

**SGI® Altix® ICE 8400 Series
System Hardware User's Guide**

Document Number 007-5635-002

COPYRIGHT

© 2010 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 software described in this document is "commercial computer software" provided with restricted rights (except as to included open/free source) as specified in the FAR 52.227-19 and/or the DFAR 227.7202, or successive sections. Use beyond license provisions is a violation of worldwide intellectual property laws, treaties and conventions. This document is provided with limited rights as defined in 52.227-14.

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 SGI, 46600 Landing Parkway, Fremont, CA 94538.

TRADEMARKS AND ATTRIBUTIONS

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

Intel, Itanium and Xeon are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

AMD is a registered trademark and Opteron is a trademark of Advanced Micro Devices, Inc.

HyperTransport is a licensed trademark of the HyperTransport Technology Consortium.

UNIX is a registered trademark in the United States and other countries, licensed exclusively through X/Open Company, Ltd.

Infiniband is a trademark of the InfiniBand Trade Association.

Voltaire is a registered trademark of Voltaire Inc.

Red Hat and all Red Hat-based trademarks are trademarks or registered trademarks of Red Hat, Inc. in the United States and other countries.

Linux is a registered trademark of Linus Torvalds.

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

Record of Revision

Version	Description
001	May, 2010 First release
002	December, 2010 Information overview on AMD based blades, service nodes and new switch blade options.

Contents

List of Figures	ix
List of Tables	xi
Audience	xiii
Important Information	xiii
Chapter Descriptions	xiv
Related Publications	xv
Conventions	xvii
Product Support	xvii
Reader Comments	xviii
1. Operation Procedures	1
Precautions	1
ESD Precaution	1
Safety Precautions	2
Console Connections	3
Powering the System On and Off	4
Preparing to Power On	5
Powering On and Off	8
Console Management Power (cpower) Commands	8
Monitoring Your Server	11
2. System Management	13
Using the IU Console Option	15
Levels of System and Chassis Control	15
Chassis Controller Interaction	15
Chassis Manager Interconnects	17
Chassis Management Control (CMC) Functions	17
Chassis Management Control Front Panel Display	17

	System Power Status	18
3.	System Overview	19
	System Models	20
	Intel System and Blade Architectures	23
	IP101 Blade Architecture Overview	23
	IP103 Blade Architecture Overview	24
	IP105 Blade Architecture Overview	24
	QuickPath Interconnect Features	25
	AMD System and Blade Architectures	25
	AMD HyperTransport Communication Features	26
	IP106 Blade Architecture Overview	26
	IP110 Blade Architecture Overview	26
	Blade Memory Features	27
	Intel Blade DIMM Memory Features	27
	AMD Blade DIMM Memory Features	27
	Memory Channel Recommendation	27
	Blade DIMM Bandwidth Factors	27
	System InfiniBand Switch Blades	30
	IRU Switch Density Choices	30
	System Features and Major Components	32
	Modularity and Scalability	32
	System Administration Server	33
	Rack Leader Controller	33
	Service Nodes	34
	Login Server Function	34
	Batch Server Node	35
	I/O Gateway Node	36
	The 4U Service Node	36
	Multiple Chassis Manager Connections	38
	Reliability, Availability, and Serviceability (RAS)	40

System Components	41
IRU (Unit) Numbering	42
Rack Numbering	43
Optional System Components	43
4. Rack Information	45
Overview	45
Altix ICE 8400 Series Rack (42U)	46
ICE Rack Technical Specifications	50
5. ICE Administration/Leader Servers	51
Overview	51
Administrative/Controller Server	53
1U Rack Leader Controller Server	55
Optional AMD 2U Service Nodes	56
Optional 3U Service Nodes.	57
Optional 4U Service Nodes.	58
6. Basic Troubleshooting	61
Troubleshooting Chart	62
LED Status Indicators	63
IRU Power Supply LEDs	63
Compute/Memory Blade LEDs.	64
Chassis Management Panel “Service Required” Notices	65
A. Technical Specifications and Pinouts	67
System-level Specifications	67
Physical and Power Specifications	68
Environmental Specifications	69
I/O Port Specifications	70
Ethernet Port	71
Serial Ports	72
B. Safety Information and Regulatory Specifications	75
Safety Information	75

Regulatory Specifications 77
CMN Number 77
CE Notice and Manufacturer’s Declaration of Conformity 77
Electromagnetic Emissions 78
FCC Notice (USA Only) 78
Industry Canada Notice (Canada Only) 79
VCCI Notice (Japan Only) 79
Chinese Class A Regulatory Notice 79
Korean Class A Regulatory Notice 79
Shielded Cables 80
Electrostatic Discharge 80
Laser Compliance Statements 81
Lithium Battery Statements 82
Index 83

List of Figures

Figure 1-1	Flat Panel Rackmount Console Option	3
Figure 1-2	Administrative Controller Video Console Connection Points	4
Figure 1-3	IRU Power Supply Cable Location Example	5
Figure 1-4	Eight-Outlet Single-Phase PDU Example	6
Figure 1-5	Three-Phase PDU Example	7
Figure 1-6	Chassis Management Front Panel Example	11
Figure 1-7	IRU Chassis Management Board Location Example	12
Figure 2-1	SGI Altix ICE System Network Access Example	14
Figure 2-2	Chassis Manager Interconnection Diagram Example	16
Figure 2-3	Chassis Management Control Interface Front Panel	17
Figure 3-1	SGI Altix ICE 8400 Series System (Single Rack)	20
Figure 3-2	IRU and Rack Components Example	22
Figure 3-3	Functional Block Diagram of an ICE 8400 Individual Rack Unit (IRU)	29
Figure 3-4	InfiniBand QDR Switch Example Numbering in IRUs	31
Figure 3-5	Example Rear View of a 1U Service Node	34
Figure 3-6	2U Service Node Rear Panel	35
Figure 3-7	3U Service Node Rear Panel Example	36
Figure 3-8	4U Service Node Rear Panel Example	37
Figure 3-9	Administration and RLC Cabling to Chassis Managers Example	39
Figure 3-10	Altix ICE 8400 Series IRU System Components Example	42
Figure 4-1	Altix ICE 8400 Series Rack Example	47
Figure 4-2	Front Lock on Tall (42U) Altix Rack	48
Figure 4-3	Optional Water-Chilled Doors on Rear of Altix Rack	49
Figure 5-1	ICE System Administration Hierarchy Example Block Diagram	52
Figure 5-2	Front View of 2U Administration/Controller or Service Node	53
Figure 5-3	Rear View of 2U Administration/Controller or Service Node	53
Figure 5-4	Administrative/Controller Server Control Panel Diagram	54

Figure 5-5	1U Rack Leader Controller (RLC) Server Front and Rear Panels	55
Figure 5-6	Optional 2U AMD Service Node Example Front and Rear View	56
Figure 5-7	SGI 3U Optional Service Node Front View	57
Figure 5-8	SGI 3U Service Node Rear View	58
Figure 5-9	4U Service Node Front Controls and Interfaces	59
Figure 5-10	4U Service Node Front Panel	60
Figure 6-1	Compute Blade Status LED Locations	64
Figure 6-2	Fan Service Required Example Message on Chassis Management Panel	65
Figure A-1	Ethernet Port	71
Figure A-2	Serial Port Connector	72
Figure B-1	VCCI Notice (Japan Only)	79
Figure B-2	Chinese Class A Regulatory Notice	79
Figure B-3	Korean Class A Regulatory Notice	79

List of Tables

Table 1-1	cpower option descriptions	8
Table 1-2	cpower example command strings	10
Table 3-1	4U Service Node Rear Panel Items	37
Table 4-1	Tall Altix Rack Technical Specifications	50
Table 5-1	System administrative server control panel functions	54
Table 5-2	4U Service Node Front Control and Interface Descriptions	59
Table 5-3	4U Service Node Front Panel Item Identification	60
Table 6-1	Troubleshooting Chart	62
Table 6-2	Power Supply LED States	63
Table A-1	Altix ICE 8400 Series Configuration Ranges	67
Table A-2	Altix ICE 8400 System Physical Specifications	68
Table A-3	Environmental Specifications	69
Table A-4	Ethernet Pinouts	71
Table A-5	Serial Port Pinout.	73

About This Guide

This guide provides an overview of the architecture, general operation and descriptions of the major components that compose the SGI® Altix® integrated compute environment (ICE) 8400 series blade systems. It also provides the standard procedures for powering on and powering off the system, basic troubleshooting information, and important safety and regulatory specifications.

Audience

This guide is written for owners, system administrators, and users of SGI Altix ICE 8400 series computer systems.

It is written with the assumption that the reader has a good working knowledge of computers and computer systems.

Important Information



Warning: To avoid problems that could void your warranty, your SGI or other approved system support engineer (SSE) should perform all the setup, addition, or replacement of parts, cabling, and service of your SGI Altix ICE 8400 series system, with the exception of the following items that you can perform yourself:

- Using your system console or network access workstation to enter commands and perform system functions such as powering on and powering off, as described in this guide.
- Adding and replacing disk drives used with your system and using the ESI/ops panel (operating panel) on optional mass storage.

Chapter Descriptions

The following topics are covered in this guide:

- Chapter 1, “Operation Procedures,” provides instructions for powering on and powering off your system.
- Chapter 2, “System Management,” describes the function of the chassis management controllers (CMC) and provides overview instructions for operating the controllers.
- Chapter 3, “System Overview,” provides environmental and technical information needed to properly set up and configure the blade systems.
- Chapter 4, “Rack Information,” describes the system’s rack features.
- Chapter 5, “ICE Administration/Leader Servers” describes all the controls, connectors and LEDs located on the front of the stand-alone administrative, rack leader and other support server nodes. An outline of the server functions is also provided.
- Chapter 6, “Basic Troubleshooting,” provides recommended actions if problems occur on your system.
- Appendix A, “Technical Specifications and Pinouts,” provides physical, environmental, and power specifications for your system. Also included are the pinouts for the non-proprietary connectors.
- Appendix B, “Safety Information and Regulatory Specifications,” lists regulatory information related to use of the blade cluster system in the United States and other countries. It also provides a list of safety instructions to follow when installing, operating, or servicing the product.

Related Publications

The following documents are relevant to and can be used with the Altix ICE 8400 series of computer systems:

- *SGI Rackable C2108-TY10 System User's Guide* (P/N 007-5688-00x)

This product was previously marketed as:

- *SGI Altix XE270 User's Guide*, (P/N 007-5535-00x)

This guide covers general operation, configuration, and servicing of the 2U Rackable C2108-TY10 server node(s) used in the SGI Altix ICE 8400 system. The C2108-TY10 is used as the default administrative server on the Altix ICE 8400 system. Additionally, the C2108-TY10 can also be used as a service node for login, batch, or other service node purposes. Note that the C2108-TY10 (Altix XE270) is externally similar to the older XE250 service node used in the Altix ICE 8200 series; however, the internal mother board is significantly different and the two units are **not** interchangeable.

- *SuperServer 6016T-URF User's Manual*, (P/N 860-0498-00x)

This guide discusses the use, maintenance and operation of the 1U server primarily used as the system's rack leader controller (RLC) server node. This stand-alone 1U compute node may also be ordered configured as a login, or batch server, or other type of support server used with the Altix ICE 8400 series of computer systems.

- *SGI Rackable C3108-TY11 System User's Guide* (P/N 007-5687-00x)

This product was previously marketed as:

- *SGI Altix XE500 System User's Guide*, (P/N 007-5572-00x)

This user's guide covers general operation, configuration, and servicing of the optional 3U-high C3108-TY11 service node(s) used in the SGI Altix ICE 8400 series. The C3108-TY11 is not used as the administrative server or rack leader controller. The 3U-system may be used as a general service node for login or batch services or more specifically as a graphics interface for the larger ICE system. The server may also be used as an I/O gateway, or a mass storage resource.

- *SGI Rackable H2106-G7 System User's Guide*, (P/N 007-5683-00x).

This document covers general operation, features, options and configuration of the 2U-high H2106-G7 AMD based service node used with ICE 8400 systems that are based on AMD processor blades. The document provides basic troubleshooting and maintenance information, BIOS information, and important safety and regulatory specifications. The server is intended as a service node for login, batch, or other service node purposes.

- *SGI Altix UV 10 System User's Guide*, (P/N 007-5645-00x)

This user's guide covers general operation, configuration, troubleshooting and a description of major components of the optional 4U-high Altix UV 10 multi-node service unit used in SGI Altix ICE 8400 systems. The Altix UV 10 cannot be used as an administrative server or rack leader controller. Uses for the system include configuration as an I/O gateway, a mass storage resource, a general service node for login or batch services or some combination of the previous functions.

- *SGI Tempo System Administrator's Guide*, (P/N 007-4993-00x)

This guide discusses system configuration and software administration operations used with the SGI Altix ICE 8400 series. At time of publication, this document is intended for people who manage the operation of ICE systems with SUSE Linux Enterprise Server 11 (SLES 11) or later.

- Man pages (online)

Man pages locate and print the titled entries from the online reference manuals.

You can obtain SGI documentation, release notes, or man pages in the following ways:

- 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 information.
- The release notes, which contain the latest information about software and documentation in this release, are in a file named README.SGI in the root directory of the SGI ProPack for Linux Documentation CD.
- You can also view man pages by typing `man <title>` on a command line.

SGI systems include a set of Linux man pages, formatted in the standard UNIX “man page” style. Important system configuration files and commands are documented on man pages. These are found online on the internal system disk (or DVD) and are displayed using the `man` command. For example, to display a man page, type the request on a command line:

```
man commandx
```

References in the documentation to these pages include the name of the command and the section number in which the command is found. For additional information about displaying man pages using the `man` command, see `man (1)`. In addition, the `apropos` command locates man pages based on keywords. For example, to display a list of man pages that describe disks, type the following on a command line:

```
apropos disk
```


Conventions

The following conventions are used throughout this document:

Convention	Meaning
Command	This fixed-space font denotes literal items such as commands, files, routines, path names, signals, messages, and programming language structures.
<i>variable</i>	The italic typeface denotes variable entries and words or concepts being defined. Italic typeface is also used for book titles.
user input	This bold fixed-space font denotes literal items that the user enters in interactive sessions. Output is shown in nonbold, fixed-space font.
[]	Brackets enclose optional portions of a command or directive line.
...	Ellipses indicate that a preceding element can be repeated.
man page(x)	Man page section identifiers appear in parentheses after man page names.
GUI element	This font denotes the names of graphical user interface (GUI) elements such as windows, screens, dialog boxes, menus, toolbars, icons, buttons, boxes, fields, and lists.

Product Support

SGI provides a comprehensive product support and maintenance program for its products, as follows:

- If you are in North America, contact the Technical Assistance Center at +1 800 800 4SGI or contact your authorized service provider.
- If you are outside North America, contact the SGI subsidiary or authorized distributor in your country. International customers can visit <http://www.sgi.com/support/>. Click on the “Support Centers” link under the “Online Support” heading for information on how to contact your nearest SGI customer support center.

Reader Comments

If you have comments about the technical accuracy, content, or organization of this document, contact SGI. Be sure to include the title and document number of the manual with your comments. (Online, the document number is located in the front matter of the manual. In printed manuals, the document number is located at the bottom of each page.)

You can contact SGI in any of the following ways:

- Send e-mail to the following address: techpubs@sgi.com
- Contact your customer service representative and ask that an incident be filed in the SGI incident tracking system.
- Send mail to the following address:

Technical Publications
SGI
46600 Landing Parkway
Fremont, California 94538

SGI values your comments and will respond to them promptly.

Operation Procedures

This chapter explains how to operate your new system in the following sections:

- “Precautions” on page 1
- “Console Connections” on page 3
- “Powering the System On and Off” on page 4
- “Monitoring Your Server” on page 11

Precautions

Before operating your system, familiarize yourself with the safety information in the following sections:

- “ESD Precaution” on page 1
- “Safety Precautions” on page 2

ESD Precaution

Caution: Observe all electro-static discharge (ESD) precautions. Failure to do so can result in damage to the equipment.

Wear an approved ESD wrist strap when you handle any ESD-sensitive device to eliminate possible damage to equipment. Connect the wrist strap cord directly to earth ground.

Safety Precautions



Warning: Before operating or servicing any part of this product, read the “Safety Information” on page 75.



Danger: Keep fingers and conductive tools away from high-voltage areas. Failure to follow these precautions will result in serious injury or death. The high-voltage areas of the system are indicated with high-voltage warning labels.



Caution: Power off the system only after the system software has been shut down in an orderly manner. If you power off the system before you halt the operating system, data may be corrupted.



Warning: If a lithium battery is installed in your system as a soldered part, only qualified SGI service personnel should replace this lithium battery. For a battery of another type, replace it only with the same type or an equivalent type recommended by the battery manufacturer, or an explosion could occur. Discard used batteries according to the manufacturer’s instructions.

Console Connections

The flat panel console option (see Figure 1-1) has the following listed features:

1. **Slide Release** - Move this tab sideways to slide the console out. It locks the drawer closed when the console is not in use and prevents it from accidentally sliding open.
2. **Handle** - Used to push and pull the module in and out of the rack.
3. **LCD Display Controls** - The LCD controls include On/Off buttons and buttons to control the position and picture settings of the LCD display.
4. **Power LED** - Illuminates blue when the unit is receiving power.

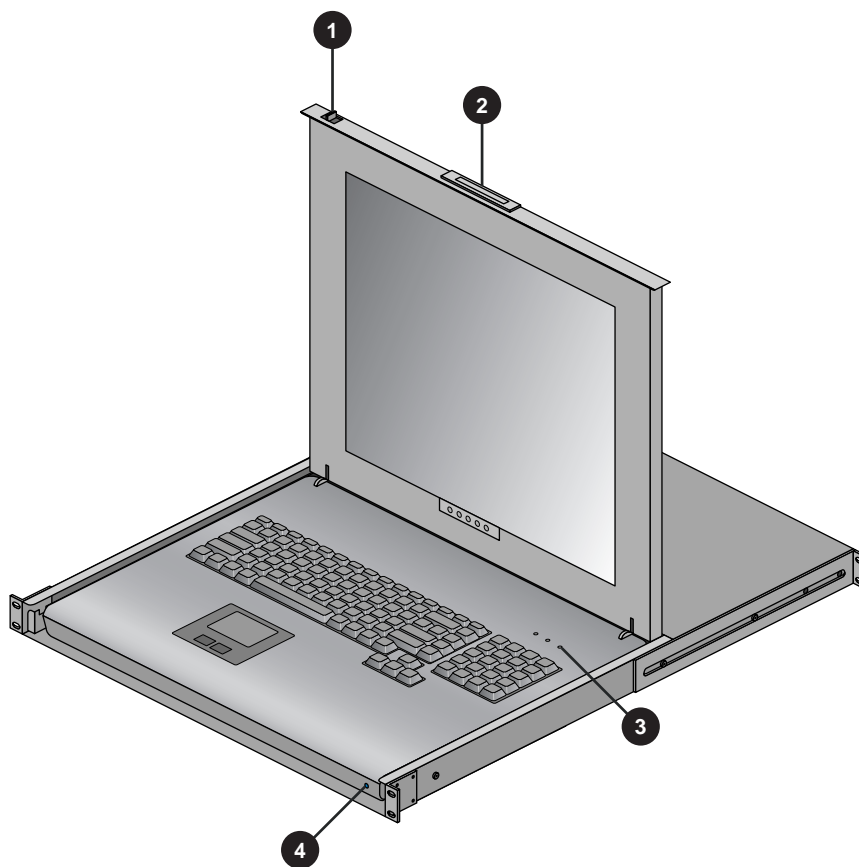


Figure 1-1 Flat Panel Rackmount Console Option

A console is defined as a connection to the system (to the administrative server) that provides administrative access to the cluster. SGI offers a rackmounted flat panel console option that attaches to the administrative node's video, keyboard and mouse connectors.

A console can also be a LAN-attached personal computer, laptop or workstation (RJ45 Ethernet connection). Serial-over-LAN is enabled by default on the administrative controller server and normal output through the RS-232 port is disabled. In certain limited cases, a dumb (RS-232) terminal could be used to communicate directly with the administrative server. This connection is typically used for service purposes or for system console access in smaller systems, or where an external ethernet connection is not used or available. Check with your service representative if use of an RS-232 terminal is required for your system.

The flat panel rackmount or other optional VGA console connects to the administration controller's video and keyboard/mouse connectors as shown in Figure 1-2.

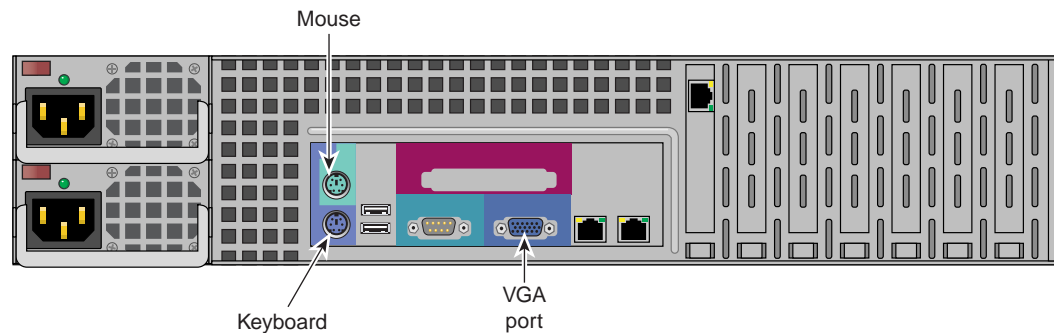


Figure 1-2 Administrative Controller Video Console Connection Points

Powering the System On and Off

This section explains how to power on and power off individual rack units, or your entire Altix ICE system, as follows:

- “Preparing to Power On” on page 5
- “Powering On and Off” on page 8

Entering commands from a system console, you can power on and power off individual IRUs, blade-based nodes, and stand-alone servers, or the entire system.

When using the SGI cluster manager software, you can monitor and manage your server from a remote location (see the *SGI Tempo System Administrator's Guide*).

You may also monitor and manage your server with tools such as the Voltaire or Intel message passing interface (MPI). For details, see the documentation for the particular tool.

Preparing to Power On

To prepare to power on your system, follow these steps:

1. Check to ensure that the cabling between the rack's power distribution units (PDUs) and the wall power-plug receptacle is secure.

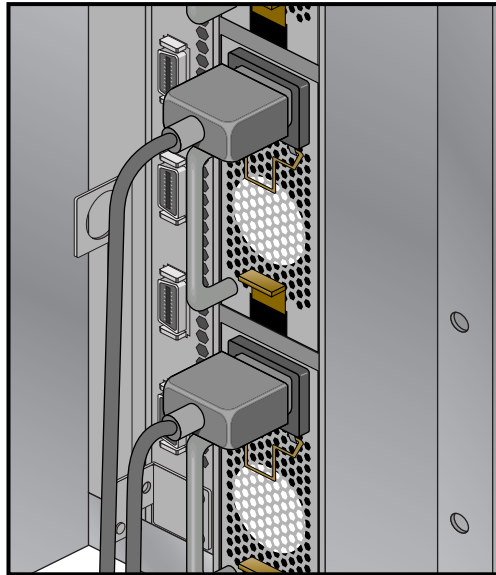


Figure 1-3 IRU Power Supply Cable Location Example

2. For each individual IRU that you want to power on, make sure that the power cables are plugged into all the IRU power supplies correctly, as shown in Figure 1-3. Setting the circuit breakers on the PDUs to the “On” position will apply power to the IRU and will start the chassis manager in each IRU. Note that the chassis manager in each IRU stays powered on as long as there is power coming into the unit. Turn off the PDU breaker switch that supplies voltage to the IRU if you want to remove all power from the unit.

3. If you plan to power on a server that includes optional mass storage enclosures, make sure that the power switch on the rear of each PSU/cooling module (one or two per enclosure) is in the **1** (on) position.
4. Make sure that all PDU circuit breaker switches (see the examples in Figure 1-4, and Figure 1-5 on page 7) are turned on to provide power when the system is booted up.

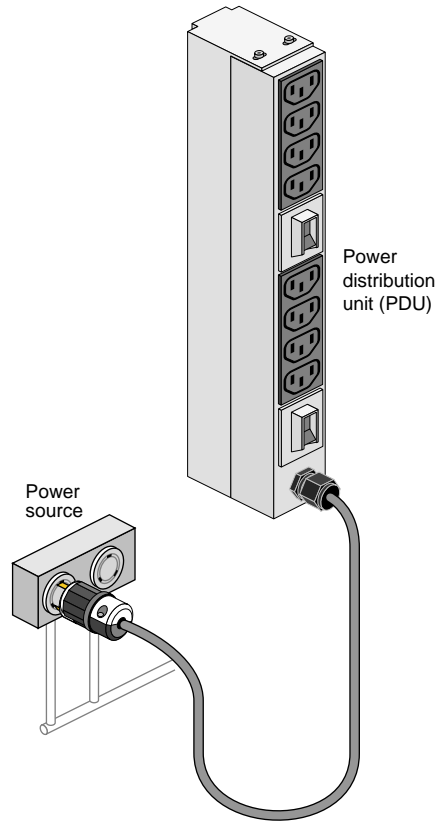


Figure 1-4 Eight-Outlet Single-Phase PDU Example

Figure 1-5 on page 7 shows an example of the three-phase PDU.

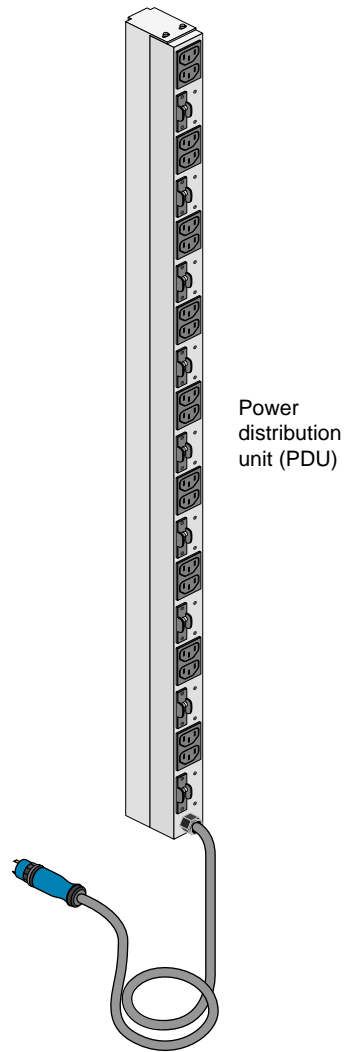


Figure 1-5 Three-Phase PDU Example

Powering On and Off

The power-on and off procedure varies with your system setup. See the *SGI Tempo System Administrator's Guide*, (P/N 007-4993-00x) for a more complete description of system commands.

Note: The `cpower` commands are normally run through the administration node. If you have a terminal connected to an administrative server with a serial interface, you should be able execute these commands.

Console Management Power (`cpower`) Commands

This section provides an overview of the console management power (`cpower`) commands for the SGI Altix ICE 8400 system.

The `cpower` commands allow you to power up, power down, reset, and show the power status of multiple or single system components or individual racks.

The `cpower` command is, as follows:

```
cpower <option...> <target_type> <action> <target>
```

The `cpower` command accepts the following arguments as described in Table 1-1.

See Table 1-2 on page 10 for examples of the `cpower` command strings.

Table 1-1 `cpower` option descriptions

Argument	Description
Option	
<code>--noleader</code>	Do not include rack leader nodes. Valid with rack and system domains only.
<code>--noservice</code>	Do not include service nodes.
<code>--ipmi</code>	Uses ipmitool to communicate.
<code>--ssh</code>	Uses ssh to communicate.
<code>--intelplus</code>	Use the “-o intelplus option” for ipmitool [default].
<code>--verbose</code>	Print additional information on command progress.

Table 1-1 (continued) cpower option descriptions

Argument	Description
--noexec	Display but do not execute commands that affect power.
Target_type	
--node	Apply the action to a node or nodes. Nodes can be blade compute nodes (inside an IRU), administration server nodes, rack leader controller nodes or service nodes.
--iru	Apply the action at the IRU level
--rack	Apply the action to all components in a rack
--system	Apply the action to the entire system. You must not specify a target with this type.
--all	Allows the use of wildcards in the target name
Action	
--status	Shows the power status of the target [default]
--up --on	Powers up the target
--down --off	Powers down the target
--cycle	Power cycles the target
--reboot	Reboot the target, even if it is already booted. Wait for all targets to boot.
--halt	Shutdown the target, but do not power it off. Wait for targets to shut down.
--help	Usage and help text

Note: If you include a rack leader controller in your wildcard specification, and a command that may take it offline, you will see a warning intended to prevent accidental resets of the RLC, as that could make the rack unreachable.

Table 1-2 cpower example command strings

Command	Status/result
# cpower --system --up	Powers up all nodes in the system (--up is the same as --on).
# cpower --rack r1	Determines the power status of all nodes in rack 1 (including the RLC), except CMCs.
# cpower --system	Provides status of every compute node in the system.
# cpower --boot --rack r1	Boots any nodes in rack 1 not already online.
# cpower --system --down	Completely powers down every node in the system. Use only if you want to shut down all nodes (see the next example).
# cpower --halt --system --noleader --noservice	Shuts down (halts) all the IRU compute nodes in the system, but not the administrative controller server, rack leader controller or other service nodes.
# cpower --boot r1i0n8	Command tries to specifically boot rack 1, IRU0, node 8.
# cpower --halt --rack r1	Will halt and then power off all of the computer nodes in parallel located in rack 1, then halts the rack leader controller. Use --noleader if you want the RLC to remain on.

See the *SGI Tempo System Administrator's Guide*, (P/N 007-4993-00x) for more information on `cpower` commands.

See the section “System Power Status” on page 18 in this manual for additional related console information.

Monitoring Your Server

You can monitor your Altix ICE 8400 server from the following sources:

- An optional flat panel rackmounted monitor with PS2 keyboard/mouse can be connected to the administration server node for basic monitoring and administration of the Altix system. See the section “Console Connections” on page 3 for more information. SLES 11 or higher is required for this option.
- You can attach an optional LAN-connected console via secure shell (ssh) to an Ethernet port adapter on the administration controller server. You will need to connect either a local or remote workstation/PC to the IP address of the administration controller server to access and monitor the system via IPMI.

See the Console Management section in the *SGI Tempo System Administrator's Guide*, (P/N 007-4993-00x) for more information on the open source console management package.

These console connections enable you to view the status and error messages generated by your Altix ICE 8400 system. You can also use these consoles to input commands to manage and monitor your system. See the section “System Power Status” on page 18, for additional information.

The chassis management front panel is an additional source of IRU status information.

Figure 1-6 shows an example of the panel and the functions it reports. Figure 1-7 on page 12 shows a front view of a system IRU and the location of the chassis management panel.

The CMC panel offers basic information on the rack (its assigned number) and various types of status information regarding the IRU chassis components. For additional information, see “Chassis Management Panel “Service Required” Notices” in Chapter 6.

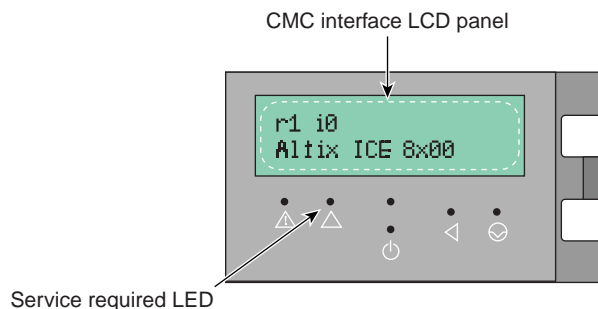


Figure 1-6 Chassis Management Front Panel Example

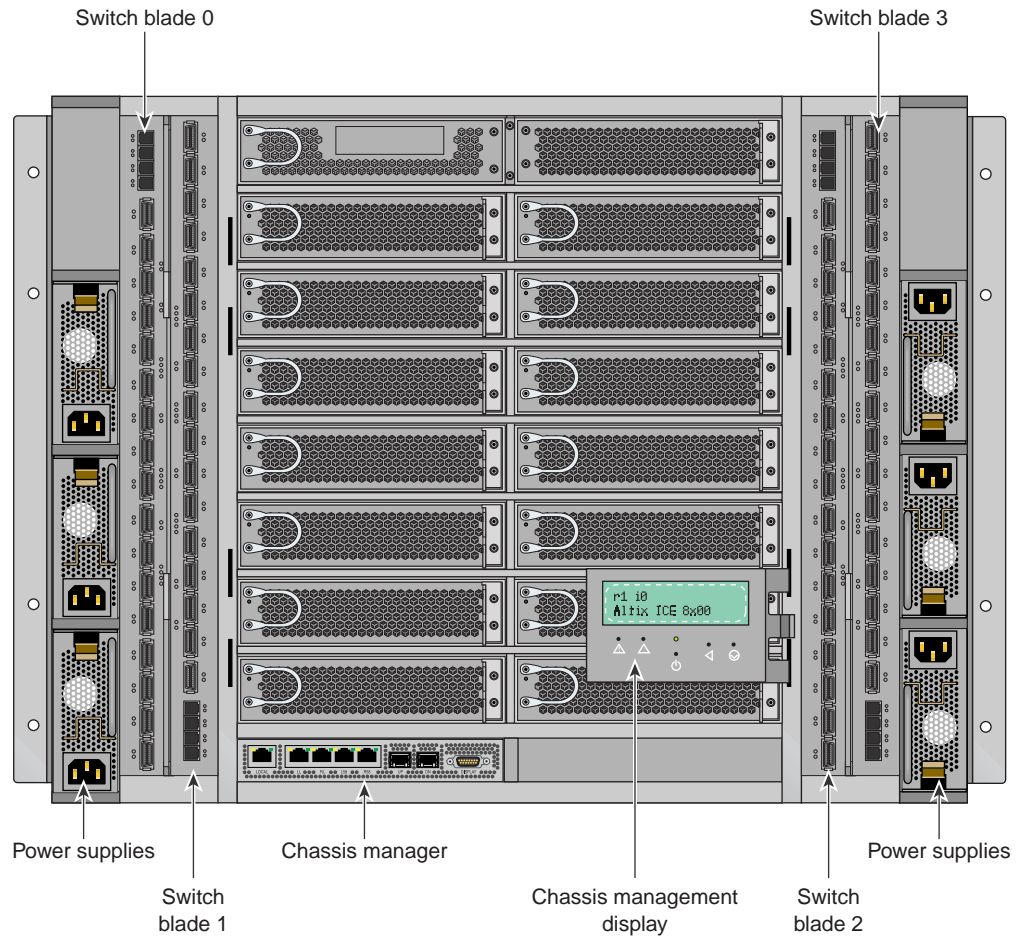


Figure 1-7 IRU Chassis Management Board Location Example

PCIe based I/O sub-systems are sited in the administrative controller server, rack leader controller and service node systems used with the IRUs. These are the primary configurable I/O system interfaces for the Altix ICE 8400 systems. See the particular server's user guide for detailed information on installing optional I/O cards or other components.

Note that each IRU is configured with two or four InfiniBand switch blades for dual-plane topologies and one or two InfiniBand switch blades for single-plane topologies.

System Management

This chapter describes the interaction and functions of system controllers in the following sections:

- “Levels of System and Chassis Control” on page 15
- “Chassis Manager Interconnects” on page 17
- “System Power Status” on page 18

Each IRU has one chassis manager, which is located directly below compute blade slot 0. The chassis manager supports power-up and power-down of the IRU’s compute/memory node blades and environmental monitoring of all units within the IRU.

Note that the stand-alone service nodes use IPMI to monitor system “health”.

Mass storage enclosures do not share a direct interconnect with the ICE 8400 chassis manager (CMC).

Figure 2-1 shows an example remote LAN-connected console used to monitor a single-rack Altix ICE 8400 series system.

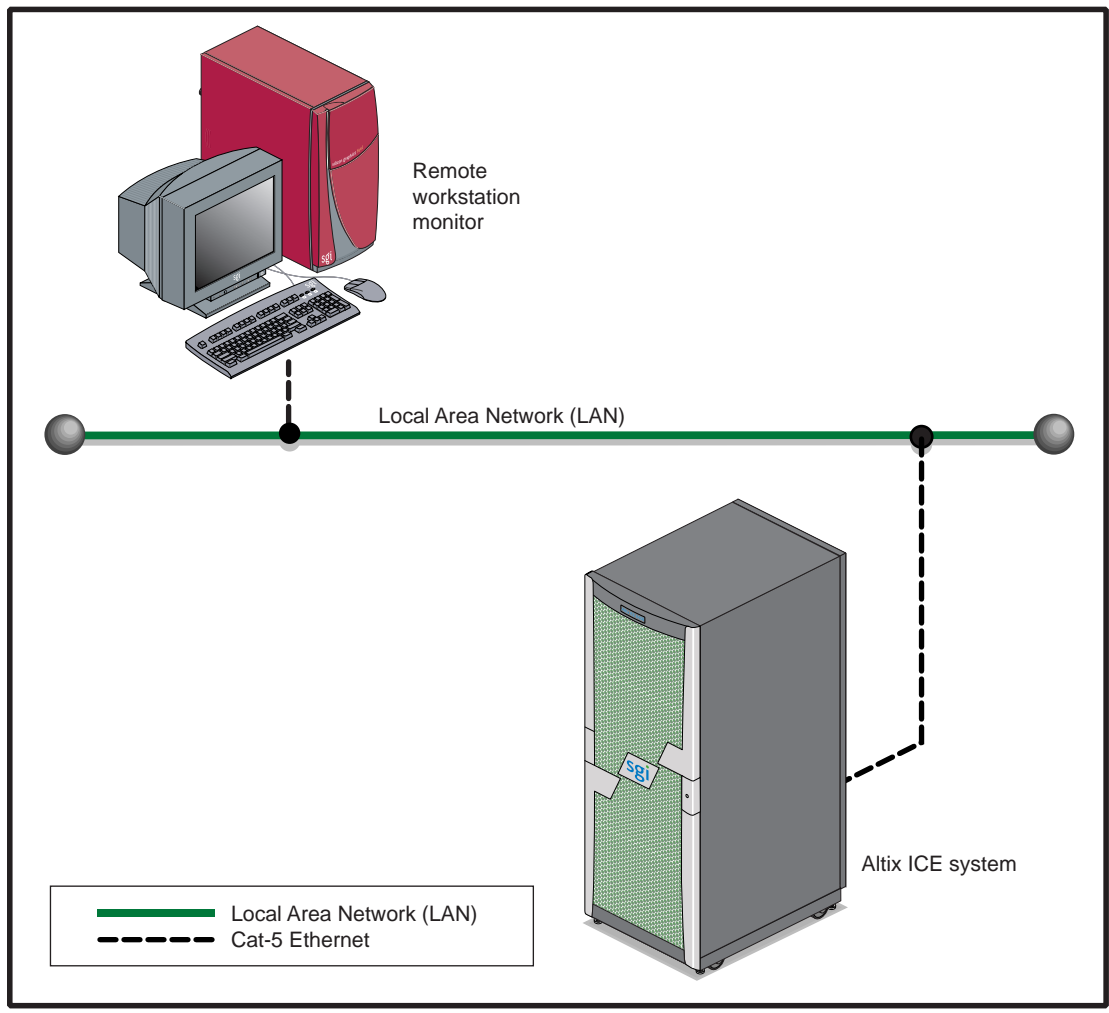


Figure 2-1 SGI Altix ICE System Network Access Example

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 administrative controller server using PS/2 and HD15M connectors or to an optional 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.

Note: While the 1U console is normally plugged into the administrative controller server in the ICE system, it can also be connected to a rack leader controller server in the system for terminal access purposes.

The 27-pound (12.27-kg) console automatically goes into sleep mode when the cover is closed.

Levels of System and Chassis Control

The chassis management control network configuration of your ICE 8400 series machine will depend on the size of the system and the control options selected. Typically, any system with multiple IRUs will be interconnected by the chassis managers in each IRU.

Note: Mass storage option enclosures are not monitored by the IRU’s chassis manager. Most optional mass storage enclosures have their own internal microcontrollers for monitoring and controlling all elements of the disk array. See the owner’s guide for your mass storage option for more information on this topic.

Chassis Controller Interaction

In all Altix ICE 8400 series systems all the system chassis management controllers communicate with each other in the following ways:

- All enclosures within a system communicate with each other through their chassis manager connections (CMC) (note the chassis managers are enlarged for clarity in Figure 2-2).
- The CMC does the environmental management for an IRU as well as power control, and provides an ethernet network infrastructure for the management of the system.

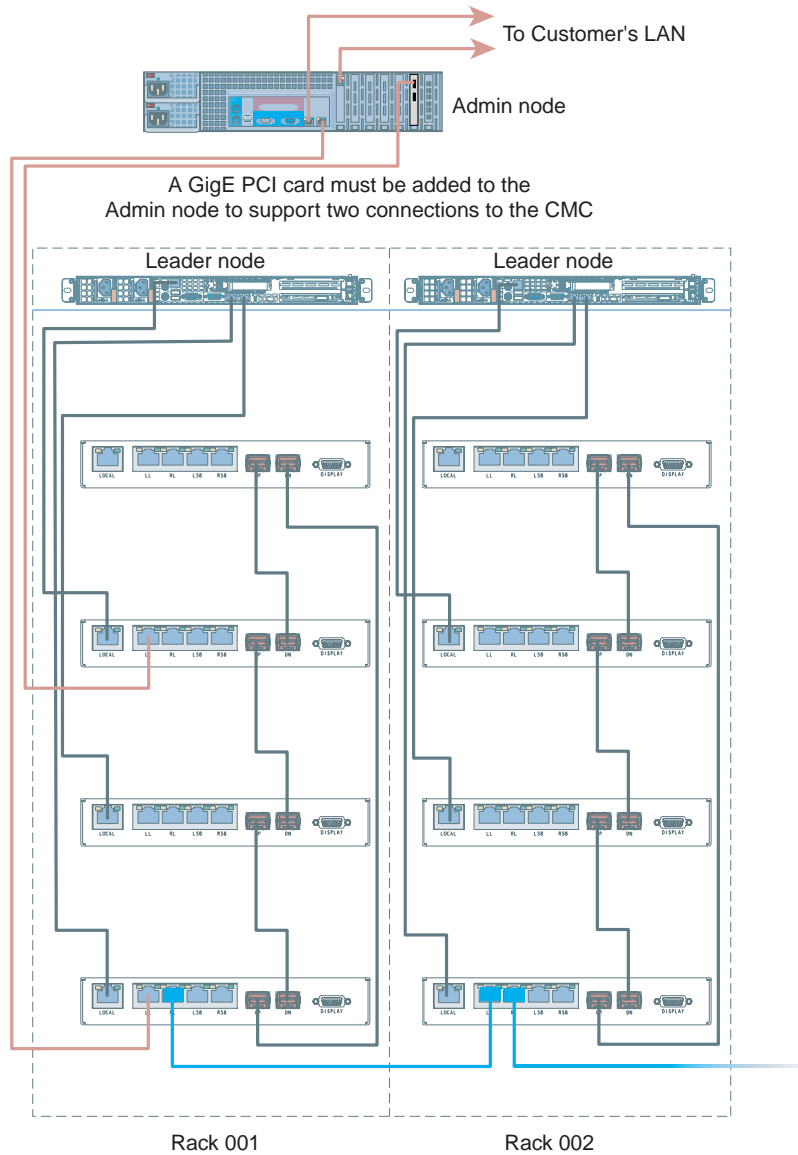


Figure 2-2 Chassis Manager Interconnection Diagram Example

Chassis Manager Interconnects

The chassis manager in the lower IRU connects to the administration and rack leader servers, see the example in Figure 2-2. See also “Multiple Chassis Manager Connections” in Chapter 3.

Chassis Management Control (CMC) Functions

The following list summarizes the control and monitoring functions that the CMC performs. Most functions are common across multiple IRUs.

- Controls and monitors IRU fan speeds
- Reads system identification (ID) PROMs
- Monitors voltage levels and reports failures
- Monitors the On/Off power sequence
- Monitors system resets
- Applies a preset voltage to switch blades and fan control boards

Chassis Management Control Front Panel Display

Figure 2-3 shows the chassis management controller display panel on the IRU.

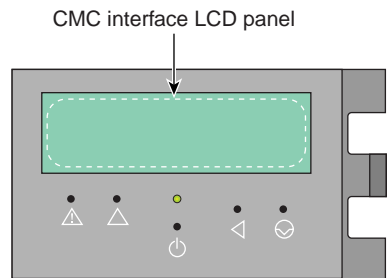


Figure 2-3 Chassis Management Control Interface Front Panel

Note: The CMC front panel display provides rack and IRU identification information and IRU hardware chassis malfunction alerts. See “Chassis Management Panel “Service Required” Notices” in Chapter 6 for additional information.

System Power Status

The `cpower` command is the main interface for all power management commands. You can request power status, power-on and power-off the system with commands entered via the administrative controller server or rack leader controller in the system rack. The `cpower` commands are communicating with BMCs using the IPMI protocol.

The `cpower` commands may require several seconds to several minutes to complete, depending on how many IRUs are being queried for status, powered-up, or shut down.

```
# cpower --system
```

This command gives the status of all compute nodes in the system.

To power on or power off a specific IRU, enter the following commands:

```
# cpower --iru --up r1i0
```

The system should respond by powering up the IRU 0 nodes in rack 1. Note that **--on** is the same as **--up**. This command does not power-up the system administration (server) controller, rack leader controller (RLC) server or other service nodes.

```
# cpower --iru --down r1i0
```

This command powers down all the nodes in IRU 0 in rack 1. Note that **--down** is the same as **--off**. This command does not power-down the system administration node (server), rack leader controller server or other service nodes.

See “Console Management Power (cpower) Commands” on page 8 for additional information on power-on, power-off and power status commands. The *SGI Tempo System Administrator's Guide*, (P/N 007-4993-00x) has more extensive information on these topics.

System Overview

This chapter provides an overview of the physical and architectural aspects of your SGI Altix Integrated Compute Environment (ICE) 8400 series system. The major components of the Altix ICE systems are described and illustrated.

Because the system is modular, it combines the advantages of lower entry-level cost with global scalability in processors, memory, InfiniBand connectivity and I/O. You can install and operate the Altix ICE 8400 series system in your lab, or server room. Each 42U SGI rack holds from one to four 10U-high individual rack units (IRUs) that support up to sixteen compute/memory cluster sub modules known as “blades.” These blades are single printed circuit boards (PCBs) with ASICS, processors, memory components and I/O chip sets mounted on a mechanical carrier. The blades slide directly in and out of the IRU enclosures. Every compute blade contains at least two dual-inline memory module (DIMM) memory units.

Each blade supports two processor sockets that can use either Intel or AMD processors. Note that a maximum system size of 64 compute/memory blades per rack is supported at the time this document was published. Optional chilled water cooling may be required for large processor count rack systems. Customers wishing to emphasize memory capacity over processor count may request blades configured with only one processor installed per blade. Contact your SGI sales or service representative for the most current information on these topics.

The SGI Altix ICE 8400 series systems can run parallel programs using a message passing tool like the Message Passing Interface (MPI). The ICE blade system uses a distributed memory scheme as opposed to a shared memory system like that used in the SGI Altix UV series of 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. This chapter consists of the following sections:

- “System Models” on page 20
- “Intel System and Blade Architectures” on page 23
- “AMD System and Blade Architectures” on page 25
- “System Features and Major Components” on page 32
- “System Components” on page 41

System Models

Figure 3-1 shows an example configuration of a single-rack Altix ICE 8400 server.

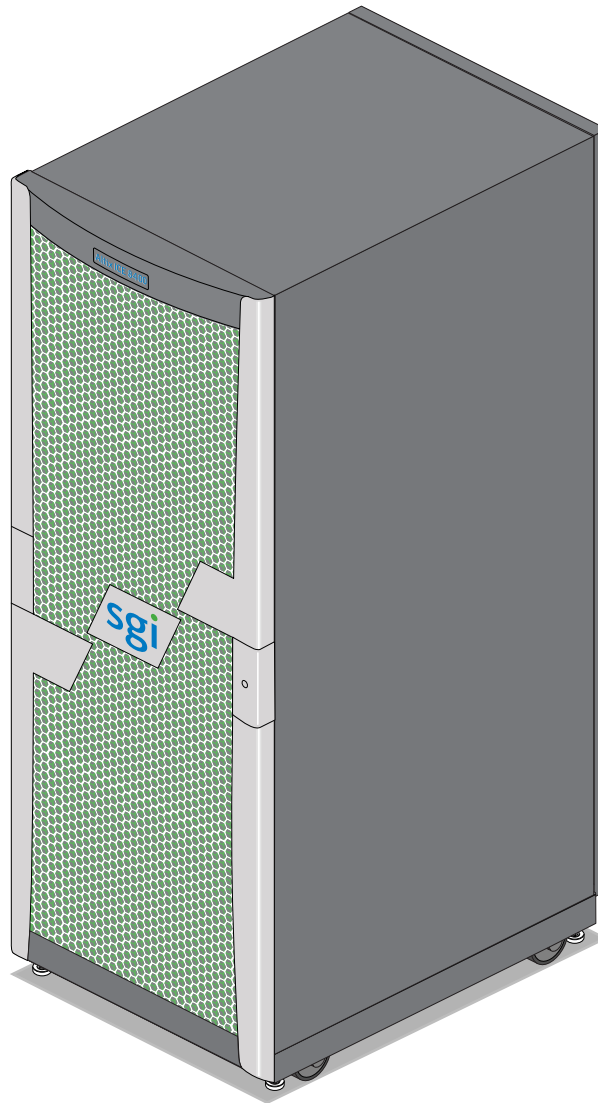


Figure 3-1 SGI Altix ICE 8400 Series System (Single Rack)

The 42U rack for this server houses all IRU enclosures, option modules, and other components; up to 1536 processor cores in a single rack. The basic enclosure within the Altix ICE system is the 10U high (17.5 inch or 44.45 cm) “individual rack unit” (IRU). The IRU enclosure supports a maximum of 16 single-wide compute/memory blades, up to eight power supplies, one chassis manager interface and one to four InfiniBand architecture I/O fabric switch interface blades. Note that optional water chilled rack cooling is available for systems in environments where ambient temperatures do not meet adequate air cooling requirements.

The system requires a minimum of one 42U tall rack with two PDUs per IRU. Each single-phase PDU has 8 outlets; two PDUs must be used with the first IRU and its support servers. Subsequent IRU’s installed into the rack are supported by two single-phase PDUs each.

Figure 3-2 shows an IRU and Rack. The optional three-phase PDU has 18 outlets and two PDUs are installed in each ICE 8400 compute rack. You can also add additional RAID and non-RAID disk storage to your rack system and this should be factored into the number of required outlets.

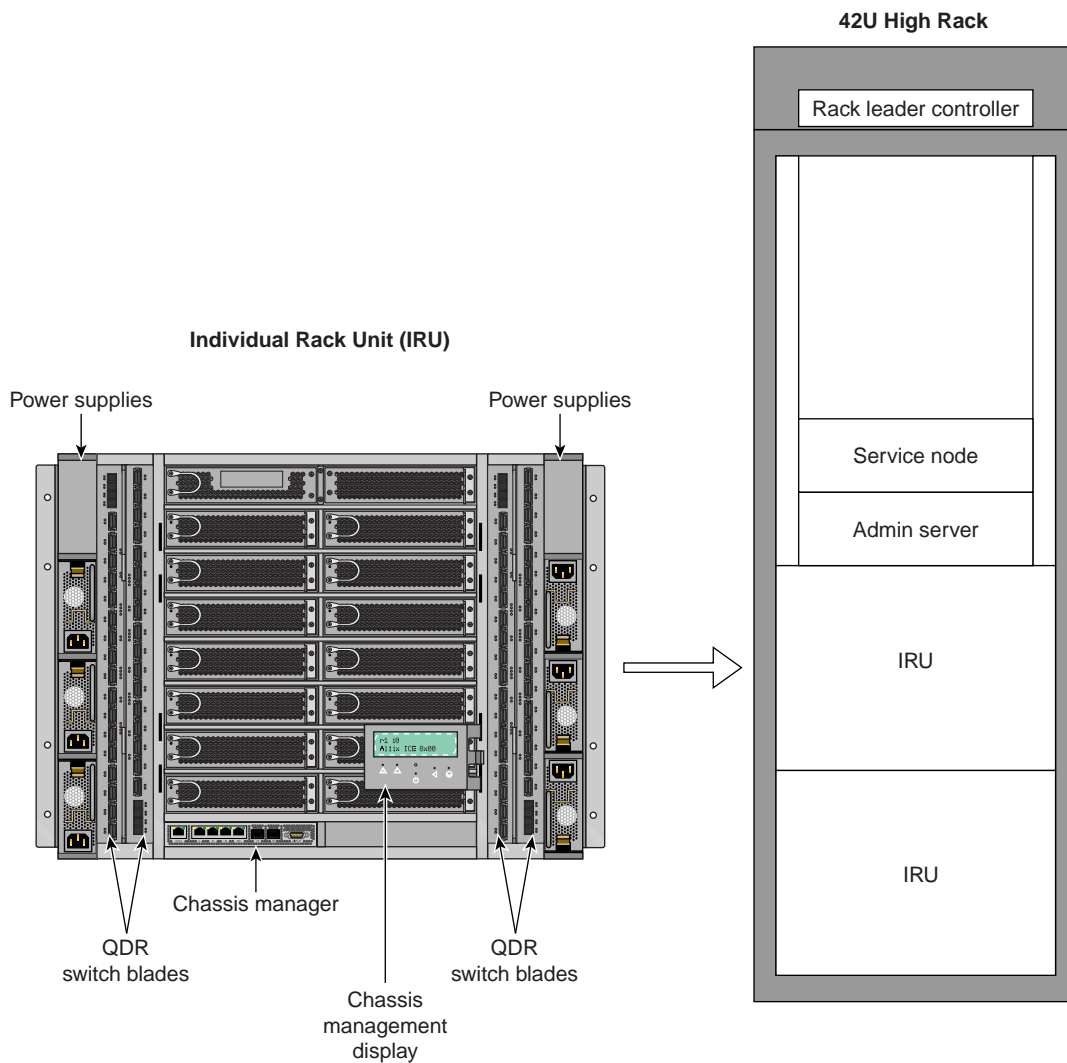


Figure 3-2 IRU and Rack Components Example

Intel System and Blade Architectures

The Altix ICE 8400 series of computer systems are based on an InfiniBand I/O fabric. This concept is supported and enhanced by using the Intel blade-level technologies described in the following subsections.

Depending on the configuration you ordered and your high-performance compute needs, your system may be equipped with different single-wide blade types. These compute blades all use Intel chip sets and different quad-data rate (QDR) InfiniBand on-board host-controller adapters (HCAs) as follows:

- The IP101 compute/memory blade uses a single-port InfiniBand host controller adapter (HCA).
- The IP103 compute/memory blade is equipped with one dual-port InfiniBand HCA.
- The IP105 version of the system blade uses two single-port InfiniBand HCAs.

Note: The IP105 compute blade is approximately 1/2-inch (12.7 mm) longer than the IP101 or IP103. The faceplate of the IP105 blade will protrude outward from the face of the IRU about 1/2-inch (12.7 mm) while the IP101 and IP103 blades will fit in the IRU with their faceplates flush to the front of the unit.

IP101 Blade Architecture Overview

An enhanced and updated version of the SGI Altix ICE compute blade is used in the ICE 8400 systems. Each blade uses one or two six-core or four-core Intel processors. The IP101 compute blade is **not** compatible with and cannot be used in “previous generation” Altix ICE 8200 series IRUs. This blade architecture is described in the following sections. The primary difference between the IP101 and IP103 blades is the single-port vs. dual port imbedded InfiniBand HCA. Note that the IP101 compute blade can only be used in single-plane configurations of the ICE 8400 system.

The compute blade contains the processors, memory, and one QDR InfiniBand single-port imbedded HCA. As previously mentioned, each compute blade is configured with one or two six-core or quad-core Intel processors - a maximum of 12 processor cores per compute blade. A maximum of twelve DDR3 memory DIMMs are supported per compute blade.

The two processors on the IP101 maintain an interactive communication link using Intel's QuickPath Interconnect (QPI) technology. This high-speed interconnect technology provides data transfers between the processors, memory and I/O hub components. Note the IP101 blade does not support a native "on-board" hard disk drive option.

IP103 Blade Architecture Overview

An enhanced and updated six-core or quad-core version of the SGI Altix ICE compute blade is used in the ICE 8400 systems. The IP103 compute blade is **not** compatible with and cannot be used in "previous generation" Altix ICE 8200 series IRUs. This blade architecture is described in the following sections.

The compute blade contains the processors, memory, and one QDR InfiniBand dual-port imbedded HCA. Each compute blade is configured with one or two six-core or quad-core Intel processors - a maximum of 12 processor cores per compute blade. A maximum of twelve DDR3 memory DIMMs are supported per compute blade.

The two processors on the IP103 maintain an interactive communication link using Intel's QuickPath Interconnect (QPI) technology. This high-speed interconnect technology provides data transfers between the processors, memory and I/O hub components. Note that the IP103 blade does not support a native "on-board" hard disk drive option.

IP105 Blade Architecture Overview

Although the Intel based compute blades used in the Altix ICE 8400 products are different physically, their basic compute architecture is nearly identical. The primary functional difference is that the larger IP105 blade supports two single-port QDR InfiniBand ASICs and an optional on-board hard disk drive or solid-state disk (SSD). The two compute blades are virtually the same in terms of available numbers of processors, (maximum of 12 cores) on-board memory control, QPI interfaces, DIMM types used and I/O control interfaces.

Figure 3-3 on page 29 shows a functional block diagram of the Altix ICE 8400 series system IRU using IP105 single-wide compute blades, InfiniBand interface, and component interconnects.

QuickPath Interconnect Features

Each processor on an Intel blade uses two QuickPath Interconnect (QPI) links. The QPI link consists of two point to point 20 bit channels - one send channel and one receive channel. The QPI link has a theoretical maximum aggregate bandwidth of 25.6 GB/s. Each blade's I/O chip set supports two processors. Each processor is connected to one of the I/O chips with a QPI channel. The two processors and the I/O chips are also connected together with a single QPI channel.

The maximum bandwidth of a single QPI link is calculated as follows:

- The QPI channel uses a 3.2 GHz clock, but the effective clock rate is 6.4 GHz because two bits are transmitted at each clock period -once on the rising edge of the clock and once on the falling edge (DDR).
- Of the 20 bits in the channel, 16 bits are data and 4 bits are error correction.
- 6.4 GHz times 16 bits equals 102.4 bits per clock period.
- Convert to bytes: 102.4 divided by 8 equals 12.8 GB/s (the maximum single direction bandwidth)
- The total aggregate bandwidth of the QPI channel is 25.6 GB/s: (12.8 GB/s times 2 channels)

AMD System and Blade Architectures

The Altix ICE 8400 series of computer systems are based on an InfiniBand I/O fabric. An AMD based eight-core or 12-core version of the SGI Altix ICE compute blade is used in the ICE 8400 systems. This concept is supported and enhanced by using the AMD blade-level technologies described in the following subsections.

Depending on the configuration you ordered and your high-performance compute needs, your system may be equipped with different single-wide blade types. These compute blades use AMD chip sets and different quad-data rate (QDR) InfiniBand on-board host-controller adapters (HCAs) as follows:

- The IP106 version of the system blade uses two single-port InfiniBand host controller adapter (HCAs).
- The IP110 version of the system blade uses one dual-port InfiniBand host controller adapter (HCA).

AMD HyperTransport Communication Features

The AMD based blade uses processors supporting the HyperTransport I/O scheme. This technology provides three 16-bit direct I/O links that provide a combined data transfer rate of up to 19.2 GB/sec at 800 MHz.

The HyperTransport links can be used for high-speed communication between the blade's two on-board processors. They can also be used as a low-latency interface to on-board high-speed I/O bridge chips. The separate memory and I/O paths within the blade help eliminate most bus contention

IP106 Blade Architecture Overview

Although the AMD based compute blades used in the Altix ICE products are slightly different physically, their basic compute architecture is nearly identical. The primary functional difference is that the IP106 blade supports two single-port QDR InfiniBand ASICs as well as an optional on-board hard disk drive or solid-state disk (SSD).

The two AMD based compute blades are virtually the same in terms of available numbers and types of processors, on-board memory control, HyperTransport interfaces, DIMM types used and I/O control interfaces.

Note that the two single-port QDR InfiniBand ASICs provide increased network bandwidth to the IRU backplane.

IP110 Blade Architecture Overview

The IP110 blade supports two sockets for a total of 24 AMD processor cores per blade. This blade cannot be mixed in an IRU with Intel based blades and is **not** compatible with and cannot be used in "previous generation" Altix ICE 8200 series IRUs. The IP110 blade architecture is described in the following paragraphs.

The compute blade contains the processors, memory, and one QDR InfiniBand dual-port imbedded HCA. The single-plane 4x QDR InfiniBand (IB) interconnect topology allows for lower-cost.

Each compute blade is configured with one or two 4-channel AMD processors. A maximum of eight DDR3 memory DIMMs are supported per populated processor compute socket (two DIMMs per channel).

The two processors on the IP110 maintain an interactive communication link using AMD HyperTransport technology. This high-speed interconnect technology provides data transfers between the processors, memory and I/O hub components.

Blade Memory Features

The memory control circuitry is integrated into the processors and provides greater memory bandwidth and capacity than previous generations of ICE compute blades.

Intel Blade DIMM Memory Features

Note that each processor on an Intel blade uses three DDR3 memory channels with one or more memory DIMMs on each channel (depending on configuration selected). Each blade can support up to 12 DIMMs. The DDR3 memory channel supports a maximum memory bandwidth of up to 10.66 GB per second. The combined maximum bandwidth for all three memory channels on a single processor is 25.6 GB per second.

AMD Blade DIMM Memory Features

Each AMD based blade supports up to sixteen DDR3 memory DIMMs. The memory controller on each Opteron processor supports four DDR3 memory channels. Each channel supports up to two DDR3 SDRAM DIMMs for a total of eight memory DIMMs. Each AMD node board can support up to 32 DIMMs or a maximum memory size of 512 GB physical memory. A total of 72 bits of data are transferred each clock period: 64 bits of data and 8 bits for error correction

Memory Channel Recommendation

It is highly recommended (though not required) that each processor on either the Intel or AMD based blades be configured with a minimum of one DIMM for each memory channel on a processor. This will help to ensure the best DIMM data throughput.

Blade DIMM Bandwidth Factors

The memory bandwidth on either the Intel or AMD based blades is generally determined by three key factors:

- The processor speed - different processor SKUs support different DIMM speeds.
- The number of DIMMs per channel.

- The DIMM speed - the DIMM itself has a maximum operating frequency or speed. At the time this document was published the ICE 8400 DIMM speed was 1333 MT/s.

Note: A DIMM must be rated for the maximum speed to be able to run at the maximum speed. For example: a single 1066 MT/s DIMM on a channel will only operate at 1066 MT/s - not 1333 MT/s.

As an example, populating one 1333 MT/s DIMM on each channel of an Intel based blade delivers a maximum of 10.66 GB/s per channel or 31.99 GB/s total memory bandwidth. The QuickPath Interconnect technology allows memory transfer or retrieval between the blade's two processors at up to 25.6 GB per second.

A minimum of one dual-inline-memory module (DIMM) is required for each processor on a blade. An example IRU with all blade slots filled is shown in Figure 3-10 on page 42. Each of the DIMMs on a blade must be the same capacity and functional speed. When possible, it is generally recommended that all blades within an IRU use the same number and capacity (size) DIMMs.

Note: Regardless of the number of DIMMs installed, a minimum of 4 GB of DIMM memory is recommended for each compute blade. Failure to meet this requirement may have impact on overall application performance.

Each blade in the IRU may have a different total DIMM capacity. For example, one blade may have 12 DIMMs, and another may have only six. Note that while this difference in capacity is acceptable functionally, it may have impact on compute "load balancing" within the system.

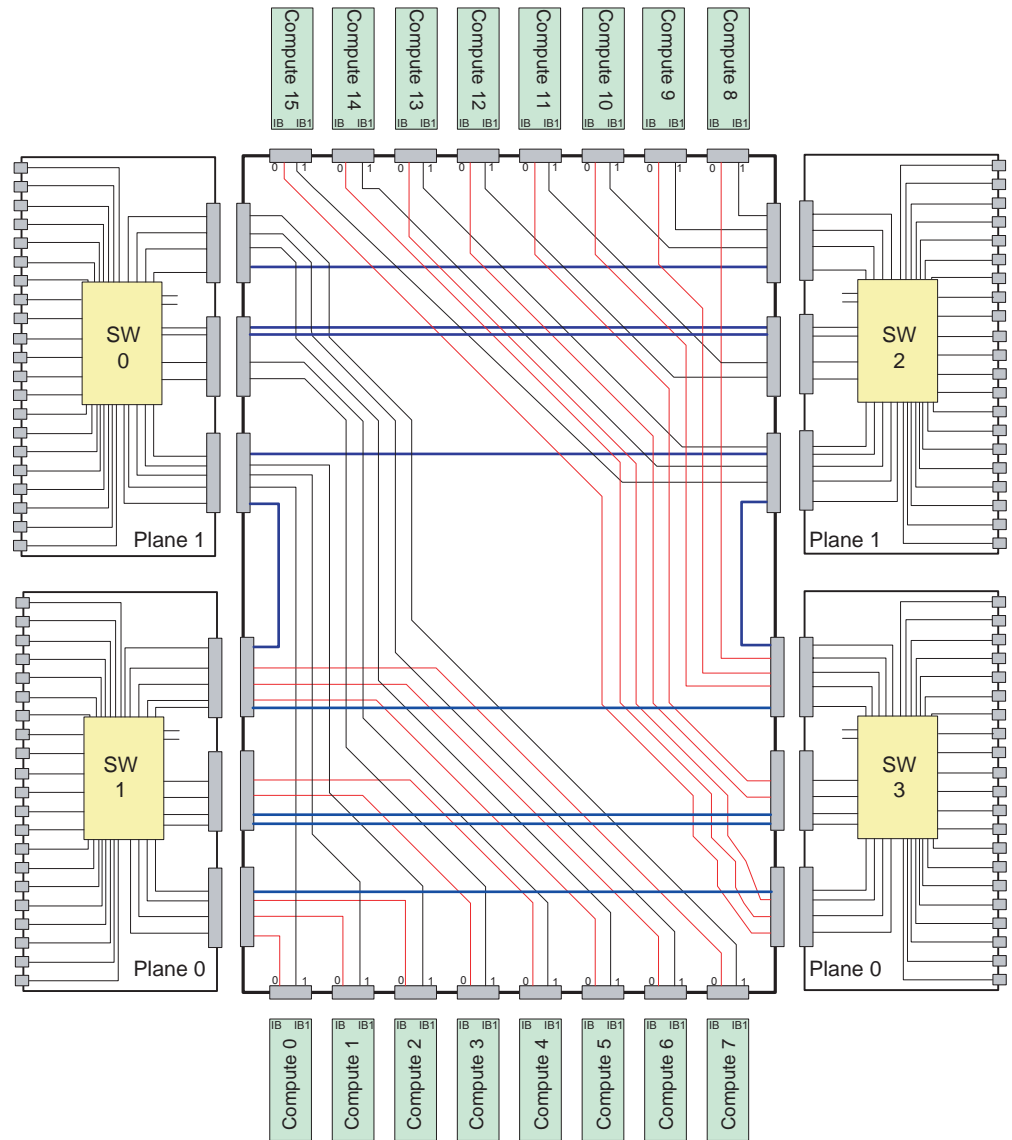


Figure 3-3 Functional Block Diagram of an ICE 8400 Individual Rack Unit (IRU)

System InfiniBand Switch Blades

One, two, or four quad-data-rate (QDR) InfiniBand switch blades can be used with each IRU configured in the Altix ICE 8400 system. IRUs with four switch blades use a dual-plane topology that provides high-bandwidth communication between compute blades inside the IRU as well as blades in other IRUs.

IRU Switch Density Choices

Each Altix ICE 8400 system comes with a choice of two IRU backplane configurations.

- Eight-compute node first level switch backplane (four switch blades maximum)
- Sixteen-compute node first level switch backplane (two switch blades maximum)

The compute blades and power supplies are interchangeable between the two IRU types. The switch blades for each IRU are **not** interchangeable. The outward appearance of the two IRU types is very similar. The difference is the number, placement, and type of supported switch blades within the IRU.

IRUs using one or two QDR switch blades are available in certain specific configurations. The single-switch blade configuration supports a single-plane QDR InfiniBand topology only; check with your SGI sales or service representative for additional information on availability.

The eight-compute node connected switching backplane uses up to four switch blades and supports a dual-plane InfiniBand switching fabric for higher throughput with most applications.

The Altix ICE 8400 QDR switch blade example shown in Figure 3-4 has 21 external ports (four of these are mini-SAS ports) to support the InfiniBand fabric. Any external switch blade ports not used to support the IB system fabric may be connected to optional service nodes or InfiniBand mass storage. Check with your SGI sales or service representative for information on available options.

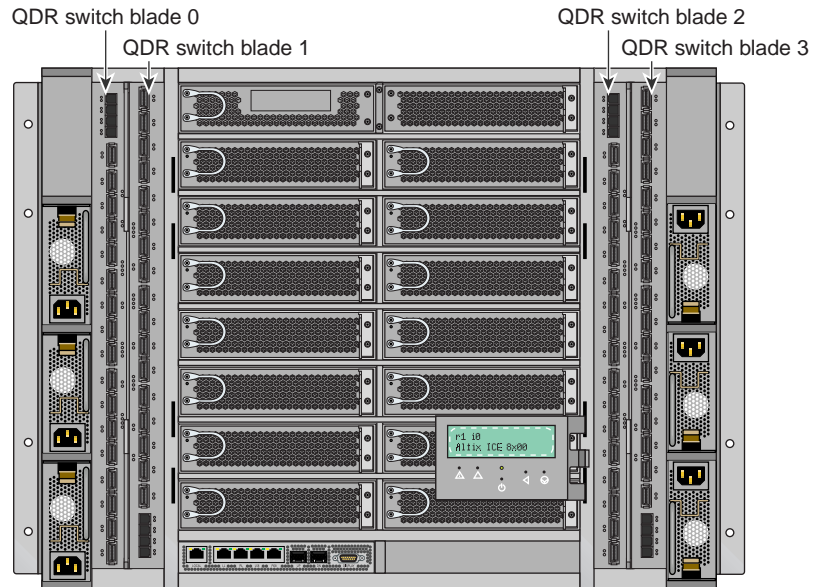


Figure 3-4 InfiniBand QDR Switch Example Numbering in IRUs

System Features and Major Components

The main features of the Altix ICE 8400 series server systems are introduced in the following sections:

- “Modularity and Scalability” on page 32
- “Reliability, Availability, and Serviceability (RAS)” on page 40
- “IRU (Unit) Numbering” on page 42

Modularity and Scalability

The Altix ICE 8400 series systems are modular, blade-based, scaleable, high-density cluster systems. The system rack components are primarily housed in building blocks referred to as individual rack units (IRUs). However, other “free-standing” SGI compute servers are used to administer, access and service the ICE 8400 series systems. Additional optional mass storage may be added to the system along with additional IRUs. You can add different types of stand-alone module options to a system rack to achieve the desired system configuration. You can configure and scale IRUs around processing capability, memory size or InfiniBand fabric I/O capability. The air-cooled IRU enclosure has redundant, hot-swap fans and redundant, hot-swap power supplies. A water-chilled rack option expands an ICE 8400 rack’s heat dissipation capability for the IRU components without requiring lower ambient temperatures in the lab or server room. See Figure 4-3 on page 49 for an example water-chilled rack configuration.

A number of free-standing (non-blade) compute and I/O servers (often referred to as nodes) are used with Altix ICE 8400 series systems in addition to the standard two-socket blade-based compute nodes. These free-standing units are:

- System administration controller
- System rack leader controller (RLC) server
- Service nodes with the following functions:
 - Fabric management service node
 - Login node
 - Batch node
 - I/O gateway node

Each ICE system will have one system administration controller, one rack leader controller (RLC) per system rack, and at least one service node.

The administration controller is a 2U server and the RLCs are integrated stand-alone 1U servers. The service nodes are integrated stand-alone non-blade 1U, 2U, 3U or 4U servers.

The following subsections further define the free-standing unit functions described in the previous list.

System Administration Server

There is one stand-alone administration controller server and I/O unit per system. The system administration controller is a non-blade SGI 2U server system. The server is used to install ICE system software, administer that software and monitor information from all the compute blades in the system. Check with your SGI sales or service representative for information on “cold spare” options that provide a standby administration server on site for use in case of failure.

The administration server on ICE 8400 systems is connected to the external network and may be set up for interactive logins under specific circumstances. However, most ICE systems are configured with dedicated “login” servers for this purpose. In this case, you might configure multiple “service nodes” and have all but one devoted to interactive logins as “login nodes”, see the “Login Server Function” on page 34 and the “I/O Gateway Node” on page 36.

Rack Leader Controller

A rack leader controller (RLC) server is generally used by administrators to provision and manage the system using SGI’s cluster management (CM) software. There is generally only one leader controller per rack and it is a non-blade “stand-alone” 1U server. The rack leader controller is guided and monitored by the system administration server. It in turn monitors, pulls and stores data from the compute nodes of all the IRUs within the rack. The rack leader then consolidates and forwards data requests received from the IRU’s blade compute nodes to the administration server. The leader controller may also supply boot and root file sharing images to the compute nodes in the IRUs.

For large systems or systems that run many MPI jobs, multiple RLC servers may be used to distribute the load (one RLC server per rack). The first RLC in the ICE system is the “master” controller server. Additional RLCs are slaved to the first RLC (normally installed in rack 1). The second RLC runs the same fabric management image as the primary “master” RLC. Check with your SGI sales or support representative for configurations that use a “cold spare” RLC or

administration server. This option can provide rapid replacement for a failed RLC or administrative unit.

In most ICE configurations the fabric management function is handled by the rack leader controller (RLC) node. The RLC is an independent server that is not part of an IRU. See the “Rack Leader Controller” on page 33 subsection for more detail. The fabric management software runs on one or more RLC nodes and monitors the function of and any changes in the InfiniBand fabrics of the system. It is also possible to host the fabric management function on a dedicated service node, thereby moving the fabric management function from the rack leader node and hosting it on an additional server(s). A separate fabric management server would supply fabric status information to the RLC server periodically or upon request. As with the rack leader controller server, only one per rack is supported.

Service Nodes

The functionality of the service “nodes” listed in this subsection are all services that can technically run on a single hardware server unit. As the system scales, you can add more servers (nodes) and dedicate them to these service functions if the size of the system requires it. However you can also have a smaller system where many of the services are combined on just a single service node. Figure 3-5 shows an example rear view of a 1U service node. Note that dedicated fabric management nodes are required on 8-rack or larger systems.

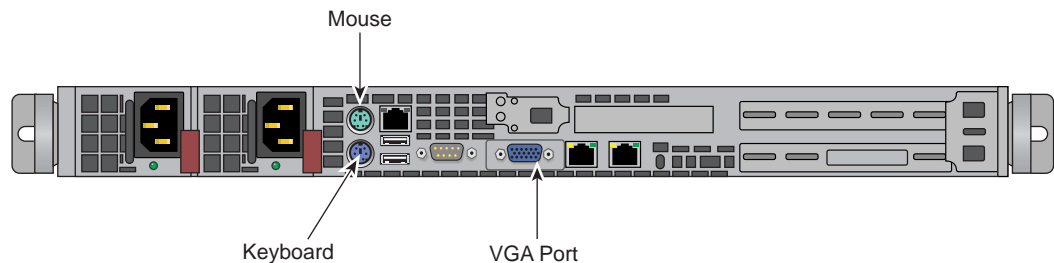


Figure 3-5 Example Rear View of a 1U Service Node

Login Server Function

The login server function within the ICE system can be functionally combined with the I/O gateway server node function in some configurations. One or more per system are supported. Very large systems with high levels of user logins may use multiple dedicated login server nodes. The login node functionality is generally used to create and compile programs, and additional login

server nodes can be added as the total number of user logins increase. The login server is usually the point of submittal for all message passing interface (MPI) applications run in the system. An MPI job is started from the login node and the sub-processes are distributed to the ICE system's compute nodes. Another operating factor for a login server is the file system structure. If the node is NFS-mounting a network storage system outside the ICE system, input data and output results will need to pass through for each job. Multiple login servers can distribute this load.

Figure 3-6 shows the rear connectors and interface slots on a 2U service node.

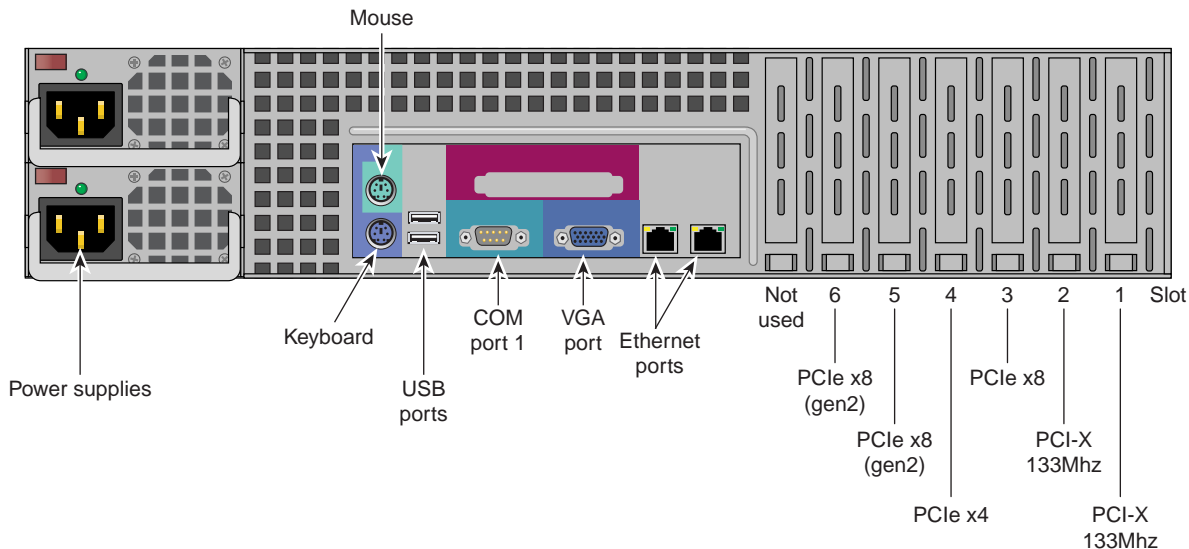


Figure 3-6 2U Service Node Rear Panel

Batch Server Node

The batch server function may be combined with login or other service nodes for many configurations. Additional batch nodes can be added as the total number of user logins increase. Users login to a batch server in order to run batch scheduler portable-batch system/load-sharing facility (PBS/LSF) programs. Users login or connect to this node to submit these jobs to the system compute nodes.

I/O Gateway Node

The I/O gateway server function may be combined with login or other service nodes for many configurations. If required, the I/O gateway server function can be an optional 1U, 2U or 3U stand-alone server within the ICE system. See Figure 3-7 for a rear view example of the 3U service node. One or more I/O gateway nodes are supported per system, based on system size and functional requirement. The node may be separated from login and/or batch nodes to scale to large configurations. Users login or connect to submit jobs to the compute nodes. The node also acts as a gateway from InfiniBand to various types of storage, such as direct-attach, Fibre Channel, or NFS.

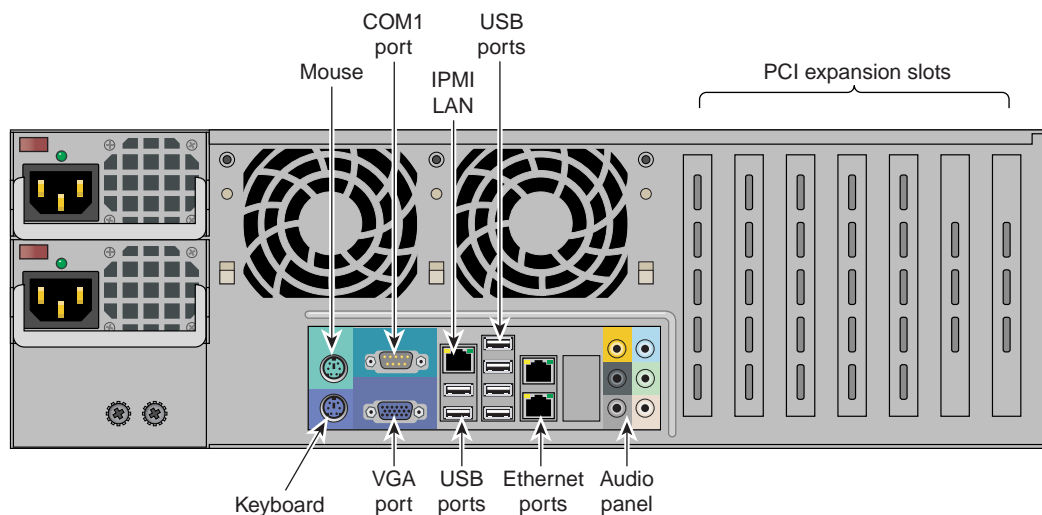


Figure 3-7 3U Service Node Rear Panel Example

The 4U Service Node

An optional 4U service node is offered with the ICE 8400 systems. This server is a higher-performance system that can contain multiple processors (up to 4) and serve multiple purposes within the ICE system. The 4U server is not used as an administrative node or rack leader controller.

Figure 3-8 on page 37 shows the rear panel of the 4U service node and Table 3-1 on page 37 identifies the functional items on the back of the unit. See the *SGI Altix UV 10 System User's Guide*, (P/N 007-5645-00x) for details on operating the 4U server.

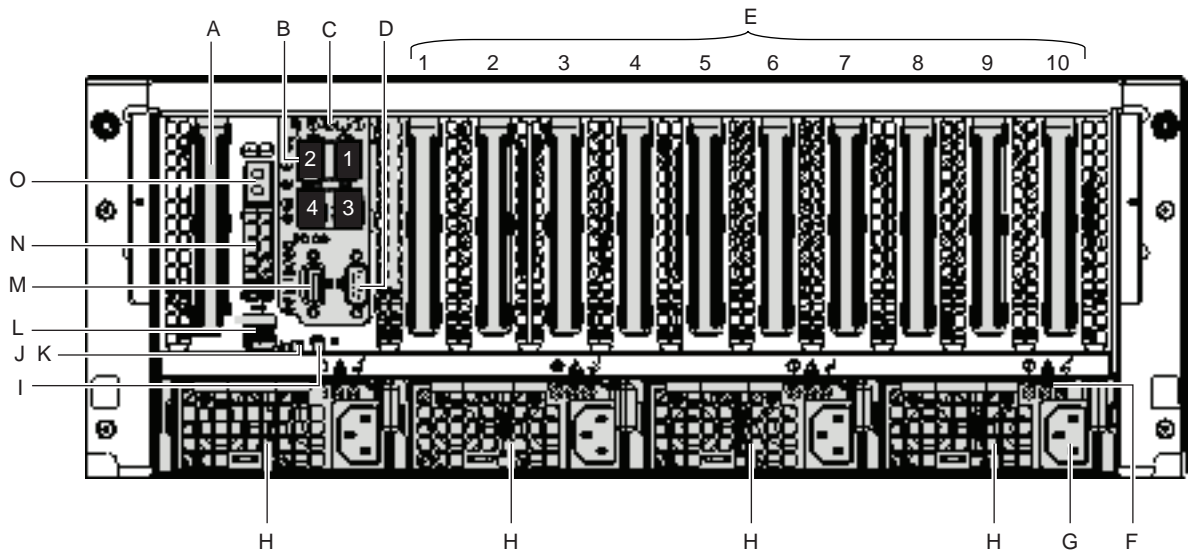


Figure 3-8 4U Service Node Rear Panel Example

Table 3-1 4U Service Node Rear Panel Items

Item	Description
A	SAS riser slot - PCIe Gen-2x8 half-height slot
B	I/O riser Gigabit Ethernet ports
C	I/O riser module
D	Serial port connector
E	PCIe Gen-2x8 slots
F	Power supply unit status LEDs
G	AC power input connectors
H	Hot-swap power supply
I	System ID on/off button
J	System status/fault LED
K	System ID LED (blue)

Table 3-1 **(continued)** 4U Service Node Rear Panel Items

Item	Description
L	USB 2.0 ports
M	VGA video port (up to 1600x1200) 15-pin connector
N	8 power on status test (POST) status LEDs
O	I/O riser management Ethernet port

Multiple Chassis Manager Connections

In certain multiple-IRU configurations the chassis managers in each IRU may be interconnected and wired to the administrative server and the rack leader controller (RLC) server. Figure 3-9 on page 39 shows an example diagram of the CMC interconnects between two ICE 8400 system racks.

Note: The unconnected chassis manager extension (blue) shown on the lower-right side of the figure illustrates a hypothetical continuation of the CMC network to a third ICE rack.

For more information on these and other topics related to the CMC, see Chapter 1 in the *SGI Tempo System Administrator's Guide*, (P/N 007-4993-00x).

Note also that the scale of the CMC drawings in Figure 3-9 on page 39 is adjusted to clarify the interconnect locations.

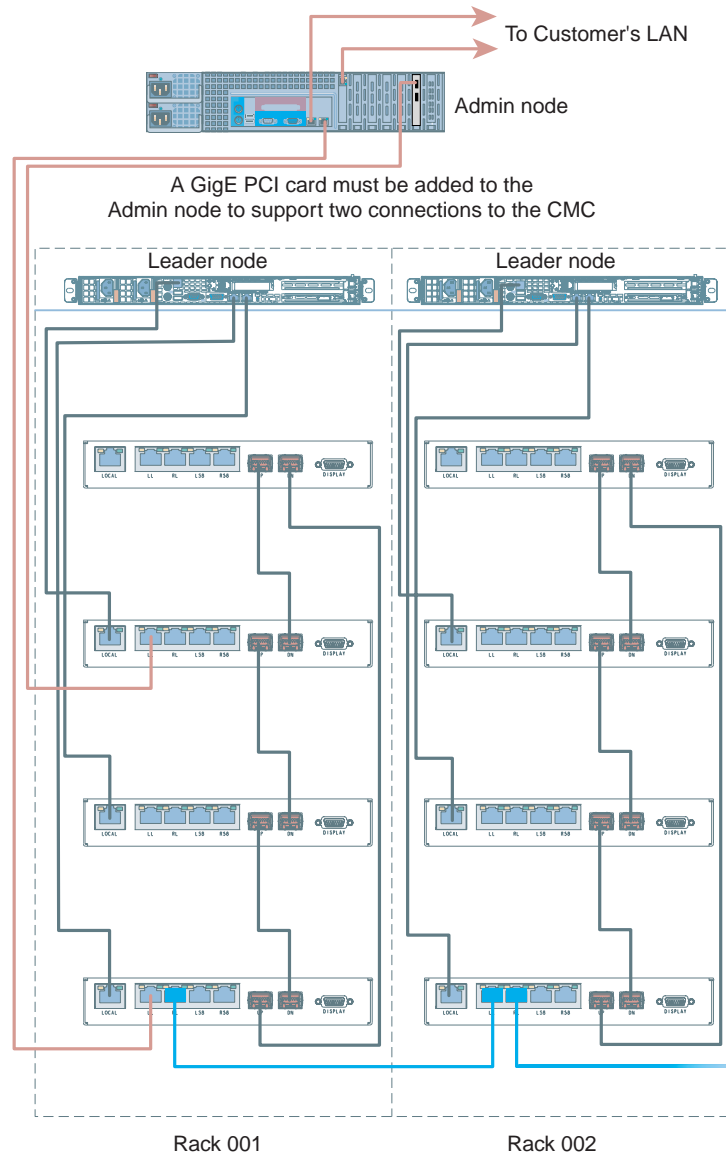


Figure 3-9 Administration and RLC Cabling to Chassis Managers Example

Reliability, Availability, and Serviceability (RAS)

The Altix ICE 8400 server series components have the following features to increase the reliability, availability, and serviceability (RAS) of the systems.

- **Power and cooling:**
 - IRU power supplies are redundant and can be hot-swapped under most circumstances. Note that this might not be possible in a “fully loaded” IRU.
 - A rack-level water chilled cooling option is available for all configurations.
 - IRUs have overcurrent protection at the blade and power supply level.
 - Fans are redundant and can be hot-swapped.
 - Fans run at multiple speeds in the IRUs. Speed increases automatically when temperature increases or when a single fan fails.
- **System monitoring:**
 - Chassis managers monitor the internal voltage, power and temperature of the IRUs.
 - Redundant system management networking is available.
 - Each IRU and each blade/node installed has failure LEDs that indicate the failed part; LEDs are readable at the front of the IRU.
 - Systems support remote console and maintenance activities.
- **Error detection and correction**
 - External memory transfers are protected by cyclical redundancy correction (CRC) error detection. If a memory packet does not checksum, it is retransmitted.
 - Nodes within each IRU exceed SECDED standards by detecting and correcting 4-bit and 8-bit DRAM failures.
 - Detection of all double-component 4-bit DRAM failures occur within a pair of DIMMs.
 - 32-bits of error checking code (ECC) are used on each 256 bits of data.
 - Automatic retry of uncorrected errors occurs to eliminate potential soft errors.
- **Power-on and boot:**
 - Automatic testing (POST) occurs after you power on the system nodes.
 - Processors and memory are automatically de-allocated when a self-test failure occurs.
 - Boot times are minimized.

System Components

The Altix ICE 8400 series system features the following major components:

- **42U rack.** This is a custom rack used for both the compute and I/O rack in the Altix ICE 8400 series. Up to 4 IRUs can be installed in each rack. Note that the primary (first) rack must have 3U of space reserved for the 2U administrative controller server and 1U rack leader controller (RLC) server.
- **Individual Rack Unit (IRU).** This enclosure contains the compute/memory blades, chassis manager, InfiniBand fabric I/O blades and front-access power supplies for the Altix ICE 8400 series computers. The enclosure is 10U high. Figure 3-10 on page 42 shows the Altix ICE 8400 series IRU system components.
- **Single-wide compute/memory blade.** Holds two (quad-core or six-core) processor sockets and up to 12 memory DIMMs.
- **1U (Rack leader controller).** The 1U rack leader server is required in each system rack.
- **2U Administrative server with PCIe expansion.** This server node supports an optional console, administrative software and PCI Express option cards. The administrative server is installed in the primary rack in the system.
- **1U Service node.** Additional 1U server(s) can be added to a system rack and used specifically as an optional login, batch, or fabric functional node. Note that these service functions cannot be incorporated as part of the system RLC server.
- **2U Service node.** An optional 2U service node may be used as a login, batch, or fabric functional node. In smaller systems these functions may be combined on one server. Note that the 2U service node function is never a shared part of the 2U administrative server.
- **3U Service node.** The optional 3U server node is offered with certain configurations needing higher performance I/O access for the ICE system. It offers multiple I/O options and graphics options not available with the 1U or 2U service nodes.
- **4U Service node.** The optional 4U server is offered as the highest overall performance service node available with the ICE 8400 system. It offers the highest processing power, best I/O performance and most flexible configuration options of the available service nodes.

Note: PCIe options may vary, check with your SGI sales or support representative.

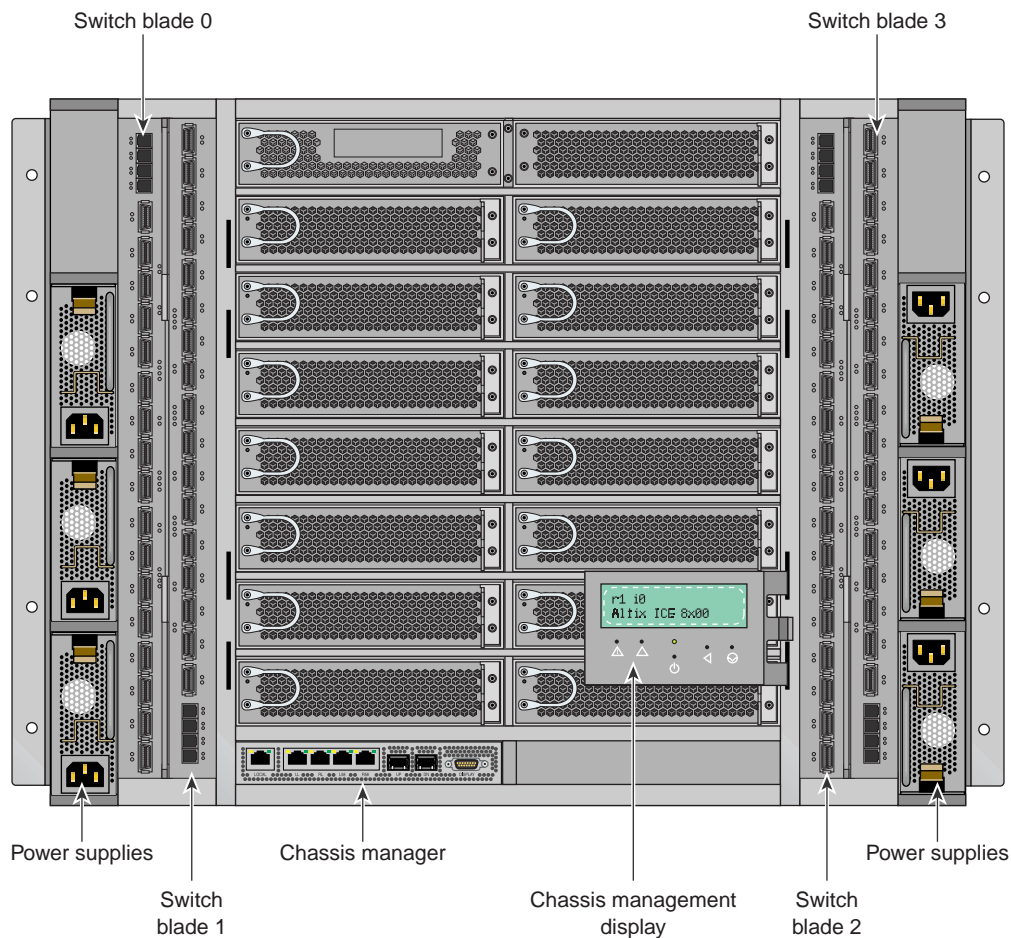


Figure 3-10 Altix ICE 8400 Series IRU System Components Example

IRU (Unit) Numbering

IRUs in the racks are not identified using standard units. A standard unit (SU) or unit (U) is equal to 1.75 inches (4.445 cm). IRUs within a rack are identified by the use of module IDs 0, 1, 2, and 3, with IRU 0 residing at the bottom of each rack. These module IDs are incorporated into the host names of the CMC (i0c, i1c, etc.) and the compute blades (r1i0n0, r1i1n0, etc.) in the rack.

Rack Numbering

Each rack in a multi-rack system is numbered with a single-digit number sequentially beginning with (001). A rack contains IRU enclosures, administrative and rack leader server nodes, service specific nodes, optional mass storage enclosures and potentially other options.

Note: In a single compute rack system, the rack number is always (1).

The number of the first IRU will always be zero (0). These numbers are used to identify components starting with the rack, including the individual IRUs and their internal compute-node blades. Note that these single-digit ID numbers are incorporated into the host names of the rack leader controller (RLC) (r1lead) as well as the compute blades (r1i0n0) that reside in that rack.

Optional System Components

Availability of optional components for the SGI ICE 8400 series of systems may vary based on new product introductions or end-of-life components. Some options are listed in this manual, others may be introduced after this document goes to production status. Check with your SGI sales or support representative for the most current information on available product options not discussed in this manual.

Rack Information

This chapter describes the physical characteristics of the tall (42U) ICE 8400 racks in the following sections:

- “Overview” on page 45
- “Altix ICE 8400 Series Rack (42U)” on page 46
- “ICE Rack Technical Specifications” on page 50

Overview

At the time this document was published only the tall (42U) Altix ICE rack (shown in Figure 4-2) was approved for ICE system racks shipped from the SGI factory.

Altix ICE 8400 Series Rack (42U)

The tall rack (shown in Figure 4-1 on page 47) has the following features and components:

- **Front and rear door.** The front door is opened by grasping the outer end of the rectangular-shaped door piece and pulling outward. It uses a key lock for security purposes that should open all the front doors in a multi-rack system (see Figure 4-2 on page 48).

Note: The front door and rear door locks are keyed differently. The optional water-chilled rear door (see Figure 4-3 on page 49) does not use a lock.

The standard rear door has a push-button key lock to prevent unauthorized access to the system. The rear doors have a master key that locks and unlocks all rear doors in a system made up of multiple racks. You cannot use the rear door key to secure the front door lock.

- **Cable entry/exit area.** Cable access openings are located in the front floor and top of the rack. Cables are only attached to the front of the IRUs; therefore, most cable management occurs in the front and top of the rack. Stand-alone administrative, leader and login server modules are the exception to this rule and have cables that attach at the rear of the rack. Rear cable connections will also be required for optional storage modules installed in the same rack with the IRU(s). Optional inter-rack communication cables pass through the top of the rack. I/O and power cables normally pass through the bottom of the rack.
- **Rack structural features.** The rack is mounted on four casters; the two rear casters swivel. There are four leveling pads available at the base of the rack. The base of the rack also has attachment points to support an optional ground strap, and/or seismic tie-downs.
- **Power distribution units in the rack.** Up to thirteen outlets are required for a single IRU system as follows:
 - up to 7 outlets for an IRU (depending on configuration)
 - 4 outlets for administration and RLC servers (in primary rack)
 - 2 outlets for a service node (server)
 - Allow six or seven more outlets for each additional IRU in the system

Two single-phase power distribution units (PDUs) are needed for a base rack system, (8 outlets per PDU). A three-phase power distribution unit has 18 outlet connections, up to 14 of them may be needed to power two IRUs.

Note: System IRUs that use high-wattage processors require seven power supplies and take up all but one of the available plugs in a single-phase PDU.

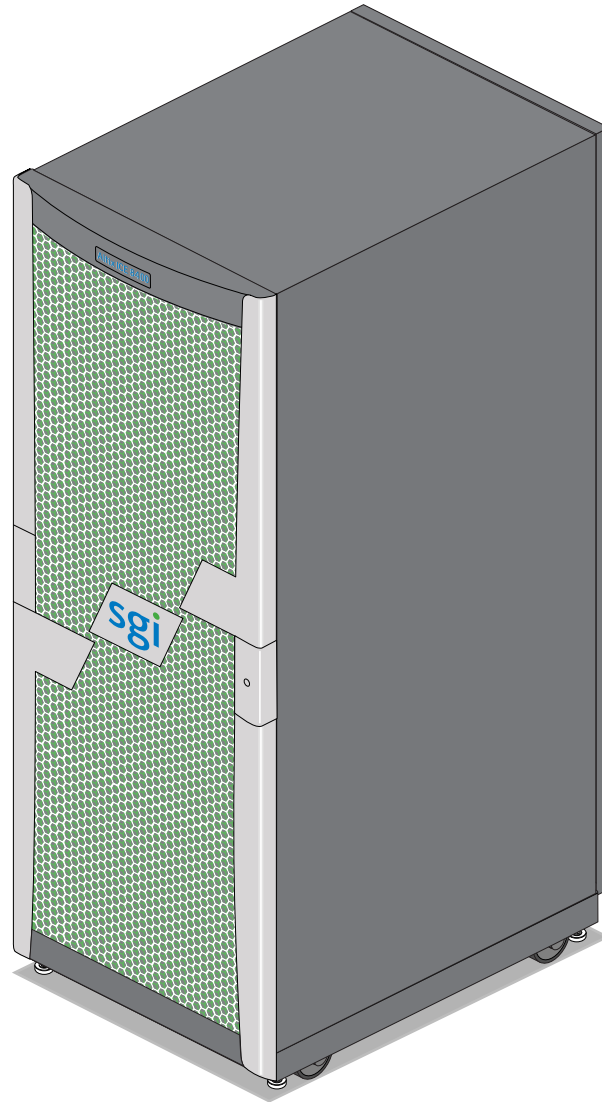


Figure 4-1 Altix ICE 8400 Series Rack Example

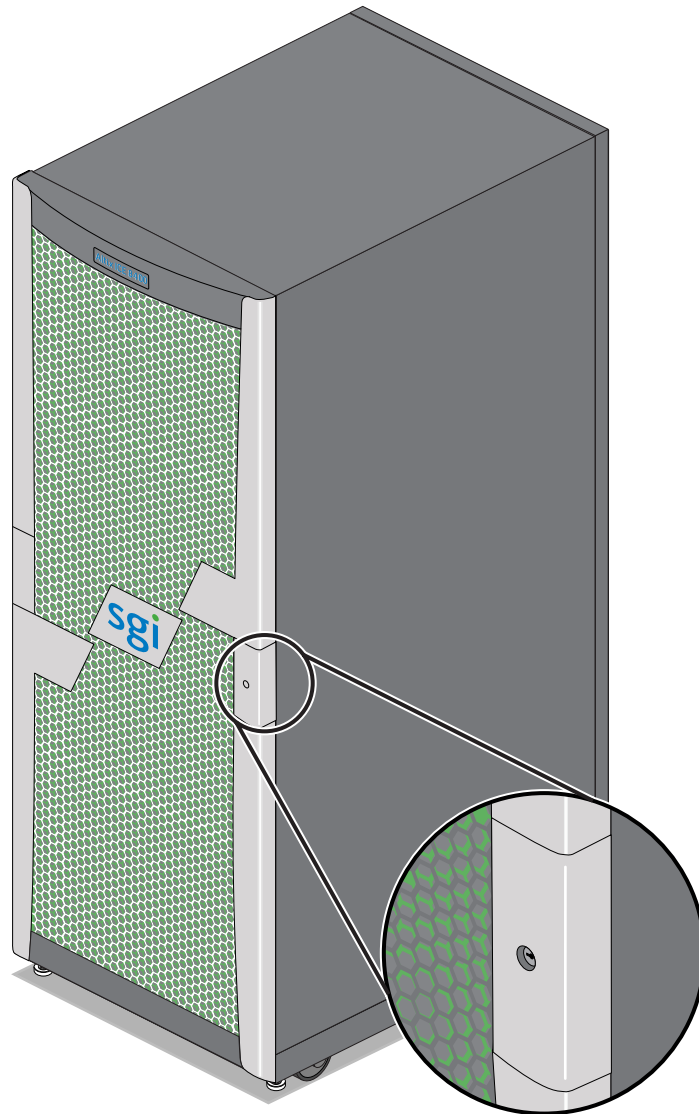


Figure 4-2 Front Lock on Tall (42U) Altix Rack

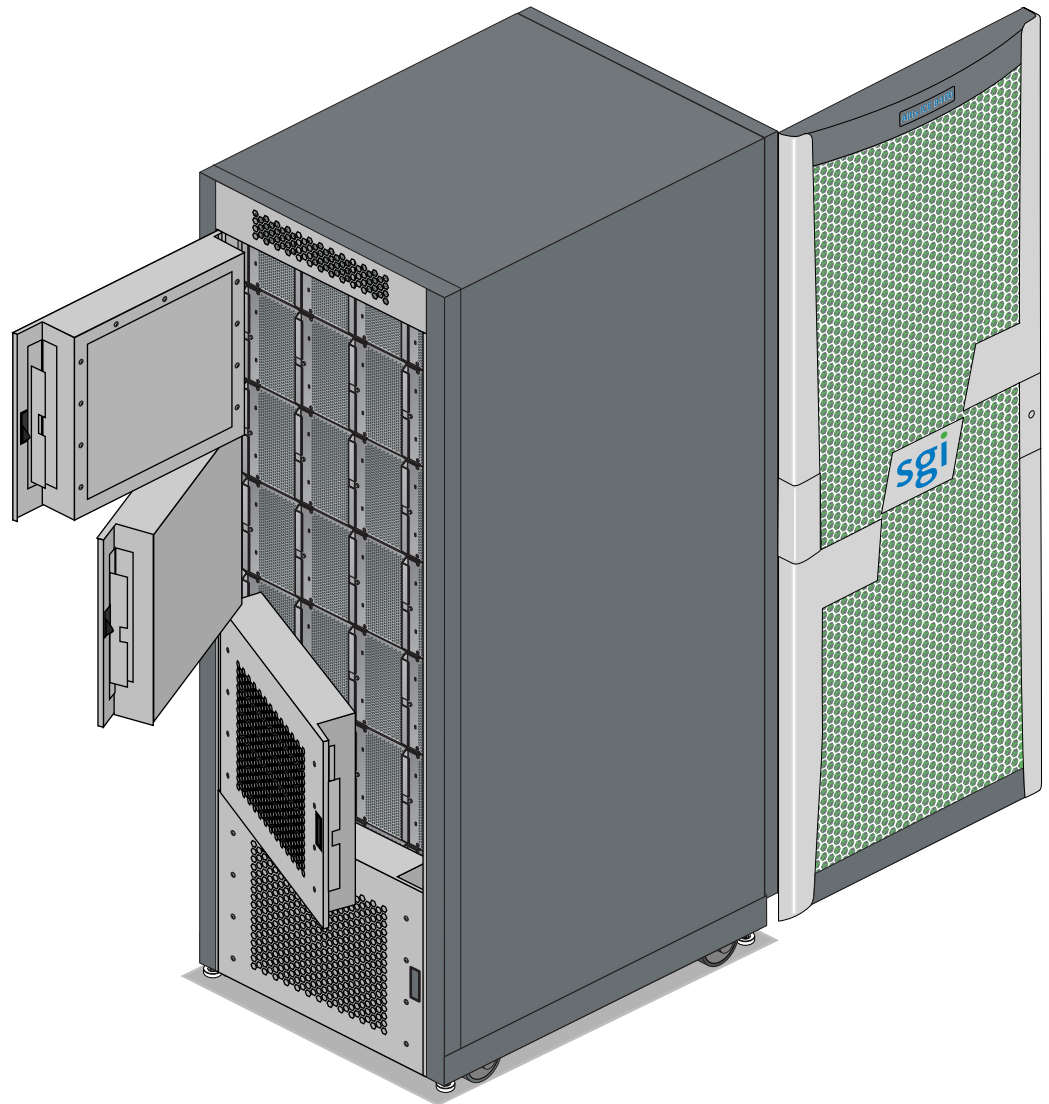


Figure 4-3 Optional Water-Chilled Doors on Rear of Altix Rack

ICE Rack Technical Specifications

Table 4-1 lists the technical specifications of the Altix ICE 8400 series tall rack.

Table 4-1 Tall Altix Rack Technical Specifications

Characteristic	Specification
Height	79.5 in. (201.9 cm)
Width	31.3 in. (79.5 cm)
Depth	45.8 in. (116.3 cm)
Weight (full)	2,490 lbs. (1,132 kg) approximate
Shipping weight (max)	2,535 lbs. (1,152 kg) approximate
Voltage range	North America/International
Nominal	200-240 VAC /230 VAC
Tolerance range	180-264 VAC
Frequency	North America/International
Nominal	60 Hz /50 Hz
Tolerance range	47-63 Hz
Phase required	Single-phase or 3-phase
Power requirements (max)	34.58 kVA (33.89 kW)
Hold time	16 ms
Power cable	12 ft. (3.66 m) pluggable cords

ICE Administration/Leader Servers

This chapter describes the function and physical components of the administrative/rack leader control servers (also referred to as nodes) in the following sections:

- “Overview” on page 51
- “Administrative/Controller Server” on page 53

For purposes of this chapter “administration/controller server” is used as a catch-all phrase to describe the stand-alone servers that act as management infrastructure controllers. The specialized functions these servers perform within the ICE system primarily include:

- Administration and management
- Rack leader controller (RLC) functions

Other servers can be configured to provide additional services, such as:

- Fabric management (usually used with 8-rack or larger systems)
- Login
- Batch
- I/O gateway (storage)

Note that these functions are usually performed by the system’s “service nodes” which are additional individual servers set up for single or multiple service tasks.

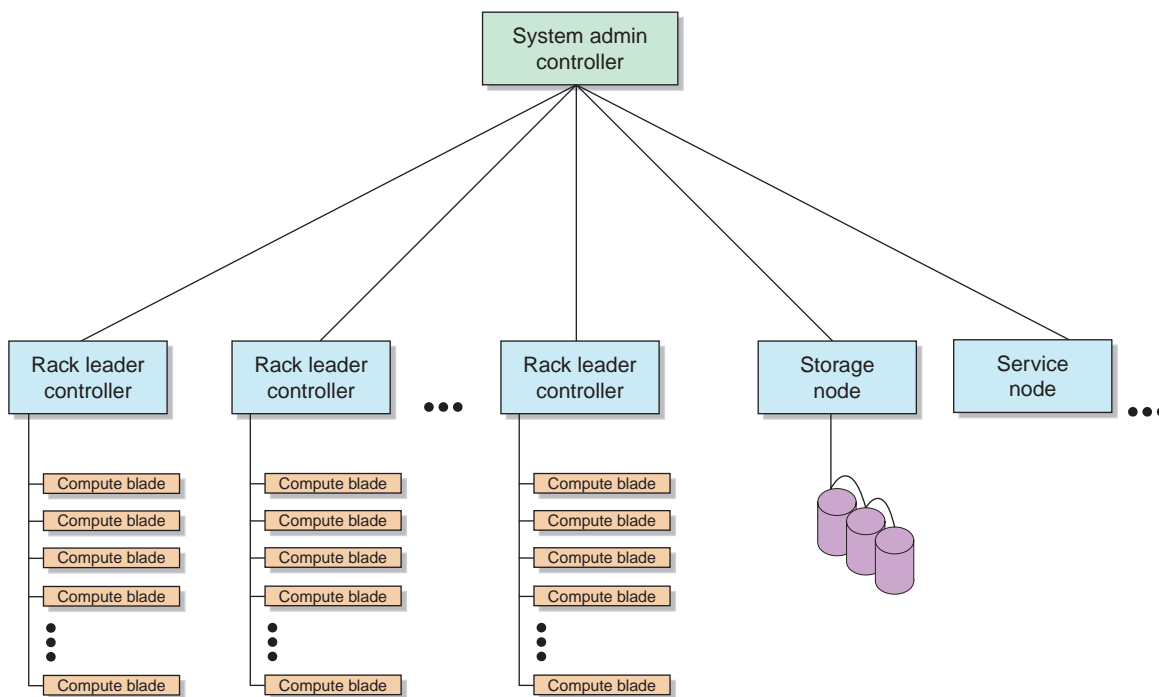
Overview

User interfaces consist of the Compute Cluster Administrator, the Compute Cluster Job Manager, and a Command Line Interface (CLI). Management services include job scheduling, job and resource management, Remote Installation Services (RIS), and a remote command environment. The 2U administrative controller server is connected to the system via a Gigabit Ethernet link, (it is not directly linked to the system’s InfiniBand communication fabric).

Note: The system management software runs on the administrative node, RLC and service nodes as a distributed software function. The system management software performs all of its tasks on the ICE system through an Ethernet network.

The administrative controller server is at the top of the distributed management infrastructure within the ICE system. The overall ICE 8400 series management is hierarchical (see Figure 5-1), with the RLC(s) communicating with the compute nodes via CMC.

System management hierarchy



A maximum of 64 compute blades per rack leader controller

Figure 5-1 ICE System Administration Hierarchy Example Block Diagram

Administrative/Controller Server

The system administrative controller unit acts as the ICE system’s primary interface to the “outside world”, typically a local area network (LAN). The server is used by administrators to provision and manage cluster functions using SGI’s cluster manager software.

For systems using a separate login, batch, I/O, fabric management, or other service node; this 2U server is also an available option. Figure 5-2 and Figure 5-3 show front and rear views of the 2U administration/service node. Note that the server uses up to 12 DIMM memory cards. This server is currently marketed as the SGI Rackable C2108-TY10; it was previously offered as the XE270.

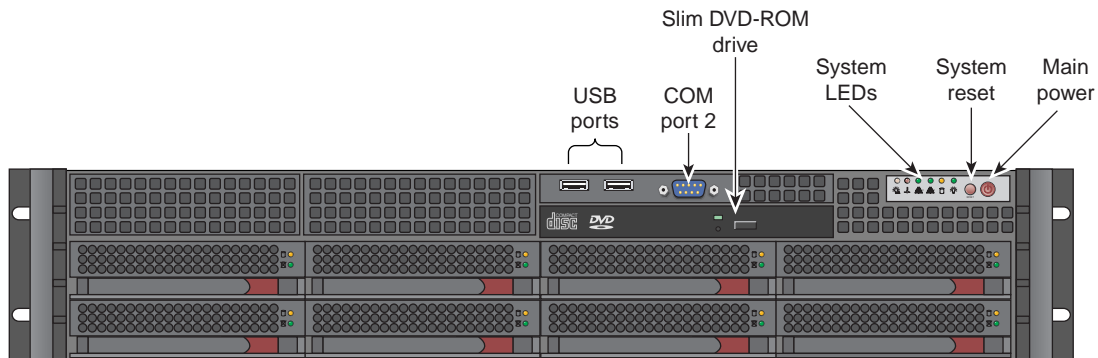


Figure 5-2 Front View of 2U Administration/Controller or Service Node

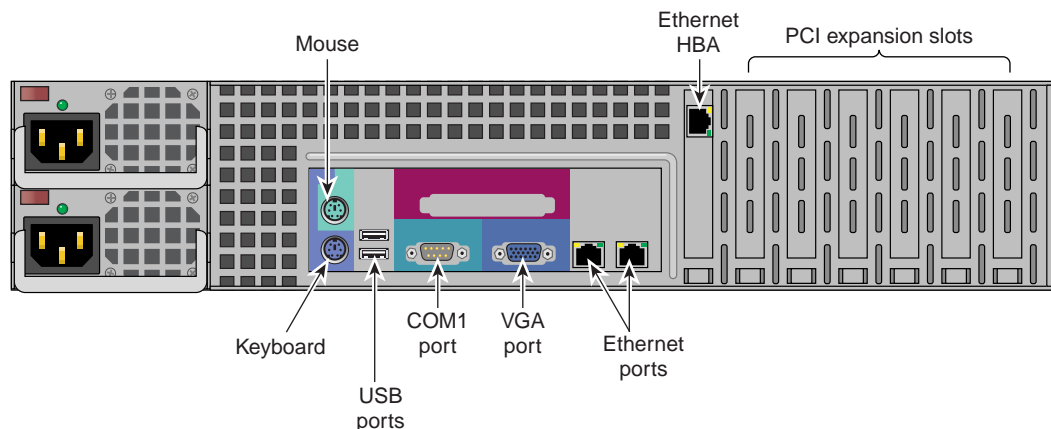


Figure 5-3 Rear View of 2U Administration/Controller or Service Node

See the *SGI Rackable C2108-TY10 System User's Guide* (P/N 007-5688-00x) for more detailed information on the 2U administration/controller or service node. For more information on the XE270, see the *SGI Altix XE270 User's Guide*, (P/N 007-5535-00x).

The administrative server's control panel features are shown in Figure 5-4.

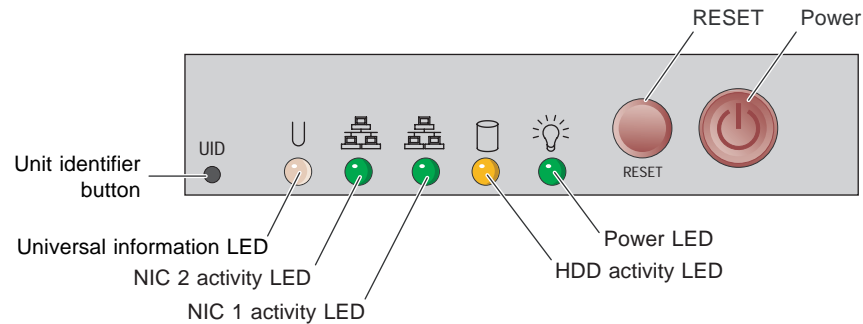


Figure 5-4 Administrative/Controller Server Control Panel Diagram

Table 5-1 System administrative server control panel functions

Functional feature	Functional description
Unit identifier button	Pressing this button lights an LED on both the front and rear of the server for easy system location in large configurations. The LED will remain on until the button is pushed a second time.
Universal information LED	This multi-color LED blinks red quickly, to indicate a fan failure and blinks red slowly for a power failure. A continuous solid red LED indicates a CPU is overheating. This LED will be on solid blue or blinking blue when used for UID (Unit Identifier).
NIC 2 Activity LED	Indicates network activity on LAN 2 when flashing green.
NIC 1 Activity LED	Indicates network activity on LAN 1 when flashing green.
Disk activity LED	Indicates drive activity when flashing.
Power LED	Indicates power is being supplied to the server's power supply units.
Reset button	Pressing this button reboots the server.
Power button	Pressing the button applies/ removes power from the power supply to the server. Turning off power with this button removes main power but keeps standby power supplied to the system.

1U Rack Leader Controller Server

An MPI job is started from the rack leader controller server and the sub-processes are distributed to the system blade compute nodes. The main process on the RLC server will wait for the sub-processes to finish. Note that every Altix ICE 8400 system is required to have at least one RLC. For multi-rack systems or systems that run many MPI jobs, multiple RLC servers are used to distribute the load (one per rack).

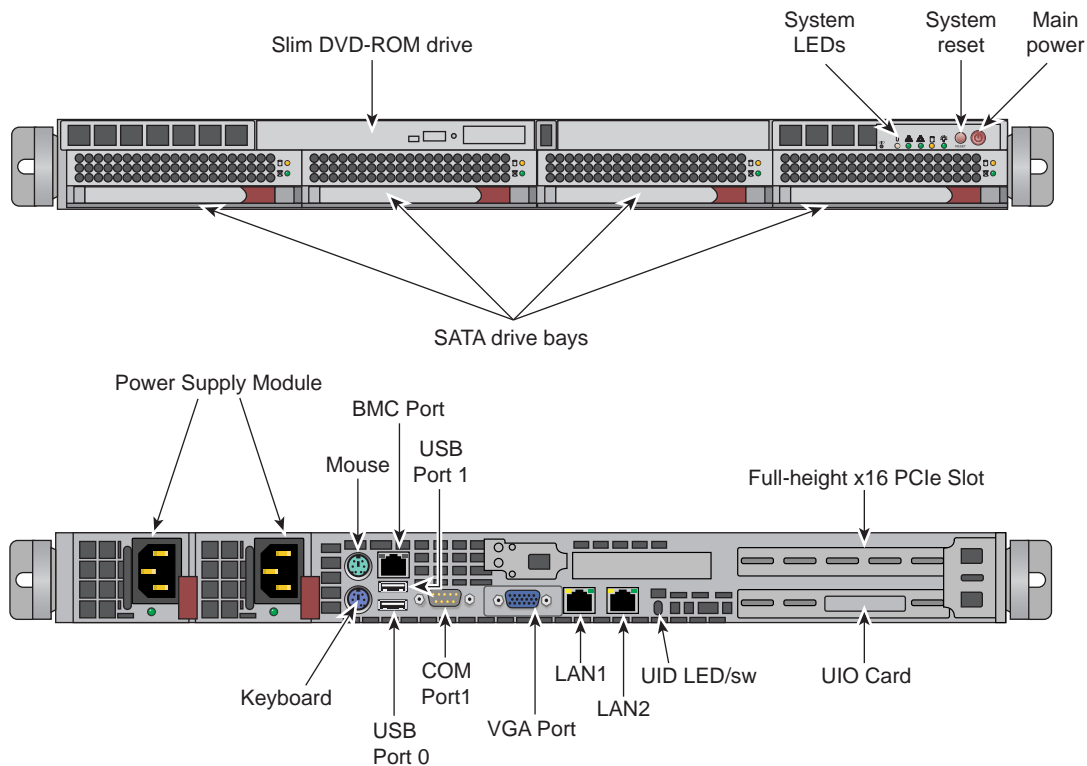


Figure 5-5 1U Rack Leader Controller (RLC) Server Front and Rear Panels

Batch or login functions most often run on individual separate service nodes, especially when the system is a large-scale multi-rack installation or has a large number of users. This 1U server may also be used as a separate (non-RLC) login, batch, I/O or fabric management node. See the section “Modularity and Scalability” on page 32 for a list of administration and support server types and additional functional descriptions.

Optional AMD 2U Service Nodes

The SGI Altix ICE 8400 offers an optional 2U AMD based service node. This server offers four AMD Opteron processors and up to 32 memory DIMMs. The unit supports up to six drives and four PCIe option slots. Figure 5-6 shows the front and rear features of the 2U service node.

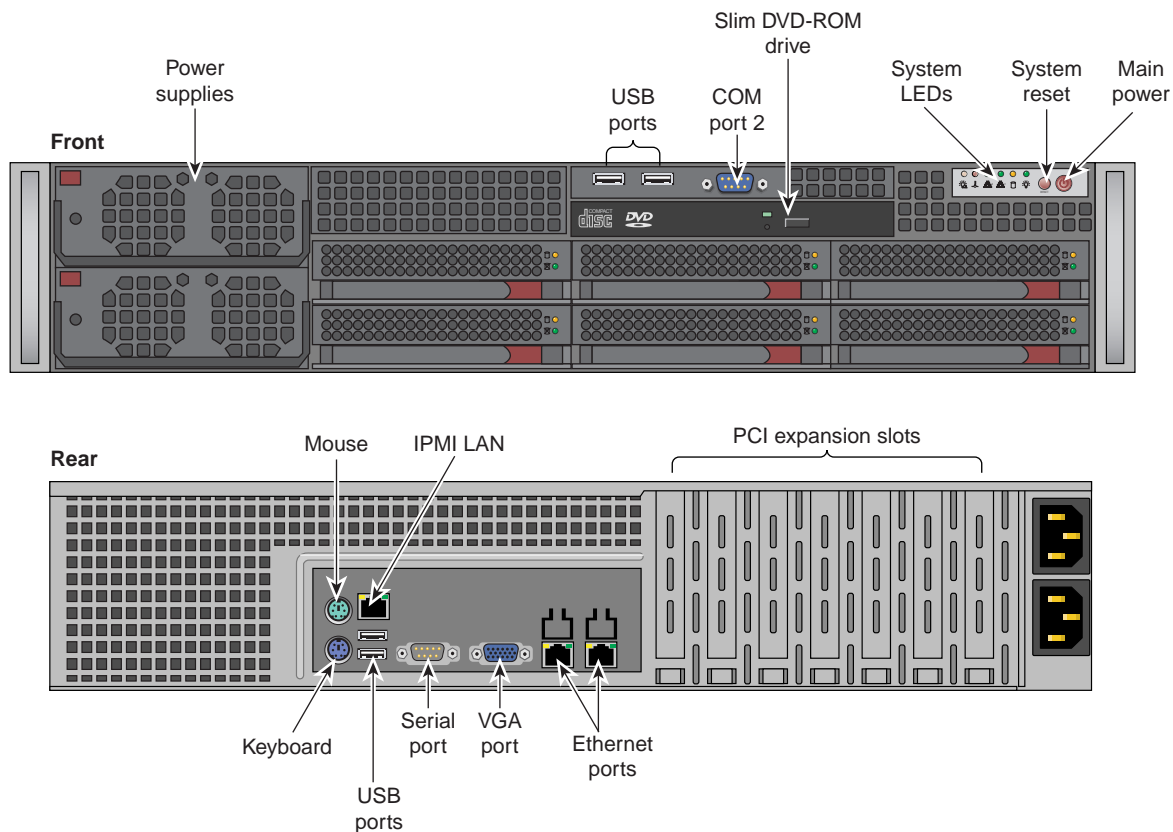


Figure 5-6 Optional 2U AMD Service Node Example Front and Rear View

For additional information on architecture, general operation, and descriptions of the major components in the 2U AMD service node, see the *SGI Rackable H2106-G7 System User's Guide*, (P/N 007-5683-00x). This document also provides basic troubleshooting and maintenance information, BIOS information, and important safety and regulatory specifications.

Optional 3U Service Nodes

The Altix ICE system also offers a 3U-high service node as a separate login, batch, I/O, fabric management, or graphics support node. Under specific circumstances the 3U server can be configured as a mass storage resource for the ICE system. Figure 5-7 shows an example front view of the optional server.

For more information on using the 3U service node, see the *SGI Rackable C3108-TY11 System User's Guide* (P/N 007-5687-00x). Note that this server was previously marketed as the Altix XE500, see the *SGI Altix XE500 System User's Guide*, (P/N 007-5572-00x).

Check with your SGI sales or service representative for more information on available graphics card options that can be used with the server in an ICE system.

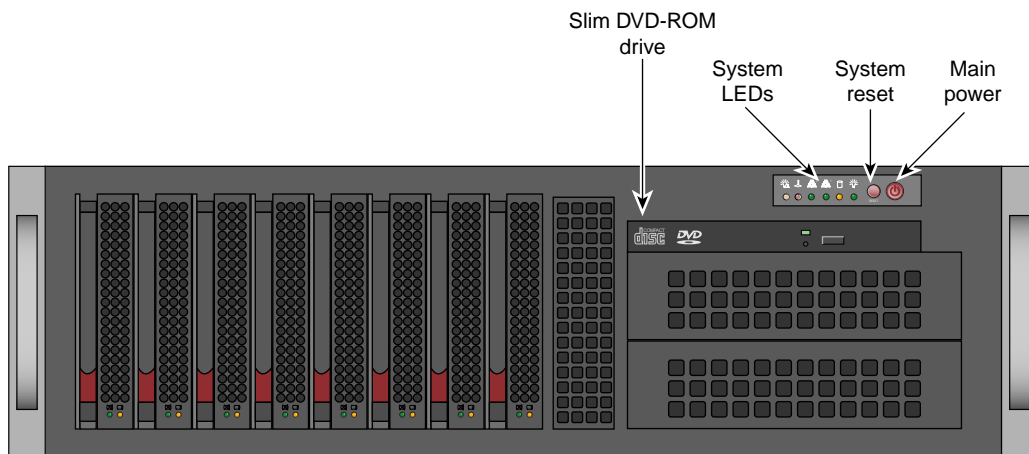


Figure 5-7 SGI 3U Optional Service Node Front View

Figure 5-8 on page 58 shows an example rear view of the 3U service node.

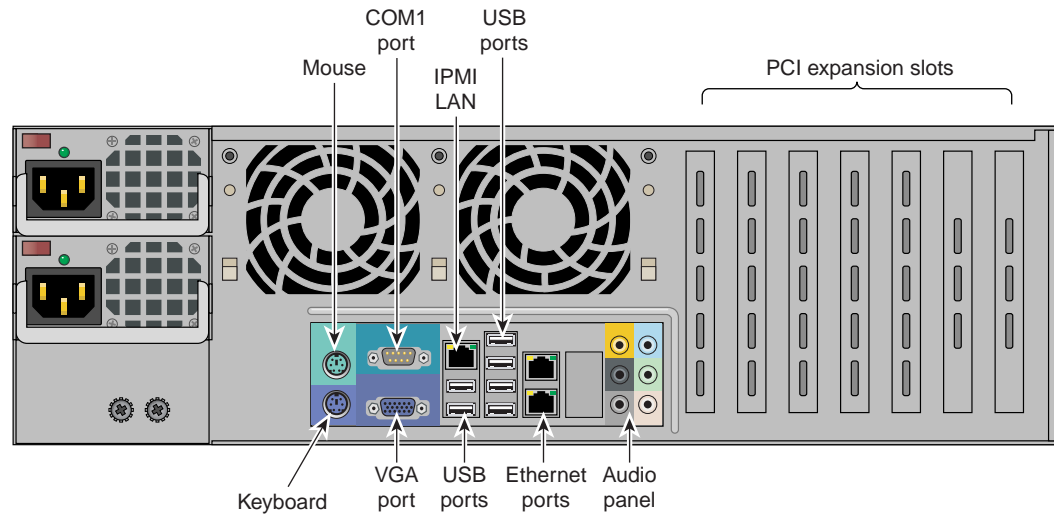


Figure 5-8 SGI 3U Service Node Rear View

Optional 4U Service Nodes

The highest performance optional service node in the Altix ICE 8400 system is offered as a 4U-high service node. It can serve as a separate login, batch, I/O, fabric management, or graphics support node, or combine several of these functions. Under specific circumstances the 4U server can be configured as a mass storage resource for the ICE system.

Figure 5-9 on page 59 shows the front controls and interfaces available on the server. Table 5-2 on page 59 describes the front panel control and interface functions on the 4U server.

Figure 5-10 on page 60 calls out the components used on the front of the 4U server. Table 5-3 on page 60 identifies the components called out in the figure. Rear components used on the 4U server are shown in Figure 3-8 on page 37.

For more information on using the 4U service node, see the *SGI Altix UV 10 System User's Guide*, (P/N 007-5645-00x).

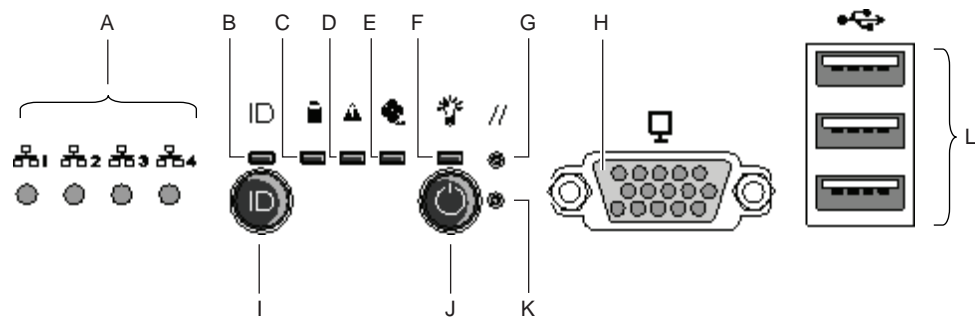


Figure 5-9 4U Service Node Front Controls and Interfaces

Table 5-2 4U Service Node Front Control and Interface Descriptions

Callout	Item function or description
A	Local area network (LAN) status LEDs (1 through 4)
B	System ID LED (blue)
C	Hard drive status LED (green)
D	System status/fault LED (green/amber)
E	Fan fault LED (amber)
F	System power LED (green) shows system power status
G	System reset button
H	VGA video connector
I	System ID button (toggles the blue identification LED - callout B)
J	System power button
K	Non-maskable interrupt (NMI) button - asserts NMI
L	USB 2.0 connector ports

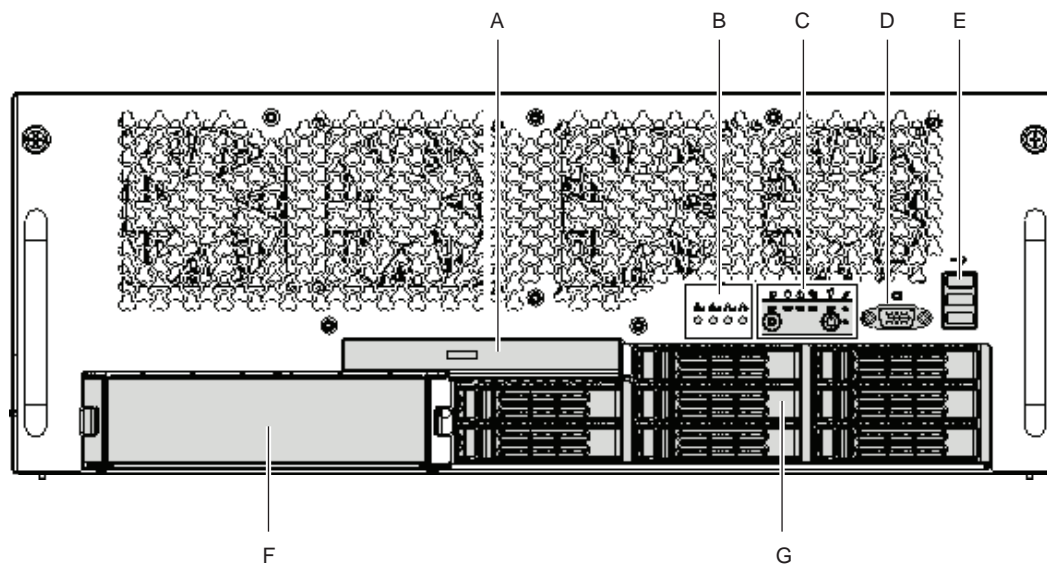


Figure 5-10 4U Service Node Front Panel

Table 5-3 4U Service Node Front Panel Item Identification

Front panel item	Functional description
A	Optional optical drive bay
B	Rear LAN LEDs
C	System control panel
D	Video connector
E	USB 2.0 connectors
F	5.25-inch peripheral bay
G	Hard drive bays

Basic Troubleshooting

This chapter provides the following sections to help you troubleshoot your system:

- “Troubleshooting Chart” on page 62
- “LED Status Indicators” on page 63

Troubleshooting Chart

Table 6-1 lists recommended actions for problems that can occur. To solve problems that are not listed in this table, contact your SGI system support engineer (SSE).

Table 6-1 Troubleshooting Chart

Problem Description	Recommended Action
The system will not power on.	<p>Ensure that the power cords of the IRU are seated properly in the power receptacles.</p> <p>Ensure that the PDU circuit breakers are on and properly connected to the wall source.</p> <p>If the power cord is plugged in and the circuit breaker is on, contact your SSE.</p>
An individual IRU will not power on.	<p>Ensure the power cables of the IRU are plugged in.</p> <p>View the CMC output from your system administration controller console.</p> <p>If the CMC is not running, contact your SSE.</p>
The system will not boot the operating system.	Contact your SSE.
The Service Required LED illuminates on an IRU.	View the CMC display of the failing IRU; contact your administrator or SSE for help as needed.
The PWR LED of a populated PCI slot in a support server is not illuminated.	Reseat the PCI card.
The Fault LED of a populated PCI slot in a support server is illuminated (on).	Reseat the PCI card. If the fault LED remains on, replace the PCI card.
The amber LED of a disk drive is on.	Replace the disk drive.

LED Status Indicators

There are a number of LEDs on the front of the IRUs that can help you detect, identify and potentially correct functional interruptions in the system.

The following subsections describe these LEDs and ways to use them to understand potential problem areas.

IRU Power Supply LEDs

Each power supply installed in an IRU has a single bi-color (green/amber) status LED.

The LED will either light green or amber (yellow), or flash green or yellow to indicate the status of the individual supply. See Table 6-2 for a complete list.

Table 6-2 Power Supply LED States

Power supply status	Green LED	Amber LED
No AC power to the supply	Off	Off
Power supply has failed	Off	On
Power supply problem warning	Off	Blinking
AC available to supply (standby) but IRU is off	Blinking	Off
Power supply on (IRU on)	On	Off

Compute/Memory Blade LEDs

Each compute/memory blade installed in an IRU has a total of eleven LED indicators arranged in a single row behind the perforated sheetmetal of the blade. The LEDs are located in the front lower section of the compute blade and are visible through the screen of the compute blade, see Figure 6-1 for an example. The functions of the LED status lights are as follows:

1. UID - Unit identifier - this blue LED is used during troubleshooting to find a specific compute node. The LED can be lit via software to aid in locating a specific compute node.
2. CPU Power OK - this green LED lights when the correct power levels are present on the processor(s).
3. IB0 link - green LED lights when a link is established on the internal InfiniBand 0 port
4. IB0 active - this amber LED flashes when IB0 is active (transmitting data)
5. IB1 link - green LED lights when a link is established on the internal InfiniBand 1 port
6. IB1 active - this amber LED flashes when IB1 is active (transmitting data)
7. Eth1 link - this green LED is illuminated when a link as been established on the system control Eth1 port
8. Eth1 active - this amber LED flashes when Eth1 is active (transmitting data)
9. Eth2 link - this LED indicates the compute blade's BMC Ethernet interface link status
10. Eth2 active - this LED indicates the compute blade's BMC Ethernet activity status
11. BMC heartbeat - this green LED flashes when the blade's BMC boots and is running normally. No illumination, or an LED that stays on solidly indicates the BMC failed.

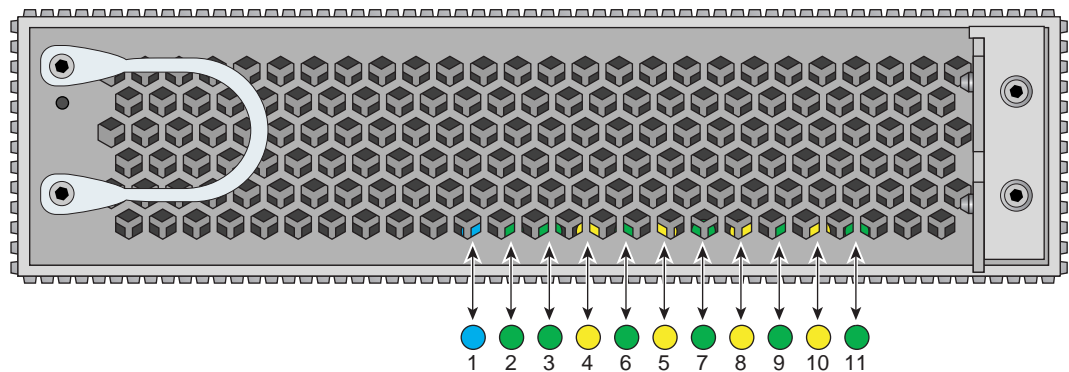


Figure 6-1 Compute Blade Status LED Locations

Chassis Management Panel “Service Required” Notices

Environmental “out-of-bounds” and chassis hardware failure conditions are reported on the chassis management panel. For individual rack units that experience a chassis-related component failure, a message appears on the CMC interface panel. This message is accompanied by the lighting of the amber “Service Required” LED on the panel’s front face (second from left). In the example shown in Figure 6-2, IRU 0 in rack 1 has experienced a fan failure.

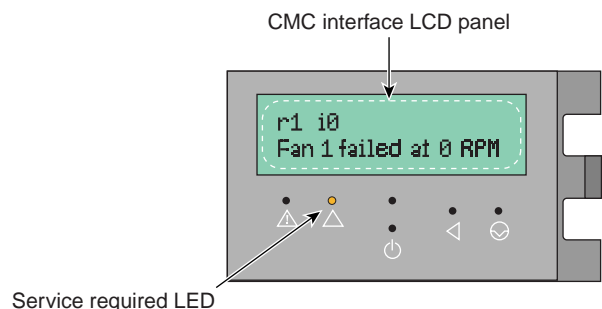


Figure 6-2 Fan Service Required Example Message on Chassis Management Panel

This type of information can be useful in helping your administrator or service provider identify and quickly correct hardware problems.

Technical Specifications and Pinouts

This appendix contains technical specification information about your system, as follows:

- “System-level Specifications” on page 67
- “Physical and Power Specifications” on page 68
- “Environmental Specifications” on page 69
- “I/O Port Specifications” on page 70

System-level Specifications

Table A-1 summarizes the Altix ICE 8400 series configuration ranges. Note that each processor socket normally houses six processor “cores” though a limited number of system configurations may be available with optional quad-core processors.

Table A-1 Altix ICE 8400 Series Configuration Ranges

Category	Minimum	Maximum
Blades	2 blades ^a	64 blades per rack
Individual Rack Units (IRUs)	1 per rack	4 per rack
Compute/memory blade DIMM capacity	1 DIMM per blade	16 DIMMs per blade
Chassis management blades	One per IRU	4 per rack
InfiniBand switch blades	1 per IRU	4 per IRU (16 per rack)

a. Compute blades support one or two stuffed sockets each. This is a total of six or 12 cores per blade maximum.

Physical and Power Specifications

Table A-2 shows the physical specifications of the Altix ICE system.

Table A-2 Altix ICE 8400 System Physical Specifications

System Features (single rack)	Specification
Height	79.5 in. (201.9 cm)
Width	31.3 in. (79.5 cm)
Depth	45.8 in. (116.3 cm)
Weight (full) maximum	~2,490 lbs. (1,132 kg) approximate
Shipping weight maximum	~2,535 lbs. (1,152 kg) approximate
Voltage range	North America/International
Nominal	200-240 VAC /230 VAC
Tolerance range	180-264 VAC /180-254 VAC
Frequency	North America/International
Nominal	60 Hz /50 Hz
Tolerance range	47-63 Hz /47-63 Hz
Phase required	Single-phase or 3-phase
Power requirements (max)	34.58 kVA (33.89 kW)
Hold time	16 ms
Power cable	12 ft. (3.66 m) pluggable cords
Shipping dimensions	Height: 81.375 in. (206.7 cm) Width: 48 in. (121.9 cm) Depth: 54 in. (137.1 cm)
Access requirements	
Front	48 in. (121.9 cm)
Rear	48 in. (121.9 cm)
Side	None

Environmental Specifications

Table A-3 lists the standard environmental specifications of the system.

Important: If you have ordered an Altix ICE 8400 HPC system using 130W processors, the proper operating ambient temperature tolerances are adjusted as follows:

+5 °C (41 °F) to +28 °C (82 °F) (up to 1500 m / 5000 ft.)

+5 °C (41 °F) to +23 °C (73 °F) (1500 m to 3000 m /5000 ft. to 10,000 ft.)

Table A-3 Environmental Specifications

Feature	Specification
Temperature tolerance (operating)	+5 °C (41 °F) to +35 °C (95 °F) (up to 1500 m / 5000 ft.) +5 °C (41 °F) to +30 °C (86 °F) (1500 m to 3000 m /5000 ft. to 10,000 ft.)
Temperature tolerance (non-operating)	-40 °C (-40 °F) to +60 °C (140 °F)
Relative humidity	10% to 80% operating (no condensation) 8% to 95% non-operating (no condensation)
Heat dissipation	
Altix ICE (rack)	Approximately 115.63 kBTU/hr maximum (based on 33.89 kW)
Cooling requirement	Ambient air or optional water cooling
Air flow: intake (front), exhaust (rear)	Approximately 2,800 CFM (normal operation)
Maximum altitude	10,000 ft. (3,049 m) operating 40,000 ft. (12,195 m) non-operating
Acoustical noise level (sound power)	8.7 bels (maximum at front of system)

I/O Port Specifications

This section contains specifications and port pinout information for the base I/O ports of your system, as follows:

- “Ethernet Port” on page 71
- “Serial Ports” on page 72

Ethernet Port

The system auto-selects the Ethernet port speed and type (duplex vs. half-duplex) when the server is booted, based on what it is connected to. Figure A-1 shows the Ethernet port.

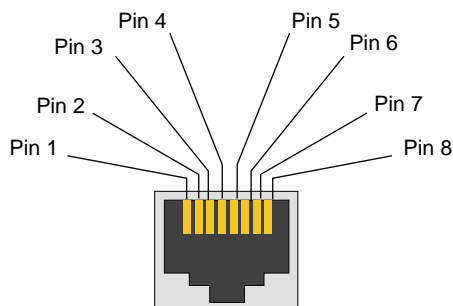


Figure A-1 Ethernet Port

Table A-4 shows the cable pinout assignments for the Ethernet port operating in 10/100-Base-T mode and also operating in 1000Base-T mode.

Table A-4 Ethernet Pinouts

Ethernet 10/100Base-T Pinouts		Gigabit Ethernet Pinouts	
Pins	Assignment	Pins	Assignment
1	Transmit +	1	Transmit/Receive 0 +
2	Transmit -	2	Transmit/Receive 0 -
3	Receive +	3	Transmit/Receive 1 +
4	NU	4	Transmit/Receive 2 +
5	NU	5	Transmit/Receive 2 -
6	Receive -	6	Transmit/Receive 1 -
7	NU	7	Transmit/Receive 3 +
8	NU	8	Transmit/Receive 3 -

NU = Not used

Serial Ports

The IRU's chassis management control boards have 9-pin serial interface connectors. These ports are capable of transferring data at rates as high as 230 kbps. Other features of the ports include the following:

- Programmable data, parity, and stop bits
- Programmable baud rate and modem control

Figure A-2 shows an example serial port.

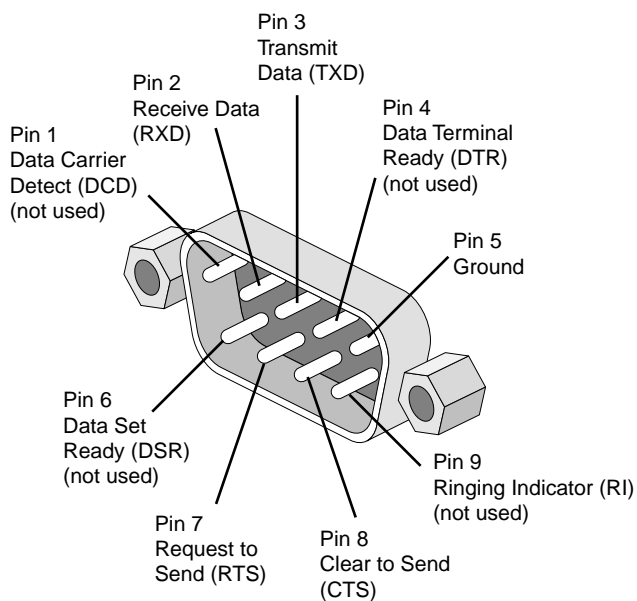


Figure A-2 Serial Port Connector

Table A-5 shows pinout assignments for the 9-pin male DB-9 connector.

Table A-5 Serial Port Pinout

Pin	Assignment	Description
1	DCD	Data carrier detect
2	RXD	Receive data
3	TXD	Transmit data
4	DTR	Data terminal ready
5	GND	Signal ground
6	DSR	Data set ready
7	RTS	Request to send
8	CTS	Clear to send
9	RI	Ring indicator

Safety Information and Regulatory Specifications

This appendix provides safety information and regulatory specifications for your system in the following sections:

- “Safety Information” on page 75
- “Regulatory Specifications” on page 77

Safety Information

Read and follow these instructions carefully:

1. Follow all warnings and instructions marked on the product and noted in the documentation included with this product.
2. Unplug this product before cleaning. Do not use liquid cleaners or aerosol cleaners. Use a damp cloth for cleaning.
3. Do not use this product near water.
4. Do not place this product or components of this product on an unstable cart, stand, or table. The product may fall, causing serious damage to the product.
5. Slots and openings in the system are provided for ventilation. To ensure reliable operation of the product and to protect it from overheating, these openings must not be blocked or covered. This product should never be placed near or over a radiator or heat register, or in a built-in installation, unless proper ventilation is provided.
6. This product should be operated from the type of power indicated on the marking label. If you are not sure of the type of power available, consult your dealer or local power company.
7. Do not allow anything to rest on the power cord. Do not locate this product where people will walk on the cord.
8. Never push objects of any kind into this product through cabinet slots as they may touch dangerous voltage points or short out parts that could result in a fire or electric shock. Never spill liquid of any kind on the product.

9. Do not attempt to service this product yourself except as noted in this guide. Opening or removing covers of node and switch internal components may expose you to dangerous voltage points or other risks. Refer all servicing to qualified service personnel.
10. Unplug this product from the wall outlet and refer servicing to qualified service personnel under the following conditions:
 - When the power cord or plug is damaged or frayed.
 - If liquid has been spilled into the product.
 - If the product has been exposed to rain or water.
 - If the product does not operate normally when the operating instructions are followed. Adjust only those controls that are covered by the operating instructions since improper adjustment of other controls may result in damage and will often require extensive work by a qualified technician to restore the product to normal condition.
 - If the product has been dropped or the cabinet has been damaged.
 - If the product exhibits a distinct change in performance, indicating a need for service.
11. If a lithium battery is a soldered part, only qualified SGI service personnel should replace this lithium battery. For other types, replace it only with the same type or an equivalent type recommended by the battery manufacturer, or the battery could explode. Discard used batteries according to the manufacturer's instructions.
12. Use only the proper type of power supply cord set (provided with the system) for this unit.
13. Do not attempt to move the system alone. Moving a rack requires at least two people.
14. Keep all system cables neatly organized in the cable management system. Loose cables are a tripping hazard that cause injury or damage the system.

Regulatory Specifications

The following topics are covered in this section:

- “CMN Number” on page 77
- “CE Notice and Manufacturer’s Declaration of Conformity” on page 77
- “Electromagnetic Emissions” on page 78
- “Shielded Cables” on page 80
- “Electrostatic Discharge” on page 80
- “Laser Compliance Statements” on page 81
- “Lithium Battery Statements” on page 82

This SGI system conforms to several national and international specifications and European Directives listed on the “Manufacturer’s Declaration of Conformity.” The CE mark insignia displayed on each device is an indication of conformity to the European requirements.



Caution: This product has several governmental and third-party approvals, licenses, and permits. Do not modify this product in any way that is not expressly approved by SGI. If you do, you may lose these approvals and your governmental agency authority to operate this device.

CMN Number

The model number, or CMN number, for the system is on the system label, which is mounted inside the rear door on the base of the rack.

CE Notice and Manufacturer’s Declaration of Conformity

The “CE” symbol indicates compliance of the device to directives of the European Community. A “Declaration of Conformity” in accordance with the standards has been made and is available from SGI upon request.

Electromagnetic Emissions

This section provides the contents of electromagnetic emissions notices from various countries.

FCC Notice (USA Only)

This equipment complies with Part 15 of the FCC Rules. Operation is subject to the following two conditions:

- This device may not cause harmful interference.
- This device must accept any interference received, including interference that may cause undesired operation.

Note: This equipment has been tested and found to comply with the limits for a Class A digital device, pursuant to Part 15 of the FCC Rules. These limits are designed to provide reasonable protection against harmful interference when the equipment is operated in a commercial environment. This equipment generates, uses, and can radiate radio frequency energy and, if not installed and used in accordance with the instruction manual, may cause harmful interference to radio communications. Operation of this equipment in a residential area is likely to cause harmful interference, in which case you will be required to correct the interference at your own expense.

If this equipment does cause harmful interference to radio or television reception, which can be determined by turning the equipment off and on, you are encouraged to try to correct the interference by using one or more of the following methods:

- Reorient or relocate the receiving antenna.
- Increase the separation between the equipment and receiver.
- Connect the equipment to an outlet on a circuit different from that to which the receiver is connected.

Consult the dealer or an experienced radio/TV technician for help.



Caution: Changes or modifications to the equipment not expressly approved by the party responsible for compliance could void your authority to operate the equipment.

Industry Canada Notice (Canada Only)

This Class A digital apparatus meets all requirements of the Canadian Interference-Causing Equipment Regulations.

Cet appareil numérique n'émet pas de perturbations radioélectriques dépassant les normes applicables aux appareils numériques de Classe A prescrites dans le Règlement sur les interférences radioélectriques établi par le Ministère des Communications du Canada.

VCCI Notice (Japan Only)

この装置は、情報処理装置等電波障害自主規制協議会 (VCCI) の基準に基づくクラス A 情報技術装置です。この装置を家庭環境で使用すると電波妨害を引き起こすことがあります。この場合には使用者が適切な対策を講ずるよう要求されることがあります。

Figure B-1 VCCI Notice (Japan Only)

Chinese Class A Regulatory Notice

警告使用者：

這是甲類的資訊產品，在居住的環境中使用時，可能會造成射頻干擾，在這種情況下，使用者會被要求採取某些適當的對策。

Figure B-2 Chinese Class A Regulatory Notice

Korean Class A Regulatory Notice

이 기기는 업무용으로 전자파적합등록을 한 기기이오니 판매자 또는 사용자는 이 점을 주의하시기 바라며 만약 잘못 판매 또는 구입하였을 때에는 가정용으로 교환하시기 바랍니다.

Figure B-3 Korean Class A Regulatory Notice

Shielded Cables

This SGI system is FCC-compliant under test conditions that include the use of shielded cables between the system and its peripherals. Your system and any peripherals you purchase from SGI have shielded cables. Shielded cables reduce the possibility of interference with radio, television, and other devices. If you use any cables that are not from SGI, ensure that they are shielded. Telephone cables do not need to be shielded.

Optional monitor cables supplied with your system use additional filtering molded into the cable jacket to reduce radio frequency interference. Always use the cable supplied with your system. If your monitor cable becomes damaged, obtain a replacement cable from SGI.

Electrostatic Discharge

SGI designs and tests its products to be immune to the effects of electrostatic discharge (ESD). ESD is a source of electromagnetic interference and can cause problems ranging from data errors and lockups to permanent component damage.

It is important that you keep all the covers and doors, including the plastics, in place while you are operating the system. The shielded cables that came with the unit and its peripherals should be installed correctly, with all thumbscrews fastened securely.

An ESD wrist strap may be included with some products, such as memory or PCI upgrades. The wrist strap is used during the installation of these upgrades to prevent the flow of static electricity, and it should protect your system from ESD damage.

Laser Compliance Statements

The DVD-ROM drive in this computer is a Class 1 laser product. The DVD-ROM drive's classification label is located on the drive.



Warning: Avoid exposure to the invisible laser radiation beam when the device is open.



Warning: Attention: Radiation du faisceau laser invisible en cas d'ouverture. Eviter toute exposition aux rayons.



Warning: Vorsicht: Unsichtbare Laserstrahlung, Wenn Abdeckung geöffnet, nicht dem Strahl aussetzen.



Warning: Advertencia: Radiación láser invisible al ser abierto. Evite exponerse a los rayos.



Warning: Advarsel: Laserstråling ved åbning se ikke ind i strålen



Warning: Varo! Lavattaessa Olet Alttina Lasersäteilylle



Warning: Varning: Laserstrålning när denna del är öppnad ålå tuijota såteeseenstirra ej in i strålen.



Warning: Varning: Laserstrålning nar denna del år öppnadstirra ej in i strålen.



Warning: Advarsel: Laserstråling nar deksel åpnesstirr ikke inn i strålen.

Lithium Battery Statements



Warning: If a lithium battery is a soldered part, only qualified SGI service personnel should replace this lithium battery. For other types, replace the battery only with the same type or an equivalent type recommended by the battery manufacturer, or the battery could explode. Discard used batteries according to the manufacturer's instructions.



Warning: Advarsel!: Lithiumbatteri - Eksplosionsfare ved fejlagtig håndtering. Udskiftning må kun ske med batteri af samme fabrikat og type. Léver det brugte batteri tilbage til leverandøren.



Warning: Advarsel: Eksplosjonsfare ved feilaktig skifte av batteri. Benytt samme batteritype eller en tilsvarende type anbefalt av apparatfabrikanten. Brukte batterier kasseres i henhold til fabrikantens instruksjoner.



Warning: Varning: Explosionsfara vid felaktigt batteribyte. Använd samma batterityp eller en ekvivalent typ som rekommenderas av apparattillverkaren. Kassera använt batteri enligt fabrikantens instruktion.



Warning: Varoitus: Päristö voi räjähtää, jos se on virheellisesti asennettu. Vaihda paristo ainoastaan laitevalmistajan suosittelemaan tyyppiin. Hävitä käytetty paristo valmistajan ohjeiden mukaisesti.



Warning: Vorsicht!: Explosionsgefahr bei unsachgemäßen Austausch der Batterie. Ersatz nur durch denselben oder einen vom Hersteller empfohlenem ähnlichen Typ. Entsorgung gebrauchter Batterien nach Angaben des Herstellers.

Index

A

Altix ICE servers
 monitoring locations, 11
Altix ICE single-rack server
 illustration, 20

B

battery statements, 82
block diagram
 system, 29

C

chassis management controller
 front panel display, 17
CMC controller
 functions, 17
CMN number, 77
Compute/Memory Blade LEDs, 64
customer service, xvii

D

documentation
 available via the World Wide Web, xvi
 conventions, xvii

E

environmental specifications, 69

F

front panel display
 L1 controller, 17

L

laser compliance statements, 81
LED Status Indicators, 63
LEDs on the front of the IRUs, 63
lithium battery warning statements, 2, 82

M

Message Passing Interface, 19
monitoring
 server, 11

N

numbering
 IRUs in a rack, 42
 racks, 43

O

optional water chilled rack cooling, 21

P

physical specifications

Altix ICE 8400 System Physical Specifications, 68

pinouts

Ethernet connector, 71

serial connector, 72

Power Supply LEDs, 63

powering on

preparation, 5

product support, xvii

R

RAS features, 40

S

server

monitoring locations, 11

system architecture, 23, 25

system block diagram, 29

system components

Altix ICE IRU front, 42

list of, 41

system features, 32

system overview, 19

T

tall rack

features, 46

technical specifications

system level, 67

technical support, xvii

three-phase PDU, 21

troubleshooting

problems and recommended actions, 62

Troubleshooting Chart, 62