Sun Cluster 3.x Hardware Administration Manual

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054
U.S.A.

Part No: 817–0168–10
April 2003, Revision A

030714@5943

Contents

Preface

The *Sun Cluster 3.x Hardware Administration Manual* provides a variety of information about how to install and administer basic Sun™ Cluster hardware components. Topics covered in this book include how to install and configure terminal concentrators, the cluster interconnect, public network hardware, campus clustering, and dynamic reconfiguration.

Use this manual with any version of Sun Cluster 3.x software. Unless otherwise noted, procedures are the same for all Sun Cluster 3.x versions. See the "Revision History" on page 8 for a list of changes to this manual.

This book does not include information about configuring servers in a Sun Cluster environment nor does it include specific storage device procedures.

Who Should Use This Book

This book is for Sun representatives who are performing the initial installation of a Sun Cluster configuration and for system administrators who are responsible for maintaining the system.

This document is intended for experienced system administrators with extensive knowledge of Sun software and hardware. Do not use this document as a planning or presales guide. You should have already determined your system requirements and purchased the appropriate equipment and software before reading this document.

How This Book Is Organized

The following chapters contain information about hardware used in a Sun Cluster environment.

Chapter 1 provides an overview of installing and administering Sun Cluster hardware.

Chapter 2 describes how to install and configure a terminal concentrator.

Chapter 3 describes how to install and maintain cluster interconnect hardware.

Chapter 4 describes how to install and maintain the public network hardware.

Appendix A provides guidelines and diagrams on how to configure a campus cluster.

Appendix B describes how to verify cluster redundancy.

Appendix C provides procedures on how to recable disks.

Revision History

The following table lists the information that has been revised or added since the initial release of this documentation. The table also lists the revision date for these changes.

TABLE P–1 Sun Cluster 3.x Hardware Administration Manual

Revision Date	New Information
July 2003	VLAN support expanded to campus cluster configurations. See "Configuring VLANs as Private Interconnect Networks" on page 53.
	Updated `scswitch` command.
	Added SAN section.
	Included references to the Sun Fire V1280 system administration documentation for dynamic reconfiguration. See"Dynamic Reconfiguration Operations For Sun Cluster Nodes" on page 16.

Related Documentation

The following books provide conceptual information or procedures to administer hardware and applications. If you plan to use this documentation in a hard-copy format, ensure that you have the following books available for your reference.

The following Sun Cluster books support the Sun Cluster 3.1 release. If you are maintaining a different version of Sun Cluster software, refer to the appropriate documentation. All Sun Cluster documentation is available on `http://docs.sun.com`.

Documentation that is not available on `http://docs.sun.com` is listed with the appropriate URL.

TABLE P–2 Hardware Documentation

Title	Part Number
Sun StorEdge SAN Foundation 4.2 Installation Guide	817-1244
Available on `http://www.sun.com/products-n-solutions/hardware/docs/Network_Storage_Solutions/SAN`	
Sun StorEdge SAN Foundation 4.2 Configuration Guide	817-1245
Available on `http://www.sun.com/products-n-solutions/hardware/docs/Network_Storage_Solutions/SAN`	
Sun StorEdge SAN Foundation 4.2 Guide to Documentation	817-1576
Available on `http://www.sun.com/products-n-solutions/hardware/docs/Network_Storage_Solutions/SAN`	
Sun StorEdge SAN Foundation 4.2 Release Notes	817-1246
Available on `http://www.sun.com/products-n-solutions/hardware/docs/Network_Storage_Solutions/SAN`	
IP Network Multipathing Administration Guide	816-0850
Sun Gigabit Ethernet/P 2.0 Adapter Installation and User's Guide	804-6140
Available on `http://www.sun.com/products-n-solutions/hardware/docs`	
Sun StorEdge Long Wave GBIC Interface Converter Guide	805-6965
Available on `http://www.sun.com/products-n-solutions/hardware/docs`	

TABLE P–2 Hardware Documentation *(Continued)*

Title	Part Number
Sun StorEdge Traffic Manager Installation and Configuration Guide Available on `http://www.sun.com/products-n-solutions/hardware/docs`	816-1420

TABLE P–3 Sun Cluster Documentation

Application	Title	Part Number
Concepts	*Sun Cluster 3.1 Concepts Guide*	816-3383
Hardware	*Sun Cluster 3.x Hardware Administration Manual*	817-0168
	Sun Cluster 3.x Hardware Administration Collection at `http://docs.sun.com/db/coll/1024.1/`	
Software Installation	*Sun Cluster 3.1 Software Installation Guide*	816-3388
Data Services	*Sun Cluster 3.1 Data Service Planning and Administration Guide*	817–1526
	Sun Cluster 3.1 Data Service Collection at `http://docs.sun.com/db/coll/573.10/`	
API Development	*Sun Cluster 3.1 Data Services Developer's Guide*	816-3385
Administration	*Sun Cluster 3.1 System Administration Guide*	816-3384
Error Messages	*Sun Cluster 3.1 Error Messages Guide*	816-3382
Man Pages	*Sun Cluster 3.1 Man Page Reference Manual*	816–5251
Release Notes	*Sun Cluster 3.1 Release Notes*	816-5317
	Sun Cluster 3.1 Release Notes Supplement	816-3381

Using UNIX Commands

This document contains information on commands used to install, configure, or upgrade a Sun Cluster configuration. This document might not contain complete information on basic UNIX® commands and procedures such as shutting down the system, booting the system, and configuring devices.

See one or more of the following sources for this information.

- Online documentation for the Solaris software environment
- Other software documentation that you received with your system

- Solaris operating environment man pages

Getting Help

If you have problems installing or using Sun Cluster, contact your service provider and provide the following information.

- Your name and email address (if available)
- Your company name, address, and phone number
- The model number and serial number of your systems
- The release number of the operating environment (for example, Solaris 8)
- The release number of Sun Cluster (for example, Sun Cluster 3.0)

Use the following commands to gather information on your system for your service provider.

Command	Function
`prtconf -v`	Displays the size of the system memory and reports information about peripheral devices
`psrinfo -v`	Displays information about processors
`showrev -p`	Reports which patches are installed
`prtdiag -v`	Displays system diagnostic information
`/usr/cluster/bin/scinstall -pv`	Displays Sun Cluster release and package version information

Also have available the contents of the `/var/adm/messages` file.

Accessing Sun Documentation Online

The docs.sun.com^{SM} Web site enables you to access Sun technical documentation online. You can browse the docs.sun.com archive or search for a specific book title or subject. The URL is `http://docs.sun.com`.

Typographic Conventions

The following table describes the typographic changes used in this book.

TABLE P–4 Typographic Conventions

Typeface or Symbol	Meaning	Example
`AaBbCc123`	The names of commands, files, and directories; on-screen computer output	Edit your `.login` file. Use `ls -a` to list all files. `machine_name% you have mail.`
`AaBbCc123`	What you type, contrasted with on-screen computer output	`machine_name%` **`su`** `Password:`
AaBbCc123	Command-line placeholder: replace with a real name or value	To delete a file, type **`rm`** *filename*.
AaBbCc123	Book titles, new words, or terms, or words to be emphasized.	Read Chapter 6 in *User's Guide*. These are called *class* options. You must be *root* to do this.

Shell Prompts in Command Examples

The following table shows the default system prompt and superuser prompt for the C shell, Bourne shell, and Korn shell.

TABLE P–5 Shell Prompts

Shell	Prompt
C shell prompt	`machine_name%`
C shell superuser prompt	`machine_name#`
Bourne shell and Korn shell prompt	`$`
Bourne shell and Korn shell superuser prompt	`#`

CHAPTER 1

Introduction to Sun Cluster Hardware

This chapter provides overview information on cluster hardware. The chapter also provides overviews of the tasks that are involved in installing and maintaining this hardware *specifically in a Sun™ Cluster environment*.

This chapter contains the following information:

Installing Sun Cluster Hardware

The following procedure lists the tasks for installing a cluster and the sources for instructions.

1. **Plan for cluster hardware capacity, space, and power requirements.**

   For more information, see the site planning documentation that shipped with your servers and other hardware.

2. **Install the nodes.**

   For server installation instructions, see the documentation that shipped with your servers.

3. **Install the administrative console.**

   For more information, see the documentation that shipped with your administrative console.

4. **Install a console access device.**

   Use the procedure that is indicated for your type of console access device. For example, Sun Enterprise E10000 servers use a System Service Processor (SSP) as a console access device, rather than a terminal concentrator.

   For installation instructions, see "Installing the Terminal Concentrator" on page 19 or the documentation that shipped with your Sun Enterprise E10000 hardware.

5. **Install the cluster interconnect and public network hardware.**

   For installation instructions, see Chapter 3.

6. **Install and configure the storage arrays.**

   Perform the service procedures that are indicated for your type of storage hardware.

7. **Install the Solaris operating environment and Sun Cluster software.**

   For more information, see Sun Cluster software installation documentation .

8. **Plan, install, and configure resource groups and data services.**

   For more information, see the Sun Cluster data services collection.

Maintaining Sun Cluster Hardware

Sun Cluster 3.x Hardware Administration Manual augments documentation that ships with your hardware components by providing information on maintaining the hardware *specifically in a Sun Cluster environment*. Table 1–1 describes some of the differences between maintaining cluster hardware and maintaining standalone hardware.

TABLE 1–1 Sample Differences Between Servicing Standalone and Cluster Hardware

Task	Standalone Hardware	Cluster Hardware
Shutting down a node	Use the `shutdown`(1M) command.	To perform an orderly node shutdown, first use the `scswitch(1M)` command to switch device groups and resource groups to another node. Then shut down the node by running the `shutdown(1M)` command.

TABLE 1–1 Sample Differences Between Servicing Standalone and Cluster Hardware *(Continued)*

Task	Standalone Hardware	Cluster Hardware
Adding a disk	Run `boot -r` or `devfsadm`(1M)to assign a logical device name to the disk. You also need to run volume manager commands to configure the new disk if the disks are under volume management control.	Use the `devfsadm(1M)`, `scgdevs(1M)`, and `scdidadm(1M)` commands. You also need to run volume manager commands to configure the new disk if the disks are under volume management control.
Adding a transport adatper or public network adapter	Perform an orderly node shutdown, then install the public network adapter. After you install the network adapter, update the `/etc/hostname.adapter` and`/etc/inet/hosts` files.	Perform an orderly node shutdown, then install the public network adapter. After you install the public network adapter, update the `/etc/hostname.adapter` and`/etc/inet/hosts` files. Finally, add this public network adapter to an IP Network Multipathing group.

Powering On and Off Sun Cluster Hardware

Consider the following when powering on and powering off cluster hardware.

- Use shut down and boot procedures in your Sun Cluster system administration documentation for nodes in a running cluster.
- Use the power-on and power-off procedures in the manuals that shipped with the hardware *only* for systems that are newly installed or are in the process of being installed.

Caution – After the cluster is online and a user application is accessing data on the cluster, do not use the power-on and power-off procedures listed in the manuals that came with the hardware.

Dynamic Reconfiguration Operations For Sun Cluster Nodes

The Sun Cluster environment supports Solaris 8 dynamic reconfiguration (DR) operations on qualified servers. Throughout the *Sun Cluster 3.x Hardware Administration Collection*, there are procedures that require that the user add or remove transport adapters or public network adapters in a cluster node. Contact your service provider for a list of storage arrays that are qualified for use with DR-enabled servers.

Note – Review the documentation for the Solaris 8 DR feature on your hardware platform *before* you use the DR feature with Sun Cluster software. All of the requirements, procedures, and restrictions that are documented for the Solaris 8 DR feature also apply to Sun Cluster DR support (except for the operating environment quiescence operation).

Documentation for DR on currently qualified server platforms are listed here.

- *Sun Enterprise 10000 Dynamic Reconfiguration User Guide*
- *Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual*
- *Sun Fire 6800, 4810, 4800, and 3800 Systems Dynamic Reconfiguration User Guide*
- *Sun Fire 6800, 4810, 4800, and 3800 Systems Dynamic Reconfiguration Release Notes*
- *Sun Fire 15K Dynamic Reconfiguration (DR) User Guide*
- *Sun Fire 15K Dynamic Reconfiguration Release Notes*
- *Sun Fire 880 Dynamic Reconfiguration User's Guide*
- *Sun Fire V1280/Netra 1280 System Administration Guide*

▼ DR Operations in a Cluster With DR-Enabled Servers

Some procedures within the *Sun Cluster 3.x Hardware Administration Collection* instruct the user to shut down and power off a cluster node before you add, remove, or replace a transport adapter or a public network adapter (PNA).

However, if the node is a server that is enabled with the DR feature, the user does *not* have to power off the node before you add, remove, or replace the transport adapter or PNA. Instead, do the following:

1. **Follow the procedure steps in** *Sun Cluster 3.x Hardware Administration Collection*, **including any steps for disabling and removing the transport adapter or PNA from the active cluster interconnect.**

   See the Sun Cluster system administration documentation for instructions about how to remove transport adapters or PNAs from the cluster configuration.

2. **Skip any step that instructs you to power off the node, where the purpose of the power-off is to add, remove, or replace a transport adapter or PNA.**

3. **Perform the DR operation (add, remove, or replace) on the transport adapter or PNA.**

4. **Continue with the next step of the procedure in** *Sun Cluster 3.x Hardware Administration Collection***.**

   For conceptual information about Sun Cluster support of the DR feature, see your Sun Cluster concepts documentation document.

Local and Multihost Disks in a Sun Cluster Environment

Two sets of storage arrays reside within a cluster: local disks and multihost disks.

- Local disks are directly connected to a single node and hold the Solaris operating environment for each node.
- Multihost disks are connected to more than one node and hold client application data and other files that need to be accessed from multiple nodes.

For more conceptual information on multihost disks and local disks, see Sun Cluster concepts documentation.

Removable Media in a Sun Cluster Environment

Removable media include tape and CD-ROM drives, which are local devices. *Sun Cluster 3.x Hardware Administration Manual* does not contain procedures for adding, removing, or replacing removable media as highly available storage arrays. Although

tape and CD-ROM drives are global devices, these drives do not have more than one port and do not have multi-initiator firmware support that would enable these devices as highly available. Thus, *Sun Cluster 3.x Hardware Administration Manual* focuses on disk drives as global devices.

Although tape and CD-ROM drives cannot be highly available at this time, in a cluster environment, you can access tape and CD-ROM drives that are not local to your system. All the various density extensions (such as `h`, `b`, `l`, `n`, and `u`) are mapped so that the tape drive can be accessed from any node in the cluster.

Install, remove, replace, and use tape and CD-ROM drives as you would in a noncluster environment. For procedures about how to install, remove, and replace tape and CD-ROM drives, see the documentation that shipped with your hardware.

SAN Solutions in a Sun Cluster Environment

You cannot have a single point of failure in a SAN configuration that is in a Sun Cluster environment. For information on how to install and configure a SAN configuration, see your SAN documentation.

CHAPTER 2

Installing and Configuring the Terminal Concentrator

This chapter provides the hardware and software procedures for installing and configuring a terminal concentrator as a console access device in a Sun Cluster environment. This chapter also includes information on how to use a terminal concentrator.

This chapter contains the following procedures:

For conceptual information on console access devices, see your Sun Cluster concepts documentation.

Installing the Terminal Concentrator

This section describes the procedure for installing the terminal concentrator hardware and for connecting cables from the terminal concentrator to the administrative console and to the cluster nodes.

▼ How to Install the Terminal Concentrator in a Cabinet

This procedure provides step-by-step instructions for rack-mounting the terminal concentrator in a cabinet. For convenience, you can rack-mount the terminal concentrator even if your cluster does not contain rack-mounted nodes.

- To rack-mount your terminal concentrator, go to the first step of the following procedure.
- If you do not want to rack-mount your terminal concentrator, place the terminal concentrator in its standalone location, connect the unit power cord into a utility outlet, and go to "How to Connect the Terminal Concentrator" on page 24.

1. **Install the terminal concentrator bracket hinge onto the primary cabinet:**

   a. **Locate the bracket hinge portion of the terminal concentrator bracket assembly (see Figure 2–1).**

   b. **Loosely install two locator screws in the right-side rail of the rear of the cabinet.**

      Thread the screws into holes 8 and 29, as shown in Figure 2–1. The locator screws accept the slotted holes in the hinge piece.

   c. **Place the slotted holes of the hinge over the locator screws, and let the hinge drop into place.**

   d. **Install the screws into holes 7 and 28.**

      Tighten these screws, and the screws in holes 8 and 29, as shown in Figure 2–1.

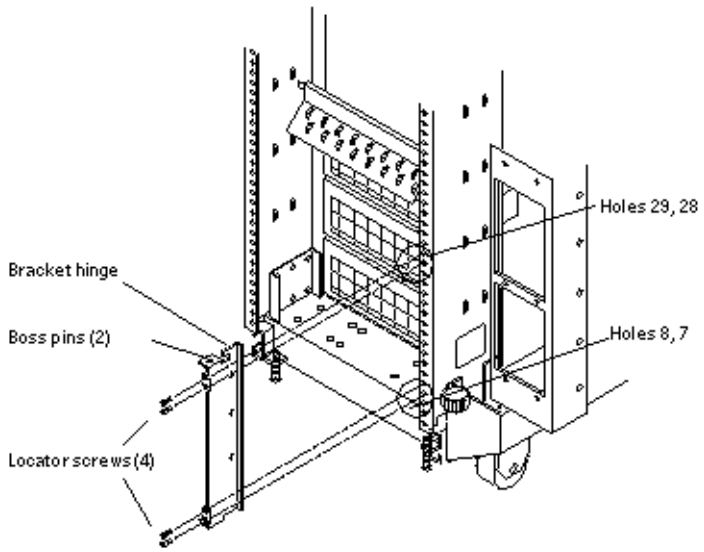

FIGURE 2–1 Installing the Terminal Concentrator Bracket Hinge to the Cabinet

2. **Install the terminal concentrator into the bracket.**

   a. **Place the side pieces of the bracket against the terminal concentrator, as shown in Figure 2–2.**

   b. **Lower the terminal concentrator (with side pieces) onto the bottom plate, aligning the holes in the side pieces with the threaded studs on the bottom plate.**

   c. **Install and tighten three nuts on the three threaded studs that penetrate through each side plate.**

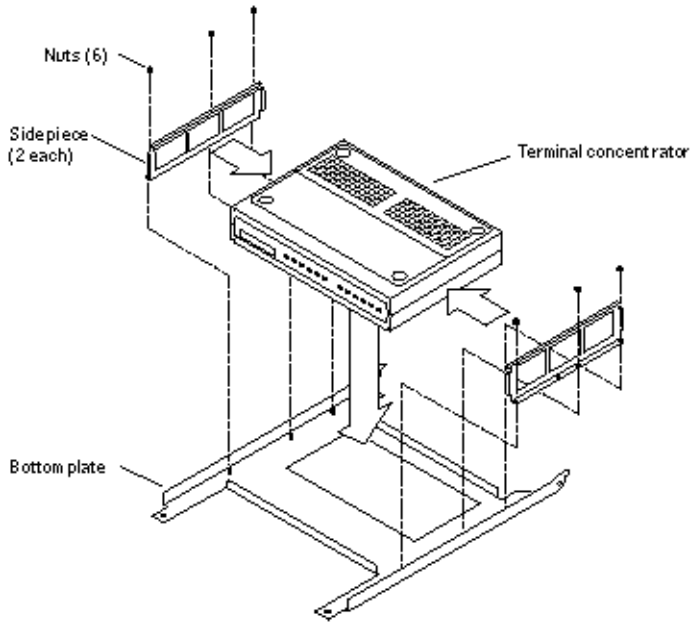

FIGURE 2–2 Installing the Terminal Concentrator Into the Bracket

3. **Install the terminal concentrator bracket onto the bracket hinge that is already installed on the cabinet.**

   a. **Turn the terminal concentrator bracket on its side so the hinge holes and cable connectors face toward the bracket hinge (see Figure 2–3).**

   b. **Align the bracket holes with the boss pins on the bracket hinge and install the bracket onto the hinge.**

   c. **Install the keeper screw in the shorter boss pin to ensure the assembly cannot be accidentally knocked off the hinge.**

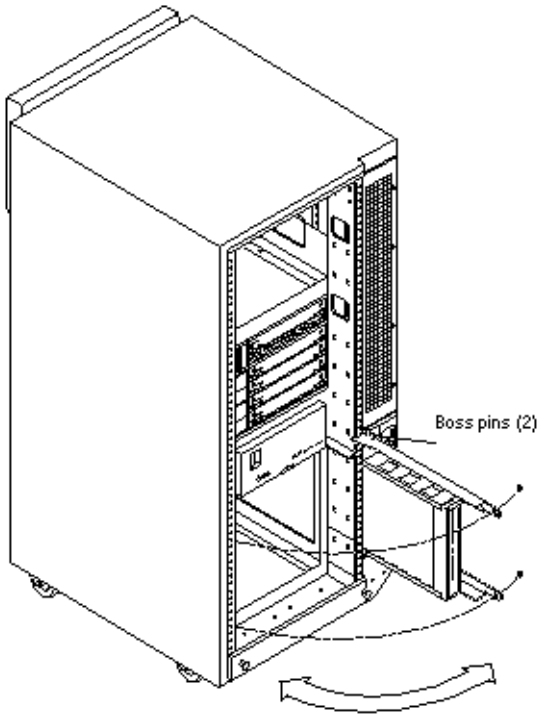

FIGURE 2–3 Terminal Concentrator Bracket Installed on the Hinge

4. **Connect one end of the power cord to the terminal concentrator, as shown in Figure 2–4. Connect the other end of the power cord to the power distribution unit.**

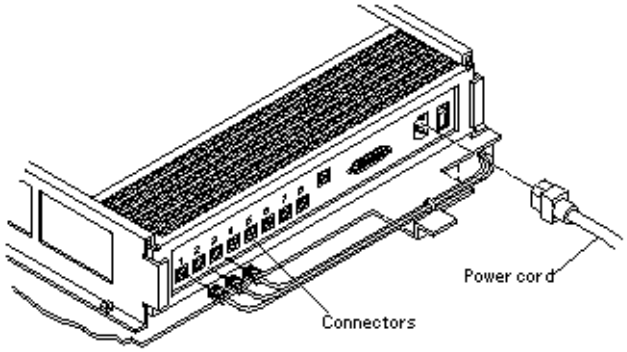

FIGURE 2–4 Terminal Concentrator Cable Connector Locations

Where to Go From Here

To cable the terminal concentrator, go to "How to Connect the Terminal Concentrator" on page 24.

▼ How to Connect the Terminal Concentrator

1. **Connect a DB-25 to RJ-45 serial cable (part number 530-2152-01 or 530-2151-01) from serial port A on the administrative console to serial port 1 on the terminal concentrator, as shown in Figure 2–5.**

   This cable connection from the administrative console enables you to configure the terminal concentrator. You can remove this connection after you set up the terminal concentrator.

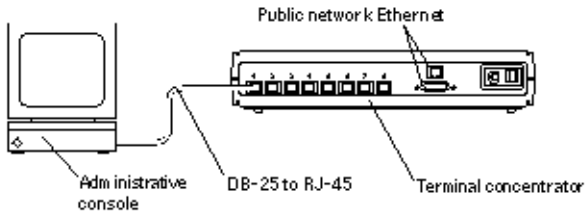

FIGURE 2–5 Connecting the Administrative Console

2. **Connect the cluster nodes to the terminal concentrator by using DB-25 to RJ-45 serial cables.**

   The cable connections from the concentrator to the nodes enable you to access the `ok` prompt or OpenBoot™ PROM (OBP) mode by using the Cluster Console windows from the Cluster Control Panel (CCP). For more information on how to use the CCP, see your Sun Cluster system administration documentation.

3. **Connect the public network Ethernet cable to the appropriate connector on the terminal concentrator.**

   Note – The terminal concentrator requires a 10-Mbit/sec Ethernet connection.

   ---

4. **Close the terminal concentrator bracket, and install screws in holes 8 and 29 on the left-side rear rail of the cabinet (see Figure 2–3).**

Where to Go From Here

Go to “Configuring the Terminal Concentrator” on page 25.

Configuring the Terminal Concentrator

This section describes the procedure for configuring the terminal concentrator’s network addresses and ports.

▼ How to Configure the Terminal Concentrator

1. **From the administrative console, add the following entry to the `/etc/remote` file.**

```
tc:\
     :dv=/dev/term/a:br#9600:
```

2. **Verify that the server and the terminal concentrator are powered on and that the cabinet keyswitch (if applicable) is in the ON position.**

3. **Establish a connection to the terminal concentrator’s serial port:**

```
# tip tc
```

4. **Hold down the terminal concentrator Test button (Figure 2–6) until the power LED flashes (about three seconds), then release the Test button.**

5. **Hold down the terminal concentrator Test button again for one second, then release it.**

   The terminal concentrator performs a self-test, which lasts about 30 seconds. Messages display on the administrative console. If the network connection is not found, press the Q key to stop the message.

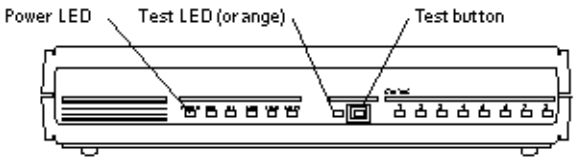

FIGURE 2–6 Terminal Concentrator Test Button and LEDs

6. **Observe the terminal concentrator front-panel LEDs:**

- If the front-panel LEDs light up as shown in Table 2–1, and the administrative console displays a `monitor::` prompt, go to Step 7.
- If the front-panel LEDs do not light up as shown in Table 2–1, or the administrative console does not display a `monitor::` prompt, use Table 2–2 and the documentation that shipped with your terminal concentrator to troubleshoot the problem.

TABLE 2–1 Front-Panel LEDs: Indicating a Successful Boot or Monitor Mode Reset

Power (Green)	Unit (Green)	Net (Green)	Attn (Amber)	Load (Green)	Active (Green)	Test (Orange)
ON	ON	ON	ON	OFF	Intermittent blinking	ON

TABLE 2–2 Front-Panel LEDs: Indicating a Failed Boot

Mode	Power (Green)	Unit (Green)	Net (Green)	Attn (Amber)	Load (Green)	Active (Green)
Hardware failure	ON	Blinking	OFF	Blinking	OFF	OFF
Network test failure	ON	ON	Blinking	OFF	OFF	Intermittent blinking
Network test aborted, or net command failed	ON	ON	OFF	Blinking	OFF	Intermittent blinking
Booted wrong image	ON	ON	ON	Blinking	OFF	OFF
Other failure	One or more Status LEDs (1-8) are ON					

7. **Use the `addr` command to assign an IP address, subnet mask, and network address to the terminal concentrator.**

   In the following example (Class B network, Class C subnet), the broadcast address is the terminal concentrator's address with the host portion set to 255 (all binary 1's).

```
monitor:: addr
Enter Internet address [<uninitialized>]:: 172.25.80.6
 Internet address: 172.25.80.6
Enter Subnet mask [255.255.0.0]:: 255.255.255.0
 Subnet mask: 255.255.255.0
Enter Preferred load host Internet address [<any host>]:: 172.25.80.6
```

```
*** Warning: Load host and Internet address are the same ***
 Preferred load host address: 172.25.80.6
Enter Broadcast address [0.0.0.0]:: 172.25.80.255
 Broadcast address: 172.25.80.255
Enter Preferred dump address [0.0.0.0]:: 172.25.80.6
 Preferred dump address: 172.25.80.6
Select type of IP packet encapsulation (ieee802/ethernet) [<ethernet>]::
         Type of IP packet encapsulation: <ethernet>
Load Broadcast Y/N [Y]:: n
            Load Broadcast: N
```

8. **After you finish the `addr` session, power-cycle the terminal concentrator.**

   The Load and Active LEDs should briefly blink, then the Load LED should turn off.

9. **Use the `ping`(1M) command to confirm that the network connection works.**

10. **Exit the `tip` utility by pressing Return and typing a tilde, followed by a period.**

```
<Return>~.
~
[EOT]
#
```

Where to Go From Here

Go to "How to Set Terminal Concentrator Port Parameters" on page 27.

▼ How to Set Terminal Concentrator Port Parameters

This procedure explains how to determine if the port type variable must be set and how to set this variable.

The port type parameter must be set to `dial_in`. If the parameter is set to `hardwired`, the cluster console might be unable to detect when a port is already in use.

1. **Locate the Sun serial number label on the top panel of the terminal concentrator (Figure 2–7).**

2. **Check if the serial number is in the lower serial-number range. The serial number consists of 7 digits, followed by a dash and 10 more digits.**

   - If the numbers after the dash start with 9520 or higher, the port type variable is set correctly. Go to Step 4.
   - If the numbers after the dash start with 9519 or lower, you must change the port type variable. Go to Step 3.

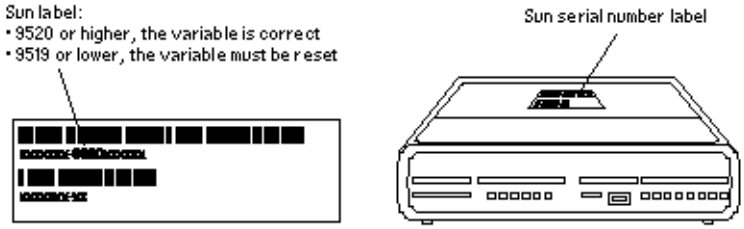

FIGURE 2–7 Determining the Version From the Serial Number Label

3. **Use the administrative console to change the port type variable to `dial_in` by setting the port parameters, then reboot the terminal concentrator as shown in the following example.**

   The boot command causes the changes to take effect. The terminal concentrator is unavailable for approximately one minute.

```
admin-ws# telnet tc_name
Trying terminal concentrator IP address
Connected to tc_name
Escape character is "^]".
Rotaries Defined:
    cli                             -
Enter Annex port name or number: cli
Annex Command Line Interpreter  *  Copyright 1991 Xylogics, Inc.
annex: su
Password: password
(The default password is the terminal concentrator IP address)
annex# admin
Annex administration MICRO-XL-UX R7.0.1, 8 ports
admin : set port=1-8 type dial_in imask_7bits Y
  You may need to reset the appropriate port, Annex subsystem or
        reboot the Annex for changes to take effect.
admin : set port=1-8 mode slave
admin : quit
annex# boot
bootfile:  <return>
warning:   <return>
```

Note – Ensure that the terminal concentrator is powered on and has completed the boot process before you proceed.

---

4. **Verify that you can log in from the administrative console to the consoles of each node.**

   For information on how to connect to the nodes' consoles, see "How to Connect to a Node's Console Through the Terminal Concentrator" on page 32.

▼ How to Correct a Port Configuration Access Error

A misconfigured port that does not accept network connections might return a `Connect: Connection refused` message when you use `telnet`(1). Use the following procedure to correct the port configuration.

1. **Connect to the terminal concentrator without specifying a port.**

   ```
   # telnet tc_name
   ```

   tc_name Specifies the hostname of the terminal concentrator

2. **Press Return again after you make the connection, then specify the port number.**

   ```
   Trying ip_address ..
   Connected to 192.9.200.1
   Escape character is "^]".
   ...
   [RETURN]
   Rotaries Defined:
         cli                              -
   Enter Annex port name or number: 2
   ```

   - If you see a `Port(s) busy, do you wish to wait? (y/n)` message, answer **`N`** and go to "How to Reset a Terminal Concentrator Port" on page 33.
   - If you see an `Error: Permission denied` message, the port mode is configured incorrectly to the command-line interface and must be set to slave. Go to Step 3.

3. **Select the terminal concentrator's command-line interface.**

   ```
   ...
   Enter Annex port name or number: cli
   annex:
   ```

4. **Type the `su` command and password.**

   The default password is the terminal concentrator's IP address.

   ```
   annex: su
   Password:
   ```

5. **Reset the port.**

   ```
   annex# admin
   Annex administration MICRO-XL-UX R7.0.1, 8 ports
   admin: port 2
   admin: set port mode slave
       You may need to reset the appropriate port, Annex subsystem or
       reboot the Annex for changes to take effect.
   admin: reset 2
   ```

Example—Correcting a Terminal Concentrator Port Configuration Access Error

The following example shows how to correct an access error on the terminal concentrator port 4.

```
admin-ws# telnet tc1
Trying 192.9.200.1 ...
Connected to 192.9.200.1.
Escape character is '^]'.
[Return]
Enter Annex port name or number: cli
...
annex: su
Password: root_password
annex# admin
Annex administration MICRO-XL-UX R7.0.1, 8 ports
admin: port 4
admin: set port mode slave
    You may need to reset the appropriate port, Annex subsystem or
    reboot the Annex for changes to take effect.
admin: reset 4
```

▼ How to Establish a Default Route for the Terminal Concentrator

---

Note – This procedure is optional. By setting a default route, you prevent possible problems with routing table overflows (see the following paragraphs). Routing table overflow is not a problem for connections that are made from a host that resides on the same network as the terminal concentrator.

A routing table overflow in the terminal concentrator can cause network connections to be intermittent or lost altogether. Symptoms include connection timeouts and routes that are reestablished, then disappear, even though the terminal concentrator itself has not rebooted.

The following procedure fixes this problem by establishing a default route within the terminal concentrator. To preserve the default route within the terminal concentrator, you must also disable the `routed` feature.

1. **Connect to the terminal concentrator.**

   # **telnet** *tc_name*

   tc_name Specifies the name of the terminal concentrator

2. **Press Return again after you make the connection, then select the command-line interface to connect to the terminal concentrator.**

```
...
Enter Annex port name or number: cliannex:
```

3. **Type the `su` command and password.**

   The default password is the terminal concentrator's IP address.

```
annex: su
Password:
```

4. **Start the editor to change the `config.annex` file.**

```
annex# edit config.annex
```

Note – The keyboard commands for this editor are Control-W:save and exit, Control-X:exit, Control-F:page down, and Control-B:page up.

   The `config.annex` file, which is created in the terminal concentrator's EEPROM file system, defines the default route. The `config.annex` file can also define rotaries that enable a symbolic name to be used instead of a port number.

5. **Add the following lines to the file.**

   Substitute the appropriate IP address for your default router.

```
%gateway
net default gateway 192.9.200.2 metric 1 active ^W
```

6. **Disable the local `routed` feature.**

```
annex# admin set annex routed n
```

7. **Reboot the terminal concentrator.**

```
annex# boot
bootfile: <reboot>
warning: <return>
```

   While the terminal concentrator is rebooting, you cannot access the node consoles.

Example—Establishing a Default Route for the Terminal Concentrator

The following example shows how to establish a default route for the terminal concentrator.

```
admin-ws# telnet tc1
Trying 192.9.200.1 ...
```

```
Connected to 192.9.200.1.
Escape character is '^]'.
[Return]
Enter Annex port name or number: cli
...
annex: su
Password: root_password
annex: edit config.annex
(Editor starts)
Ctrl-W:save and exit Ctrl-X:exit Ctrl-F:page down Ctrl-B:page up
%gateway
net default gateway 192.9.200.2 metric 1 active ^W
annex# admin set annex routed n
You may need to reset the appropriate port, Annex subsystem or
reboot the Annex for changes to take effect.
annex# boot
```

Using the Terminal Concentrator

This section describes the procedures about how to use the terminal concentrator in a cluster.

TABLE 2–3 Task Map: Using the Terminal Concentrator

Task	For Instructions, Go To
Connect to a node's console through the terminal concentrator	"How to Connect to a Node's Console Through the Terminal Concentrator" on page 32
Reset a terminal concentrator port	"How to Reset a Terminal Concentrator Port" on page 33

▼ How to Connect to a Node's Console Through the Terminal Concentrator

The following procedure enables remote connections from the administrative console to a cluster node's console by first connecting to the terminal concentrator.

1. **Connect to a node by starting a session with the terminal concentrator port that the node is cabled to.**

   # **telnet** *tc_name tc_port_number*

   tc_name Specifies the name of the terminal concentrator.

tc_*port_number* — Specifies the port number on the terminal concentrator. Port numbers are configuration dependent. Typically, ports 2 and 3 (5002 and 5003) are used for the first cluster that is installed at a site.

Note – If you set up node security, you are prompted for the port password.

---

2. **Log in to the node's console.**

   After establishing the telnet connection, the system prompts you for the login name and password.

3. **Set the terminal type, based on the type of window that was used in Step 1.**

```
# TERM=xterm
# export TERM
```

Example—Connecting to a Node's Console Through the Terminal Concentrator

The following example shows how to connect to a cluster node in a configuration that uses a terminal concentrator. A Shell tool has already been started by using an `xterm` window.

```
admin-ws# telnet tc1 5002
Trying 192.9.200.1 ...
Connected to 192.9.200.1.
Escape character is '^]'.
[Return]
pys-palindrome-1 console login: root
password: root_password
(for sh or ksh)
phys-palindrome-1# TERM=xterm; export TERM
(for csh)
phys-palindrome-1# set term=xterm
```

▼ How to Reset a Terminal Concentrator Port

When a port on the terminal concentrator is busy, you can reset the port to disconnect its user. This procedure is useful if you need to perform an administrative task on the busy port.

A busy port returns the following message when you try to connect to the terminal concentrator.

```
telnet: Unable to connect to remote host: Connection refused
```

If you use the port selector, you might see a `port busy` message. See "How to Correct a Port Configuration Access Error" on page 29 for details on the `port busy` message.

1. **Connect to the terminal concentrator port.**

   ```
   # telnet tc_name
   ```

   tc_name Specifies the name of the terminal concentrator

2. **Press Return again after you make the connection and select the command-line interface to connect to the terminal concentrator.**

   ```
   ...
   Enter Annex port name or number: cli
   annex:
   ```

3. **Type the `su` command and password.**

   The default password is the terminal concentrator's IP address.

   ```
   annex: su
   Password:
   ```

4. **Determine which port to reset.**

   The `who` command shows ports that are in use.

   ```
   annex# who
   ```

5. **Reset the port that is in use.**

   ```
   annex# admin reset port_number
   ```

6. **Disconnect from the terminal concentrator.**

   ```
   annex# hangup
   ```

   You can now connect to the port.

Example—Resetting a Terminal Concentrator Connection

The following example shows how to reset the terminal concentrator connection on port `2`.

```
admin-ws# telnet tc1
Trying 192.9.200.1 ...
Connected to 192.9.200.1.
Escape character is '^]'.
[Return]
...
Enter Annex port name or number: cli
...
annex: su
```

```
Password: root_password
annex: who
Port    What    User    Location        When    Idle    Address
2       PSVR    ---    ---            ---    1:27    192.9.75.12
v1        CLI    ---    ---            ---            192.9.76.10
annex# admin reset 2
annex# hangup
```

CHAPTER 3

Installing and Maintaining Cluster Interconnect Hardware

This chapter describes the procedures to install and maintain cluster interconnect hardware. Where appropriate, this chapter includes separate procedures for the two supported varieties of Sun Cluster interconnect: Ethernet, peripheral component interconnect-scalable coherent interface (PCI-SCI), and Sun Fire Link.

This chapter contains information about how to maintain cluster interconnect hardware. This chapter contains the following procedures.

For conceptual information about cluster interconnects, see your Sun Cluster concepts documentation.

For information about how to administer cluster interconnects, see your Sun Cluster system administration documentation.

Installing Ethernet Cluster Interconnect Hardware

Table 3–1 lists procedures about how to install Ethernet cluster interconnect hardware. Perform the procedures in the order that the procedures are listed. This section contains a procedure about how to install cluster hardware during an *initial installation* of a cluster, before you install Sun Cluster software.

TABLE 3–1 Task Map: Installing Ethernet Cluster Interconnect Hardware

Task	For Instructions, Go To
Install the transport adapters.	The documentation that shipped with your nodes and host adapters
Install the cluster transport cables. For clusters with more than two nodes, also install the transport junctions.	"How to Install Ethernet Transport Cables and Transport Junctions" on page 38

▼ How to Install Ethernet Transport Cables and Transport Junctions

Use this procedure to install Ethernet transport cables and transport junctions (switches).

1. **If not already installed, install transport adapters in your cluster nodes.**

   For the procedure about how to install transport adapters, see the documentation that shipped with your host adapters and node hardware.

2. **Install the transport cables. If necessary, install transport junctions.**

   - As Figure 3–1 shows, a cluster with only two nodes can use a point-to-point connection, requiring no transport junctions. Use a point-to-point (crossover) Ethernet cable if you are connecting 100BaseT or TPE ports of a node directly to ports on another node. Gigabit Ethernet uses the standard fiber-optic cable for both point-to-point and switch configurations.

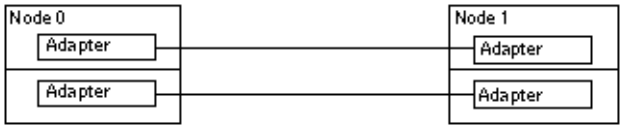

FIGURE 3–1 Typical Two-Node Cluster Interconnect

Note – If you use a transport junction in a two-node cluster, you can add additional nodes to the cluster. You can add additional nodes without bringing the cluster offline to reconfigure the transport path.

- As Figure 3–2 shows, a cluster with more than two nodes require two transport junctions. These transport junctions are Ethernet switches (customer-supplied).

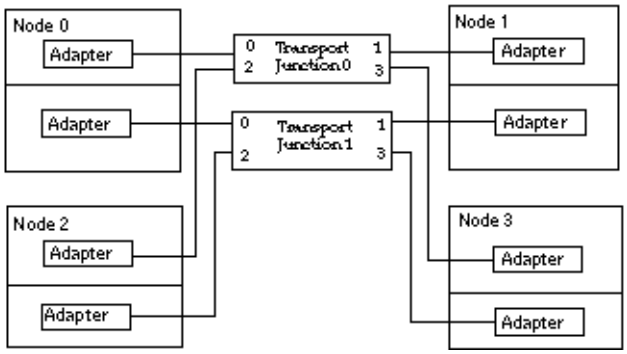

FIGURE 3–2 Typical Four-Node Cluster Interconnect

Where to Go From Here

Install the cluster software. Configure the interconnect after you install all other hardware. To review the task map about how to install cluster hardware and software, see Table 3–1.

Installing PCI-SCI Cluster Interconnect Hardware

Table 3–2 lists procedures about how to install PCI-SCI cluster interconnect hardware. Perform the procedures in the order that the procedures are listed. This section contains a procedure about how to install cluster hardware during an *initial installation* of a cluster before you install Sun Cluster software.

TABLE 3–2 Task Map: Installing PCI-SCI Cluster Interconnect Hardware

Task	For Instructions, Go To
Install the PCI-SCI transport cables. For four-node clusters, also install PCI-SCI transport junction.	"How to Install PCI-SCI Transport Cables and Transport Junction" on page 40

▼ How to Install PCI-SCI Transport Cables and Transport Junction

Use this procedure to install PCI-SCI transport cables and transport junctions (switches).

When you perform this procedure, the following error messages display on your console.

- Solaris 8

```
Nov 13 20:11:43 e04a ip: ip_rput_dlpi(scid0): DL_ERROR_ACK for
DL_ENABMULTI_REQ(29), errno 7, unix 0
Nov 13 20:11:43 e04a ip: ip: joining multicasts failed (7) on scid0
 - will use link layer broadcasts for multicast
```

  These error messages display because the SCI `dlpi` driver does not support the multicast feature. These error messages display when the `ip` module probes the driver. Sun Cluster software does not use the multicast feature on the private interconnect. You can safely ignore these error messages.

- Solaris 9

```
Dec  4 17:40:14 e03a in.routed[132]: write(rt_sock) RTM_ADD 172.17.0.128/25
-->172.17.0.129    metric=0 flags=0: File exists
Dec  4 17:40:19 e03a in.routed[132]: interface scid0 to 172.17.0.129 broken:
in=0 ierr=0 out=0 oerr=4
```

  These error messages are responses to the way Solaris 9 handles SCI `dlpi` interfaces. Solaris 9 uses the `in.routed` routing protocol as the default routing protocol. You can safely ignore these error messages. The `in.routed` routing

protocol is the source of these error messages.

1. **If not already installed, install PCI-SCI transport adapters in your cluster nodes.**

   For the procedure about how to install PCI-SCI transport adapters and set their DIP switches, see the documentation that shipped with your PCI-SCI host adapters and node hardware.

Note – Sbus-SCI host adapters are not supported by Sun Cluster software. If you are upgrading from a Sun Cluster 2.2 cluster, remove any Sbus-SCI host adapters from the cluster nodes. If you do not remove these adapters, you might see panic error messages during the SCI self test.

---

2. **Install the PCI-SCI transport cables and optionally, transport junctions, depending about how many nodes are in your cluster.**

   - A two-node cluster can use a point-to-point connection, requiring no transport junction. See Figure 3–3.

     Connect the ends of the cables marked `SCI Out` to the `O` connectors on the adapters.

     Connect the ends of the cables marked `SCI In` to the `I` connectors of the adapters as shown in Figure 3–3.

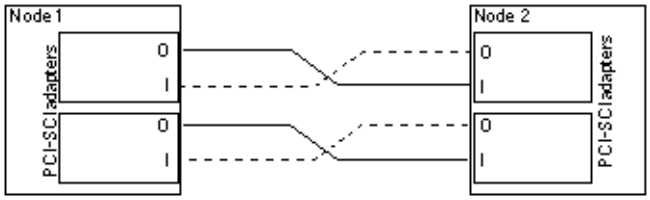

FIGURE 3–3 Typical Two-Node PCI-SCI Cluster Interconnect

   - A four-node cluster requires SCI transport junctions. See Figure 3–4 for a cabling diagram. For the procedure about how to install and cable, see the SCI switch documentation that shipped with your hardware the switches.

     Connect the ends of the cables that are marked `SCI Out` to the `O` connectors on the adapters and the `Out` connectors on the transport junctions.

     Connect the ends of the cables that are marked `SCI In` to the `I` connectors of the adapters and `In` connectors on the transport junctions, as shown in Figure 3–4.

Note – Set the Unit selectors on the fronts of the SCI transport junctions to `F`. Do not use the `X-Ports` on the SCI transport junctions.

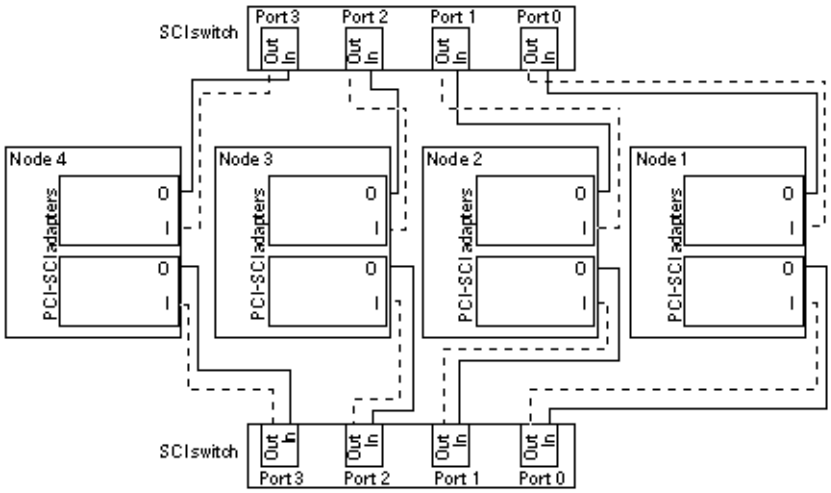

FIGURE 3–4 Typical Four-Node PCI-SCI Cluster Interconnect

Troubleshooting PCI-SCI Interconnects

If you have problems with your PCI-SCI interconnect, check the following items.

- Verify that the LED on the PCI-SCI transport adapter is blinking green rapidly. For detailed LED interpretations and actions, see the documentation that shipped with your host adapter.
- Verify that the PCI-SCI transport adapter card's DIP switch settings are correct. For more information, see the documentation that shipped with your PCI-SCI host adapter.
- Verify that the PCI-SCI cables are correctly connected. The PCI-SCI cable connects to the connector that is marked `SCI In` on the PCI-SCI adapter cards and, if you are using transport junctions, to the `Out` ports on the SCI transport junctions.
- Verify that the PCI-SCI cables are correctly connected. The PCI-SCI cable connects to the connector that is marked `SCI Out` on the PCI-SCI adapter cards and, if you are using transport junctions, to the `In` ports on the SCI transport junctions.
- Verify that the PCI-SCI switch Unit selectors are set to `F`.

Where to Go From Here

Install the Sun Cluster software. Configure the interconnect after you install all other hardware. To review the task map about how to install cluster hardware, see Table 3–1.

Installing Sun Fire Link Cluster Interconnect Hardware

Table 3–3 lists procedures about how to install Sun Fire cluster interconnect hardware. Perform the procedures in the order that the procedures are listed.

TABLE 3–3 Task Map: Installing Sun Fire Link Cluster Interconnect Hardware

Task	For Instructions, Go To
Install the Sun Fire cluster interconnect hardware.	*Sun Fire Link Interconnect Hardware Installation Guide*
Perform the Sun Fire Link software installation.	*Sun Fire Link Software Installation Guide*
Create and activate the Sun Fire Link fabric.	*Sun Fire Link Fabric Administrator's Guide*

Maintaining Ethernet and PCI-SCI Interconnect Hardware in a Running Cluster

The following table lists procedures about how to maintain Ethernet and PCI-SCI cluster interconnect and public network hardware.

TABLE 3–4 Task Map: Maintaining Cluster Interconnect Hardware

Task	For Instructions, Go To
Add transport adapters.	"How to Add Transport Adapters" on page 44

TABLE 3–4 Task Map: Maintaining Cluster Interconnect Hardware *(Continued)*

Task	For Instructions, Go To
Replace transport adapters.	"How to Replace Transport Adapters" on page 45
Remove transport adapters.	"How to Remove Transport Adapters" on page 48
Upgrade transport adapter firmware	"How to Upgrade Transport Adapter Firmware" on page 48
Add transport cables and transport junctions.	"How to Add Transport Cables and Transport Junctions" on page 49
Replace transport cables and transport junctions.	"How to Replace Transport Cables and Transport Junctions" on page 50
Remove transport cables and transport junctions.	"How to Remove Transport Cables and Transport Junctions" on page 51

▼ How to Add Transport Adapters

This section contains the procedure about how to add transport adapters to nodes in a running cluster. For conceptual information on transport adapters, see your Sun Cluster concepts documentation.

1. **Shut down the node in which you are installing the transport adapter.**

   For the procedure about how to shut down a node, see your Sun Cluster system administration documentation.

2. **Power off the node.**

   For the procedure about how to power off a node, see the documentation that shipped with your node.

3. **Install the transport adapter.**

   For the procedure about how to install transport adapters and set transport adapter DIP switches, see the documentation that shipped with your host adapter and node hardware.

4. **Perform a reconfiguration boot to create the new Solaris device files and links.**

Where to Go From Here

After you add your interconnect hardware, you might want to reconfigure Sun Cluster software with the new interconnect components. If you want to reconfigure Sun Cluster software with the new interconnect components, see your Sun Cluster system administration documentation.

▼ How to Replace Transport Adapters

This section contains the procedure about how to replace a failed transport adapter in a node in a running cluster. For conceptual information on transport adapters, see your Sun Cluster concepts documentation.

Caution – You must maintain at least one cluster interconnect between the nodes of a cluster. The cluster does not function without a working cluster interconnect. You can check the status of the interconnect with the command, `scstat -W`. For more details on checking the status of the cluster interconnect, see your Sun Cluster system administration documentation.

1. **Shut down the node with the transport adapter you want to replace.**

   For the procedure about how to shut down a node, see your Sun Cluster system administration documentation.

2. **Power off the node.**

   For the procedure about how to power off your node, see the documentation that shipped with your node.

3. **Disconnect the transport cable from the transport adapter and other devices.**

   For the procedure about how to disconnect cables from transport adapters, see the documentation that shipped with your host adapter and node.

4. **Replace the transport adapter.**

   For the procedure about how to replace transport adapters, see the documentation that shipped with your host adapter and node.

5. **Reconnect the transport cable to the new transport adapter.**

   For the procedure about how to connect cables to transport adapters, see the documentation that shipped with your host adapter and node.

6. **Perform a reconfiguration boot to create the new Solaris device files and links.**

Where to Go From Here

After you replace your interconnect hardware, you might want to reconfigure Sun Cluster software with the new interconnect components. If you want to reconfigure Sun Cluster software with the new interconnect components, see your Sun Cluster system administration documentation.

How to Replace a PCI-SCI Transport Adapter

This section contains the procedure about how to replace a PCI-SCI transport adapter in a running cluster.

Caution – You must maintain at least one cluster interconnect between the nodes of a cluster. The cluster does not function without a working cluster interconnect. You can check the status of the interconnect with the command. `scstat –W`. For more details on checking the status of the cluster interconnect, see your Sun Cluster system administration documentation.

1. **Identify the adapter that you want to replace.**

   ```
   # scstat -W
   ```

2. **Remove the adapter and the cable from the cluster configuration.**

   ```
   # scsetup
   ```

   See the Sun Cluster system administration documentation for the procedure on removing host adapters.

3. **Verify that the removed adapter and cable path has been removed from the cluster configuration.**

   ```
   # scstat -W
   ```

4. **Shut down the node that contains the host adapter you plan to replace.**

   For the full procedure on shutting down and powering off a node, see "Shutting Down and Booting a Single Cluster Node" in the *Sun Cluster 3.1 System Administration Guide*.

5. **Replace the host adapter.**

   For the procedure on replacing the host adapter, see the documentation that shipped with your SCI-PCI host adapter.

   Note – Ensure that you insert the replacement adapter into the original adapter's slot. If the new adapter is inserted into a different slot, the system will assign a new device instance number to the new adapter.

   If you decide to insert the adapter into a different slot, note the adapter's new name. You will need to use the new name when you add the adapter to the configuration with `scsetup`.

6. **Boot the node.**

   For more information on booting nodes, see the Sun Cluster system administration documentation.

For more information on booting nodes, see “Shutting Down and Booting a Single Cluster Node” in the *Sun Cluster 3.1 System Administration Guide*.

7. **Verify that the cluster configuration did not change after you shut down the node.**

```
# scstat -W
```

8. **Determine the adapter’s instance number.**

   a. **Based on the slot number, find the adapter’s name.**

   The following screen shot is an example and might not reflect the hardware that you plan to configure.

```
# prtdiag
...
======================== IO Cards =========================

                    Bus  Max
 IO  Port Bus       Freq Bus  Dev,
Type  ID  Side Slot MHz  Freq Func State Name Model
---- ---- ---- ---- ---- ---- ---- ----- --------------------------------
PCI   8    B    2    33   33  2,0  ok    pci11c8,0-pci11c8,d665.11c8.0.0
PCI   8    B    3    33   33  3,0  ok    pci11c8,0-pci11c8,d665.11c8.0.0
...
```

   b. **By using the adapter’s name and slot number, find the adapter’s instance number.**

   The following screen shot is an example and might not reflect the hardware that you plan to configure.

```
# prtconf
...
    pci, instance #0
        pci11c8,0, instance #0
        pci11c8,0, instance #1
...
```

9. **Add the new adapter card and cable to the configuration.**

Note – It is important that you use the correct SCI adapter name when you add the new adapter through `scsetup` . The adapter name is a combination of the type of the adapter and the adapter’s instance number, for example `sci0`.

```
# scsetup
```

10. **Verify that the paths that connect to the new host adapter are online.**

```
# scstat -W
```

▼ How to Remove Transport Adapters

This section contains the procedure about how to remove an unused transport adapter from a node in a running cluster. For conceptual information on transport adapters, see your Sun Cluster concepts documentation.

Caution – You must maintain at least one cluster interconnect between the nodes of a cluster. The cluster does not function without a working cluster interconnect.

1. **Verify that the transport adapter that you want to remove is not configured in the Sun Cluster configuration.**
   - If the transport adapter that you want to remove appears in the Sun Cluster configuration, remove the transport adapter from the Sun Cluster configuration. To remove a transport path, follow the procedures in your Sun Cluster system administration documentation before going to Step 2.
   - If the transport adapter that you want to remove does not appear in the Sun Cluster configuration, go to Step 2.
2. **Shut down the node that contains the transport adapter that you want to remove.**
   For the procedure about how to shut down a node, see your Sun Cluster system administration documentation.
3. **Power off the node.**
   For the procedure about how to power off a node, see the documentation that shipped with your node.
4. **Disconnect the transport cables from the transport adapter you want to remove.**
   For the procedure about how to disconnect cables from transport adapters, see the documentation that shipped with your host adapter and node.
5. **Remove the transport adapter.**
   For the procedure about how to remove transport adapters, see the documentation that shipped with your host adapter and node.
6. **Perform a reconfiguration boot to create the new Solaris device files and links.**

How to Upgrade Transport Adapter Firmware

Use this procedure to update transport adapter firmware.

1. **Determine the resource groups and the device groups that are running on the node. This node is the node on which you are upgrading transport adapter firmware.**

```
# scstat
```

2. **Migrate the resource groups and device groups off the node that you plan to upgrade the firmware on.**

   ```
   # scswitch -S -h from-node
   ```

3. **Perform the firmware upgrade.**

   For the procedure about how to upgrade your transport adapter firmware, see the patch documentation.

4. **Return the resource groups and device groups to their original node.**

   ```
   # scswitch -z -g resource-group -h nodename
   # scswitch -z -D device-group -h nodename
   ```

▼ How to Add Transport Cables and Transport Junctions

This section contains the procedure about how to add transport cables, transport junctions (switches), or transport cables and transport junctions in a running cluster.

If you plan to use virtual local area networks (VLANs) in your cluster interconnect, configure the transport junction. For more information, see the documentation that shipped with your hardware.

For more information about how configure VLANs, see "Configuring VLANs as Private Interconnect Networks" on page 53.

1. **Shut down the node that is to be connected to the new transport cable, transport junction, or transport cable and transport junction.**

   For the procedure about how to shut down a node, see your Sun Cluster system administration documentation.

2. **Install the transport cable, transport junction, or transport cable and transport junction.**

   - If you are using an Ethernet interconnect, see "How to Install Ethernet Transport Cables and Transport Junctions" on page 38 for cabling diagrams and considerations.
   - If you are using a PCI-SCI interconnect, see "How to Install PCI-SCI Transport Cables and Transport Junction" on page 40 for cabling diagrams and considerations.

3. **Perform a reconfiguration boot to create the new Solaris device files and links.**

Where to Go From Here

After you add your interconnect hardware, you might want to reconfigure Sun Cluster software with the new interconnect components. If you want to reconfigure Sun Cluster software with the new interconnect components, see your Sun Cluster system administration documentation.

▼ How to Replace Transport Cables and Transport Junctions

This section contains the procedure about how to replace failed transport cable, transport junction (switch), or transport cable and transport junction in a running cluster.

If Virtual Local Area Networks (VLANs) are configured, more than one cluster might be impacted by removing a transport junction. Ensure that all clusters are prepared for the removal of a transport junction. Also, record the configuration information of the transport junction you plan to replace and configure the new transport junction accordingly.

For more information about how to configure VLANs, see "Configuring VLANs as Private Interconnect Networks" on page 53.

Caution – You must maintain at least one cluster interconnect between the nodes of a cluster. The cluster does not function without a working cluster interconnect.

1. **Shut down the node that is connected to the transport cable, transport junction, or transport cable and transport junction that you are replacing.**

   For the procedure about how to shut down a node, see your Sun Cluster system administration documentation.

2. **Disconnect the failed transport cable and transport junction from the other devices.**

   For the procedure about how to disconnect cables from transport adapters, see the documentation that shipped with your host adapter and node.

3. **Connect the new transport cable, transport junction, or transport cable and transport junction to the other cluster devices.**

   - If you are replacing an Ethernet interconnect, see "How to Install Ethernet Transport Cables and Transport Junctions" on page 38 for cabling diagrams and considerations.

- If you are replacing a PCI-SCI interconnect, see "How to Install PCI-SCI Transport Cables and Transport Junction" on page 40 for cabling diagrams and considerations.

4. **Perform a reconfiguration boot to create the new Solaris device files and links.**

Where to Go From Here

After you replace your interconnect hardware, you might want to reconfigure Sun Cluster software with the new interconnect components. If you want to reconfigure Sun Cluster software with the new interconnect components, see your Sun Cluster system administration documentation.

▼ How to Remove Transport Cables and Transport Junctions

This section contains the procedure about how to remove an unused transport cable or transport junction (switch) from a node in a running cluster.

If Virtual Local Area Networks (VLANs) are configured, more than one cluster might be impacted by removing a transport junction. Ensure that all clusters are prepared for the removal of a transport junction.

Caution – You must maintain at least one cluster interconnect between the nodes of a cluster. The cluster does not function without a working cluster interconnect.

1. **Check to see whether the transport cable, transport junction, or transport cable and transport junction that you are removing appears in the Sun Cluster configuration.**
   - If the interconnect component that you want to remove appears in the Sun Cluster configuration, remove the interconnect component from the Sun Cluster configuration. To remove an interconnect component, follow the interconnect administration procedures in your Sun Cluster system administration documentation before going to Step 2.
   - If the interconnect component that you want to remove does not appear in the Sun Cluster configuration, go to Step 2.
2. **Shut down the node that is connected to the transport cable, transport junction, or transport cable and transport junction that you are removing.**

   For the procedure about how to shut down a node, see your Sun Cluster system administration documentation.

3. **Disconnect the transport cable, transport junction, or transport cable and transport junction from the other cluster devices.**

   For the procedure about how to disconnect cables from transport adapters, see the documentation that shipped with your host adapter and node.

4. **Perform a reconfiguration boot to create the new Solaris device files and links.**

Maintaining Sun Fire Link Interconnect Hardware in a Running Cluster

This section contains the procedure about how to maintain Sun Fire Link interconnect hardware.

▼ How to Replace Sun Fire Link Interconnect Hardware

Use this procedure to replace any component of the Sun Fire Link interconnect. This procedure requires you to switch data path loads before you replace interconnect hardware.

Caution – You must maintain at least one cluster interconnect between the nodes of a cluster. The cluster does not function without a working cluster interconnect. You can check the status of the interconnect with the command. `scstat –W`. For more details on checking the status of the cluster interconnect, see your Sun Cluster system administration documentation.

1. **Determine whether your node is enabled with the Solaris dynamic reconfiguration (DR) feature.**

   - If your node is enabled with DR, go to Step 2.
   - If your node does not have DR enabled, shutdown the node.

     For the full procedure about how to shut down a node, see your Sun Cluster system administration documentation.

2. **Disconnect the data path component from the other cluster devices.**

   For the procedure about how to disconnect Sun Fire Link components, see the *Sun Fire Link Service Manual*.

3. **Connect the new data path component to the other cluster devices.**

For the procedure about how to connect Sun Fire Link components, see your *Sun Fire Link Service Manual*.

4. **If you shut down the node in Step 1, boot the node (otherwise, skip this step).**
   For more information, see your Sun Cluster system administration documentation.

Configuring VLANs as Private Interconnect Networks

Sun Cluster software supports the use of private interconnect networks over switch-based virtual local area networks (VLAN). In a switch-based VLAN environment, Sun Cluster software enables multiple clusters and nonclustered systems to share Ethernet transport junction (switch) in two different configurations.

The implementation of switch-based VLAN environments is vendor-specific. Since each switch manufacturer implements VLAN differently, the following guidelines address Sun Cluster software requirements about how to configure VLANs with cluster interconnects.

- You must understand your capacity needs before you set up a VLAN configuration. You must know the minimum bandwidth necessary for your interconnect and application traffic.

  For the best results, set the Quality of Service (QOS) level for each VLAN to accommodate basic cluster traffic plus the desired application traffic. Ensure that the bandwidth allocated to each VLAN extends from node to node.

  To determine the basic cluster traffic requirements, use the following equation. In this equation, `n` equals the number of nodes in the configuration, and `s` equals the number of switches per VLAN.

  ```
  n (s-1) x 10Mb
  ```

- Interconnect traffic must be placed in the highest priority queue.
- All ports must be equally serviced, similar to a round robin or first in first out model.
- You must verify that you have properly configured your VLANs to prevent path timeouts.

The first VLAN configuration enables nodes from multiple clusters to send interconnect traffic across one pair of Ethernet transport junctions. Sun Cluster software requires the use of at least two transport junctions to eliminate a single point of failure. The following figure is an example of the first VLAN configuration in a two-node cluster. VLAN configurations are not limited to two-node clusters.

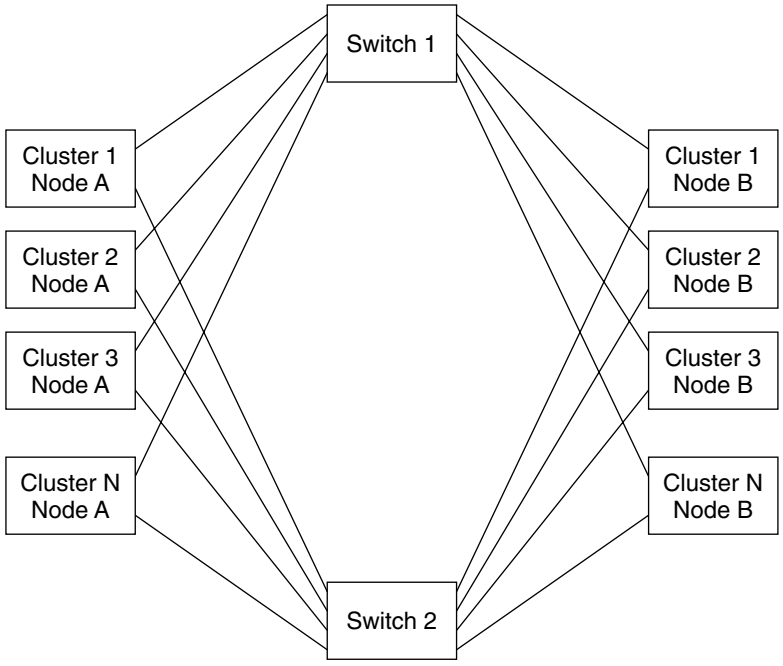

FIGURE 3–5 First VLAN configuration

The second VLAN configuration also uses the same transport junctions for the interconnect traffic of multiple clusters. However, the second VLAN configuration requires two pairs of transport junctions that are connected by links. This configuration allows VLANs to be supported in a campus cluster configuration with the same restrictions as other campus cluster configurations. The following figure illustrates the second VLAN configuration.

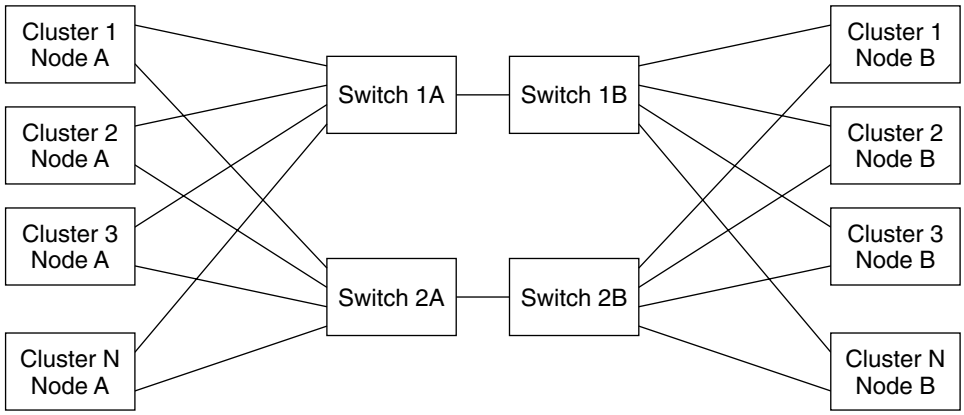

FIGURE 3–6 Second VLAN configuration

CHAPTER 4

Installing and Maintaining Public Network Hardware

This chapter describes the procedures about how to install and maintain public network hardware.

This chapter contains information about how to maintain public network hardware. This chapter contains the following procedures.

For conceptual information on cluster interconnects and public network interfaces, see your Sun Cluster concepts documentation.

For information on how to administer public network interfaces, see your Sun Cluster system administration documentation

Installing Public Network Hardware

This section covers installing cluster hardware during an *initial cluster installation, before Sun Cluster software is installed*.

Physically installing public network adapters to a node in a cluster is no different from adding public network adapters in a noncluster environment.

For the procedure about how to add public network adapters, see the documentation that shipped with your nodes and public network adapters.

Where to Go From Here

Install the cluster software and configure the public network hardware after you have installed all other hardware. To review the task map about how to install cluster hardware, see "Installing Sun Cluster Hardware" on page 13.

Maintaining Public Network Hardware in a Running Cluster

The following table lists procedures about how to maintain cluster interconnect and public network hardware. The interconnect maintenance procedures in this section are for both Ethernet-based and PCI-SCI interconnects.

TABLE 4–1 Task Map: Maintaining Public Network Hardware

Task	Information
Add public network adapters.	"How to Add Public Network Adapters" on page 58
Replace public network adapters.	"How to Replace Public Network Adapters" on page 58
Remove public network adapters.	"How to Remove Public Network Adapters" on page 59

How to Add Public Network Adapters

Physically adding public network adapters to a node in a cluster is no different from adding public network adapters in a noncluster environment.

For the procedure about how to add public network adapters, see the hardware documentation that shipped with your node and public network adapters.

Where to Go From Here

To add a new public network adapter to an IP Network Multipathing group, see the *IP Network Multipathing Administration Guide*.

How to Replace Public Network Adapters

For cluster-specific commands and guidelines about how to replac public network adapters, see your Sun Cluster system administration documentation.

For procedures about how to administer public network connections, see the *IP Network Multipathing Administration Guide*.

For the procedure about remove public network adapters, see the hardware documentation that shipped with your node and public network adapters.

Where to Go From Here

To add the new public network adapter to a IP Network Multipathing group, see your Sun Cluster system administration documentation.

How to Remove Public Network Adapters

For cluster-specific commands and guidelines about how to remove public network adapters, see your Sun Cluster system administration documentation.

For procedures about how to administer public network connections, see the *IP Network Multipathing Administration Guide*.

For the procedure about how to remove public network adapters, see the hardware documentation that shipped with your node and public network adapters.

Sun Gigabit Ethernet Adapter Considerations

Some Gigabit Ethernet switches require some device parameter values to be set differently than the defaults. Chapter 3 of the *Sun Gigabit Ethernet/P 2.0 Adapter Installation and User's Guide* describes the procedure about how to change device parameters. The procedure that you use on nodes that are running Sun Cluster software varies slightly from the procedure that is described in the guide. In particular, the difference is in how you derive parent names for use in the `ge.conf` file from the `/etc/path_to_inst` file.

Chapter 3 of the *Sun Gigabit Ethernet/P 2.0 Adapter Installation and User's Guide* describes the procedure on how to change `ge` device parameter values. This change occurs through entries in the `/kernel/drv/ge.conf` file. The procedure to derive the parent name from the `/etc/path_to_inst` listing, which is be used in `ge.conf` entries, appears in *Setting Driver Parameters Using a ge.conf File*. For example, from the following `/etc/path_to_inst` line, you can derive the parent name for `ge2` to be `/pci@4,4000`.

```
"/pci@4,4000/network@4" 2 "ge"
```

On Sun Cluster nodes, a `/node@`*nodeid* prefix appears in the `/etc/path_to_inst` line. Do *not* consider the `/node@`*nodeid* prefix when you derive the parent name. For example, on a cluster node, an equivalent `/etc/path_to_inst` entry would be the following:

```
"/node@1/pci@4,4000/network@4" 2 "ge"
```

The parent name for `ge2`, to be used in the `ge.conf` file is still `/pci@4,4000` in this instance.

`ce` Sun Ethernet Driver Considerations

The Sun Cluster software supports the `ce` Sun Ethernet driver for cluster interconnect and public network applications. Consult your Sun sales representative for details on which network interface products are supported. These network interfaces can only be used as new interconnects and public network interfaces. As such, you cannot upgrade from another type of adapter card to those network interfaces by using the `ce` Sun Ethernet driver.

When you use the `ce` Sun Ethernet driver for public network connections, add the following line to your `/etc/system` file.

```
set ce:ce_reclaim_pending=1
```

Apply any required patches, and then reboot the system. See the Sun Cluster system administration documentation for instructions on rebooting a cluster node.

The `ce` driver version must be at least 1.115. To determine the `ce` driver version, run the following command.

```
# modinfo | grep CE
```

Note – For the current list of patches that are required for the `ce` Sun Ethernet driver, refer to SunSolve. SunSolve is available online to Sun service providers and to customers with SunSolve service contracts at the SunSolve site: http://sunsolve.sun.com.

APPENDIX **A**

Campus Clustering with Sun Cluster Software — Concepts

This appendix introduces some of the basic concepts of campus clustering and provides some configuration and setup examples. This appendix does not attempt to explain clustering, provide information on clustering administration, or furnish details about hardware installation and configuration. For conceptual information on clustering, see your Sun Cluster concepts documentation and your Sun Cluster system administration documentation.

Introduction

The only significant difference between traditional clustering and campus clustering is that of distance. In campus clustering, the nodes of a cluster configuration can be up to several kilometers apart. This increases the likelihood that, in the case of a catastrophe such as a fire or earthquake, at least one server and its storage will survive.

Sun Cluster software supports up to three nodes in a campus cluster configuration, and both two and three-room configuration are supported. A **room** can be thought of as a functionally independent hardware grouping (such as a node and its attendant storage, or a quorum device physically separated from any nodes) that has been separated from other rooms to increase the likelihood of failover and redundancy in case of accident or failure. The definition of a room therefore depends on the type of failures to be safeguarded against, as indicated by Table A–1.

TABLE A–1 Definitions of "Room"

Failure Scenario	Sample Definitions of Separate Rooms
Power line failure	Isolated and independent power supplies

TABLE A–1 Definitions of "Room" *(Continued)*

Failure Scenario	Sample Definitions of Separate Rooms
Minor accidents, furniture collapse, seepage, etc.	Different parts of physical room
Small fire. Sprinklers (fire) starting	Different physical areas (for example, sprinkler zone)
Structural failure (building-wide fire, for example)	Different buildings
Large-scale natural disaster (for example, earthquake or flood)	Different corporate campuses up to several kilometers apart

In two-room configurations, the quorum disk occupies the same room as one node (see "Example Two-Room Configuration With Host-Based Data Replication" on page 63). In the case of a three-room configuration, the third room is used for the quorum disk ("Example Three-Room Configuration" on page 64) or a small server.

In a two-room configuration, the quorum disk should be located in the room that is more likely to survive an accident or catastrophe in the event that all cluster transport and disk connectivity is lost between rooms. (If **only** cluster transport is lost, the node sharing a room with the quorum disk will not necessarily be the room that reserves the quorum disk first.) For more information about quorum and quorum devices, see the Sun Cluster concepts documentation.

The advantage to a three-room cluster is that, if any one of the three rooms is lost, automatic failover should typically be possible; whereas with a two-room cluster, if an entire room is lost, automatic failover is possible only if the surviving room contains the quorum disk. **Only a three-room configuration guarantees system availability in the event of complete loss of an entire room (in the absence of any other failures).**

Note – As with noncampus configurations, data integrity will be compromised if there are other unrecovered I/O failures present when a room is destroyed, and the most up-to-date submirror was in the destroyed room.

In a campus cluster configuration, each of the two rooms used for nodes should have an equal number of shared disks. (In a two-room configuration, one room can have a separate quorum disk, so the two rooms need not have the same number of **total** disks.) Data replication between shared disks must always be performed across rooms, rather than within rooms. In other words, both copies of the data should never be in the same room. Mirroring is required for all campus cluster configurations, since RAID-5 alone does not lend itself to providing data redundancy across rooms.

If you use Solstice DiskSuite/Solaris Volume Manager as your volume manager for shared device groups, pay special consideration to the distribution of replicas. In two-room configurations, all disksets should be configured with an additional replica

in the room that houses the cluster quorum disk. Further, all Solstice DiskSuite/Solaris Volume Manager device groups should be configured to use the node in the quorum disk room as their default primary room. In three-room configurations, the third room should not only house the quorum disk, but also include at least one extra disk configured into each of the disksets. Each diskset should include a third-room disk with an extra Solstice DiskSuite/Solaris Volume Manager replica per diskset. Quorum disks can be used as `metadb` replicas in a metaset.

Sun Cluster software supports campus cluster configurations with rooms up to 10km apart by using FC over single-mode fiber. Sun Cluster supports campus cluster configurations separated by greater distances by using Nortel OPTera wavelength division multiplexing (WDM) products. For more information regarding Nortel OPTera WDM products, see the Nortel OPTera product documentation.

Campus Cluster Configuration Examples

This section provides examples of two and three-room campus clustering configurations:

- Figure A–1 shows a two-room configuration that is using Sun StorEdge T3 partner groups for host-based mirroring.
- Figure A–2 shows a three-room campus cluster that is using Sun StorEdge A5x00 disk arrays for host-based mirroring.

Note – The examples in this chapter illustrate general configurations and are not intended to indicate required or recommended setups. The type of storage array shown, for example, is purely for illustrative purposes. For simplicity's sake, the diagrams and explanations concentrate only on features unique to understanding campus clustering, so that, for example, public-network Ethernet connections are not shown.

See "Additional Campus Cluster Configuration Examples" on page 67 for other example setups.

Example Two-Room Configuration With Host-Based Data Replication

A two-room configuration with host-based data replication is defined as follows.

- Two separate rooms

- Both rooms with one node each and disk subsystems
- Data mirrored across disk subsystems in these rooms
- At least one disk subsystem, attached to both hosts, used as a quorum device, located in one of the rooms

In the event of loss of the room containing the quorum disk, the system will be unable to recover automatically. Recovery will require operator intervention.

Figure A–1 shows a sample two-room configuration by using a partner-group of Sun StorEdge T3/T3+ disk arrays in each room.

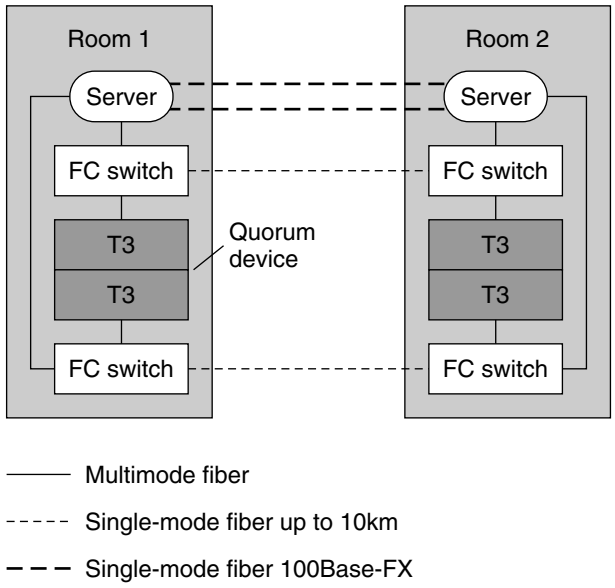

FIGURE A–1 Example Two-Room Campus Cluster without a multipathing solution implemented

Figure A–1 is similar to a standard noncampus configuration; the most obvious difference is that FC switches have been added to switch from multimode to single-mode fibers.

Although not so depicted in this chapter, campus clustering enables for configurations by using multiple storage arrays. Where large storage environments are required, additional SAN switches for Sun StorEdge T3/T3+ arrays might be required.

Example Three-Room Configuration

Sun Cluster supports two different three-room configurations.

- Two rooms with one node each and an equal number of disk arrays, in this case, Sun StorEdge A5x00 disk subsystems; data is mirrored across disk-subsystems in these rooms

  Third room with at least one disk subsystem attached to both hosts to be used as a quorum device
- Two rooms contain one node each and an equal number of disk arrays. The third room contains a small server. The small server removes the need for a storage array to be used as a quorum device.

Figure A–2 shows this configuration by using Sun StorEdge A5x00 disk arrays. Note that, unlike Figure A–1, which used Sun StorEdge T3 disk trays, this configuration does not use FC switches to connect Sun StorEdge A5x00 disk arrays. (Switches are not needed, as long-wave GBICs are already present in the A5x00s and should also be present in the servers' host bus adapters.)

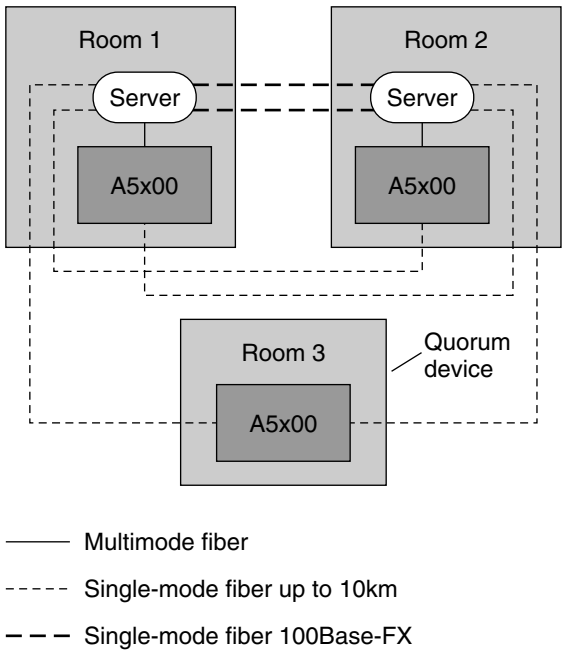

FIGURE A–2 Example Three-Room Configuration

In this configuration, as long as at least two rooms are up and communicating, recovery will be automatic. This is the only configuration in which loss of any one room is guaranteed to be automatically handled. Loss of two rooms requires the replacement or rebuilding of one room and typically requires Sun Service intervention.

Campus Cluster Interconnect Hardware

The following components make up the campus cluster interconnect hardware configuration:

- SunFastEthernet Adapter. The SunFastEthernet Adapter provides 10/100 Mbps Ethernet functionality with an RJ-45 connector.
- Media converters. Use RJ-45 media converters to convert from copper to optical fiber connectors. Consult your Sun sales representative for additional information about media converters.
- Two 9/125-micrometer single-mode fiber pairs.

Figure A–3 shows the setup for the cluster interconnect:

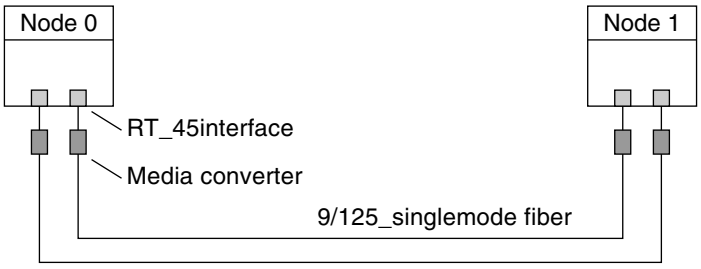

FIGURE A–3 100BASE-FX Setup

Installing and Configuring Interconnect, Storage, and Fibre Channel Hardware

For the most part, using interconnect, storage, and FC hardware does not differ markedly from noncampus cluster configurations.

- The steps for installing Ethernet-based campus cluster interconnect hardware are very similar to those for noncampus clusters. Refer to "Installing Ethernet Cluster Interconnect Hardware" on page 38.

  When installing the media converters, consult the documentation that came with them, including requirements for fiber connections.
- The guidelines for installing virtual local area networks s interconnect networks are very similar to those for noncampus clusters. See "Configuring VLANs as Private Interconnect Networks" on page 53.

- The steps for installing Sun StorEdge A5x00 and Sun StorEdge T3/T3+ arrays are very similar to those for noncampus clusters. For information about how to install a Sun StorEdge A5x00 array, see the *Sun Cluster 3.1 With Sun StorEdge A5x00 Array Installation and Service Manual*. For information about how to install a Sun StorEdge T3/T3+ array, see the *Sun Cluster 3.1 With Sun StorEdge T3 or T3+ Array Partner-Group Configuration Installation and Service Manual* and the *Sun Cluster 3.1 With Sun StorEdge T3 or T3+ Array Single Controller Configuration Installation and Service Manual*.

  However, when installing Sun StorEdge A5x00 arrays at distances greater than 500m, install the Sun Long Wave GBICs as indicated in the *Sun StorEdge Long Wave GBIC Interface Converter Guide*. This manual also includes single-mode fiber specifications.
- Campus clusters by using Sun StorEdge T3/T3+ arrays require FC switches to mediate between multimode and single-mode fibers. The steps for configuring the settings on the FC switches are very similar to those for noncampus clusters.

  Cascaded-switch SAN setups have requirements that are distinct from those of noncascades setups, especially with regard to mirroring and zoning. For cascaded-switch SAN setup requirements, see the documentation that shipped with your switch hardware. For more information about how to use the FC switches in a campus cluster configuration, also see the documentation that shipped with your switch hardware.

Additional Campus Cluster Configuration Examples

"Campus Cluster Configuration Examples" on page 63 showed three possible configurations for campus clustering: a two-room configuration that uses Sun StorEdge T3 partner groups for host-based mirroring, a two-room configuration that uses Sun StorEdge 9900 series storage devices for storage-based data replication, and a three-room configuration by using Sun StorEdge A5x00 arrays. While detailing all of the configurations possible in campus clustering is far beyond the scope of this document, the following illustrations depict some other variations on the setups previously shown.

The following are two-room examples:

- Two-room campus cluster with a multipathing solution implemented (Figure A–4)
- Two-room campus cluster without a multipathing solution implemented (Figure A–5)
- Two-room cluster by using Sun StorEdge A5x00 disk arrays (Figure A–6)

The following are three-room examples:

- Three-room campus cluster by using a StorEdge A5x00 with a multipathing solution implemented (Figure A–7)
- Three-room campus cluster by using a StorEdge A5x00 without a multipathing solution implemented (Figure A–8)
- Three-room campus cluster by using Sun StorEdge A5x00 disk arrays (Figure A–9)
- Three-room campus cluster by using a server to replaces the storage array in the third room. (Figure A–10)
- Three-room campus cluster by using StorEdge T3 arrays and FC switches. (Figure A–11)

Two-Room Examples

Figure A–1 depicted a two-room campus cluster by using Sun StorEdge T3 partner groups and four FC switches. Figure A–4 shows a similar setup with a multipathing solution implemented. The four switches are added to the cluster for greater redundancy and potentially better I/O throughput. This configuration could be implemented by using Sun StorEdge T3 partner groups or Sun StorEdge 9910/9960 arrays with Sun StorEdge Traffic Manager software installed.

For information on Sun StorEdge Traffic Manager software, see the *Sun StorEdge Traffic Manager Installation and Configuration Guide* at `http://www.sun.com/products-n-solutions/hardware/docs/`.

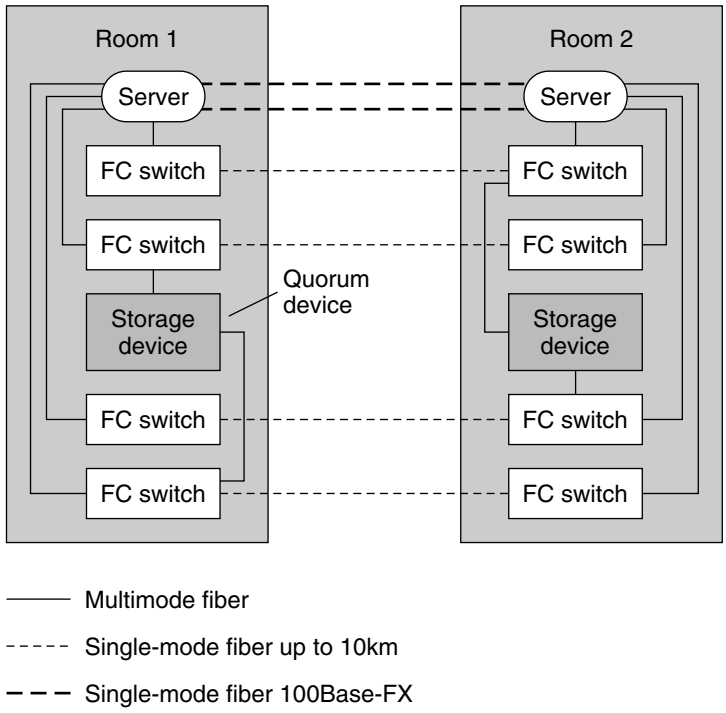

FIGURE A–4 Two-Room Campus Cluster With a Multipathing Solution Implemented

Figure A–5 is similar to Figure A–1, except that it does not have a multipathing solution implemented. This configuration could be implemented by using Sun StorEdge T3/T3+ arrays in single-controller configurations, rather than partner-groups.

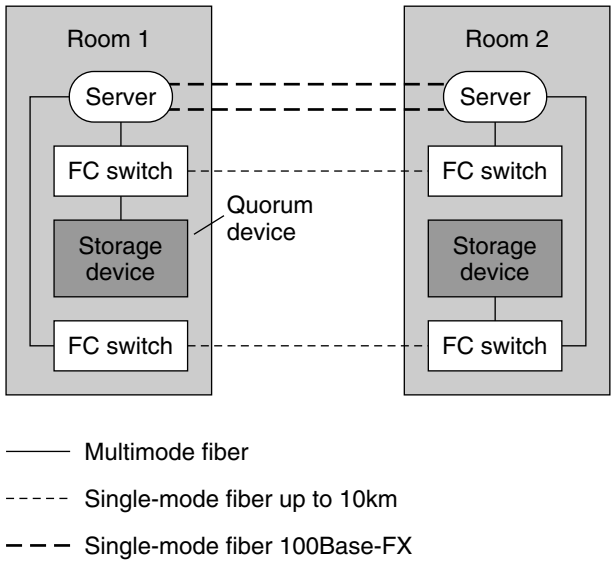

FIGURE A–5 Two-Room Campus Cluster Without a Multipathing Solution Implemented

Figure A–6 depicts a two-room campus cluster by using Sun StorEdge A5x00s. Note the absence of switches.

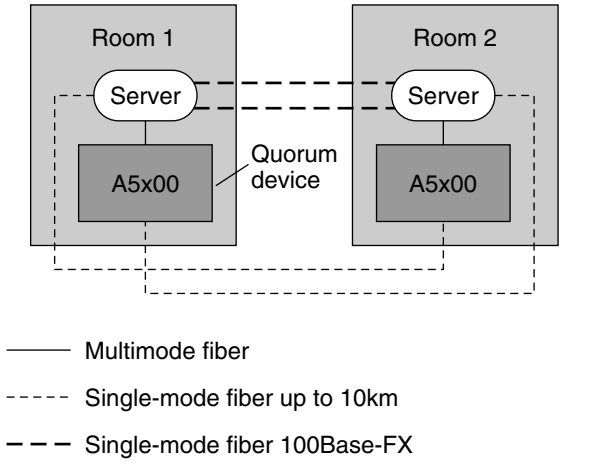

FIGURE A–6 Two-Room Configuration (Sun StorEdge A5x00s)

Three-Room Examples

Sun Cluster supports several three-room campus cluster configurations.

Figure A–7, Figure A–8, and Figure A–9 show Sun StorEdge A5x00 disk arrays in the quorum room. The StorEdge A5x00 arrays do not require switches as the long-wave GBICS are already present in the A5x00 and should be present in the server's host bus adapters.

Figure A–7 is the same as Figure A–1, except that the quorum disk is separate and in a third room.

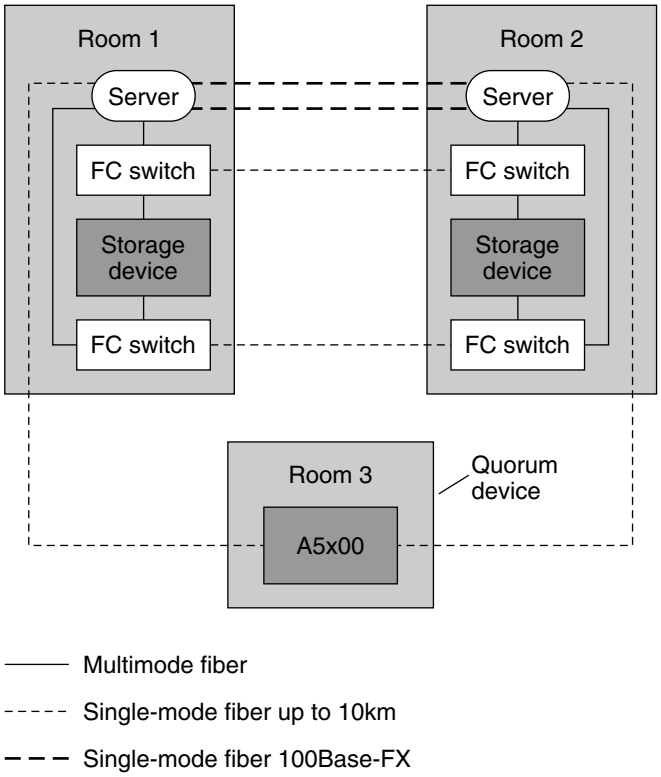

FIGURE A–7 Three-Room Campus Cluster With a Multipathing Solution Implemented

Like Figure A–5, Figure A–8 depicts a campus cluster without a multipathing solution implemented. However, in Figure A–8 the quorum disk is separate, in a third room.

Because the configurations in Figure A–7 and Figure A–8 use heterogeneous storage setups (Sun StorEdge T3/T3+ or Sun StorEdge 9910/9960 and Sun StorEdge A5x00 arrays), the host bus adapter types will be mixed. The Sun StorEdge A5x00 connects to

the server by using a SOC+-based SBus host adapter, so the server must be SBus-based, meaning that the connection to the storage arrays must also be SBus-based. Therefore, the storage array connection must be via a Sun StorEdge SBus Dual FC Network Adapter.

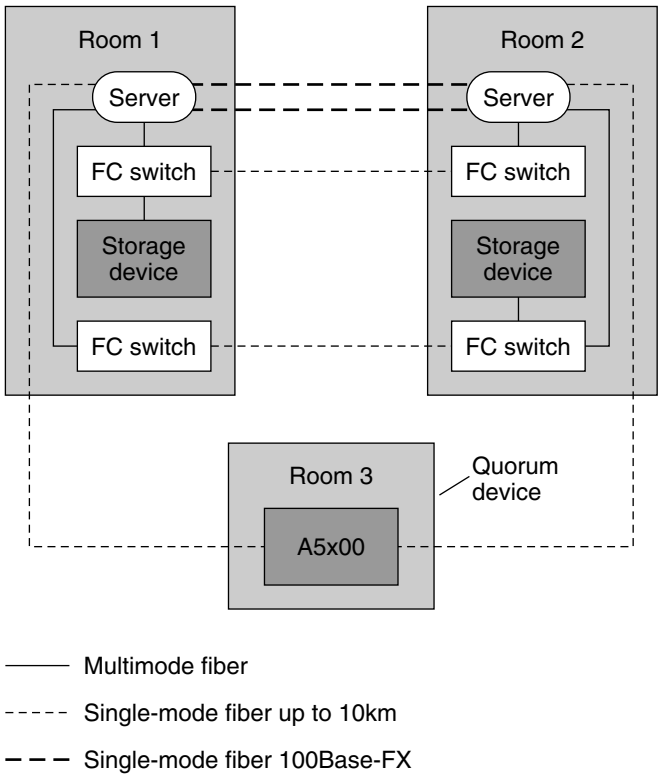

FIGURE A–8 Three-Room Campus Cluster Without a Multipathing Solution Implemented

Figure A–9, like Figure A–6, depicts a campus cluster by using Sun StorEdge A5x00 arrays; however, in Figure A–9, the quorum disk is now separate and in its own room.

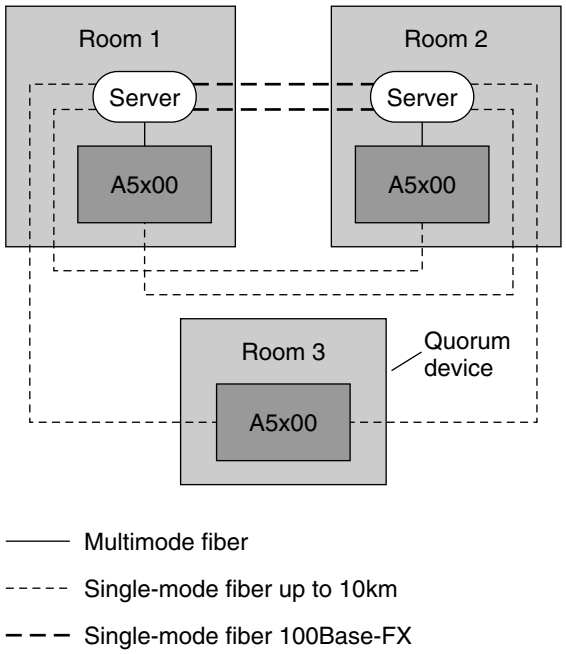

FIGURE A–9 Three-Room Campus Cluster (Sun StorEdge A5x00s)

Figure A–10 depicts a three-room campus cluster where the third room contains a server instead of a storage array. By including the third server in the configuration, you remove the need for a quorum device.

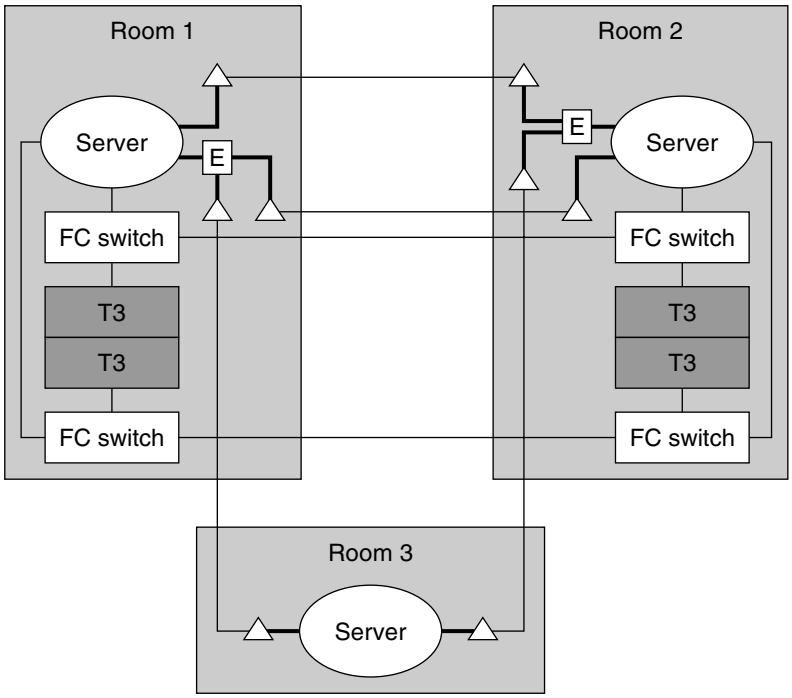

FIGURE A–10 Three-Room Campus Cluster (Third-room Server Replaces Storage Array)

Figure A–11 depicts a three-room campus cluster where the third room contains a StorEdge T3 and is connected to the other rooms through FC switches.

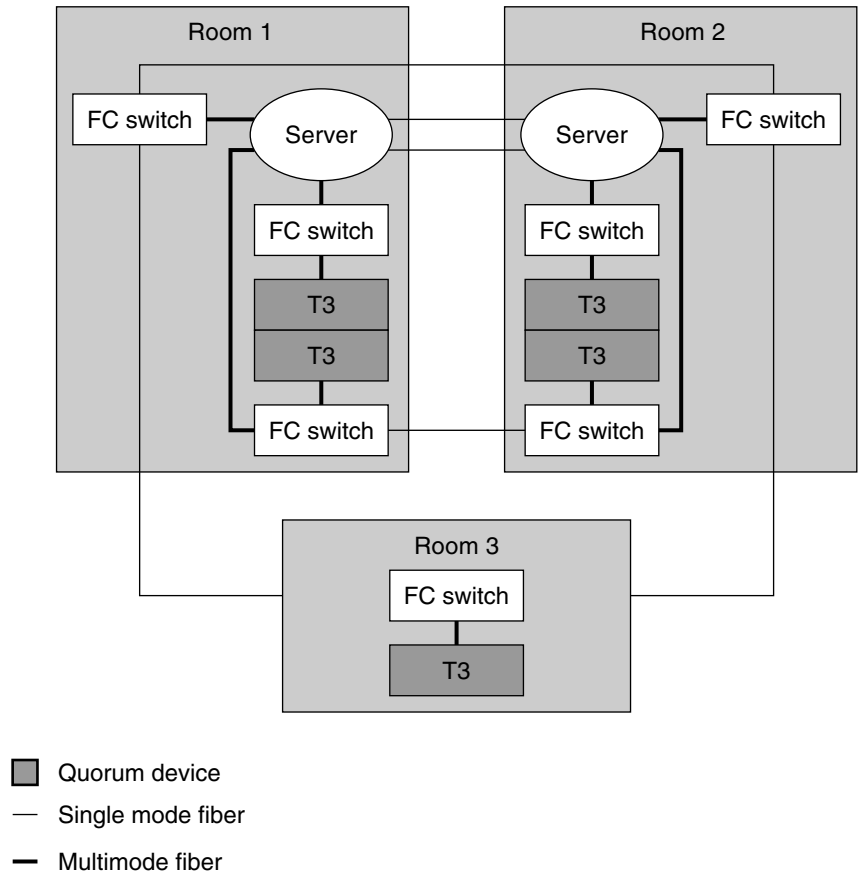

FIGURE A–11 Three-Room Campus Cluster (StorEdge T3 Arrays)

APPENDIX **B**

Verifying Sun Cluster Hardware Redundancy

This appendix describes the tests for verifying and validating the high availability (HA) of your Sun Cluster configuration. The tests in this appendix assume that you installed Sun Cluster hardware, the Solaris operating environment, and Sun Cluster software. All nodes should be booted as cluster members.

This appendix contains the following procedures.

If your cluster passes these tests, your hardware has adequate redundancy. This redundancy means that your nodes, cluster transport cables, and Network Adapter Failover (NAFO) groups are not single points of failure.

To perform the tests in "How to Test Nodes Using a Power-off Method" on page 78 and "How to Test Cluster Interconnects" on page 79, you must first identify the device groups that each node masters. Perform these tests on all cluster pairs that share a disk device group. Each pair has a primary and a secondary for a particular device group. Use the `scstat`(1M) command to determine the initial primary and secondary.

For conceptual information on primary, secondary, failover, device groups, or cluster hardware, see your Sun Cluster concepts documentation.

Testing Node Redundancy

This section provides the procedure about how to test node redundancy and high availability of device groups. Perform the following procedure to confirm that the secondary takes over the device group that is mastered by the primary when the primary fails.

▼ How to Test Nodes Using a Power-off Method

1. **Power off the primary node.**

   Cluster interconnect error messages appear on the consoles of the existing nodes.

2. **On another node, verify that the secondary took ownership of the device group that is mastered by the primary.**

   Look for the output that shows the device group ownership.

   ```
   # scstat
   ```

3. **Power on the initial primary. Boot the node into cluster mode.**

   Wait for the system to boot. The system automatically starts the membership monitor software. The node then rejoins the Sun Cluster configuration.

4. **Do you have the device group failback option enabled?**

   Use the `scconf -p` command to determine if your device group has the device group failback option enabled.

   - If yes, skip to Step 6.

     The system boot process moves ownership of the device group back to the initial primary.

   - If no, proceed to Step 5.

5. **From the initial primary, move ownership of the device group back to the initial primary.**

   ```
   # scswitch -S -h nodename
   ```

6. **Verify that the initial primary has ownership of the device group.**

   Look for the output that shows the device group ownership.

   ```
   # scstat
   ```

Testing Cluster Interconnect and IP Network Multipathing Group Redundancy

This section provides the procedure about how to test cluster interconnect and IP Network Multipathing group redundancy.

▼ How to Test Cluster Interconnects

1. **Disconnect one of the cluster transport cables from a primary node that masters a device group.**

   Messages appear on the consoles of each node, and error messages appear in the `/var/adm/messages` file. If you run the `scstat`(1M) command, the Sun Cluster software assigns a `faulted` status to the cluster transport path that you disconnected. This fault does not result in a failover.

2. **Disconnect the remaining cluster transport cable from the primary node that you identified in Step 1.**

   Messages appear on the consoles of each node, and error messages appear in the `/var/adm/messages` file. If you run the `scstat` command, the Sun Cluster software assigns a `faulted` status to the cluster transport path that you disconnected. This action causes the primary node to shutdown. This shutdown results in a partitioned cluster.

   For conceptual information on failure fencing or *split brain*, see your Sun Cluster concepts documentation.

3. **On another node, verify that the secondary node took ownership of the device group that was mastered by the primary.**

   ```
   # scstat
   ```

4. **Reconnect all cluster transport cables.**

5. **Boot the initial primary, which you identified in Step 1, into cluster mode.**

   For the procedure about how to boot a node, see your Sun Cluster system administration documentation.

6. **Do you have the device group failback option enabled?**

   Use the `scconf -p` command to determine if your device group has the device group failback option enabled.

   - If yes, skip to Step 8.

     The system boot process returns ownership of the resource groups and device groups to the initial primary.

   - If no, proceed to Step 7.

7. **Move all resource groups and device groups off the current primary.**

   ```
   # scswitch -S -h from-node
   ```

8. **Verify that the Sun Cluster software assigned a `path online` status to each cluster transport path that you reconnected in Step 4.**

   ```
   # scstat
   ```

▼ How to Test IP Network Multipathing Groups

If you perform this test, you can verify that IP addresses failover from one adapter to another adapter within the same IP Network Multipathing group.

1. **Verify that all network adapters are active.**

   ```
   # scstat -i
   ```

2. **Disconnect one public network cable from an active network adapter.**

3. **Verify that the network adapter status displays as `Failed`.**

   ```
   # scstat -i
   ```

4. **Ensure the IP address failed over to another adapter.**

   ```
   # ifconfig -a
   ```

5. **Reconnect the public network cable to the network adapter.**

   - If `failback` is set to `yes` in the `/etc/default/mpathd` file, the IP address automatically returns to the original network adapter.
   - If `failback` is not set to `yes`, the IP address remains with the new adapter until you perform a manual switchover.

     For more procedures about how to move an IP address to another adapter, see your Sun Cluster system administration documentation.

6. **Repeat steps Step 1 to Step 5 for each active adapter.**

APPENDIX C

Recabling and Replacing Disk Devices

This appendix contains the procedures for recabling and replacing disk devices. This appendix provides the following procedures.

Replacing a Failed Boot Disk

This section contains the procedure about how to replace a failed boot disk.

How to Replace a Failed Boot Disk in a Running Cluster

Use this procedure to replace a failed boot disk that is mirrored and that is on a node in a running cluster. Use this procedure for both Solstice DiskSuite/Solaris Volume Manager and VERITAS Volume Manager. This procedure also assumes that one mirror is available. This procedure defines Node N as the node on which you are replacing the failed boot disk.

If you do not have a mirror that is available, see your Sun Cluster system administration documentation to restore data on the boot disk.

1. **Is Node N up and running?**
   - If no, skip to Step 5

- If yes, proceed to Step 2

2. **Is your disk drive hot-pluggable?**
   - If yes, skip to Step 5
   - If no, proceed to Step 3

3. **Determine the resource groups and device groups running on Node N.**

   Record this information because you use this information in Step 6 of this procedure to return resource groups and device groups to these nodes.

   ```
   # scstat
   ```

   For more information, see your Sun Cluster system administration documentation.

4. **Move all resource groups and device groups off Node N.**

   ```
   # scswitch -S -h from-node
   ```

   For more information, see your Sun Cluster system administration documentation.

5. **Replace the failed boot disk by using the procedure that is outlined in your volume manager documentation.**

   For more information, see your Solstice DiskSuite/Solaris Volume Manager or VERITAS Volume Manager documentation. For the procedure about how to replace a boot disk, see "Boot Problems" in *Solaris Volume Manager Administration Guide*.

6. **If you moved all resource groups and device groups off Node N in Step 4, return the resource groups and device groups that you identified in Step 3 to Node N.**

   ```
   # scswitch -z -g resource-group -h nodename
   # scswitch -z -D device-group-name -h nodename
   ```

   For more information, see your Sun Cluster system administration documentation.

Moving a Disk Cable

You can move a disk cable to a different host adapter on the same bus because of a failed host adapter. However, the best solution is to replace the failed host adapter rather than recable the disk cable to a different host adapter. You might want to move a disk cable to a different host adapter on the same bus to improve performance.

This section provides the following two procedures for moving a disk cable.

- "How to Move a Disk Cable to a New Host Adapter" on page 83
- "How to Move a Disk Cable From One Node to Another Node" on page 84

Use one of these two procedures to prevent interference with normal operation of your cluster when you want to move a disk cable to a different host adapter on the same bus. If you do not follow these procedures correctly, you might see an error the next time you run the `scdidadm -r` command or the `scgdevs` command. If you see an error message that says `did reconfiguration discovered invalid diskpath`, go to "How to Update Sun Cluster Software to Reflect Proper Device Configuration" on page 85.

▼ How to Move a Disk Cable to a New Host Adapter

Use this procedure to move a disk cable to a new host adapter within a node.

Caution – Failure to follow this cabling procedure might introduce invalid device IDs (DIDs) and render the devices inaccessible.

1. **Stop all I/O to the affected disk(s).**

   For more information, see your Solstice DiskSuite/Solaris Volume Manager or VERITAS Volume Manager documentation.

2. **Unplug the cable from the old host adapter.**

3. **From the local node, unconfigure all drives that are affected by the recabling.**

   ```
   # cfgadm
   ```

   Or reboot the local node.

   ```
   # reboot -- -r
   ```

4. **From the local node, update the Solaris device link.**

   ```
   # devfsadm -C
   ```

5. **From the local node, update the DID device path.**

   ```
   # scdidadm -C
   ```

6. **Connect the cable to the new host adapter.**

7. **From the local node, configure the drives in the new location.**

   ```
   # cfgadm
   ```

   Or reboot the local node.

   ```
   # reboot -- -r
   ```

8. **Add the new DID device path.**

   ```
   # scgdevs
   ```

Where to Go From Here

If you did not follow this procedure correctly, you might see an error the next time you run the `scdidadm -r` command or the `scgdevs` command. If you see an error message that says `did reconfiguration discovered invalid diskpath`, go to "How to Update Sun Cluster Software to Reflect Proper Device Configuration" on page 85.

▼ How to Move a Disk Cable From One Node to Another Node

Use this procedure to move a disk cable from one node to another node.

Caution – Failure to follow this cabling procedure might introduce invalid device IDs (DIDs) and render the devices inaccessible.

1. **Delete all references to the DID device path that you are removing from all volume manager and data service configurations.**

   For more information, see your Sun Cluster data services collection and your Solstice DiskSuite/Solaris Volume Manager or VERITAS Volume Manager documentation.

2. **Stop all I/O to the affected disk(s).**

   For more information, see your Solstice DiskSuite/Solaris Volume Manager or VERITAS Volume Manager documentation.

3. **Unplug the cable from the old node.**

4. **From the old node, unconfigure all drives that are affected by the recabling.**

   ```
   # cfgadm
   ```

   Or reboot the old node.

   ```
   # reboot -- -r
   ```

5. **From the old node, update the Solaris device link.**

   ```
   # devfsadm -C
   ```

6. **From the old node, update the DID device path.**

   ```
   # scdidadm -C
   ```

7. **Connect the cable to the new node.**

8. **From the new node, configure the drives in the new location.**

```
# cfgadm
```

Or reboot the new node.

```
# reboot -- -r
```

9. **From the new node, create the new Solaris device links.**

```
# devfsadm
```

10. **From the new node, add the new DID device path.**

```
# scgdevs
```

11. **Add the DID device path on the new node to any volume manager and data service configurations that are required.**

When you configure data services, check that your node failover preferences are set to reflect the new configuration.

For more information, see your Sun Cluster data services collection and your Solstice DiskSuite/Solaris Volume Manager or VERITAS Volume Manager documentation.

Where to Go From Here

If you did not follow this procedure correctly, you might see an error the next time you run the `scdidadm -r` command or the `scgdevs` command. If you see an error message that says `did reconfiguration discovered invalid diskpath`, go to "How to Update Sun Cluster Software to Reflect Proper Device Configuration" on page 85.

▼ How to Update Sun Cluster Software to Reflect Proper Device Configuration

If you see the following error when you run the `scdidadm -r` command or the `scgdevs` command, the Sun Cluster software does not reflect the proper device configuration because of improper device recabling.

```
did reconfiguration discovered invalid diskpath.
This path must be removed before a new path
can be added. Please run did cleanup (-C)
then re-run did reconfiguration (-r).
```

Use this procedure to ensure that the Sun Cluster software becomes aware of the new configuration and to guarantee device availability.

1. **Ensure that your cluster meets the following conditions.**

- The cable configuration is correct.

- The cable that you are removing is detached from the old node.
- The old node is removed from any volume manager or data service configurations that are required.

For more information, see your Sun Cluster data services collection and your Solstice DiskSuite/Solaris Volume Manager or VERITAS Volume Manager documentation.

2. **From all nodes, one node at a time, unconfigure all drives.**

   ```
   # cfgadm
   ```

   Or reboot all nodes, one node at a time.

   ```
   # reboot -- -r
   ```

3. **From all nodes, one node at a time, update the Solaris device link.**

   ```
   # devfsadm -C
   ```

4. **From all nodes, one node at a time, update the DID device path.**

   ```
   # scdidadm -C
   ```

5. **From all nodes, one node at a time, configure all drives.**

   ```
   # cfgadm
   ```

   Or reboot all nodes, one node at a time.

   ```
   # reboot -- -r
   ```

6. **From any node, add the new DID device path.**

   ```
   # scgdevs
   ```

7. **From all nodes that are affected by the recabling, verify that SCSI reservations are in the correct state.**

   ```
   # scdidadm -R device
   ```

Index

F

H

I

L

M

N

P

R