SGI® ICE™ X System Hardware User Guide

Document Number 007-5806-001

Record of Revision

Version	Description
-001	March, 2012 First release

Contents

List of Figures

List of Tables

About This Guide

This guide provides an overview of the architecture, general operation and descriptions of the major components that compose the SGI® Integrated Compute Environment (ICE™) X series blade enclosure systems. It also provides the standard procedures for powering on and powering off the system, basic troubleshooting information, customer maintenance procedures and important safety and regulatory specifications.

Audience

This guide is written for owners, system administrators, and users of SGI ICE X 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 ICE X 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.
- Removing and replacing power supplies and fans as detailed in this document.
- Adding and replacing disk drives in optional storage systems and using the operator's 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, “SGI ICE X 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.
- Chapter 7, “Maintenance Procedures,” covers end-user service procedures that do not require special skills or tools to perform. Procedures not covered in this chapter should be referred to SGI customer support specialists or in-house trained service personnel.
- 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 ICE X series of computer systems:

- *SuperServer 6017R-N3RF4+ User's Manual,* (P/N 007-5849-00*x*)

This guide discuses 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 is also used as the default administrative server on the ICE X system. It may also be ordered configured as a login, or batch server, or other type of support server used with the ICE X series of computer systems.

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

This guide covers general operation, configuration, and servicing of the 2U Rackable C2108-TY10 server node(s) used in the SGI ICE X system. The C2108-TY10 can be used as a service node for login, batch, or other service node purposes.

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

This user's guide covers general operation, configuration, and servicing of the optional 3U-high C3108-TY11 service node(s) used in the SGI ICE X 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 X system. The server may also be used as an I/O gateway, or a mass storage resource.

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

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 ICE X 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 Management Center for SGI ICE X*, (P/N 007-5787-00*x*)

This guide discuses system configuration and software administration operations used with the SGI ICE X series. At time of publication, this document is intended for people who manage the operation of ICE X 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.
variable	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

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

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 85.

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.

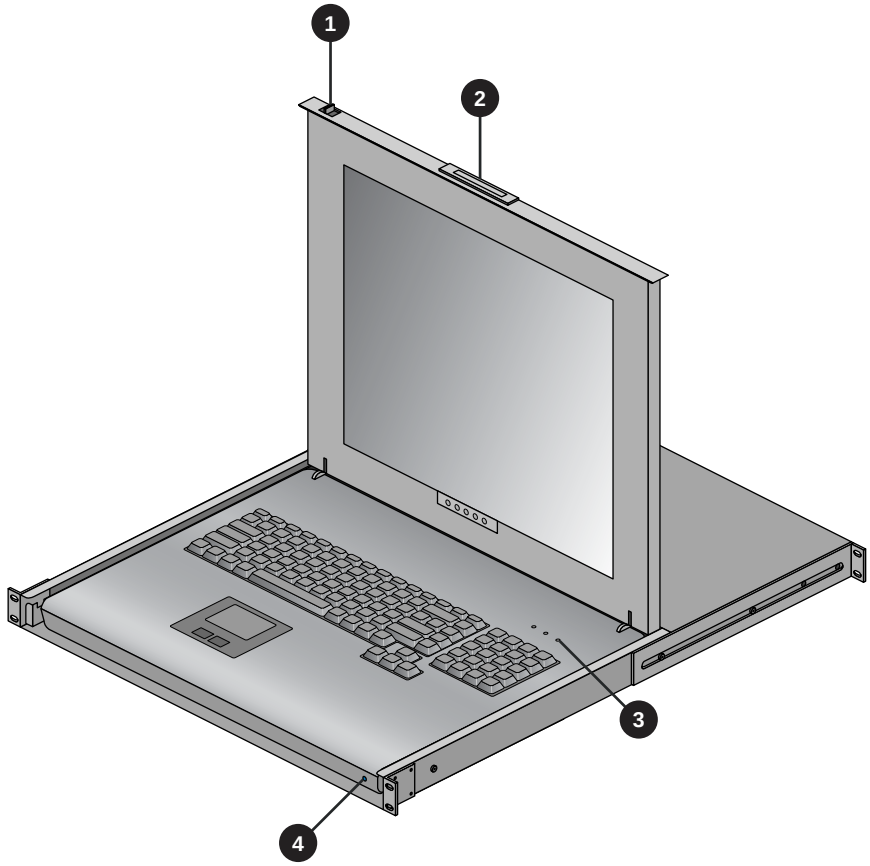

Figure 1-1 Flat Panel Rackmount Console Option

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.

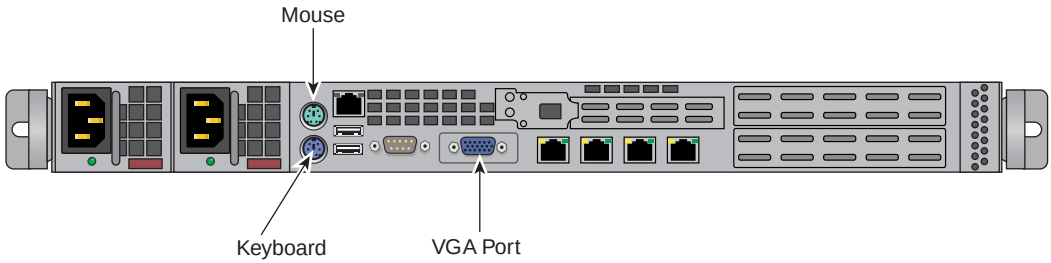

Figure 1-2 Administrative Controller Video Console Connection Points

Powering the System On and Off

This section explains how to power on and power off individual rack units, or your entire SGI ICE X 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 blade enclosures, 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 Management Center for SGI ICE X* (P/N 007-5787-00*x*).

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.
2. For each individual blade enclosure pair that you want to power on, make sure that the power cables are plugged into all the blade enclosure power supplies correctly, see the example in Figure 1-3. Setting the circuit breakers on the PDUs to the "On" position will apply power to the blade enclosure supplies and will start each of the chassis managers in each enclosure. Note that the chassis managers in each blade enclosure stay powered on as long as there is power coming into the unit. Turn off the PDU breaker switch that supplies voltage to the enclosure pair if you want to remove all power from the unit.

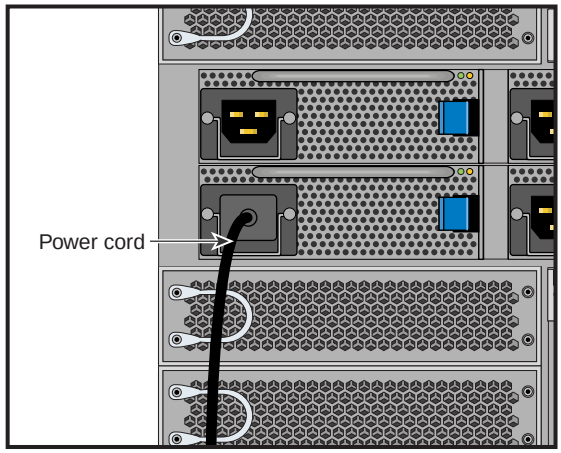

Figure 1-3 Blade Enclosure Power Supply Cable Example

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.

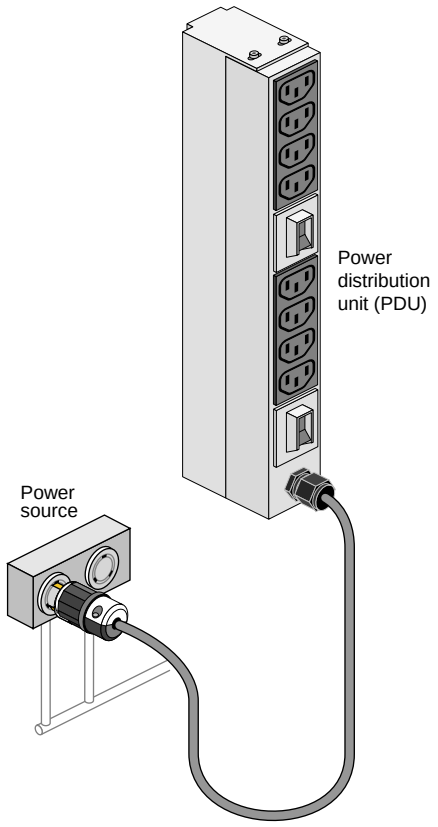

Figure 1-4 Eight-Outlet Single-Phase PDU Example

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

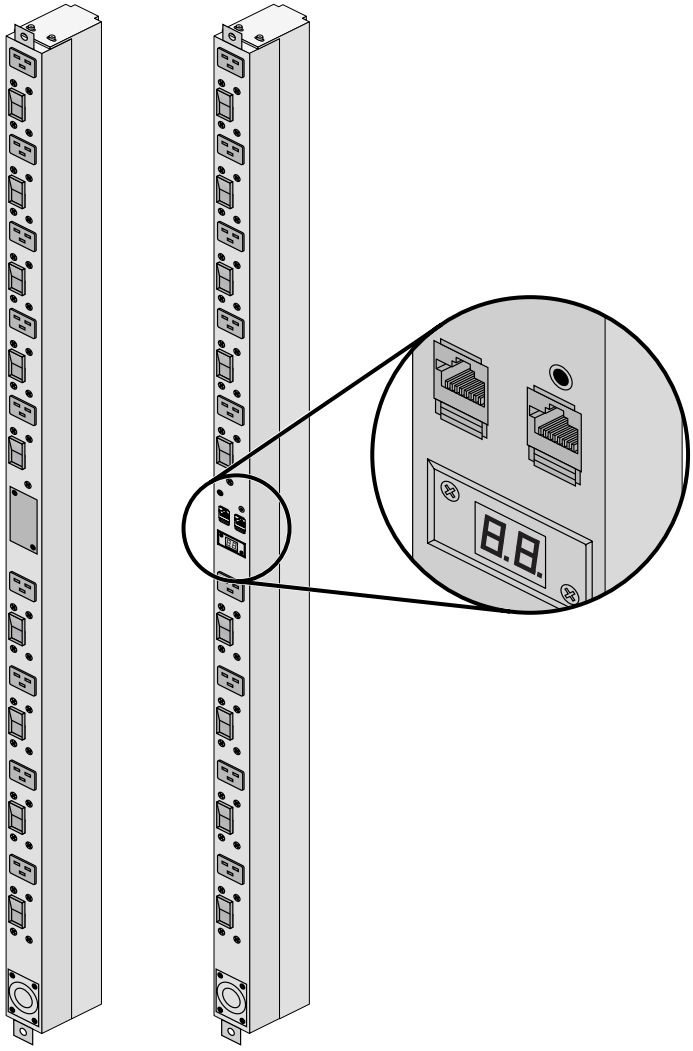

Figure 1-5 Three-Phase PDU Examples

Powering On and Off

The power-on and off procedure varies with your system setup. See the *SGI Management Center for SGI ICE X* (P/N 007-5787-00*x*) 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 ICE X 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	
`--noleader`	Do not include rack leader nodes. Valid with rack and system domains only.
`--noservice`	Do not include service nodes.
`--ipmi`	Uses ipmitool to communicate.
`--ssh`	Uses ssh to communicate.
`--intelplus`	Use the “-o intelplus option” for ipmitool [default].
`--verbose`	Print additional information on command progress.
`--noexec`	Display but do not execute commands that affect power.

Table 1-1 (continued) cpower option descriptions

Argument	Description
Target_type	
`--node`	Apply the action to a node or nodes. Nodes can be blade compute nodes (inside a blade enclosure), administration server nodes, rack leader controller nodes or service nodes.
`--IRU`	Apply the action at the blade enclosure 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`	Display 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 blade enclosure 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 the **`--noleader`** argument to the command string if you want the RLC to remain on.

See the *SGI Management Center for SGI ICE X* (P/N 007-5787-00*x*) 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 SGI ICE X server from the following sources:

- An optional flat panel rackmounted monitor with PS/2 keyboard/mouse can be connected to the administration server node for basic monitoring and administration of the SGI ICE X 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 Management Center for SGI ICE X*, (P/N 007-5787-00*x*) 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 SGI ICE X 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.

  Figure 1-6 on page 12 shows an example of the CMC board front panel locations in a blade enclosure. Note that a system using single-node ICE X blades will have one CMC board per blade enclosure (installed in the lower position in the enclosure). An ICE X system using dual-node blades must use two CMC boards.

  See Figure 2-4 on page 18 for an example illustration of the connectors and indicators used on the CMC board.

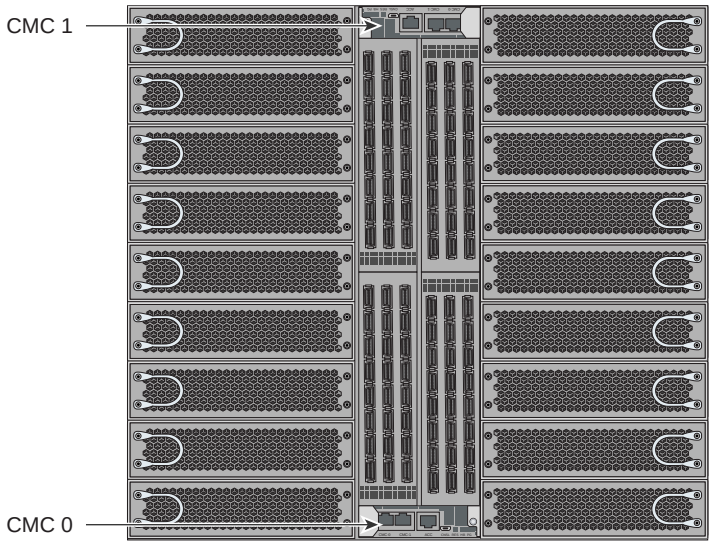

Figure 1-6 Blade Enclosure Chassis Management Board Locations

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

Note that each blade enclosure pair is configured with either two or four InfiniBand switch blades.

System Management

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

One or two chassis management controllers (CMCs) are used in each blade enclosure. A single CMC is used with single-node blades and two CMCs are needed when the enclosure uses dual-node blades. The first CMC is located directly below the enclosure's switch blade(s) and the other directly above. The chassis manager supports power-up and power-down of the blade enclosure's compute node blades and environmental monitoring of all units within the enclosure.

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

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

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

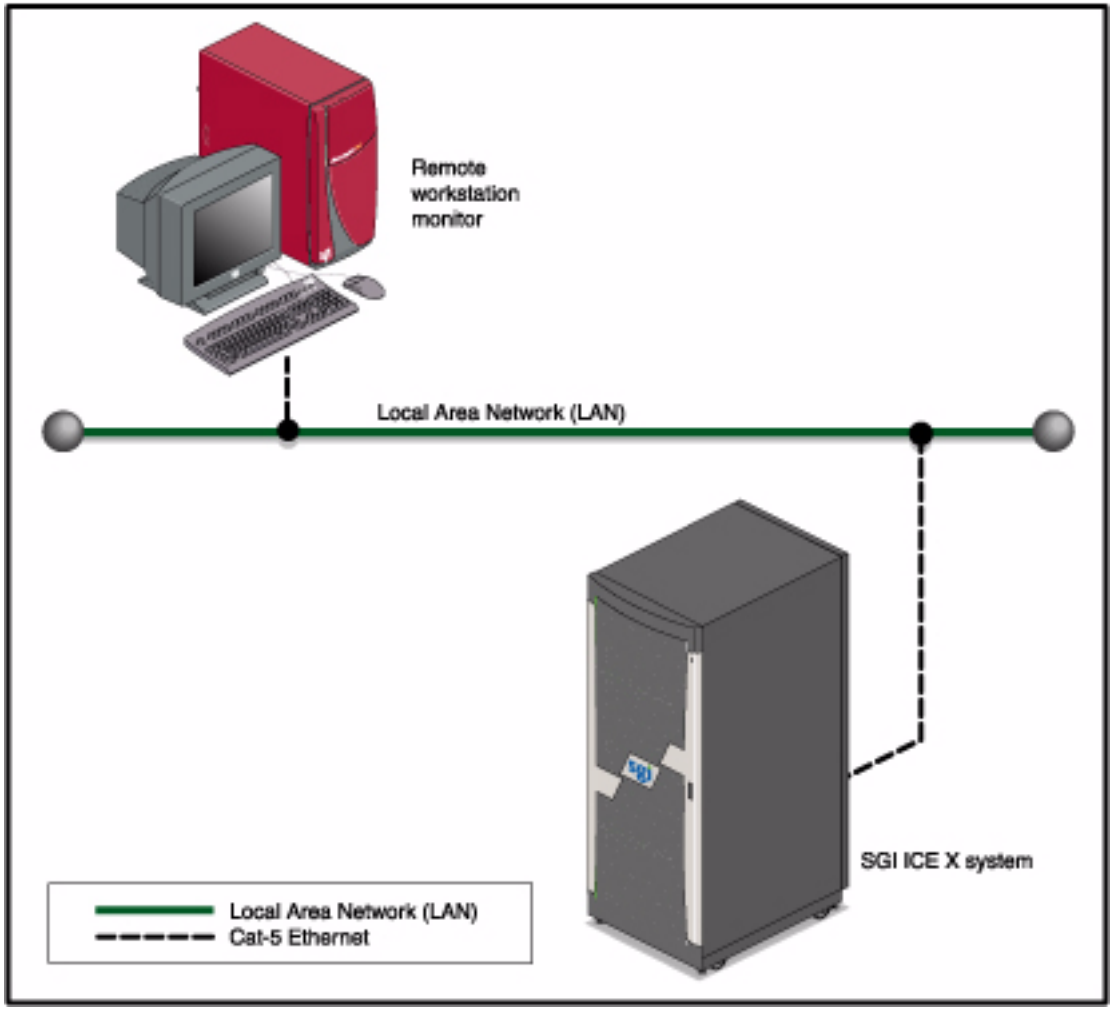

Figure 2-1 SGI ICE X System Network Access Example

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 SGI ICE X 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 X series machine will depend on the size of the system and the control options selected. Typically, any system with multiple blade enclosures will be interconnected by the chassis managers in each blade enclosure.

Note: Mass storage option enclosures are not monitored by the blade enclosure’s chassis manager. Most optional mass storage enclosures have their own internal microcontrollers for monitoring and controlling all elements of the disk array.

Chassis Controller Interaction

In all SGI ICE X series systems the system chassis management controllers communicate in the following ways:

- All blade enclosures within a system are polled for and provide information to the administrative node and RLC through their chassis management controllers (CMCs). Note that the CMCs are enlarged for clarity in Figure 2-3.
- The CMC does the environmental management for each blade enclosure, as well as power control, and provides an ethernet network infrastructure for the management of the system.

Chassis Manager Interconnects

The chassis managers in each blade enclosure connect to the system administration, rack leader and service node servers via gigabit Ethernet switches. See the redundant switch example in Figure 2-2 and the non-redundant example in Figure 2-3 on page 17.

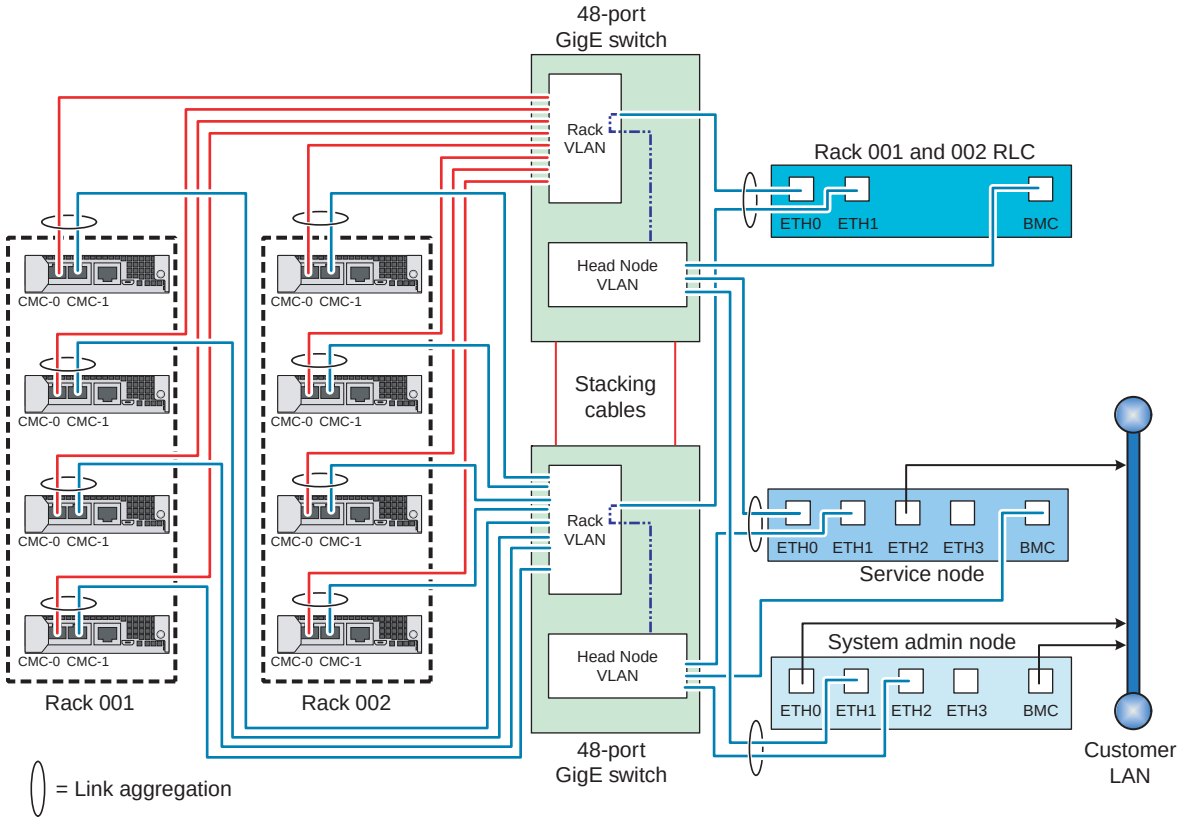

Figure 2-2 Redundant Chassis Manager Interconnect Diagram Example

Note that the non-redundant example (shown in Figure 2-3 on page 17) is a non-standard chassis management configuration with only a single virtual local area network (VLAN) connect line from each CMC to the internal LAN switch. See also “Multiple Chassis Manager Connections” in Chapter 3.

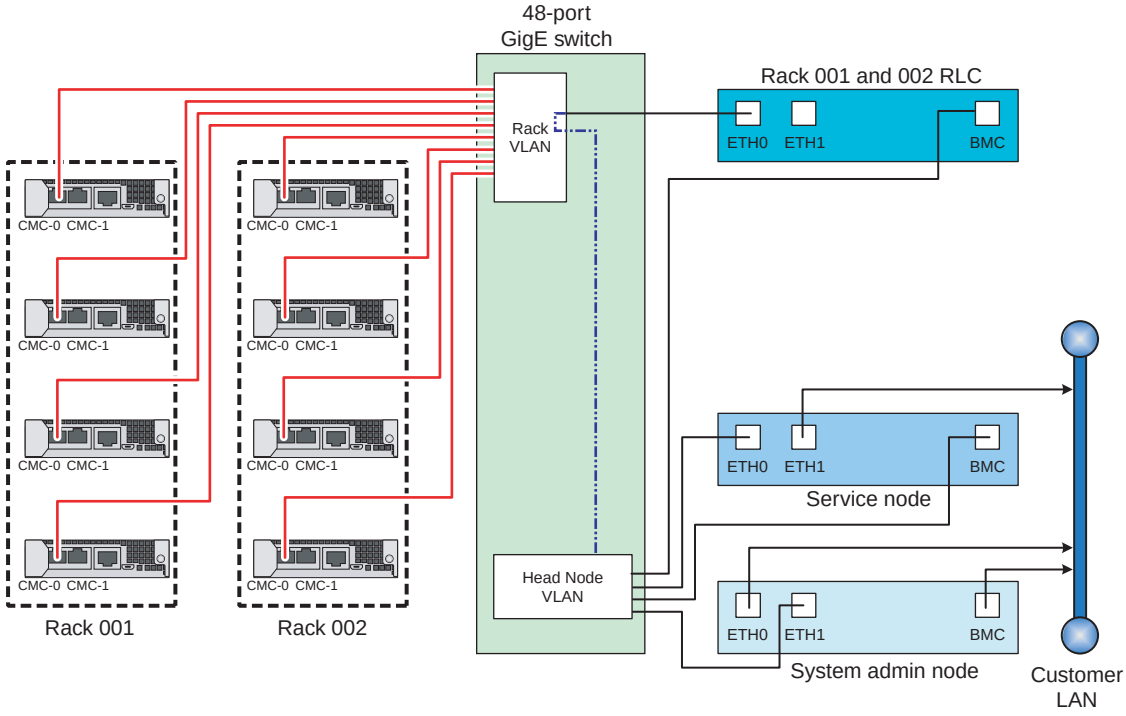

Figure 2-3 Non-redundant Chassis Manager Interconnection Diagram Example

Chassis Management Control (CMC) Functions

The following list summarizes the control and monitoring functions that the CMCs perform. Most functions are common across multiple blade enclosures:

- Controls and monitors blade enclosure 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

CMC Connector Ports and Indicators

The ports on the CMC board are used as follows:

- CMC-0 - Primary CMC connection, connects to the RLC via the 48-port management switch
- CMC-1 - Secondary CMC connection to the RLC via the 48-port management switch (used with redundant VLAN switch configurations)
- ACC - Accessory port, used as a direct connection to the microprocessor for service
- CNSL - Console connection - used for service troubleshooting
- RES - RESET switch, depress this switch to reset the CMC microprocessor
- HB - Heartbeat LED, green flashing LED indicates CMC is running
- PG - Power Good LED, this LED is illuminated green when power is present

Figure 2-4 shows the chassis management controller front panel in the blade enclosure.

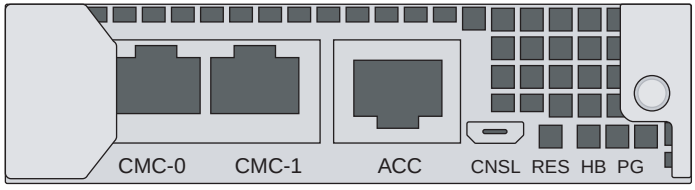

Figure 2-4 Chassis Management Controller Board Front Panel Ports and Indicators

System Power Status

The `cpower` command is the main interface for all power management commands. You can request power status and power-on or 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. Note that the term “IRU” represents a single blade enclosure within a blade enclosure pair.

The `cpower` commands may require several seconds to several minutes to complete, depending on how many blade enclosures 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 blade enclosure, 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 Management Center for SGI ICE X* (P/N 007-5787-00*x*) has more extensive information on these topics.

System Overview

This chapter provides an overview of the physical and architectural aspects of your SGI Integrated Compute Environment (ICE) X series system. The major components of the SGI ICE X 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 SGI ICE X series system in your lab or server room. Each 42U SGI rack holds one or two 21U-high (blade enclosure pairs). An enclosure pair is a sheetmetal assembly that consists of two 18-blade enclosures (upper and lower). The enclosures are separated by two power “shelves” that each hold three power supplies (shared by the blade enclosures). Each enclosure also has an internal InfiniBand communication backplane. The 18 blades supported in each enclosure 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 enclosures. Every compute blade contains four or eight dual-inline memory module (DIMM) memory units per processor socket. Optional hard disk or solid-state (SSD) drives may be available with specific blade configurations.

Each blade supports two processor sockets. Note that a maximum system size of 72 compute 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. Contact your SGI sales or service representative for the most current information on these topics.

The SGI ICE X series systems can run parallel programs using a message passing tool like the Message Passing Interface (MPI). The SGI ICE X blade system uses a distributed memory scheme as opposed to a shared memory system like that used in the SGI 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

Figure 3-1 shows an example configuration of a single-rack SGI ICE X server.

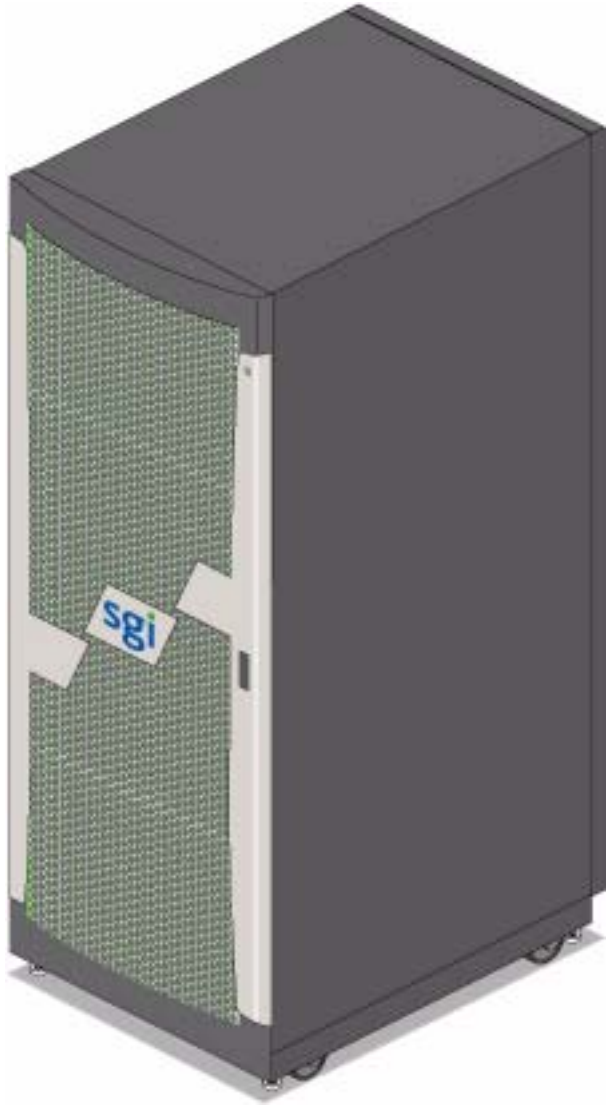

Figure 3-1 SGI ICE X Series System (Single Rack)

The 42U rack for this server houses all blade enclosures, option modules, and other components; up to 1152 processor cores in a single rack. The basic enclosure within the SGI ICE X system is the 21U-high (36.75 inch or 93.35 cm) blade enclosure pair. The enclosure pair supports a maximum of 36 compute blades, up to six power supplies, up to four chassis management controllers (CMCs) and two to four InfiniBand architecture I/O fabric switch interface blades. Note that two additional power supplies used in the enclosure pair are installed at the rear of the unit and dedicated to running the unit's cooling fans (blowers). 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 PDUs installed to support each blade enclosure pair and any support servers or storage units.

Figure 3-2 shows a blade enclosure pair and rack. The optional three-phase 208V PDU has nine outlets and two PDUs are installed in each SGI ICE X 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. An optional single-phase PDU has 8 outlets and can be used in an optional I/O support rack.

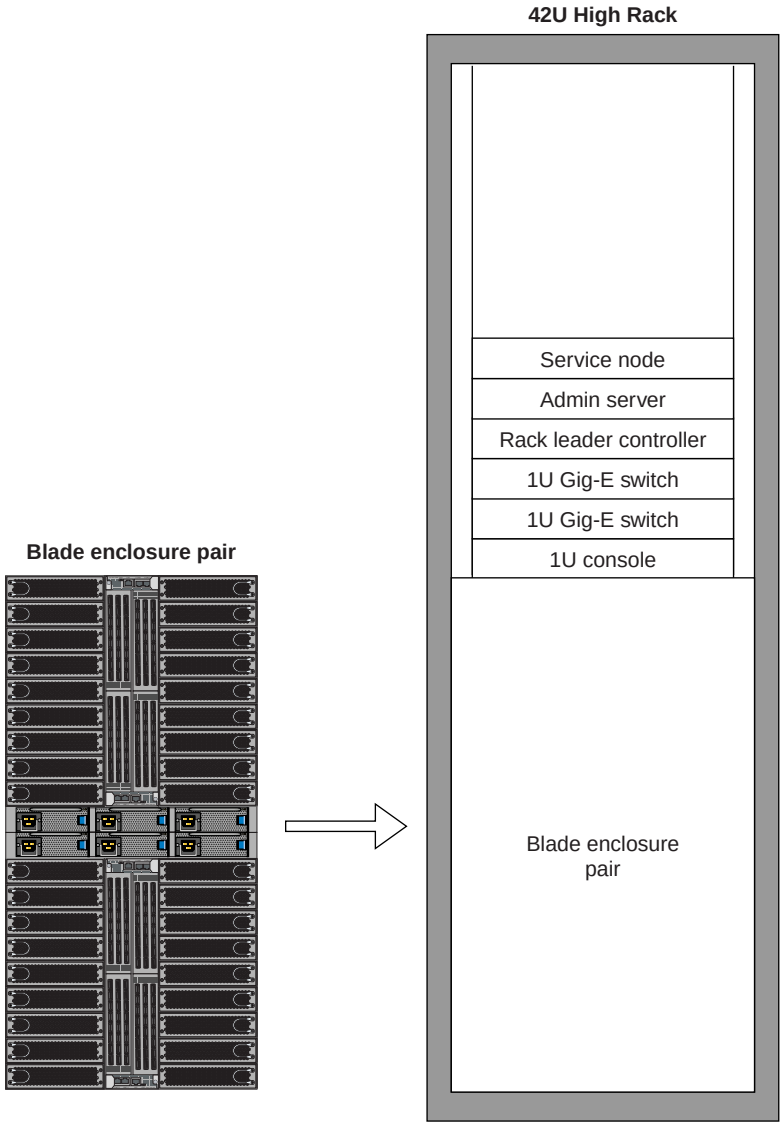

Figure 3-2 Blade Enclosure and Rack Components Example

Intel System and Blade Architectures

The SGI ICE X series of computer systems are based on an FDR 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 blades using a choice of one of three host-channel adapter (HCA) cards, see “IP113 Blade Architecture Overview”.

IP113 Blade Architecture Overview

An enhanced and updated four, six or eight-core version of the SGI ICE compute blade is used in the ICE X systems. The IP113 compute blade **cannot** be plugged into and cannot be used in “previous generation” SGI Altix ICE 8200 or 8400 series blade enclosures. Multi-generational system interconnects can be made through the InfiniBand fabric level. Check with your SGI service or sales representative for additional information on this topic. The IP113 blade architecture is described in the following sections.

The compute blade contains the processors, memory, and one of the following fourteen-data rate (FDR) InfiniBand imbedded HCA selections:

- One single-port IB HCA
- One dual-port IB HCA
- One HCA with two single-port IB connectors

Each compute blade is configured with two four-core, six-core or eight-core Intel processors - a maximum of 16 processor cores per compute blade. A maximum of 16 DDR3 memory DIMMs are supported per compute blade.

The two processors on the IP113 maintain an interactive communication link using the Intel QuickPath Interconnect (QPI) technology. This high-speed interconnect technology provides data transfers between the processors, memory and I/O hub components. Note that the IP113 blade can optionally support one or two native “on-board” hard disk or SSD drive options for local swap/scratch usage.

QuickPath Interconnect Features

Each processor on an Intel-based 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)

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.

Blade DIMM Memory Features

Note that each processor on an Intel blade uses four DDR3 memory channels with one or more memory DIMMs on each channel (depending on configuration selected). Each blade can support up to 16 DIMMs. The DDR3 memory channel supports a maximum memory bandwidth of up to 12.8 GBs per second. The combined maximum bandwidth for all memory channels on a single processor is 51.2 GBs per second.

Memory Channel Recommendation

It is highly recommended (though not required) that each processor on a system blade 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 the Intel 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, such as 1600MT/s or 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 1333 MT/s DIMM on a channel will only operate at 1333 MT/s - not 1600 MT/s.

Populating one 1600 MT/s DIMM on each channel of an Intel based blade delivers a maximum of 12.8 GB/s per channel or 51.2 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; four DIMMs per processor are recommended. An example blade enclosure with all blade slots filled is shown in Figure 3-9 on page 41. 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 enclosure use the same number and capacity (size) DIMMs.

Each blade in the enclosure pair may have a different total DIMM capacity. For example, one blade may have 16 DIMMs, and another may have only eight. Note that while this difference in capacity is acceptable functionally, it may have impact on compute "load balancing" within the system.

System InfiniBand Switch Blades

Two or four fourteen-data-rate (FDR) InfiniBand switch blades can be used with each blade enclosure pair configured in the SGI ICE X system. There are two switch blades in an enclosure pair for single-plane InfiniBand topologies. Enclosure pairs with four switch blades use a dual-plane topology that provides high-bandwidth communication between compute blades inside the enclosure as well as blades in other enclosures.

Enclosure Switch Density Choices

Each SGI ICE X system comes with a choice of two switch configurations.

- Single 36-port FDR IB ASIC (standard) with 18 ports external in each enclosure
- Dual 36-port FDR IB ASIC (premium) with a total of 48 external ports

The single-switch ASIC and dual-switch ASIC switch blades for each enclosure pair are **not** interchangeable without re-configuration of the system. The outward appearance of the two types is very similar, but differs in regards to the number and location of QSFP ports.

Enclosures using one or two FDR switch blades are available in certain specific configurations. A single-switch blade within a blade enclosure supports a single-plane FDR InfiniBand topology only; check with your SGI sales or service representative for additional information on availability.

The SGI ICE X FDR switch blade locations example is shown in Figure 3-3. 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.

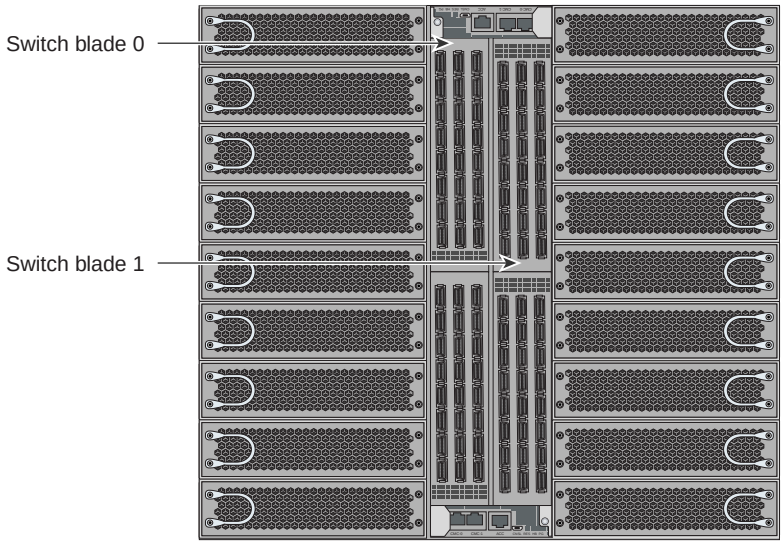

Figure 3-3 InfiniBand 48-port (Premium) FDR Switch Numbering in Blade Enclosures

System Features and Major Components

The main features of the SGI ICE X series server systems are introduced in the following sections:

- "Modularity and Scalability" on page 30
- "Reliability, Availability, and Serviceability (RAS)" on page 38

Modularity and Scalability

The SGI ICE X series systems are modular, blade-based, scaleable, high-density cluster systems. The system rack components are primarily housed in building blocks referred to as blade enclosure pairs. Each enclosure pair consists of a sheetmetal housing with internal IB backplanes and six (shared) power supplies that serve two "blade enclosures".

However, other "free-standing" SGI compute servers are used to administer, access and service the SGI ICE X series systems. Additional optional mass storage may be added to the system along with additional blade enclosures. 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 blade enclosures around processing capability, memory size or InfiniBand fabric I/O capability. The air-cooled blade enclosure enclosure has redundant, hot-swap fans and redundant, hot-swap power supplies. A water-chilled rack option expands an ICE X rack's heat dissipation capability for the blade enclosure 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 (also referred to as nodes) are used with SGI ICE X 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
  - MDS or OSS nodes (used in optional Lustre configurations)

Each SGI ICE X system will have one system administration controller, one rack leader controller (RLC) and at least one service node. All ICE X systems require one RLC for every eight CMCs in the system.

The administration 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 1U server system (node). The server is used to install SGI ICE X 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 X systems is connected to the external network and may be set up for interactive logins under specific circumstances. However, most ICE X 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 33 and the “I/O Gateway Node” on page 34.

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. One rack leader controller is required for every eight CMC boards used in a system and it is a non-blade “stand-alone” 1U server. The rack leader controllers are guided and monitored by the system administration server. Each RLC in turn monitors, pulls and stores data from the compute nodes of all the blade enclosures within the SSI. The rack leader then consolidates and forwards data requests received from the blade enclosure’s compute nodes to the administration server. A rack leader controller may also supply boot and root file sharing images to the compute nodes in the enclosures.

For large systems, multiple RLC servers may be used to distribute the job load. Note that a high-availability RLC configuration is available that doubles the number of RLCs used in a system. In high-availability (HA) RLC configurations two RLCs are paired together. The primary RLC is backed up by an identical “backup” RLC server. The second (backup) RLC runs the same fabric management image as the primary RLC. Check with your SGI sales or support

representative for configurations that use a “spare” RLC or administration server. This option can provide rapid “fail-over” replacement for a failed RLC or administrative unit.

Multiple Chassis Manager Connections

In multiple-rack configurations the chassis managers (up to eight CMCs) may be interconnected to the administrative server and the rack leader controller (RLC) server via one or two Ethernet switches. Figure 3-4 shows an example diagram of the CMC interconnects between two ICE X system racks using a virtual local area network (VLAN).

For more information on these and other topics related to the CMC, see the *SGI Management Center for SGI ICE X* (P/N 007-5787-00*x*).

Note also that the scale of the CMC drawings in Figure 3-4 is adjusted to clarify the interconnect locations.

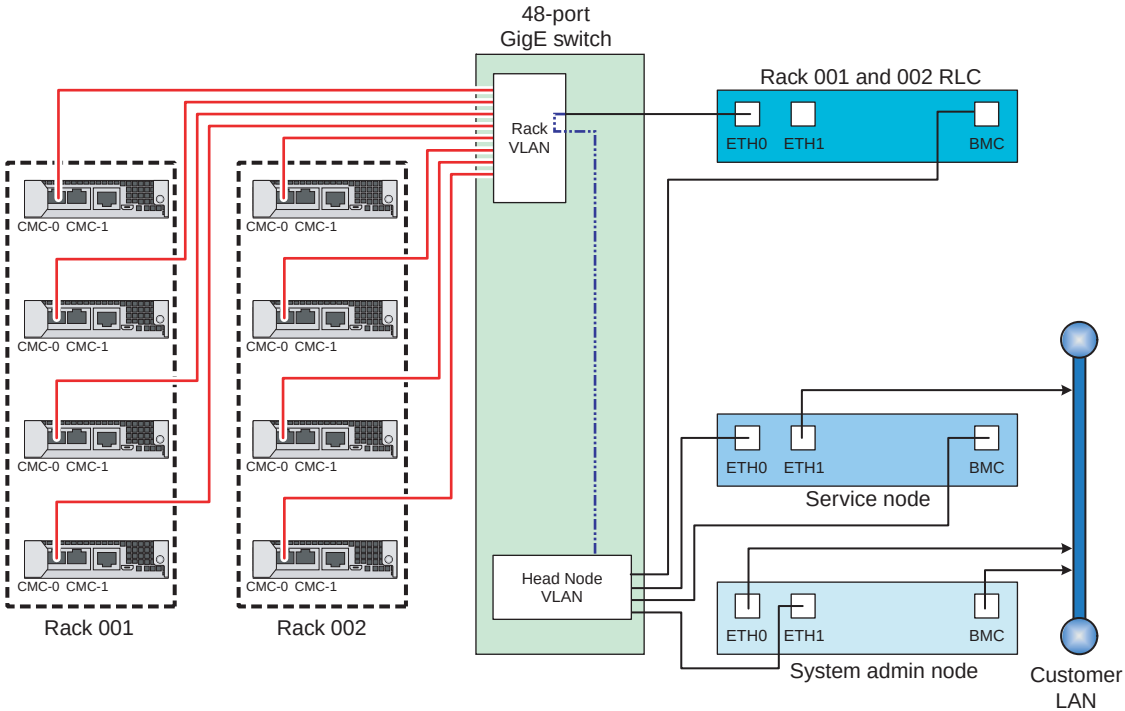

Figure 3-4 Administration and RLC Cabling to Chassis Managers Via Ethernet Switch

The RLC as Fabric Manager

In some SGI ICE X 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 the blade enclosure pair. See the “Rack Leader Controller” on page 31 subsection for more detail. The fabric management software runs on one or two 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.

Service Nodes

The functionality of the service “nodes” listed in this subsection are all services that can technically be shared on a single hardware server unit. System scale, configuration and number of users generally determines when you add more servers (nodes) and dedicate them to these service functions. However, you can also have a smaller system where several 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.

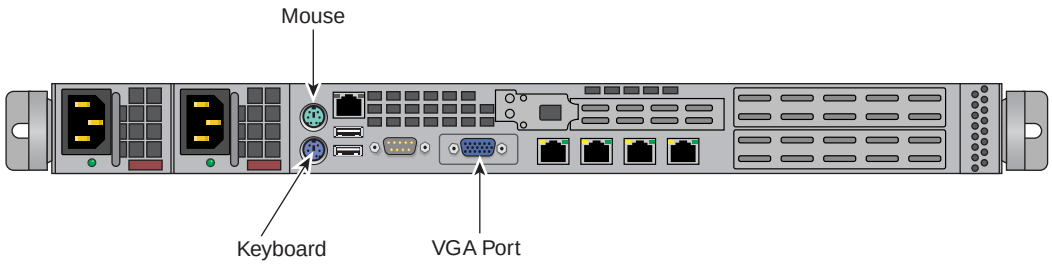

Figure 3-5 Example Rear View of a 1U Service Node

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.

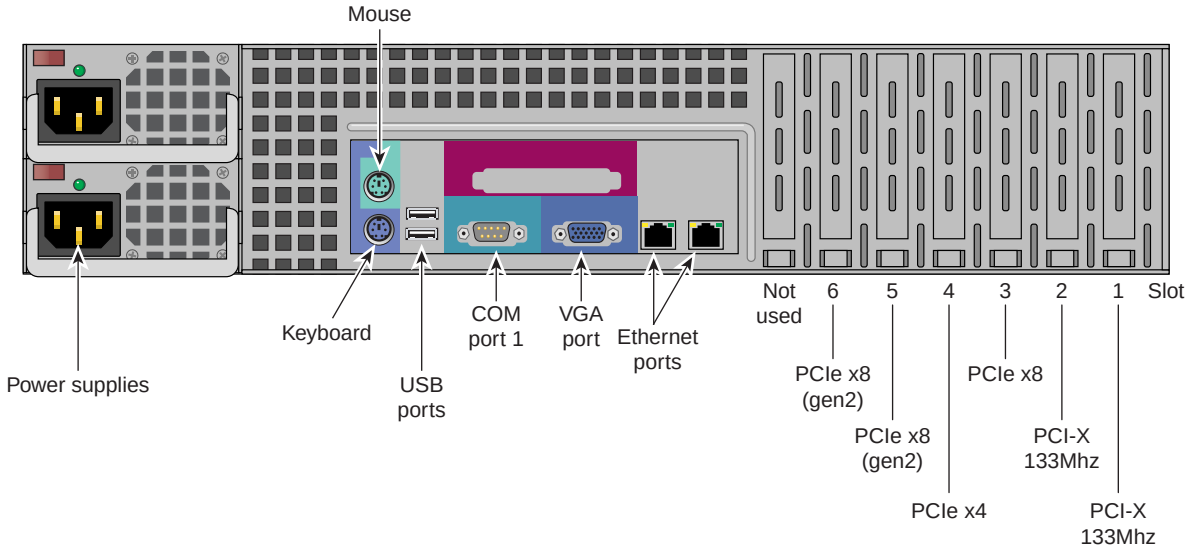

Figure 3-6 2U Service Node Rear Panel

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 on page 35 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.

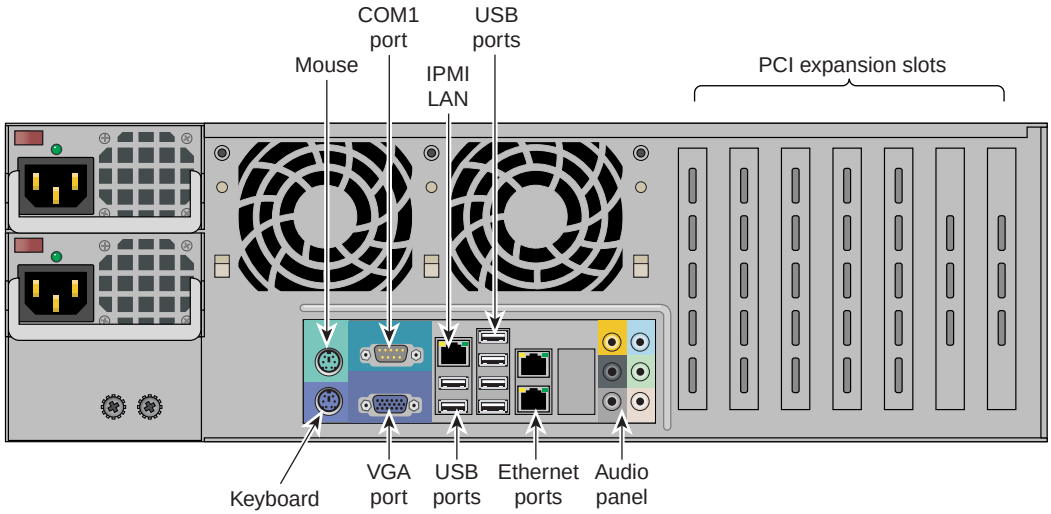

Figure 3-7 3U Service Node Rear Panel Example

The 4U Service Node

An optional 4U service node is offered with the SGI ICE X systems. This server is a higher-performance system that can contain multiple processors (up to 4) and serve multiple purposes within the SGI ICE X system. The 4U server is not used as an administrative node or rack leader controller. Figure 3-8 on page 36 shows the rear panel of the 4U service node and Table 3-1 identifies the functional items on the back of the unit. See the *SGI Altix UV 10 System User's Guide* (P/N 007-5645-00*x*) for details on operating the 4U server.

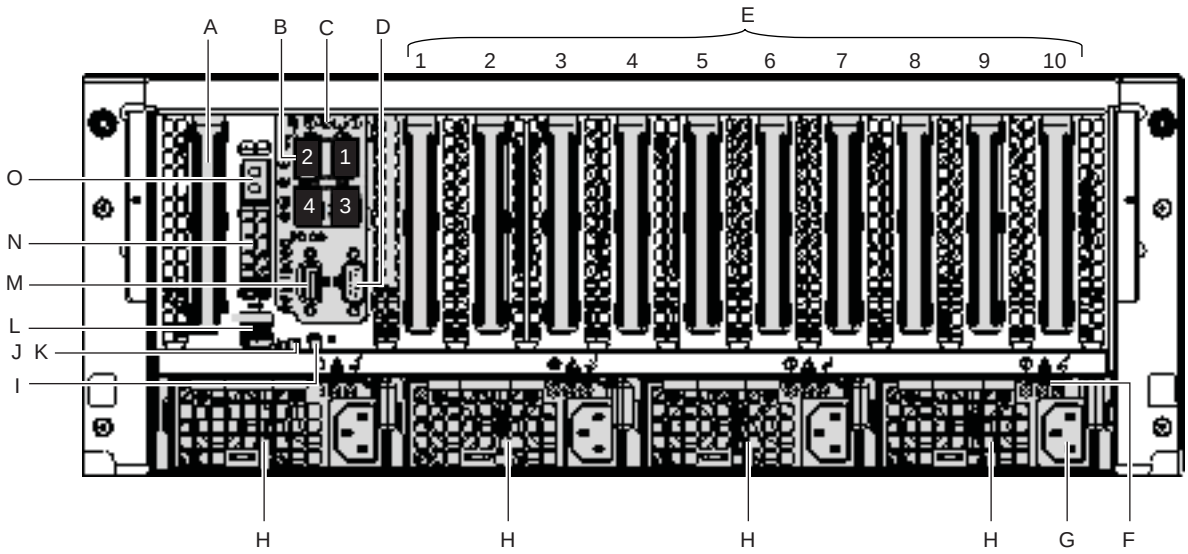

Figure 3-8 4U Service Node Rear Panel Example

Table 3-1 4U Service Node Rear Panel Items

Item	Description
A	SAS riser slot - PCIe Gen-2 x8 half-height slot
B	I/O riser Gigabit Ethernet ports
C	I/O riser module
D	Serial port connector
E	PCIe Gen-2 x8 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

Optional Lustre Nodes Overview

The nodes in the following subsections are used when the SGI ICE X system is set up as a Lustre file system configuration. In SGI ICE X installations the MDS and OSS functions are generally on separate nodes within the ICE X system and communicating over a network.

Lustre clients access and use the data stored in the OSS node's object storage targets (OSTs). Clients may be compute nodes within the SGI ICE X system or Login, Batch or other service nodes. Lustre presents all clients with a unified namespace for all of the files and data in the filesystem, using standard portable operating system interface (POSIX) semantics. This allows concurrent and coherent read and write access to the files in the OST filesystems. The Lustre MDS server (see "MDS Node") and OSS server (see "OSS Node"), will read, write and modify data in the format imposed by these file systems. When a client accesses a file, it completes a filename lookup on the MDS node. As a result, a file is created on behalf of the client or the layout of an existing file is returned to the client. For read or write operations, the client then interprets the layout in the logical object volume (LOV) layer, which maps the offset and size to one or more objects, each residing on a separate OST within the OSS node.

MDS Node

The metadata server (MDS node) uses a single metadata target (MDT) per Lustre filesystem. Two MDS nodes can be configured as an active-passive failover pair to provide redundancy. The metadata target stores namespace metadata, such as filenames, directories, access permissions and file layout. The MDT data is usually stored in a single localized disk filesystem. The storage used for the MDT (a function of the MDS node) and OST (located on the OSS node) backing filesystems is partitioned and optionally organized with logical volume management (LVM) and/or RAID. It is normally formatted as a fourth extended filesystem, (a journaling file system for Linux). When a client opens a file, the file-open operation transfers a set of object pointers and their layout from the MDS node to the client. This enables the client to directly interact with the

OSS node where the object is stored. The client can then perform I/O on the file without further communication with the MDS node.

OSS Node

The object storage server (OSS node) is one of the elements of a Lustre File Storage system. The OSS is managed by the SGI ICE X management network. The OSS stores file data on one or more object storage targets (OSTs). Depending on the server's hardware, an OSS node typically serves between two and eight OSTs, with each OST managing a single local disk filesystem. An OST is a dedicated filesystem that exports an interface to byte ranges of objects for read/write operations. The capacity of each OST on the OSS node can range from a maximum of 24 to 128 TB depending on the SGI ICE X operating system and the Lustre release level. The data storage capacity of a Lustre file system is the available storage total of the capacities provided by the OSTs.

Reliability, Availability, and Serviceability (RAS)

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

- **Power and cooling:**
  - Power supplies within the blade enclosure pair chassis are redundant and can be hot-swapped under most circumstances.
  - A rack-level water chilled cooling option is available for all configurations.
  - Blade enclosures have overcurrent protection at the blade and power supply level.
  - Fans (blowers) are redundant and can be hot-swapped.
  - Fans can run at multiple speeds. Speed increases automatically when temperature increases or when a single fan fails.
- **System monitoring:**
  - Chassis managers monitor blade enclosure internal voltage, power and temperature.
  - Redundant system management networking is available.
  - Each blade/node installed has status LEDs that can indicate a malfunctioning or failed part; LEDs are readable at the front of the system.
  - Systems support remote console and maintenance activities.

- **Error detection and correction**
  - External memory transfers are protected by cyclic redundancy check (CRC) error detection. If a memory packet does not checksum, it is retransmitted.
  - Nodes within each blade enclosure 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 SGI ICE X 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 SGI ICE X series. Up to two blade enclosure pairs can be installed in each rack. Note that multi-rack systems will often have a dedicated I/O rack holding GigE switches, RLCs, Admin servers and additional service nodes.
- **Blade enclosure pair.** This sheetmetal enclosure contains the two enclosures holding up to 36 compute blades, up to four chassis manager boards, up to four InfiniBand fabric I/O blades and six front-access power supplies for the SGI ICE X series computers. The enclosure pair is 21U high. Figure 3-9 on page 41 shows the SGI ICE X series blade enclosure pair system front components.
- **Fan (blower) enclosure.** This sheetmetal enclosure is installed back-to-back with each blade enclosure pair. The fan enclosure consists of two 6-blower enclosures and two dedicated power supplies. Figure 7-3 on page 73 shows an example of the enclosure.
- **Single-wide compute blade.** Holds two processor sockets and up to 16 memory DIMMs. See Figure 3-10 on page 42 for an example of blade number assignments.
- **1U RLC (rack leader controller).** One 1U rack leader server is required for each eight CMCs in a system. High-availability configurations using redundant RLCs are supported.
- **1U Administrative server with PCIe expansion.** This server node supports an optional console, administrative software and two PCIe option cards. The administrative server is generally installed in a dedicated I/O rack in any multi-rack ICE X system.
- **1U Service node.** Additional 1U server(s) can be added to a system rack and used specifically as an optional login, batch, MDS, OSS or other service node. Note that these service functions cannot be incorporated as part of the system RLC or administration server.
- **2U Service node.** An optional 2U service node may be used as a login, batch, MDS, OSS or fabric node. In smaller systems, multiple functions may be combined on one server.
- **3U Service node.** The optional 3U server node is offered with certain configurations needing higher performance I/O access for the SGI ICE X 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 SGI ICE X system. It offers the highest processing power, best I/O performance and most flexible configuration options of the available service nodes.

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

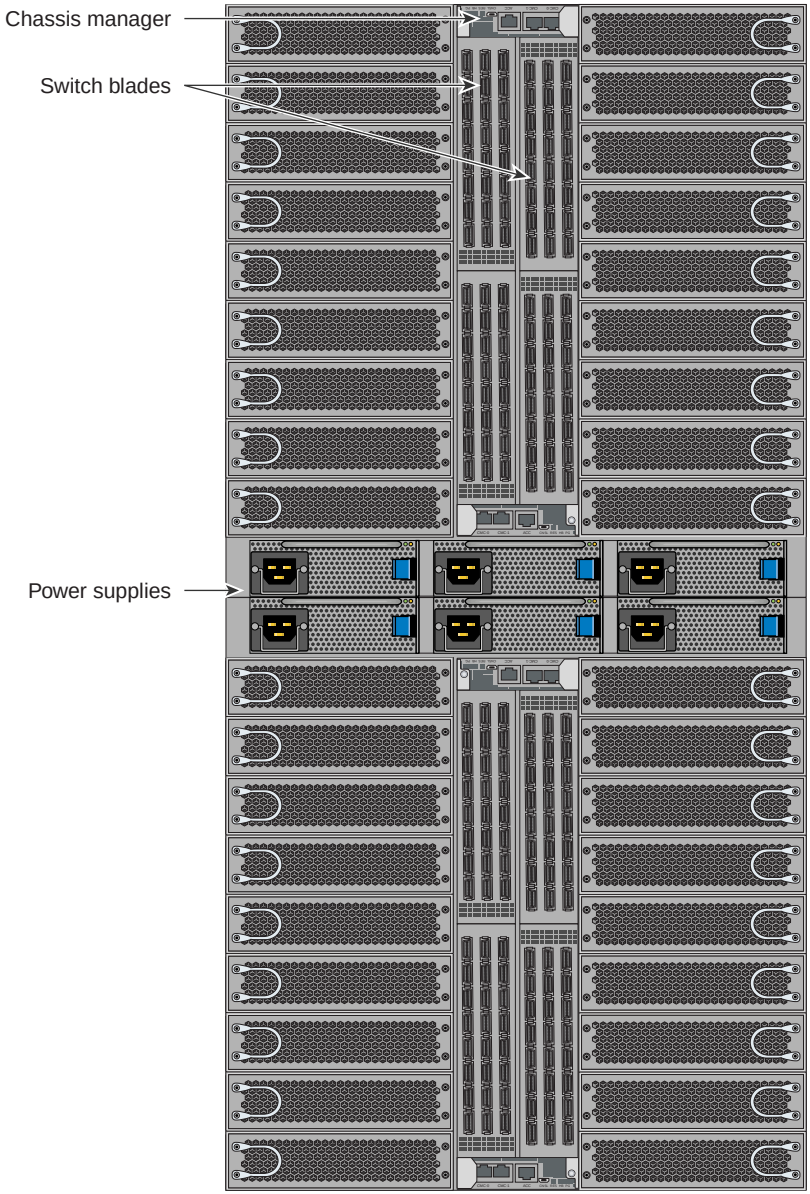

Figure 3-9 SGI ICE X Series Blade Enclosure Pair Components Example

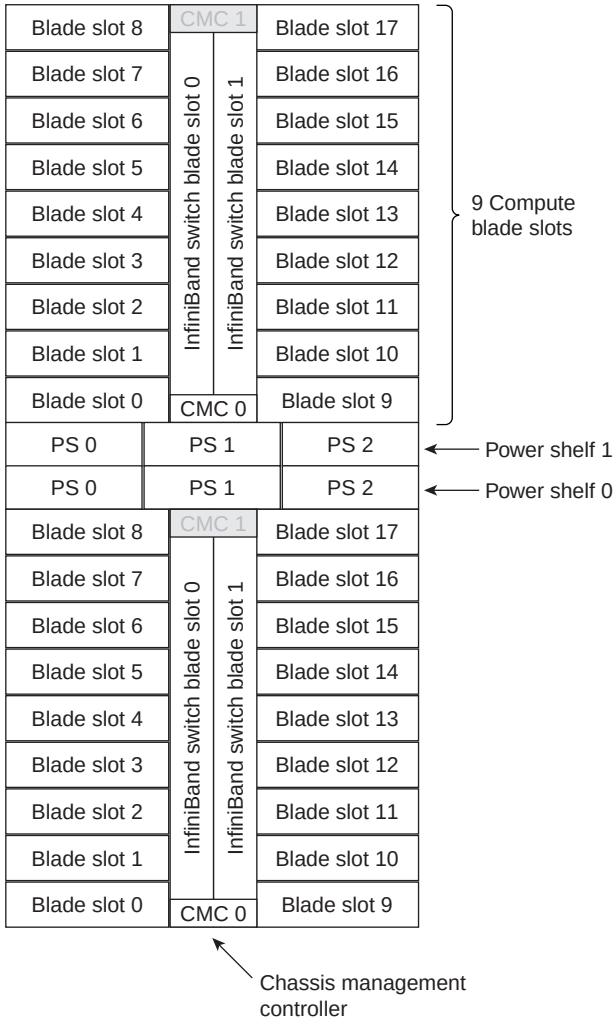

Figure 3-10 Single-node Blade Enclosure Pair Component Front Diagram

Note: Enclosures using single-node blades use one CMC, enclosures using dual-node blades must use two CMC boards.

Unit Numbering

Blade enclosures 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). Enclosures within a rack are identified by the use of module IDs 0, 1, 2, and 3, with enclosure 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 blade 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 (001).

The number of the first blade enclosure will always be zero (0). These numbers are used to identify components starting with the rack, including the individual blade enclosures 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) as well as the compute blades that reside in that rack.

Optional System Components

Availability of optional components for the SGI ICE X 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 X racks in the following sections:

Overview

At the time this document was published only the tall (42U) SGI ICE X rack (shown in Figure 4-1 on page 47) was approved for ICE X system racks shipped from the SGI factory.

SGI ICE X Series Rack (42U)

The SGI 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). A front door is required on every rack.

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

  Up to four optional 10.5 U-high (18.25-inch) water-cooled doors can be installed on the rear of the SGI ICE X rack.

  Each air-cooled rack has a key lock to prevent unauthorized access to the system via the rear door, see Figure 4-4 on page 50. In a system made up of multiple air-cooled racks, rear doors have a master key that locks and unlocks all rear doors in a system. 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 enclosure(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 fourteen outlets are required for a single enclosure pair system as follows:

  – up to 6 outlets for an enclosure pair (depending on configuration)

  – two outlets for the rear fan (blower) enclosure power supplies

  – 4 outlets for administration and RLC servers (in primary rack)

  – 2 outlets for a service node (server)

  – Allow eight or more outlets for an additional enclosure pair in the system

Note that up to 12 power outlets may be needed to power a single blade enclosure pair and supporting servers installed in a single rack. Optional single-phase PDUs can be used in SGI ICE X racks dedicated to I/O functionality.

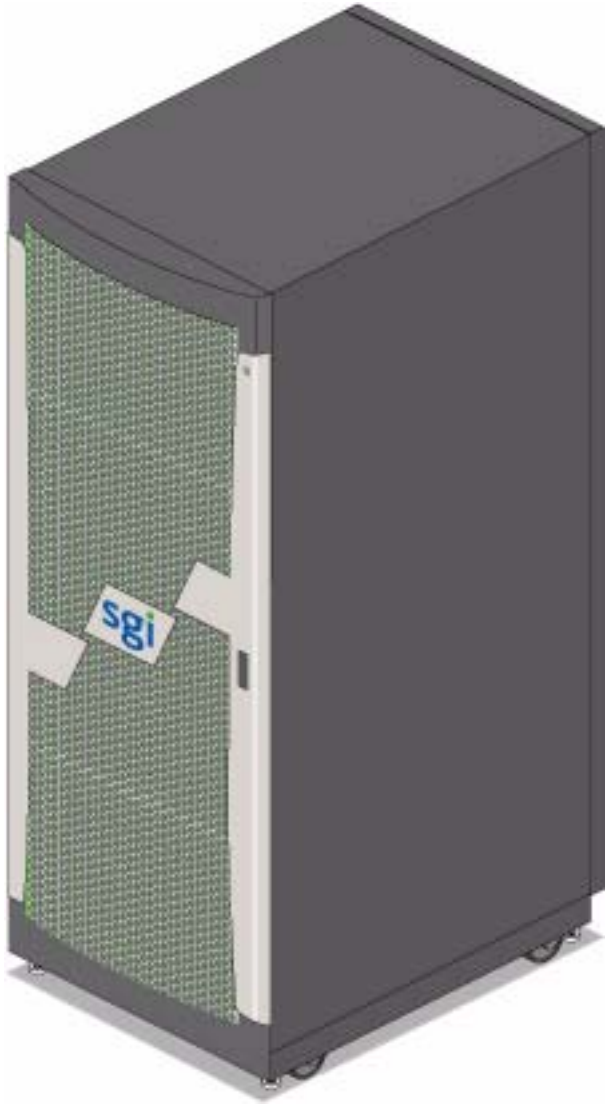

Figure 4-1 SGI ICE X Series Rack Example

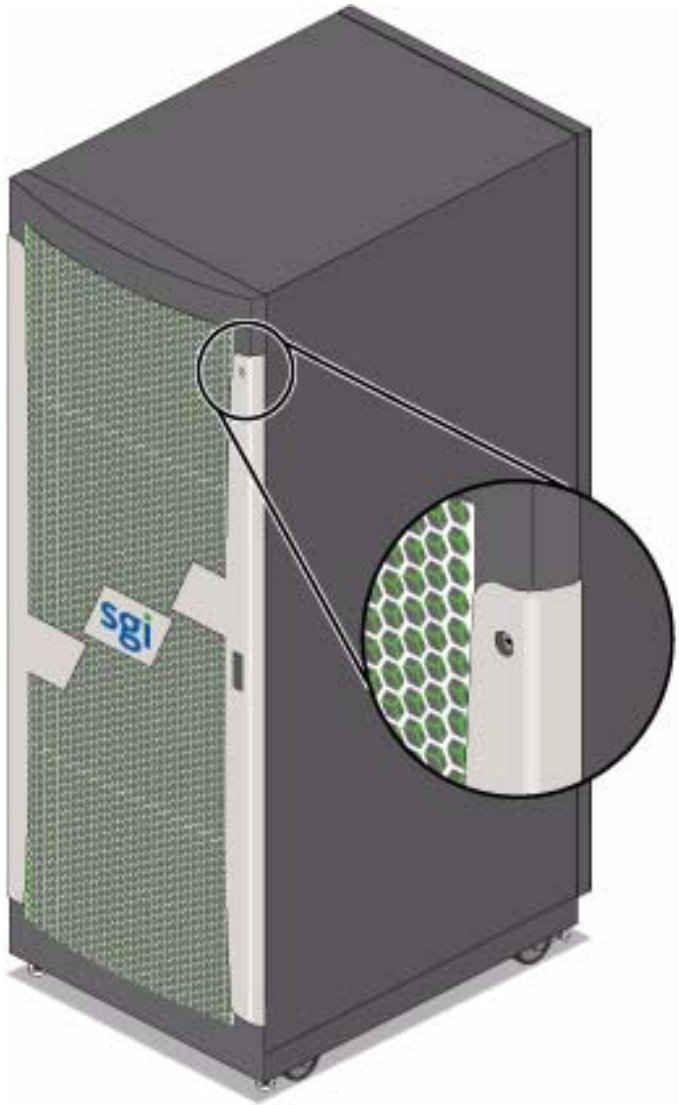

Figure 4-2 Front Lock on Tall (42U) Rack

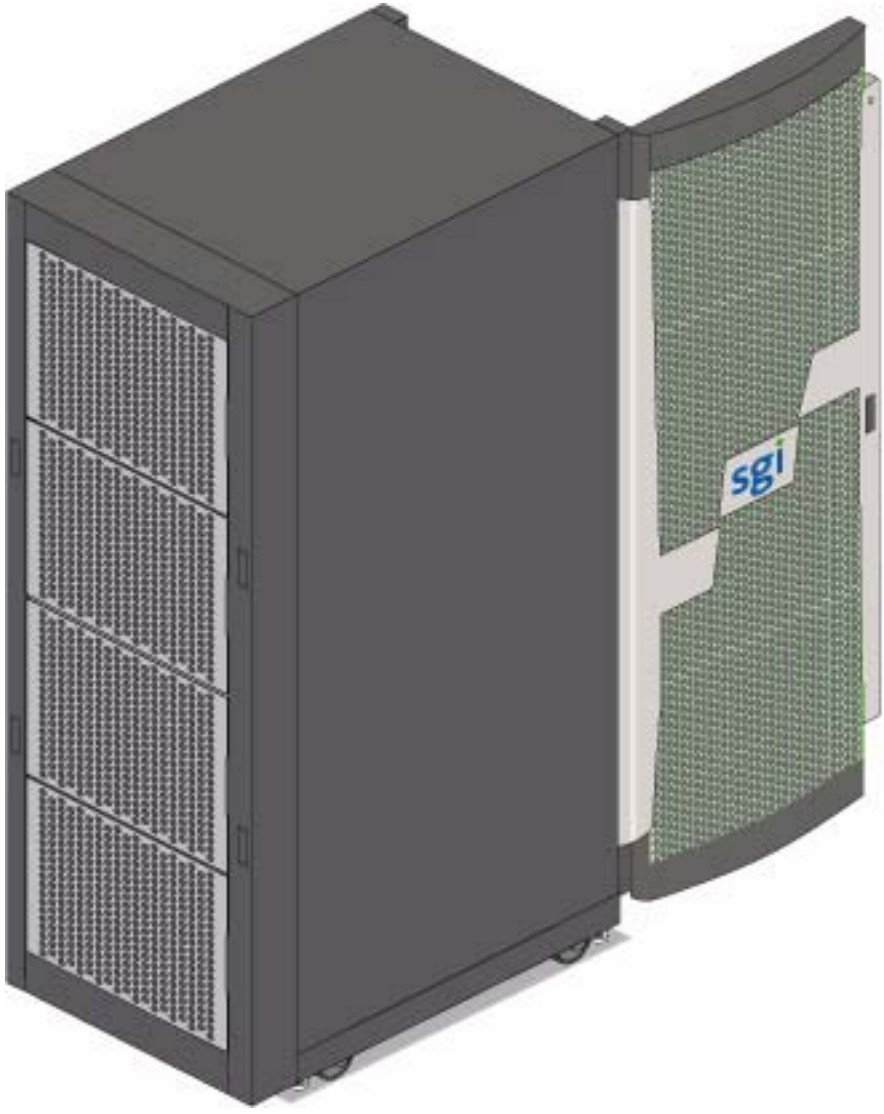

Figure 4-3 Optional Water-Chilled Door Panels on Rear of ICE X Rack

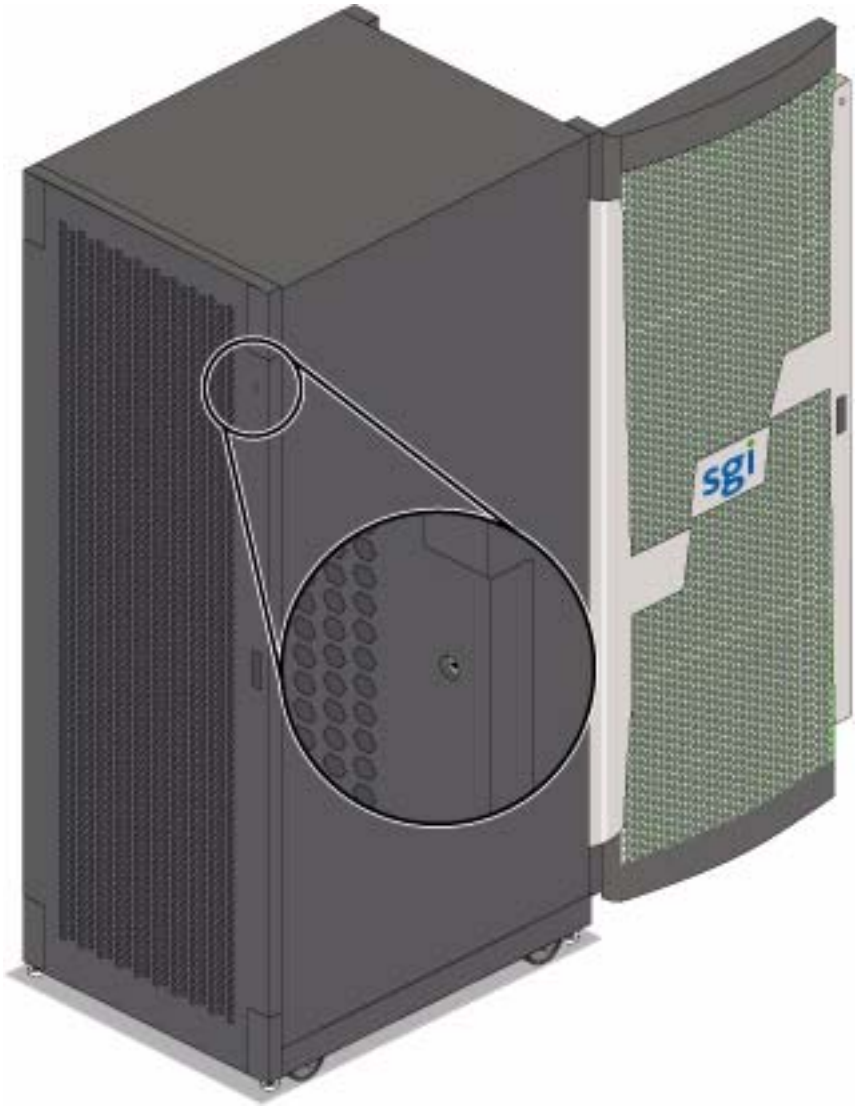

Figure 4-4 Air-Cooled Rack Rear Door and Lock Example

ICE X Rack Technical Specifications

Table 4-1 lists the technical specifications of the SGI ICE X series tall rack.

Table 4-1 Tall SGI ICE X Rack Technical Specifications

Characteristic	Specification
Height	79.5 in. (201.9 cm) 82.25 in (208.9 cm) with 2U top
Width	24 in. (61 cm) - optionally expandable
Depth	49.5 in. (125.7 cm) - air cooled; 50.75 in. (128.9 cm) - water cooled
Weight (full)	~2,500 lbs. (1,136 kg) approximate (water cooled)
Shipping weight (max)	~2,970 lbs. (1,350 kg) approximate maximum
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	3-phase (optional single-phase available in I/O rack)
Power requirements (max)	34.58 kVA (33.89 kW)
Hold time	16 ms
Power cable	12 ft. (3.66 m) pluggable cords

Important: The rack's optional water-cooled door panels only provide cooling for the bottom 42U of the rack. If the top of the rack is "expanded" 2U, 4U, or 6U, to accommodate optional system components, the space in the extended zone is **not** water cooled.

See "System-level Specifications" in Appendix A for a more complete listing of SGI ICE X system operating specifications and environmental requirements.

SGI ICE X 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:

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 SGI ICE X system primarily include:

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

Other servers described in this chapter 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)
- MDS node (Lustre configurations)
- OSS node (Lustre configurations)

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 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 that 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 X system through an Ethernet network.

System management hierarchy

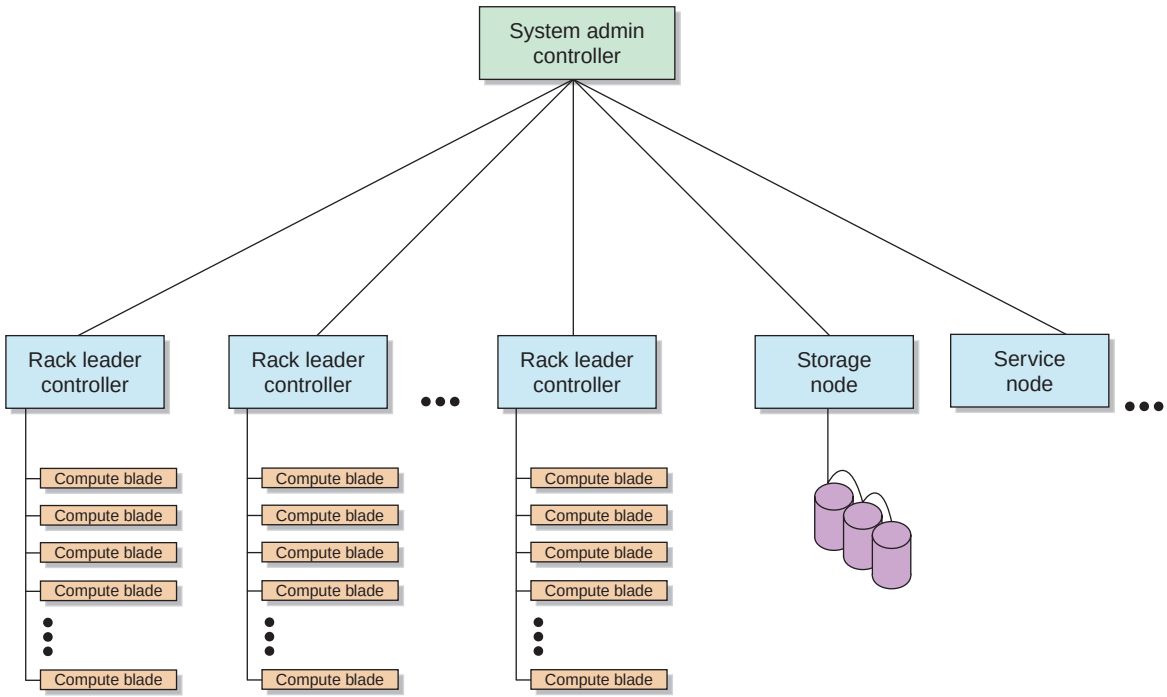

Figure 5-1 SGI ICE X System Administration Hierarchy Example Diagram

The administrative controller server is at the top of the distributed management infrastructure within the SGI ICE X system. The overall SGI ICE X series management is hierarchical (see Figure 5-1 on page 54), with the RLC(s) communicating with the compute nodes via CMC interconnect.

1U Rack Leader Controller and Administration 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 SGI ICE X 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 for every two racks).

The system administrative controller unit acts as the SGI ICE X 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.

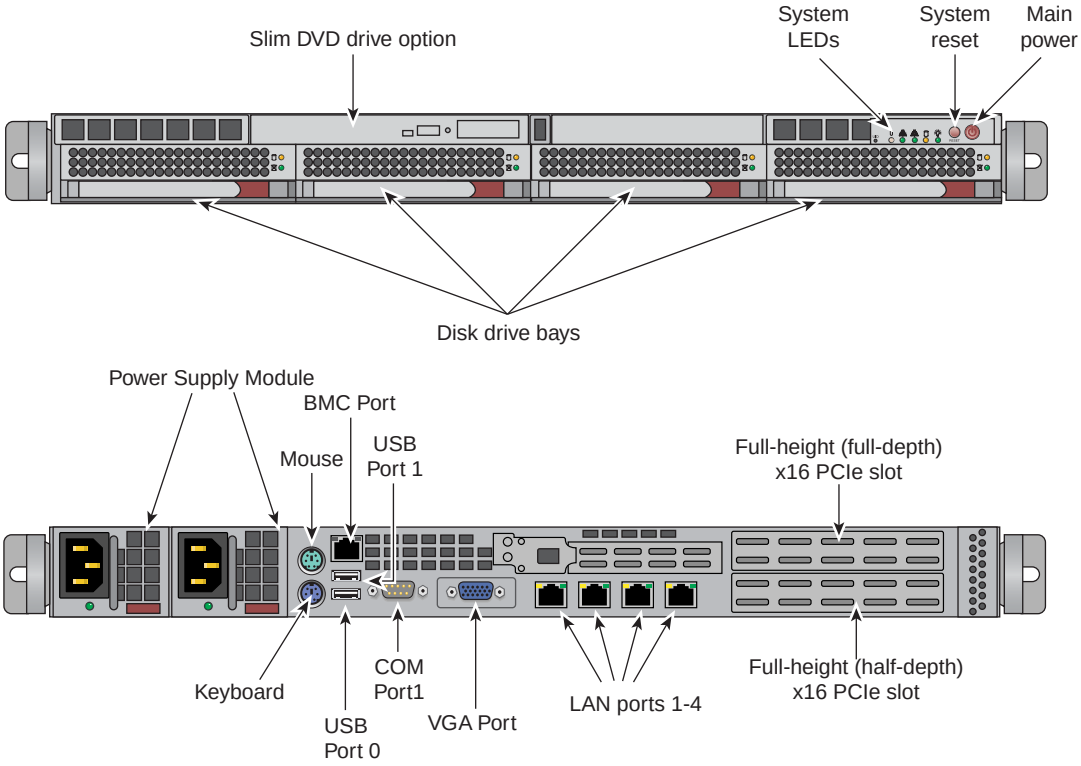

Figure 5-2 1U Rack Leader Controller (RLC) Server Front and Rear Panels

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. The 1U server may also be used as a separate (non-RLC/admin) login, batch, I/O, MDS, OSS or fabric management node. See the section “Modularity and Scalability” on page 30 for a list of administration and support server types and additional functional descriptions.

2U Service Node

For systems using a separate login, batch, I/O, fabric management, or other service node; this 2U server is also an available option. Figure 5-3 and Figure 5-4 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.

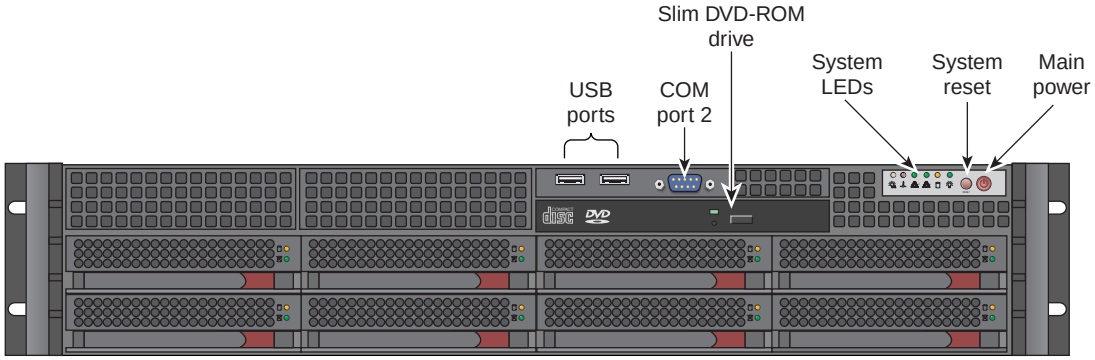

Figure 5-3 Front View of 2U Service Node

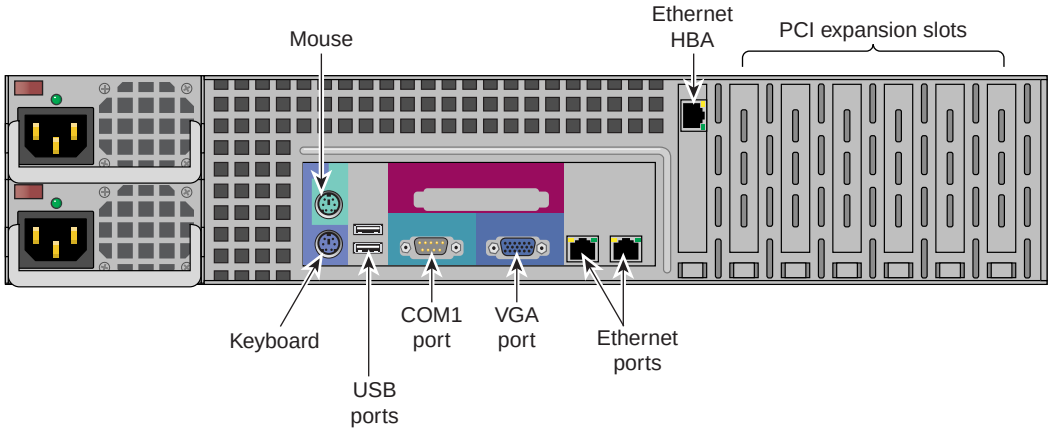

Figure 5-4 Rear View of 2U Service Node

See the *SGI Rackable C2108-TY10 System User's Guide* (P/N 007-5688-00*x*) for more detailed information on the 2U service node. The 2U server's control panel features are shown in Figure 5-4.

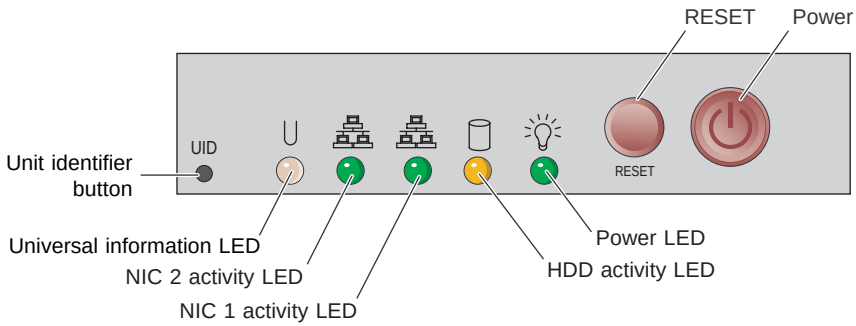

Figure 5-5 2U Service Node Control Panel Diagram

Table 5-1 2U 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.

Optional 3U Service Nodes

The SGI ICE X system also offers a 3U-high service node as a separate login, batch, I/O, fabric management, MDS, OSS or graphics support node. Under specific circumstances the 3U server can be configured as a mass storage resource for the SGI ICE X system. Figure 5-6 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 007-5687-00*x*).

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 SGI ICE X system.

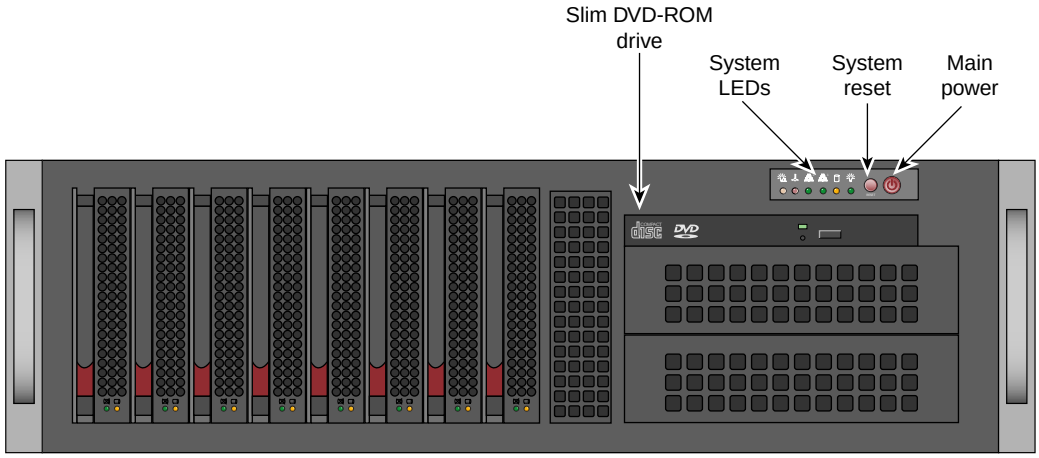

Figure 5-6 SGI 3U Optional Service Node Front View

Figure 5-7 on page 60 shows an example rear view of the 3U service node.

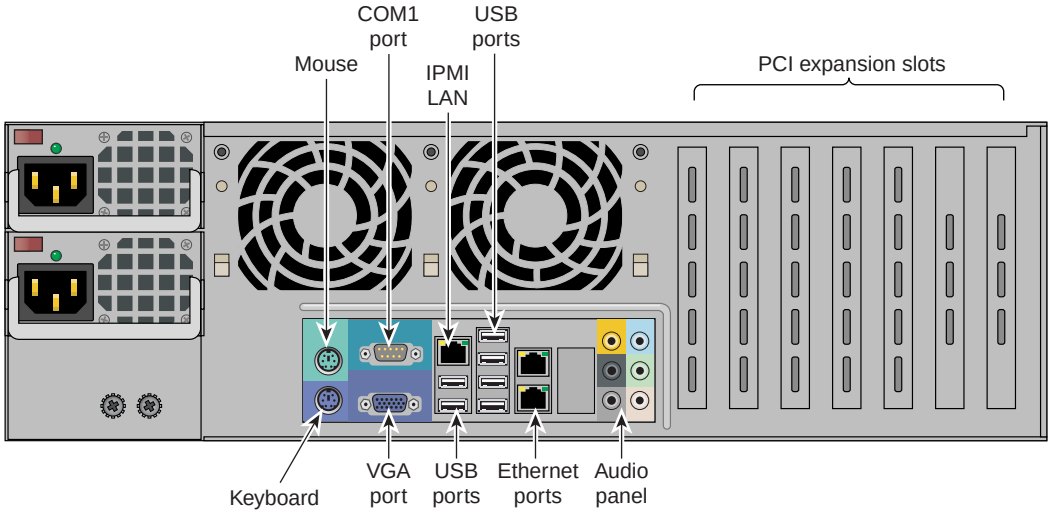

Figure 5-7 SGI 3U Service Node Rear View

Optional 4U Service Nodes

The highest performance optional service node in the SGI ICE X system is offered as a 4U-high service node. It can serve as a separate login, batch, I/O, fabric management, MDS, OSS 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 SGI ICE X system.

Figure 5-8 on page 61 shows the front controls and interfaces available on the server. Table 5-2 on page 61 describes the front panel control and interface functions on the 4U server.

Figure 5-9 on page 62 calls out the components used on the front of the 4U server. Table 5-3 on page 62 identifies the components called out in the figure. Rear components used on the 4U server are shown in Figure 3-8 on page 36.

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

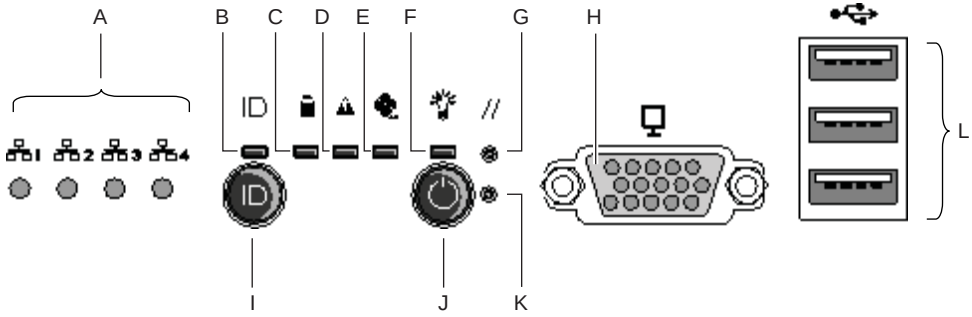

Figure 5-8 4U Service Node Front Controls and Interfaces

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

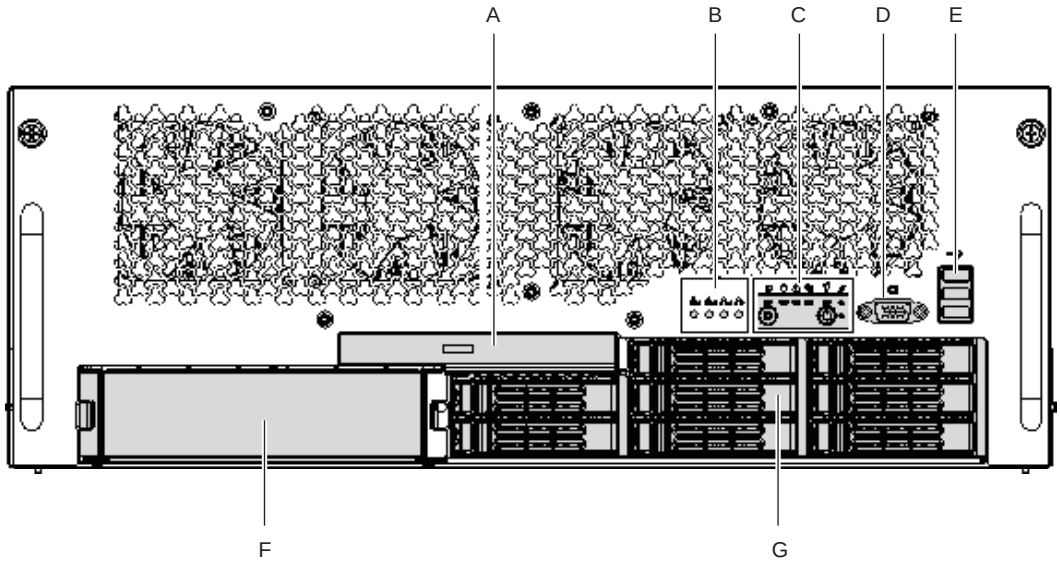

Figure 5-9 4U Service Node Front Panel

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

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.	Ensure that the power cords of the enclosure are seated properly in the power receptacles. Ensure that the PDU circuit breakers are on and properly connected to the wall source. If the power cord is plugged in and the circuit breaker is on, contact your SSE.
An enclosure pair will not power on.	Ensure the power cables of the enclosure are plugged in and the PDU is turned on. View the CMC output from your system administration controller console. If the CMC is not running, contact your SSE.
The system will not boot the operating system.	Contact your SSE.
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.
The amber LED of a system power supply is on.	Replace the power supply.

LED Status Indicators

There are a number of LEDs visible on the front of the blade enclosures 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.

Blade Enclosure Pair Power Supply LEDs

Each power supply installed in a blade enclosure pair (six total) has one green and one amber status LED located at the right edge of the supply. Each of the LEDs (see Figure 6-1) will either light green or amber (yellow), stay dark, or flash green or yellow to indicate the status of the individual supply. See Table 6-2 for a complete list.

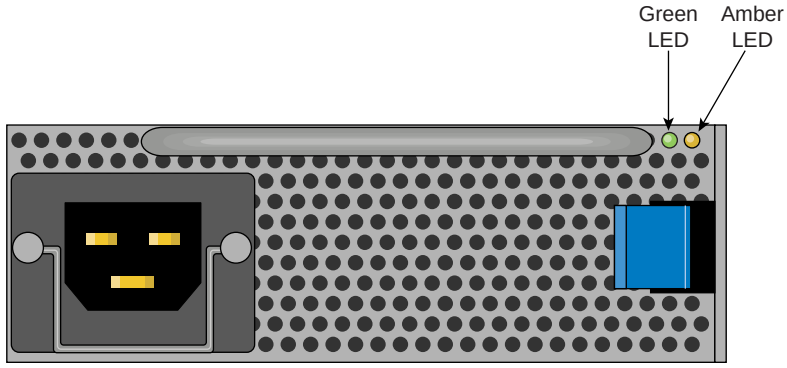

Figure 6-1 Power Supply Status LED Indicator Locations

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 - solid
Power supply problem warning	Off	Blinking
AC available to supply (standby) but enclosure is powered off	Blinking	Off
Power supply on - function normal	On	Off

Compute Blade LEDs

Each compute blade installed in an enclosure has status 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 6-2 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 blade.
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. 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.

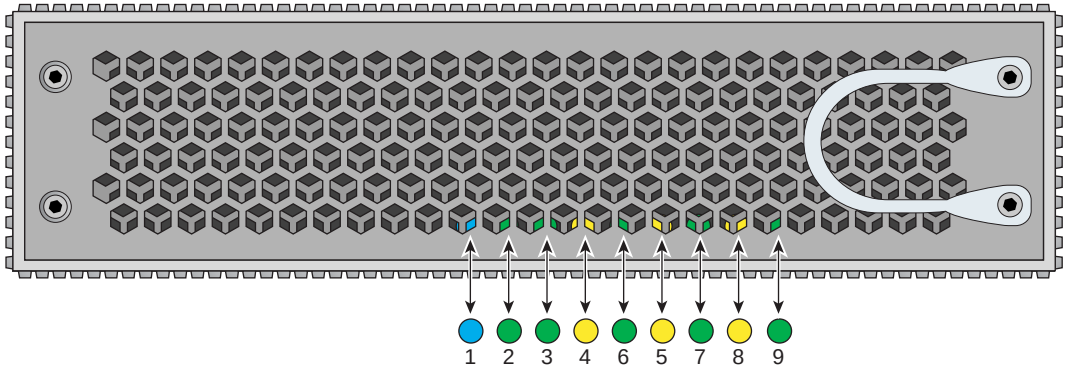

Figure 6-2 Compute Blade Status LED Locations Example

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

Maintenance Procedures

This chapter provides information about installing or removing components from your SGI ICE X system, as follows:

Maintenance Precautions and Procedures

This section describes how to access the system for specific types of customer approved maintenance and protect the components from damage. The following topics are covered:

Preparing the System for Maintenance or Upgrade

To prepare the system for maintenance, you can follow the guidelines in “Powering On and Off” on page 8 and power down the affected blade enclosure pair. The section also has information on powering-up the enclosure after you have completed the maintenance/upgrade required.

If your system does not boot correctly, see Chapter 6 for troubleshooting procedures.

Installing or Removing Internal Parts

Caution: The components inside the system are extremely sensitive to static electricity. Always wear a wrist strap when you work with parts inside your system.

To use the wrist strap, follow these steps:

1. Unroll the first two folds of the band.
2. Wrap the exposed adhesive side firmly around your wrist, unroll the rest of the band, and then peel the liner from the copper foil at the opposite end.
3. Attach the copper foil to an exposed electrical ground, such as a metal part of the chassis.

Caution: Do not attempt to install or remove components that are not listed in Table 7-1. Components not listed must be installed or removed by a qualified SGI field engineer.

Table 7-1 lists the customer-replaceable components and the page on which you can find the instructions for installing or removing the component.

Table 7-1 Customer-replaceable Components and Maintenance Procedures

Component	Procedure
Blade enclosure power supply	“Removing and Replacing a Blade Enclosure Power Supply” on page 69
Enclosure fans (blowers)	“Removing and Replacing Rear Fans (Blowers)” on page 72
Enclosure blower power supplies	“Removing a Fan Assembly Power Supply” on page 76

Replacing ICE X System Components

While many of the blade enclosure components are not considered end-user replaceable, a select number of components can be removed and replaced. These include:

- Blade enclosure pair power supplies (front of system)
- Rear-mounted blade enclosure cooling fans (also called blowers)
- Cooling fan power supplies (rear of system)

Removing and Replacing a Blade Enclosure Power Supply

To remove and replace power supplies in a blade enclosure, you do not need any tools.

Under most circumstances a single power supply in a blade enclosure pair can be replaced without shutting down the enclosure or the complete system. In the case of a fully configured (loaded) enclosure, this may not be possible.

Caution: The body of the power supply may be hot; allow time for cooling and handle with care.

Use the following steps to replace a power supply in the blade enclosure box:

1. Open the front door of the rack and locate the power supply that needs replacement.
2. Disengage the power-cord retention clip and disconnect the power cord from the power supply that needs replacement.
3. Press the retention latch of the power supply toward the power connector to release the supply from the enclosure, see Figure 7-1 on page 70.
4. Using the power supply handle, pull the power supply straight out until it is partly out of the chassis. Use one hand to support the bottom of the supply as you fully extract it from the enclosure.

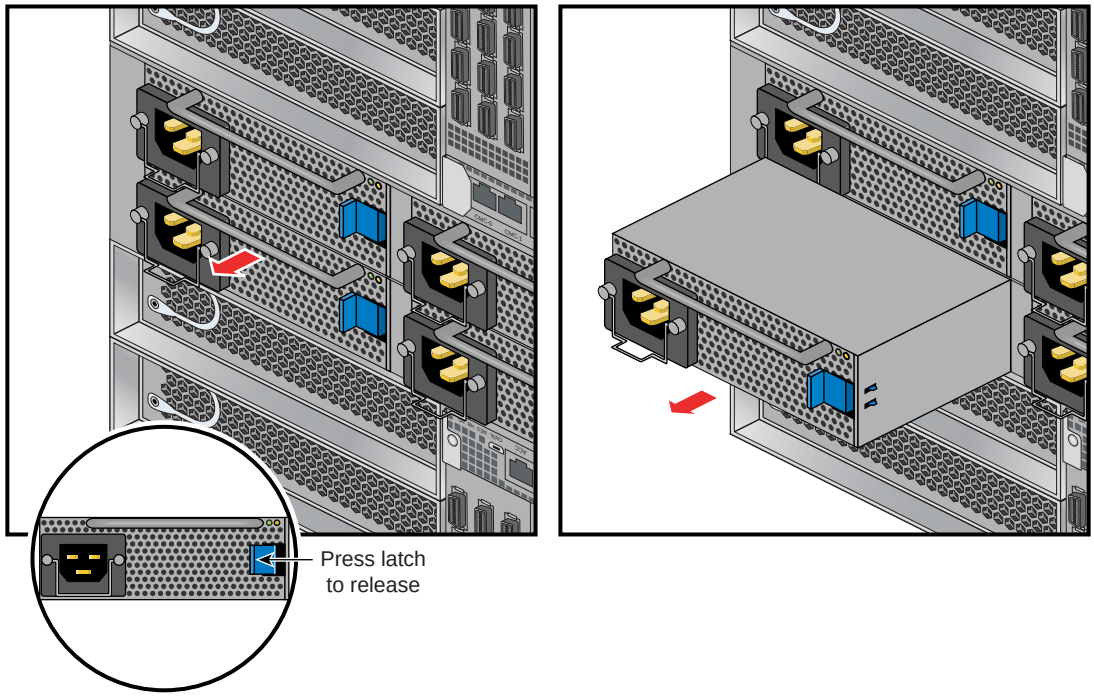

Figure 7-1 Removing an Enclosure Power Supply

5. Align the rear of the replacement power supply with the enclosure opening.
6. Slide the power supply into the chassis until the retention latch engages.
7. Reconnect the power cord to the supply and engage the retention clip.

Note: When AC power to the rear fan assembly is disconnected prior to the replacement procedure, all the fans will come on and run at top speed when power is reapplied. The speeds will readjust when normal communication with the blade pair enclosure CMC is fully established.

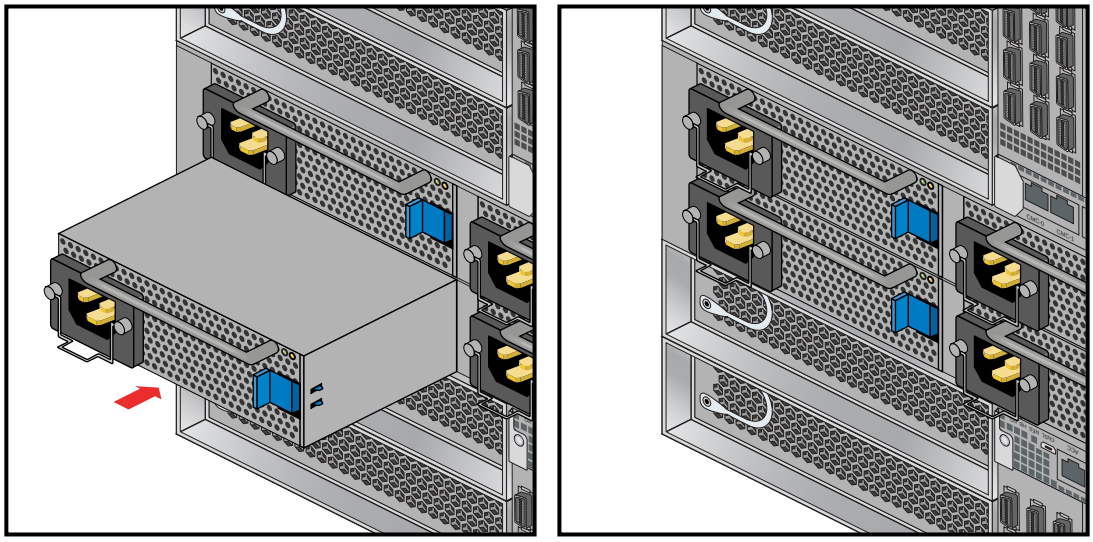

Figure 7-2 Replacing an Enclosure Power Supply

Removing and Replacing Rear Fans (Blowers)

The blade enclosure cooling fan assembly (blower enclosure) is positioned back-to-back with the blade enclosure pair. You will need to access the rack from the back to remove and replace a fan. The enclosure's system controller issues a warning message when a fan is not running properly. This means the fan RPM level is not within tolerance. When a cooling fan fails, the following things happen:

1. The system console will show a warning indicating the rack and enclosure position

   `001c01 L2> Fan (`*`number`*`) warning limit reached @ 0 RPM`

2. A line will be added to the L1 system controller's log file indicating the fan warning.
3. If optional SGI Embedded Support Partner (ESP) is used, a warning message will be sent to it also.

The chassis management controller (CMC) monitors the temperature within each enclosure. If the temperature increases due to a failed fan, the remaining fans will run at a higher RPM to compensate for the missing fan. The system will continue running until a scheduled maintenance occurs.

The fan numbers for the enclosure (as viewed from the rear) are shown in Figure 7-3 on page 73.

Note that under most circumstances a fan can be replaced while the system is operating. You will need a #1 Phillips-head screw driver to complete the procedure.

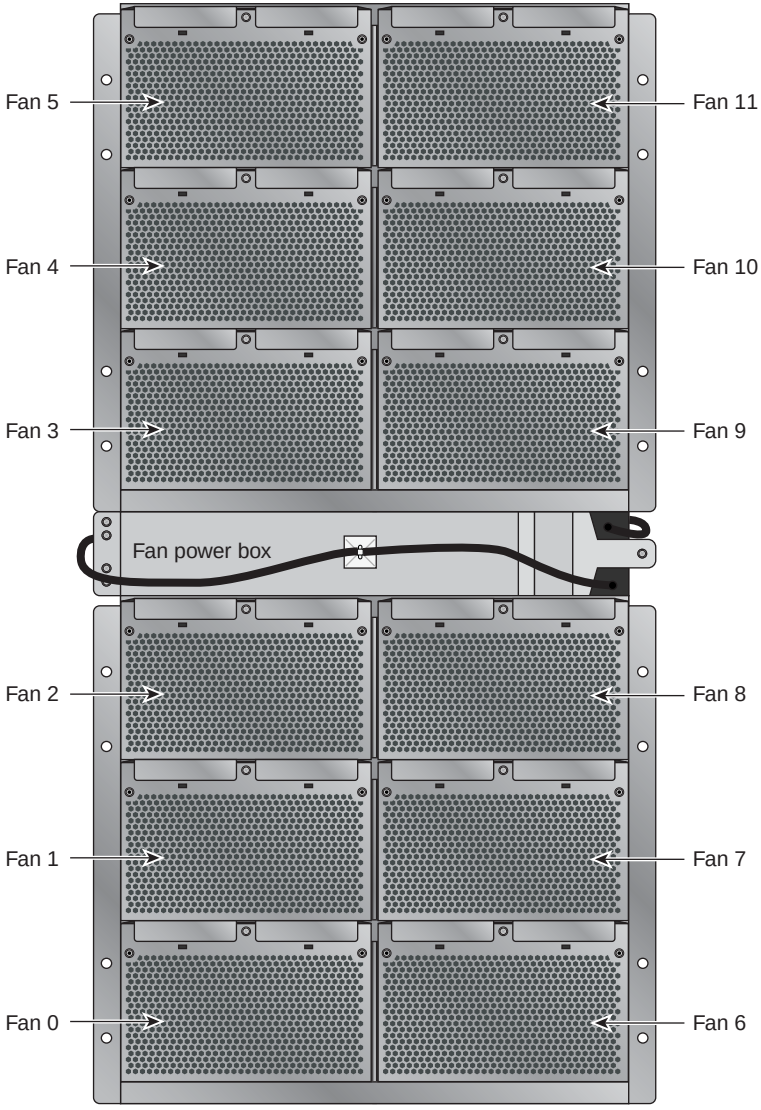

Figure 7-3 Enclosure-Pair Rear Fan Assembly (Blowers)

Use the following steps and illustrations to replace an enclosure fan:

1. Using the #1 Phillips screwdriver, undo the (captive) screw (located in the middle of the blower assembly handle). The handle has a notch for the screw access, see Figure 7-4.
2. Grasp the blower assembly handle and pull the assembly straight out.

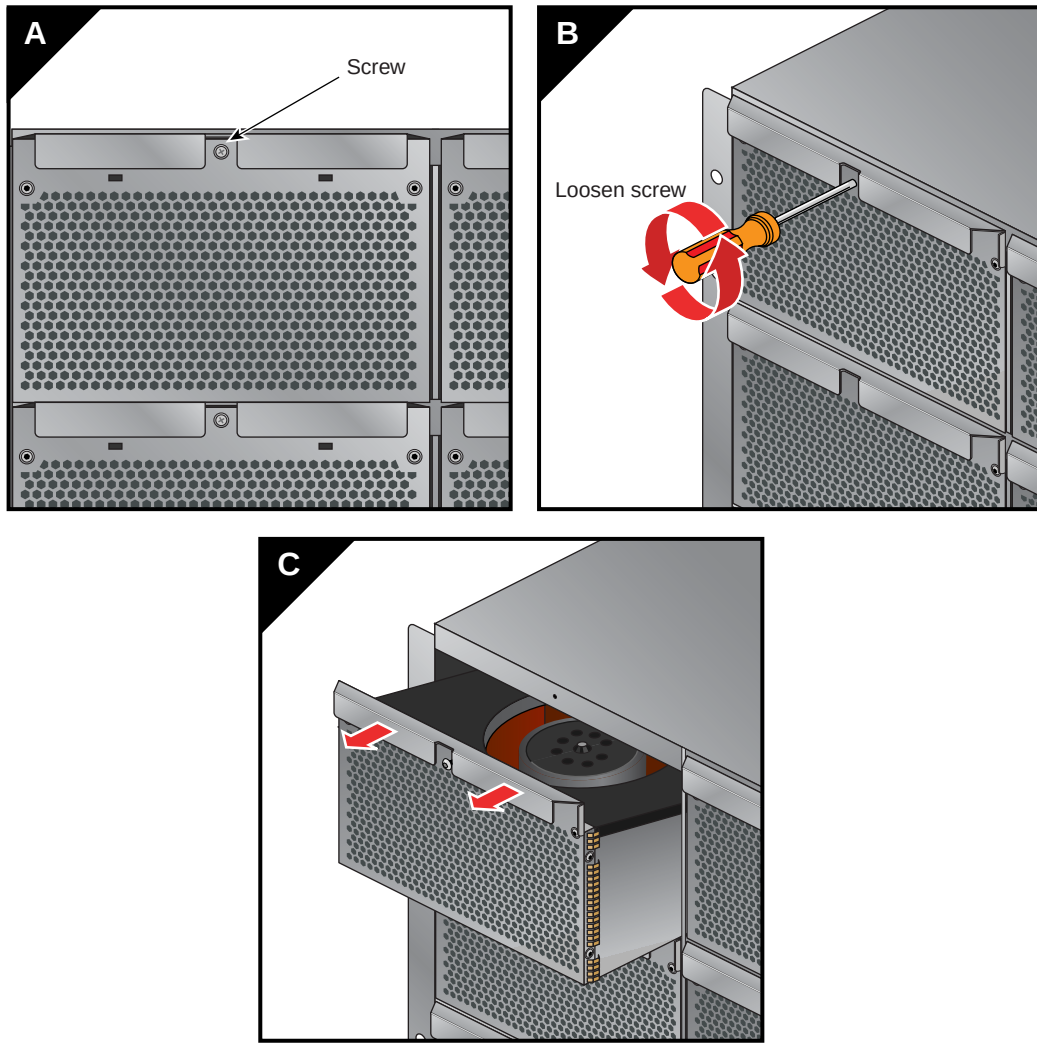

Figure 7-4 Removing a Fan From the Rear Assembly

3. Slide a new blower assembly completely into the open slot, see Figure 7-5.
4. Tighten the blower assembly screw to secure the new fan.

Note: If you disconnected the AC power to the rear fan assembly prior to the replacement procedure, all the fans will come on and run at top speed when power is reapplied. The speeds will readjust when normal communication with the blade pair enclosure CMC is fully established.

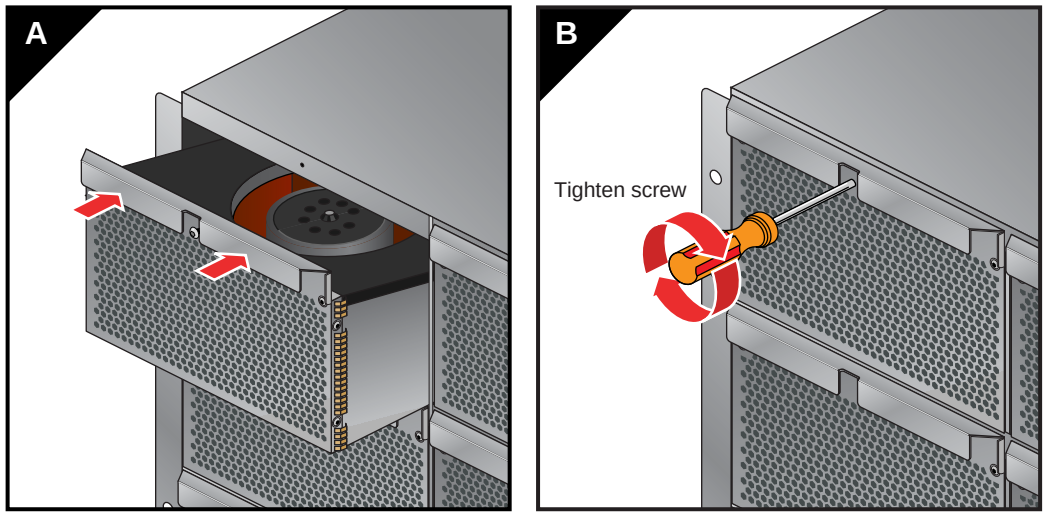

Figure 7-5 Replacing an Enclosure Fan

Removing or Replacing a Fan Enclosure Power Supply

The 12-fan (blower) assembly that is mounted back-to-back with the blade enclosure pair to provide cooling uses two power supplies to provide voltage to the blowers. Removal and replacement of a blower assembly power supply requires the use of a T-25 torx driver.

Removing a Fan Assembly Power Supply

Use the following information and illustrations to remove a power supply from the fan (blower) assembly enclosure:

1. Open the rear door of the rack and locate the fan power supply access door. The access door will be located between the upper and lower blower sets.
2. Use a T-25 torx driver to undo the screw that holds the supply access door (on the right) to the fan enclosure chassis.

Note: You may have to adjust or move power or other cables to enable the access door to swing outward.

---

3. Move the fan power box outward so that the front of the supply is fully accessible.
4. Disconnect the power cord from the supply that is to be replaced. If the supply has been active, allow several minutes for it to cool down.
5. Push the power supply retention tab towards the center of the supply to release it from the fan power box.
6. Pull the supply out of the fan power box while supporting it from beneath.

Replacing a Fan Power Supply

Use the following steps to replace a fan power supply:

1. Align the rear of the power supply with the empty fan power box.
2. Slide the unit all the way in until the supply's retention tab snaps into place.
3. Reconnect the power cable to the supply and secure the cable retention clip.
4. Move the fan power box inward until the access door is again flush with the rear of the rack.
5. Use the T-25 torx driver to secure the power box door screw to the rear of the fan enclosure.

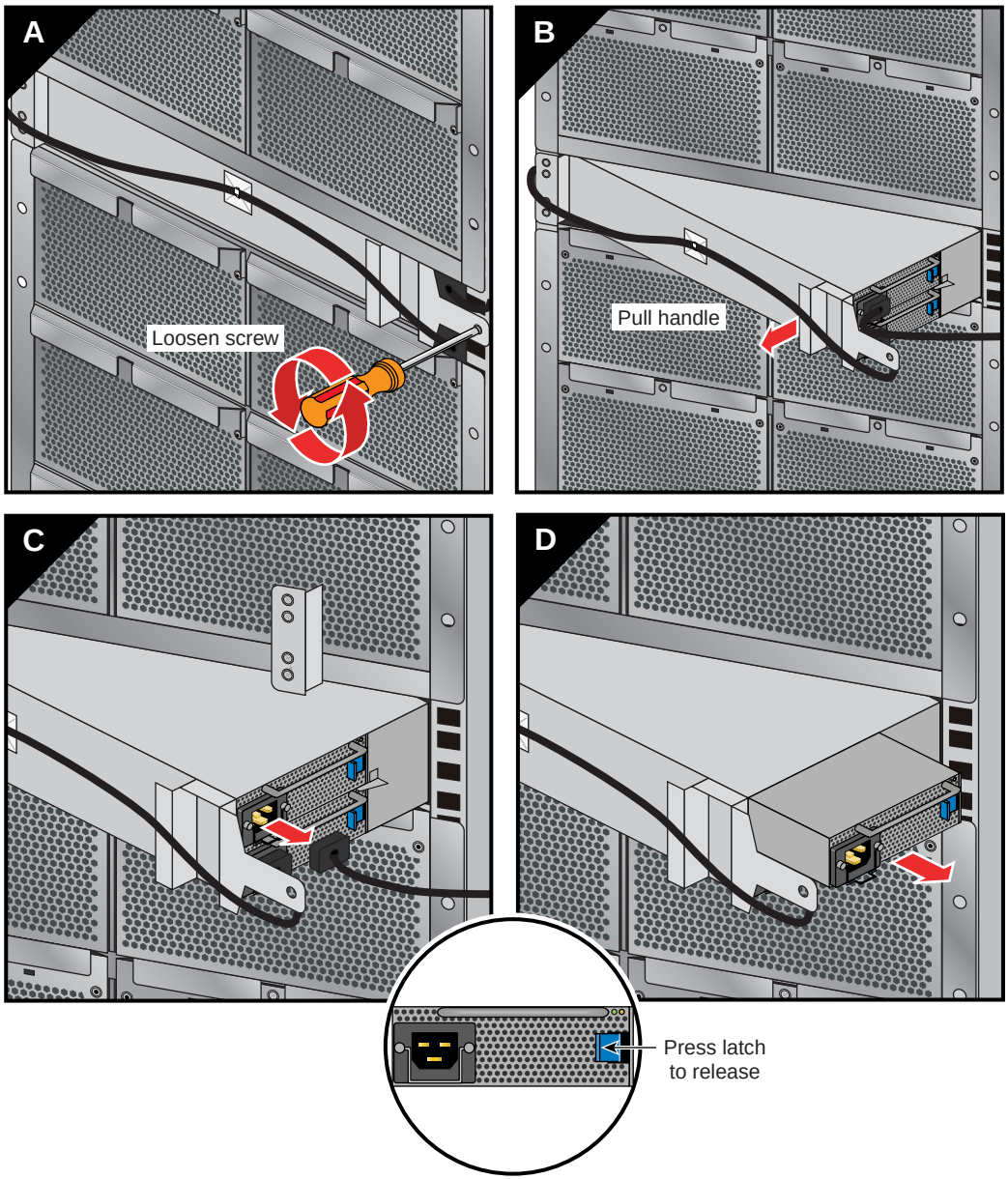

Figure 7-6 Removing a Power Supply From the Fan Power Box

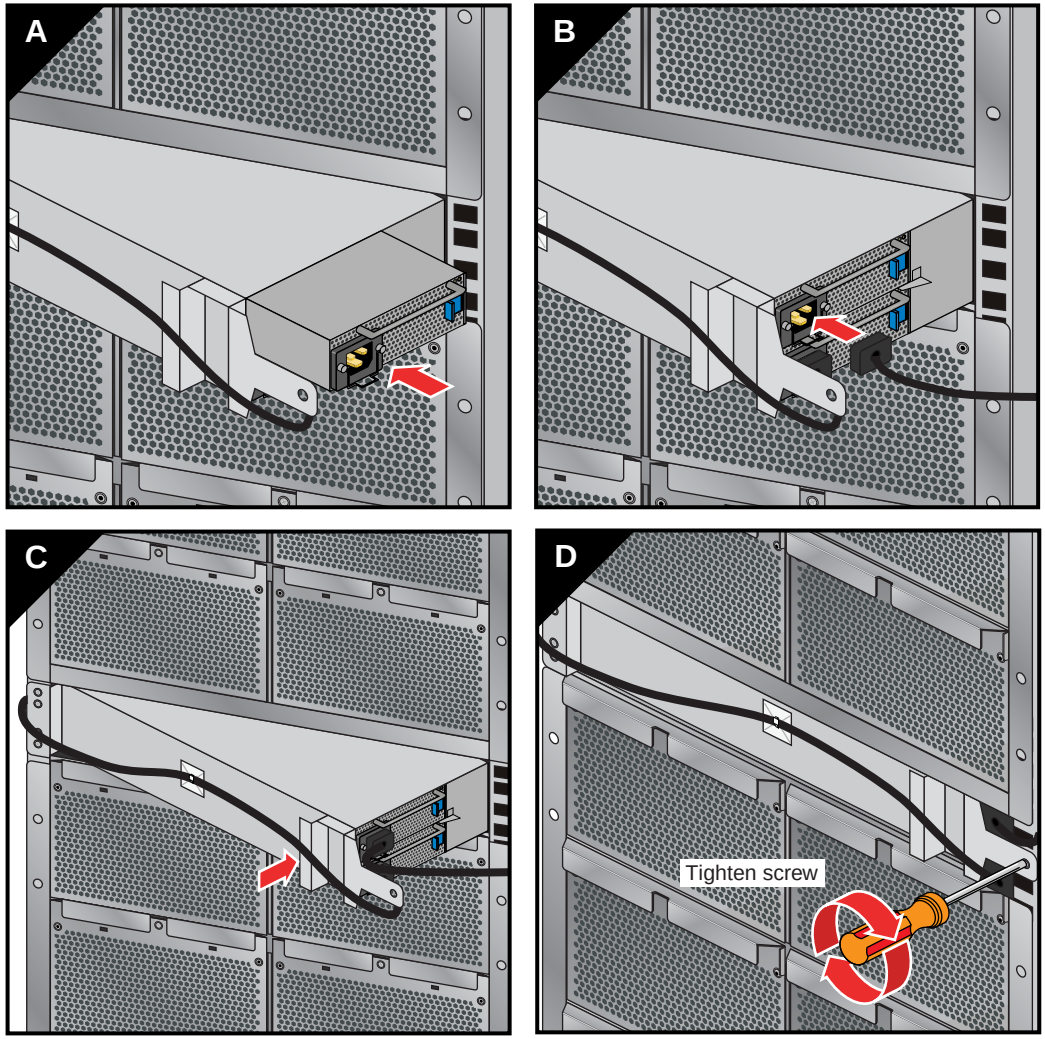

Figure 7-7 Replacing a Power Supply in the Fan Power Box

Overview of PCI Express Operation

This section provides a brief overview of the PCI Express (PCIe) technology that will be available as an option with your system's stand-alone administration, RLC and service nodes. PCI Express has both compatibility and differences with older PCI/PCI-X technology. Check with your SGI sales or service representative for more detail on PCI Express board options available with your SGI ICE X system.

PCI Express is compatible with PCI/PCI-X in the following ways:

- Compatible software layers
- Compatible device driver models
- Same basic board form factors
- PCIe controlled devices appear the same as PCI/PCI-X devices to most software

PCI Express technology is different from PCI/PCI-X in the following ways:

- PCI Express uses a point-to-point serial interface vs. a shared parallel bus interface used in older PCI/PCI-X technology
- PCIe hardware connectors are not compatible with PCI/PCI-X (see Figure 7-8)
- Potential sustained throughput of x16 PCI Express is approximately four times that of the fastest PCI-X throughputs

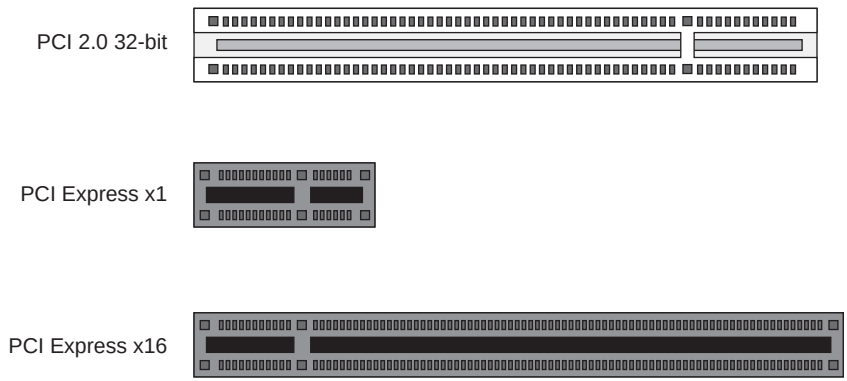

Figure 7-8 Comparison of PCI/PCI-X Connector with PCI Express Connectors

PCI Express technology uses two pairs of wires for each transmit and receive connection (4 wires total). These four wires are generally referred to as a lane or x1 connection (also called “by 1”). SGI administrative node PCIe technology uses x16, x8 and x4 connector technology in the PCI Express card slots (see Figure 1-2 on page 4 for an example). The PCIe technology will support PCIe boards that use connectors up to x16 in size. Table 7-2 shows this concept.

Table 7-2 SGI Administrative Server PCIe Support Levels

SGI Admin PCIe Connectors	
x1 PCIe cards	Supported
x2 PCIe cards	Supported
x4 PCIe cards	Supported
x8 PCIe cards	Supported
x16 PCIe cards	Two supported
x32 PCIe cards	Not supported

If you need more specific information on installing PCIe cards in an administrative, leader, or other standalone server, see the user documentation for that particular unit. After installing or removing a new PCIe card, do the following:

1. Return the server to service.
2. Boot your operating system software. (See your software operation guide if you need instructions to boot your operating system.)
3. Run the `lspci` PCI hardware inventory command to verify the installation. This command lists PCI hardware that the operating system discovered during the boot operation.

Technical Specifications and Pinouts

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

System-level Specifications

Table A-1 summarizes the SGI ICE X series configuration ranges.

Table A-1 SGI ICE X Series Configuration Ranges

Category	Minimum	Maximum
Blades per enclosure pair	2 blades[a]	36 blades
Blade enclosure pair	1 per rack	2 per rack
Compute blade DIMM capacity	8 DIMMs per blade	16 DIMMs per blade
Chassis management blades	2 per enclosure pair	4 per enclosure pair
InfiniBand switch blades	2 per enclosure pair	4 per enclosure pair

a. Compute blades support two stuffed sockets each.

Physical and Power Specifications

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

Table A-2 ICE X System Rack Physical Specifications

System Features (single rack)	Specification
Height	79.5 in. (201.9 cm) 82.25 in (208.9 cm) with 2U top
Width	24.0 in. (61 cm) - air and water cooled
Depth	49.5 in. (125.7 cm) - air cooled; 50.75 in. (128.9 cm) - water cooled
Weight (full) maximum	~2,500 lbs. (1,136 kg) approximate (water cooled)
Shipping weight maximum	~2,970 lbs. (1,350 kg) approximate maximum
Shipping height maximum	88.75 in. (225.4 cm)
Shipping width	44 in. (111.8 cm)
Shipping depth	62.75 in. (159.4 cm)
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	3-phase (optional single-phase available in I/O rack)
Power requirements (max)	34.58 kVA (33.89 kW)
Hold time	16 ms
Power cable	12 ft. (3.66 m) pluggable cords
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.

Table A-3 Environmental Specifications (Single Rack)

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)
Rack cooling requirements	Ambient air or optional water cooling
Heat dissipation to air Air-cooled ICE X (rack)	Approximately 115.63 kBTU/hr maximum (based on 33.89 kW - 100% dissipation to air)
Heat dissipation to air Water-cooled ICE X (rack)	Approximately 5.76 kBTU/hr maximum (based on 33.89 kW - 5% dissipation to air)
Heat dissipation to water	Approximately 109.85 kBTU/hr maximum (based on 33.89 kW - 95% dissipation to water)
Air flow: intake (front), exhaust (rear)	Approximately 3,200 CFM (typical air cooled) (2400 CFM - water cooled) Approximately 4,800 CFM (maximum air cooled)
Maximum altitude	10,000 ft. (3,049 m) operating 40,000 ft. (12,195 m) non-operating
Acoustical noise level (sound power)	Approximately 72 dBA (at front of system) - 82 dBA (at system rear)

Ethernet Port Specification

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.

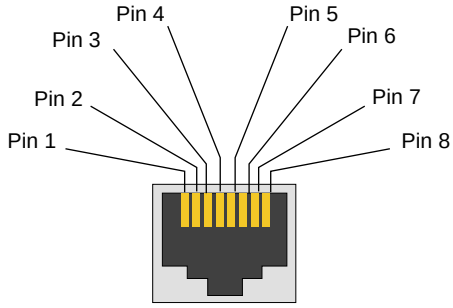

Figure A-1 Ethernet Port

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

Safety Information and Regulatory Specifications

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

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:

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 német pas de perturbations radioélectriques dépassant les normes applicables aux appareils numériques de Classe A préscrites dans le Règlement sur les interferences 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. Evitter 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 vedåbning se ikke ind i strålen

Warning: Varo! Lavattaessa Olet Alttina Lasersåteilylle

Warning: Varning: Laserstrålning når denna del år 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. Anvãnd samma batterityp eller en ekvivalent typ som rekommenderas av apparattillverkaren. Kassera anvãnt batteri enligt fabrikantens instruktion.**

Warning: **Varoitus: Päristo 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

O

P

R

S

T