SGI 9000 Storage System Owner's Guide

Document Number 007-4068-001

CONTRIBUTORS

Written by Carolyn Curtis
Illustrated by Dan Young and Carolyn Curtis
Production by Amy Swenson
Engineering contributions by Sammy Wilborn, Paul Stolle, Tim Piper, and Jaideep Deshmukh
St. Peter's Basilica image courtesy of ENEL SpA and InfoByte SpA. Disk Thrower image courtesy of Xavier Berenguer, Animatica.

The author gratefully acknowledges the assistance of the publications group at Xyratex, Inc., in developing this manual.

FCC Warning
This equipment has been tested and found compliant 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 the user will be required to correct the interference at his own expense.

Attention
This product requires the use of external shielded or optical cables in order to maintain compliance pursuant to Part 15 of the FCC Rules.

SGI 9000 Storage System Owner's Guide
Document Number 007-4068-001

European Union Statement

This device complies with the European Directives listed on the "Declaration of Conformity" which is included with each product. The CE mark insignia displayed on the device is an indication of conformity to the aforementioned European requirements.

International Special Committee on Radio Interference (CISPR)

This equipment has been tested to and is in compliance with the Class A limits per CISPR publication 22, Limits and Methods of Measurement of Radio Interference Characteristics of Information Technology Equipment; Germany's BZT Class A limits for Information Technology Equipment; and Japan's VCCI Class 1 limits.

Canadian Department of Communications Statement

This digital apparatus does not exceed the Class A limits for radio noise emissions from digital apparatus as set out in the Radio Interference Regulations of the Canadian Department of Communications.

Attention

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

VCCI Class 1 Statement for Japan

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。

SGI 9000 Storage System Owner's Guide
Document Number 007-4068-001

Chinese Class A Warning

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.

SGI 9000 Storage System Owner's Guide
Document Number 007-4068-001

Contents

List of Figures

List of Tables

About This Guide

This guide explains how to operate and maintain the SGI 9000 fibre channel storage system. Part of the SGI Total Performance Series of fibre channel storage, this standalone tower provides compact, high-capacity, high-availability RAID and JBOD (“just a bunch of disks”) storage for the SGI 1000 series of servers.

The storage system can be connected to one or two fibre channel PCI boards (host bus adapters, or HBAs) in the SGI server separately or in combination (loop). Software interfaces from a third party are shipped with the storage system.

Audience

This guide is written for owners and users of the SGI 9000 fibre channel storage system. It presumes general knowledge of fibre channel technology and knowledge of the host SGI server, of the PCI HBA, and other fibre channel devices to which the storage system might be cabled.

Structure of This Document

This guide consists of the following chapters:

- Chapter 1, “Storage System Features,” describes storage system formats and the modules in the storage system.
- Chapter 2, “Connecting to a Host and Powering On and Off,” explains how to cable the storage system to a host, how to connect the power cord, and how to power the storage system on and off.
- Chapter 3, “Features of the RAID Controller Module,” describes SCSI Enclosure Services, configuration on disk, drive roaming, Mylex Online RAID Expansion, and data caching.

- Chapter 4, "Using the RAID Controller," introduces software tools for the controller, gives configuration information, and explains RAID levels and criteria for selecting them, storage system drives and drive state management, and automatic rebuild.
- Chapter 5, "Troubleshooting," describes storage system problems and suggests solutions. It explains how to use storage system LEDs and the storage system alarm for troubleshooting.
- Chapter 6, "Installing and Replacing Disk Drive Modules," explains how to add a new disk drive module and how to replace a defective disk drive module.
- Appendix A, "Technical Specifications," gives specifications for the storage system in general and for specific modules.

An index completes this guide.

Other Required Documentation

Besides this manual and the manuals for the storage system third-party software, have available the latest versions of the owner's and installation guide for the server:

- *SGI 1400 Server Family Maintenance and Upgrades Guide*
- *SGI 1400 Server Family User's Guide*

Also have available the owner's guides for any other fibre channel devices to which you are attaching the storage (such as the SGI Fibre Channel Hub or switch).

If you do not have these guides, you can find the information online in the following locations:

- IRIS InSight Library: from the Toolchest, choose Help > Online Books > SGI EndUser or SGI Admin, and select the applicable owner's or hardware guide.
- Technical Publications Library: if you have access to the Internet, enter the following URL in your Web browser location window: http://techpubs.sgi.com/library/

Conventions Used in This Guide

These type conventions and symbols are used in this guide:

Helvetica Bold	Labels on hardware, such as for ports and LEDs
Italics	Filenames, manual or book titles, new terms, program variables, tools, utilities, and variables to be supplied by the user in examples, code, and syntax statements
“”	(Double quotation marks) References in text to document section titles

Chapter 1

Storage System Features

This chapter gives information on the basic features of the SGI 9000 fibre channel RAID storage system. Figure 1-1 shows the front of the storage system.

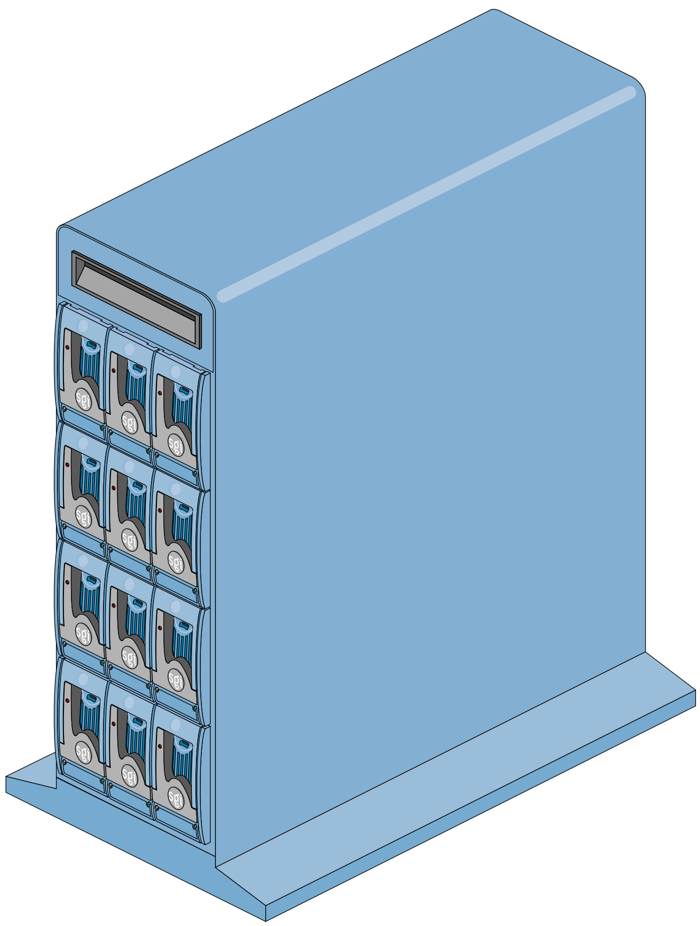

Figure 1-1 Tower (Deskside), Front View

The SGI 9000 storage system uses disk technology that allows you to replace a disk while the system continues to run. This “hot swapping” must be done only under controlled circumstances.

The disk drives are installed at the front of the tower; all other modules are installed at the rear, as shown in Figure 1-2. Blanking plates cover unoccupied bays.

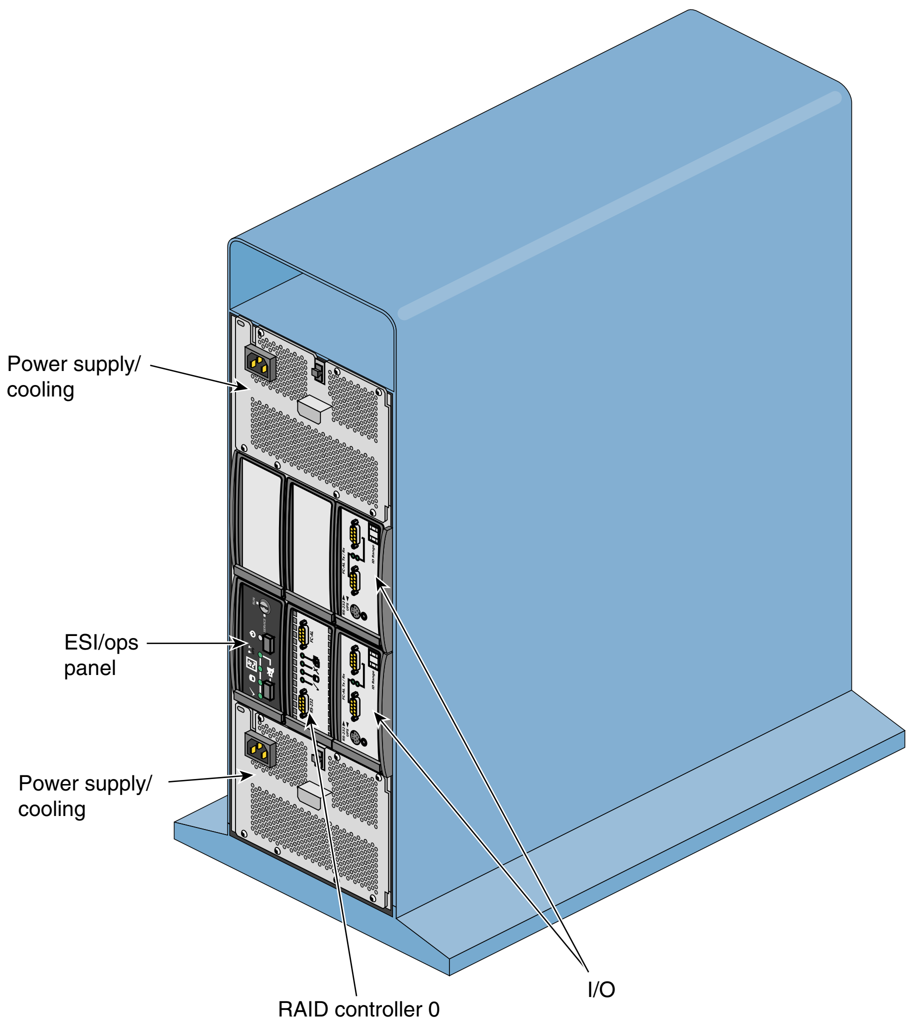

Figure 1-2 Tower, Rear View

The storage system can be connected to one or two fibre channel PCI boards (host bus adapters, or HBAs) in an SGI 1000 series server separately or in combination (loop).

The chassis consists of two segmented aluminum assemblies with a midplane PCB between them. Each chassis assembly contains 12 bays; a bay is defined as the space required to house a single 3.5-inch disk drive (1.6 inches high, or 1 inch high with a foam filler) in its carrier module. Larger modules—the power supplies—are accommodated in multiple bay spaces. The midplane PCB provides logic level signal and low-voltage power distribution paths.

Component modules in the storage system are

- Two power supply/cooling modules
- Enclosure system interface/operators panel (ESI/ops)
- One RAID controller module with FC-AL input/output
- Two loop resiliency circuit (LRC) I/O modules
- From two to 12 FC-AL disk drive carrier modules

The modules are described in separate sections in this chapter:

- "Power Supply/Cooling Module" on page 4
- "Enclosure System Interface/Operator (ESI/Ops) Panel Module" on page 5
- "RAID Controller Module" on page 7
- "FC-AL Loop Resiliency Circuit (LRC) I/O Modules" on page 9
- "Disk Drive Modules" on page 10

This chapter concludes with a section describing chassis bay numbering: "Storage System Bay Numbering" on page 12.

Power Supply/Cooling Module

Two 550-watt power supply/cooling modules are mounted in the rear of the storage system. Figure 1-3 shows this module. Each module contains two fans as well as the power supply and its associated electronics. The fans derive power from the chassis midplane, not from the power supply itself.

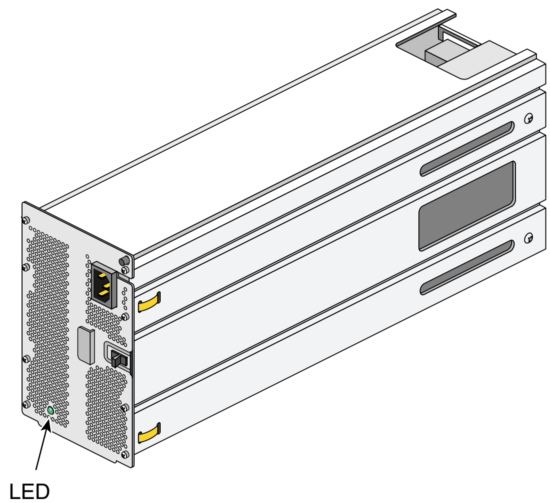

Figure 1-3 Power Supply/Cooling Module

Power supply voltage operating ranges are nominally 115V or 230V, selected automatically.

An LED on the front panel of the power supply/cooling module indicates the status of the power supply and the fan.

The modules operate together; if one fails, the other maintains the power supply and cooling while the faulty unit is replaced. The faulty module, however, still provides proper airflow for the storage system.

Power cords are included with the storage systems; they are described in "Connecting the Power Cord" on page 14 in Chapter 2. Specifications for the cables and for the power supply are in Table A-4 on page 68 in Appendix A, "Technical Specifications."

Enclosure System Interface/Operator (ESI/Ops) Panel Module

The ESI/ops panel provides the storage system with a microcontroller for monitoring and controlling all elements of the storage system. Each element (power, cooling, temperature, device status) is interfaced to the processor using an I^2C (I square C) bus.

Figure 1-4 shows details of the ESI/ops panel; Figure 1-2 shows its location in the tower.

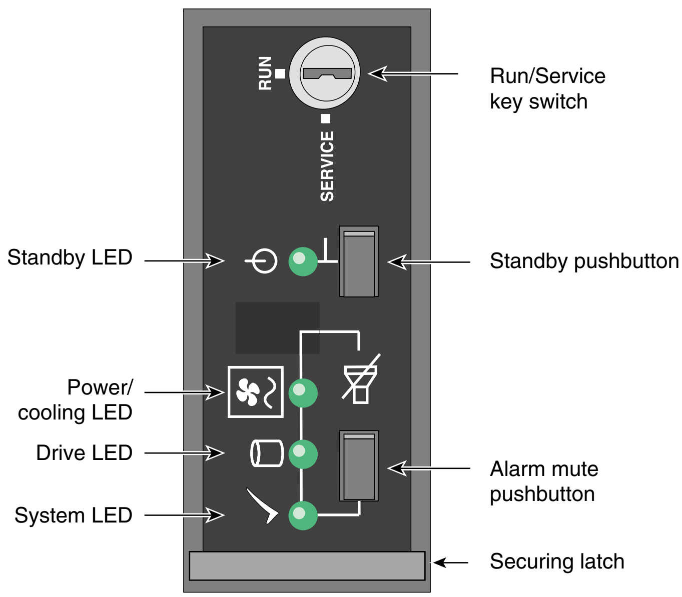

Figure 1-4 ESI/Ops Panel Indicators and Switches

The ESI/ops panel shows a consolidated status for all modules. Table 1-1 summarizes the function of the LEDs on the ESI/ops panel, as well as those on other modules in the storage system.

Table 1-1 Storage System LED Indications

LED	Meaning
Green, steady	Positive condition
Green, flashing	Noncritical condition; Disk drives: data activity
Amber, flashing	Noncritical condition
Amber, steady	Fault

In addition to the indicators on the ESI/ops panel, each module type has its own status LED(s). Using the LEDs for troubleshooting is explained in “Using Storage System LEDs for Troubleshooting” on page 45 in Chapter 5.

Other features of the ESI/ops panel are as follows:

- Key switch: activates run or service mode, as explained in “Key Switch for Starting the Storage System”
- Standby pushbutton: activates standby mode when the key switch is in the Service position; see “Powering On the Storage System” on page 17 in Chapter 2
- Alarm mute pushbutton: mutes the alarm; see “Audible Alarm” on page 7

The ESI/ops panel module firmware includes SCSI Enclosure Services, which monitors the modules in the storage system and controls the ESI/ops panel LEDs. The ESI/ops panel requires two disk drives in specific drive bays to serve as conduits for information from the storage system to the ESI/ops panel module. SES is also required for the RAID controller module(s); see “SCSI Enclosure Services (SES) and Disk Drive Control” on page 21 in Chapter 3 for more information on this aspect.

Key Switch for Starting the Storage System

For general use, the storage system is started by turning the key switch on the far right (top) of the ESI/ops panel to the **RUN** position; the storage system ships with the key in this position. For specific instructions on starting and stopping the storage system, see “Powering On the Storage System” on page 17 in Chapter 2.

Audible Alarm

The audible alarm on the ESI/ops panel sounds when a fault state is present. Pressing the alarm mute pushbutton reduces the volume of the alarm, but leaves a beep at approximately 10-second intervals to indicate that a fault state is still present. The mute pushbutton is beneath the indicators on the ESI/ops panel (see Figure 1-4).

“Using the Alarm for Troubleshooting” on page 52 in Chapter 5 explains alarm function.

RAID Controller Module

Each storage system contains a RAID controller that brings RAID functionality and high-speed fibre data transfer performance for connection to a fibre channel host bus adapter (HBA)* installed in a host system.

The intelligent caching controller supports industry-standard RAID levels (0, 1, 3, 5, and 0+1) for multiple-drive arrays, and JBOD for single-drive control functionality. The controller also supports RAID 10, 30, and 50 in spanned disk arrays. (“RAID Levels” on page 30 in Chapter 4 has information on supported RAID levels.)

* The QLogic QLA 2100 PCI Fibre Channel HBA is recommended for use with the controller.

Figure 1-5 shows features of the RAID controller module panel.

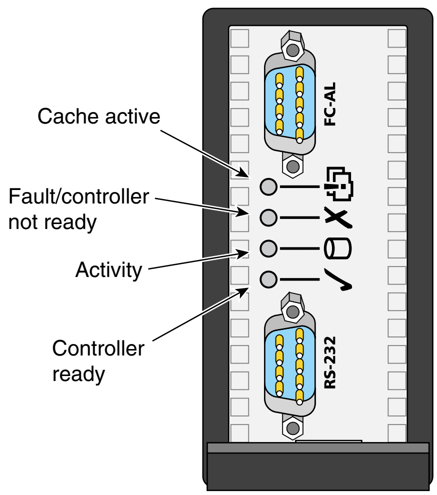

Figure 1-5 RAID Controller Panel Features

LEDs on the controller panel give information on controller activity; see Figure 1-5. The fault/controller not ready LED is amber; the other three LEDs are green.

RAID devices, such as host bus adapters in servers, are cabled to the ports on the I/O module. The ports on the RAID controller itself are used as follows:

- FC-AL port: connection to host system (SGI 1400 server)
- RS-232 port: serial connection

For more information on the controller, see Chapter 3 and Chapter 4.

FC-AL Loop Resiliency Circuit (LRC) I/O Modules

The storage system is shipped with two FC-AL loop resiliency circuit (LRC) I/O modules, which make dual (redundant) data paths possible. Each module provides connection between the RAID controller and one loop of up to 12 drives.

Figure 1-6 shows features and location of the I/O module panel in a tower.

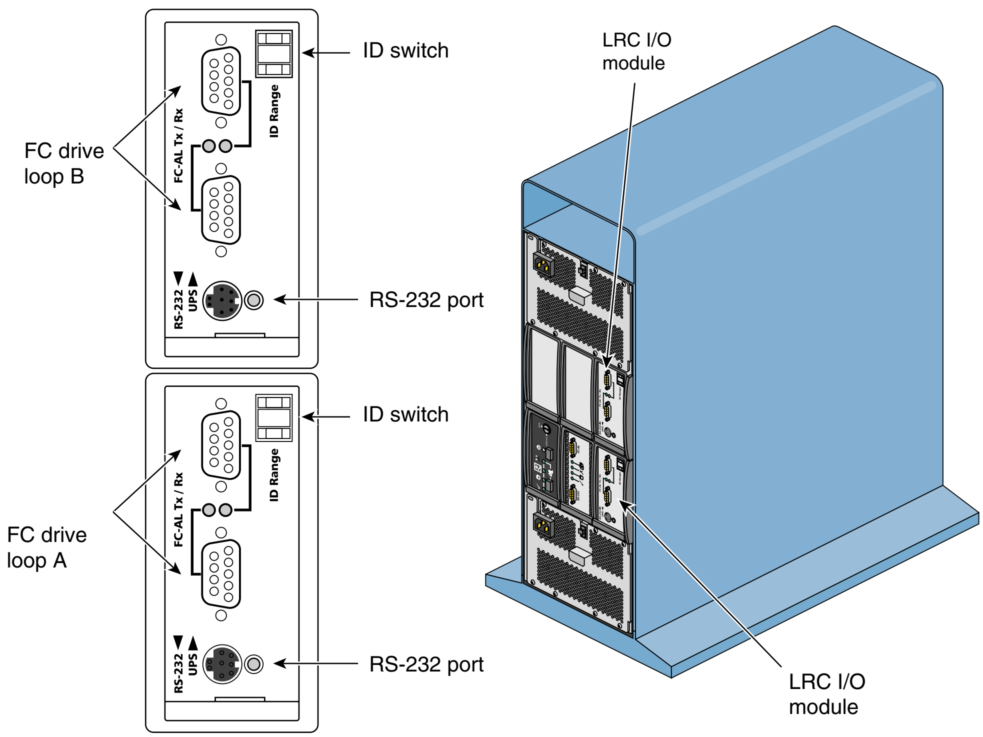

Figure 1-6 I/O Module Panel Features and Location, Tower

Each I/O module has an ID selector switch for the system ID, which is set to 0 on both modules at the factory. Disk drive IDs are based on this ID.

Each port on the I/O module panel has an LED indicating its status.

At initial startup, each I/O module controls a separate loop of six drives; see "Loop Configuration" on page 36 in Chapter 4.

Disk Drive Modules

A disk drive module (FC-AL) consists of a hard disk mounted in a carrier. Each drive bay in the front of the storage system houses a single 1.6-inch high, 3.5-inch disk drive in its carrier, or a 1-inch high drive, which has a foam filler. Figure 1-7 shows the 1.6-inch disk drive module.

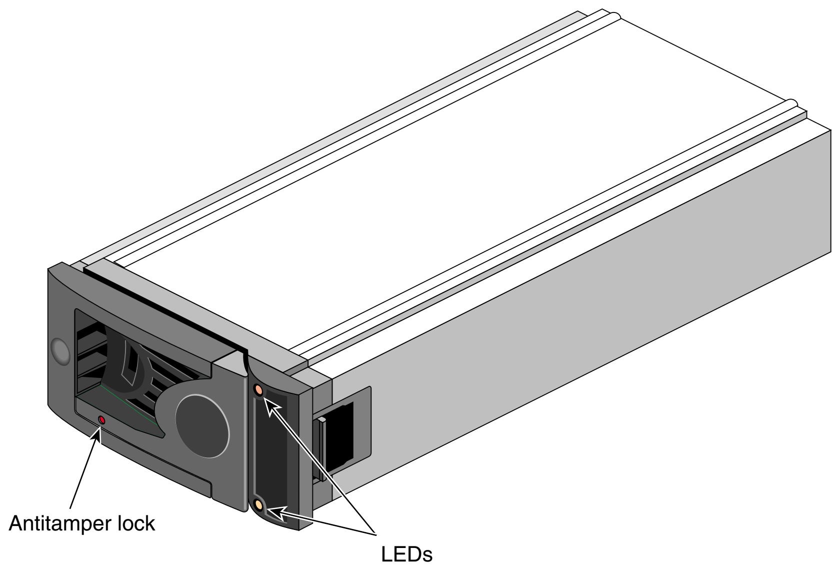

Figure 1-7 Disk Drive Module

The drives in the left top and left bottom bays of the tower are required for storage system management; see “Storage System Bay Numbering” on page 12 for specific information on their location.

The drives are dual-ported and can be configured as two separate loops; see “Loop Configuration” on page 36 in Chapter 4.

Disk Drive Carrier

The extruded aluminum carrier provides thermal conduction and radio frequency and electromagnetic induction protection, and affords the drive maximum physical protection.

The drive module cap has an ergonomic handle that provides the following functions:

- Camming of the carrier into and out of the drive bays
- Positive spring loading of the drive/backplane connector

Disk Drive LEDs

Each drive carrier has two LEDs, an upper (green) and lower (amber). In normal operation the green LED is on, and flickers as the drive operates. The amber LED illuminates when the drive is faulty.

Disk Drive Antitamper Lock

The drive carrier has antitamper lock that disables the normal push/push latch action of the handle; it is fitted in the drive carrier handle, as shown in Figure 1-7. The lock is set through the hole in the lower part of the handle trim. An indicator shows the setting:

- When the drive is locked, a red indicator is visible in the center rectangular aperture in the handle.
- When the drive is unlocked, a black indicator is visible.

The antitamper lock setting is changed with a key through the small round cutout in the lower part of the handle trim piece. (A key is included with each tower.) Changing the setting is explained in “Adding a Disk Drive Module” on page 58 and in “Replacing the Disk Drive Module” on page 66 in Chapter 6.

Drive Fillers

Drive fillers or dummy drives (front dummy fascias) are provided for all unused drive bays and are required to maintain a balanced airflow. They are designed as integral drive module front caps with handles.

Storage System Bay Numbering

Figure 1-8 shows the correct positions of the modules and the enclosure bay numbering convention. A bay is defined as the space required to house a single 3.5-inch disk drive in its carrier module.

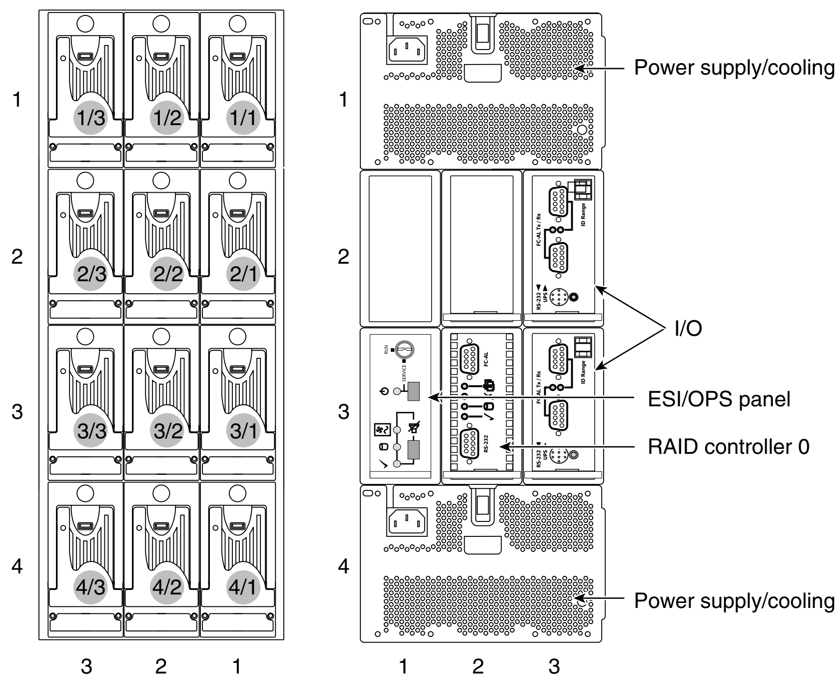

Figure 1-8 Module Locations and Bay Numbering (Tower Configuration)

The tower is 3 x 4, that is, 3 bays wide by 4 bays high.

- The disk drive bays in front are numbered 1 to 3 from right to left and 1 to 4 from top to bottom.
- The rear bays are numbered 1 to 3 from left to right and 1 to 4 from top to bottom.

Module locations are identified by combining the column and row numbers (top and side numbers in Figure 1-8). For example, the ESI/ops panel on the tower is in rear bay 3/1 (third row, the leftmost column).

Chapter 2

Connecting to a Host and Powering On and Off

This chapter consists of the following sections:

Connecting to a Host

Physical attachment to the host is from the FC-AL DB-9 port on the RAID controller module to an FC-AL port in a host (for the SGI 1400, the PCI HBA card).

Besides cabling directly to an HBA in a host, you can connect the storage system to an SGI Fibre Channel Hub using fibre channel copper cabling with DB-9 connectors at each end, or to a SGI fibre channel 8-port or 16-port switch (using a copper cable and copper GBIC). See Table A-7 in Appendix A for information on these cables.

Note: The I/O module current limit for MIA support is 1.5A.

For more information on configurations, see "Loop Configuration" on page 36 in Chapter 4.

Connecting the Power Cord

The tower requires 115-220 volts (autoranging), and is shipped with

- One bifurcated power cord, shown in Figure 2-1; this cable gives you the advantage of a single point of disconnect

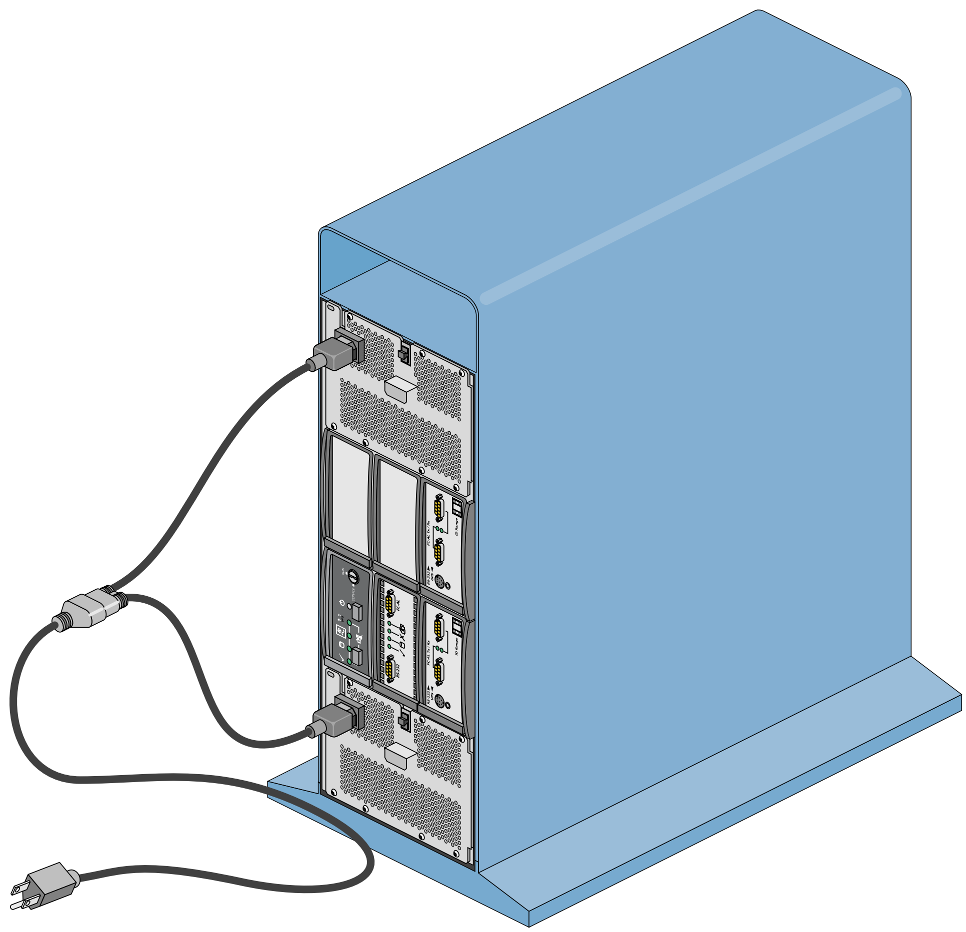

Figure 2-1 Bifurcated Power Cord for the Tower

The bifurcated power cord has dual IEC 320 receptacles and an IEC plug inline cord extension.

- Two country-specific power cords:
  - Attach one at the free end of the bifurcated power cord to connect to a power source.

  or

  - Attach one to each power supply/cooling module power socket instead of the bifurcated power cord, and connect the other end to a power source.
- Two IEC 320 power cords for cabling to the rack power distribution units

  These are included in case the tower is converted to a rackmountable enclosure.

Country-specific power cords are available from SGI that are appropriate to local standards; contact your service provider.

Caution: Use power cords supplied with the storage system or ones that match the specification quoted in Table A-4 on page 68 in Appendix A.

Checking AC Power

The standby LED on the ESI/ops panel turns green if AC power is present; see Figure 2-2 for its location.

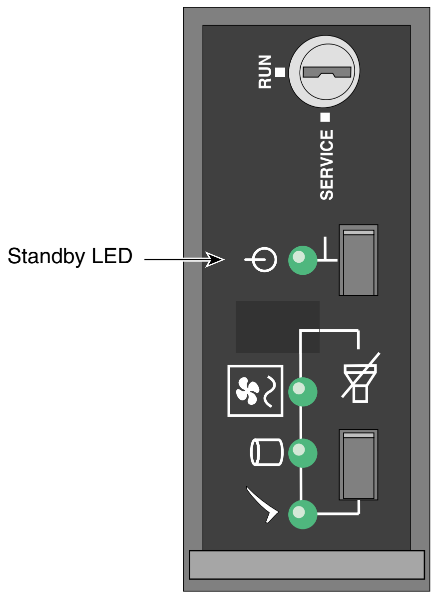

Figure 2-2 EPI/Ops Panel: Standby LED Indicating AC Power Is Present

Checking Grounding

To ensure that a safe grounding system is provided for a tower, follow these steps:

1. Make sure that the power cord is connected to the storage system, but not plugged in to a power source.

   Warning: Some electrical circuits could be damaged if external signal cables or power control cables are present during the grounding checks.

2. Check for continuity between the earth pin of the power cord on the rack power cord on one of the power supply /cooling modules and any exposed metal surface of the storage system.

Powering On the Storage System

Note the following:

- Before powering on the storage system, ensure that all modules are firmly seated in their correct bays and that blank plates are fitted in any empty bays.
- Do not operate the storage system unless the ambient temperature is within the specified operating range of 10˚ C to 40˚ C (50˚ F to 104˚ F). If drives have been recently installed, make sure that they have had time to acclimatize before operating them.

For an unattended installation where the storage system is required to power on whenever AC power is present, the **RUN**/**SERVICE** key switch should be in the **RUN** position. The storage system ships from the factory with this setting. See Figure 2-3.

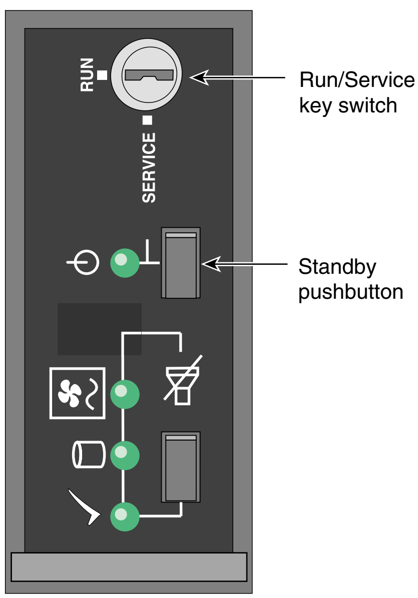

Figure 2-3 ESI/Ops Panel: Powering On

If the **RUN**/**SERVICE** key switch is in the **RUN** position, the tower powers on when the power cord is plugged into a power source.

If the **RUN**/**SERVICE** key switch is in the **RUN** position and power is interrupted, the storage system returns to the power-on condition when power is restored.

In run mode, the standby pushbutton is disabled.

Note: Setting the **RUN**/**SERVICE** key switch on the ESI/ops panel to the **SERVICE** position puts the storage system in service mode; this mode is used by SGI System Support Engineers only. In service mode, if power is lost for any reason, the storage system defaults to standby mode when power is restored. Press the standby pushbutton to power on the storage system.

Checking LEDs at Power-On

At power-on, check storage system LEDs for system status. Under normal conditions, the LEDs should all be illuminated constant green. Figure 2-4 shows details of the ESI/ops panel.

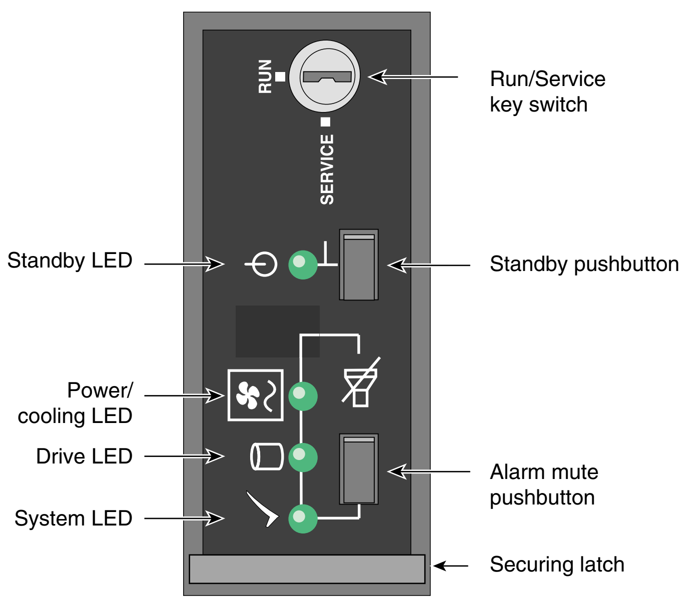

Figure 2-4 ESI/Ops Panel Indicators and Switches

If a problem is detected, the ESI processor in the operator panel changes the relevant LED from green to amber. See "Solving Initial Startup Problems" on page 44 and "Using Storage System LEDs for Troubleshooting" on page 45 in Chapter 5.

Other modules in the storage system also have LEDs, which are described in "Using Storage System LEDs for Troubleshooting."

Powering Off

The storage system toggles between standby mode (all drives are powered off) and run (operation) mode whenever the standby pushbutton is pressed.

Power off the storage system in one of two ways:

- Turn the key switch to the **SERVICE** position and then press the standby pushbutton for 2 seconds.

  To power the storage system back on again from this setting, press the standby pushbutton for 2 seconds again.
- Remove power at the power source (unplug the storage system).

Note the following:

- If the key switch is in the **RUN** position when power is restored to the storage system, the system powers on.
- If the key switch is in the **SERVICE** position when power is restored to the storage system, the system does not power on.

Chapter 3

Features of the RAID Controller Module

This chapter describes features and operation of the RAID controller, in these sections:

SCSI Enclosure Services (SES) and Disk Drive Control

The RAID controllers use SCSI-3 SCSI Enclosure Services (SES) commands to manage the physical storage system. SES provides support for disk drives, power supply, temperature, door lock, alarms, and the controller electronics for the enclosure services. The storage system ESI/ops panel firmware includes SES.

Note: These services are performed by drives installed in bays 1/3 and 4/3; these drives must be present for the system to function. See Figure 1-8 on page 12 in Chapter 1 for their location.

SES access is through an enclosure services device, which is included in the ESI/ops module. SCSI commands are sent to a direct access storage device, namely, the drives in bays 1/3 and 4/3, and are passed through to the SES device.

During controller initialization, each device attached to each fibre loop is interrogated, and the inquiry data is stored in controller RAM. If SES devices are detected, the SES process is started. The SES process polls and updates

- Disk drive insertion status
- Power supply status

- Cooling element status
- Storage system temperature

The LEDs on the ESI/ops panel show this status.

Configuration on Disk (COD)

Configuration on disk (COD) retains the latest version of the saved configuration at a reserved location on each disk drive in the storage system: the same information is written to a reserved location of the NVRAM of the controller(s).

The configuration information includes the configuration ID number and the configuration sequence number. Both numbers are reset to 0 when the configuration is changed. The configuration sequence number increments whenever there is a physical disk drive state change. This feature allows the configuration to be maintained through a controller replacement.

COD plays a significant role during the power-up sequence after a controller is replaced. The replacement controller progresses through a series of checks and validation steps to determine the validity of any configuration currently present in NVRAM, followed by additional checks and validation steps to determine the validity of the COD information on all disk drives in the storage system. The final configuration is determined by one of the following criteria:

- The majority of COD information with the same configuration ID
- The COD available on the disk drive (when there is one disk drive and one controller in the storage system)
- The first configuration encountered (when there is an equal number of valid configurations)

At startup, the RAID controller checks if the storage system configuration matches that stored on disk (see “Configuration on Disk (COD)”); if it does not match, the storage system takes steps to bring it into conformity with the stored configuration.

Drive Roaming

Drive roaming allows disk drives to be moved to other channel/target ID locations while the storage system is powered off. Drive roaming allows for easier disassembly and assembly of systems; it also enhances performance by optimizing channel usage.

Drive roaming uses the Configuration on Disk (COD) information stored on the physical disk drive and in the storage system drive configuration structure. When the storage system reboots, a table is generated for each disk drive's current location and its location before the storage system was powered off. The controller uses this information to remap the physical disk drives into the proper location in the system, or logical, drive.

If a drive fails in a RAID level that uses a hot spare, drive roaming allows the controller to keep track of the new hot spare, which is the replacement for the failed drive.

Drive roaming is designed for use within one system environment, for example, a single storage system or a cluster of systems sharing a single-controller or a dual-active controller configuration.

Note: Disk drives containing valid COD information from other systems must not be introduced into a storage system. If the COD information on a replacement disk drive is questionable or invalid, the disk drive is labeled DEAD. If you use GAM to change that drive's state the invalid COD is overwritten automatically. If the drive with the invalid COD is inserted into the same slot as one that died and was removed, it starts a rebuild. If it is placed in any other nonconfigured slot, it comes up dead and must be marked as standby.

Mylex Online RAID Expansion (MORE)

The Mylex Online RAID Expansion operation adds capacity to an existing LUN while the controller is online with the host. For example, a storage system using a LUN with five disk drives can add another disk drive to create a LUN with six drives. The MORE operation can be performed on all RAID levels except JBOD. The procedures mentioned in this section are performed using the GAM interface. The GAM online help and other GAM documentation describe the processes fully. See “Global Array Manager (GAM), the Software Tool for the Controller” on page 29 in Chapter 4 for an introduction to this software.

The controller continues to service host I/O requests during the LUN expansion process, which includes restriping data from the old (smaller) LUN to the new (expanded) LUN.

The MORE operation provides two options for configuring the added capacity: Add Logical Drive (Add System Drive) option and Expand Logical Drive (Expand or Enlarge System Drive) option. These options and the requirements for disks to be added are explained in separate subsections:

- “Disk Drive Requirements and Restrictions for MORE Options” on page 24
- “Caveats for MORE Options” on page 25
- “Add Logical Drive Option” on page 26
- “Expand Logical Drive Option” on page 26

Disk Drive Requirements and Restrictions for MORE Options

For either MORE option, the disk drives must meet these criteria:

- One to six disk drives can be added to a LUN at one time. The maximum number of physical disk drives in the new storage system drives must not exceed eight.
- The minimum number of disk drives in the source LUN is two; the maximum is seven.
- The disk drive(s) being added must be in standby mode.
- The disk drive(s) being added must not already be part of a LUN.
- The capacity of each of the added disk drives must be greater than or equal to that of the smallest disk drive in the LUN.
- The system (logical) drive to be expanded must be online.

Caveats for MORE Options

For either MORE option (Add Logical Drive or Expand Logical Drive), the following caveats apply:

- Add capacity input parameters and execution parameters are stored in the controller's NVRAM. If a power failure occurs during the process, the controller firmware automatically restarts the process.
- In the event of a disk drive failure, the process continues to completion in critical mode. After migration is completed in critical mode, if automatic rebuild mode is enabled and if a spare disk drive is available, a rebuild starts immediately after migration is complete, restoring the critical system drive to an online state.

  However, if the disk drive fails during the initialization of the newly created system drive (Add System Drive option), the migration operation stops immediately and is considered complete. At this time the rebuild operation begins if applicable.

  The migration process is aborted if a disk drive fails when capacity is added to a RAID 0 set or when a two disk drives fail.
- If a MORE process must be terminated (for example, two disk drives fail and cannot be recovered), the controller automatically aborts the migration process.
- Initialize, Rebuild, Consistency Check, and MORE are mutually exclusive operations. If one of the first three is already running, MORE fails to start.
- No configuration update commands issued from the host are allowed during the MORE process.
- Data is not lost if the RAID controller becomes disabled (for example, it is powered off by mistake or a power supply fails). The process resumes after power-on. Full recovery occurs after a power failure and the operation continues with no data loss.
- Write-back cache is disabled during the MORE process but resumes at the end of the operation. (That is, write cache is enabled during the MORE process, but is disabled at the end of the operation.)
- The following parameters do not change as a result of a MORE operation:
  - existing system drive write policy
  - LUN affinity/LUN mapping
  - cache size
  - all other controller parameters

- Removing and replacing a controller during migration is fatal. A hardware error that renders the controller unusable or the NVRAM unreadable is also fatal during migration.
- An add capacity operation does not start if the target LUN is in critical mode.

Add Logical Drive Option

The Add Logical Drive (Add System Drive) option adds a new system drive with the increased capacity, rather than by increasing the size of the existing system drive(s). The system drive's data is striped across a larger number of physical disk drives than before the expansion operation.

The Add Logical Drive option is rejected if the maximum of eight system drives is specified. When the initialization is complete, you must create and format the new system drive partition in accordance with Windows NT requirements.

Using the Add Logical Drive option, a LUN consisting of more than one system drive processes a MORE request on all system drives with one migration request, processing the system drives one at a time until all have been processed. Allowing one migration service to process all system drives on the same set creates one new extra capacity region when the migration is finished. The new system drive is added at the end of the LUN (selected drive pack).

Expand Logical Drive Option

The Expand Logical Drive (Expand or Enlarge System Drive) option appends the added capacity to the system drive specified in the MORE operation request. This option requires that only one system drive be defined using the set of physical disk drives; if more than one system drive exists on the same set of disk drives, the MORE operation is rejected.

Creating and formatting the added capacity requires backing up all data on the system drive to be enlarged (if appropriate), deleting the current partition, and creating the new, larger system drive.

After the MORE operation, you must delete, recreate, and reformat the system drive partition according to the operating system in use. The data can be saved back onto the system drive following the operating system format process.

Using the Expand Logical Drive option, the system drive specified is enlarged and its data is striped across a larger number of physical disk drives. For this option, the system drive specified must be the only system drive on the pack of disk drives being migrated.

Data Caching

RAID controllers can be operated with write cache enabled or disabled. This section describes them in the following subsections:

- "Write Cache Enabled (Write-Back Cache Mode)" on page 27
- "Write Cache Disabled (Write-Through or Conservative Cache Mode)" on page 27

Write caching is set independently for each system drive in GAM.

Write Cache Enabled (Write-Back Cache Mode)

If write cache is enabled (write-back cache mode), a write completion status is issued to the host initiator when the data is stored in the controller's cache, but before the data is transferred to the disk drives. In dual-active controller configurations with write cache enabled, the write data is always copied to the cache of the second controller before completion status is issued to the host initiator.

Enabling write cache enhances performance significantly for data write operations; there is no effect on read performance. However, in this mode a write complete message is sent to the host system as soon as data is stored in the controller cache; some delay may occur before this data is written to disk. During this interval there is risk of data loss if

- Only one controller is present and this controller fails
- Power to the controller is lost and its internal battery fails or is discharged

Write Cache Disabled (Write-Through or Conservative Cache Mode)

If write cache is disabled (write-through data caching is enabled), write data is transferred to the disk drives before completion status is issued to the host initiator. In this mode, system drives configured with the write cache enabled policy are treated as though they were configured with write cache disabled, and the cache is flushed.

Disabling write cache (enabling write-through or conservative mode) provides a higher level of data protection after a critical storage system component has failed. When the condition disabling write cache is resolved, the system drives are converted to their original settings.

Conditions that disable write cache are

- The Enable Conservative Cache controller parameter is enabled in GAM for a dual-active controller configuration, and a controller failure has occurred.
- A MORE operation is initiated. During a MORE operation, the storage system automatically enters conservative cache mode. After MORE completes, the system drives are restored to their original settings. "Mylex Online RAID Expansion (MORE)" on page 24 has information about MORE.
- A power supply has failed (not simply that a power supply is not present[*]).
- An out-of-limit temperature condition exists.[*]
- The controller receives an indication of an AC failure.

To protect against single controller failure, certain releases of the storage system support dual controllers. To protect against power loss, an internal battery in the controller module maintains the data for up to 72 hours.

[*] SES functions.

Chapter 4

Using the RAID Controller

This chapter explains aspects of RAID controller operation, in the following sections:

Global Array Manager (GAM), the Software Tool for the Controller

The Global Array Manager configures, manage, and monitors RAID arrays connected to the controllers, even across remote servers. For more information on GAM components, requirements, and operation, refer to the documentation included on the Mylex RAID Management Software CD-ROM.

The Global Array Manager "wizard" for setting up and configuring new logical drives and disk arrays is called RAID Assist. In its simplest form, RAID Assist provides an Auto Configuration option that immediately configures all available drives into an optimal RAID 5 configuration. RAID Assist's Assisted Configuration sets up a new array according to predefined parameters, and asks questions to gather the information necessary to build the array. If the configuration needs to go beyond these options, the Manual Configuration option allows additional control over logical drive setup parameters.

For detailed information on using GAM for configuration, see Chapter 4 of the Mylex *RAID Management Software CD-ROM* or the *Global Array Manager Client Software - Installation and User Guide*.

RAID Levels

RAID stands for Redundant Array of Inexpensive Disks. In a RAID storage system multiple disk drives are grouped into *arrays*. Each array is configured as a single system drive consisting of one or more disk drives.

Correct installation of the disk array and the controller requires a proper understanding of RAID technology and concepts. The controllers implement several versions of the Berkeley RAID technology, as summarized in Table 4-1.

Note: Although JBOD is not strictly a RAID level, it is included at various points in this discussion for comparison to RAID levels. It is sometimes referred to as RAID 7.

Table 4-1 Supported RAID Levels

RAID Level	Description	Minimum Drives	Maximum Drives	Fault-Tolerant
0	Block striping is provided, which yields higher performance than is possible with individual disk drives. No redundancy is provided.	2	8	No
1	Disk drives are paired and mirrored. All data is duplicated 100% on an equivalent disk drive.	2	2	Yes
3	Data is striped across several physical disk drives. Parity protection is used for data redundancy. This level provides a larger bandwidth for applications that process large files.	3	8	Yes
5	Data and parity information is striped across all physical disk drives. Parity protection is used for data redundancy.	3	8	Yes
0+1	Combination of RAID levels 0 and 1. Data is striped across several physical disk drives. This level provides redundancy through mirroring.	4	8	Yes
JBOD (7)	“Just a bunch of disks.” Each disk drive is operated independently like a normal disk drive, or multiple disk drives can be spanned and seen as a single large drive. This level does not provide data redundancy.	1	1	No
30 and 50	Allow storage capacity to be increased by a factor of 4 from RAID 3 and RAID 5 configurations by combining four RAID 3 or RAID 5 drive groups into a superdrive group across which data is striped. Each of the four drive groups must contain the same number of physical devices, use the same stripe block size, and be the same RAID level. The appended 0 denotes that the RAID level is using striping (RAID 0) across drive groups.	3	8	Yes

You must select an appropriate RAID level when you define or create system drives. This decision is based on how you prioritize the following:

- Disk capacity utilization (number of disk drives)
- Data redundancy (fault tolerance)
- Disk performance

The controllers make the RAID implementation and the disk drives' physical configuration transparent to the host operating system. This transparency means that the host operating logical drivers and software utilities are unchanged, regardless of the RAID level selected.

Although a system drive may have only one RAID level, RAID levels can be mixed within a drive pack (LUN), as illustrated in Figure 4-1.

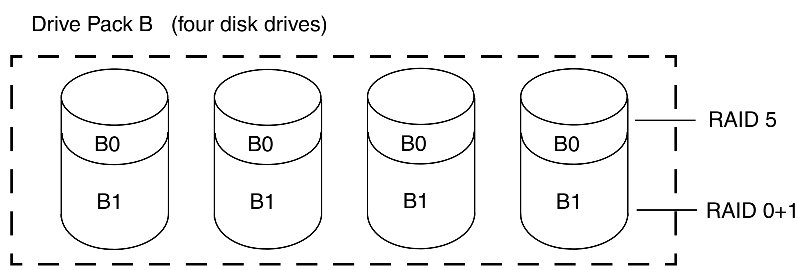

Figure 4-1 RAID Levels Within a Drive Pack (LUN): Example

In Figure 4-1, the smaller system drive (B0) is assigned a RAID 5 level of operation, while the larger system drive (B1) is assigned a RAID 0+1 level of operation.

Remember that different RAID levels exhibit different performance characteristics for a particular application or environment. The controller affords complete versatility in this regard by allowing multiple RAID levels to be assigned to a drive pack.

Drives are fault-tolerant when you use a RAID level providing redundancy. In the simplex configuration, however, if the controller or host bus adapter fails, the data is not accessible until the failure is corrected.

CAP Strategy for Selecting a RAID Level

Capacity, availability, and performance are three benefits, collectively known as CAP, that should characterize your expectations of the disk array subsystem.

It is impossible to configure an array optimizing all of these characteristics; that t is a limitation of the technology. For example, maximum capacity and maximum availability cannot exist in a single array. Some of the disk drives must be used for redundancy, which reduces capacity. Similarly, configuring a single array for both maximum availability and maximum performance is not an option.

The best approach is to prioritize requirements. Decide which benefit is most important for the operating environment. The controller in the TPS 9000 storage system is versatile enough to offer any of these preferences, either singly or in the most favorable combination possible.

The three benefits are further explained in these subsections:

- “Configuring for Maximum Capacity” on page 33
- “Configuring for Maximum Availability” on page 34
- “Configuring for Maximum Performance” on page 35

Configuring for Maximum Capacity

Table 4-2 shows the relationship between RAID levels and effective capacities offered for the quantity *X* disk drives of *N* capacity. As an example, it provides computed capacities for six disk drives of size 2 GB each.

Table 4-2 RAID Level Maximum Capacity

RAID Level	Effective Capacity	Example: Capacity in GB
0	X*N	6*2 = 12
1	(X*N)/2	6*2/2 = 6
3	(X-1)*N	(6-1)*2 = 10
5	(X-1)*N	(6-1)*2 = 10
0+1	(X*N)/2	(6*2)/2 = 6
JBOD	X*N	6*2 = 12

The greatest capacities are provided by RAID 0 and JBOD, with the entire capacity of all disk drives being used. Unfortunately, with these two solutions, there is no fault tolerance.

RAID 3 and RAID 5 give the next best capacity, followed by RAID 1 and RAID 0+1.

Configuring for Maximum Availability

Table 4-3 presents definitions of array operational conditions.

Table 4-3 Array Operating Conditions

Array Condition	Meaning
Normal (online)	The array is operating in a fault-tolerant mode, and can sustain a disk drive failure without data loss.
Critical	The array is functioning and all data is available, but the array cannot sustain a second disk drive failure without potential data loss.
Degraded	The array is functioning and all data is available, but the array cannot sustain a second disk drive failure without potential data loss. Additionally, a reconstruction or rebuild operation is taking place, reducing the performance of the array. The rebuild operation takes the array from a critical condition to a normal condition.
Offline	The array is not functioning. If the array is configured with a redundant RAID level, two or more of its member disk drives are not online. If the array is configured as a RAID 0 or JBOD, one or more of its member disk drives are not online.
Not fault-tolerant	No fault-tolerant RAID levels have been configured for any of the disk drives in the array.

You can achieve an additional measure of fault tolerance (or improved availability) with a hot spare, or standby disk drive. This disk drive is powered on but idle during normal array operation. If a failure occurs on a disk drive in a fault-tolerant set, the hot spare takes over for the failed disk drive, and the array continues to function in a fully fault-tolerant mode after it completes its automatic rebuild cycle. Thus the array can suffer a second disk drive failure after rebuild and continue to function before any disk drives are replaced.

Controller Cache and Availability

The RAID controller has a write cache of 128 MB. This physical memory is used to increase the performance of data retrieval and storage operations.

The controller can report to the operating system that a write is complete as soon as the controller receives the data. Enabling write cache (write-back cache) improves performance, but exposes the data to loss if a system crash or power failure occurs before the data in the cache is written to disk.

RAID Levels and Availability

Table 4-4 summarizes RAID levels offered by the RAID controller and the advantages (and disadvantages) of the RAID levels as they apply to availability.

Table 4-4 RAID Levels and Availability

RAID Level	Fault Tolerance Type	Availability
0	None	Data is striped across a set of multiple disk drives. If a disk drive in the set ceases to function, all data contained on the set of disk drives is lost. (This configuration is not recommended if fault tolerance is needed.)
1	Mirrored	Data is written to one disk drive, and then the same data is written to another disk drive. If either disk drive fails, the other one in the pair is automatically used to store and retrieve the data.
3 and 5	Striped	Data and parity are striped across a set of at least three disk drives. If any fail, the data (or parity) information from the failed disk drive is computed from the information on the remaining disk drives.
0+1	Mirrored and striped	Data is striped across multiple disk drives, and written to a mirrored set of disk drives.
JBOD	None	This configuration offers no redundancy and is not recommended for applications requiring fault tolerance.

Configuring for Maximum Performance

Table 4-5 presents the relative performance advantages of each RAID level.

Table 4-5 RAID Levels and Performance

RAID Level	Access Profile Characteristics
0	Excellent for all types of I/O activity
1	Excellent for write-intensive applications
3	Excellent for sequential or random reads and sequential writes
5	Excellent for sequential or random reads and sequential writes
0+1	Excellent for write-intensive applications
JBOD	Mimics normal, individual disk drive performance characteristics

Loop Configuration

After you have determined the RAID level to use, determine the loop configuration. Note the following:

- The largest RAID 3 or RAID 5 group that can be created is 7+1 drives.
- The maximum configuration is 12 drives total, those in the tower itself; no expansion to another tower or enclosure is possible.
- A maximum of 12 drives can be attached to a drive loop.
- A maximum of 8 system drives can be created (see “System Drives” on page 39 for information on these).

The disk drive modules are dual-ported. The RAID controller sees 12 drives on each loop (A and B), because it finds both ports of each drive. Via the I/O modules, it allocates the drives to each channel alternatingly, so that the drive addresses are available for failover. Thus six drives are on channel 0 via their A port and six are on channel 1 via their B port.

At initial startup, each I/O module controls a separate loop of six drives, as diagrammed in Figure 4-2.

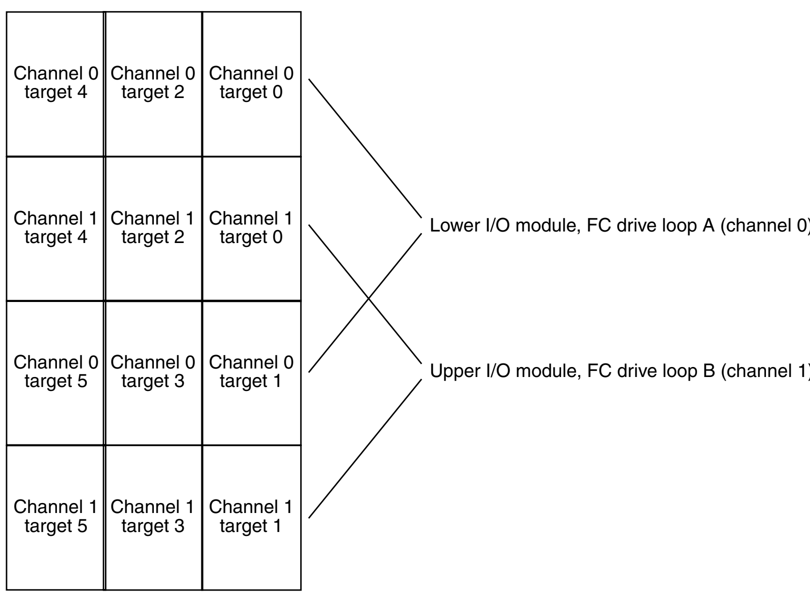

Figure 4-2 I/O Modules, Channels, and Loops

However, you can use GAM to reassign target drives in accordance with your CAP strategy to channels 0 and 1, which correspond to the I/O modules. Check and confirm if the controller parameters need to be modified for the intended application; see the documentation for the management software included with the storage system for information on controller parameters.

Note: Changes to the controller parameter settings take effect after the controller is rebooted.

Figure 4-3 diagrams an example relationship of the RAID controller, the I/O modules, and loops. In this configuration, all drives are on both channels for availability. In a configuration with two loops, if one loop fails, access is through the alternative path, which might affect performance.

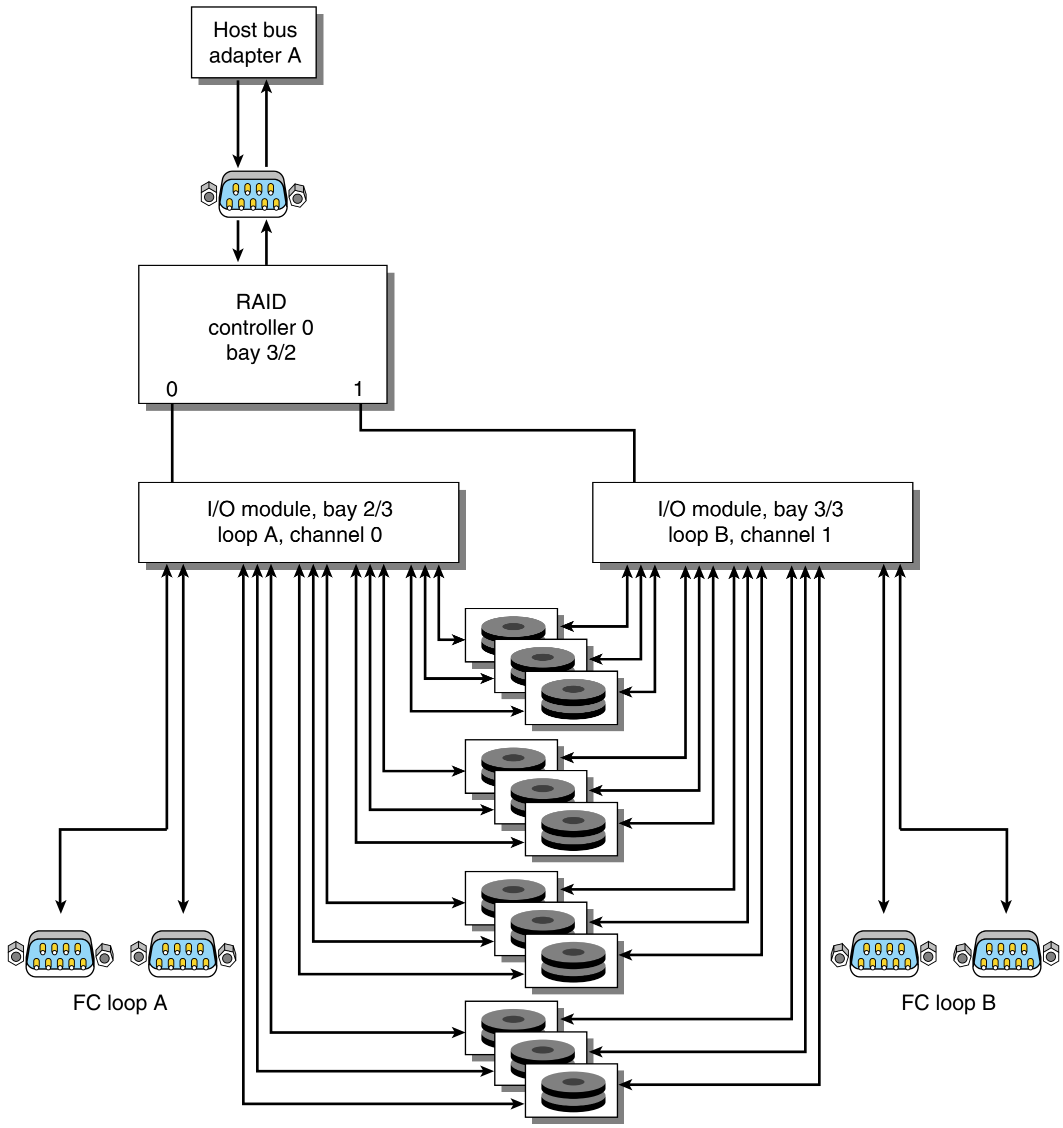

Figure 4-3 Single-Controller/Point-to-Point/Single-Host Example

In a configuration with two loops, if one loop fails, access is through the alternative path, which might affect performance.

System Drives

System drives are the logical devices that are presented to the operating system. During the configuration process, after physical disk drive packs are defined, one or more system drives must be created from the drive packs. This section discusses system drives in these subsections:

- “System Drive Properties” on page 39
- “System Drive Affinity and Programmable LUN Mapping” on page 39

System Drive Properties

System drives have the following properties:

- The minimum size of a system drive is 8 MB; the maximum size is 2 TB.
- Up to eight system drives can be created.
- Each system drive has a RAID level that is selectable (subject to the number of disk drives in the system drive’s pack).
- Each system drive has its own write policy (write-back or write-through); see “Data Caching” on page 27 in Chapter 3 for an explanation of this feature.
- Each system drive has its own LUN affinity. This capability is further discussed in “System Drive Affinity and Programmable LUN Mapping” on page 39.
- More than one system drive can be defined on a single drive pack.

System Drive Affinity and Programmable LUN Mapping

System drive affinity and programmable LUN mapping are configuration features that work together to define how the host accesses the available storage space.

System drive affinity allows system drives to be assigned to any combination of controller and host ports:

- Configurations with one RAID controller that has two host ports can use system drive affinity to define affinity of each system drive to one or both host ports.
- System drives that are not owned by a controller/host port are not accessible.

Programmable LUN mapping lets you assign any LUN ID (even multiple LUN IDs) to any system drive on each port, or configure system drive assignments without specifying the LUN, defaulting to the current mapping algorithm. System drives with the "all" affinity are mapped to a LUN ID on every controller/host port.

Drive State Reporting

The controller provides information to the configuration tool or utility to help it maintain operational information for each physical disk drive. The controller keeps a record of each physical disk drive's operational state and information on available target ID addresses. The controller first determines the presence or absence of the physical disk drive or available target ID location and then determines the operational state of the disk drives present.

Table 4-6 describes possible physical disk drive states. This information applies only to physical disk drives, not to system drives.

Table 4-6 Physical Disk Drive States

State	Description
Online	The disk drive is powered on, has been defined as a member of a drive pack, and is operating properly.
Standby	The disk drive is powered on, is able to operate properly, has not been defined as part of any drive pack, and has been defined as a standby or hot spare.
Dead	The disk drive is not present, is present but not powered on, is a newly inserted replacement disk drive, or fails to operate properly and is labeled Dead by the controller (whether or not it has been defined as a member of a drive pack).
Rebuild	The disk drive is in the process of being rebuilt, that is, data is being copied from the mirrored disk drive to the replacement disk drive in a RAID 1 array, or data is being regenerated via the exclusive OR (XOR) redundancy algorithm and written to the replacement disk drive in a RAID 3, 5, or 0+1 array.
Unconfigured	This location is unconfigured.
Environmental	An environmental device is present at this address.

For more information, see the GAM documentation and online help.

Automatic Rebuild

The controller provides automatic rebuild capabilities in the event of a physical disk drive failure. The controller performs a rebuild operation automatically when a disk drive fails and both of the following conditions are true:

- A standby or hot spare disk drive of identical or larger size is attached to the same controller.
- All system drives that are dependent on the failed disk drive are configured as a redundant array: RAID 1, RAID 3, RAID 5, or RAID 0+1.

Note: The controller always attempts first to locate a replacement (standby or hot spare) disk drive that is exactly the same size as the failed disk drive. If none is found, the controller attempts to locate a replacement disk drive that is at least the same size as the failed disk drive.

During the automatic rebuild process, storage system activity continues as normal; however, storage system performance may degrade slightly.

Note: The priority of rebuild activity can be adjusted using the GAM controller parameters to adjust the rebuild rate.

To use the automatic rebuild feature, you must maintain a hot spare or standby disk drive in the storage system. The number of hot spare disk drives in a storage system is limited only by the maximum number of disk drives available on each disk channel.

You can create a hot spare disk drive in one of two ways:

- A disk drive can be labeled as a spare or standby using the GAM disk information dialog.
- The configuration is created or changed using the GAM RAID Assist dialog.

The RAID controller supports hot swap disk drive replacement while the storage system is online: depending on the RAID level, a disk drive can be disconnected, removed, or replaced with another disk drive without taking the storage system offline.

Chapter 5

Troubleshooting

The TPS 9000 storage system includes a processor and associated monitoring and control logic that allows it to diagnose problems within the storage system's power, cooling, and drive systems.

SES (SCSI Enclosure Services) communications are used between the storage system and the RAID controllers. Status information on power, cooling, and thermal conditions is communicated to the controllers and is also displayed in the GAM management utilities.

The enclosure services processor is housed in the ESI/ops panel module. The sensors for power, cooling, and thermal conditions are housed within the power supply/cooling modules. Each module in the storage system is monitored independently.

This chapter consists of the following sections:

Note: If a fault is indicated on the ESI/ops panel, refer first to Table 5-1 in "Using Storage System LEDs for Troubleshooting" on page 45 and then to the section referenced within that table.

Solving Initial Startup Problems

If cords are missing or damaged, plugs are incorrect, or cables are too short, contact your supplier for a replacement.

If the RAID module fault LED is also on, there is a RAID controller fault; check that the RAID module is correctly seated.

If the alarm sounds when you power on the storage system, one of the following conditions exists:

- A fan is slowing down: see “ESI/Ops Panel Power Supply/Cooling Module LED: Faults and Remedies” on page 47 for further checks to perform.
- Voltage is out of range; the tower requires 115/220 volts (autoranging).
- There is an overtemperature or thermal overrun condition; see “ESI/Ops Panel Power Supply/Cooling Module LED: Faults and Remedies” on page 47.
- There is a storage system fault; see “ESI/Ops Panel System LED: Faults and Remedies” on page 47.

If the SGI server does not recognize the storage system, check the following:

- Make sure that the device driver for the host bus adapter board has been installed. If the HBA was installed at the factory, this software is in place; if not, check the HBA and the server documentation for information on the device driver.
- Make sure the FC-AL interface cables from the RAID module to the fibre channel board in the host computer are fitted correctly.
- Check the selector switches on the I/O modules in the storage system; both I/O modules should be set to 0.
- See if the LEDs on all installed drive carrier modules are green. Note that the drive LEDs flash during drive spinup.
- Check that all disk drive modules are correctly installed.
- Check that the RAID controller has completed its startup and the green controller ready LED is illuminated; see “RAID Controller LEDs and Problem Solutions” on page 50.

If an amber disk drive module LED drive fault is on, there is a drive fault: see Table 5-2 on page 51.

Using Storage System LEDs for Troubleshooting

This section summarizes LED functions and gives instructions for solving storage system problems in these subsections:

- "ESI/Ops Panel LEDs" on page 45
- "Power Supply/Cooling Module LED and Problem Solution" on page 48
- "FC LRC I/O Module LEDs and Problem Solutions" on page 49
- "RAID Controller LEDs and Problem Solutions" on page 50
- "Disk Drive LEDs and Problem Solutions" on page 51

ESI/Ops Panel LEDs

Figure 5-1 shows details of the ESI/ops panel.

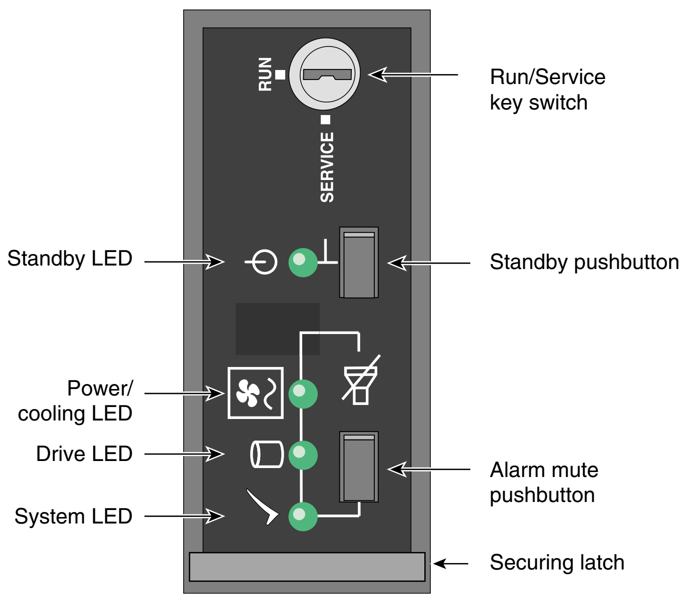

Figure 5-1 ESI/Ops Panel Indicators and Switches

Table 5-1 summarizes functions of the LEDs on the ESI/ops panel.

Table 5-1 ESI/Ops Panel LEDs and Troubleshooting

LED	Off	Green Steady	Green or Amber Flashing	Amber Steady	Remedy
System status	Power is off, or standby mode is selected.	All ESI functions OK	System configuration error	ESI / I^2C function failure	See “ESI/Ops Panel System LED: Faults and Remedies” on page 47.
Drive status	Power is off, or standby mode is selected.	All drives OK	Drive fault (alternating green and amber)	N/A	Replace the drive as explained in Chapter 6, “Installing and Replacing Disk Drive Modules”.
Power/ Cooling	Power is off, or standby mode is selected.	All power and fans OK	Not used	Power supply/cooling module or fan fault	See “ESI/Ops Panel Power Supply/Cooling Module LED: Faults and Remedies” on page 47. Contact service provider for replacement.
Standby	No AC power is supplied.	Power is good, or standby status is selected	Not used	Storage system in standby mode due to ESI thermal shutdown	See “Thermal Alarm” on page 54, “Thermal Shutdown” on page 54.

Note the following:

- If all LEDs on the ESI/ops panel flash simultaneously, see “Using Test Mode” on page 55.
- If test mode has been enabled (see “Using Test Mode” on page 55), the amber and green drive bay LEDs flash for any non-muted fault condition.
- The LED for the uninterruptible power supply (UPS) is covered with a sticker. This LED is lit whether or not a UPS is present. Use of a UPS is not supported in this release. Leave the sticker in place.

ESI/Ops Panel System LED: Faults and Remedies

If the system status LED on the ESI/ops panel is amber and the alarm is sounding, the ESI processor has detected an internal fault, such as failure of an internal communications path. Follow these steps:

1. Check if the LEDs on the power supply/cooling modules are amber. If a PSU error is present, there might be a communications problem with that power supply/cooling module. Contact your service representative for a replacement.
2. Check if the disk drive module LEDs are amber. If none are, there might be either an ESI processor problem or a chassis midplane problem. Contact your service provider.

See also “Thermal Shutdown” on page 54.

ESI/Ops Panel Power Supply/Cooling Module LED: Faults and Remedies

If the alarm is sounding, the power supply/cooling LED on the ESI/ops panel is amber, and the LED on one or both power supply/cooling modules is amber, the cause is one of the following:

- Power fault
- Fan failure
- Thermal condition that is causing the power supply/cooling module to overheat

To troubleshoot the problem, follow these steps:

1. Check the power connection to each power supply/cooling module.
2. If possible, reduce the ambient temperature. If the problem persists, contact your service provider.

If one of the power supply/cooling modules is missing or defective, all of the following occur:

- The alarm sounds
- The system LED on the ESI/ops panel flashes
- The power supply/cooling module LED remains illuminated

If these warnings occur, check the LEDs on the power supply/cooling modules. If one or both are amber, contact your service representative.

Power Supply/Cooling Module LED and Problem Solution

Figure 5-2 shows the location of the LED on the power supply/cooling module.

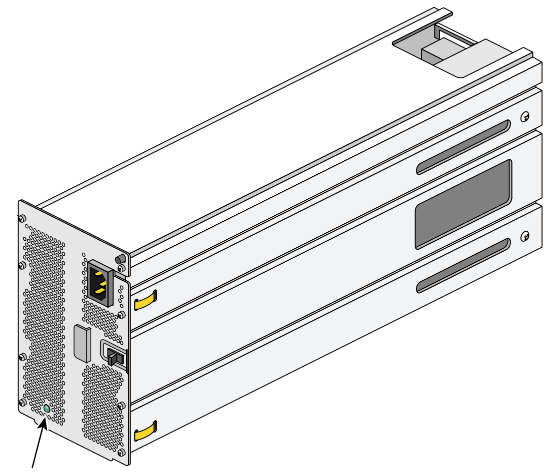

Figure 5-2 Power Supply/Cooling Module LED

If this LED is amber, or if the power/cooling LED on the ESI/ops panel is amber and the alarm is sounding, contact your service provider.

FC LRC I/O Module LEDs and Problem Solutions

Two LEDs between the DB-9 connectors on the I/O panel glow steady green to indicate a good FC-AL signal on the cables attached to the adjacent connectors. Figure 5-3 shows the LEDs on the panel.

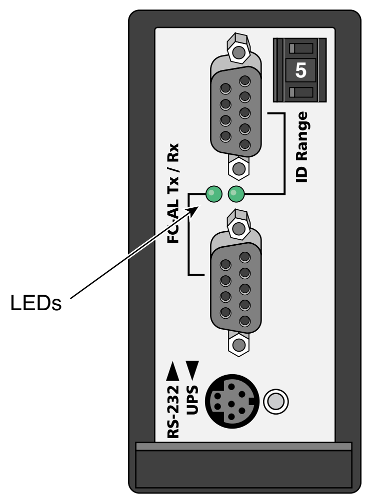

Figure 5-3 FC LRC I/O Module LEDs

If the LED for a connector is amber, check the connection. If the problem persists, contact your service provider for a replacement.

RAID Controller LEDs and Problem Solutions

Figure 5-4 shows the LEDs on the RAID controller panel.

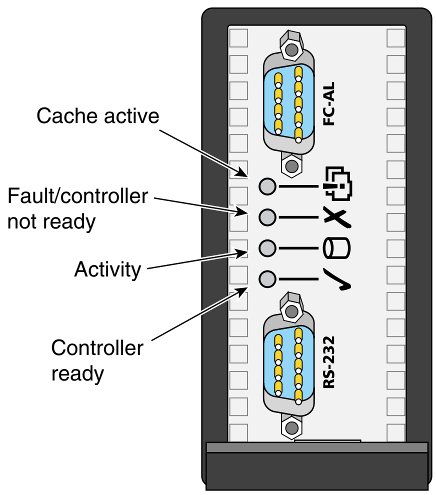

Figure 5-4 RAID Controller Panel LEDs

If the fault/controller not ready LED is on, check that the module is properly seated. If you cannot find the fault, contact your service provider.

Disk Drive LEDs and Problem Solutions

Each disk drive module has two LEDs, an upper (green) and a lower (amber), as shown in Figure 5-5.

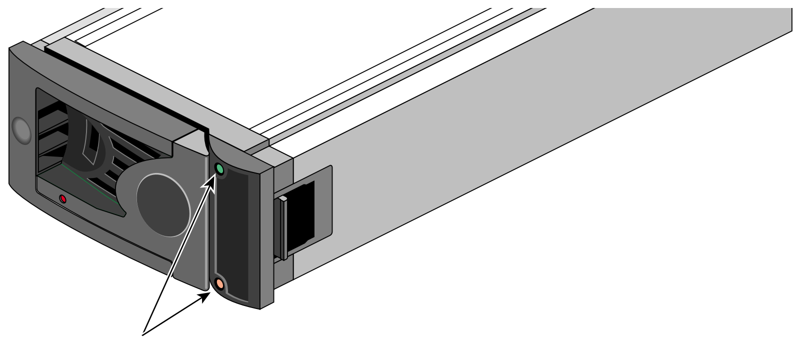

Figure 5-5 Disk Drive Module LEDs

The LEDs on the disk drive modules provide the information summarized in Table 5-2.

Table 5-2 Disk Drive LED Function

State	Green LED	Amber LED	Remedy
No disk drive fitted; that is, the drive is not fully seated	Off	Off	Check that drive is fully seated
Disk drive power is on, but the drive is not active	On	Off	N/A
Disk drive is active (LED might be off during power-on)	Blinking	Off	N/A
Disk drive fault (SES function)	Flashing at 2-second intervals	On	Contact your service provider for a replacement drive and follow instructions in Chapter 6.
Disk drive identify (SES function)	N/A	Flashing at half-second intervals	N/A

In addition, the amber drive LED on the ESI/ops panel alternates between on and off every 10 seconds when a drive fault is present.

Using the Alarm for Troubleshooting

The ESI/ops panel includes a n audible alarm that indicates when a fault state is present. The following conditions activate the audible alarm:

- RAID controller fault
- Drive fault
- Fan slows down
- Voltage out of range
- Overtemperature
- Thermal overrun
- Storage system fault

You can mute the audible alarm by pressing the alarm mute pushbutton for about a second, until you hear a double beep. The mute pushbutton is beneath the indicators on the ESI/ops panel (see Figure 5-1).

When the alarm is muted, it continues to sound with short intermittent beeps to indicate that a problem still exists. It is silenced when all problems are cleared.

Note: If a new fault condition is detected, the alarm mute is disabled.

Storage System Temperature Conditions, Problems, and Solutions

This section explains storage system temperature conditions and problems in these subsections:

- "Thermal Control" on page 53
- "Thermal Alarm" on page 54
- "Thermal Shutdown" on page 54

Thermal Control

The storage system uses extensive thermal monitoring and ensures that component temperatures are kept low and acoustic noise is minimized. Airflow is from front to rear of the storage system.

If the ambient air is cool (below 25° C or 77° F) and you can hear that the fans have sped up by their noise level and tone, then some restriction on airflow might be raising the storage system's internal temperature. The first stage in the thermal control process is for the fans to automatically increase in speed when a thermal threshold is reached. This might be a normal reaction to higher ambient temperatures in the local environment. The thermal threshold changes according to the number of drives and power supplies fitted.

If fans are speeding up, follow these steps:

1. Check that there is clear, uninterrupted airflow at the front and rear of the storage system.
2. Check for restrictions due to dust buildup; clean as appropriate.
3. Check for excessive recirculation of heated air from the rear of the storage system to the front.
4. Check that all blank plates and dummy disk drives are in place.
5. Reduce the ambient temperature.

Thermal Alarm

The thermal alarm is a combination of all these:

- The alarm sounds.
- The power supply/cooling LED on the ESI/ops panel is amber.
- The LED on one or both power supply/cooling modules is amber.
- The air temperature of the air exiting the PSU is above 55˚ C (131˚ F).

This alarm occurs when the internal temperature (measured in the airflow through the storage system exceeds a threshold. Follow these steps:

1. Check that local ambient environment temperature is below the upper temperature specification of 40˚ C (104˚ F).
2. Check that there is clear, uninterrupted airflow at the front and rear of the storage system.
3. Check for restrictions due to dust buildup; clean as appropriate.
4. Check for excessive recirculation of heated air from the rear to the front.

Thermal Shutdown

The storage system automatically shuts down to protect itself and the disk drives from damage. Thermal shutdown occurs:

- After the thermal alarm condition is present; shutdown takes place at a higher threshold than the thermal alarm
- When all fans fail
- When only one fan is operating and the storage system internal temperature reaches 40˚ C (104˚ F)

The storage system shuts down ten seconds after any of these conditions is present.

When a thermal shutdown is imminent, all amber LEDs on the ESI/ops panel and on all disk drives flash and the alarm sounds continuously and cannot be muted.

When thermal shutdown conditions are present, follow these steps (if possible):

1. Check for airflow restrictions.
2. Check for power supply/cooling module faults as detailed in “ESI/Ops Panel Power Supply/Cooling Module LED: Faults and Remedies” on page 47.
3. Check for excessive ambient temperatures (over 40˚ C (104˚ F))

If the overheating problem is not solved, thermal shutdown occurs. The standby LED on the ESI/ops panel and the power supply status LED are amber and the storage system powers itself off.

Follow these steps:

1. Correct the source of the overheating.
2. Allow the storage system to cool down.
3. Unplug the power cord from the storage system and leave it unplugged for at least 30 seconds to reset the shutdown condition.
4. Reconnect the power cord and restart the storage system following the normal procedure (see “Powering On the Storage System” on page 17 in Chapter 2).
5. Check for cooling faults that persist, particularly fan failure. If a fan has failed, contact your service provider for a replacement.

Using Test Mode

When no faults are present in the storage system, you can run test mode to check the LEDs and the audible alarm on the ESI/ops panel. In this mode, the amber and green LEDs on each of the drive carrier modules and the ESI/ops panel flash on and off in sequence.

To activate test mode, press the alarm mute pushbutton until you hear a double beep. The LEDs then flash until the storage system is reset, either when you press the alarm mute pushbutton again or if an actual fault occurs.

Chapter 6

Installing and Replacing Disk Drive Modules

This chapter explains how to replace a faulty disk drive module or install a new one, in the following sections:

Note: The RAID controller supports hot swap disk drive replacement while the storage system is online: depending on the RAID level, a disk drive can be disconnected, removed, or replaced with another disk drive without taking the storage system offline.

Adding a Disk Drive Module

Note the following:

- All disk drive bays must be filled with either a disk drive module or a dummy drive; no bay should be left completely empty.
- The drives in bays 1/3 and 4/3 are required for enclosure management; these bays must always be occupied.
- Install drives in the order shown in Figure 6-1. Generally, after the two required bays are occupied, drive bays are filled starting at the bottom of the storage system.

2 (required)	8	12
4	7	11
3	6	10
1 (required)	5	9

Figure 6-1 Drive Installation Order

Follow these steps to add a new disk drive module to the storage system:

1. Make sure that enough disk drives and dummy drives have been ordered to occupy all bays.
2. For each new drive, carefully open the bag containing the drive.

Warning: The disk drive handle might have come unlatched in shipment and might spring open when you open the bag. As you open the bag, keep it a safe distance from your face.

3. Place the disk drive modules on a work surface, and make sure that their antitamper locks are disengaged (unlocked). A disk drive module cannot be installed if its antitamper lock is activated outside the enclosure.

   Drives are shipped with the locks set in the unlocked position. However, if a drive is locked, insert the key (included with the disk drive) into the socket in the lower part of the handle trim and turn it 90 degrees clockwise until the indicator visible in the center aperture of the handle shows black. See Figure 6-2.

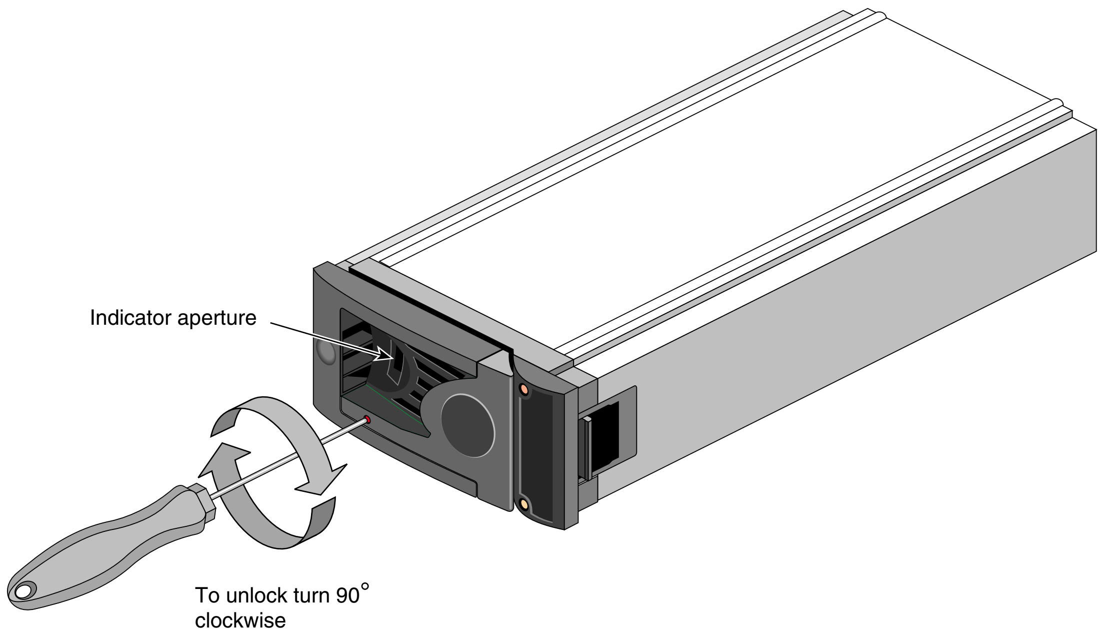

Figure 6-2 Unlocking the Disk Drive Module (Disengaging the Antitamper Lock)

4. Release the disk drive module handle by pressing on the small indentation on the left of the module front; see Figure 6-1.

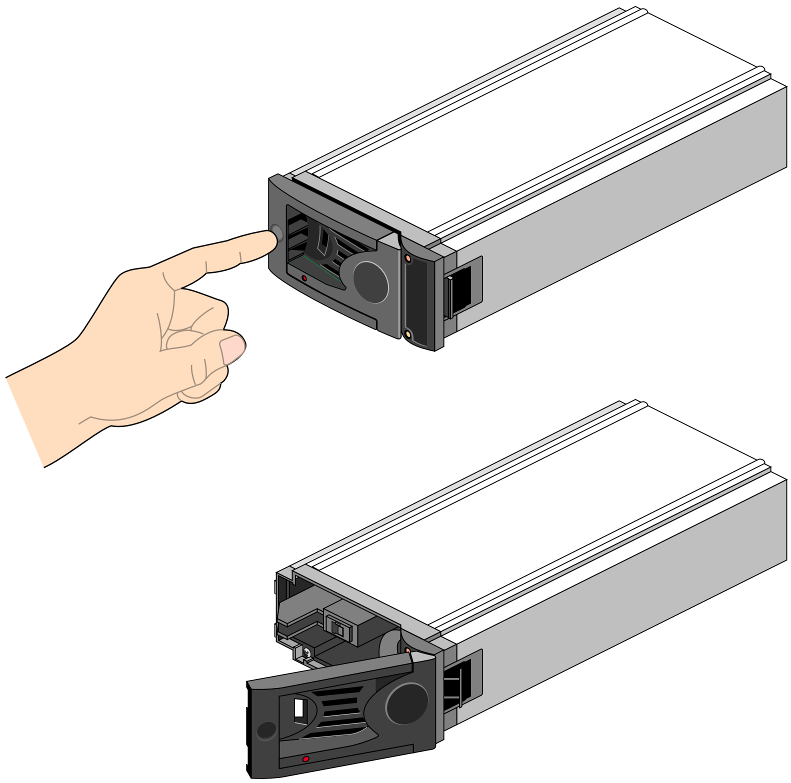

Figure 6-3 Releasing the Module Handle

5. Insert the disk drive module into the chassis; see Figure 6-4.

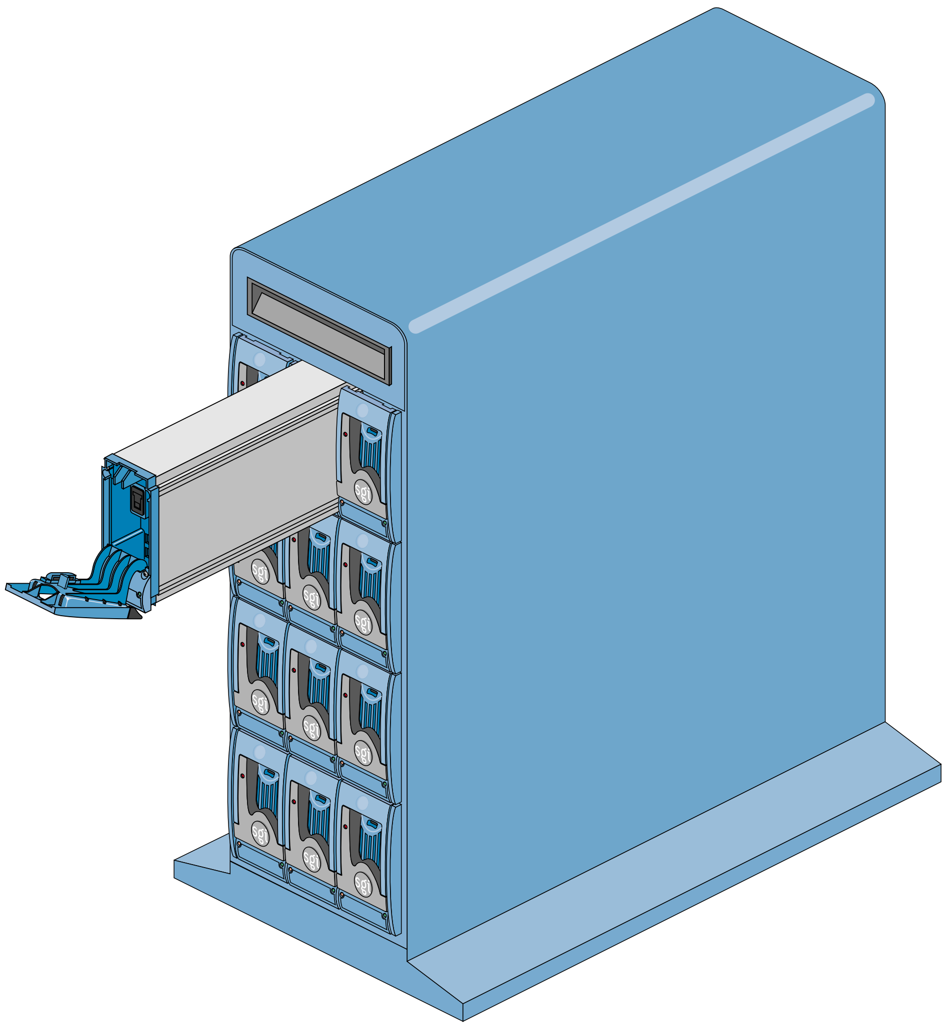

Figure 6-4 Inserting the Disk Drive Module

Note: The drive modules are not keyed. Make sure you are inserting the drive module in the correct orientation.

6. Push the disk drive module gently into the chassis until it is stopped by the camming lever on the right of the module; see Figure 6-5.

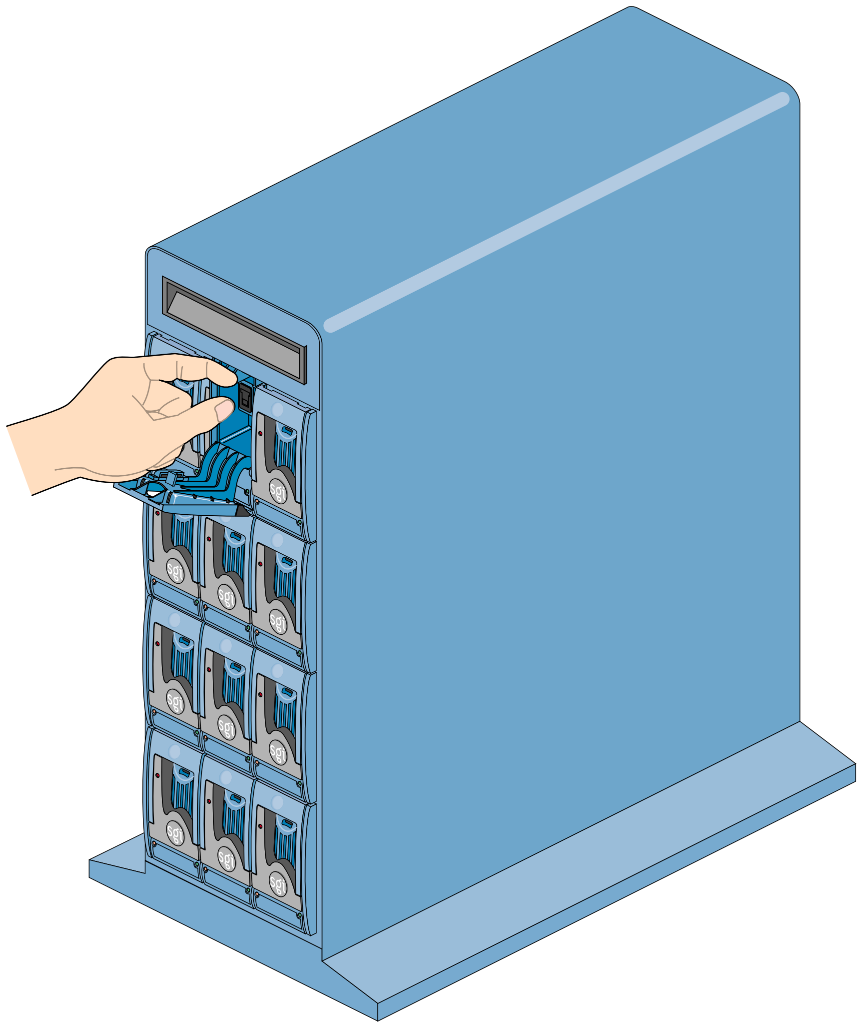

Figure 6-5 Inserting the Disk Drive Module: Drive Seated

7. Press on the drive handle to seat the drive. The camming lever on the right of the module engages with a slot in the chassis; see Figure 6-6.

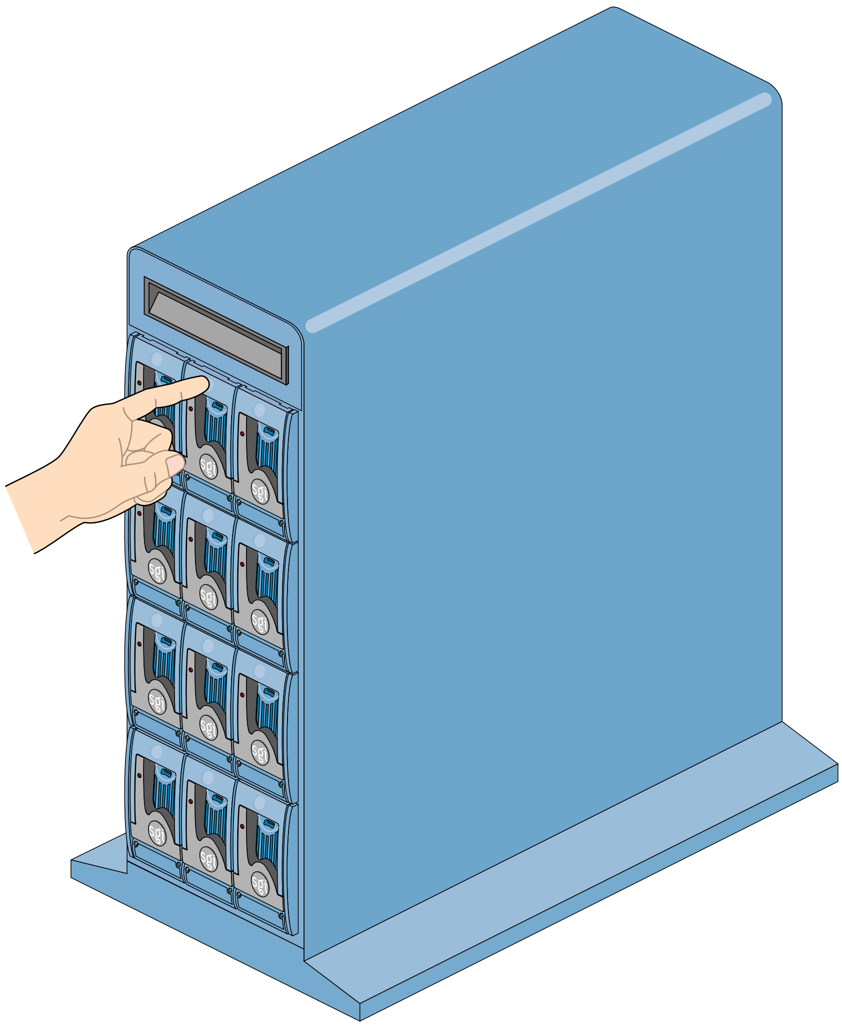

Figure 6-6 Inserting the Disk Drive Module: Engaging the Camming Lever

Continue to push firmly until the handle fully engages with the module cap. You should hear a click as the latch engages and holds the handle closed.

8. Repeat steps 1 through 7 for all drive modules to be installed.
9. When you have finished installing all drives, activate the drive antitamper lock on each one:
   - Carefully insert the lock key provided into the cutout in the handle trim. See Figure 6-7.

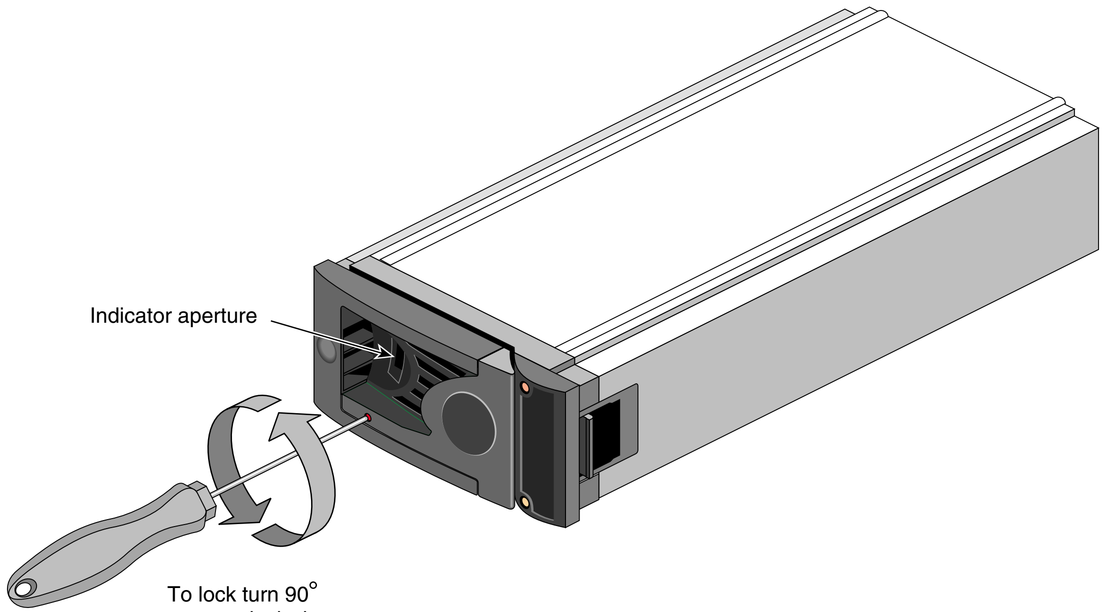

Figure 6-7 Locking the Disk Drive Module (Setting the Antitamper Lock)

   - Rotate the key 90 degrees counterclockwise until the indicator in the handle shows red.
   - Remove the key and store it.
10. Fit all empty drive bays with dummy drives. The drive handle and camming mechanism operate the same way as for disk drives.

Replacing a Disk Drive Module

This section explains how to replace a defective disk drive module, in the following sections

- "LUN Integrity and Disk Drive Module Failure" on page 65
- "Replacing the Disk Drive Module" on page 66

LUN Integrity and Disk Drive Module Failure

When a disk drive fails in a RAID 5, 3, 1, or 1/0 LUN, the green LEDs on all disks in the LUN (except the failed one) alternate on/off every 1.2 seconds until the fault condition is cleared. The amber LED on the failed disk remains lit.

For a RAID 5, 3, 1, or 0+1 LUN, you can replace the disk module without powering off the array or interrupting user applications. If the array contains a hot spare on standby, the controller automatically rebuilds the failed module on the hot spare. A RAID 0 must be taken offline to be replaced if a single disk module fails.

A hot spare is a special LUN that acts as a global disk spare that can be accessed by any RAID 5, 3, 1, or 1/0 LUN. A hot spare is unowned until it becomes part of a LUN when one of the LUN's disk modules fails.

Note: If a disk fails in a LUN and the storage system puts the hot spare into the LUN, use the software included with the storage system to check disk module status, and replace the failed disk as soon as possible. The replacement becomes the new hot spare; this arrangement (drive roaming) differs from that of other RAID systems. Therefore, it is important to keep track of the location of the hot spare.

In these cases, the LUN's data integrity is compromised and it becomes unowned (not accessible by the controller). After you replace the failed disk modules (one at a time), you delete and then recreate the affected LUN(s). If the data on the failed disks was backed up, restore it to the new disks.

Note: Before replacing a disk module, use the software included with the storage system to check disk module status.

Replacing the Disk Drive Module

If an LED indicates that a disk drive is defective, follow these steps to remove the faulty drive:

1. Make sure that enough disk drives and dummy drives are available to occupy all bays.
2. Make sure users are off the affected systems; back up data if necessary.
3. Using the key, turn the antitamper lock to the off position:

   Note: Replace disk drive modules one at a time,

   - Carefully insert the lock key provided into the cutout in the handle trim, as shown in Figure 6-2.
   - Rotate the key 90 degrees clockwise until the indicator in the handle shows black.
   - Remove the key and store it.
4. Make sure the faulty drive has spun down.

   Caution: Damage can occur to a drive if it is removed while still spinning.
5. Release the handle by pressing on the small indentation area on the left of the module front.
6. Gently withdraw the module approximately 25 mm (one inch) and wait 30 seconds.
7. Withdraw the module from the drive bay. Replace the drive immediately; follow instructions in “Adding a Disk Drive Module” on page 58.
8. If you are replacing a module in a LUN that uses a hot spare, note the location of the replacement module; it is the new hot spare.

Appendix A

Technical Specifications

This appendix contains the following sections:

Storage System Physical Specifications

Table A-1 gives the physical specifications for the tower.

Table A-1 Physical Specifications and Environmental Requirements

Dimension	Storage system
Height	57 cm (22.4 in.), including wheels
Width	30.2 cm (11.8 in.) at base 183 cm (71.9 in.) at top
Depth	59.2 cm (23.2 in.)
Weight	
Maximum configuration	About 56 kg (about l123 lb)
Power supply/cooling module	4 kg (8.8 lb)
ESI/Ops panel module	Under 1 kg (under 2.2 lb)
RAID controller module	About 0.8 kg (about 1.76 lb)

Table A-2 gives temperature and humidity requirements for both storage system configurations.

Table A-2 Ambient Temperature and Humidity Requirements

Factor	Temperature	Relative Humidity	Maximum Wet Bulb
Operating temperature	10˚ C to 40˚ C (50˚ F to 104˚ F)	20% to 80% noncondensing	23˚ C (73.4˚ F)
Nonoperating temperature	0˚ C to 50˚ C (32˚ F to 122˚ F)	8% to 80% noncondensing	27˚ C (80.6˚ F)
Shipping	-40˚ C to +60˚ C (-40˚ F to 140˚ F)	5% to 100% nonprecipitating	29˚ C (84.2˚ F)

Table A-3 gives other environmental requirements for the storage system.

Table A-3 Environmental Requirements

Factor	Requirement
Altitude	0 to 2133 m (6996.2 feet or 1.3 miles)
Operational shock	Vertical axis 5 g peak 1/2 sine, 10ms
Nonoperational shock	20 g 20 ms square wave
Operational vibration	Random vibration power spectrum available on request
Acoustics	Freestanding enclosure declared ‘A’ weighted sound power level equal to or less than 68 deciBels

Table A-4 gives minimum storage system power requirements.

Table A-4 Minimum Power Requirements

Factor	Requirement
Voltage	100 to 120 or 200 to 240 VAC
Frequency	47 to 63 Hz
Maximum power consumption	800 VA
Typical power consumption	400 VA or less
Inrush current (25º C (77˚ F) cold start 1 PSU)	40/80 A @ 110/230VAC

RAID Controller Module Specifications

Table A-5 gives specifications for the RAID controller module.

Table A-5 RAID Module Specifications

Factor	Requirement
Controller type	Mylex Corp. DACFFX
CPU	Two Intel 960RD RISC 32 bit microprocessors
Host interface connectors	One DB-9 FCAL with MIA support, up to 100MB/sec
External FC-AL signal cables	SGI dual-port HBAs: 25 m (82 ft) Storage area network (SAN) and SGI single-port HBAs: maximum 12 m (39.36 ft) copper or 100 m (328 ft) optical See Table A-7 for information on cables
Drive interface	2 x FC-AL loops, connected internally to FCAL LRC I/O
Mode of operation	Simplex: single controller
Current limit	1.5 A for MIA
RAID levels	0, 1, 3, 5, and 0+1 10, 30, and 50 JBOD (RAID level 7)
Configuration and management	Via FC-AL using Global Array Manager (GAM) utility
Indicators	LEDs for Controller Ready, Activity, Failed, and Cache
Memory	Default 128 MB; approved memory options available
Cache	Selectable write-through or write-back Read always enabled
Battery	NiCd cache battery protects 128 MB data for up to 72 hours

Disk Drive Module Specification

Please refer to your supplier for details of disk drives which are supported for use with the RAID storage system. Table A-6 gives specifications for a typical drive carrier module.

Table A-6 Drive Carrier Module Specifications (1.6-Inch 18 GB Drive)

Factor	Requirement
Weight	1.2 kg (2.64 lb)
Operating temperature	10˚ C to 40˚ C (50˚ F to 104˚ F) when installed
Power dissipation	28 watts maximum

SGI Cables for the SGI 9000 Storage System

Table A-7 lists SGI cable options that can be connected to the SGI 9000 product.

Table A-7 SGI Fibre Channel Fabric Cabling Options for SGI 9000

Cable	Length	Marketing Code	Part Number
FC copper cable with DB-9 connector at each end	1 m (39.3 in)	X-F-COP-1M	TBD
	10 m (32.8 ft)	X-F-COP-10M	018-0570-00x
	25 m (82 ft)	X-F-COP-25M	018-0571-00x
FC copper cable, high-speed style-2 balanced cable connector and DB-9 connector	3 m (9.8 ft)	X-FS-COP-3M	N/A
	12 m (39.36 ft)	X-FS-COP-12M	N/A
FC optical cable (62.5 µm)	3 m (9.8 ft)	X-F-OPT-3M	018-0656-001
	10 m (32.8 ft)	X-F-OPT-10M	018-0656-101
	25 m (82 ft)	X-F-OPT-25M	018-0656-201
	100 m (328 ft)	X-F-OPT-100M	018-0656-301
	300 m (980 ft)[a]	X-F-OPT-300M	018-0656-401
Two media interface adapter (MIA) modules (optical-electrical converters)	N/A	X-F-OE-KIT	9980952

a. This cable is not authorized for use with SGI fibre channel switches.

Index

A

B

C

D

E

F

G

H

I

J

K

L

M

O

P

R

W

Tell Us About This Manual

As a user of Silicon Graphics products, you can help us to better understand your needs and to improve the quality of our documentation.

Any information that you provide will be useful. Here is a list of suggested topics:

- General impression of the document
- Omission of material that you expected to find
- Technical errors
- Relevance of the material to the job you had to do
- Quality of the printing and binding

Please send the title and part number of the document with your comments. The part number for this document is 007-4068-001.

Thank you!

Three Ways to Reach Us

- To send your comments by **electronic mail**, use either of these addresses:
  - On the Internet: techpubs@sgi.com
  - For UUCP mail (through any backbone site): *[your_site]*!sgi!techpubs
- To **fax** your comments (or annotated copies of manual pages), use this fax number: 650-932-0801
- To send your comments by **traditional mail**, use this address:

  Technical Publications
  Silicon Graphics, Inc.
  1600 Amphitheatre Parkway, M/S 535
  Mountain View, California 94043-1351