SGI® InfiniteData™ Cluster Hardware User Guide

Document Number 007-6308-001

Record of Revision

Version	Description
001	February, 2014 First release

Contents

Overview

This chapter provides an overview of the main features of the SGI InfiniteData Cluster platform, shown in Figure 1-1. The InfiniteData Cluster is a custom-configured, high-performance, compute server that features high-density computing (up to 160 processors per cabinet) and low power consumption (as low as 10W per node for cooling). The last section of this chapter describes operating precautions. Before operating your system, familiarize yourself with the safety information in that section.

System Features

This section describes the flexible architecture and technical features of the SGI InfiniteData platform in the following sections:

SGI InfiniteData Cabinet (Rack)

The SGI InfiniteData cabinet is a stand-alone 42U rack. The 42U rack can hold a variable number of *compute trays*, up to 80 1U trays or 40 2U trays. Figure 1-1 on page 2 shows the front of the

42U InfiniteData cabinet. Note that an optional top-mounted extension is available (see Figure 1-1) to add room for switches or optional full-width equipment.

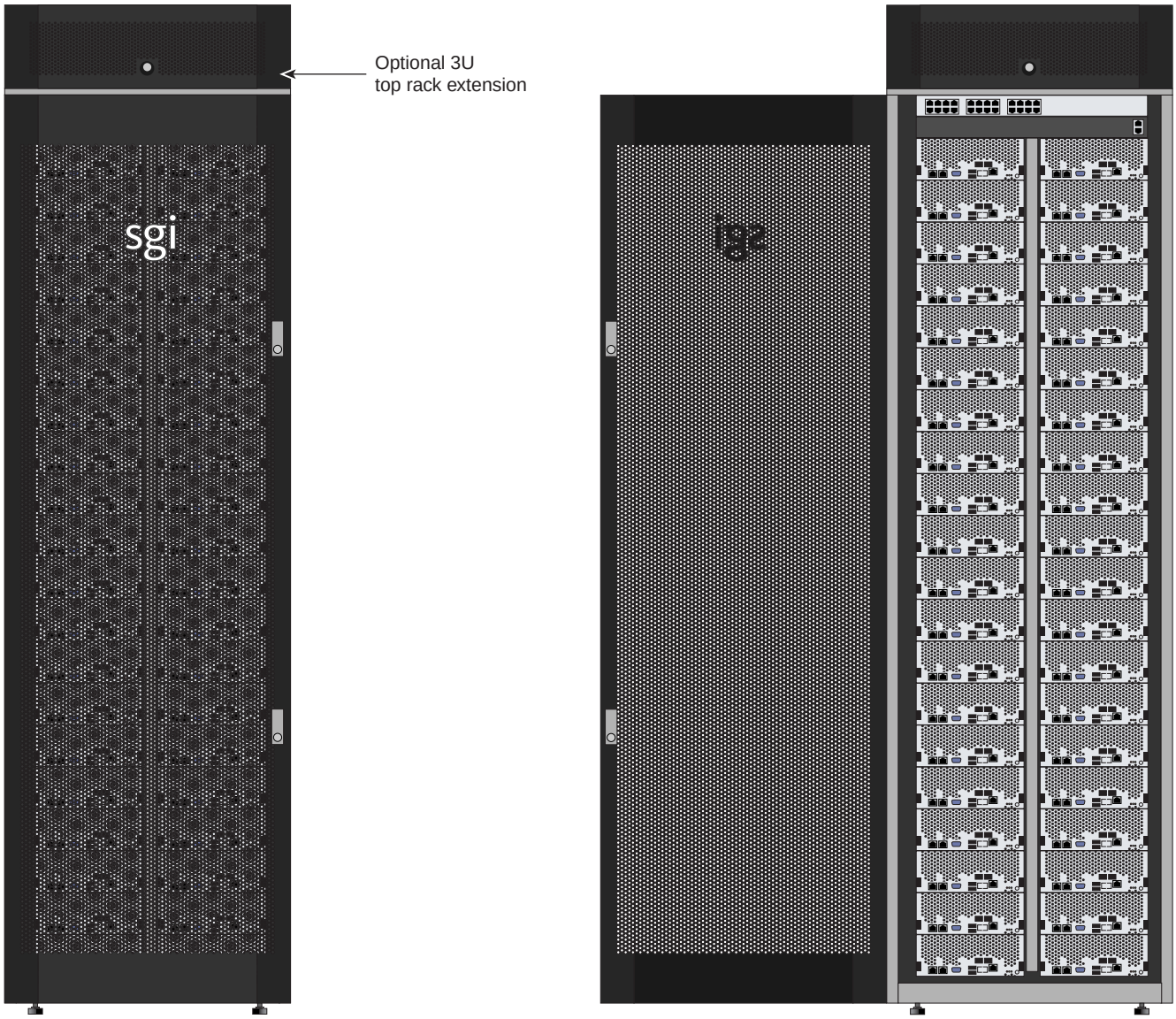

Figure 1-1 SGI InfiniteData Cluster Cabinet Examples

Note that an optional 2U head node spans the width of the rack interior, see the subsection “Optional Head Node Server” on page 9 for additional detail. The cabinet’s 24” x 48” footprint approximates that of standard floor tiles and an optional 3U extension is available for additional head nodes, networking or ATS equipment.

Hadoop Specialized Rack Option

The InfiniteData system can be optionally specialized for Hadoop Reference Implementation (RI) usage. This version of the rack allocates 36U of space in the lower section of the cabinet for 1U or 2U InfiniteData compute trays. A 2U space is reserved for switch installation and the remaining 4U of space (U38 through U42) is used for the installation of 1U full-width (full-depth) nodes. Only one Hadoop RI rack is used in each Hadoop cluster configuration.

Figure 1-2 on page 4 shows an example front view of the Hadoop Reference Implementation specialized rack with the front door removed.

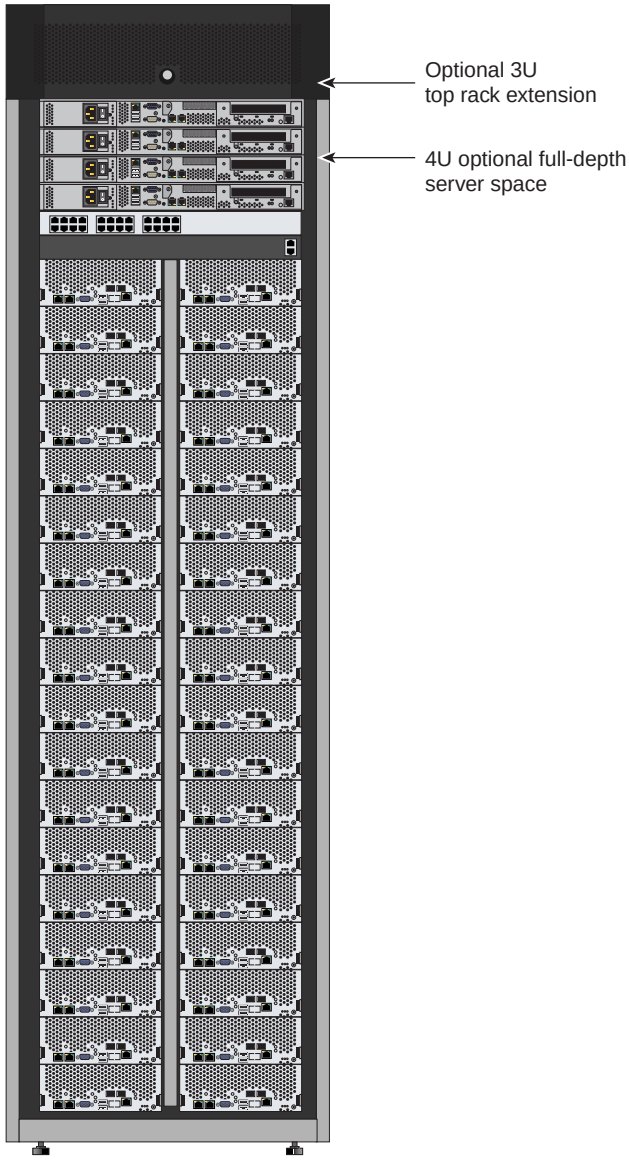

Figure 1-2 Hadoop Specialized InfiniteData Rack Option Example

Compute Trays

At the heart of each InfiniteData compute tray is a multi-processor based nodeboard (serverboard). The InfiniteData Cluster tray serverboards documented in this guide are based on Intel chipsets. The compute trays also house hard drives, graphics or I/O options, as well as fans and power supplies (see Figure 1-2 for an example 2U tray).

Tip: Customers requiring AMD™ clusters can check with an SGI sales or service representative for information on non-InfiniteData AMD based cluster offerings available from SGI.

As stated earlier, the InfiniteData Cluster rack architecture supports up to 80 1U trays or 40 2U trays (72 1U and 36 2U trays in a Hadoop RI rack), along with the following features:

- Ability to mix and match 1U and 2U trays
- Ability to use a 1U standard-depth head node (multiple nodes in Hadoop configuration)
- Individual node serviceability

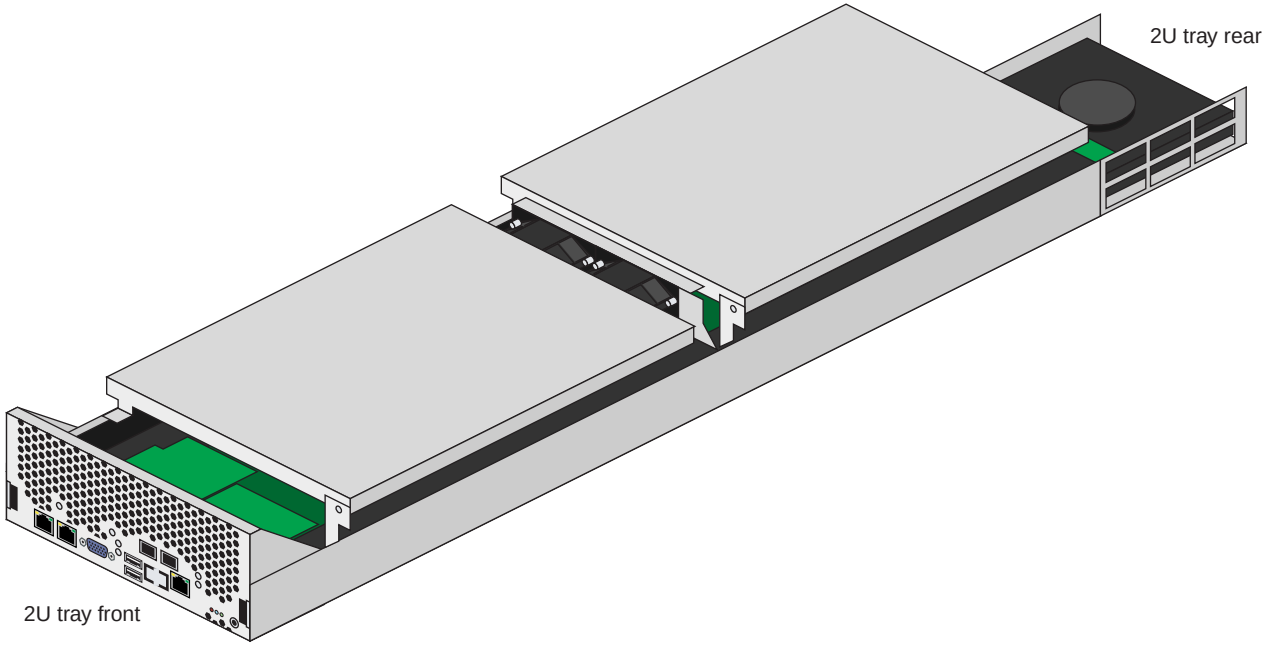

Figure 1-3 2U Compute Tray Example

Processors

The InfiniteData Cluster can support the following minimum/maximum configuration of processors:

- 42U-system minimum: 4 processors (2 nodes)
- 42U-system maximum: 160 processors (80 nodes)

The exact type of processors provided with your system depends on the specific configuration you ordered. Check with your sales or service representative for information on processor availability, upgrades and compatibility. The following are examples of types available for the SGI InfiniteData Cluster platform:

- One dual-socket Intel® Xeon® E5-2600 series processor board per tray
- One dual-socket Intel® Xeon® E5-2600v2 series processor board per tray

DIMM Memory

Memory configuration varies depending on the processor board type ordered with your compute trays. DIMM population requirements are eight or 16 per serverboard and standard DIMM speeds are1600 MHz and 1866 MHz. Note that use of more than one DIMM per channel will result in the processor operating that channel at 1600 MHz by default regardless of the DIMM's operational capacity.

Power Supplies

Your InfiniteData Cluster uses one high-efficiency power supply per compute tray. Depending on your system configuration, each power supply provides a minimum of 450W of power to the compute node tray.

Caution: The cabinet power supply cord is used as the main disconnect device. Ensure that a socket outlet is located or installed near the equipment and is easily accessible.

Compute Tray Drive Subsystem

Each compute tray was designed to support SATA or SAS hard disk drives. Specific optional 2U trays may be available with up to 12 drives configured for each tray. Note that the InfiniteData standard trays only support SATA disk drives. However, by adding an optional SAS HBA certain versions of the compute trays can support SAS disk drives. The compute trays will also support software RAID 0, 1 and 10 on both SAS and SATA disk drives. Hardware supported RAID 5 and 6 are also optionally available in each compute tray.

Compute Tray Front I/O Panel

Each compute tray installs via the front of the InfiniteData cabinet. Figure 1-4 shows an example compute tray front panel. Its I/O panel typically provides COM ports, status LEDs, and Ethernet ports.

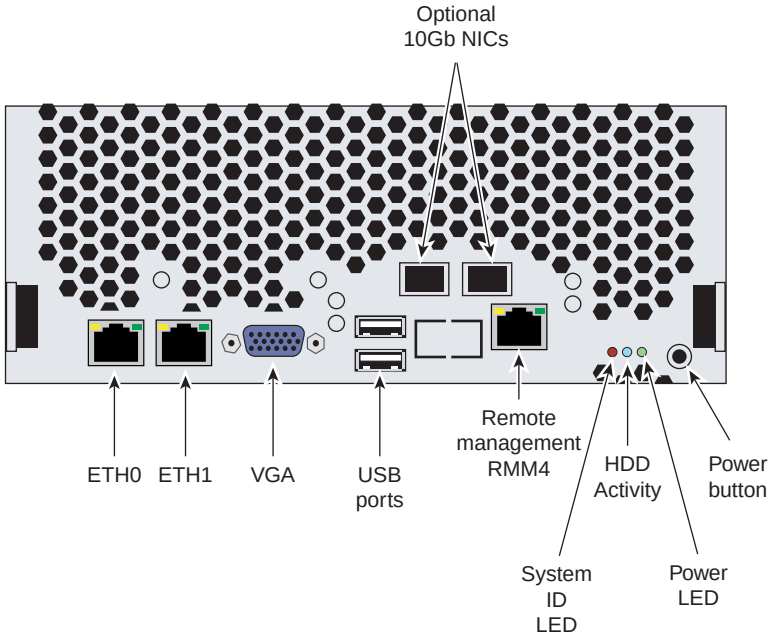

Figure 1-4 Compute Tray Front Panel Example

Cooling System

The hardware architecture of both the compute trays and the racks provide for optimal front-to-back cooling. Each compute tray has an enclosed fan set, see Figure 1-5 for an example. 1U compute trays use 40mm and 2U trays use 80mm cooling fans. A sensor on the main board can increase or decrease the fan speed based on the ambient conditions. Note that maximum operational ambient air temperature is +30ºC (86ºF). See Table A-2 on page 24 for more environmental information.

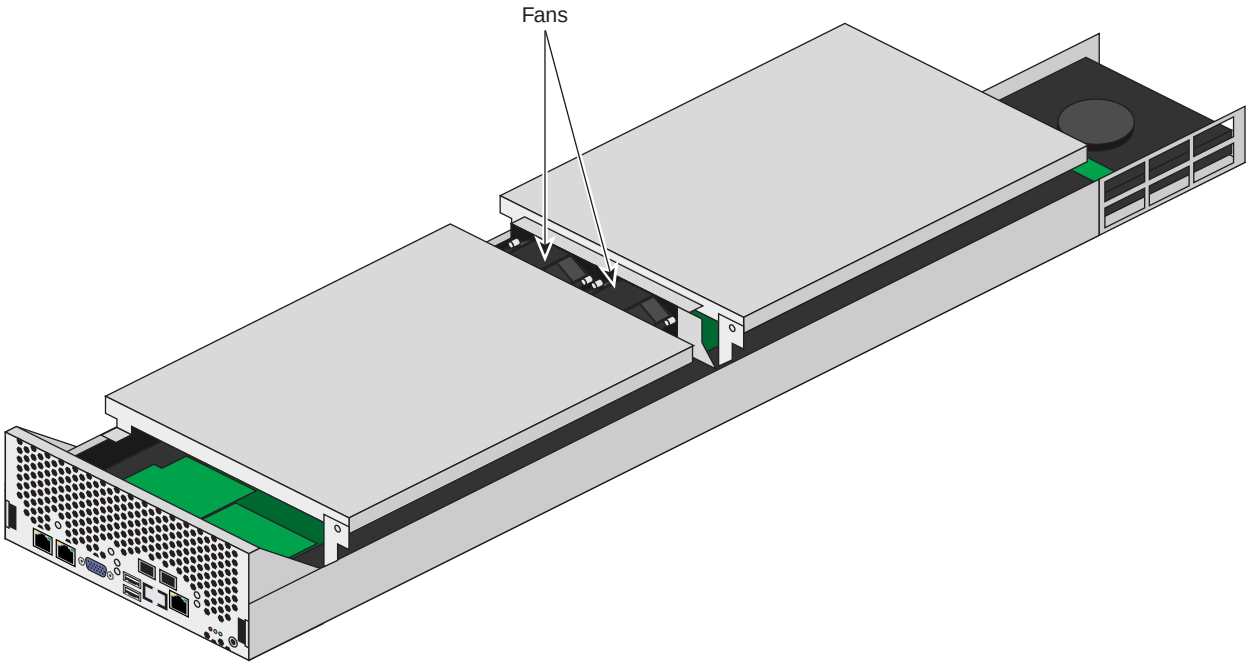

Figure 1-5 Compute Tray Fans Example

Optional Hardware

In addition to the InfiniteData cabinet and the compute trays, various other hardware components may be included as part of your InfiniteData configuration as noted below:

- SATA accessories:
  - A minimum of one disk drive (per tray) is required for operation.

  - One (1) internal SATA port per compute node
  - One (1) SATA cable set per compute node.
- SAS hard drive options
- One or two Gbit Ethernet or (optional) InfiniBand switches

Note: InfiniteData Hadoop clusters do not offer optional InfiniBand connectivity.

Optional Head Node Server

The SGI Rackable C1110-RP6 system is a compute-optimized head node server comprised of two main subsystems: the 1U chassis and a serverboard (motherboard). Check with your sales or service representative before loading any operating system on your server not provided by the SGI factory or service organization.

Various hardware components may be included as part of your SGI Rackable C1110-RP6 system head node as listed below:

- Serverboard with:
  - Two Intel Xeon processors in LGA2011 sockets
  - Intel C602 chipset and Quick Path Interconnect system bus support the processors
  - Sixteen DDR3 memory DIMM slots
  - Two external Gen3 x16 PCIe slots (one full-height and one low-profile)
- Two disk drive configurations include:
  - Four 3.5-inch hard disk drives plus DVD option drive
  - Ten 2.5-inch hard disk drives (no DVD option available)
- Seven 4-cm counter-rotating cooling fans for the serverboard
- One 450 Watt auto-ranging power supply (with single power connector)
- Optional 650 Watt redundant auto-ranging power supply (with dual power connectors)

Figure 1-6 on page 10 shows an example of the rear panel features of the SGI Rackable C1110-RP6 optional head node server. Your SGI Rackable C1110-RP6 system uses a single-cord high-efficiency 450-Watt cold-swap power supply (shown on the left side of Figure 1-6). Optionally, a dual-cord 650-Watt redundant power supply is available. Both types of supply are

80+ Gold rated. The AC power supply cord(s) should always be removed from the C1110-RP6 head node before servicing or replacing the power supply.

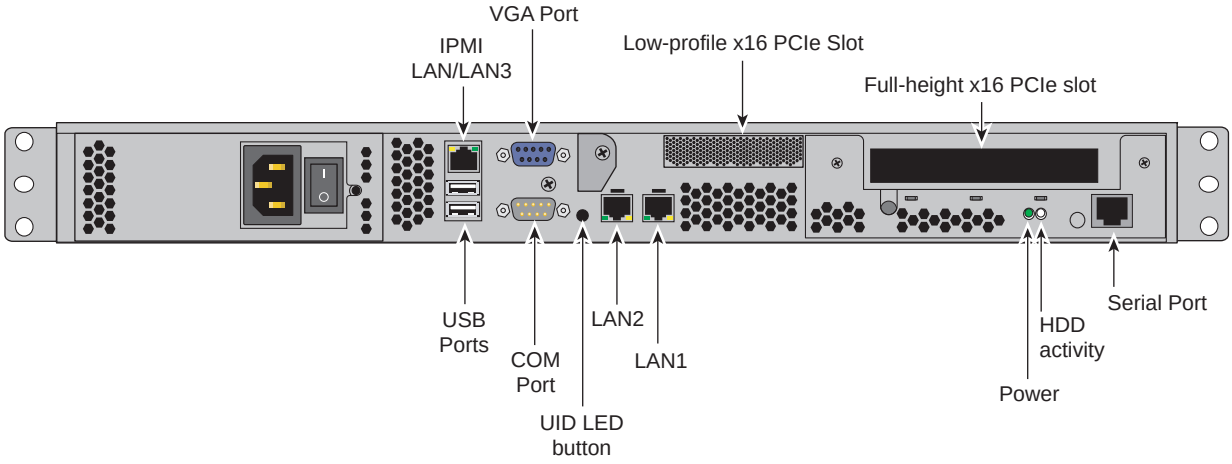

Figure 1-6 SGI Rackable C1110-RP6 Head Node System Rear Panel and I/O Ports

Operating System Support

SGI supports Linux based industry-leading, high-performance operating systems. At time of publication, operating systems supported, tested and installed by SGI included Red Hat Linux and SUSE Linux. SGI can optionally install CentOS, but does not support or test it.

Check with your SGI sales or service representative before loading any operating system on your server not provided by the SGI factory or service organization.

Electrostatic Discharge (ESD) and Safety Precautions

Caution: Observe all ESD precautions. Failure to do so can result in damage to the equipment.

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

Warning: Before operating or servicing any part of this product, read the safety precautions.

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.

System Operation and Troubleshooting

The first half of this chapter describes the basic steps needed to get your SGI InfiniteData Cluster up and running. Following these steps in the order given should enable you to have the system operational within a minimal amount of time. The second half of this chapter provides you with some basic troubleshooting advice. Use these sections to eliminate simple problems or obtain information that may be needed by your service provider.

Unpacking the System and Choosing an Operating Location

You should inspect the InfiniteData Cluster shipping crate and note if it was damaged in any way. If the server itself shows damage you should file a damage claim with the carrier who delivered it.

When you decide on a suitable location for the system, it should be situated in a clean, dust-free area that is well ventilated. Avoid areas where heat, electrical noise and electromagnetic fields are generated. You will also need it placed near a dedicated, 200-240 Volt, and grounded three-phase power outlet.

The InfiniteData Cluster is designed to fit into a server room environment. Take care to maintain the following operating conditions:

- The system should have a six-inch (15 cm) minimum top air clearance.
- The system should be protected from harsh environments that produce excessive vibration and heat.
- The system should be kept in a clean, dust-free location to reduce maintenance problems.
- Available power must be rated for large computer operation (30 amps at 200-240 Volts).

Choosing a Setup Location

When choosing a setup location, observe the following guidelines:

- Leave enough clearance in front of the cabinet to enable opening the front door completely ~48 inches (1.2 meters).
- Leave sufficient clearance in the back of the cabinet to allow for adequate airflow and ease in servicing.
- High-performance systems are generally placed in a restricted access location (dedicated equipment rooms, service closets, or the like).

Cabinet Precautions

Observe the following guidelines:

- Compute trays should be installed in the lowest part of the cabinet possible.
- To maintain proper cooling, always keep the cabinet front door and all panels and components on the compute trays closed when not servicing.

Compute Tray Precautions

Observe the following guidelines:

- Review the electrical and general safety precautions in Chapter 1.
- For extra protection, use a regulating uninterruptible power supply (UPS) to protect the InfiniteData Cluster power supplies from power surges, voltage spikes and to keep your system operating in case of a power failure. This is an optional device not provided by SGI with your system.
- Service personnel should always allow the hot-plug disk drives and power supply modules to cool before touching them. To maintain proper cooling, always keep the cabinet door closed when it is not being serviced.
- Make sure all power and data cables are properly connected and not blocking the enclosure's airflow.

Compute Tray Operating Considerations

Use the guidelines in the following subsections to properly use and maintain the compute trays in the cabinet.

Ambient Operating Temperature

The ambient operating temperature of the rack environment may be greater than the ambient temperature of the room. Therefore, consideration should be given to installing the equipment in an environment compatible with the manufacturer's maximum rated ambient temperature. Note that maximum operational ambient air temperature is +30ºC (86ºF). See Table A-2 on page 24 for more environmental operating information.

Reduced Airflow

Equipment should be mounted into the cabinet so that the amount of airflow required for safe operation is not compromised.

Mechanical Loading

Equipment should be mounted into the cabinet so that a hazardous condition does not arise due to uneven mechanical loading.

Circuit Overloading

Consideration should be given to the connection of the equipment to the power supply circuitry and the effect that any possible overloading of circuits might have on over-current protection and power supply wiring. Appropriate consideration of equipment nameplate ratings should be used when addressing this concern.

Reliable Ground

A reliable ground for the system must be maintained at all times. To ensure this, the rack itself should be grounded. Particular attention should be given to power supply connections other than the direct connections to the branch circuit (for example, the use of power strips). Note that all

power and data cables should be routed in such a way that they do not block the airflow generated by the compute tray fans.

Providing Power

Plug the power cord from the power supply array into a rack power distribution unit (PDU) or high-quality power source that offers protection from electrical noise and power surges.

For higher availability, it is recommended that you use an optional uninterruptible power supply (UPS) with the cluster server (not a standard product provided by SGI).

Unlike many rack enclosures, the back of the InfiniteData Cluster cabinet does not have a master power switch. The rack's individual rear AC breakers should all be switched on prior to booting the system.

Note also that the rack does not have an opening rear door and access is limited. See Figure 2-1 for an example view of the cabinet rear.

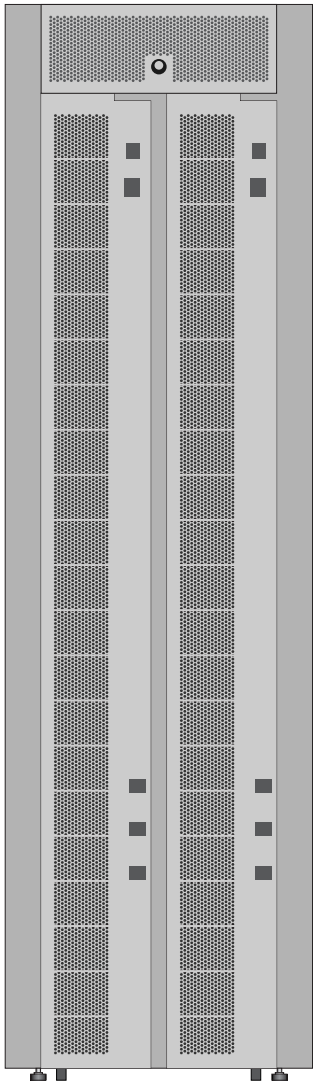

Figure 2-1 InfiniteData Cabinet—Back View Example

You can power on each compute tray individually using it's power button, but this is normally accomplished remotely using the interface in the SGI Management Center software. See the lower-right portion of Figure 2-2 for an example of the power button location.

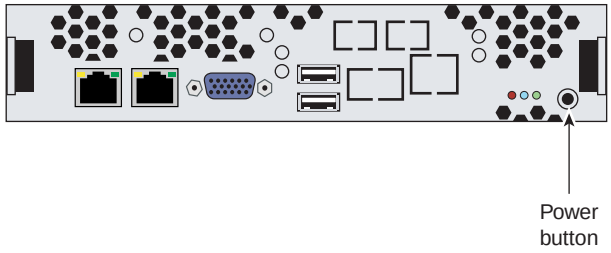

Figure 2-2 Compute Tray Front Power Button Location Example

The compute tray power connector and rocker switch are located on the rear of each unit, see Figure 2-3 for an example. Note that a remote power-on is the normal bring-up process for trays.

Figure 2-3 Rear Tray Power Switch Location Example

As listed in the previous chapter, the SGI InfiniteData Cluster can also be configured as a Hadoop Reference Implementation system. The cabinet's rear sheetmetal is slightly different from the standard InfiniteData compute rack, see Figure 2-4 on page 19 for an example.

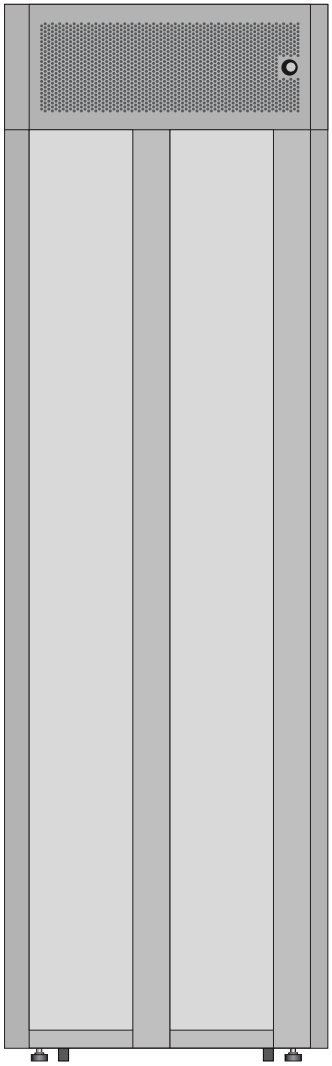

Figure 2-4 Hadoop Reference Implementation Cabinet Example Rear View

Upgrading or Replacing Components

The following InfiniteData Cluster components are field-replaceable:

- Memory DIMMs
- PCIe cards
- Disk drives
- Fans
- Power supplies

They are industry-standard components and should be managed in the conventional manner.

Warning: If a lithium battery is installed in your system as a soldered part, only service personnel qualified by SGI 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.

Troubleshooting the System

The following table lists recommended actions for problems that can occur. To solve problems that are not listed in this table or in another section of this chapter, contact your SGI or other approved service provider.

Table 2-1 Troubleshooting Chart

Problem Description	Recommended Action
The system will not power on.	Ensure that the power cord of the cabinet is seated properly in the power receptacle. Did you push the power button on the “head node” compute tray as well as the “compute node” trays? If the power cord is plugged in and all the power switches are on, contact your support organization.
An individual compute/memory tray will not power on.	View the LED outputs on the front of the tray, (see also Figure 2-5 on page 22). Ensure the unit is plugged in and switched on If the LEDs are not lit, contact your support provider.
The system will not boot the operating system.	Contact your support provider.
The PWR LED of a populated PCI slot in a tray is not illuminated.	Reseat the PCI card, reboot the tray and recheck the LED.
The Fault LED of a populated PCI slot is illuminated (on).	Reseat the PCI card. If the fault LED remains on, replace the PCI card.

Individual Tray LEDs and Buttons

Each server tray installed in an InfiniteData Cluster cabinet has LED indicators (lower-right) to show the operational status of the tray. See the example in Figure 2-5.

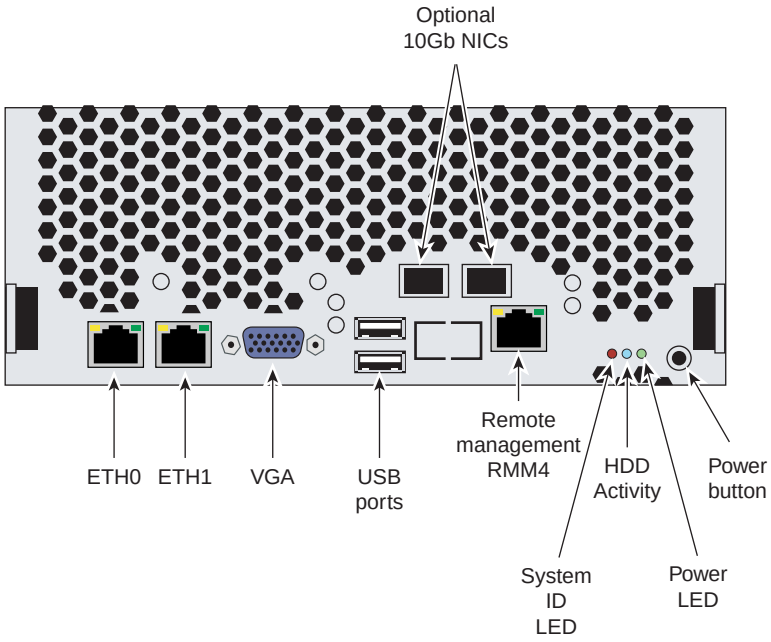

Figure 2-5 Example Compute Tray Status LEDs and Buttons

The status LEDs are located on the lower-right front section of the tray and are visible when the front door of the cabinet is open. Table 2-2 describes the tray's front status LEDs.

Table 2-2 Compute Tray LEDs and Buttons

LED/Button	Display	Description
Eth *x*/NIC *x*		When flashing, the NIC LED indicates network activity on LAN *x*.
System ID	Red	This programmable LED is used for selection and identification of a particular tray in an InfiniteData Cluster system rack.
HDD Status	Blue/White	When lit, this LED indicates that functionality is established on the hard disk drive or solid state drive.
Power Status	Green	When lit, this LED indicates that power is being supplied to the tray.
Power Button		Pushing this button applies or turns off power to the compute tray but keeps standby power supplied to the tray.

Technical Specifications

This appendix contains technical specification information about your system.

System Specifications and Features

Table A-1 lists the physical and electrical specifications of the SGI InfiniteData Cluster.

Table A-1 SGI InfiniteData Cluster Physical Specifications

System Features	Specification
Cabinet height	85.6 in. (217.4 cm)
Cabinet width	24 in. (61 cm)
Cabinet depth	47.4 in. (120.4)
Weight (full) maximum	Approximately 2135 lbs. Shipping weight will be greater.
Phase required	Two three-phase PDUs per rack
Voltage and CPS range	North America - nominal voltage is 208V at 24A per phase (50/60hz for 3W+PE) Europe - nominal voltage is 415V at 30A per phase (50/60hz for 3W+N+PE)
Power cables	Two three-phase pluggable cords per rack
Power supplies	One per compute tray (450W minimum per power supply)
Cooling	Compute tray fans (40mm x4 for 1U tray or 80mm x2 for 2U tray)
Hard drive bays	Up to 4 per 1U compute tray, up to 12 per 2U compute tray
PCIe slots	Optional low-profile (x16) PCI-Express slots on HPC compute trays, plus two (x16) PCIe slots on optional C1110-RP6 head node

Environmental Specifications

Table A-2 lists the environmental specifications of the InfiniteData system.

Table A-2 Environmental Specifications

Feature	Specification
Temperature tolerance (operating)	+5 °C (41 °F) to +30 °C (86 °F) (up to 1500 m / 5000 ft.) +5 °C (41 °F) to +25 °C (77 °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)
Cooling requirement	Ambient air cooling
Maximum altitude	10,000 ft. (3,049 m) operating 40,000 ft. (12,195 m) non-operating
Acoustical noise level (sound pressure)	Less than 84 dBa maximum

Regulatory Overview

This appendix provides a summary of the product regulatory compliance.

Table B-1 Product Regulatory Compliance

Category	Compliance Information
Electromagnetic Emissions	FCC Part 15 Class A, EN 55022 Class A, ICES-003 Class A, AS/NZS CISPR 22 Class A, VCCI V-3
Electromagnetic Immunity	EN 55024
Safety	CSA/EN/IEC/UL 60950-1 Compliant, UL or CSA Listed (USA and Canada)
Perchlorate Materials	California Best Management Practices Regulations for Perchlorate Materials: ”This Perchlorate warning applies only to products containing CR (Manganese Dioxide) Lithium coin cells. Special handling for perchlorate material may apply. See www.dtsc.ca.gov/hazardouswaste/perchlorate.”