This document is intended as an overview of some of the common operations that system administrators may have to perform to set-up, boot, re-configure (upgrade) or troubleshoot an SGI Altix XE cluster.

The SGI Altix XE cluster is a set of SGI Altix 1U or 2U-high servers (compute nodes), and one or more SGI Altix 2U-high servers (head nodes) networked together and running parallel programs using a message passing tool like the Message Passing Interface (MPI). Systems ordered prior to the second quarter of 2008 generally use SGI Altix 310 servers as compute nodes and SGI Altix XE240 servers as administrative head nodes. Most XE clusters ordered starting in the second half of 2008 use SGI Altix XE320 servers as compute nodes and SGI Altix XE250 servers as administrative head nodes.

Note: Altix XE cluster configurations requiring higher levels of I/O may use XE250 servers as compute nodes. It is possible to have an XE cluster that uses all SGI Altix XE250 servers, or a combination of XE250 and XE320 servers used as compute nodes. In all of these cases the head node will be an SGI Altix XE250 server.

Consult with your SGI support representative before swapping nodes between pre-existing and newer clusters.

The XE cluster is a distributed memory system as opposed to a shared memory system like that used in the SGI Altix 450 or SGI Altix 4700 high-performance compute servers. Instead of passing pointers into a shared virtual address space, parallel processes in an application pass messages and each process has its own dedicated processor and address space. Just like a multi-processor shared memory system, a cluster can be shared among multiple applications. For

instance, one application may run on 16 processors in the cluster while another application runs on a different set of 8 processors. Very large clusters may run dozens of separate, independent applications at the same time.

Typically, each process of an MPI job runs exclusively on a processor. Multiple processes can share a single processor, through standard Linux context switching, but this can have a significant effect on application performance. A parallel program can only finish when all of its sub-processes have finished. If one process is delayed because it is sharing a processor and memory with another application, then the entire parallel program is delayed. This gets slightly more complicated when systems have multiple processors (and/or multiple cores) that share memory, but the basic rule is that a process is run on a dedicated processor core.

These are the primary hardware component types in the rackmounted cluster:

- Head node(s)
(SGI Altix XE240 or XE250 2U-high servers)
- Compute nodes
(SGI Altix XE310 or XE320 1U-high servers)
(SGI Altix XE250 2U-high servers)
- Network interconnect components
(Gigabit Ethernet switches, InfiniBand switches, PCI cards, and cables)

The head node is connected to the interconnect network and also to the “outside world”, typically via the local area network (LAN). The head node is the point of submittal for all MPI application runs in the cluster. An MPI job is started from the head node and the sub-processes are distributed to the cluster compute nodes from the head node. The main process on the head node will wait for the sub-processes to finish. For large clusters or clusters that run many MPI jobs, multiple head nodes may be used to distribute the load.

The compute nodes are identical or compatible computing systems that run the primary processes of MPI applications. These compute nodes are connected to each other through the interconnect network.

The network interconnect components are typically Gigabit Ethernet or InfiniBand. The MPI messages are passed across this network between the processes. This compute node network does not connect directly to the “outside world” because mixing external and internal cluster network traffic could impact application performance.

Site Plan Verification

Ensure that all site requirements are met before you install and boot your system. If you have questions about the site requirements or you would like to order full-size floor templates for your site, contact an SGI site planning representative by e-mail (site@sgi.com).

Additional helpful site planning information can be found in the *SGI Altix XE Cluster Site Planning Guide*, (P/N 007-5456-001).

Unpacking and Installing a Cluster Rack

When your system is housed in a single rack, the cluster components come rackmounted and cabled together and a document describing how to unpack and install the rack should be included with the system. Refer to the *SGI Altix XE System Rack Installation Instructions* (P/N 007-4902-00x). Follow the instructions provided in that manual to safely and properly unpack and install your rack system. Ensure all rack power distribution units are properly plugged in and the circuit breakers are switched to **(On)**. All units within the rack should be connected to power before booting.

Multi-rack cluster systems require connection of special interconnect cables between racks. The *Manufacturing System Diagram* document (P/N 007-4944-00x) shipped with your cluster system describes the inter-rack cable connections. If you have arranged for SGI field personnel to install the system rack(s), contact your service representative. After your cluster rack(s) are installed, refer back to this guide to continue working with your SGI cluster system.

Booting the XE Cluster

Power on any mass storage units attached to your cluster, then press the power button on the front of the head node and let it fully boot. Repeat the process on all the other nodes (compute nodes) in the cluster. See the applicable subsection for your system configuration.

SGI Altix XE250 Node Front Controls and Indicators

The front control panel on the SGI Altix XE250 head node or compute node (see Figure 1-1) has six LED indicators to the left of the power and reset buttons. The LEDs provide critical server related information.

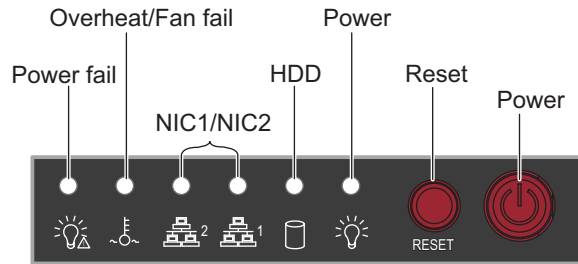
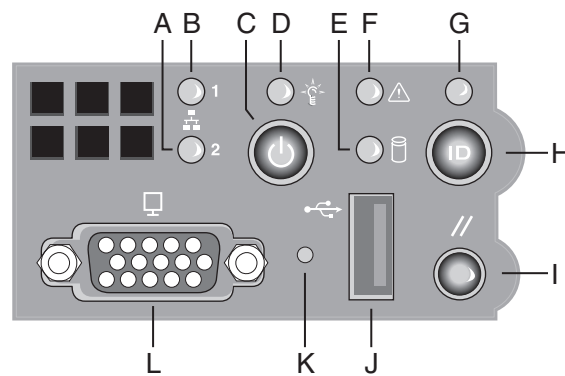


Figure 1-1 SGI Altix XE250 Front Control Panel

- **HDD:** Channel activity for the hard disk drive (HDD). This light indicates drive activity on the node board when flashing.
- **NIC1/NIC2:** Indicates network activity on the LAN1 or LAN2 interconnect when flashing.
- **Overheat/Fan fail:** When the Overheat/Fan Fail LED flashes, it indicates that a fan has failed. When the Overheat/Fan Fail LED is on continuously, it indicates that an overheat condition has occurred, which may be caused by cables obstructing the airflow in the system, covers removed, or the ambient room temperature being too warm.
- **Power Fail:** Indicates power is being supplied to the system’s power supply unit. This LED should normally be illuminated when the system is operating.

SGI Altix XE240 Head Node Front Controls and Indicators



AF000030

Figure 1-2 SGI Altix XE240 Head Node Controls and Indicators

Table 1-1 SGI Altix XE240 Head Node Controls and Indicators

Callout	Feature	Description
A	NIC 2 Activity LED	Continuous green light indicates a link between the system and the network interface card to which it is connected.
B	NIC 1 Activity LED	Blinking green light indicates network interface card 1 activity
C	Power/Sleep button	Powers the system On/Off. Puts the system in an ACPI sleep state.
D	Power/Sleep LED	Constant green light indicates the system has power applied to it. Blinking green indicates the system is in S1 sleep state. No light indicates the power is off or is in ACPI S4 or S5 state.
E	Hard disk drive activity LED	Blinking green light indicates hard disk activity (SAS or SATA). Unlighted LED indicates no hard disk drive activity.
F	System status LED	Solid green indicates normal operation. Blinking amber indicates degraded performance. Solid amber indicates a critical or non-recoverable condition. No light indicates the system POST is running or the system is off.
G	System Identification LED	Solid blue indicates system identification is active. No light indicates system identification is not active.
H	System Identification Button/LED	Press this button once to activate the System Identification LED. Press the button again to de-activate the System Identification LED. Solid blue indicates system identification is active. No light indicates system identification is not active.
I	Reset Button	Reboots and initializes the server.
J	USB 2.0 port	Allows attachment of a USB component to the front of the node.
K	NMI button	Puts the node in a halt-state for diagnostic purposes.
L	Video Port	Allows attachment of a video monitor to the front of the chassis. Note the front and rear video ports cannot be used at the same time.

Altix XE310 or XE320 Compute Node Controls and Indicators

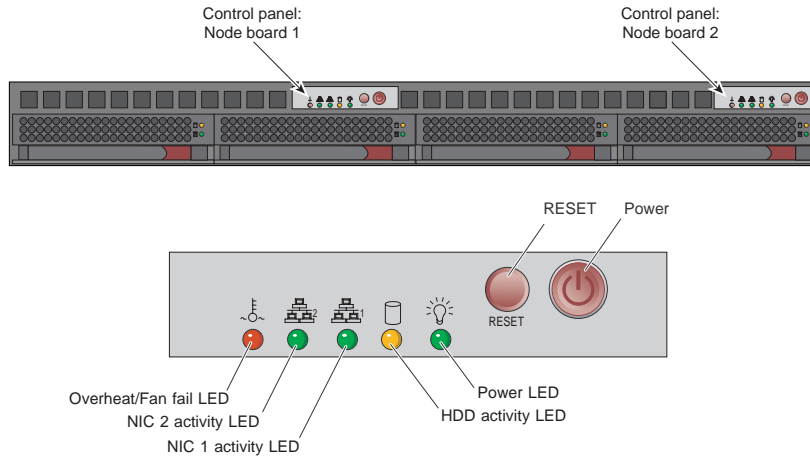


Figure 1-3 SGI Altix XE310/XE320 Compute Node Controls and Indicators

Table 1-2 SGI Altix XE310/XE320 Compute Node Controls and Indicators

Feature	Description
RESET	Press the reset button to reboot only the node board controlled by that control panel.
POWER	Press power button to apply or remove power only to the node board controlled by that control panel. Pressing this button removes the main power but keeps standby power supplied to the node board.
Overheat/Fan fail	When the Overheat/Fan Fail LED flashes, it indicates that a fan has failed. When the Overheat/Fan Fail LED is on continuously, it indicates that an overheat condition has occurred, which may be caused by cables obstructing the airflow in the system or the ambient room temperature being too warm.
NIC2	Indicates network activity on LAN2 when flashing.
NIC1	Indicates network activity on LAN1 when flashing.
HDD	Channel activity for the hard disk drive (HDD). This light indicates SATA drive activity on the node board when flashing.
Power	Indicates power is being supplied to the system's power supply unit. This LED should normally be illuminated when the system is operating.

Cluster Configuration Overview

The following figures are intended to represent some of the general types of cluster configurations used with SGI Altix XE cluster systems.

Note: These configuration drawings are for informational purposes only and are not meant to represent any specific cluster system.

Figure 1-4 on page 8 diagrams a basic Gigabit Ethernet configuration using a single Ethernet switch for node-to-node communication.

Figure 1-5 on page 9 shows an example cluster configuration with a single Gigabit Ethernet switch supporting node-to-node communication between 2U (XE250) servers used as compute nodes.

Figure 1-6 on page 10 illustrates a dual-switch cluster configuration with one switch handling MPI traffic and the other used for basic cluster administration and communication.

Figure 1-7 on page 11 is an example configuration using one Ethernet switch for general administration and one InfiniBand switch for MPI traffic.

Figure 1-8 on page 12 shows a configuration with one Ethernet switch used for administration, one Ethernet switch for NAS, and an Infiniband switch used for handling MPI traffic.

Figure 1-9 on page 13 is an example configuration using a mix of 1U (XE320) compute nodes and 2U (XE250) compute nodes. This configuration uses one Ethernet switch for general administration and one InfiniBand switch for MPI traffic.

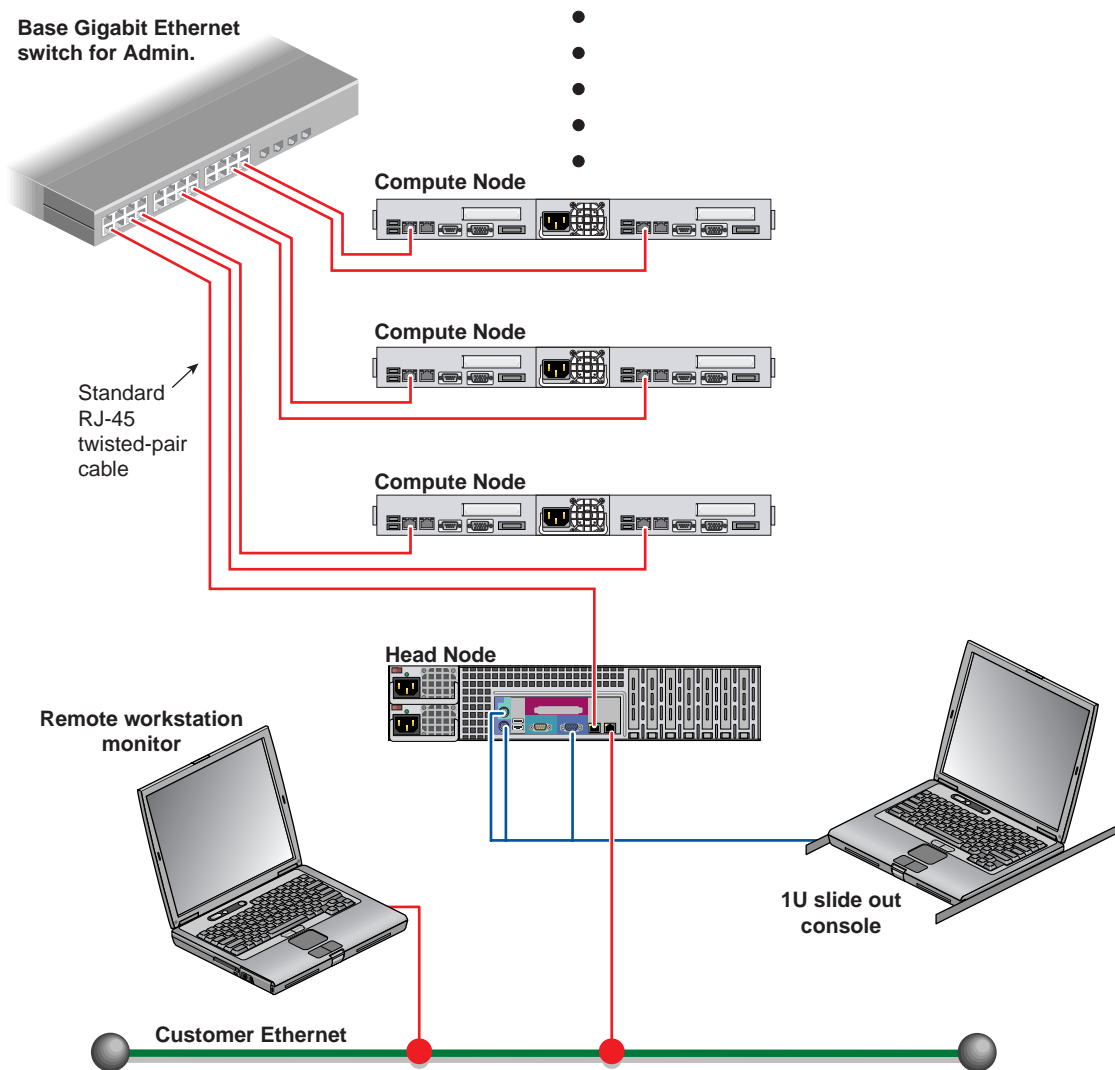


Figure 1-4 Basic Cluster Configuration Example Using a Single Ethernet Switch

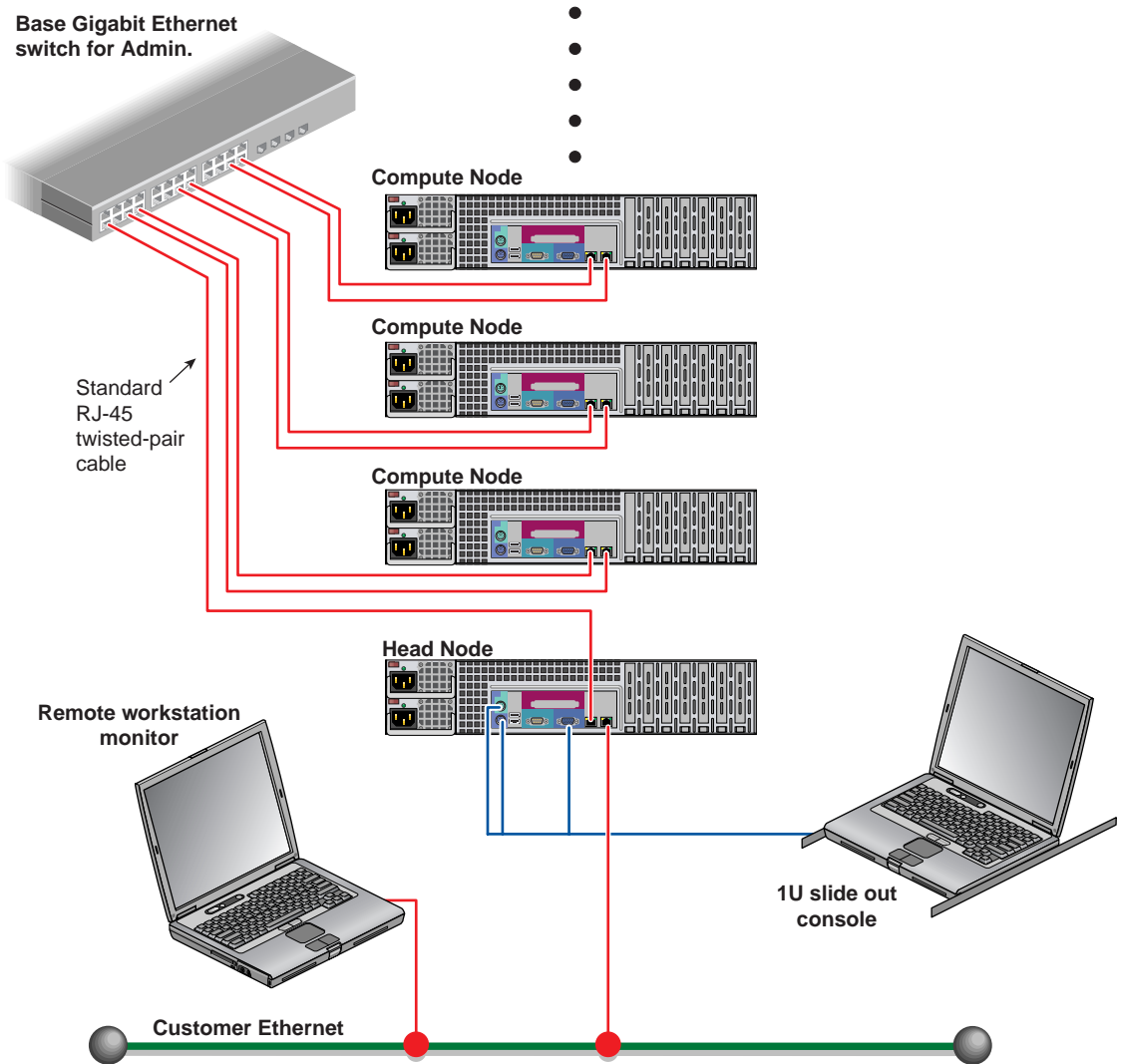


Figure 1-5 Single Ethernet Switch Cluster with 2-U Compute Nodes Example

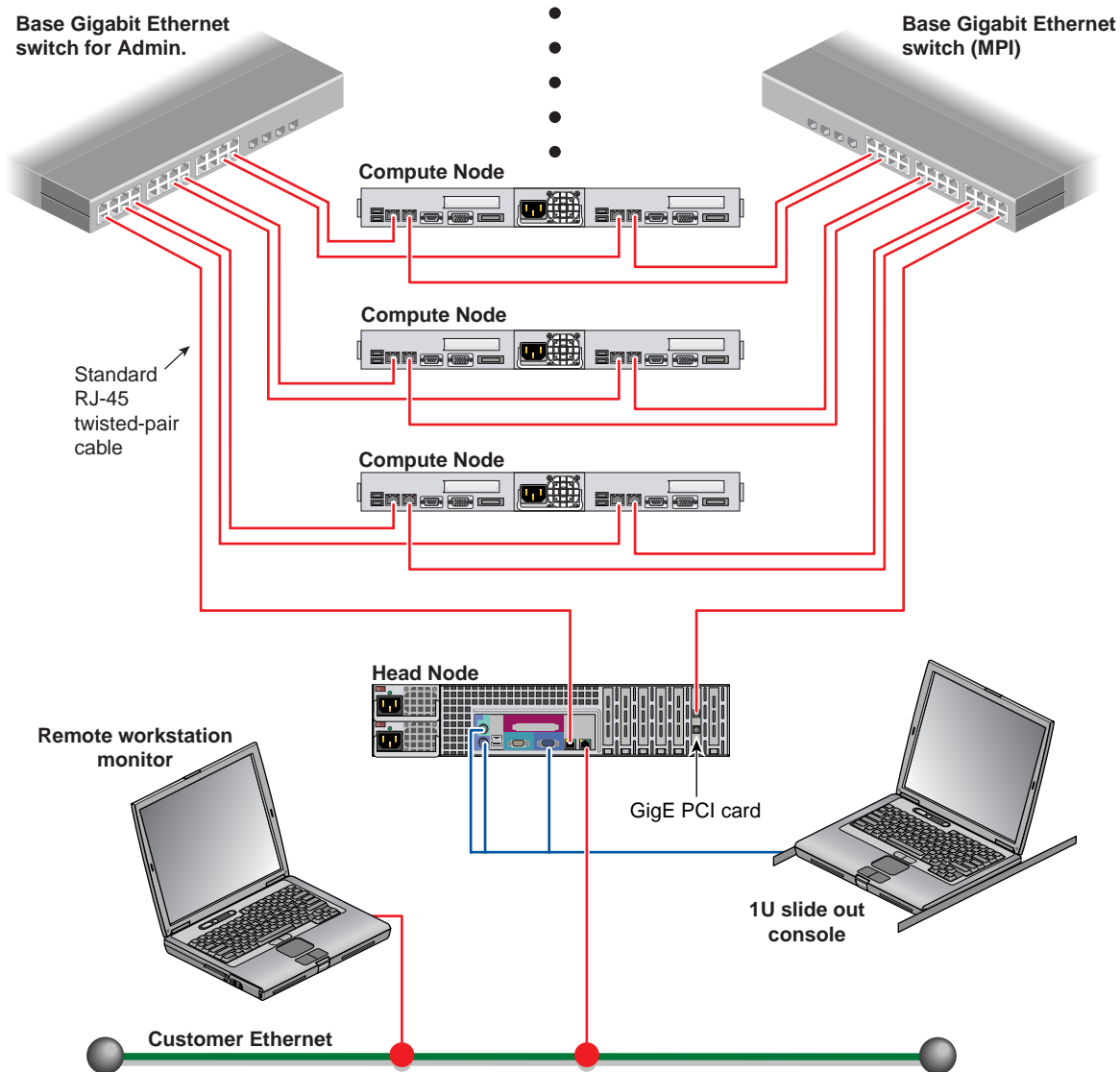


Figure 1-6 Dual-Ethernet Switch Based Cluster Example

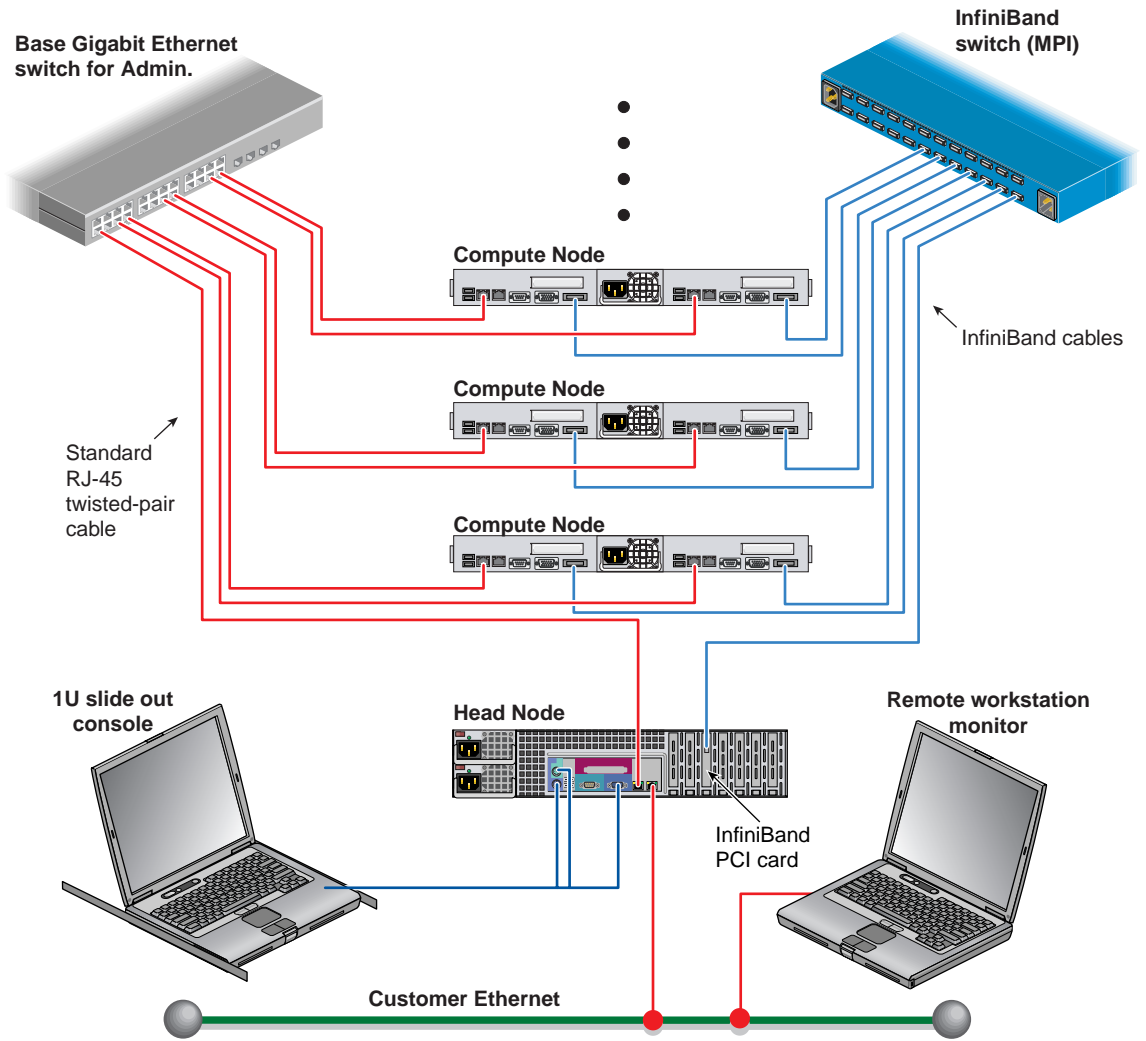


Figure 1-7 Single Ethernet and Single InfiniBand Switch Configuration Example

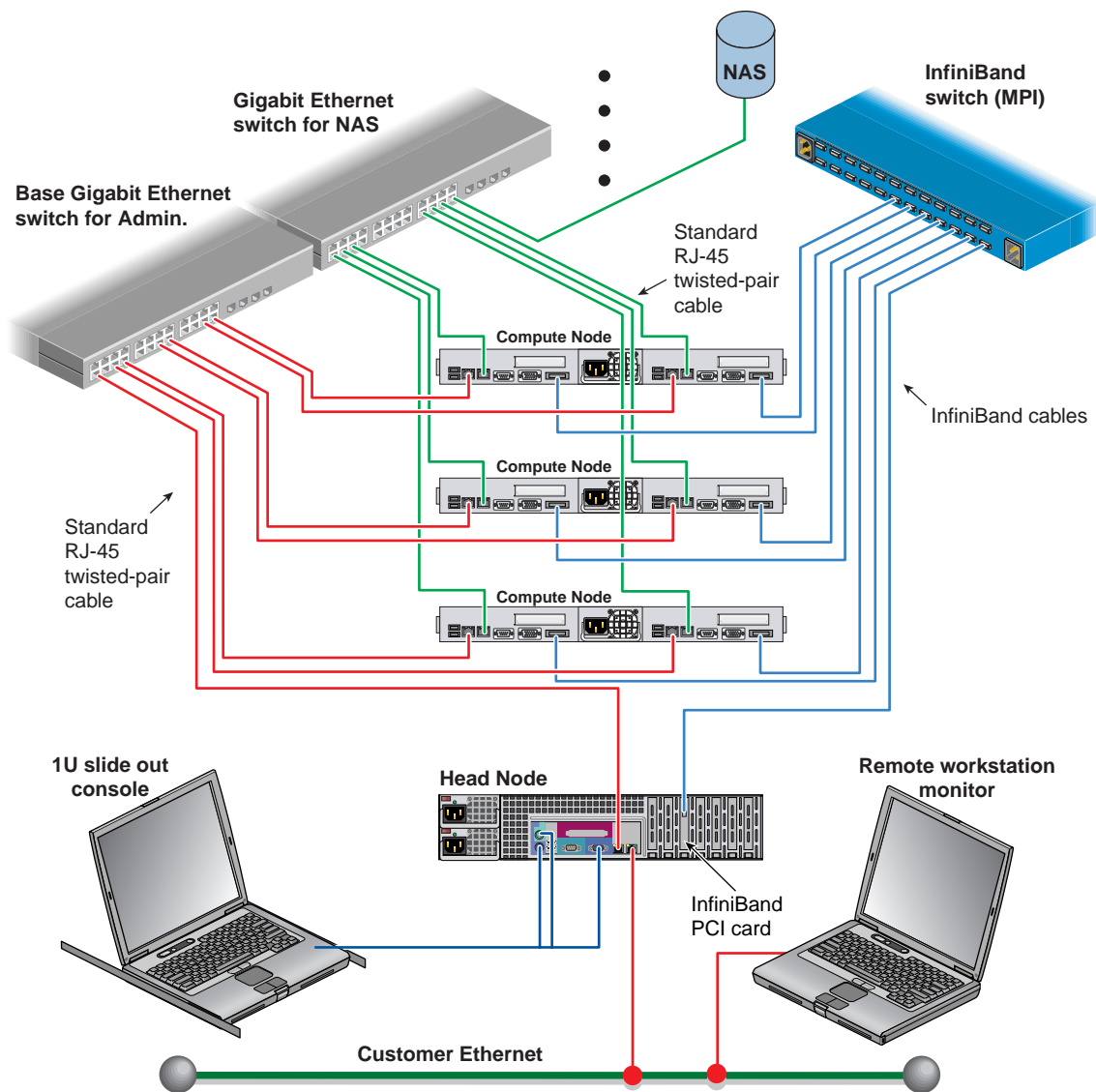


Figure 1-8 Dual Ethernet Plus Infiniband Switch Cluster Configuration Example



Figure 1-9 Single Ethernet Plus InfiniBand in Mixed 1U and 2U Compute Node System

Power Down the Cluster

Note: You can also use the baseboard management controller (BMC) interface to perform power management and other administrative functions. Refer to the *SGI Altix XE310 User's Guide*, publication number 007-4960-00x, for more information about the BMC interface. See the *SGI Altix XE320 User's Guide*, publication number 007-5466-00x for information on its BMC.

Remote power management is done via Platform Manager's GUI or CLI. Remote power command communicate directly with the BMC via the ipmitool. Login to the head node as root, and use the commands to manage the system. Note that the `scash` command is used under the operating system only for remote shutdown or reboots of all or some of the cluster servers.

Typical usage:

```
scash /opt/scali/sbin/power [option] [nodelist <on|off|cycle|status>]
```

Example:

The following command shuts down (halts) cluster nodes 001 through 004:

```
scash -p -n c11n /opt/scali/sbin/power [001-004] -halt
```

If all the system compute servers in the cluster are called "c11", you can use the cluster name to shutdown (halt) them:

```
scash -p -n c11 /opt/scali/sbin/power -halt
```

If your cluster uses the Platform Manage administrative software (release 5.7.x) or later, you can power-off, power-on or power-cycle specific nodes or the entire system using the graphical user interface. Select the cluster or individual nodes or the entire system using the graphical user interface. Select the cluster or individual nodes from the cluster, then select Management Menu>Power Mgt>Power Off/On/Cycle.

The compute nodes can be halted from the Platform GUI by selecting the nodes and choosing "halt system" and "power down" from the System Management menu. As referenced previously, a command line interface (CLI) is also available to power-on/off or check status.

If the operating system is running the selected nodes, they can be halted or rebooted from the Platform Manager GUI by choosing "Shutdown" or "Reboot" from the Management menu. Alternatively, the Platform Manager `scash` parallel command can also be used from the management headnode to halt or reboot a node or nodes. Refer to the *Platform Manage User's Guide* for more information. You must have root privileges to perform these types of tasks.

Powering Off Manually

To power off your cluster system manually, follow these steps:



Caution: If you power off the cluster before you halt the operating system, you can lose data.

1. Shut down the operating system by entering the following command:

```
# init 0
```
2. Press the power button on the head node(s) that you want to power off. You may have to hold the button down for up to 5 seconds. You may power off the nodes in any order.
3. To power off the compute nodes, press the power button (for up to 5 seconds) on the front panel of each unit (refer to Figure 1-2 on page 4).
4. To power off optional storage units in the cluster rack, press the power button(s) on their rear panel to the OFF (●) position.

Ethernet Network Interface Card (NIC) Guidelines

While Ethernet ports are potentially variable in a cluster, the following rules generally apply to the cluster head node:

- The server motherboard's nic1 is always a public IP in the head node.
- The server motherboard's nic2 is always a private administrative network connection.
- Nic3 is always a PCI expansion controller port. It is typically used to handle MPI traffic.

Cluster Management (Head Node) IP Addresses

The primary head node of the cluster (head node1) is also known as the cluster management head node. Head node 1 is where the cluster management software is installed and it has the following technical attributes:

- On-board network interface (nic1) IP address is variable (used as public Ethernet access).

Important: The on-board network interface 1 (nic1) IP address is the factory IP address setting. This setting needs to be changed to reflect the customer domain IP address before connection to the LAN. Refer to the section “Changing the NIC1 (Customer Domain) IP Address” on page 17.

- On-board network interface 2 (nic2) (10.0.10.1) is always used as the management and administration (internal) network port on the primary head node of the cluster.

Note: In the case of a Gigabit Ethernet solution, nic3 is used for MPI traffic. In this case nic3 is on a PCI expansion card.

- The optional Infiniband HCA IP address is 192.168.10.1.
- Board Management Control (BMC) static IP address. The Intelligent Platform Management Interface (IPMI) uses IP address 10.0.30.1 to make controller connections to the other BMCs that exists in the cluster.

It is possible to have additional head nodes on the cluster. Table 1-3 lists examples of the head node port IP address information for more than one head node. Baseboard Management Control routes through nic1 in any additional head nodes added to the cluster. Each fourth octet number in an address iterates by one number as a head node is added.

Table 1-3 Head Node Ethernet Address Listings

Head node number	Internal management IP address nic2	(GigEnet) MPI NAS/SAN option nic3	Infiniband IP address	Baseboard Management Control or IPMI address nic1
1	10.0.10.1	172.16.10.1	192.168.10.1	10.0.30.1
2	10.0.10.2	172.16.10.2	192.168.10.2	10.0.30.2
3	10.0.10.3	172.16.10.3	192.168.10.3	10.0.30.3
4	10.0.10.4	172.16.10.4	192.168.10.4	10.0.30.4

Changing the NIC1 (Customer Domain) IP Address

The “external” IP address assigned to NIC1 must be changed to reflect the new network environment. In addition, a set of network parameters specific to your networking environment need to be specified.

Note: A README file covering this process is also available in:
 /opt/sgi/Factory-Install/Scripts

1. Open the Platform manage GUI using the command **pmgui**
2. Login with password: **sgisgi** - A Platform Manager screen appears.
3. Right click on the “IP Networks” icon and select “Create New Subnet”.
4. Enter the new subnet information and click the “Create New Subnet” box (lower right), then click “OK” to confirm the change.
5. Open the “Independent Servers” tree by clicking on the left arrow. Right click on Platformmanage headnode, then mouse over “Configure...” and select “System Settings”.
6. Right click on “Hostname” and change the content to reflect your network setup.
7. Under “Independent Servers”, Right click on the Platformmanage headnode, then mouse over “Configure...” and select “Network”.
8. Click on the “Network Interfaces” tab, then make the following changes:
 - Click the “Interface Hostname” boxes for eth0, eth1, etc. and change the hostnames

- Click the “IP Address” box for device eth0 and change the IP address
 - Click the “Subnet” box for each network and select (arrow) the new subnet
9. Click in the “Default Gateway” tab. Click on the “Gateway IP Address” and change it to your network address
 10. Click on the “NAT Settings” tab and configure any NAT settings (if applicable). See the Add and Remove buttons (lower right) in the window.
 11. Click on the “DNS Settings” tab to enable, disable and configure DNS.
 - Right click on the DNS entry to make selected appropriate changes
 12. When all changes have been made:
 - Click on the “Network Interfaces” tab
 - Click “Save” (lower right)
 - Click “OK” to confirm
 - Click “Apply Changes” when prompted to “Update configuration files now”
 - Wait for the node configuration task to complete

You may see some errors with the Platform manage GUI. If this occurs, you can troubleshoot the problem by bringing up a Terminal window and running the following commands:

```
/etc/init.d/scance restart
```

You may have to edit the `/etc/hosts` file and manually change any entries with the old hostname/IP to the new one. Then enter the following command:

```
platformmanage-cli reconfigure all
```

Wait for the job to complete and return back to the prompt.

It is best to reboot the headnode also.

When you re-launch the `pmgui` interface, you will have to enter the new hostname, user (root), and password (factory default is **sgisgi**).

Cluster Compute Node IP Addresses

The cluster system can have multiple compute nodes that each use up to three IP address points (plus the Infiniband IP address). As with the head nodes, each fourth octet number in an address iterates by one number as a compute node is added to the list. Table 1-4 shows the factory assigned IP address settings for compute nodes one through four.

Table 1-4 Compute node Ethernet address listings

Compute node number	Management IP address nic1	Infiniband IP address	Gigabit Ethernet solution nic2	Baseboard Management (BMC) or IPMI address nic1
Compute node1	10.0.1.1	192.168.1.1	172.16.1.1	10.0.40.1
Compute node2	10.0.1.2	192.168.1.2	172.16.1.2	10.0.40.2
Compute node3	10.0.1.3	192.168.1.3	172.16.1.3	10.0.40.3
Compute node4	10.0.1.4	192.168.1.4	172.16.1.4	10.0.40.4

Note: The management (internal cluster administration port) IP address and the BMC/IPMI address are shared by the same network interface port (nic1). The circuitry allows the same physical Ethernet port to share two separate IP address references.

Switch Connect and IP Address

The following subsections list the factory IP address for switches that may be used with your cluster.

Web or Telnet Access to Maintenance Port on the Gigabit Ethernet Switch

Your switch(s) setup is configured in the factory before shipment and should be accessible via telnet or a web browser. The switch can be a single switch or a stacked master/slave combination. You can connect to a console directly from the head node through the administration network using telnet.

To access the switch via telnet:

```
telnet 10.0.20.1
```

Login as the administrator:

```
login admin
```

```
passwd: admin
```

Web access would be:

```
http://10.0.20.1
```

Web or Telnet Access to the Compute-Traffic Gigabit Ethernet Switch

The SMC Gigabit Ethernet switch is configured with the IP address shown below when used with a NAS/SAN option or message passing interface (MPI) traffic. The switch can be a single switch or a stacked master/slave combination.

To access the switch via telnet:

```
telnet 172.16.20.1
```

Login as the administrator:

```
login admin
```

```
passwd: admin
```

Web access would be:

```
http://172.16.20.1
```

Serial Access to the SMC Switch

Use of a serial interface to the switch should only be needed if the factory assigned IP address for the switch has been somehow deleted, altered or corrupted. Otherwise, use of the web or telnet access procedure is recommended. To use a serial interface with the switch, connect a laptop, or PC to the switch's console port. Refer to Figure 1-10 for the location of the console port and use the steps that follow for access.

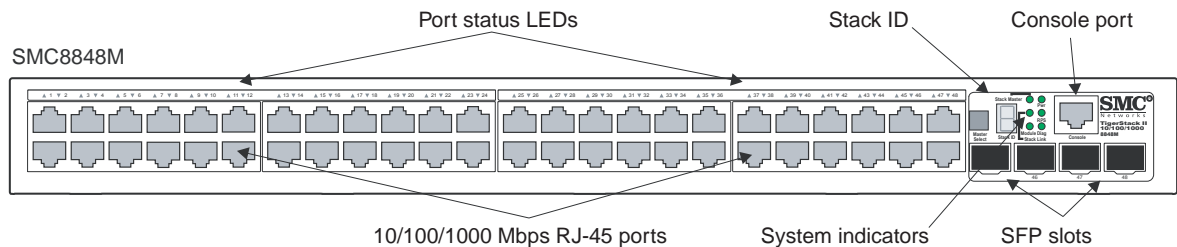


Figure 1-10 SMC Switch Connectors Example

1. Establish a command line interface (CLI) and list the port connection settings:

```
Port Settings
  Bits Per Second=19200
  Data bits=8
  Parity=None
  Stop Bits=1
  Flow Control=none
```

2. In order to verify and save any new settings type the following:

```
console# show running-config (make sure your settings are intact)
console# copy running-config startup-config (it will ask for a file name)
console# file name? startup
```

Note: Any changes made to the switch port settings through the serial interface or Web interface are **not** saved unless the previous steps have been executed.

3. Power cycle the switch by disconnecting and reconnecting its power cable.

InfiniBand Switch Connect and IP Address

The subsection “Web or Telnet Access to the InfiniBand Switch” on page 22 lists the factory IP address settings for your InfiniBand switch or switch “stack” used with the cluster. For clusters with greater than 288 network ports, consult SGI Professional Services for specific IP address configuration information.

Web or Telnet Access to the InfiniBand Switch

Your InfiniBand switch(s) setup is configured in the factory before shipment and should be accessible via telnet or a web browser.

Note: There might be only one managed InfiniBand switch when multiple InfiniBand switches are used in blocking configurations.

To access the managed InfiniBand switch via telnet:

telnet 10.0.21.1

Login as the administrator:

login **admin**

passwd: **123456**

Web access would be:

http://10.0.21.1

javaws (java Webstart) is required for use of the InfiniBand fabric GUI.

SLES 9 service pack 3 location of javaws is:

/usr/java/j2re1.4.2_12/javaws/javaws

SLES 10 location of javaws is: /usr/bin/javaws or

/usr/java/jre1.5.0_11/javaws

Serial Access to the Switch

You should connect a Voltaire serial cable (either DV-9 to DB-9 or DB-9 to DB-9) that comes with the 24-port switch, from a PC/laptop directly to the switch for serial access. Use of a serial interface to the switch should only be needed if the factory assigned IP address for the switch has been somehow deleted, altered or corrupted. Otherwise, use of the web or telnet access procedure is recommended.

Note: For Voltaire switches 96-ports or larger, always use a DB-9 serial cable.

To interface with the switch, use the connected laptop or other PC to:

1. List the port connection settings. Default settings are:

```
Port Settings
Bits Per Second=38400
Data bits=8
Parity=None
Stop Bits=1
Flow Control=xon/xoff
```

2. Click “ok” if the settings are acceptable. In the serial interface window on the PC, press enter several times until the `ISR-xxxx login:` prompt displays, then enter the following:
`ISR-xxxx login: admin`
`ISR-xxxx login: Password: 123456`
`ISR-xxxx> enable`
`ISR-xxxx> Password: voltaire`

3. Set up the network for your InfiniBand switch cluster configuration using the following information and the IP reference provided in “Web or Telnet Access to the InfiniBand Switch” on page 22.

Enter the following commands to set up the network:

```
ISR-xxxx# config
ISR-xxxx(config)# interface fast
ISR-xxxx(config-if-fast)# ip-address-fast set [10.0.21.1] 255.255.0.0
ISR-xxxx(config-if-fast)# broadcast-fast set 10.0.255.255
ISR-xxxx(config-if-fast)# exit
ISR-xxxx(config)# exit
ISR-xxxx# reset software (This reboots the 24-port InfiniBand switch)
For a 96-port or larger switch:
```

4. ISR-xxxx# **reload software**
ISR-xxxx# **fast-interface show** (This command lists the IP address)
5. Power cycle the switch by disconnecting its power cable from the power connector, and then plug it back in.

Using the 1U Console Option

The SGI optional 1U console is a rackmountable unit that includes a built-in keyboard/touchpad, and uses a 17-inch (43 cm) LCD flat panel display of up to 1280 x 1024 pixels. The 1U console attaches to the headnode using PS/2 and HD15M connectors or to a KVM switch (not provided by SGI). The 1U console is basically a “dumb” VGA terminal, it cannot be used as a workstation or loaded with any system administration program. The 27-pound (12.27kg) console automatically goes into sleep mode when the monitor cover is closed down.

Note: While the 1U console is normally plugged into the head node on the cluster, it can be connected to any node in the system for terminal access purposes.

Installing or Updating Software

Platform Manage offers a mechanism to upload and install software across the cluster. This upload and installation process requires that the software installation be in RPM format. Tarball software distributions can be installed across a cluster. Please see the Platform `scarcop` (cluster remote copy) and the `scash` (cluster remote shell) commands in the *Platform Manage User's Guide*.

Instructions for installing software options or uploading additional software for your cluster using the Platform GUI are covered in Chapter 3 of the *Platform Manage User's Guide*.

Your integrated cluster also comes with an NFS mounted filesystem. The head node exports a `/data1` directory. Each compute node mounts this exported filesystem on `/cluster`. This can be used as a mechanism to install software across the cluster as well.

Customers with support contracts needing BIOS or Firmware updates should check the SGI Supportfolio Web Page at:

<https://support.sgi.com/login>

Accessing BIOS Information

BIOS Setup Utility options are used to change server configuration defaults. You can run BIOS Setup with or without an operating system being present. You can enter and start the BIOS Setup Utility after you apply power to a head node or compute node (with a console attached) and the Power-On Self Test (POST) completes the memory test. During the POST, the following prompt displays may be encountered:

SGI Altix XE240 Head Node:

Press <F2> to enter SETUP

SGI Altix XE310 Compute Node:

Press to enter SETUP

SGI Altix XE250 Head Node or Compute Node:

Press to enter SETUP

SGI Altix XE320 Compute Node:

Press to enter SETUP

Note: The DEL key and F2 key work only if the proper ACSII terminal settings are in place. Many Linux distributions default to varied ASCII settings. In the case of the SGI Altix XE310 or XE320 compute node, or the Altix XE250 head node, the DEL key should always generate an “ACSII DEL”. If it does not, type Ctrl-Backspace to enter BIOS setup menu.

Important: The BIOS comes preconfigured with the SGI recommended settings. Changes to any of the BIOS settings can impact the performance of your cluster.

If CMOS/NVRAM becomes corrupted on an SGI Altix XE240 server, you may see other prompts but not the <F2>prompt:

```
Warning: CMOS checksum invalid
Warning: CMOS time and date not set
```

Under these circumstances, you should contact your SGI service representative. Refer to the:

SGI Altix XE250 User's Guide (P/N 007-5467-00x)

SGI Altix XE320 User's Guide (P/N 007-5466-00x)

SGI Altix XE240 User's Guide (P/N 007-4873-00x)

SGI Altix XE310 User's Guide (P/N 007-4960-00x)

for specific information about BIOS settings.

Platform Manage Troubleshooting Tips

This section describes some general guidelines as well as emergency procedures.

Whenever a Platform cluster parameter is changed, it is necessary to apply the configuration. This can be done either through the GUI (Provisioning > Apply All Configuration Changes) or via CLI: `platformmanage-cli reconfigure all`. Changes can be made in batches and then applied all at once.

There are situations when the GUI does not reflect the cluster configuration properly. Restarting the GUI may solve this problem.

In rare cases the Platform product enters an inconsistent state. In this state it shows abnormal behavior and refuses to take any input. In this case try to re-initialize the head node via `/etc/init.d/scance restart`.

This command must be run on the head node. If this does not change Platform's state, then you should reboot the head node. This should ensure that Platform will be in a consistent state. If you have trouble that is more hardware related, see "Customer Service and Removing Parts" on page 31.

NFS Quick Reference Points

The cluster head node exports an NFS, compute nodes import NFS on the head node. The cluster comes with a pre-configured NFS mount. The headnode exports the `/data` filesystem. The compute nodes mount head node `/data1` on `/cluster`.

You need to execute the following commands to export a filesystem via NFS from the head node:

```
# platformmanage-cli addnfsexport <head_node> <filesystem>
# /etc/init.d/scance restart
```

To import this filesystem on a particular compute node:

```
# platformmanage-cli addremotefs compute_node nfs
head_node:/filesystem mount_point
# platformmanage-cli reconfigure compute_node
```

If the compute nodes need to mount filesystems located outside the cluster, then NAT must be enabled on the head node. You need to execute the following commands on the head node:

```
# platformmanage-cli addnatSERVICE head_node ethernet_dev
# /etc/init.d/scance restart
```

Now you can access nodes outside the cluster from your compute nodes.

To mount a remote filesystem residing outside the cluster on a particular compute node you need to do the following:

```
# platformmanage-cli addremotefs compute_node nfs
external_node:/filesystem mount_point
# platformmanage-cli reconfigure compute_node
```

Related Publications

The following SGI system documents may be useful or necessary while configuring and operating your Altix XE cluster system:

- *Manufacturing Audit Checklist* (P/N 007-4942-00x)
This document contains the network configuration/validation switch IP addresses for your system.
- *Manufacturing Configuration Summary* (P/N 007-4943-00x)
- *Manufacturing System Diagram* (P/N 007-4944-00x)
- *SGI Altix XE System Rack Installation Instructions* (P/N 007-4902-00x)
This manual provides instructions for moving, unpacking and installing a rack at your location.
- *SGI Altix XE250 User's Guide* (P/N 007-5467-00x)
This guide covers general operation, configuration, and servicing of the Altix XE250 head node(s) within the Altix XE cluster.
- *SGI Altix XE320 User's Guide* (P/N 007-5466-00x)
This guide covers general operation, configuration, and servicing of the SGI Altix XE320 compute modules within the SGI Altix XE cluster.
- *SGI Altix XE310 User's Guide* (P/N 007-4960-00x)
This guide covers general operation, configuration, and servicing of the SGI Altix XE310 compute modules within the SGI Altix XE cluster.
- *SGI Altix XE240 User's Guide* (P/N 007-4873-00x)
This guide covers general operation, configuration, and servicing of the Altix XE240 head node(s) within the Altix XE cluster.
- *Porting IRIX® Applications to SGI® Altix® Platforms: SGI ProPack™ for Linux®*, Publication Number 007-4674-00x
This manual may be useful for any administrator porting existing IRIX applications to Linux.
- `/opt/sgi/Factory-Install/` on the head node
This online directory on the head node contains useful information specific to your system configuration.

- *SGI Altix® Systems Dual-Port Gigabit Ethernet Board User's Guide*, Publication Number 007-4326-00x

This guide describes the two versions of the optional SGI dual-port Gigabit Ethernet board, shows you how to connect the boards to an Ethernet network, and explains how to operate the boards. You can use the dual-port Gigabit Ethernet board to replace or supplement the built-in Ethernet network adapters in your system.

- *SGI 10-Gigabit Ethernet Network Adapter User's Guide*, Publication Number 007-4669-00x

This guide describes the SGI 10-Gigabit (Gbit) Ethernet network adapter. It shows you how to connect the adapter to an Ethernet network and explains how to operate the adapter. The manual also provides information on how to performance tune this high-speed interface card.

- *SGI ProPack 5 for Linux Start Here*, Publication Number 007-4837-00x

- *SGI ProPack 6 for Linux Start Here*, Publication Number 007-5482-00x

These documents provide information about the SGI ProPack for Linux release, including the major features of the release, flowcharts of disk partitions and file configurations. Select the release used on your cluster.

- *Guide to Administration, Programming Environments, and Tools Available on SGI Altix XE Systems*, Publication Number 007-4901-00x.

For internet access to these SGI documents, see the SGI Technical Publications Library at:
<http://docs.sgi.com>

Various formats are available. This library contains the most recent and most comprehensive set of online books, release notes, man pages, and other customer information.

Third-Party Clustering Documents

The SGI Altix XE Cluster is provided in different configurations and not all the third-party documents listed here will be applicable to every system. Note that Linux is the only operating system supported with the SGI Altix XE cluster.

- *Voltaire® HCA-4x0 User Manual*, Publication Number 399Z00007

This document describes how to install and configure the Voltaire HCA-4x0 Linux based InfiniBand software and the hardware interface cards needed for a cluster. The HCA-4x0 User Manual covers a number of technical features that are not supported with the Altix XE cluster products. These include SDP, CM and GSI.

For a complete listing of supported and non-supported host stack components, see the release notes section on: “Voltaire InfiniBand Stack for SGI Altix Systems”. The release notes for the stack is included in the Voltaire software CD.

- *Voltaire® ISR 9024S/D Installation Manual*, Publication Number 399Z00002

This manual covers unpacking, installation, configuration, and power-up information as well as basic troubleshooting information for the 24-port InfiniBand Switch Routers.

- *Voltaire ISR 9288/ISR 9096 Installation Manual*, Publication Number 399Z40000

This manual covers unpacking, installation, configuration, and power-up information as well as basic troubleshooting information for the 96-port and 288-port InfiniBand Switch Routers.

- *Voltaire® Switch User Manual*, Publication Number 399Z00038

This manual covers technical descriptions and use of different switches; ISR9024, ISR9600 and ISR9288. Only the ISR9024, ISR9600 and ISR9288 are supported options for the Altix XE cluster product. Information on managing switch usage via the Voltaire Device Manager CLI or GUI, and basic diagnostics are also covered in this manual.

- *SMC® TigerStack™ II Gigabit Ethernet Switch Installation Guide*

Use this guide to install or configure the SMC8824M 24-port switch and the SMC8848M 48-port switch. These are stackable 10/100/1000 base-T switches used with Altix XE clusters.

- *SMC® TigerStack™ II Gigabit Ethernet Switch Management Guide*

Use this guide to manage the operations of your SMC8824M 24-port switch or SMC8848M 48-port switch.

- *Platform Manage™ User's Guide*

This document provides an overview of a Platform managed system in terms of instructions for building a Platform Manager administered cluster system. Configuration guidelines for hardware and software are covered along with instructions on use and general management of the cluster system.

- *Platform Manage™ Installation Guide*

This document describes the Platform Manage software installer that helps the user with installation of the OS, Platform Manager software and third-party applications that are installed as RPMs.

Customer Service and Removing Parts

If you are experiencing trouble with the cluster and determine that a replacement part will be needed, please contact your SGI service representative using the information in “Contacting the SGI Customer Service Center” on page 32. Return postage information is included with replacement parts.

Removal and replacement of the hardware components that make up the head and compute nodes within the cluster are fully documented in:

- *SGI Altix XE250 User's Guide* (P/N 007-5467-00x)
- *SGI Altix XE320 User's Guide* (P/N 007-5466-00x)
- *SGI Altix XE240 User's Guide*, Publication number 007-4873-00x
- *SGI Altix XE310 User's Guide*, Publication number 007-4960-00x

These documents can be used to help troubleshoot node-level hardware problems and are included as soft copy (PDF format) on the head node's system disk at:

`/opt/sgi/Factory-Install/Docs`

You can also down-load these documents via internet, from the SGI publications library at: <http://docs.sgi.com>

If you need to replace a node within your cluster, go to the SGI Supportfolio web page: <https://support.sgi.com/login>

Contacting the SGI Customer Service Center

To contact the SGI Customer Service Center, call 1-800-800-4SGI, or visit:
<http://www.sgi.com/support/customerservice.html>
From outside the United States contact your local SGI sales office.

To reach SGI for other purposes, use the following contact information:

SGI Corporate Office

1140 E. Arques Avenue

Sunnyvale, CA 94085

(650) 960-1980

<http://www.sgi.com>

North America +1 800.800.7441

Latin America +55 11.5185.2860

Europe +44 118.912.7500

Japan +81 3.5488.1811

Asia Pacific +1 650.933.3000

Cluster Administration Training from SGI

SGI offers customer training classes covering all current systems, including clusters. If you have a maintenance agreement in place with SGI, contact SGI Customer Education at 1-800-361-2621 for information on the time, location and cost of the applicable training course you are interested in. Or, go to the following URL site for more education information:
<http://www.sgi.com/support/custeducation/>

Customers with support contracts can also obtain information from:
<https://support.sgi.com/login>

Administrative Tips and Adding a Node

This chapter provides general administrative information as well as basic instructions on starting and using the Platform Manage GUI to add a node in a Platform managed cluster. For information on using the Platform Manage command line interface to add a node, refer to the *Platform Manage User's Guide*. Basic information on starting Platform Manager, administrative passwords and factory installed files and scripts are covered in the first section of this chapter, “Administrative Tips” on page 34.

Add a node to the cluster using the following sections and accompanying screen snaps:

- “Start the Platform Manager GUI” on page 36
- “Head Node Information Screen” on page 37
- “Adding a Node Starting from the Main GUI Screen” on page 38
- “Adding a Cluster Compute Node” on page 39
- “Selecting the Server Type” on page 40
- “Network BMC Configuration” on page 41
- “Select Preferred Operating System” on page 42
- “Node Network Configuration Screen” on page 43
- “DNS and NTP Configuration Screen” on page 45
- “NIS Configuration Screen” on page 46
- “Platform Manager Options Screen” on page 47
- “Configuration Setup Complete Screen” on page 48
- “Checking the Log File Entries (Optional)” on page 49

Set a node failure “alarm” using the information in:

- “Setting a Node Failure Alarm on Platform Manage” on page 50

Administrative Tips

Root password and administrative information includes:

- Root password = **sgisgi** (head node and compute nodes)
- Ipmitool user/password info: User = **admin** Password = **admin**

Refer to Table 1-3 on page 17 and Table 1-4 on page 19 for listings of the IPMI IP addresses for nodes.

Note: The `ipmitool` command syntax for SGI Altix XE240 head nodes (run via the Platform Manage head node):

```
-ipmitool -I lanplus -o intelplus -H ip address command
```

The `ipmitool` command syntax for SGI Altix XE250 head nodes and SGI Altix XE310 and XE320 compute nodes

(run via the Platform Manage head node):

```
-ipmitool -I lanplus -o supermicro -H ip address command
```

SGI Altix XE systems that run SLES10, release 4, can use the following service:

```
checkconfig ipmi on  
/etc/init.d/ipmi start
```

Following is the `ipmitool` command syntax for running directly from the head node (or any node) on itself. (Note the absence of the `-I -o` and `-H` options):

```
ipmitool command
```

The Platform Manage installer directory (`/usr/local/Platform###`) is the location of the code used to install Platform Cluster management Software.

The `Factory-Install` directory is located on the head node server at `/usr/local/Factory-Install`. The `/Factory-Install` directory contains software files that support the cluster integration and many files and scripts that may be helpful, including:

Under `/usr/local/`

`/Factory-Install/Apps`

Platform, ibhost, Intel compilers, MPI runtime libraries, ipmitool, etc.

`/Factory-Install/ISO:`

CD ISO images of the base OS for installing Platform Cluster Manage software

`/Factory-Install/Docs`

Cluster documentation manuals (Platform, PBS Professional, Voltaire, SMC, SGI)

`/Factory-Install/Firmware`

Voltaire HCA and Voltaire switch firmware files, etc.

`/Factory-Install/CFG`

Cluster configuration files

`/Factory-Install/Scripts`

Miscellaneous utility scripts

Start the Platform Manager GUI

Login to the Platform Manager interface as root, the factory password is **sgisgi**. Use your system name and log in as root. Refer to Figure 2-1 for an example.

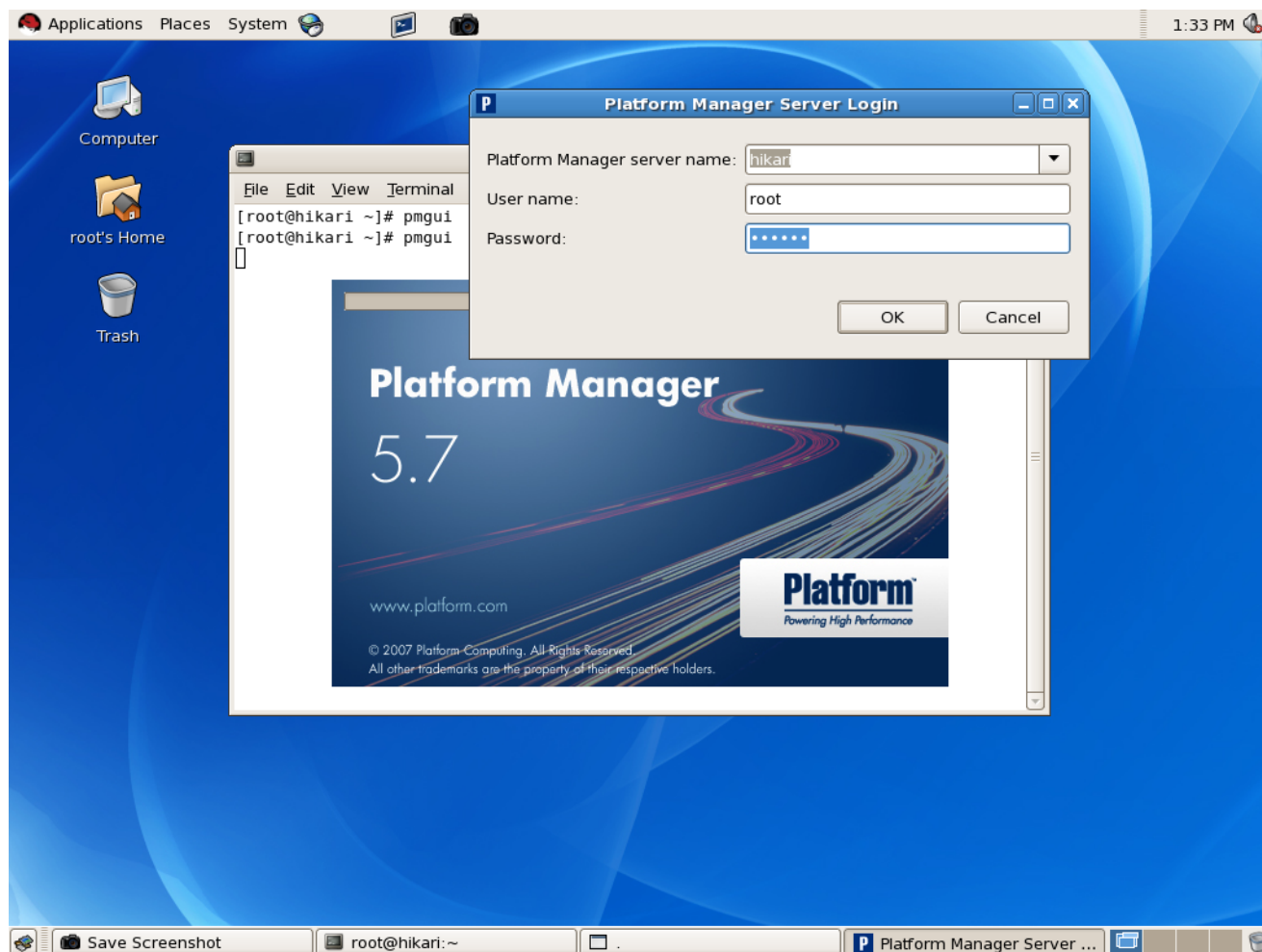


Figure 2-1 Example Starting Screen for the Platform Manager GUI

Head Node Information Screen

You can view and confirm the head node information from the main GUI screen. Click on the node icon (cl1n001 in the example below) for name and subnet information on your cluster head node.

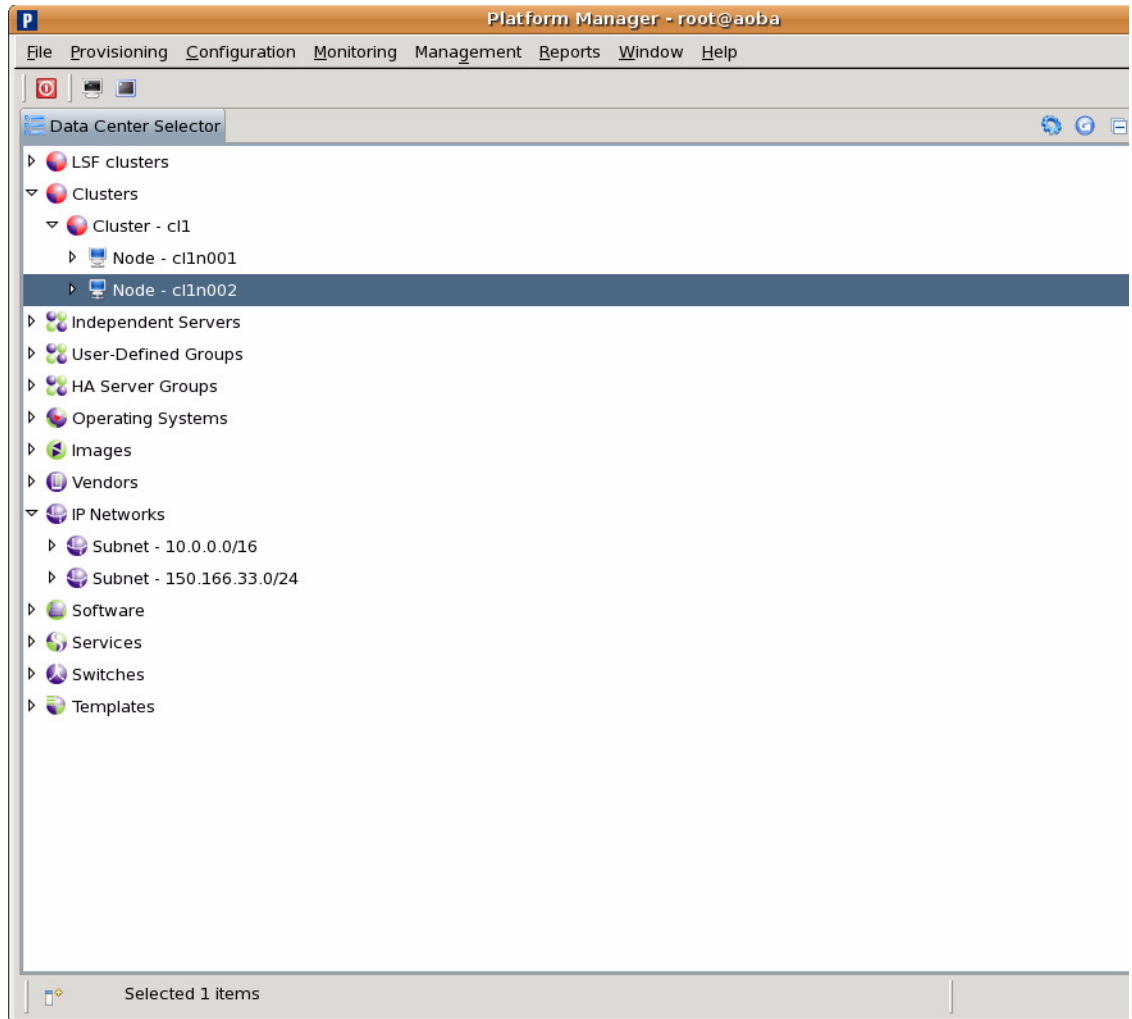


Figure 2-2 Head Node Information Screen Example

Adding a Node Starting from the Main GUI Screen

Add a node when you need to upgrade. To add a cluster node, open the Clusters tree by clicking the right mouse button. Move your cursor over the cluster tree (cluster cl1 in the example screen), and click the right mouse button. Then click the left mouse button on “New” in the popup window. Refer to Figure 2-3.

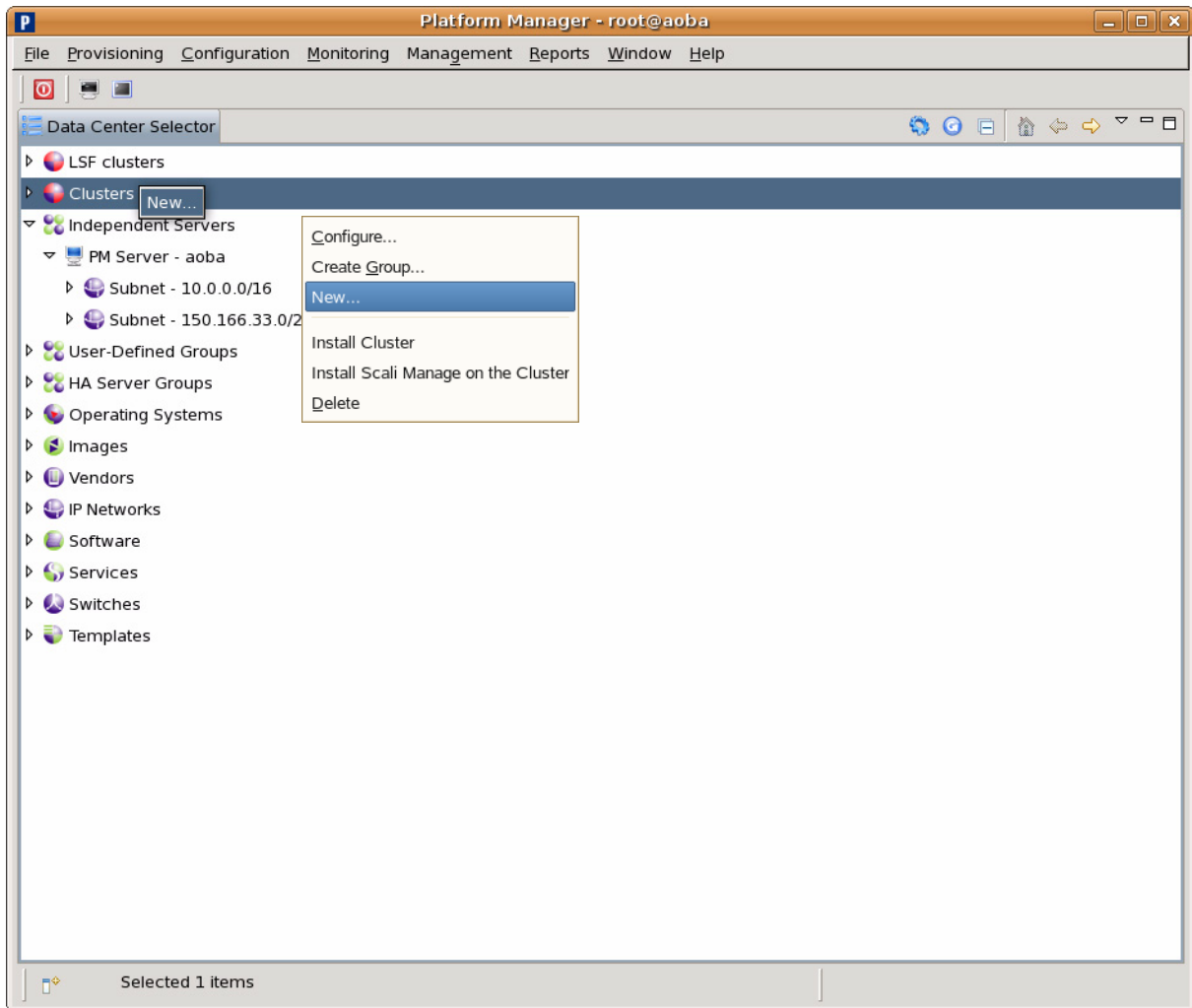


Figure 2-3 Platform Manage Main Screen Selections Example

Adding a Cluster Compute Node

These steps should only be taken if the cluster needs to be upgraded or re-created. Select the option “Extend existing cluster” and provide the number of new servers (2 in the example). Then select the “Cluster Name” (cl1 in the example). Select the server template and click “Next” to move to the following screen.

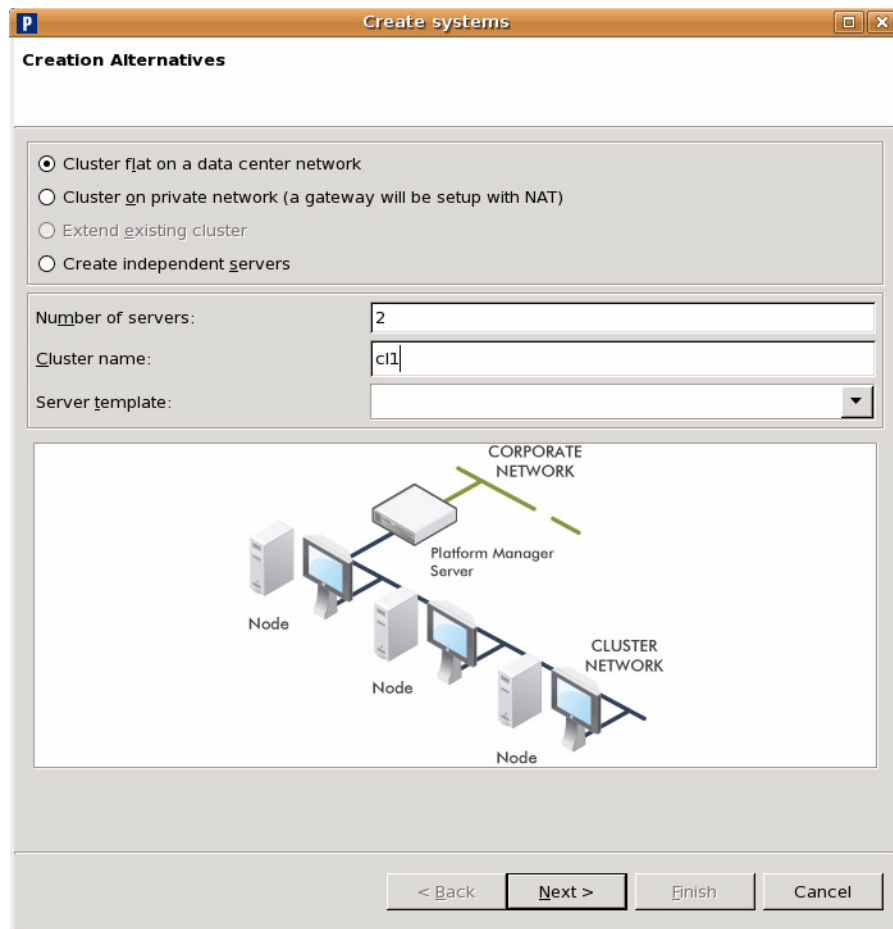


Figure 2-4 New Cluster Node Selection Example

Selecting the Server Type

Click on “Edit” to bring up the “Node Hardware Configuration” network panel.

Scroll down the menu and select the server type you are adding. Then enter the BMC user ID (**admin**) and the password (**admin**).

Create systems

Node Hardware Configuration

Server Model on Nodes

Server Models

- SGI Altix ICE Controller
- SGI Altix XE 210
- SGI Altix XE 240
- SGI Altix XE 250
- SGI Altix XE 310**
- SGI Altix XE 320
- SGI Virtio VM 200

BMC Configuration

Enable BMC configuration

Host name preview: cl1n001-bmc ... cl1n002-bmc

IP addresses: 10.0.40.1-10.0.40.2

Subnet: 10.0.0.0/16

Edit...

BMC user ID: admin

BMC password:

Confirm BMC password:

< Back Next > Finish Cancel

Figure 2-5 Node Server Type Selection Screen Example

Network BMC Configuration

Click on the “Edit” button. Assign the new BMC IP address, stepping and BMC host name. Click OK when the appropriate information is entered. Click “Next” to move to the following screen.

The screenshot displays the BMC Network Configuration interface. The main window, titled "Create systems" and "Node Hardware Configuration", shows a list of server models under "Server Models on Nodes". The "SGI Altix XE 310" model is selected. Below this, the "BMC Configuration" section includes a checked "Enable BMC configuration" option and a preview of host names, IP addresses, and subnets. A modal dialog titled "Configure BMC network" is overlaid, showing the following configuration:

- Network Configuration:**
 - Subnet: 10.0.0.0/16
 - IP address: 10.0.40.1
 - IP stepping: 1
- BMC Host Names:**
 - Prefix: cl1n
 - Start number: 1
 - Zero padding: 3
 - Suffix: -bmc

A preview of BMC host names is shown as "cl1n001-bmc ... cl1n002-bmc". The dialog includes "OK" and "Cancel" buttons. The main window also features navigation buttons: "< Back", "Next >", "Finish", and "Cancel".

Figure 2-6 BMC Network Configuration Screen Example

Select Preferred Operating System

Select the option to provision the new node's operating system. Enter the **sgisgi** factory password or whatever new password may have been assigned. Click "Next" to move to the following screen.

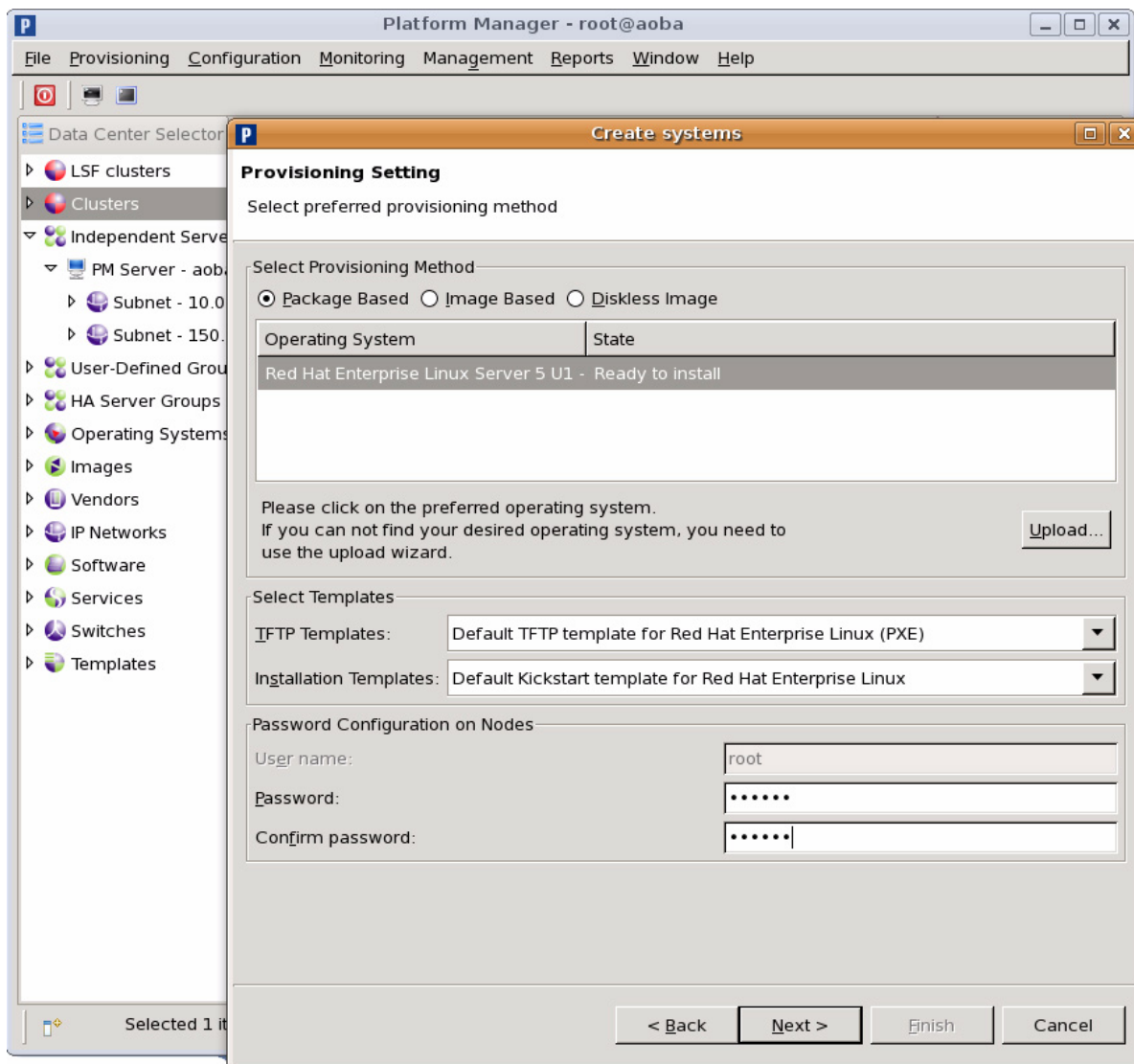


Figure 2-7 Provisioning the Preferred Operating System Screen Example

Node Network Configuration Screen

Use this screen to assign Ethernet 0 (eth0) as your network interface port. Fill in the additional information as it applies to your local network. Click “OK” to continue.

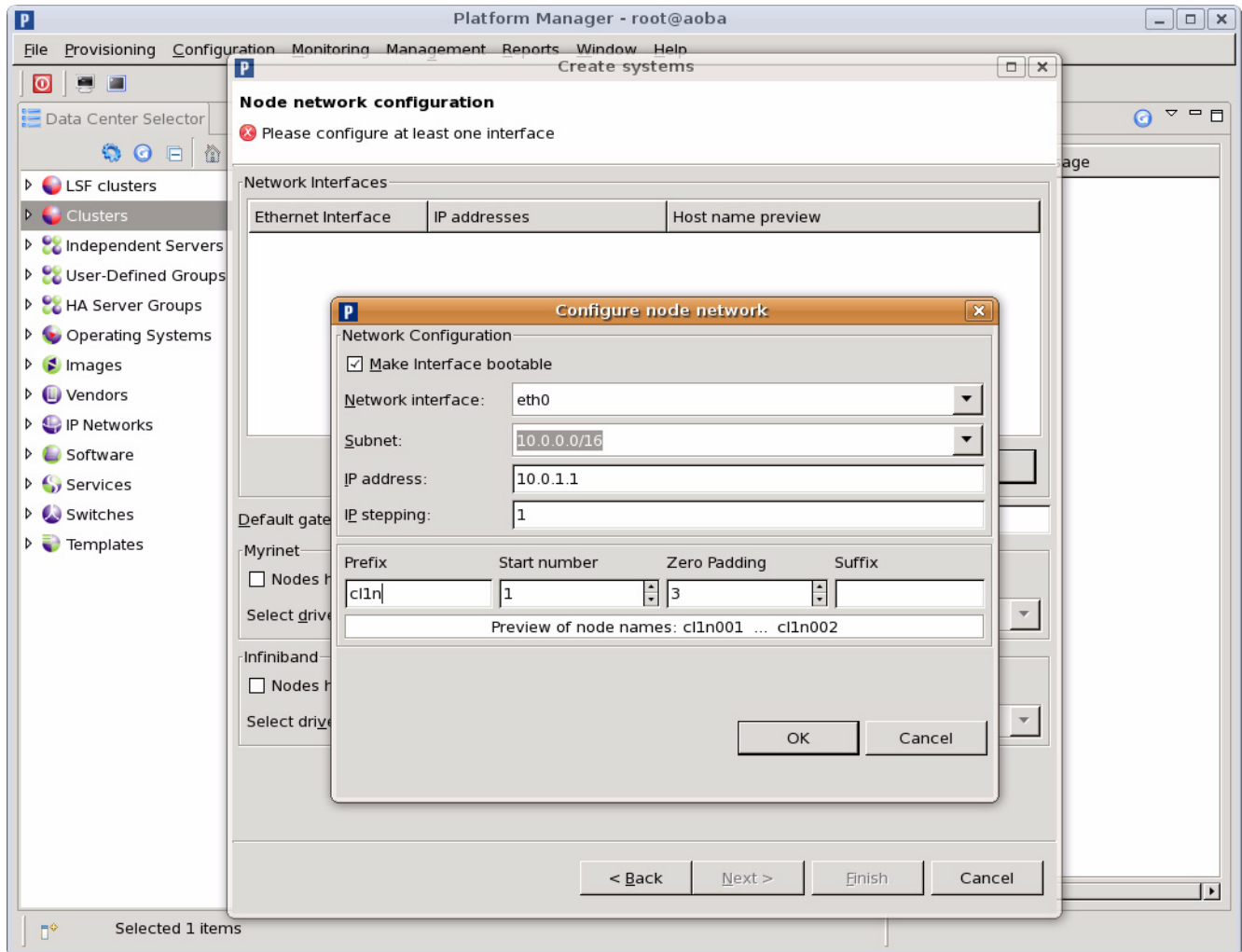


Figure 2-8 Node Network (Ethernet 0) Screen Example

Enter the default gateway information (refer to the example in Figure 2-9) and select “Next” to continue.

The screenshot shows a window titled "Create systems" with a sub-tab "Node network configuration". The window contains a table for "Network Interfaces" with the following data:

Ethernet Interface	IP addresses	Host name preview
eth0	10.0.1.1-10.0.1.2	cl1n001...cl1n002

Below the table are buttons for "Remove", "Edit...", and "Add...". The "Default gateway:" field is set to "10.0.10.2". There are sections for "Myrinet" and "Infiniband", each with a checkbox for "Nodes have ... cards installed" (both unchecked) and a dropdown menu for "Select driver software to install" (both set to "Install manually"). At the bottom are buttons for "< Back", "Next >", "Finish", and "Cancel".

Figure 2-9 Default Gateway Example Screen

DNS and NTP Configuration Screen

This screen extracts the name server numbers for use with the system configuration files. Enter the appropriate domain name enabling information or disable the function by un-checking the box. Click “Next” when complete.

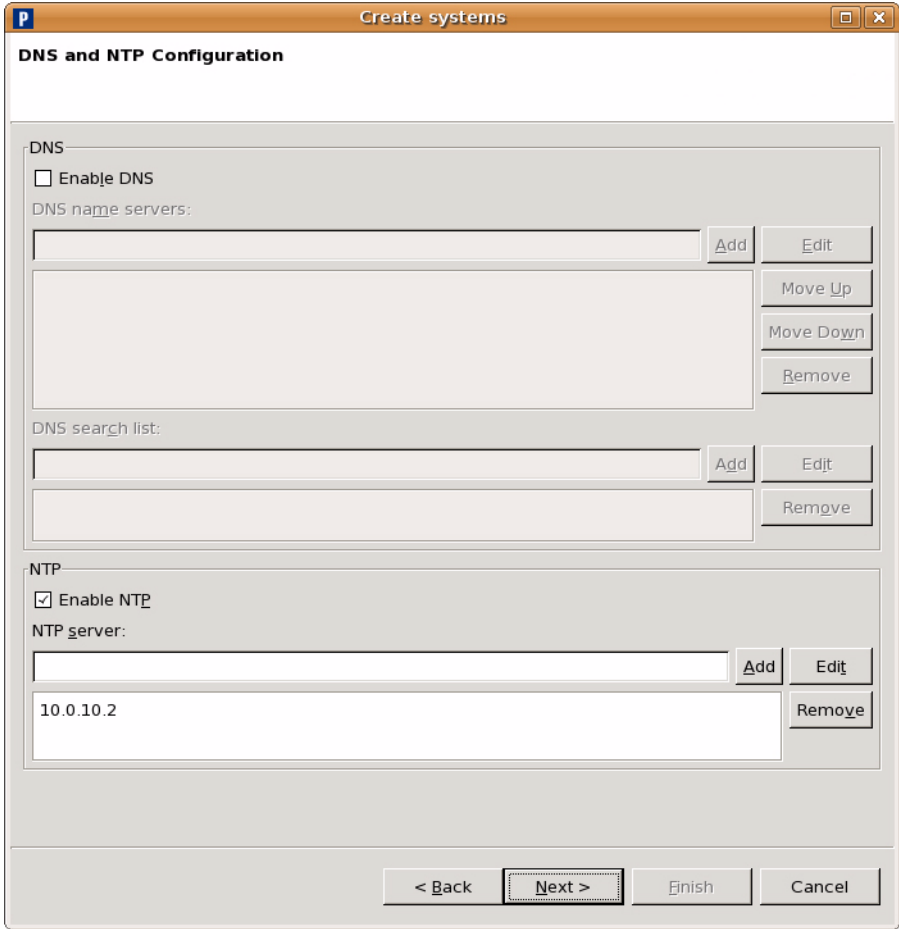
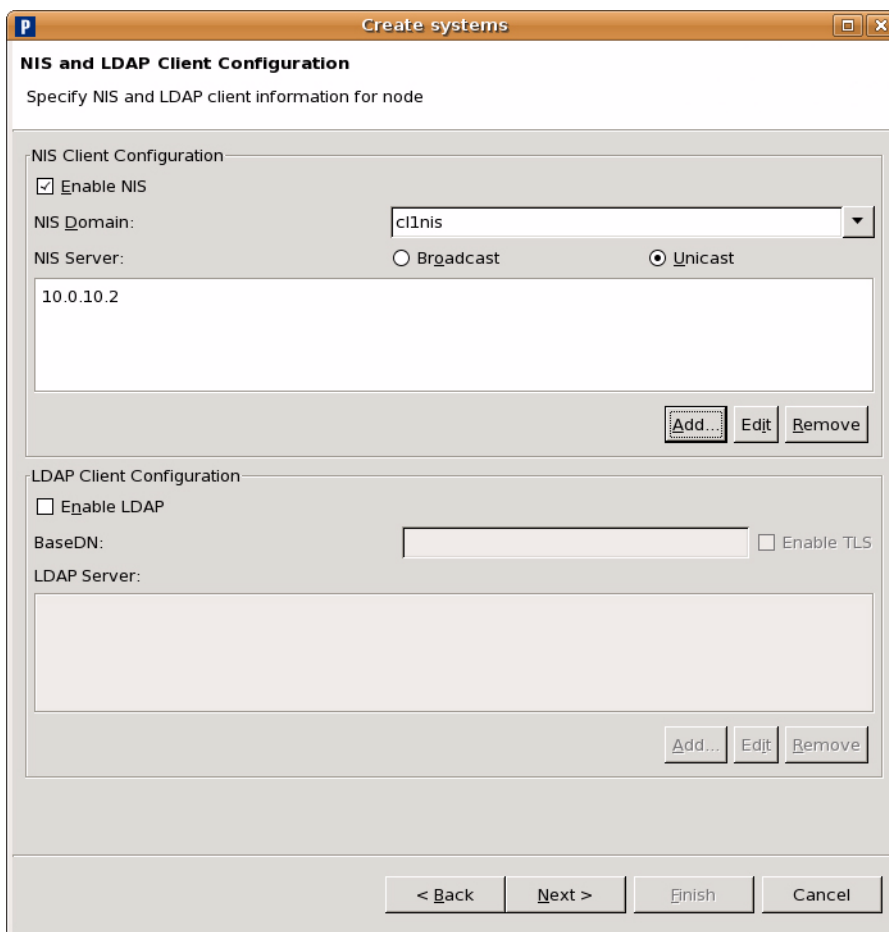


Figure 2-10 DNS and NTP Configuration Screen Example

NIS Configuration Screen

This screen allows you to specify, enable or disable a Network Information Service (NIS) for the new node. Assign your domain name (see Figure 2-11 for an example) and click “Next” to go to the following screen.



The screenshot shows a window titled "Create systems" with a sub-header "NIS and LDAP Client Configuration". Below the sub-header is the instruction "Specify NIS and LDAP client information for node".

The "NIS Client Configuration" section includes:

- Enable NIS
- NIS Domain: cl1nis (selected from a dropdown menu)
- NIS Server: Broadcast Unicast
- 10.0.10.2 (text input field)
- Buttons: Add..., Edit, Remove

The "LDAP Client Configuration" section includes:

- Enable LDAP
- BaseDN: (empty text input field)
- Enable TLS
- LDAP Server: (empty text input field)
- Buttons: Add..., Edit, Remove

At the bottom of the window are navigation buttons: < Back, Next >, Finish, and Cancel.

Figure 2-11 NIS Configuration Screen Example

Platform Manager Options Screen

This screen provides the options shown, including installation of MPI, your software version, monitor options and more. Click “Next” to move to the following screen.

Platform Manager Options

Install MPI Connect

Install MPI

Software Version : Scali MPI Connect 5.6.0, RHEL5 - x86_64

Monitor Options

Install vendor-specific monitoring software

Software Version:

Check to install	Affects	Software
<input checked="" type="checkbox"/>	Affects no nodes	PP5SP5-rhel-general
<input checked="" type="checkbox"/>	Affects no nodes	PP5SP5-rhel-headnode

If the desired software is not in the list, you must upload through the upload wizard [Upload...](#)

< Back Next > Finish Cancel

Figure 2-12 Platform Manage Options Screen Example

Configuration Setup Complete Screen

This screen allows you to install the operating system and Platform Manager immediately, or store the configuration for later use. Click “Finish” after you make your selection.

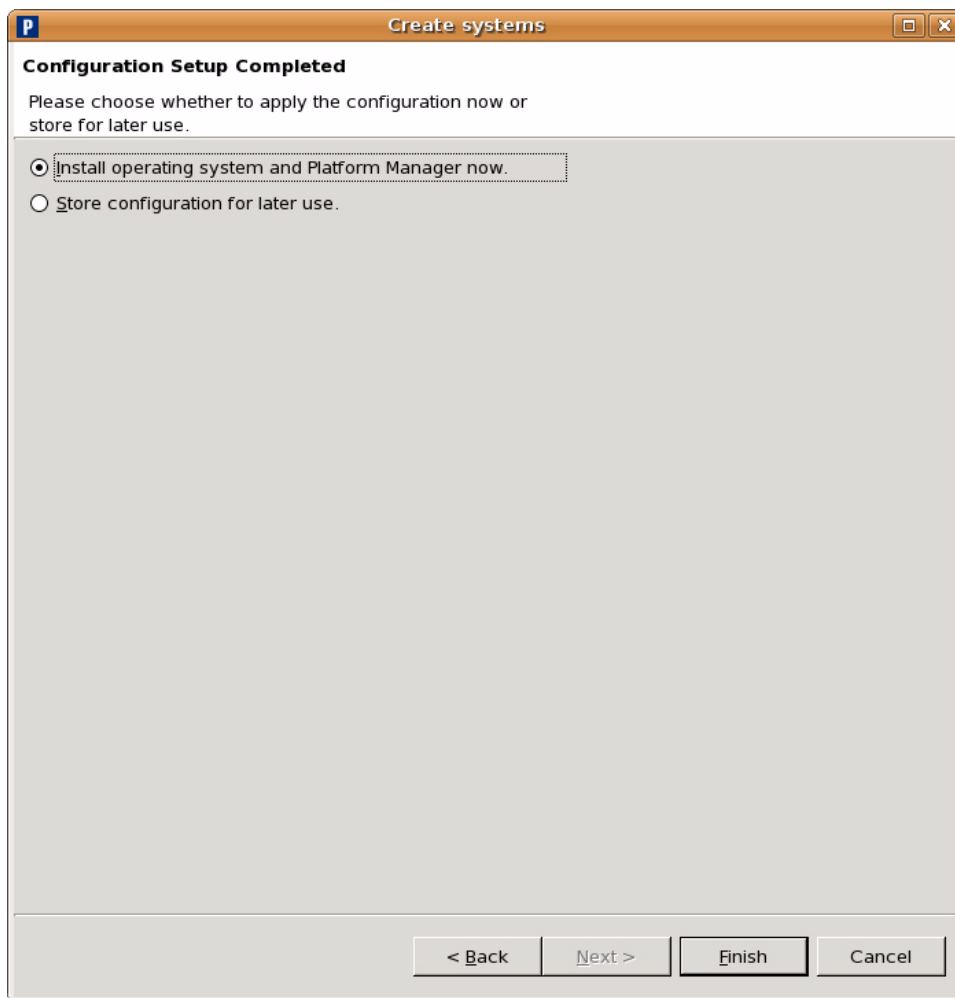


Figure 2-13 Configuration Setup Complete Screen Example

Checking the Log File Entries (Optional)

You can check the log file entries during configuration of the new node(s) to confirm that a log file has been created and to view the entries.

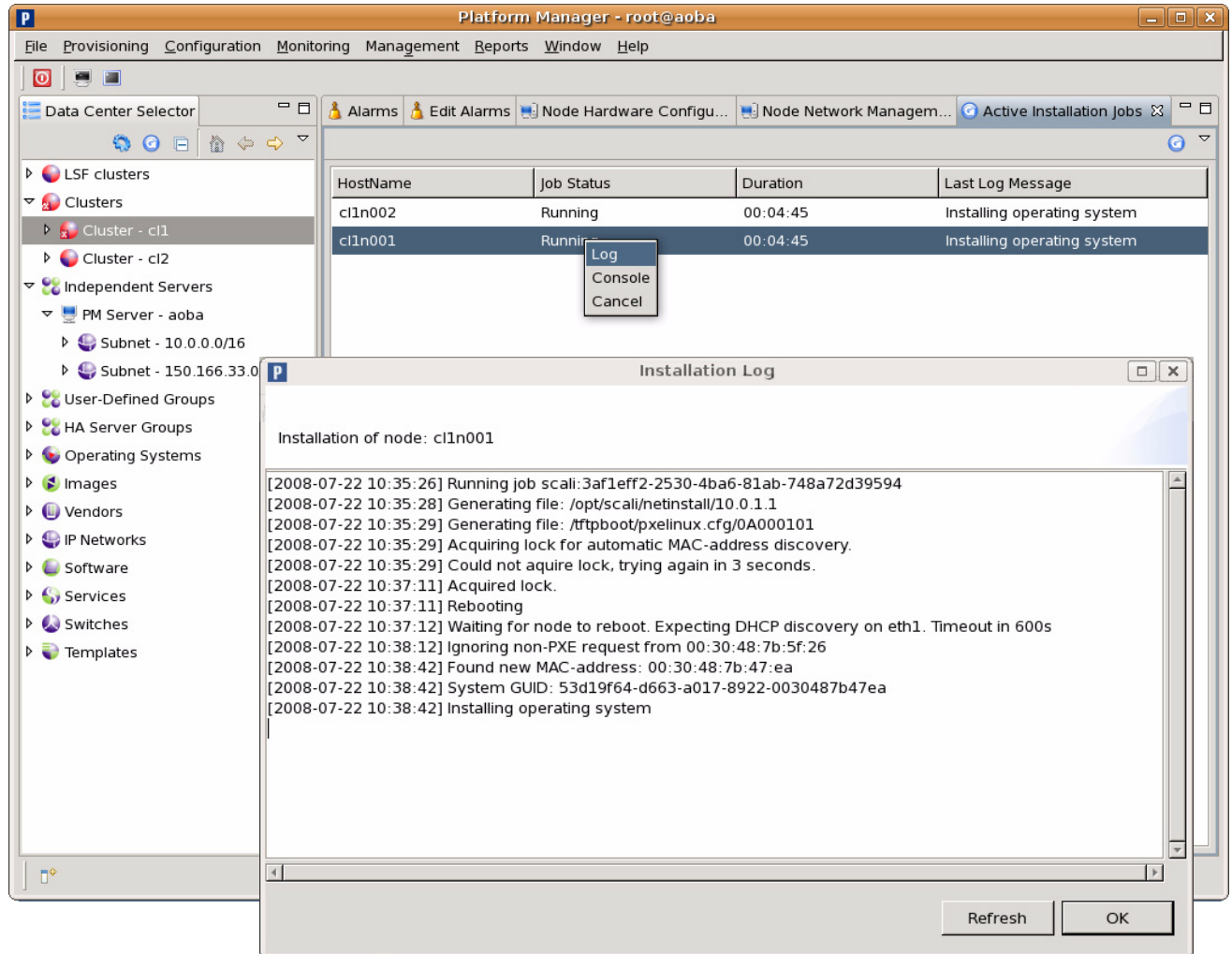


Figure 2-14 Optional Log File Screen Example

Setting a Node Failure Alarm on Platform Manage

This section shows how to create an alarm using a “Node Down” alarm as an example:

1. Start the GUI. Refer to “Start the Platform Manager GUI” on page 36 if needed.
2. Using the mouse, select the “Edit Alarms” submenu from the “Monitoring” menu item.
3. Select a node (or list of nodes) for which you want to define the alarm.
4. Then select “Add Alarm” to add the alarm and a pop-up window appears, see Figure 2-15 .
5. The popup offers input for the alarm name and an optional description for your approval.

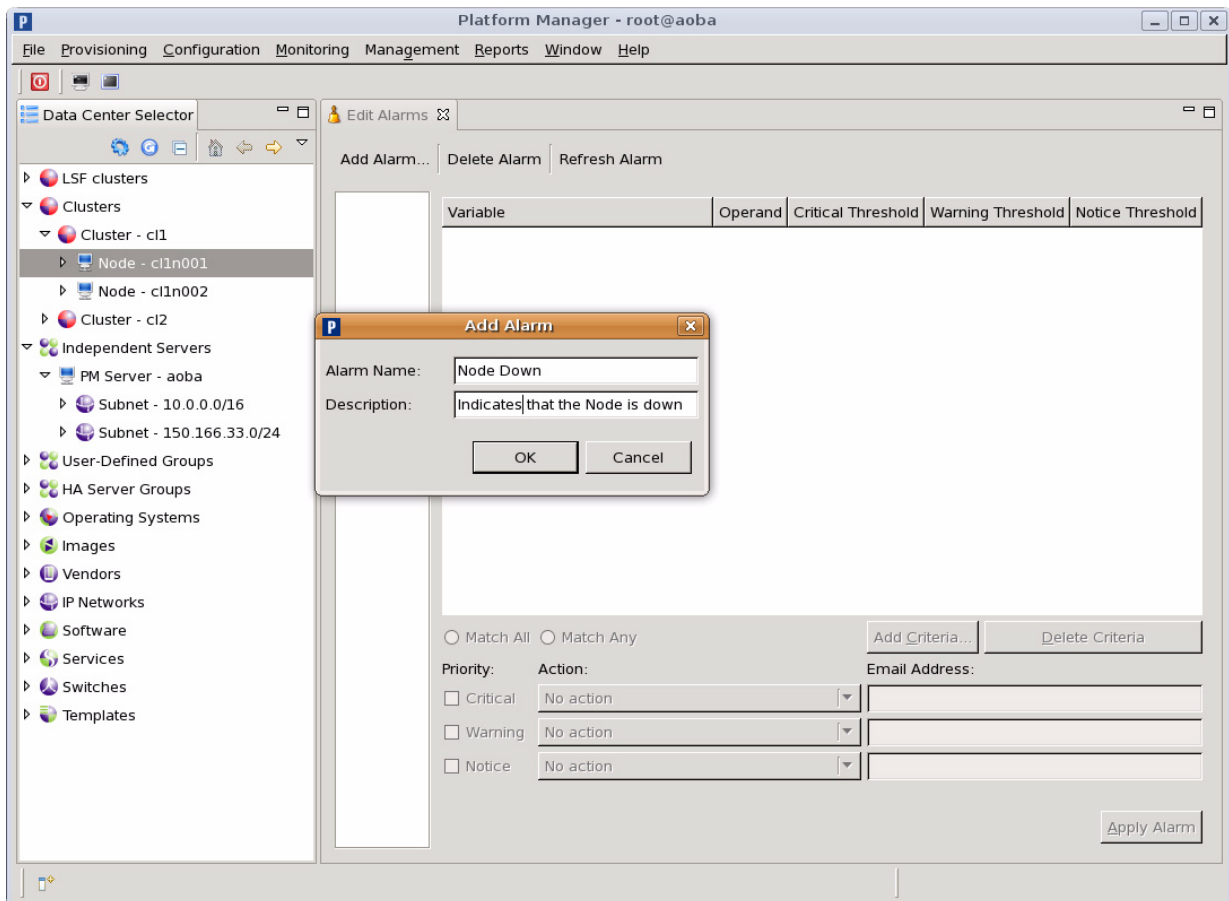


Figure 2-15 Alarm Description Popup Example

- At this time you must enter the criteria that trigger the alarm. Click on “Add Criteria” (refer to Figure 2-16.)

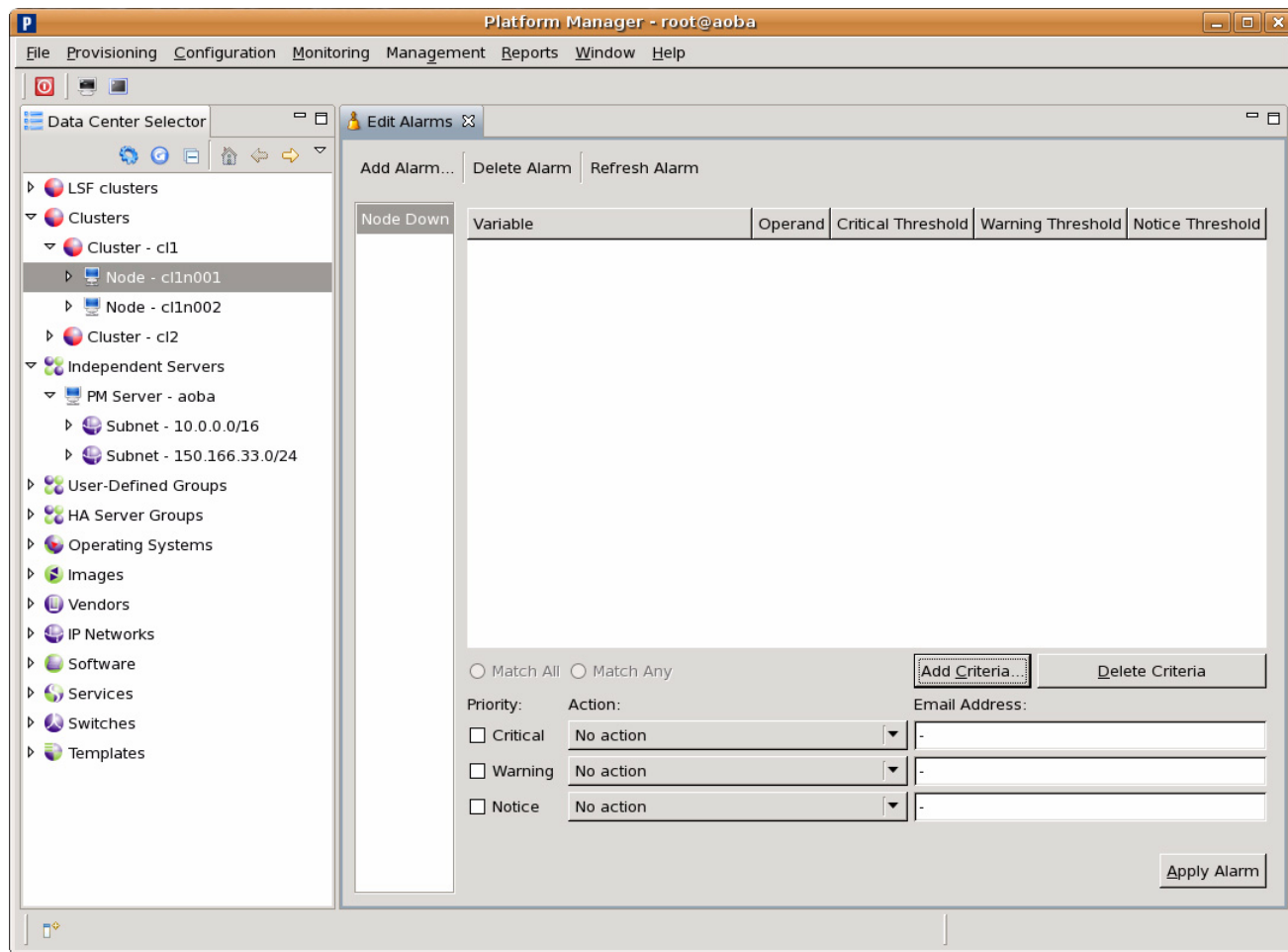


Figure 2-16 Add Criteria Screen Example

- Another popup presents itself. For this example we picked a “Filter” criteria for the node status. See the example in Figure 2-17.

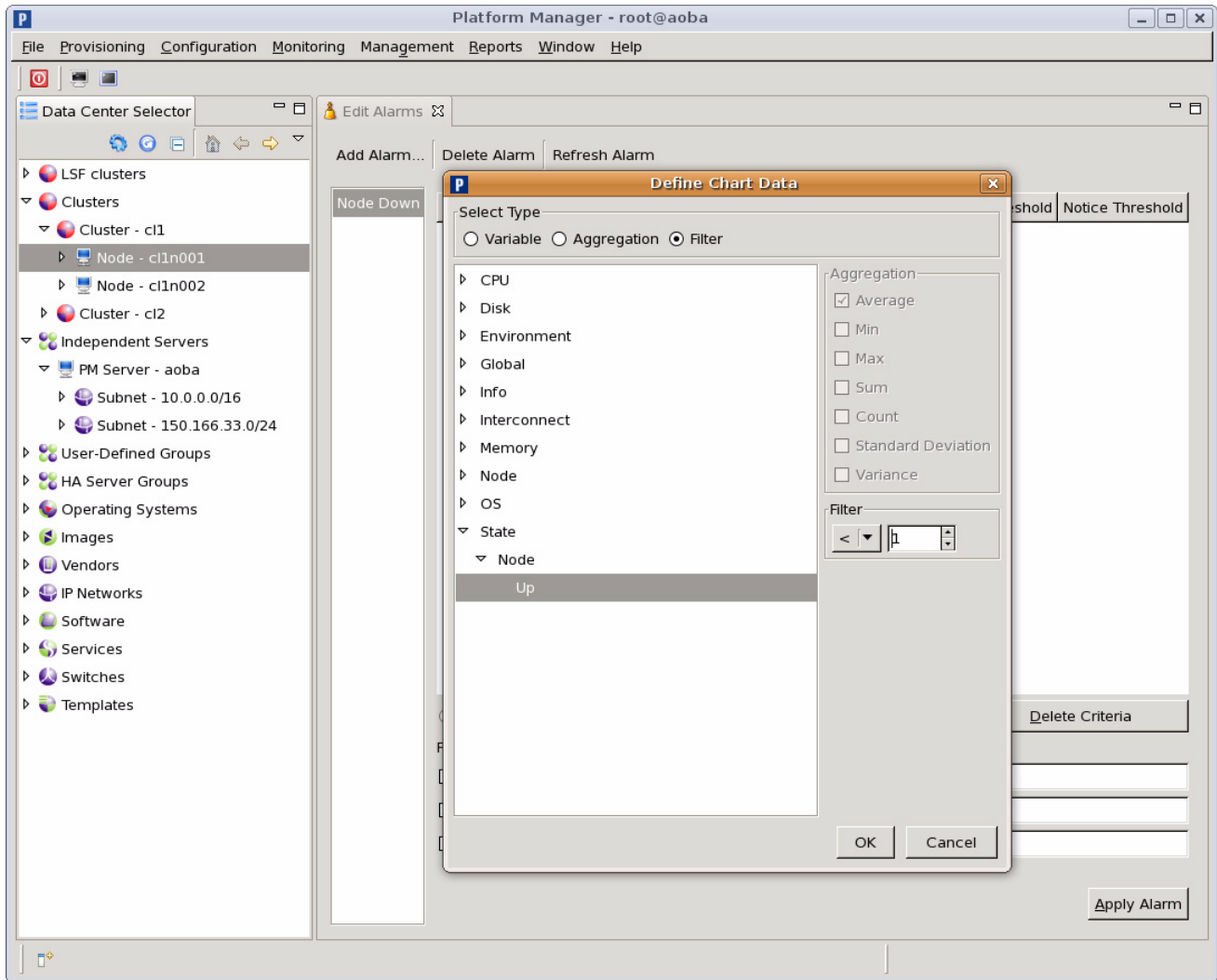


Figure 2-17 Define Chart Data Popup Example (Filter Selected)

Next we need to choose the priority for this alarm. The example assigns a critical priority for the “Node Down” alarm. We want this alarm to be triggered at most once. To enable this alarm, click on “Apply Alarm”, refer to Figure 2-18 on page 53. This alarm does not define any action to be taken when the alarm fires. This can be easily done by selecting a predefined action. As an

example, Platform Manager can send an email to a system administrator or e-mail alias. You must pick the appropriate action and supply the e-mail address or alias.

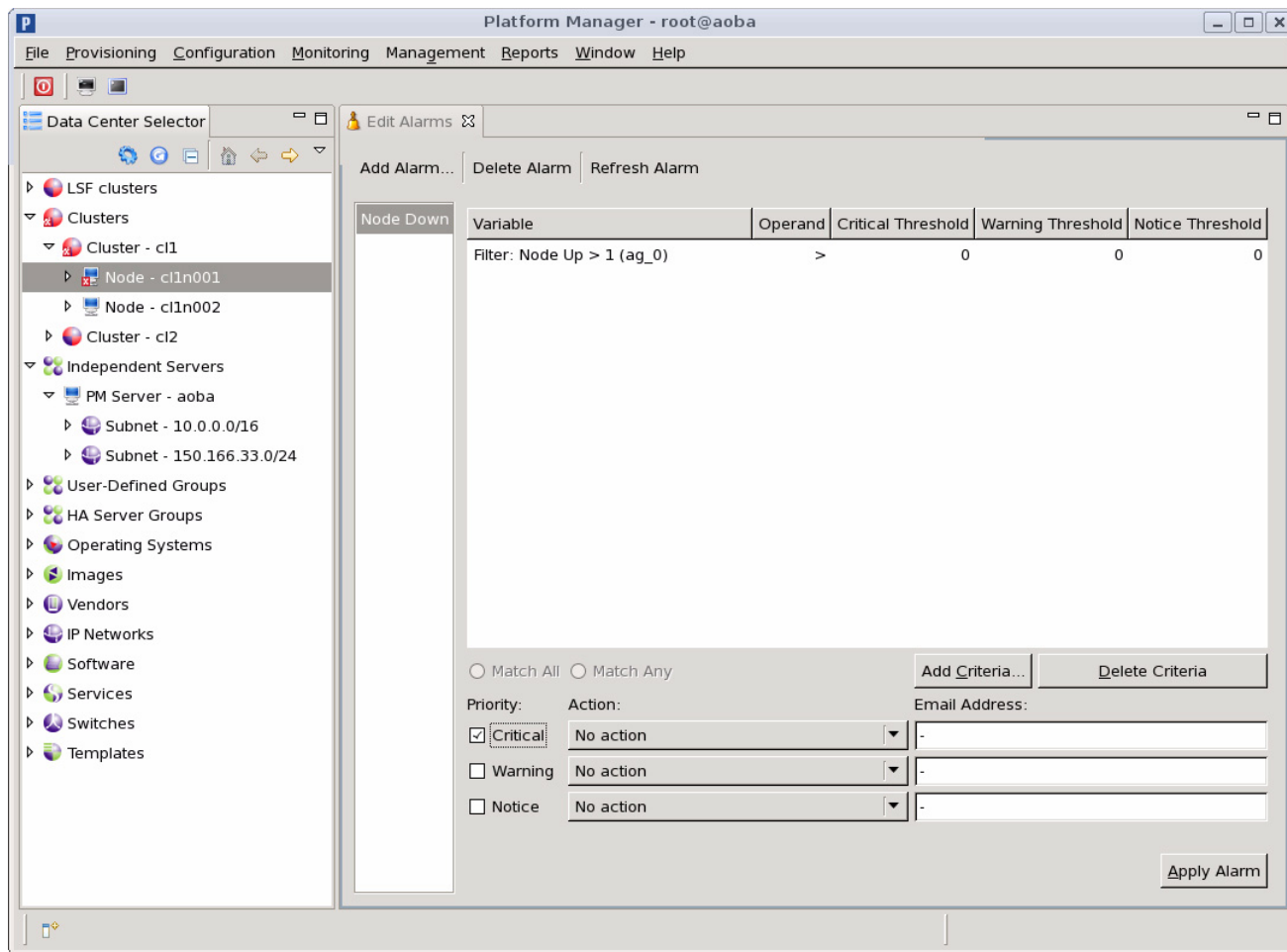


Figure 2-18 Applying the Alarm (Example Screen)

To illustrate how an alarm makes its appearance we have intentionally brought down the node. A few seconds thereafter the GUI indicates a node failure by changing the node icon in the cluster tree, refer to Figure 2-19. A few seconds later the alarm gets triggered and shows up in the alarm log, see Figure 2-20 on page 55.

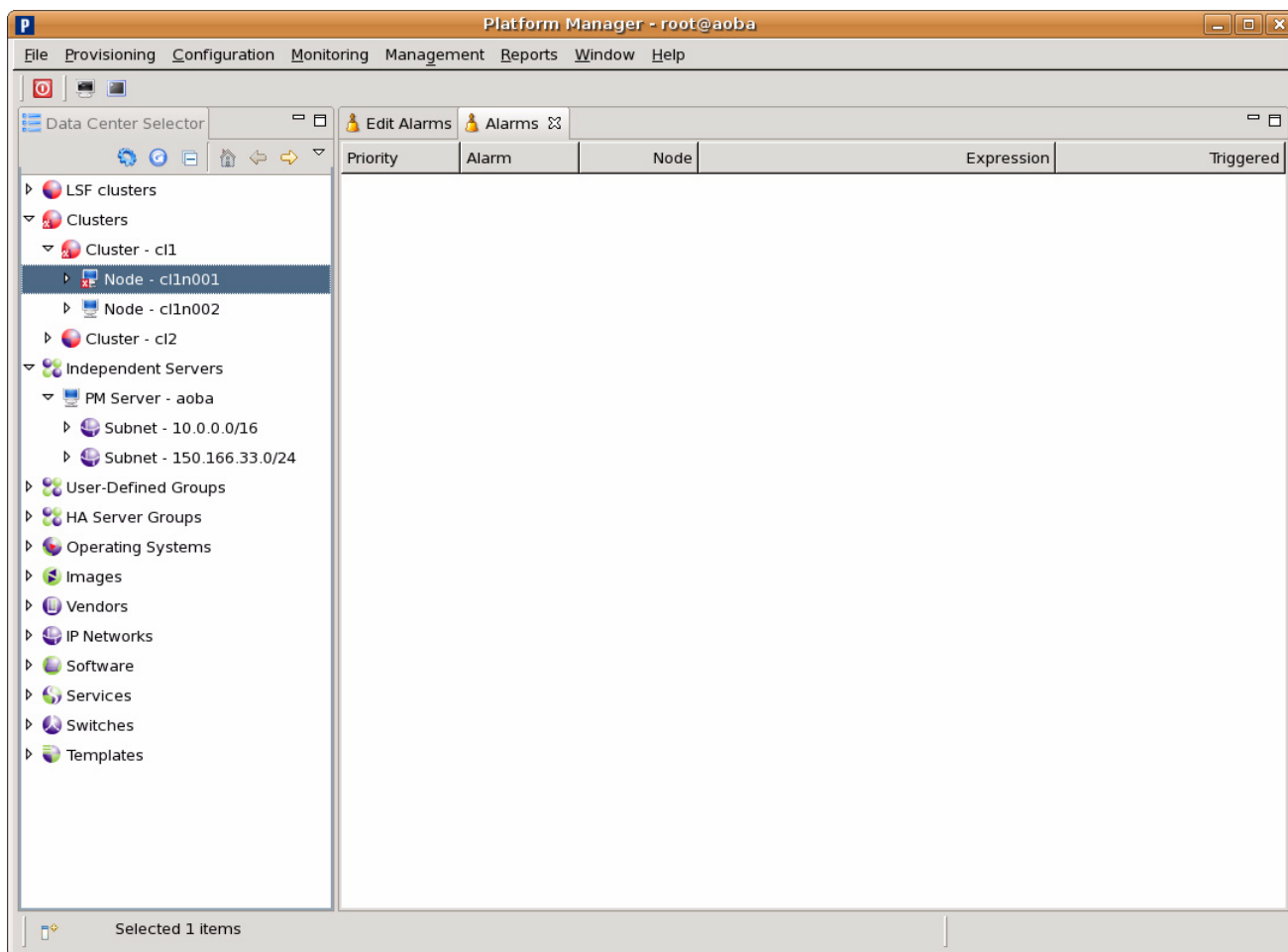


Figure 2-19 Node Failure Icon Example Screen

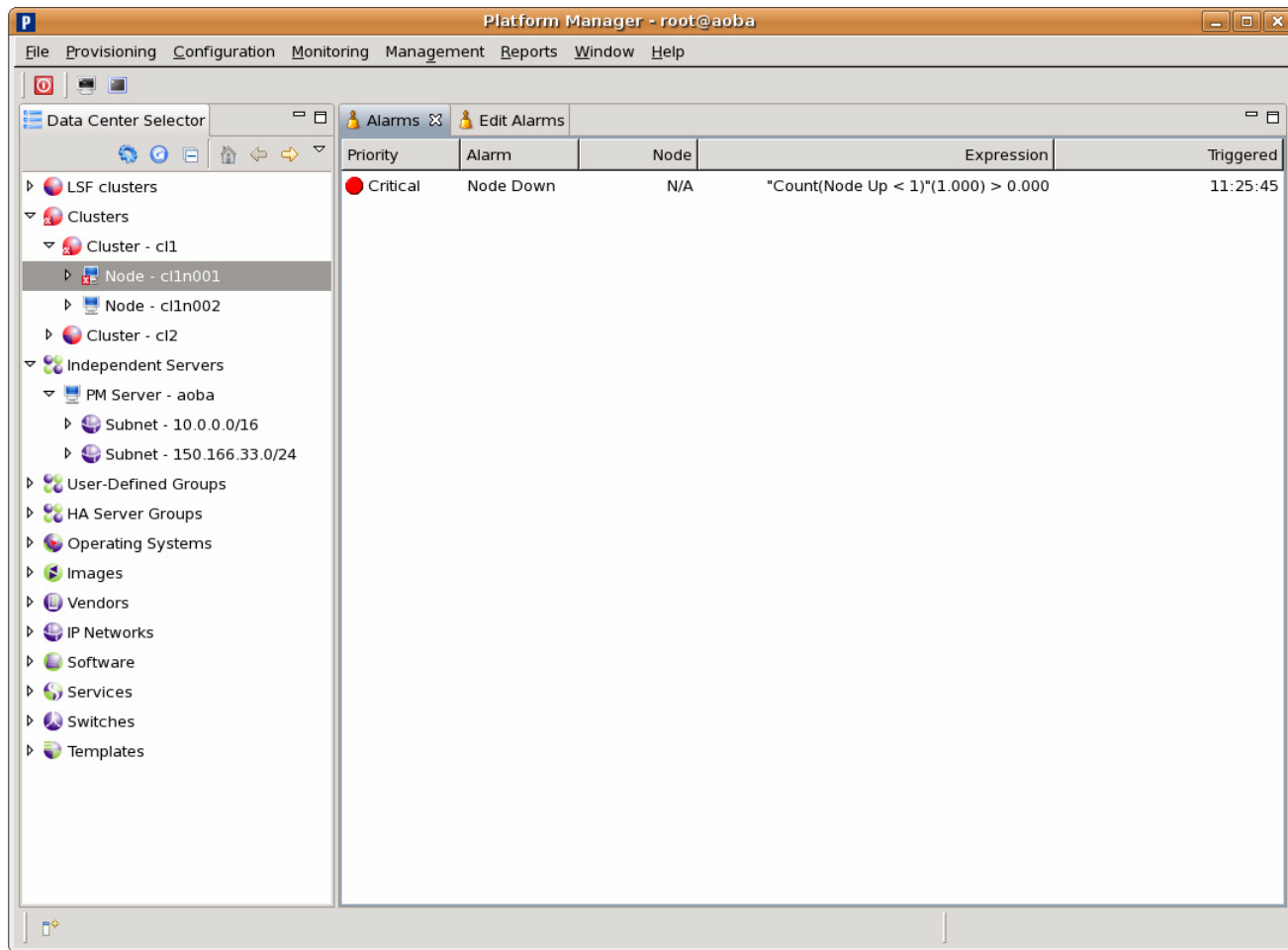


Figure 2-20 Node Down Alarm Screen Example

IPMI Commands Overview

This chapter provides a set of example IPMI commands, and is not meant to be a comprehensive guide in the use of ipmitool. Its purpose is to briefly describe some of the commonly used IPMI commands to help you get started with your cluster administration.

Command-line utility for issuing common IPMI requests allows remote operation usage:

```
ipmitool [-v] [-I interface] [-o oemtype] [-H bmc-ip-address] [-k key]
[-U user] [-P password] [-E] command...
-v : Verbosity, can be specified multiple times -vv
-I interface : IPMI interface to use
-o oemtype : Select OEM type to support
```

Note: Use `-o supermicro` for the SGI Altix XE250 head node or compute node, or the SGI Altix XE310 or XE320 compute nodes. Use `-o intelplus` for an SGI Altix XE240 head node.

Use `-o` list to see a list of current supported OEM types.

```
open - OpenIPMI driver (default)
lan - LAN connection (remote connection, requires -H/-U/-P arguments)
lanplus - LANplus connection (IPMI 2.0) Requires -H/-U/-P arguments be
supplied
```

```
-H bmc-ip-address : Hostname or IP address of remote BMC (-I lan only)
-k key : KG Key (System password) (-I lanplus only)
-U user : Username on remote system (-I lan only)
-P pass : Password for user on remote system (-I lan only)
-E : Read password from IPMI_PASSWORD environment variable
```

If `-E` and `-P` are not specified on a remote connection, the utility prompts for a password.

User Administration

BMC Supports multiple users, username/password is required for remote connections. The cluster is shipped with a factory username and password set on user id 2:

Username = **admin**

Password = **admin**

Typical ipmitool Command Line

```
ipmitool -I lanplus -o <oemtype> -H <bmc_ip_address> -U admin -P admin  
<command>
```

<opts> references in this document refer to the following command line arguments:

```
-I lanplus -o intelplus|supermicro -H <bmc-ip-address> -U admin -P admin
```

Adding a User to the BMC

```
ipmitool <opts> user set name <user ID> <username>
```

```
ipmitool <opts> user set password <user id> <password>
```

```
ipmitool <opts> user enable <user id>
```

Configuring a NIC

Display a current LAN Configuration

```
ipmitool <opts> lan print 1
```

Configure a Static IP Address

Static IP addresses are already set in the factory on LAN channel 1 of each node. Refer to Table 1-3 on page 17 and Table 1-4 on page 19 for the BMC static IP assignments.

The following commands show how to reconfigure the BMC static IP's. The "1" in the following examples indicate "channel 1" onboard nic1 controller.

```
ipmitool <opts> lan set 1 ipsrc static
```

```
ipmitool <opts> lan set 1 ipaddr x.x.x.x
```

```
impitool <opts> lan set 1 netmask x.x.x.x  
impitool <opts> lan set 1 arp respond on  
impitool <opts> lan set 1 arp generate on
```

To check your lan settings:

```
impitool <opts> lan print 1
```

Serial-over-lan Commands

Serial-Over-Lan (SOL) comes preconfigured and enabled on each node of your cluster.

Configuring SOL

SGI recommends the following parameter settings for the SGI Altix XE head node:

```
impitool <opts> sol set character-send-threshold 50 1  
impitool <opts> sol set character-accumulate-level 004 1  
impitool <opts> sol set retry-interval 20 1  
impitool <opts> sol set retry-count 6 1  
impitool <opts> sol set non-volatile-bit-rate 115.2
```

Note: Some systems were set to a 115.2 baud rate. To see your configuration, enter the following:

```
impitool <opts> sol info
```

SGI recommends the following parameter settings for the SGI Altix XE compute nodes:

```
impitool <opts> sol set character-send-threshold 80 1  
impitool <opts> sol set character-accumulate-level 012 1  
impitool <opts> sol set retry-interval 100 1  
impitool <opts> sol set retry-count 7 1  
impitool <opts> sol set non-volatile-bit-rate 115.2
```

Connecting to Node Console via SOL

```
ipmitool <opts> sol activate
```

Deactivating an SOL Connection

In certain cases using the Platform Manager GUI to access a console, you may need to deactivate the SOL connection from the command line to free up the SOL session.

```
ipmitool <opts> sol deactivate
```

Sensor commands

Sensor commands may be used to display objects, individual sensors, or all sensors in a system.

Displaying all Objects in SDR

```
ipmitool <opts> sdr list  
Ipmitool <opts> sdr dump <filename> (Dump SDR contents to a file)
```

Displaying all Sensors in the System

```
ipmitool <opts> sensor list
```

Displaying an Individual Sensor

```
ipmitool <opts> sensor get "Temp"  
Changing sensor threshold  
ipmitool <opts> sensor thresh "Temp" ucr 100
```

Thresholds are: unr, ucr, unc, lnc, lcr, lnr.

Chassis Commands

Use the following chassis commands to administer the cluster. Note that you can also use the BMC interface to perform chassis power commands on cluster nodes.

Chassis Identify

Note: The following `ipmitool chassis identify` command works only on the SGI Altix XE head node.

```
ipmitool <opts> chassis identify (defaults to 15 seconds)
ipmitool <opts> chassis identify off
```

Controlling System Power

```
ipmitool <opts> chassis power status
ipmitool <opts> chassis power off
ipmitool <opts> chassis power on
ipmitool <opts> chassis power cycle
ipmitool <opts> chassis power soft (Performs safe OS shutdown)
```

Changing System Boot Order

```
ipmitool <opts> chassis bootdev pxe
ipmitool <opts> chassis bootdev harddisk
ipmitool <opts> chassis bootdev cdrom
```

SEL Commands

The following command displays the date/time of last event, last log clear time, and number of entries in the system event log (SEL):

```
ipmitool <opts> sel info
```

To display the SEL, enter:

```
ipmitool <opts> sel list
```

To clear the SEL, enter:

```
ipmitool <opts> sel clear
```

