SGI® ICE™ XA System Hardware User Guide

Record of Revision

Version	Description
-001	May, 2015 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™) XA 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 XA 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 service technician should perform all the setup, addition, or replacement of parts, cabling, and service of your SGI ICE XA 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 service nodes in the air-cooled D-rack.

- 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 XA 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 XA series of computer systems:

- *SGI Rackable C1104-GP2 and C1110-GP2 System User Guide,* (P/N 007-6388-00*x*)

This guide discusses the use, maintenance and operation of the 1U server primarily used as the system's rack leader controller (RLC) server node. This stand-alone 1U compute node is also used as the default administrative server on the ICE XA system. It may also be ordered configured as a login, or batch server, or other type of support server used with the ICE XA series of computer systems. The C1104-GP2 variation uses four 3.5-inch internal drives and the C1110-GP2 uses ten internal 2.5-inch drives.

- *SGI Rackable C1104-GP1 System User Guide,* (P/N 007-6364-00*x*)

This user guide covers an overview of the installation, architecture, general operation, and descriptions of the major components in the SGI Rackable C1104-GP1 server. It also provides basic troubleshooting and maintenance information, and important safety and regulatory specifications. This 1U server is used only as an optional service node for login, batch, MDS or other service node purposes. This server is **not** used as a system RLC or administrative server.

- *SGI Rackable C2112-GP2 System User Guide* (P/N 007-6362-00*x*)

This guide covers general operation, installation, configuration, and servicing of the 2U Rackable C2112-GP2 server node used in the SGI ICE XA system. The 2U server can be used as a service node for login, batch, I/O gateway, MDS, or other service node purposes.

- *SGI UV 30 System User Guide,* (P/N 007-6419-00*x*)

This user guide covers general operation, configuration, and troubleshooting. Also included is a description of major components of the optional 2U-high SGI UV 30 four-socket server node unit used in SGI ICE XA systems. The UV 30 server 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 Installation and Configuration Guide for Clusters*, (P/N 007-6359-00*x*)

This guide discusses software installation and system configuration operations used with the SGI ICE XA series servers. The management center software is also used to provision other non-ICE clusters or other SGI systems.

- *SGI Management Center Administration Guide for Clusters*, (P/N 007-6358-00x)

This document is intended for people who manage and administer the operation of SGI ICE XA systems. The management center software is also used to administer other non-ICE SGI clusters or systems.

- 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 distribution media.

- 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 UNIXA "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 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.

SGI values your comments and will respond to them promptly.

Operation Procedures

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

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

Precautions

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

- "ESD Precaution" on page 1
- "Safety Precautions" on page 2

ESD Precaution

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

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

Safety Precautions

Warning: Before operating or servicing any part of this product, read the "Safety Information" on page 69.

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.
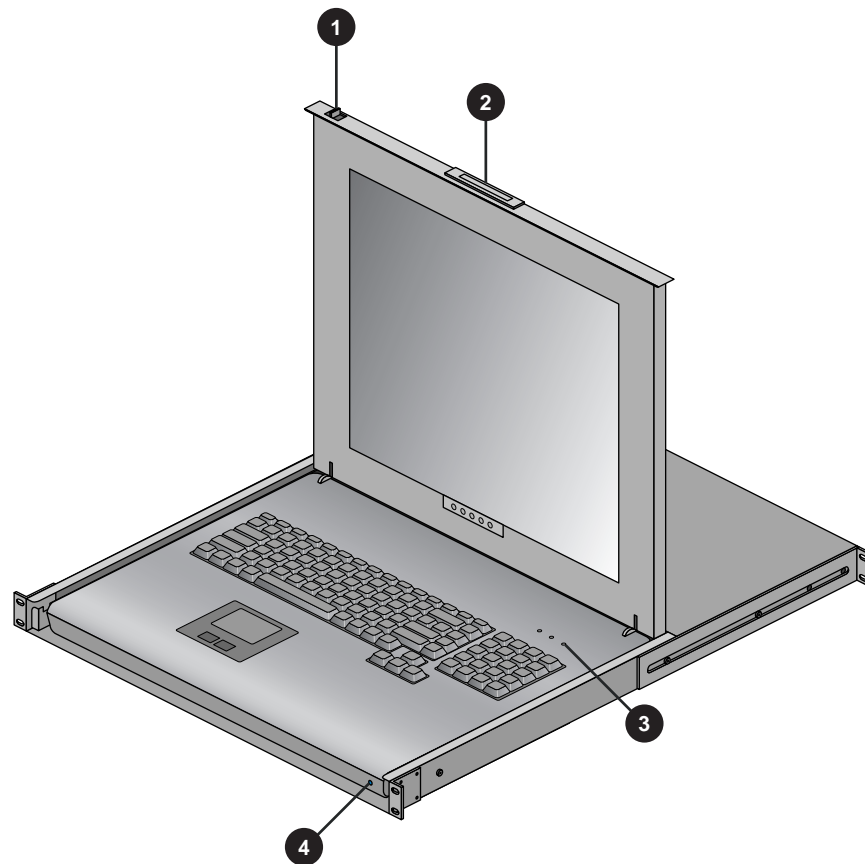
Figure 1-1   Flat Panel Rackmount Console Option

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.

The rackmounted console option is always installed in the air-cooled rack used for administrative, RLC and service node units.

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. 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 (VGA) and USB connectors on the back of the system (see Figure 1-2).
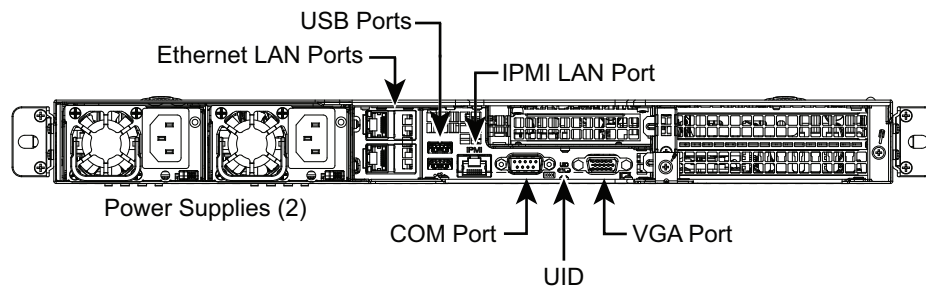
Figure 1-2   Administrative Controller Video Console Connection Points

Powering the System On and Off

This section explains how to power on and power off individual rack units, or your entire SGI ICE XA 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 Administration Guide for Clusters*, (P/N 007-6358-00x).

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. Setting the circuit breakers on the PDUs to the "On" position will apply power to the blade enclosure supplies and will start the chassis manager board in each enclosure.

   Tip: The chassis manager in each blade enclosure will 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 if you want to remove all power from the enclosure.

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 of I/O-rack PDUs in Figure 1-3 on page 6, and Figure 1-4 on page 7) are turned on to provide power when the system is booted up.

Power
distribution
unit (PDU)
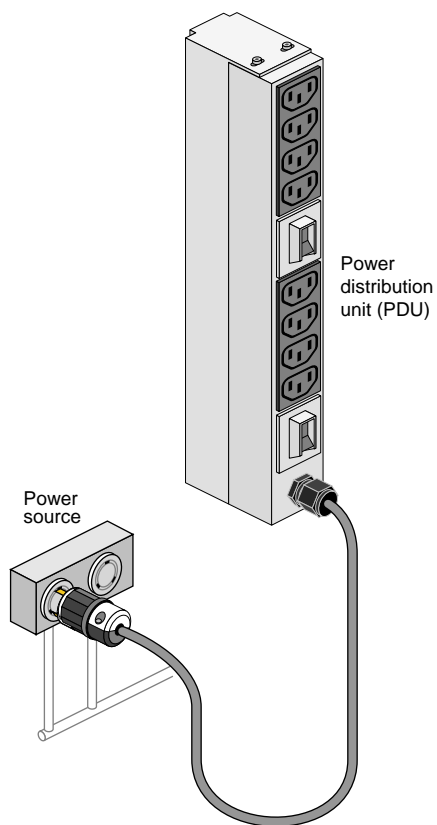
Power
source

Figure 1-3   Eight-Outlet Single-Phase PDU Example

Figure 1-4 on page 7 shows an example of the three-phase PDU.

Figure 1-4   Three-Phase PDU Example

Powering On and Off

The power-on and off procedure varies with your system setup. See the *SGI Management Center Administration Guide for Clusters*, (P/N 007-6358-00*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, you should be able to execute these commands.

Console Management Power (cpower) Commands

This section provides an overview of the console management power (cpower) commands for the SGI ICE XA 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.

Table 1-1 (continued)   cpower option descriptions

Argument	Description
`-n --noexec`	Display but do not execute commands that affect power.
Target_type	
`--node`	Apply the action to a node or nodes. Nodes can be blade compute nodes (inside a blade enclosure), admin 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`	Halts and then powers-off the target(s).
`-h, --help`	Display usage and help text.
`--shutdown`	Shuts down the target but does not power it off. Waits for target(s) to shut down.

Important: The target must always be specified except when the --system option is used. Wildcards may be used, but if wildcard use affects any RLC, the command fails with an error. Accidental resets of the RLC could make the rack unreachable.

Table 1-2   cpower example command strings

Command	Status/result

`# cpower node system up`
`or`
`# cpower node up r*i*n*`
	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 (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 --halt --system`
`or`
`# cpower --halt r*i*n*`
	Completely powers down *every* node in the system except the admin. If you are administering the system remotely through the RLC it may become unreachable (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 --reboot --system`	Reboots all rack leaders and nodes in a system.
`# 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 compute nodes in parallel located in rack 1. Use the **--noleader** argument to the command string if you want the RLC to remain on.
`# cpower node off r2i*n*`	This command example issues an ipmi tool power-off command to all of the nodes in rack 2.

See the *SGI Management Center Administration Guide for Clusters*, (P/N 007-6358-00*x*) for more information on cpower commands and related ipmi style commands.

See the section "System Power Status" on page 17 in this manual for additional related console information.

Monitoring Your Server

You can monitor your SGI ICE XA server from the following sources:

- An optional flat panel rackmounted monitor with keyboard/mouse can be connected to the administration server node for basic monitoring and administration of the SGI ICE XA system. See the section "Console Connections" on page 3 for more information.

- 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 *SGI Management Center Administration Guide for Clusters*, (P/N 007-6358-00*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 XA system. You can also use these consoles to input commands to manage and monitor your system. See the section "System Power Status" on page 17, for additional information.

Chassis Management Monitoring and Control

There is one chassis management controller board (CMC) slot in each blade enclosure. The slot is located directly below the four network switch blade slots of each blade enclosure in the rear of the rack. The CMC board supports:

- Powering up and down of the compute blades

- Performance of environmental control and monitoring of the IRU that it is located in

- Monitoring and communication with switch blades and IRU power supplies

The CMC controls master power to the compute blades under the direction of the RLC. The CMC is powered on when there is power applied to the blade enclosure.

The physical CMC board (and its status indicators) is not visible from outside the rack.

PCIe Subsystem Control

Each compute blade has a variable capacity to use a combination of PCIe cards and/or SATA disks.

Each mezzanine card is connected to the compute node via one x16 Gen3 PCIe channel. PCIe configurations on blades must be ordered from the factory or installed by trained service personnel.

Additional 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 most easily configurable I/O system interfaces for the SGI ICE XA systems. See the particular server's user guide for detailed information on installing optional I/O cards or other components.

System Management

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

- "Levels of System and Chassis Control" on page 15
- "Chassis Manager Interconnects" on page 16
- "System Power Status" on page 17

One chassis management controller (CMC) is used in each blade enclosure. The CMC is located directly below the enclosure's switch blade(s). 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. The CMC takes direction from the RLC and stays powered-on as long as there is power applied to the blade enclosure.

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

Mass storage enclosures are not managed by the SGI ICE XA system controller network.

Figure 2-1 shows an example remote LAN-connected console communication path to an SGI (D-rack) containing the administration, RLC and other service nodes for the SGI ICE XA series system. The E-Cell compute and cooling rack assembly is not shown in this diagram.
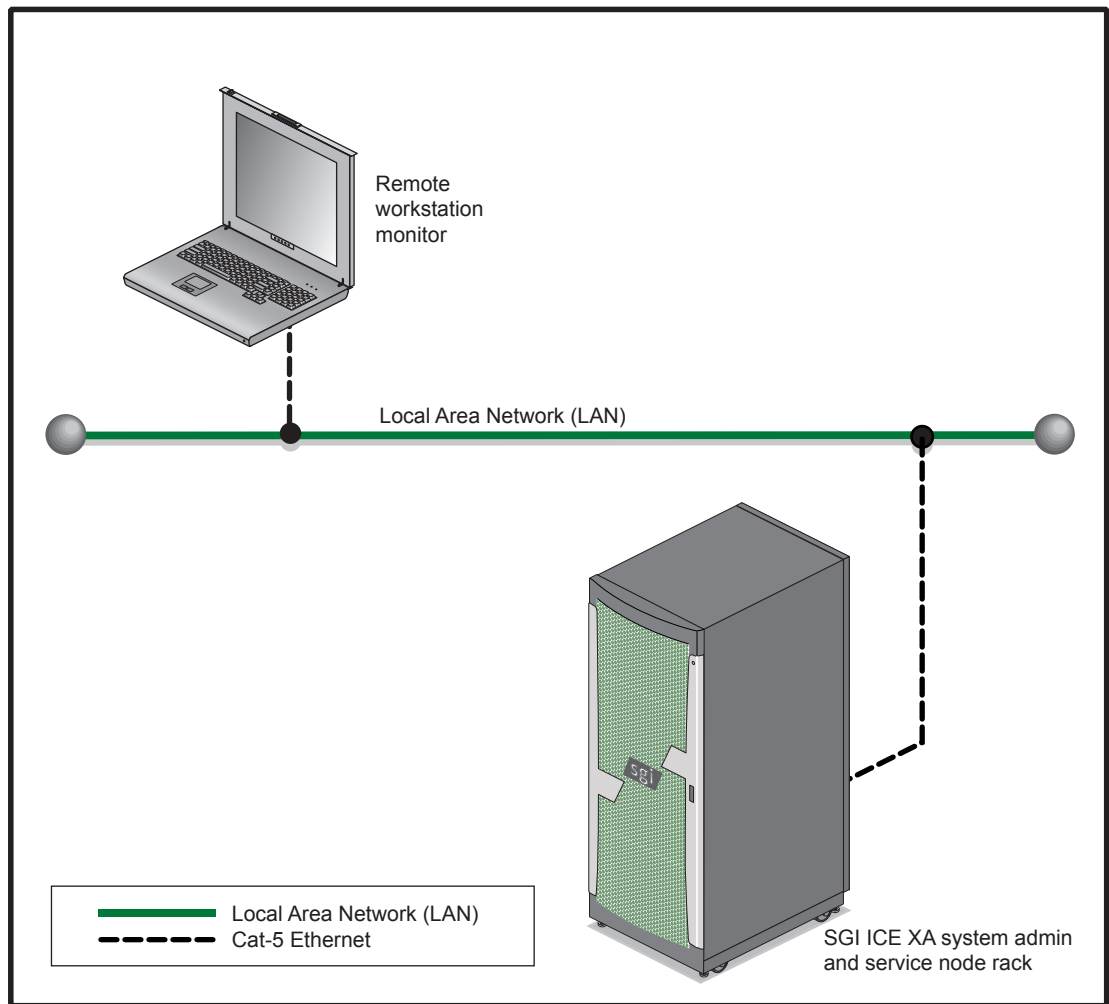
Figure 2-1   SGI ICE XA System Admin and Service Node Network Access Example

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 USB 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 XA 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 XA 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 CMCs.

Chassis Controller Interaction

In all SGI ICE XA 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 2-2 on page 16.

- 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. For an overview of how all the primary system components interact within the Ethernet network infrastructure, see the section "System Hierarchy" in Chapter 5.

Chassis Manager Interconnects

The chassis manager in each blade enclosure connects to the system administration, rack leader and service node servers via Gigabit Ethernet switches. See the example in Figure 2-2.

Use of a CDU is required in all E-Cell configurations as of the time of publication of this document. The VLAN3 interface allows the CMCs to monitor and adjust the activity of the cooling racks in an E-Cell as well as the external cooling distribution unit (CDU). The CDU rack supplies liquid cooling to the individual blades within the E-racks.
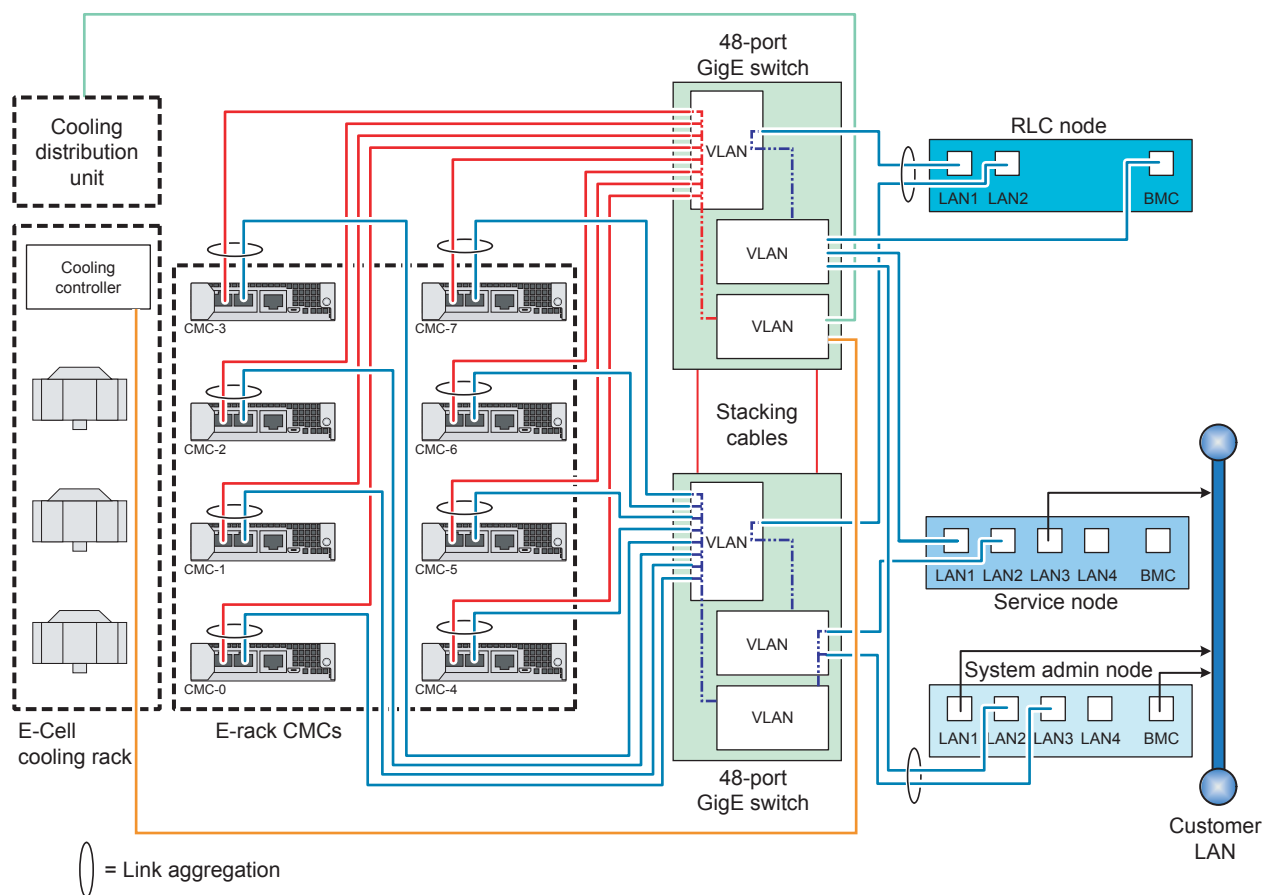
Figure 2-2   E-rack System Chassis Management Diagram (With CDU)

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

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 rack. This is a legacy identification from the ICE "individual rack unit" nomenclature.

Note that system-level power-up commands are applied first to service nodes, then to RLCs, then to blade enclosures and compute blades.

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 (blade enclosure) 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 (blade enclosure) 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 Administration Guide for Clusters*, (P/N 007-6358-00x) 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) XA series system. The major components of the SGI ICE XA 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 XA series system in your lab or server room.

Each 42U SGI E-Cell rack holds up to four 10.5U high blade enclosures. Each blade enclosure provides power, system control, and the network fabric for nine compute blades using a non-blocking backplane. The ICE XA E-Cell rack assembly consists of two custom designed 42U high racks paired with a dedicated cooling rack (between the two compute racks). The cooling rack circulates air and removes heat from the system and passes it to water-cooled extractors that carry away the heated air from the assembly. Each blade enclosure also has an internal InfiniBand communication backplane. Each of the (up to) nine compute blades supported in an enclosure are equipped with up to four two-socket nodes, with ASICs, memory components and I/O chip sets mounted on them. The blades slide directly in and out of the enclosures. Every processor socket on a blade supports dual-inline memory module (DIMM) memory units using DDR4 bandwidths. Optional PCIe slots, hard disk or solid-state (SSD) drives and MIC or GPU option boards are available.

Each compute blade supports two or four individual node boards. Note that a maximum system size of 72 compute blades per E-Cell rack assembly (9 blades x 8 enclosures) is supported at the time this document was published. Non-blade enclosures such as the system administration node, rack leader controller (RLC), and optional service nodes or storage units are housed in 42U D-rack configurations. 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.

Internal IB switch ASICs located in switch blades in the blade enclosure eliminate the requirement for external IB switches when deployed with hypercube, enhanced hypercube or all-to-all topologies. The InfiniBand technology provides for fast communication between compute blades

within the same chassis and also between compute blades in separate enclosures or separate rack assemblies.

The SGI ICE XA series systems can run parallel programs using a message passing tool like the Message Passing Interface (MPI). The SGI ICE XA 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 an SGI ICE XA server rack assembly.
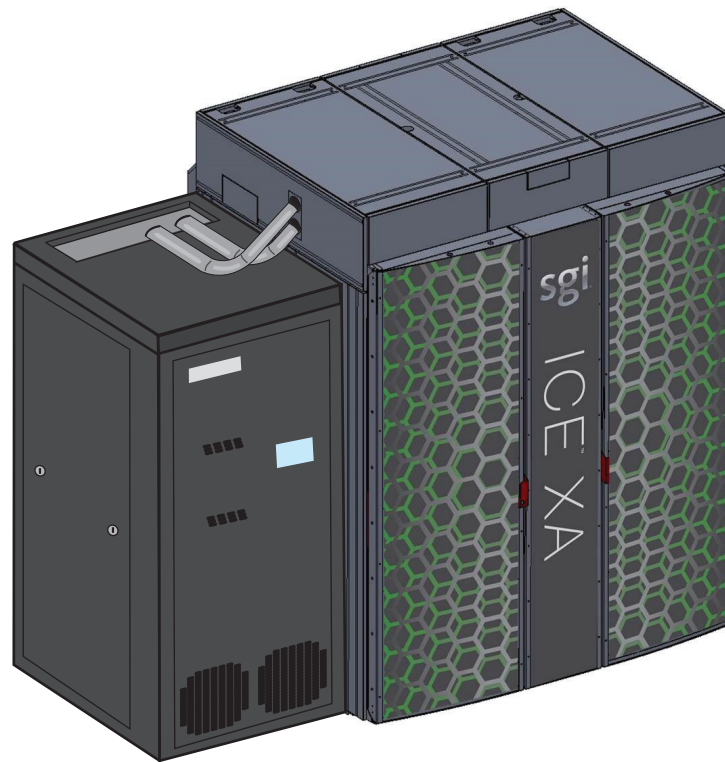
Figure 3-1   SGI ICE XA Series System E-Cell Assembly Example

The compute rack enclosures within the SGI ICE XA system are 10.5U 9-blade units. Each enclosure supports between two and nine compute blade assemblies, up to nine power supplies, one chassis management controller (CMC) and up to four InfiniBand based I/O fabric switch interface blades. Note that internal IB switch ASICs located in switch blades in the blade enclosure eliminate the requirement for external IB switches. The InfiniBand technology provides for fast communication between compute blades within the same chassis and also between compute blades in separate enclosures.

Optional water chilled D-rack cooling is available for support nodes and storage. Note that ICE XA blade enclosures require liquid-cooling and must reside in E-Cell racks and always require water cooling systems to operate. See the section Chapter 4, "Rack Information" for information on water-cooled ICE XA E-Cell systems.

SGI ICE XA systems require a minimum of one 42U tall D-rack with PDUs installed to support standard and optional servers or storage units used with the ICE XA system.

Figure 3-2 shows a diagram of the D-rack and typical support hardware that would be installed in the rack. The rack would include single-phase PDUs or optional three-phase PDUs with enough outlets to support all the installed support hardware. 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.

42U High Rack

Service node
Admin server
Rack leader controller
1U Gig-E switch
1U console kbd/monitor
Optional service node

Optional storage units
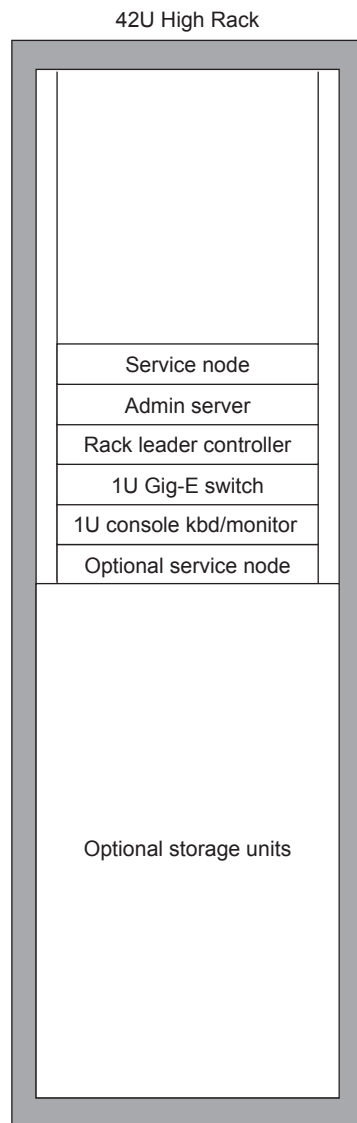
Figure 3-2   D-rack Standard and Optional Components Example

SGI ICE XA System and Blade Architectures

The SGI ICE XA series of computer systems are based on an FDR InfiniBand I/O fabric. This concept is supported and enhanced by using the SGI ICE XA 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 different InfiniBand host-channel adapter (HCA) cards, see "IP125 Blade Architecture Overview" for an example.

IP125 Blade Architecture Overview

Each IP125 blade contains four two-socket compute nodes. The IP125 blade architecture is described in more detail in the following paragraphs.

The compute blade contains the processors, memory, and four of the following fourteen-data rate (FDR) InfiniBand embedded HCAs:

- Single-port IB HCA

- Dual-port IB HCA

Each of the two physical boards within the IP125 blade is divided into two (logically independent) compute nodes equipped with two processor sockets and eight DIMM slots. Features of the IP125 blade include:

- Eight Intel processors total within each blade

- The eight processors are cooled with liquid "Cold Sink" technology

- Each processor slot supports four DDR4 memory RDIMMs at up to 2133 MT/s (thirty-two memory DIMM slots total per IP125 blade)

- Dual or single port IB mezzanine cards are supported (but not mixed) within the blade

- Four board management controllers (BMCs) - one per logical node

- The blade can be fitted with optional expansion kits that support:

    – Up to eight SATA 2.5-inch SSD or HDD (scratch/swap) drives - (note that you cannot mix SSD and HDD drive types on the same blade)

    – Up to eight low-profile PCIe cards (or four cards and four drives)

    – One or two liquid-cooled accelerator cards

The two processors on each node in an IP125 blade maintain an interactive communication link using the Intel QuickPath Interconnect (QPI) technology. This high-speed interconnect technology provides data transfers between the on-board processors. See the section "QuickPath Interconnect Features" on page 26 for an overview of the link functionality and bandwidth capability.

The IP125 compute blade cannot be plugged into and cannot be used in "previous generation" SGI ICE X, Altix ICE 8200 or 8400 series blade enclosures. Certain types of multi-generational system interconnects can be made through the InfiniBand fabric level. Check with your SGI service or sales representative for additional information on these topics.

IP139 Blade Architecture Overview

The IP139 compute blade contains two 2-socket nodes. The IP139 blade architecture is described in more detail in the following paragraphs.

The compute blade contains the processors, memory, and uses two of the following fourteen-data rate (FDR) InfiniBand embedded HCAs:

- Single-port IB HCA mezzanine cards (one per dual-socket node board)

Each of the two physical boards within the IP139 blade is a (logically independent) two-processor compute node equipped with eight DIMM slots. Features of the IP139 blade include:

- Four Intel processors total within each blade

- The four processors are cooled with liquid "Cold Sink" technology

- Each processor socket supports four DDR4 memory RDIMMs at up to 2133 MT/s (sixteen memory DIMM slots total per IP139 blade)

- Two dual-socket node boards are always installed in each IP139 blade

- Each blade has two baseboard management controllers (BMCs) - one BMC per node

- The blade can be fitted with optional expansion kits that support:

  – Up to eight SATA 2.5-inch SSD or HDD (scratch/swap) drives - (note that you cannot mix SSD and HDD drive types on the same blade)

  – Up to four full-height full-width PCIe cards

  – Up to four liquid-cooled accelerator cards

The processors on each node board in the IP139 blade maintain an interactive communication link using the Intel QuickPath Interconnect (QPI) technology. This high-speed interconnect technology provides data transfers between the on-board processors. See the section "QuickPath Interconnect Features" for an overview of the link functionality and bandwidth capability.

The IP139 compute blade cannot be plugged into and cannot be used in "previous generation" SGI ICE X, Altix ICE 8200 or 8400 series blade enclosures. Usage of ICE XA blade enclosures in non SGI ICE XA racks may be restricted/unsupported due to thermal requirements.

Multi-generational system interconnects can be made through the InfiniBand fabric level. Check with your SGI service or sales representative for additional information on these topics.

QuickPath Interconnect Features

Each processor socket on an ICE XA system node board is interconnected using two QuickPath Interconnect (QPI) links. Each QPI link consists of two point-to-point 20-bit channels - one send channel and one receive channel. Each QPI channel is capable of sending and receiving at the same time – on the same 20 bit channel. The QPI link has a theoretical maximum aggregate bandwidth of 25.6 GB/s using a 3.2 GHz clock rate and 38.4 GB/s using a 4.8 GHz clock rate. Each blade's I/O chip set supports one or two processor sockets.

QPI Bandwidth Overview

The maximum bandwidth of a single QPI link used in each blade node board 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 Gbits per second.

- Convert to bytes: 102.4 divided by 8 equals 12.8 GB/s (max single direction bandwidth)

- The total aggregate bandwidth of the QPI channel is 25.6 GB/s: (12.8 GB/s x 2 channels)

Blade Memory Features

The memory control circuitry in each Intel processor is integrated into the chip and provides greater memory bandwidth and capacity than previous generations of ICE compute blades. Note that all ICE XA blades use DDR4 DIMM technology.

Blade DIMM Memory Features

The IP125 compute blade uses four logically independent dual-processor node boards and each node board supports a maximum of eight DDR4 memory DIMMs. The IP125 compute blade supports a maximum of 32 DDR4 RDIMMs. Memory increments are in groups of four DIMMs per processor socket.

Each Intel processor in an ICE XA IP139 blade uses four DDR4 memory channels with one or more memory DIMMs on each channel (depending on the configuration selected). Each of these blades uses four processors and up to 16 DIMMs. Each 64-bit DDR4 memory channel supports one memory DIMM.

DIMM Bandwidth Factors

Note the following factors regarding processors and DIMMs on a blade:

- Different processor SKUs may support different maximum DIMM speeds, such as 1866, or 2133.

- Each DIMM has a maximum operating frequency or speed regardless of the processor it is connected to or the number of companion DIMMs.

- Each of the DIMMs on a blade's node board must be the same capacity and functional speed. Using DIMMs with different speeds and/or memory capacities to support the same processor is not recommended.

A minimum of one dual-inline-memory module (DIMM) is required for each processor on a node board; four DIMMs per processor are recommended. When possible, it is generally recommended that all blade node boards within an enclosure use the same number and capacity (size) DIMMs.

Each blade in the ICE XA enclosure 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 an impact on compute "load balancing" within the system.

Memory Channel Recommendation

It is highly recommended (though not required) that each processor on a system node board 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.

System InfiniBand Switch Blades

One, two or four fourteen-data-rate (FDR) InfiniBand switch blades can be used with each blade enclosure pair configured in the SGI ICE XA system. The switch blades plug into the rear of the blade enclosure, which reduces latency and improves air flow. The switch blades provide the interface between compute blades within the same blade enclosure and also between compute blades in separate blade enclosures.

The system supports a standard 36-port and premium 60-port FDR switch and future versions will be capable of supporting a standard 36-port and premium 60-port EDR switch.

- The Standard switch has a single 36-port ASIC with 18 ports connecting to compute nodes and 18 ports for connecting to external targets.

- The Premium switch has two 36-port ASICS with nine ports from each ASIC connecting to compute nodes and 18 ports from each ASIC for connecting to external targets.

The difference is in connectivity. The premium switch has twice the external ports to use for expanding the topology and creating larger configurations with a higher-bandwidth interconnect fabric.

Note that All-to-All and Hypercube topologies do not require external InfiniBand switches, but are built by interconnecting the internal InfiniBand switch ASICs (switch blades) in the blade enclosure. The single-switch ASIC and dual-switch ASIC switch blades for each enclosure 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 InfiniBand topology. A blade-enclosure using dual-node blades must use four switch blades to support a single-plane topology. Check with your SGI sales or service representative for additional information on availability. 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.

System Features and Major Components

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

- "Modularity and Scalability" on page 29
- "Reliability, Availability, and Serviceability (RAS)" on page 36

Modularity and Scalability

The SGI ICE XA series systems are modular, blade-based, scalable, high-density cluster systems. The system rack components are primarily housed in building blocks referred to as blade enclosures. Each enclosure consists of a sheetmetal housing with internal IB backplanes and nine (shared) power supplies.

However, other "free-standing" SGI compute servers are used to administer, access and service the SGI ICE XA 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 blade enclosure has redundant power supplies. A cooling distribution unit (CDU) rack expands an ICE XA rack's heat dissipation capability for the blade enclosure components without requiring lower ambient temperatures in the lab or server room.

A number of free-standing (non-blade) compute and I/O servers (also referred to as service nodes) are used with SGI ICE XA series systems in addition to the standard 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 XA system will have one system administration controller, at least one rack leader controller (RLC) and at least one service node. All ICE XA systems require one RLC for every eight CMCs in the system. Figure 3-3 shows an overview of the SGI ICE XA system management and component network interaction.
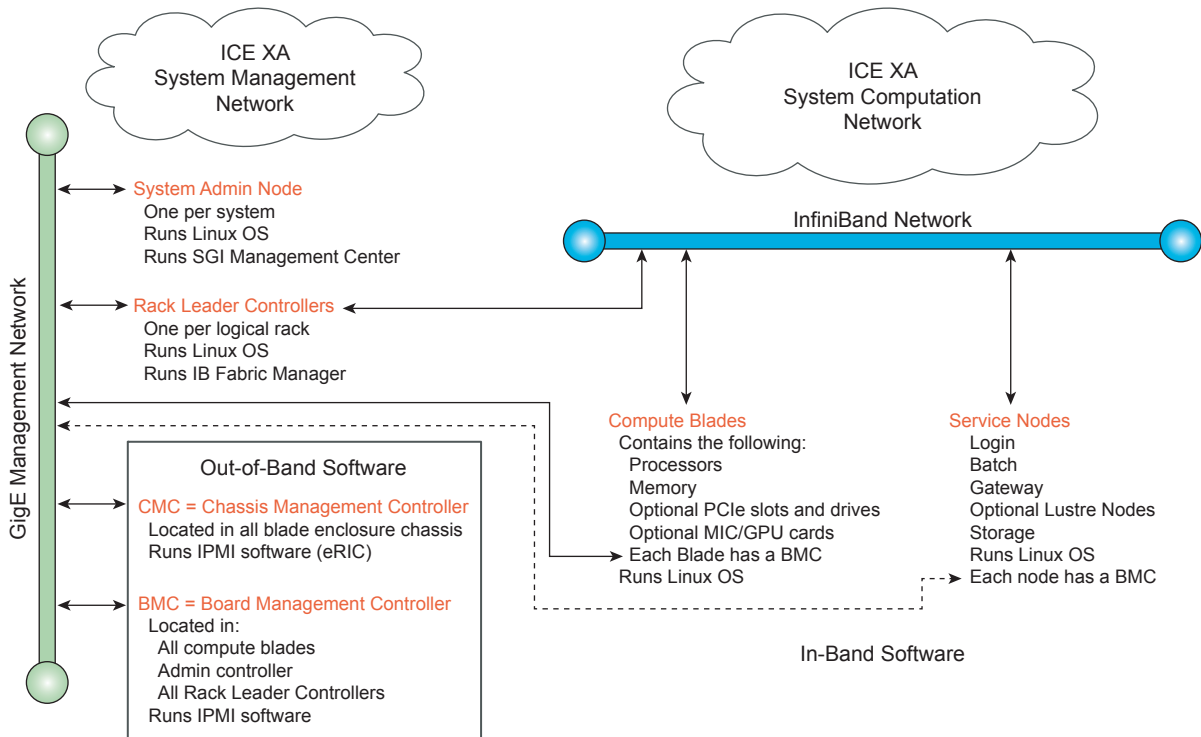
Figure 3-3 SGI ICE XA System and Network Components Overview

The administration server and the RLCs are integrated stand-alone 1U servers. The service nodes are integrated stand-alone non-blade 1U or 2U servers. The following subsections further define the free-standing server unit functions described in the previous list.

System Administration Server

There is a minimum of 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). Note that a high-availability administration server configuration is available that doubles the number of

administrative servers used in a system. In high-availability (HA) administration server configurations, two servers are paired together. The primary admin server is backed up by an identical "backup" admin server. The second (backup) server runs the same system management image as the primary server.

The server is used to install SGI ICE XA 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 XA systems is connected to the external network. All ICE XA systems are configured with dedicated "login" servers for multiple access accounts. You can 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 all the blade enclosures within the logical rack that it monitors. The rack leader then consolidates and forwards data requests received from the blade enclosures to the administration server. A rack leader controller also supplies boot and root file sharing images to the compute nodes in the enclosures.

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 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 XA 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 (SMC) Administration Guide for Clusters*, (P/N 007-6358-00*x*).

Note also that the scale of the CMC drawings in Figure 3-4 is adjusted to clarify the interconnect.
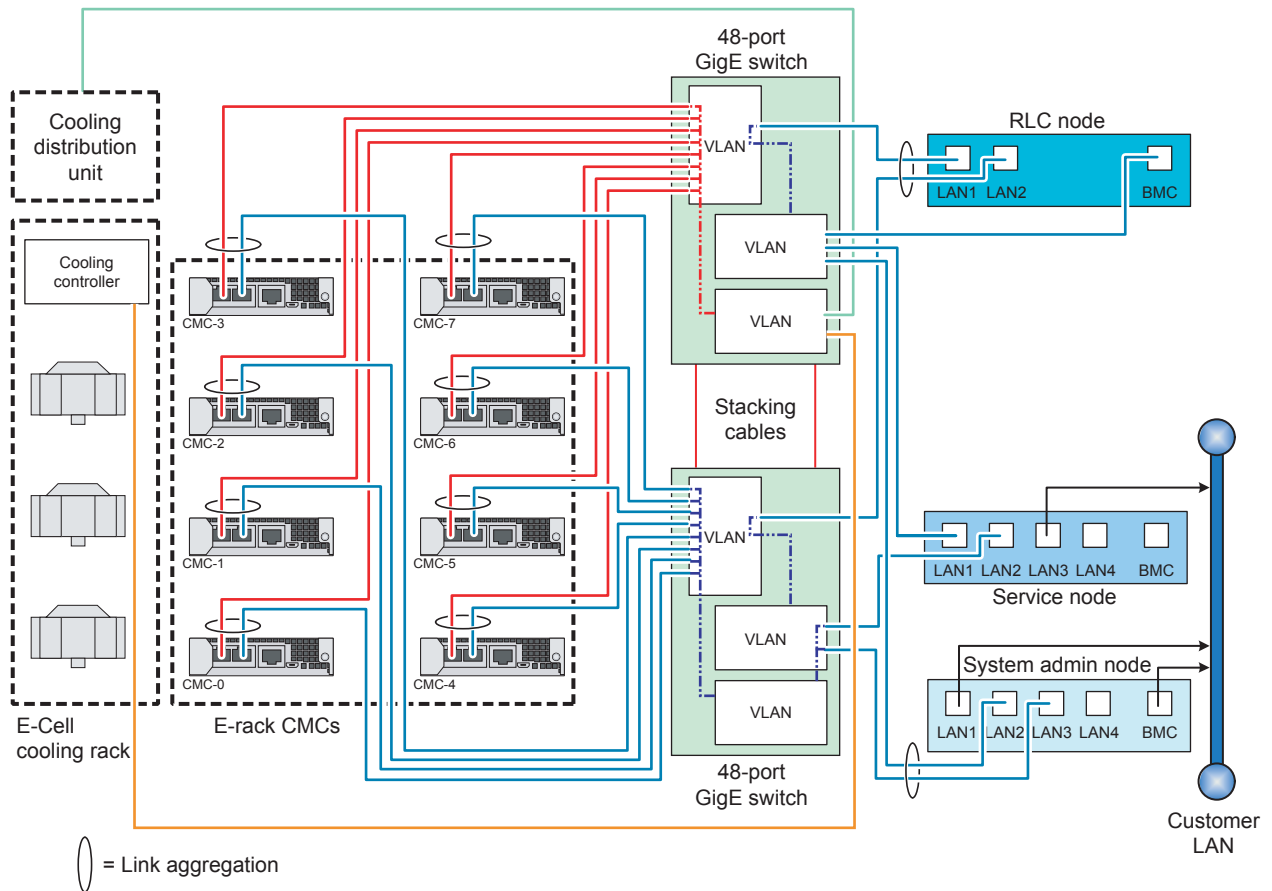
Figure 3-4   Administration and RLC Cabling to Chassis Managers

The RLC as Fabric Manager

In some SGI ICE XA 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 system's administration 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 recommended on 8-rack or larger systems.
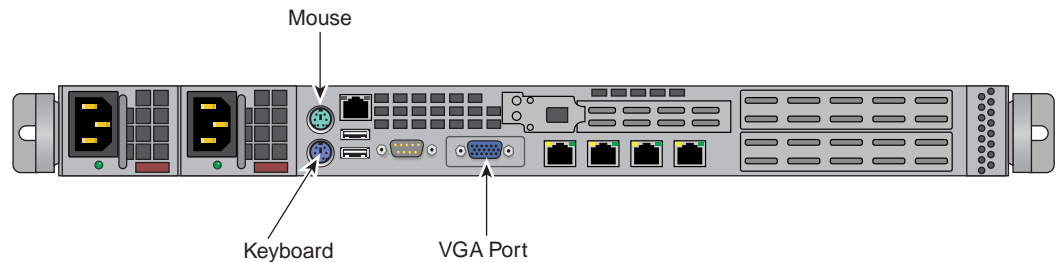
Figure 3-5   Example Rear View of a 1U Service Node

Login Server Function

The login server function within the ICE XA 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 XA 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 front and rear connectors and interface slots on a 2U service node.
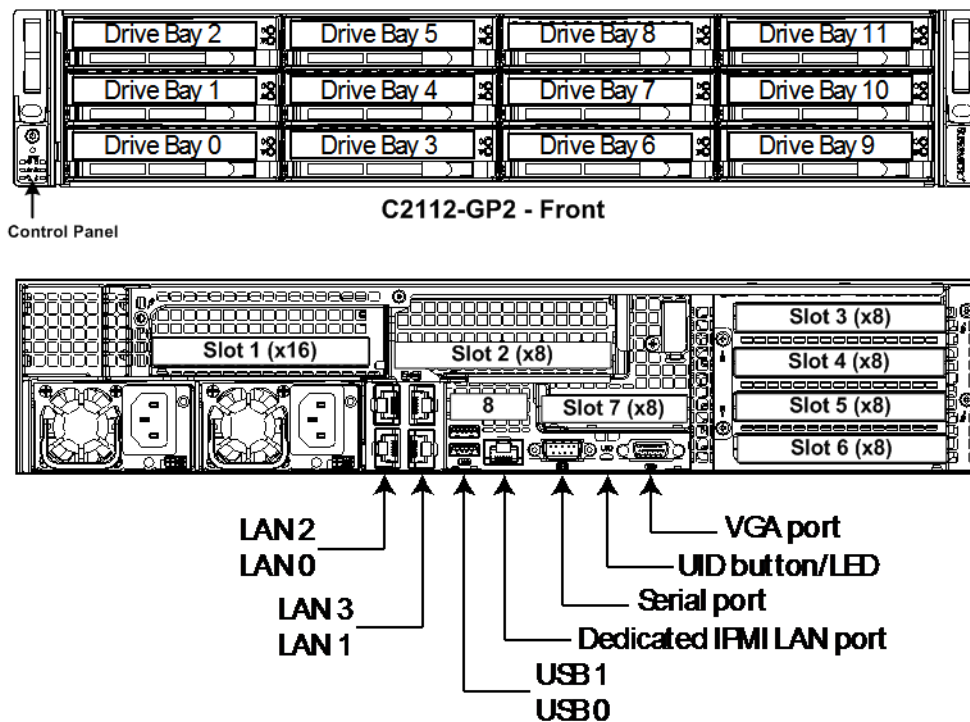
Figure 3-6   2U Service Node Front and Rear Panel Example

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 hosted on an optional 1U or 2U stand-alone server within the ICE XA system.

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.

Optional Lustre Nodes Overview

The nodes in the following subsections are used when the SGI ICE XA system is set up as a Lustre file system configuration. In SGI ICE XA installations the MDS and OSS functions are generally on separate nodes within the ICE XA 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 XA 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 XA 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 XA 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 XA 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 are redundant and can be hot-swapped under most circumstances.

    - A rack-level water chilled cooling option is available for D-racks.

    - Blade enclosures have overcurrent protection at the blade and power supply level.

    - 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.

    - 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.

Optional SGI Remote Solutions (SGI RS)

The optional SGI RS system automatically detects system conditions that indicate potential future problems and then notifies the appropriate personnel. This enables you and SGI global support teams to pro-actively support systems and resolve issues before they develop into actual failures.

SGI Remote Services provides a secure connection to SGI Customer Support - on demand. This can ensure business continuance with SGI systems management and optimization.

SGI Remote Solutions Primary Capabilities

- 24x7 remote monitoring and data gathering of SGI customer systems

- Alerts and notification on changes, failures and potential failures

- Log files immediately available

- Configuration fingerprint

- Secure file transfer

- Optional secure remote access to customer systems

SGI Remote Solutions Benefits

- Improved uptime and system availability

- Proactive identification of issues before they create an outage

- Increase system stability by monitoring hardware and software version compatibility

- Reduced time to resolve support cases

- Greater operational efficiency

- Less involvement of customer staff during troubleshooting

- Faster support case resolution

- Improved productivity

Proactive potential problem identification can result in higher system availability.

Automated Alerts and, in some instances, Case Opening results in faster problem resolution time and less direct involvement required by Customer Support Teams. SGI Remote Services is available for all currently shipping SGI ICE, UV and Rackable systems and also other specific SGI systems. Check with your SGI sales or service representative for more details.

SGI Remote Service Operations Overview

An SGI Support Services Software Agent runs on each SGI system at your location, enabling remote system monitoring and secure communication to SGI Support staff. Your basic hardware and software configuration as well as system health information is captured and stored in the Cloud. Figure 3-7 shows an example visual overview of the monitoring and response process.

Cloud intelligence automatically reviews select Event Logs around the clock (every five minutes) to identify potential failure information. If the Cloud intelligence detects a critical Event, it notifies SGI support personnel.

This monitoring requires no changes to customer systems or firewalls as long as the SGI Agent can send HTTPS messages to highly secure Cloud and Global Access Servers. It will also have no impact on customer network or system performance. All communication between SGI global support and customer systems is kept secure using Secure Socket Layer (SSL) encryption. All communication with SGI is initiated from the customer site using HTTPS protocol on port 443.
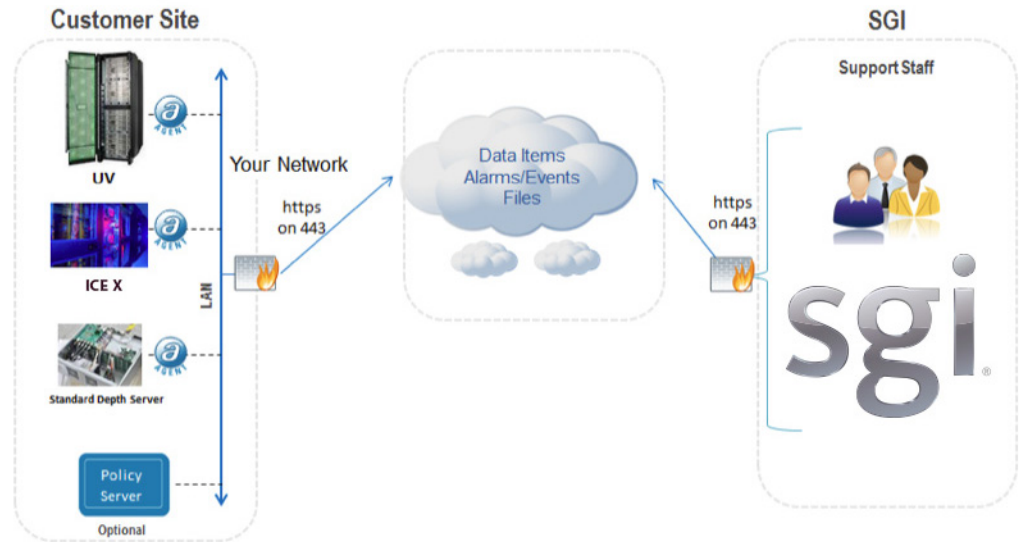
Figure 3-7   SGI Remote Solutions Process Overview

System Components

The SGI ICE XA series system features the following major components:

- **42U E-racks**. These multi-rack (E-Cell) rack assemblies use a dedicated cooling rack for each two compute racks used. Water cooling of the individual nodes is accomplished by a separate dedicated cooling-distribution rack (CDU). See "SGI ICE XA E-Cell Rack Assemblies" in Chapter 4 for additional information.

- **42U D-rack.** This is an SGI rack used for support nodes, optional storage, switches and I/O in the SGI ICE XA series. Note that each ICE XA system must have a dedicated I/O rack holding GigE switches, RLCs, Admin servers and additional service nodes. Water-cooled D-racks are optionally available.

- **Blade enclosure.** This sheetmetal enclosure contains the enclosures holding up to nine compute blades, one chassis manager board, up to four InfiniBand fabric I/O blades and front-access power supplies for the SGI ICE XA series computers. The enclosure is 10.5U high. The blade enclosure employs front-mounted power supplies.

- **Compute blade.** Holds up to four nodes and up to 24 memory DIMMs (depending on blade type).

- **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.** This server node supports an optional console and administrative software.

- **1U Service node.** Additional 1U server(s) can be added to a system support 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.

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

E-Cell Rack Numbering

Each compute (E-rack) in a multi-rack system is numbered with a three-digit number sequentially beginning with (001). A compute rack (E-rack) contains blade enclosures, optional disks and accelerators, PDUs and the cooling infrastructure that extracts heat.

> **Note:** In an E-Cell compute rack system (viewed from the front), the compute rack number on the left 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.

Note that an air cooled D-rack is used to support administrative and rack leader server nodes, service specific nodes, optional mass storage enclosures and potentially other options.

Optional System Components

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

Overview

At the time this document was published only specific SGI ICE XA racks were approved for ICE XA systems shipped from the SGI factory. See Figure 4-4 on page 52 for an example. Contact your SGI sales or support representative for more information on configuring SGI ICE XA systems in non-ICE XA factory rack enclosures.

SGI ICE XA Series D-Rack (42U)

The SGI tall D-Rack (shown in Figure 4-1 on page 45) 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 46). 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 4-2 on page 46) do not use a lock.

  Up to four optional 10.5 U-high (18.25-inch) water-cooled panels can be installed on the rear of the SGI ICE XA D-Rack.

  Note that an air-cooled rack (that does not use water-cooling panels) has a rear key lock to prevent unauthorized access to the system via the rear door. 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. Stand-alone administrative, leader and login server modules 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 service node 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 sixteen outlets may be required for a single rack as follows:

  - 4 outlets for administration and RLC servers
  - 2 outlets for each service node (minimum of one server)
  - two outlets for each optional storage unit

  Optional single-phase PDUs can be used in SGI ICE XA racks dedicated to service node or I/O functionality.
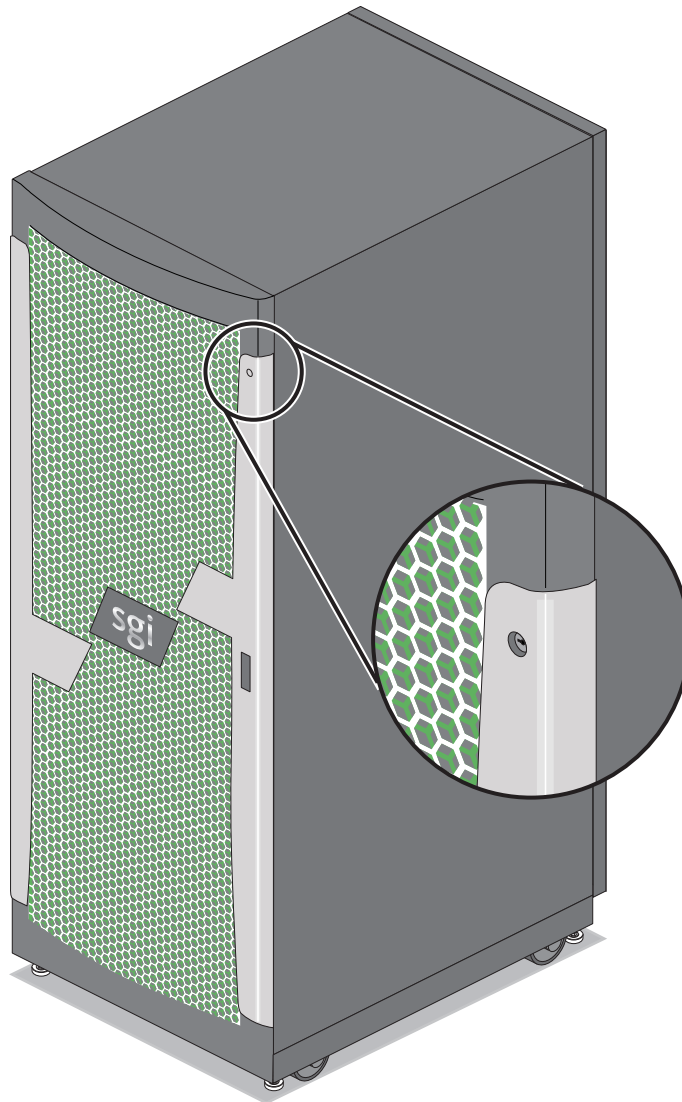
Figure 4-1   D-Rack Front Lock Example
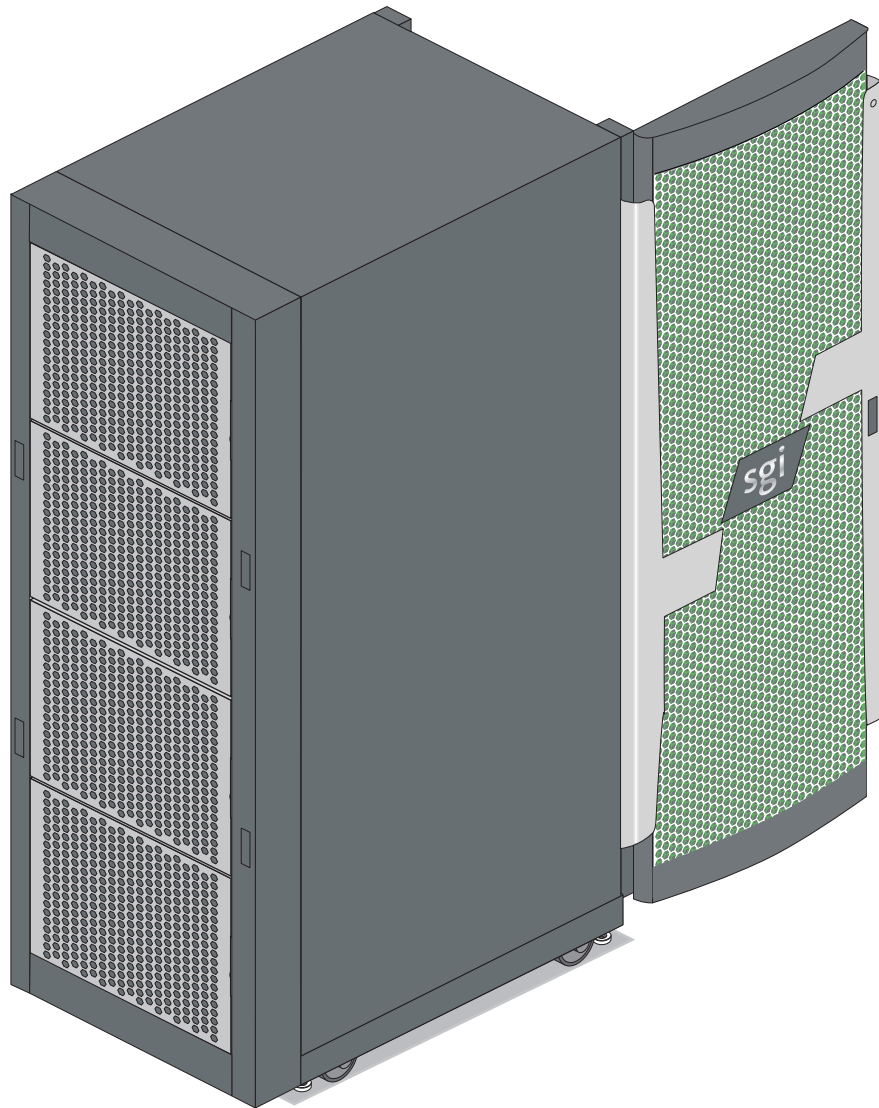
Figure 4-2   Optional Water-Chilled Door Panels on Rear of ICE XA D-Rack

ICE XA D-Rack Technical Specifications

Table 4-1 lists the technical specifications of the SGI ICE XA series D-Rack.

Table 4-1   Tall SGI ICE XA D-Rack Technical Specifications

Characteristic	Specification
Height	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)	~2,650 lbs. (1,205 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	20 ms
Power cable	10 ft. (3.0 m) pluggable cords

Important: The D-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 XA system operating specifications and environmental requirements.

SGI ICE XA E-Cell Rack Assemblies

SGI ICE XA system configurations require the use of enhanced "closed-loop" cooling and the compute rack assemblies are generally referred to as an "E-Cell". A complete E-Cell assembly consists of two compute racks and an air-to-water heat exchanger (the cooling tower) for cooling the compute racks, see Figure 4-3 for an example. The racks are connected together to create a sealed unit to support closed-loop cooling. The E-Cell rack system does not exhaust heated air into the surrounding environment; this means an E-Cell does not add to the heat load of the computer room. Multiple E-Cells can be interconnected and configured to create very large systems. Many E-Cell configurations also require the use of a separate cooling distribution rack unit (CDU rack) (not shown in Figure 4-4 on page 52).
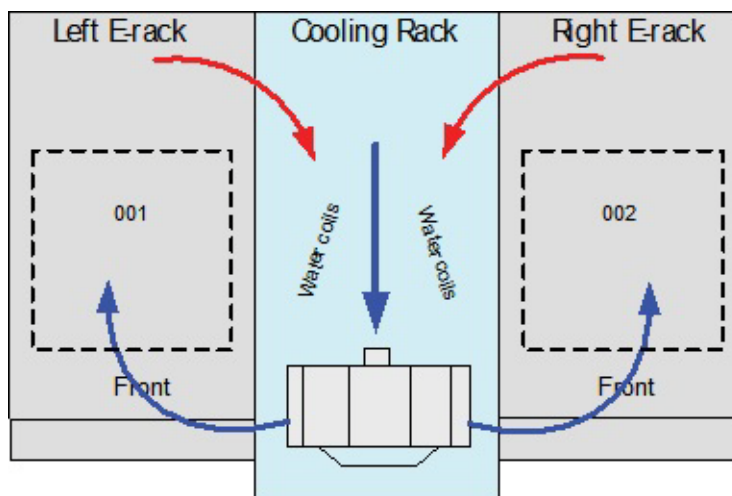
Figure 4-3 E-Cell Rack Configuration Example

The smallest E-Cell assembly consists of two compute racks with a cooling rack in between. E-Cell systems use liquid-cooled compute/graphics/InfiniBand boards and also require the use of a cooling distribution unit CDU (tower). See Figure 4-4 on page 52 for an example E-Cell configuration with single CDU.

E-Cell Functional Overview

An E-Cell consists of two E-racks and a dedicated cooling rack and also a special cooling distribution rack (CDU). The separate CDU supplies water that dissipates heat off the system CPUs, optional graphics accelerator cards and any other liquid-cooled components.

E-Cell Compute Rack Specifications

The compute racks used in an E-Cell configuration are 23.6-inches (60 cm) wide; other size and weight differences are noted in Table 4-2. In each E-cell compute rack there are four shelves for 9-slot blade enclosures. Note that the left and right-side compute racks are internally different because of air-flow requirements.

Table 4-2   SGI ICE XA E-Rack Technical Specifications

Characteristic	Specification
Height	89.63 in. (227.7 cm)
E-Cell compute rack width	23.6 in. (60 cm)
E-Cell Assembly Width	70.9 in. (180 cm) not including trim
Depth	47.25 in. (120 cm)
Empty compute rack weight	482 lbs. (219 kg)
Weight (full)	~2,300 lbs. (1,043 kg) approximate
Shipping weight (max)	~2,350 lbs. (1,066 kg) approximate
Voltage range	North America/International
Nominal	200-240 VAC /230 VAC
Tolerance range	180-264 VAC /180-254 VAC
Frequency	North America/International
Nominal	60 Hz / 50 Hz
Tolerance range	47-63 Hz / 47-63
Phase required	single-phase or 3-phase
Power requirements (max)	79.8 kVA (78.2 kW)
Hold time	20 ms
Power cable	10 ft. (3.0 m) pluggable cords

E-Cell Cooling Rack Specifications

There is one dedicated cooling rack for every two E-racks in an E-Cell assembly. The cooling rack circulates conditioned air through the E-racks to cool the components within the E-rack assembly. Figure 4-3 on page 48 shows the cooling rack at the center of the array. Note that the cooling rack does **not** house any compute or storage components and is strictly for cooling the E-Cell assembly. Table 4-3 lists the major characteristics and specifications of the E-Cell cooling rack.

Table 4-3   SGI ICE XA E-Cell Cooling-Rack Specifications

Characteristic	Specification
Cooling rack height	90 in. (236.2 cm)
E-Cell cooling rack width	23.6 in. (60 cm)
E-Cell cooling rack depth	47.25 in. (120 cm)
Cooling rack weight	Approx. 1,400 lbs. (635 kg)
Voltage range	North America/International
Nominal	220-240 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
Phase required	single-phase or 3-phase
Power requirements (max)	4.13 to 4.44 kW
Power cable	10 ft. (3.0 m) pluggable cords

Cooling Distribution Unit (CDU) Rack Specifications

The cooling distribution unit (CDU) is an interface between the building chilled water system and the processor cold sinks in the E-rack assembly. It is designed to circulate and control chilled water to the water manifolds that are in the E-rack with the compute blades. The CDU is rated for up to 305kW (1040.6 BTU/H) of cooling. The CDU consists of a rack-style cabinet that houses a heat exchanger, circulating pump(s), control valve, receiver, controls, valves, and piping. The CDU monitors room conditions and prevents coil condensation by maintaining the chilled water being pumped to the cooling modules at a temperature above the room's dew point.

All functions, such as switching pumps (if applicable), controlling water temperature, etc., are automatic.

Table 4-4   SGI ICE XA Cooling Distribution Unit (CDU) Specifications

| Characteristic	Specification
CDU height	74.2 in. (188.5 cm)
CDU width	23.6 in. (60 cm)
CDU depth	39.4 in. (100 cm)
CDU weight	727.5 lbs. (330 kg) dry | 846.5 lbs. (384 kg) water filled
CDU shipping weight	Approximately 952.4 lbs. (432 kg)
Voltage range Nominal Tolerance range	North America/International 220-240 VAC /230 VAC 180-264 VAC /180-254 VAC
Frequency Nominal Tolerance range	North America/International 60 Hz / 50 Hz 47-63 Hz / 47-63
Phase required	single-phase or 3-phase
Power requirements (max)	4.3 kW
Power cable	10 ft. (3.0 m) pluggable cords
Sound power	Approximately 55 dBA
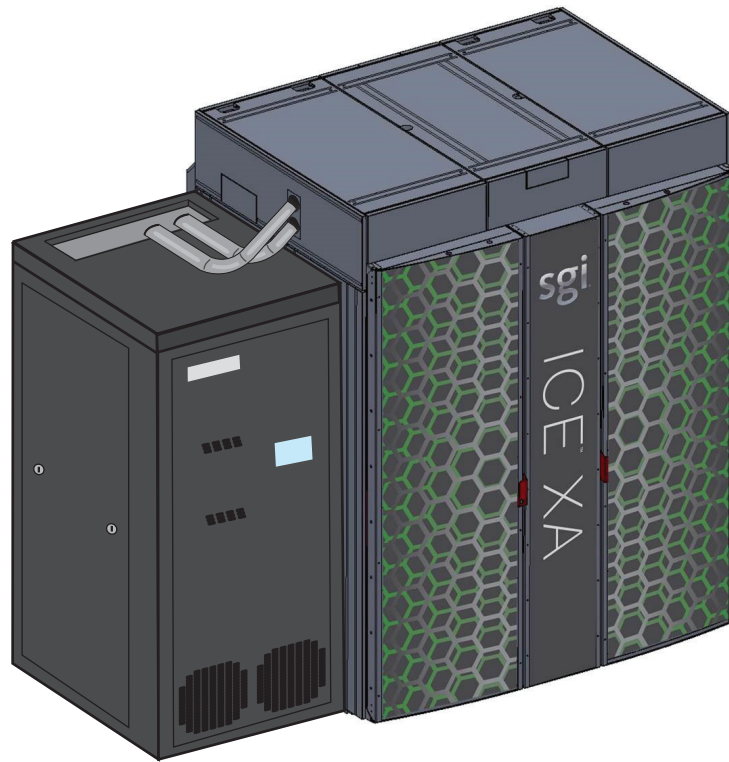
Figure 4-4   E-Cell Rack Assembly Example (CDU on Left)

SGI ICE XA 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 XA 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 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 XA system through an Ethernet network.

The administrative controller server (also known as the system admin controller) is at the top of the distributed management infrastructure within the SGI ICE XA system. The overall SGI ICE XA series management is hierarchical (see the following subsection "System Hierarchy" and also Figure 5-1 on page 56), with the RLC(s) communicating with the compute nodes via CMC interconnect.

System Hierarchy

The SGI ICE XA system has a four-tier, hierarchical management framework. The ICE XA systems contain hardware and software components as follows:

- System Admin Controller (SAC) - one per system

- Rack Leader Controller (RLC) - one per 8 CMCs

- Chassis Management Controller (CMC) - one per blade enclosure

- 48-port Gigabit Ethernet switch (note that some system racks may contain more than one 48-port switch)

- Baseboard Management Controller (BMC) - one for each of the following:

    - Compute node (note that some blades may contain more than one logical node)

    - SAC

    - RLC

    - Service node

    - 48-port Gigabit Ethernet switch

Communication Hierarchy

Communication within the overall management framework is as follows:

Admin Node

The Admin node communicates with the following:

- All rack leader controllers (RLCs)
- All cooling rack controllers (CRCs)
- All optional cooling distribution units (CDUs)

Rack Leader Controller (RLC)

The RLCs within a "logical rack" (see tip that follows) communicate with the following:

- All chassis management controllers (CMCs) within the same logical rack as the RLC.

Chassis Management Controllers

The CMCs within a logical rack communicate with the following:

- The rack leader controllers (RLCs) within the same logical rack

The following two components communicate with CMCs in E-Cell rack systems:

- The cooling rack controller (CRC) is assigned to the same logical rack as the CMCs
- The cooling distribution unit (CDU) is assigned to the same logical rack (and VLAN) as the CMCs

Tip: A logical rack can be one or two physical racks. The logical rack is based on the RLC. There is one RLC in each logical rack. An RLC supports a maximum of eight CMCs. Therefore, one logical rack of enclosures can occupy two physical racks.

System management hierarchy
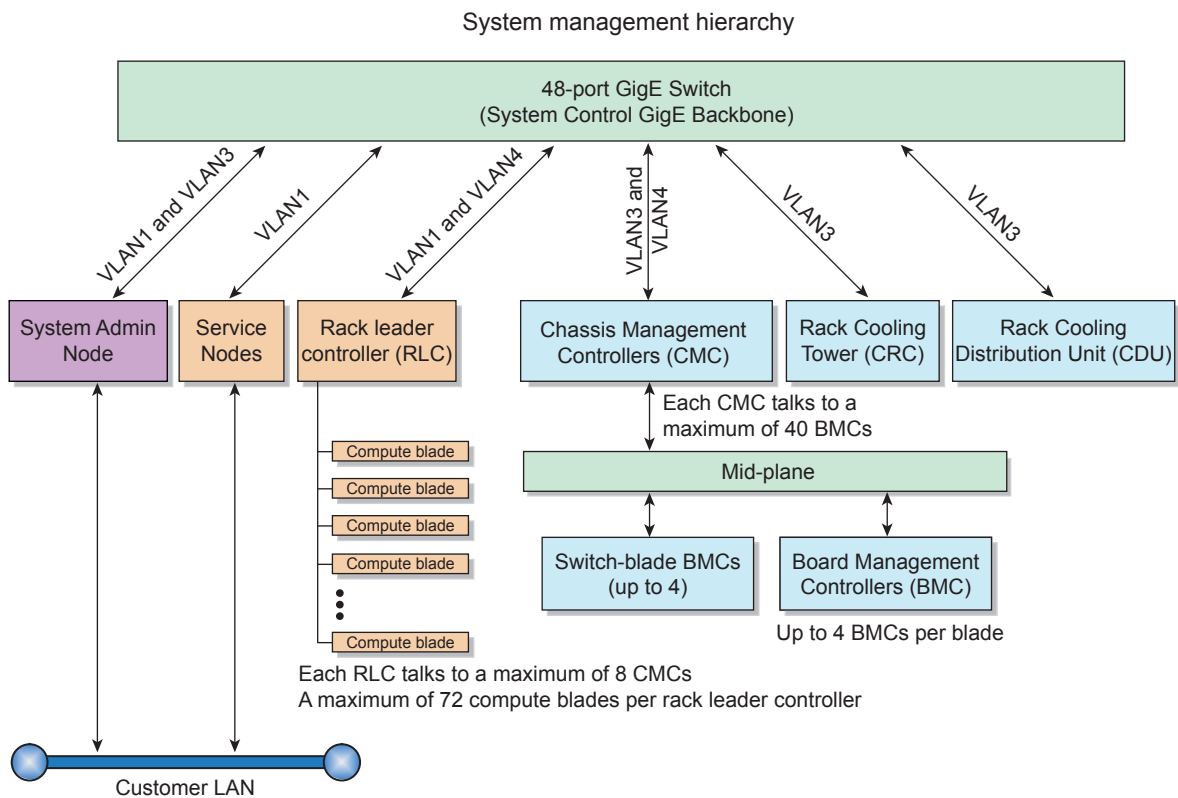
Figure 5-1   SGI ICE XA System Administration Hierarchy Example Diagram

1U Rack Leader Controller and Administration Server

Figure 5-2 shows the front of the 1U server used as a Rack Leader Controller (RLC) and also used separately as an administration server for the ICE XA system.
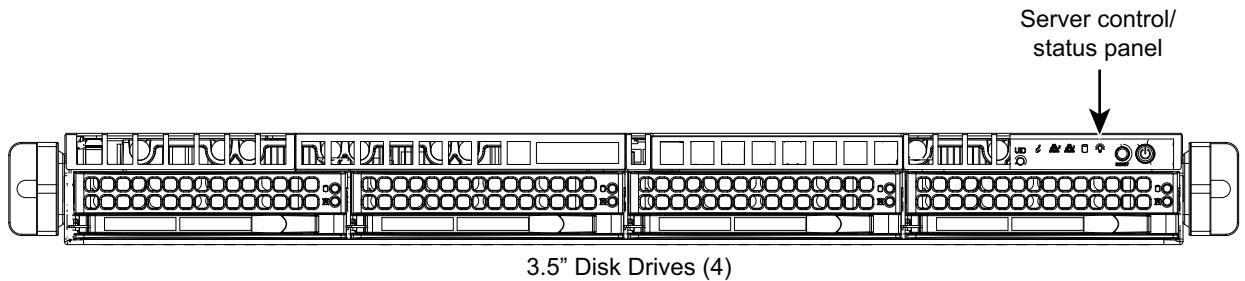
Figure 5-2   1U Rack Leader Controller (RLC)/Administration Server Front Panel

The system administrative controller unit acts as the SGI ICE XA 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 cluster management software. Refer to the *SGI Management Center Administration Guide for Clusters*, (P/N 007-6358-00*x*) if you need more detailed information.

The RLC can query the CMCs for monitored environmental data (temperatures, fan speeds, and voltages) for each blade enclosure. Power control for each blade within an enclosure is handled by the Baseboard Management Controller (BMC) also under the direction of the RLC. The RLC can also query each logical node's BMC to obtain environmental and error log information about each blade. The RLC is an appliance node and always runs software specified by SGI.
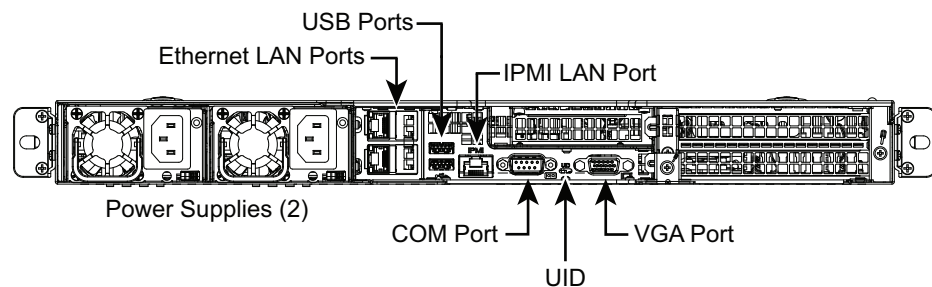
Figure 5-3   1U RLC/Administration Server Rear Components Example

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 29 for a list of administration and support server types and additional functional descriptions.

1U Service Nodes

The 1U rack leader controller server (shown in Figure 5-2 on page 57) can be optionally used as a non-RLC/admin service node. The following subsection ("C1104-GP1 1U Service Node") describes an additional 1U service node that is **never** used as the system administrator or RLC node in an ICE XA system.

C1104-GP1 1U Service Node

The Rackable C1104-GP1 server is a 1U rackmount service node used as a login, batch, fabric management, I/O, MDS, or OSS system. At the heart of the system is a dual-processor serverboard based on the Intel C610 chipset. The serverboard (motherboard) supports two Intel Xeon E5-2600(v3) series processors. Four DDR4 memory channels are available per CPU socket with two DIMM slots per channel. A direct media interface connects the node's PCH ASIC to processor 1; while 40-lane Gen-3 PCIe interconnect lines link both processors directly to the motherboard LAN ports. Separate QPI link pairs connect the two processors and the I/O hub in a network on the motherboard.

The serverboard has eight DIMM slots per processor (16 total) that support DDR4 2133/1866/1600/1333 MHz RDIMMs.

The system supports four hard disk drives and up to three internal optional GPUs. An external low-profile PCIe 3.0 x8 option card slot is also supported. Available GPU and PCIe option cards may be limited, check with your SGI sales or service representative for additional information.

Figure 5-4 on page 59 shows the front and rear panel features of the C1104-GP1 service node.

Figure 5-4 on page 59 shows an example of the server's front control panel, see the *SGI Rackable C1104-GP1 System User Guide* (P/N 007-6364-00*x*) for more detailed information on this 1U service node.
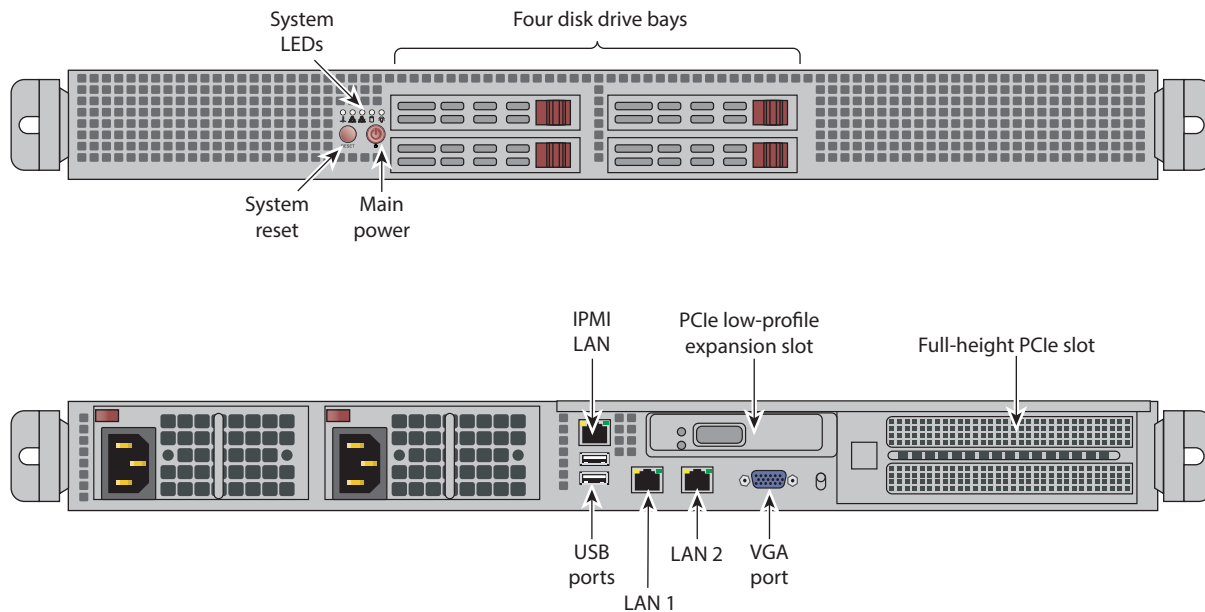
Figure 5-4   SGI Rackable C1104-GP1 1U Service Node Front and Rear Panel Example
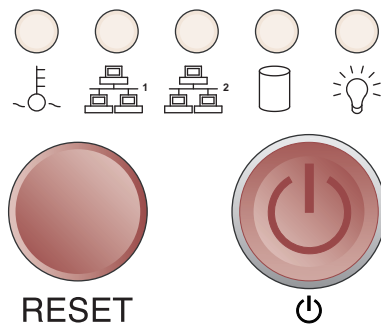
Figure 5-5   SGI Rackable C1104-GP1 System Control Panel and LEDs

From left to right the LED indicators are:

- Overheat/fan fail/UID

- LAN1 and LAN2 network indicators

- Hard drive activity and power good LEDs

2U Service Nodes

For systems using a separate login, batch, I/O, fabric management, or Luster service node; the following SGI 2U servers are also available as options.

C2112-GP2 Service Node

The SGI Rackable C2112-GP2 standard-depth servers are 2U rackmount service nodes. Each model of the server has two main subsystems: the 2U server chassis and a dual-processor serverboard. The C2112-GP2 system offered uses up to twelve hard disk drives. Figure 5-6 shows a front and rear view example of the C2112-GP2 service node's components. See the *SGI Rackable C2112-GP2 System User Guide* (P/N 007-6362-00*x*) for more detailed information on these 2U service nodes.
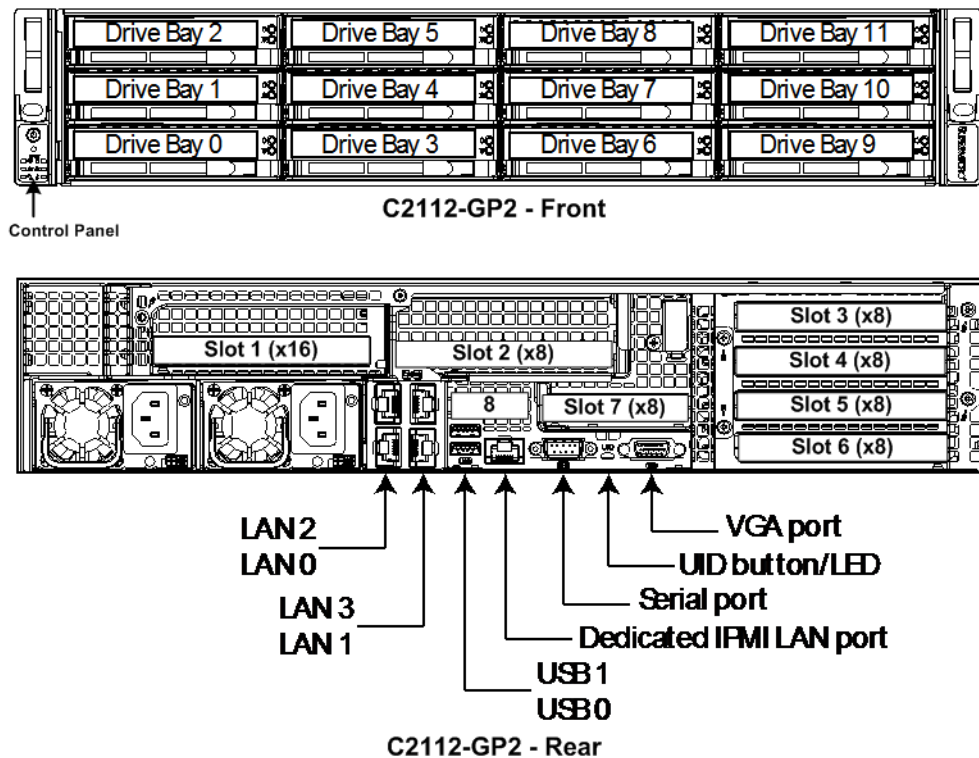
Figure 5-6   SGI Rackable C2112-GP2 Service Node Front and Rear Panel Example

SGI UV 30 2U Service Node

The 2U-high SGI UV 30 server is available as a four-processor ICE XA service node. This 2U service node uses four Intel E5-4600 v3 Xeon processors and supports up to 48 DIMM memory modules plus multiple I/O modules and storage adapters. The server supports up to 24 2.5-inch SATA drives. Figure 5-7 shows an example front and rear view of the SGI UV 30 service node and describes the unit's external components.
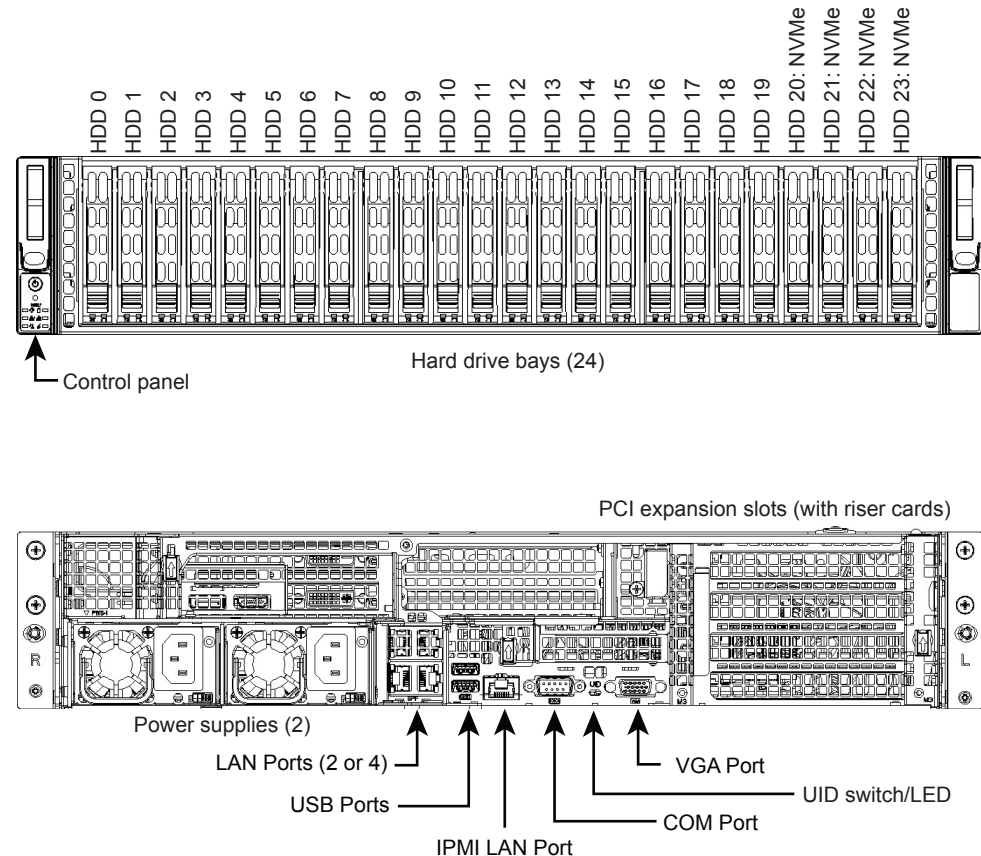
Figure 5-7   SGI UV 30 Service Node Front and Rear Panel Example

For more information on the SGI UV 30 server, see the *SGI UV 30 System User Guide*, (P/N 007-6419-00*x*).

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 XA series configuration ranges.

Table A-1   SGI ICE XA Series Configuration Ranges

Category	Minimum	Maximum
Blades per enclosure	2 blades[a]	9 blades
Blades per E-Cell	18 blades	72
Compute nodes per blade	4 compute nodes	4 compute nodes
Blade enclosure	1 per rack	4 per rack
Compute blade DIMM capacity	8 DIMMs per blade	24 DIMMs per blade
Chassis management boards	1 per enclosure	4 per rack (eight per E-Cell)
InfiniBand switch blades	1 per enclosure	4 per enclosure (16 per rack)

a. Compute blades support four stuffed sockets each.

D-Rack Physical and Power Specifications

Table A-2 shows the physical specifications of the SGI ICE XA service and storage node D-Rack.

Table A-2   ICE XA System D-Rack Physical Specifications

System Features (single rack)	Specification
Height	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	~2,650 lbs. (1,205 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)
Power requirements (max)	34 kW
Hold time	20 ms
Power cable	10 ft. (3.0 m) pluggable cords
Access requirements	
Front	48 in. (121.9 cm)
Rear	48 in. (121.9 cm)
Side	None

D-Rack System Environmental Specifications

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

Table A-3   Environmental Specifications (Single D-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 XA (rack)	Approximately 115.63 kBTU/hr maximum (based on 33.89 kW - 100% dissipation to air)
Heat dissipation to air Water-cooled ICE XA (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 2,800 CFM (typical air cooled) (2400 CFM - water 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)

ICE XA E-Rack Technical Specifications

Table A-4 provides information on the individual physical specifications of the compute racks used in an E-Cell assembly. Table A-5 on page 67 lists the environmental specifications for an individual E-rack.

Table A-4   SGI ICE XA E-Rack Physical Specifications

Characteristic	Specification
Height	93 in. (236.2 cm)
Width	33 in. (83.8 cm)
Depth	48.4 in. (121.9 cm)
Weight (full)	~2,426 lbs. (1,103 kg) approximate
Shipping weight (max)	~2,850 lbs. (1,295 kg) approximate
Voltage range Nominal Tolerance range	North America/International 200-240 VAC /230 VAC 180-264 VAC /180-254 VAC
Frequency Nominal Tolerance range	North America/International 60 Hz / 50 Hz 47-63 Hz / 47-63
Phase required	single-phase or optional 3-phase
Power requirements (max)	76 kVA (77.47 kW)
Hold time	20 ms
Power cable	10 ft. (3.0 m) pluggable cords
Power receptacle Single power option Three-phase option	North America/Japan | International Maximum 8, 30-Amp | Maximum 8, 32-Amp NEMA L6-R30 _____| NEMA IEC60309 (2) 60-Amp 4-wire __| (2) 32-Amp 5-wire IEC60309 _____| IEC60309

Table A-5   Environmental Specifications (Single E-Rack)

Feature	Specification
Temperature tolerance (operating with 95-Watt processors)	+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 (operating with 135-Watt processors)	+5 °C (41 °F) to +28 °C (82.4 °F) (up to 1500 m / 5000 ft.) +5 °C (41 °F) to +23 °C (73.4 °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 95% operating (no condensation) 10% to 95% non-operating (no condensation)
Rack cooling requirements	Chilled water cooling
Heat rejection (dissipation) to air	Zero BTUs
Heat rejection (dissipation) to water	Approximately 246 kBTU/hr maximum (21 tons) (based on 100% dissipation to water)
Maximum altitude	10,000 ft. (3,049 m) operating 40,000 ft. (12,195 m) non-operating
E-Cell rack acoustical noise level (sound power)	Approximately 80 dBA (at front of system)
Optional cooling distribution rack (CDU) acoustical noise level (sound power)	Approximately 65 dBA (at front of unit)

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.
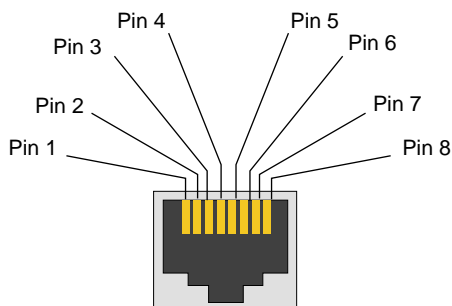
Figure A-1   Ethernet Port

Table A-6 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-6   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" on page 69
- "Regulatory Specifications" on page 71

Safety Information

Read and follow these instructions carefully:

1. Follow all warnings and instructions marked on the product and noted in the documentation included with this product.

2. Unplug this product before cleaning. Do not use liquid cleaners or aerosol cleaners. Use a damp cloth for cleaning.

3. Do not use this product near water.

4. Do not place this product or components of this product on an unstable cart, stand, or table. The product may fall, causing serious damage to the product.

5. Slots and openings in the system are provided for ventilation. To ensure reliable operation of the product and to protect it from overheating, these openings must not be blocked or covered. This product should never be placed near or over a radiator or heat register, or in a built-in installation, unless proper ventilation is provided.

6. This product should be operated from the type of power indicated on the marking label. If you are not sure of the type of power available, consult your dealer or local power company.

7. Do not allow anything to rest on the power cord. Do not locate this product where people will walk on the cord.

8. Never push objects of any kind into this product through cabinet slots as they may touch dangerous voltage points or short out parts that could result in a fire or electric shock. Never spill liquid of any kind on the product.

9. Do not attempt to service this product yourself except as noted in this guide. Opening or removing covers of node and switch internal components may expose you to dangerous voltage points or other risks. Refer all servicing to qualified service personnel.

10. Unplug this product from the wall outlet and refer servicing to qualified service personnel under the following conditions:

    •   When the power cord or plug is damaged or frayed.

    •   If liquid has been spilled into the product.

    •   If the product has been exposed to rain or water.

    •   If the product does not operate normally when the operating instructions are followed. Adjust only those controls that are covered by the operating instructions since improper adjustment of other controls may result in damage and will often require extensive work by a qualified technician to restore the product to normal condition.

    •   If the product has been dropped or the cabinet has been damaged.

    •   If the product exhibits a distinct change in performance, indicating a need for service.

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

12. Use only the proper type of power supply cord set (provided with the system) for this unit.

13. Do not attempt to move the system alone. Moving a rack requires at least two people.

14. Keep all system cables neatly organized in the cable management system. Loose cables are a tripping hazard that cause injury or damage the system.

Regulatory Specifications

The following topics are covered in this section:

- "CMN Number" on page 71
- "CE Notice and Manufacturer's Declaration of Conformity" on page 71
- "Electromagnetic Emissions" on page 71
- "Shielded Cables" on page 74
- "Electrostatic Discharge and Laser Compliance" on page 74
- "Lithium Battery Statements" on page 75

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 and Laser Compliance

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.

Any optional DVD drive used with this computer is a Class 1 laser product. The DVD 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. Använd samma batterityp eller en ekvivalent typ som rekommenderas av apparattillverkaren. Kassera använt batteri enligt fabrikantens instruktion.

Warning: Varoitus: Pääristo 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.