Oracle® Solaris Cluster 3.3 3/13 Hardware Administration Manual

Part No: E37735–01
March 2013

Contents

Preface

The *Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual* provides a variety of information about how to install and administer basic Oracle Solaris 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 Oracle Solaris Cluster 3.3 software. Unless otherwise noted, procedures are the same for all Oracle Solaris Cluster 3.3 versions.

Note – This Oracle Solaris Cluster release supports systems that use the SPARC and x86 families of processor architectures: UltraSPARC, SPARC64, and AMD64. In this document, the label x86 refers to systems that use the AMD64 family of processor architectures. The information in this document pertains to both platforms unless otherwise specified in a special chapter, section, note, bulleted item, figure, table, or example.

In this document, references to Oracle Real Application Clusters also apply to Oracle Parallel Server unless otherwise stated.

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 an Oracle Solaris Cluster environment nor does it include specific storage device procedures.

Who Should Use This Book

This book is for Oracle representatives who are performing the initial installation of an Oracle Solaris 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 Oracle software and hardware. Do not use this document as a planning or a pre-sales 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 an Oracle Solaris Cluster environment.

Chapter 1, “Introduction to Oracle Solaris Cluster Hardware,” provides an overview of installing and administering Oracle Solaris Cluster hardware.

Chapter 2, “Installing and Configuring the Terminal Concentrator,” describes how to install and configure a terminal concentrator.

Chapter 3, “Installing Cluster Interconnect Hardware and Configuring VLANs,” describes how to install cluster interconnect hardware and configure VLANs.

Chapter 4, “Maintaining Cluster Interconnect Hardware,” describes how to maintain cluster interconnect hardware.

Chapter 5, “Installing and Maintaining Public Network Hardware,” describes how to install and maintain the public network hardware.

Chapter 6, “Maintaining Platform Hardware,” describes how to perform cluster-specific procedures on your cluster node hardware.

Chapter 7, “Campus Clustering With Oracle Solaris Cluster Software,” provides guidelines and diagrams about how to configure a campus cluster.

Chapter 8, “Verifying Oracle Solaris Cluster Hardware Redundancy,” describes how to verify cluster redundancy.

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 Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual

Revision Date	New Information
April 2007	Specifications-Based Campus Clusters, which are described in Chapter 7, “Campus Clustering With Oracle Solaris Cluster Software,” now support a wider range of distance configurations. These clusters support such configurations by requiring compliance to a latency and error rate, rather than to a rigid set of distances and components.

TABLE P–1 Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual *(Continued)*

Revision Date	New Information
July 2007	Specifications-Based Campus Clusters, which are described in Chapter 7, “Campus Clustering With Oracle Solaris Cluster Software,” now support an even wider range of distance configurations, including x64. These clusters support such configurations by requiring compliance to a latency and error rate, rather than to a rigid set of distances and components.
March 2008	Corrected a number of incorrect statements about InfiniBand support, jumbo frames VLANs, and cluster interconnect in Chapter 3, “Installing Cluster Interconnect Hardware and Configuring VLANs,” and Chapter 6, “Maintaining Platform Hardware.”
November 2008	Updated “Interconnect: Requirements When Using Jumbo Frames” section at “Requirements When Using Jumbo Frames” on page 36.
January 2009	Updated links in Preface to cluster documentation.
August 2009	Added index entries for using jumbo frames on an interconnect cluster.
October 2009	Corrected the number of required transport junctions in “Configuring VLANs as Private Interconnect Networks” from two to one.
September 2010	Updated release to 3.3, incorporated Oracle name change, removed old CLI commands, and removed instructions for PCI-SCI and Sun Fire Link because both of these are no longer supported.
May 2011	Updated release to 3.3 5/11, updated links to Oracle sites.
March 2013	Updated release to 3.3 3/13, updated links to Oracle sites.

Related Documentation

The Oracle Solaris Cluster documentation provides conceptual information or procedures to administer hardware and applications. If you plan to use this documentation in a hardcopy format, ensure that you have these books available for your reference. All Oracle Solaris Cluster documentation is available at http://www.oracle.com/technetwork/indexes/documentation/index.html#sys_sw.

For information specifically about your hardware, see the documentation that shipped with the various products. Much of this documentation is also available at http://www.oracle.com/technetwork/indexes/documentation/index.html.

Using UNIX Commands

This document contains information about commands that are used to install, configure, or upgrade an Oracle Solaris Cluster configuration. This document might not contain complete information about 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 Oracle Solaris Operating System (Oracle Solaris OS)
- Other software documentation that you received with your system
- Oracle Solaris Operating System man pages

Getting Help

If you have problems installing or using Oracle Solaris 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, Oracle Solaris 10)
- The release number of Oracle Solaris Cluster (for example, Oracle Solaris Cluster 3.3)

Use the following commands to gather information about 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/clnode show-rev -v`	Displays Oracle Solaris Cluster release and package version information for each node

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

Access to Oracle Support

Oracle customers have access to electronic support through My Oracle Support. For information, visit `http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info` or visit `http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs` if you are hearing impaired.

Typographic Conventions

The following table describes the typographic conventions that are used in this book.

TABLE P–2 Typographic Conventions

Typeface	Description	Example
`AaBbCc123`	The names of commands, files, and directories, and onscreen computer output	Edit your `.login` file. Use `ls -a` to list all files. `machine_name% you have mail.`

`AaBbCc123`	What you type, contrasted with onscreen computer output	`machine_name%` **`su`**
		`Password:`

aabbcc123	Placeholder: replace with a real name or value	The command to remove a file is `rm` *filename*.
AaBbCc123	Book titles, new terms, and terms to be emphasized	Read Chapter 6 in the *User's Guide*. A *cache* is a copy that is stored locally. Do *not* save the file. **Note:** Some emphasized items appear bold online.

Shell Prompts in Command Examples

The following table shows UNIX system prompts and superuser prompts for shells that are included in the Oracle Solaris OS. In command examples, the shell prompt indicates whether the command should be executed by a regular user or a user with privileges.

TABLE P–3 Shell Prompts

Shell	Prompt
Bash shell, Korn shell, and Bourne shell	`$`

TABLE P–3 Shell Prompts *(Continued)*

Shell	Prompt
Bash shell, Korn shell, and Bourne shell for superuser	`#`
C shell	`machine_name%`
C shell for superuser	`machine_name#`

CHAPTER 1

Introduction to Oracle Solaris 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 an Oracle Solaris Cluster environment.

This chapter contains the following information:

Installing Oracle Solaris Cluster Hardware

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

TABLE 1–1 Task Map: Installing Cluster Hardware

Task	For Instructions
Plan for cluster hardware capacity, space, and power requirements.	The site planning documentation that shipped with your nodes and other hardware
Install the nodes.	The documentation that shipped with your nodes
Install the administrative console.	The documentation that shipped with your administrative console

TABLE 1–1 Task Map: Installing Cluster Hardware (Continued)

Task	For Instructions
Install a console access device. Use the procedure that is indicated for your type of console access device. Your server might use a System Service Processor (SSP) as a console access device, rather than a terminal concentrator.	"Installing the Terminal Concentrator" on page 19 or The documentation that shipped with your hardware.
Install the cluster interconnect hardware.	Chapter 3, "Installing Cluster Interconnect Hardware and Configuring VLANs"
Install the public network hardware.	Chapter 5, "Installing and Maintaining Public Network Hardware"
Install and configure the shared disk storage arrays.	Refer to the Oracle Solaris Cluster manual that pertains to your storage device as well as to the device's own documentation.
Install the Oracle Solaris Operating System and Oracle Solaris Clustersoftware.	Oracle Solaris Cluster software installation documentation
Configure the cluster interconnects.	Oracle Solaris Cluster software installation documentation

▼ Installing Oracle Solaris Cluster Hardware

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. See "Hardware Restrictions" on page 18 for critical information about hardware restrictions with Oracle Solaris Cluster.

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, your server might 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 server.

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

For installation instructions, see Chapter 3, "Installing Cluster Interconnect Hardware and Configuring VLANs."

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

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

7 **Install the Oracle Solaris Operating System and Oracle Solaris Cluster software.**

For more information, see the *Oracle Solaris Cluster Software Installation Guide*.

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

For more information, see the Oracle Solaris Cluster data services collection.

Maintaining Oracle Solaris Cluster Hardware

Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual augments documentation that ships with your hardware components by providing information on maintaining the hardware *specifically in an Oracle Solaris Cluster environment*. Table 1–2 describes some of the differences between maintaining cluster hardware and maintaining standalone hardware.

TABLE 1–2 Sample Differences Between Servicing Standalone and Cluster Hardware

Task	Standalone Hardware	Cluster Hardware
Shutting down a node	Use the `shutdown` command.	To perform an orderly node shutdown, first use the `clnode evacuate` to switch device groups and resource groups to another node. Then shut down the node by running the `shutdown(1M)` command.
Adding a disk	Perform a reconfiguration boot or use `devfsadm` 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`, `cldevice populate`, and `cldevice` 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 adapter 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 IPMP group.

Powering Oracle Solaris Cluster Hardware On and Off

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

- Use shut down and boot procedures in the *Oracle Solaris Cluster System Administration Guide* 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 Oracle Solaris Cluster Nodes

The Oracle Solaris Cluster environment supports Oracle Solaris dynamic reconfiguration (DR) operations on qualified servers. This book provides procedures that require that you 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 Oracle Solaris DR feature on your hardware platform *before* you use the DR feature with Oracle Solaris Cluster software. All of the requirements, procedures, and restrictions that are documented for the Oracle Solaris DR feature also apply to Oracle Solaris Cluster DR support (except for the operating environment quiescence operation).

▼ DR Operations in a Cluster With DR-Enabled Servers

Some procedures within this book instruct you 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 steps, including any steps for disabling and removing the transport adapter or PNA from the active cluster interconnect.**

   See the *Oracle Solaris Cluster System Administration Guide* 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.**

   For conceptual information about Oracle Solaris Cluster support of the DR feature, see the *Oracle Solaris Cluster Concepts Guide*.

Local and Multihost Disks in an Oracle Solaris 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 Oracle Solaris Operating System and other nonshared data.
- 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 the *Oracle Solaris Cluster Concepts Guide*.

Removable Media in an Oracle Solaris Cluster Environment

Removable media include tape and CD-ROM drives, which are local devices. *Oracle Solaris Cluster 3.3 3/13 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 are not supported as highly available. Thus, this manual focuses on disk drives as global devices.

Although tape and CD-ROM drives are not supported as highly available 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 an Oracle Solaris Cluster Environment

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

Hardware Restrictions

The following restrictions apply to hardware in all Oracle Solaris Cluster configurations.

- Multihost tape, CD-ROM, and DVD-ROM are not supported.
- Alternate pathing (AP) is not supported.
- Storage devices with more than a single path from a given cluster node to the enclosure are not supported except for the following storage devices:
  - Oracle's Sun StorEdge A3500, for which two paths are supported to each of two nodes.
  - Devices using Solaris I/O multipathing, formerly Sun StorEdge Traffic Manager.
  - EMC storage devices that use EMC PowerPath software.
  - Oracle's Sun StorEdge 9900 storage devices that use HDLM.
- System panics have been observed in clusters when UDWIS I/O cards are used in slot 0 of a board in a Sun Enterprise 10000 server; do not install UDWIS I/O cards in slot 0 of this server.
- Sun VTS software is not supported.

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 an Oracle Solaris Cluster environment. This chapter also includes information about how to use a terminal concentrator.

This chapter contains the following procedures:

- "How to Install the Terminal Concentrator in a Cabinet" on page 20
- "How to Connect the Terminal Concentrator" on page 24
- "How to Configure the Terminal Concentrator" on page 25
- "How to Set Terminal Concentrator Port Parameters" on page 27
- "How to Correct a Port Configuration Access Error" on page 29
- "How to Establish a Default Route for the Terminal Concentrator" on page 30
- "How to Connect to a Node's Console Through the Terminal Concentrator" on page 32
- "How to Reset a Terminal Concentrator Port" on page 33

For conceptual information on console access devices, see the *Oracle Solaris Cluster Concepts Guide*.

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.**

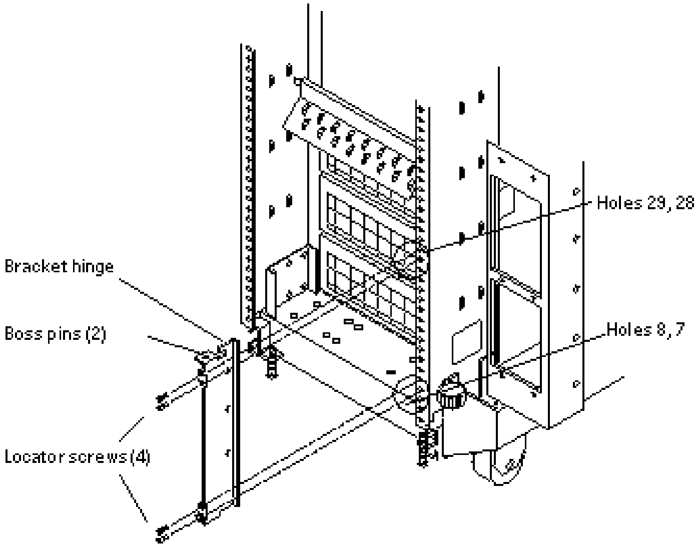
   Thread the screws into holes 8 and 29, as shown in Figure 2–1. The locator screws accept the slotted holes in the hinge piece.

   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



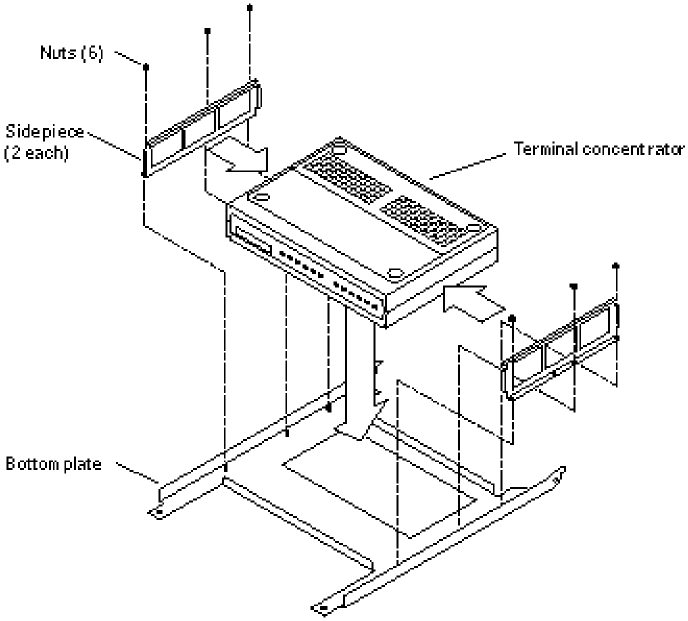
2 **Install the terminal concentrator into the bracket.**

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



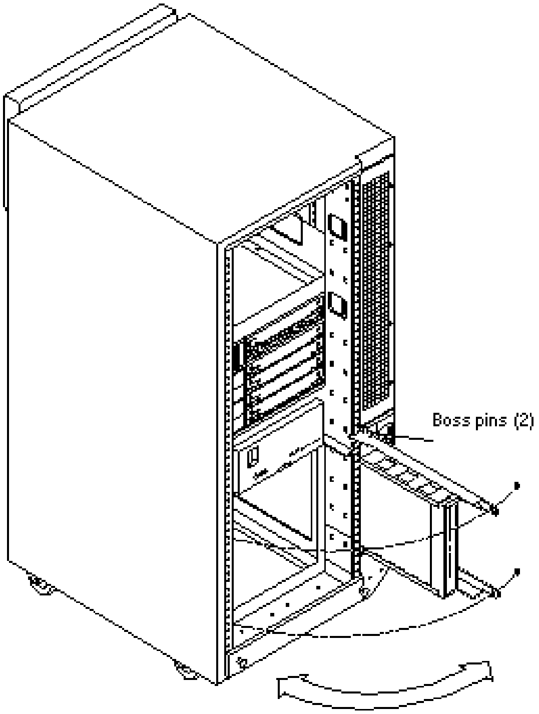
3 **Install the terminal concentrator bracket onto the bracket hinge that is already installed on the cabinet.**

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.**

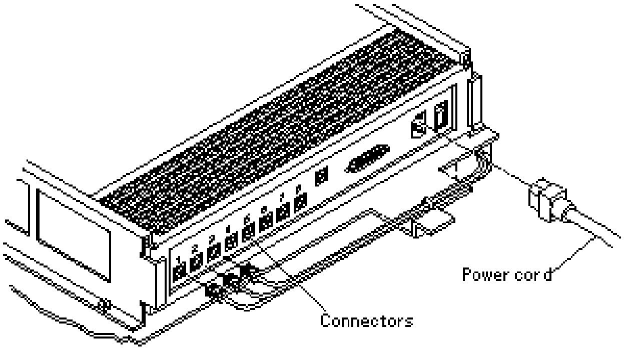
c. **Install the keeper screw in the shorter boss pin to ensure the assembly cannot be accidentally knocked off the hinge.**

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



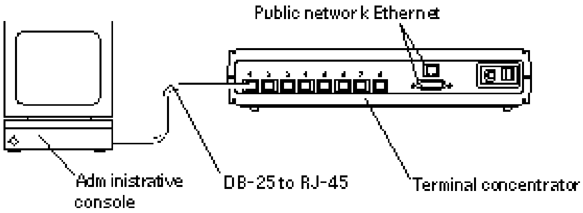
See Also To cable the terminal concentrator, go to "How to Connect the Terminal Concentrator" on page 24.

▼ 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 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). Boot subsystems are described in more detail in "Boot Subsystems" in *Oracle Solaris Administration: Basic Administration*.

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).**

Next Steps 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:**

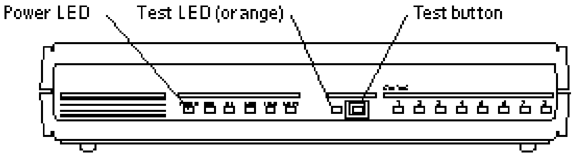
   ```
   # tip tc
   ```

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 and use the information in the following table to decide your course of action.**

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

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

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]
#
```

Next Steps 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.

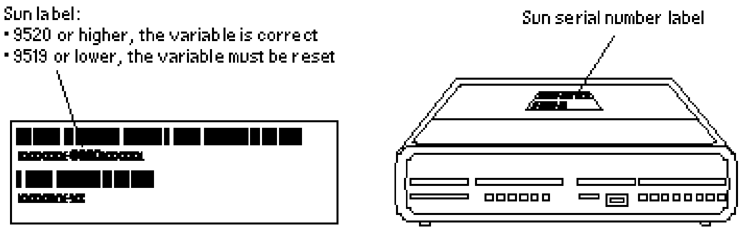
The port type parameter must be set to dial_in. If the parameter is set to hardwired, the cluster console might be unable to detect when a port is already in use.

1 **Locate Oracle's 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 at least 9520, 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


Sun label:
• 9520 or higher, the variable is correct
• 9519 or lower, the variable must be reset
Sun 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 about 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 the message `Port(s) busy, do you wish to wait? (y/n)`, answer **`n`** and go to "How to Reset a Terminal Concentrator Port" on page 33.
   - If you see the message `Error: Permission denied`, 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 2–1 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 2–2 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 2–1 Task Map: Using the Terminal Concentrator

Task	For Instructions
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 into 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 2–3 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 2–4 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 Cluster Interconnect Hardware and Configuring VLANs

This chapter describes the procedures to install cluster interconnect hardware. Where appropriate, this chapter includes separate procedures for the interconnects that Oracle Solaris Cluster software supports:

- Ethernet
- InfiniBand

This chapter contains the following information:

- "Installing Ethernet or InfiniBand Cluster Interconnect Hardware" on page 37
- "Configuring VLANs as Private Interconnect Networks" on page 39

Use the following information to learn more about cluster interconnects:

- For conceptual information about cluster interconnects, see "Cluster Interconnect" in *Oracle Solaris Cluster Concepts Guide*.
- For information about how to administer cluster interconnects, see Chapter 7, "Administering Cluster Interconnects and Public Networks," in *Oracle Solaris Cluster System Administration Guide*.

Interconnect Requirements and Restrictions

This section contains requirements on interconnect operation when using certain special features.

Cluster Interconnect and Routing

Heartbeat packets that are sent over the cluster interconnect are not IP based. As a result, these packets cannot be routed. If you install a router between two cluster nodes that are connected through cluster interconnects, heartbeat packets cannot find their destination. Your cluster consequently fails to work correctly.

To ensure that your cluster works correctly, you must set up the cluster interconnect in the same layer 2 (data link) network and in the same broadcast domain. The cluster interconnect must be located in the same layer 2 network and broadcast domain even if the cluster nodes are located in different, remote data centers. Cluster nodes that are arranged remotely are described in more detail in Chapter 7, “Campus Clustering With Oracle Solaris Cluster Software.”

Cluster Interconnect Speed Requirements

An interconnect path is one network step in the cluster private network: from a node to a node, from a node to a switch, or from the switch to another node. Each path in your cluster interconnect must use the same networking technology.

All interconnect paths must also operate at the same speed. This means, for example, that if you are using Ethernet components that are capable of operating at different speeds, and if your cluster configuration does not allow these components to automatically negotiate a common network speed, you must configure them to operate at the same speed.

Ethernet Switch Configuration When in the Cluster Interconnect

When configuring Ethernet switches for your cluster private interconnect, disable the spanning tree algorithm on ports that are used for the interconnect

Requirements When Using Jumbo Frames

If you use Scalable Data Services and jumbo frames on your public network, ensure that the Maximum Transfer Unit (MTU) of the private network is the same size or larger than the MTU of your public network.

Note – Scalable services cannot forward public network packets that are larger than the MTU size of the private network. The scalable services application instances will not receive those packets.

Consider the following information when configuring jumbo frames:

- The maximum MTU size for an InfiniBand interface is typically less than the maximum MTU size for an Ethernet interface.
- If you use switches in your private network, ensure they are configured to the MTU sizes of the private network interfaces.

For information about how to configure jumbo frames, see the documentation that shipped with your network interface card. See your Oracle Solaris OS documentation or contact your Oracle sales representative for other Oracle Solaris restrictions.

Requirements and Restrictions When Using Sun InfiniBand from Oracle in the Cluster Interconnect

The following requirements and guidelines apply to Oracle Solaris Cluster configurations that use Sun InfiniBand adapters from Oracle:

- A two-node cluster must use InfiniBand switches. You cannot directly connect the InfiniBand adapters to each other.
- Sun InfiniBand switches support up to nine nodes in a cluster.
- Jumbo frames are not supported on a cluster that uses InfiniBand adapters.
- If only one InfiniBand adapter is installed on a cluster node, each of its two ports must be connected to a different InfiniBand switch.
- If two InfiniBand adapters are installed in a cluster node, leave the second port on each adapter unused. For example, connect port 1 on HCA 1 to switch 1 and connect port 1 on HCA 2 to switch 2.
- VLANs are not supported on a cluster that uses InfiniBand switches.

Installing Ethernet or InfiniBand Cluster Interconnect Hardware

The following table lists procedures for installing Ethernet or InfiniBand cluster interconnect hardware. Perform the procedures in the order that they are listed. This section contains the procedure for installing cluster hardware during an *initial installation* of a cluster, before you install Oracle Solaris Cluster software.

TABLE 3–1 Installing Ethernet Cluster Interconnect Hardware

Task	For Instructions
Install the transport adapters.	The documentation that shipped with your nodes and host adapters
Install the transport cables.	"How to Install Ethernet or InfiniBand Transport Cables and Transport Junctions" on page 38
If your cluster contains more than two nodes, install a transport junction (switch).	"How to Install Ethernet or InfiniBand Transport Cables and Transport Junctions" on page 38

▼ How to Install Ethernet or InfiniBand Transport Cables and Transport Junctions

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

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

   See the documentation that shipped with your host adapters and node hardware.

2. **If necessary, install transport junctions and optionally configure the transport junctions' IP addresses.**

   Note – **(InfiniBand Only)** If you install one InfiniBand adapter on a cluster node, two InfiniBand switches are required. Each of the two ports must be connected to a different InfiniBand switch.

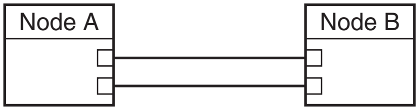
   If two InfiniBand adapters are connected to a cluster node, connect only one port on each adapter to the InfiniBand switch. The second port of the adapter must remain disconnected. Do not connect ports of the two InfiniBand adapters to the same InfiniBand switch.

   ---

3. **Install the transport cables.**

   - **(Ethernet Only) As the following figure shows, a cluster with only two nodes can use a point-to-point connection, requiring no transport junctions.**

     FIGURE 3–1 **(Ethernet Only)** Typical Two-Node Cluster Interconnect

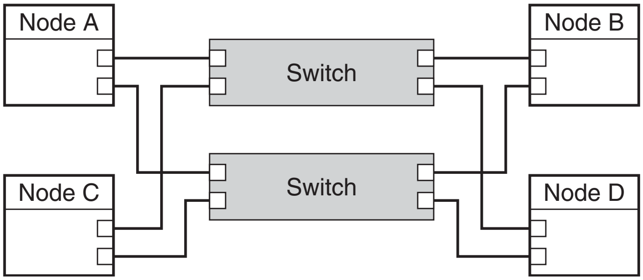
     

     (Ethernet Only) For a point-to-point connection, you can use either UTP or fibre. With fibre, use a standard patch cable. A crossover cable is unnecessary. With UTP, see your network interface card documentation to determine whether you need a crossover cable.

     Note – **(Ethernet Only)** You can optionally use transport junctions in a two-node cluster. If you use a transport junction in a two-node cluster, you can more easily add additional nodes later. To ensure redundancy and availability, always use two transport junctions.

- **As the following figure shows, a cluster with more than two nodes requires transport junctions. These transport junctions are Ethernet or InfiniBand switches (customer-supplied).**

FIGURE 3–2 Typical Four-Node Cluster Interconnect



See Also To install and configure the Oracle Solaris Cluster software with the new interconnect, see Chapter 2, "Installing Software on Global-Cluster Nodes," in *Oracle Solaris Cluster Software Installation Guide*.

(Ethernet Only) To configure jumbo frames on the interconnect, review the requirements in "Requirements When Using Jumbo Frames" on page 36 and see the Sun GigaSwift documentation for instructions.

Configuring VLANs as Private Interconnect Networks

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

Note – Even if clusters share the same switch, create a separate VLAN for each cluster.

By default, Oracle Solaris Cluster uses the same set of IP addresses on the private interconnect. Creating a separate VLAN for each cluster ensures that IP traffic from one cluster does not interfere with IP traffic from another cluster. Unless you have customized the default IP address for the private interconnect, as described in "How to Change the Private Network Address or Address Range of an Existing Cluster" in *Oracle Solaris Cluster System Administration Guide*, create a separate VLAN for each cluster.

The implementation of switch-based VLAN environments is vendor-specific. Because each switch manufacturer implements VLAN differently, the following guidelines address Oracle Solaris Cluster software requirements with regard to configuring 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.

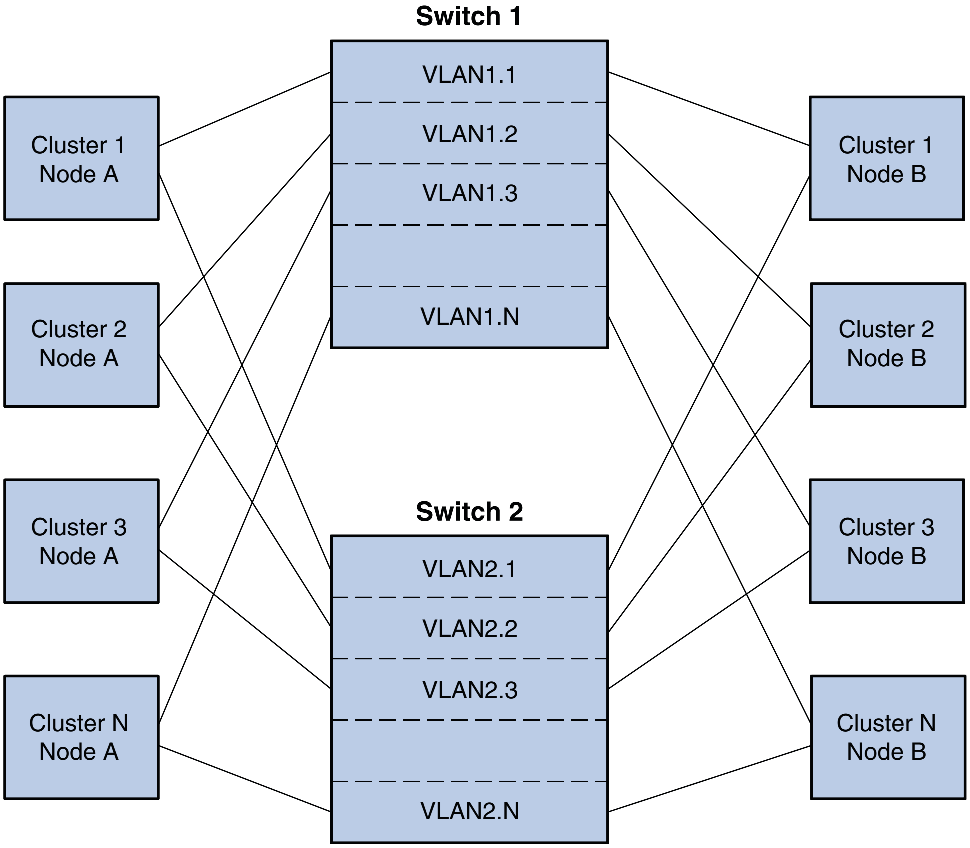
  For the best results, set the Quality of Service (QOS) level for each VLAN to accommodate basic cluster traffic and the desired application traffic. Ensure that the bandwidth that is allocated to each VLAN extends from node to node.

  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 correctly configured your VLANs to prevent path timeouts.

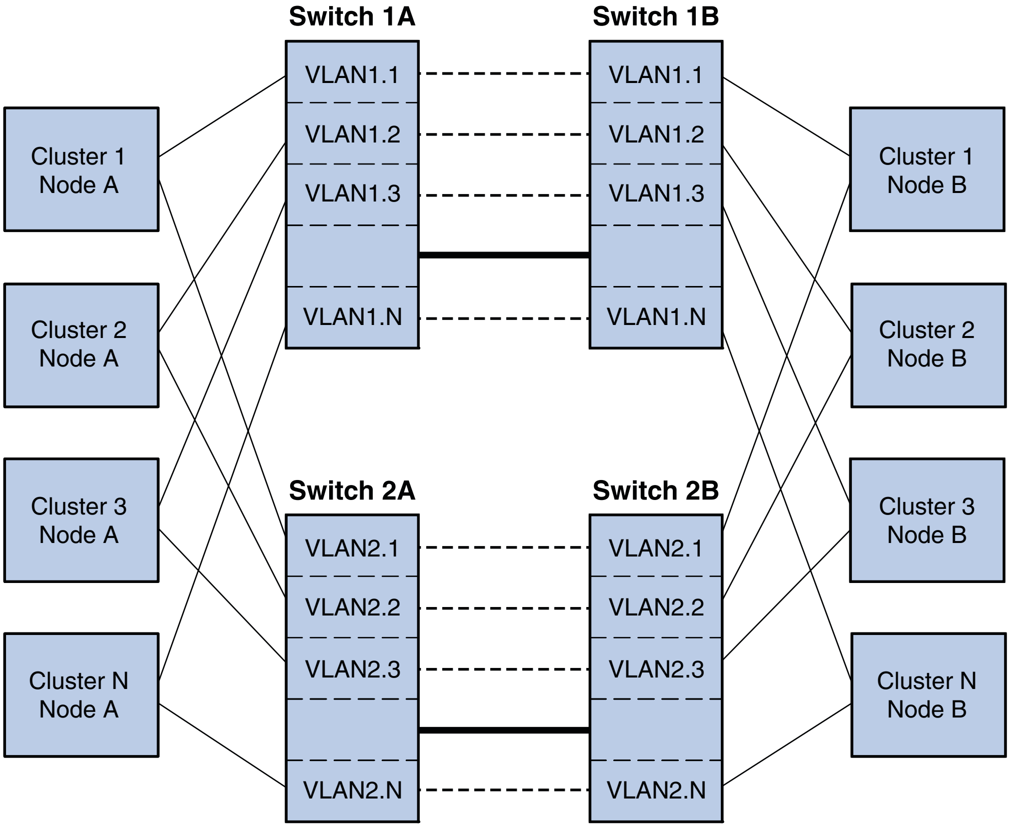
The first VLAN configuration enables nodes from multiple clusters to send interconnect traffic across one pair of Ethernet transport junctions. Oracle Solaris Cluster software requires a minimum of one transport junction, and each transport junction must be part of a VLAN that is located on a different switch. 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 3–3 First VLAN Configuration



The second VLAN configuration uses the same transport junctions for the interconnect traffic of multiple clusters. However, the second VLAN configuration has two pairs of transport junctions that are connected by links. This configuration enables 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 3–4 Second VLAN Configuration



◆ ◆ ◆ CHAPTER 4

Maintaining Cluster Interconnect Hardware

This chapter describes the procedures to maintain cluster interconnect hardware. The procedures in this chapter apply to all interconnects that Oracle Solaris Cluster software supports:

- Ethernet
- InfiniBand

This chapter contains the following procedures:

- "How to Add an Interconnect Component" on page 44
- "How to Replace an Interconnect Component" on page 45
- "How to Remove an Interconnect Component" on page 47
- "How to Upgrade Transport Adapter Firmware" on page 49

For more information, see the following documentation:

- For conceptual information about cluster interconnects, see "Cluster Interconnect" in *Oracle Solaris Cluster Concepts Guide*.
- For information about administering cluster interconnects, see "Administering the Cluster Interconnects" in *Oracle Solaris Cluster System Administration Guide*.

Maintaining Interconnect Hardware in a Running Cluster

The following table lists procedures about maintaining cluster interconnect hardware.

TABLE 4–1 Task Map: Maintaining Cluster Interconnect Hardware

Task	Instructions
Add an interconnect component.	"How to Add an Interconnect Component" on page 44
Replace an interconnect component.	"How to Replace an Interconnect Component" on page 45

TABLE 4–1 Task Map: Maintaining Cluster Interconnect Hardware *(Continued)*

Task	Instructions
Remove an interconnect component.	"How to Remove an Interconnect Component" on page 47
Upgrade transport adapter firmware	"How to Upgrade Transport Adapter Firmware" on page 49

Interconnect components include the following components:

- Transport adapter
- Transport cable
- Transport junction (switch)

▼ How to Add an Interconnect Component

This procedure defines interconnect component as any one of the following components:

- Transport adapter
- Transport cable
- Transport junction (switch)

This section contains the procedure for adding interconnect components to nodes in a running cluster.

Before You Begin This procedure relies on the following prerequisites and assumptions:

- Your cluster is operational and all nodes are powered on.
- 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 39.

1. **Determine if you need to shut down and power off the node that is to be connected to the interconnect component you are adding.**
   - If you are adding a transport junction, you do not need to shut down and power off the node. Proceed to Step 2.
   - If you are adding a transport cable, you do not need to shut down and power off the node. Proceed to Step 2.
   - If your node has Dynamic Reconfiguration (DR) enabled and you are replacing a transport adapter, you do not need to shut down and power off the node. Proceed to Step 2.
   - If your node does *not* have DR enabled and you are adding a transport adapter, shut down and power off the node with the transport adapter you are adding.

For the full procedure about shutting down a node, see Chapter 3, "Shutting Down and Booting a Cluster," in *Oracle Solaris Cluster System Administration Guide*.

2 **Install the interconnect component.**

- If you are using an Ethernet or InfiniBand interconnect, see "How to Install Ethernet or InfiniBand Transport Cables and Transport Junctions" on page 38 for cabling diagrams and considerations.
- For the procedure about installing transport adapters or setting transport adapter DIP switches, see the documentation that shipped with your host adapter and node hardware.
- If your interconnect uses jumbo frames, review the requirements in "Requirements When Using Jumbo Frames" on page 36 and see the Sun GigaSwift documentation for instructions.

3 **If you shut down the node in Step 1, perform a reconfiguration boot to update the new Oracle Solaris device files and links. Otherwise, skip this step.**

See Also

- To reconfigure Oracle Solaris Cluster software with the new interconnect component, see Chapter 7, "Administering Cluster Interconnects and Public Networks," in *Oracle Solaris Cluster System Administration Guide*.

How to Replace an Interconnect Component

This procedure defines interconnect component as any one of the following components:

- Transport adapter
- Transport cable
- Transport junction (switch)

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 `clinterconnect status`command.

For more details about checking the status of the cluster interconnect, see "How to Check the Status of the Cluster Interconnect" in *Oracle Solaris Cluster System Administration Guide*.

You might perform this procedure in the following scenarios:

- You need to replace a failed transport adapter.
- You need to replace a failed transport cable.
- You need to replace a failed transport junction.

For conceptual information about transport adapters, transport cables, and transport junction, see "Cluster Interconnect" in *Oracle Solaris Cluster Concepts Guide*.

Before You Begin This procedure relies on the following prerequisites and assumptions.

- Your cluster has another functional interconnect path to maintain cluster communications while you perform this procedure.
- Your cluster is operational and all nodes are powered on.
- Identify the interconnect component that you want to replace. Remove that interconnect component from the cluster configuration by using the procedure in "How to Remove Cluster Transport Cables, Transport Adapters, and Transport Switches" in *Oracle Solaris Cluster System Administration Guide*.
- 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 39.

1. **Determine if you need to shut down and power off the node that is connected to the interconnect component you are replacing.**

   - If you are replacing a transport junction, you do not need to shut down and power off the node. Proceed to Step 2.
   - If you are replacing a transport cable, you do not need to shut down and power off the node. Proceed to Step 2.
   - If your node has DR enabled and you are replacing a transport adapter, you do not need to shut down and power off the node. Proceed to Step 2.
   - If your node does *not* have DR enabled and you are replacing a transport adapter, shut down and power off the node with the transport adapter you are replacing.

     For the full procedure about how to shut down a node, see Chapter 3, "Shutting Down and Booting a Cluster," in *Oracle Solaris Cluster System Administration Guide*.

2. **Disconnect the failed interconnect component from 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.

3. **Connect the new interconnect component to other cluster devices.**

   - If you are replacing an Ethernet or InfiniBand interconnect, see "How to Install Ethernet or InfiniBand Transport Cables and Transport Junctions" on page 38 for cabling diagrams and considerations.
   - If your interconnect uses jumbo frames, review the requirements in "Requirements When Using Jumbo Frames" on page 36 and see the Sun GigaSwift documentation for instructions. Refer to "ce Sun Ethernet Driver Considerations" on page 54 for details of how to edit the `ce.conf` file according to the GigaSwift documentation's instructions.

4 If you shut down the node in Step 1, perform a reconfiguration boot to update the new Oracle Solaris device files and links. Otherwise, skip this step.

See Also To reconfigure Oracle Solaris Cluster software with the new interconnect component, see "How to Add Cluster Transport Cables, Transport Adapters, or Transport Switches" in *Oracle Solaris Cluster System Administration Guide*.

▼ How to Remove an Interconnect Component

This procedure defines interconnect component as any one of the following components:

- Transport adapter
- Transport cable
- Transport junction (switch)

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 `clinterconnect status`command.

For more details about checking the status of the cluster interconnect, see "How to Check the Status of the Cluster Interconnect" in *Oracle Solaris Cluster System Administration Guide*.

You might perform this procedure in the following scenarios:

- You need to remove an unused transport adapter.
- You need to remove an unused transport cable.
- You need to remove an unused transport junction.
- You want to migrate from a two–node cluster that uses switches to a point-to-point configuration.

For conceptual information about transport adapters, transport cables, and transport junctions, see "Cluster Interconnect" in *Oracle Solaris Cluster Concepts Guide*.

Before You Begin This procedure assumes that your cluster is operational and all nodes are powered on.

Before you perform this procedure, perform the following tasks:

- If you are migrating from a two–node cluster that uses switches to a point-to-point configuration, install a crossover cable before you remove a switch.
- Identify the interconnect component that you want to remove. Remove that interconnect component from the cluster configuration by using the procedure in "How to Remove Cluster Transport Cables, Transport Adapters, and Transport Switches" in *Oracle Solaris Cluster System Administration Guide*.
- If you plan to use virtual local area networks (VLANs) in your cluster interconnect, configure the transport junction. For more information about how configure VLANs, see "Configuring VLANs as Private Interconnect Networks" on page 39.

1 **Determine if you need to shut down and power off the node that is connected to the interconnect component you are removing.**

- If you are removing a transport junction you, do not need to shut down and power off the node. Proceed to Step 2.
- If you are removing a transport cable you, do not need to shut down and power off the node. Proceed to Step 2.
- If your node has DR enabled and you are removing a transport adapter, you do not need to shut down and power off the node. Proceed to Step 2.
- If your node does *not* have DR enabled and you are removing a transport adapter, shut down and power off the node with the transport adapter you are removing.

  For the full procedure about shutting down a node, see Chapter 3, "Shutting Down and Booting a Cluster," in *Oracle Solaris Cluster System Administration Guide*.

2 **Disconnect the interconnect component from 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.

3 **Remove the interconnect component.**

For the procedure about how to remove interconnect component, see the documentation that shipped with your host adapter, nodes, or switch.

4 **If you shut down the node in Step 1, perform a reconfiguration boot to update the new Oracle Solaris device files and links. Otherwise, skip this step.**

See Also To reconfigure Oracle Solaris Cluster software with the new interconnect component, see "How to Add Cluster Transport Cables, Transport Adapters, or Transport Switches" in *Oracle Solaris Cluster System Administration Guide*.

▼ How to Upgrade Transport Adapter Firmware

You might perform this procedure in the following scenarios:

- You want to use firmware bug fixes.
- You want to use new firmware features.

Use this procedure to update transport adapter firmware.

Before You Begin To perform this procedure, become superuser or assume a role that provides `solaris.cluster.read` and `solaris.cluster.modify` role-based access control (RBAC) authorization.

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

   Use the following command:

   ```
   # clresourcegroup status -n nodename
   # cldevicegroup status -n nodename
   ```

   Note the device groups, the resource groups, and the node list for the resource groups. You will need this information to restore the cluster to its original configuration in Step 4.

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

   ```
   # clnode evacuate fromnode
   ```

3. **Perform the firmware upgrade.**

   This process might require you to boot into noncluster mode. If it does, boot the node into cluster mode before proceeding. For the procedure about how to upgrade your transport adapter firmware, see the patch documentation.

4. **If you moved device groups off their original node in Step 2, restore the device groups that you identified in Step 1 to their original node.**

   Perform the following step for each device group you want to return to the original node.

   ```
   # cldevicegroup switch -n nodename devicegroup1[ devicegroup2 ...]
   ```

   `-n` *nodename* — The node to which you are restoring device groups.

   devicegroup1[*devicegroup2* …] — The device group or groups that you are restoring to the node.

   In these commands, *devicegroup* is one or more device groups that are returned to the node.

5 **If you moved resource groups off their original node in Step 2 restore the resource groups that you identified in Step 1 to their original node.**

Perform the following step for each resource group you want to return to the original node.

```
# clresourcegroup switch -n nodename resourcegroup1[ resourcegroup2 ...]
```

nodename
For failover resource groups, the node to which the groups are returned. For scalable resource groups, the node list to which the groups are returned.

resourcegroup1[*resourcegroup2* …]
The resource group or groups that you are returning to the node or nodes.

resourcegroup
The resource group that is returned to the node or nodes.

◆ ◆ ◆ CHAPTER 5

Installing and Maintaining Public Network Hardware

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

- "Public Network Hardware: Requirements When Using Jumbo Frames" on page 51
- "Installing Public Network Hardware" on page 52
- "Maintaining Public Network Hardware in a Running Cluster" on page 52
- "SPARC: Sun Gigabit Ethernet Adapter Considerations" on page 54
- "ce Sun Ethernet Driver Considerations" on page 54

For conceptual information on cluster interconnects and public network interfaces, see the *Oracle Solaris Cluster Concepts Guide*.

For information on how to administer public network interfaces, see the *Oracle Solaris Cluster System Administration Guide*.

Public Network Hardware: Requirements When Using Jumbo Frames

If you use Scalable Data Services and jumbo frames on your public network, ensure that the Maximum Transfer Unit (MTU) of the private network is the same size or larger than the MTU of your public network.

Note – Scalable services cannot forward public network packets that are larger than the MTU size of the private network. The scalable services application instances will not receive those packets.

Consider the following information when configuring jumbo frames:

- The maximum MTU size for an InfiniBand interface is typically less than the maximum MTU size for an Ethernet interface.

- If you use switches in your private network, ensure they are configured to the MTU sizes of the private network interfaces.

For information about how to configure jumbo frames, see the documentation that shipped with your network interface card. See your Oracle Solaris OS documentation or contact your Oracle sales representative for other Oracle Solaris restrictions.

Installing Public Network Hardware

This section covers installing cluster hardware during an *initial cluster installation, before Oracle Solaris 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.

Installing Public Network Hardware: 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 Oracle Solaris Cluster Hardware" on page 13.

If your network uses jumbo frames, review the requirements in "Public Network Hardware: Requirements When Using Jumbo Frames" on page 51 and see the Sun GigaSwift documentation for information about how to configure jumbo frames.

Maintaining Public Network Hardware in a Running Cluster

The following table lists procedures about how to maintain public network hardware.

TABLE 5–1 Task Map: Maintaining Public Network Hardware

Task	Information
Add public network adapters.	"Adding Public Network Adapters" on page 53
Replace public network adapters.	"Replacing Public Network Adapters" on page 53
Remove public network adapters.	"Removing Public Network Adapters" on page 53

Adding 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.

Once the adapters are physically installed, Oracle Solaris Cluster requires that they be configured in an IPMP group.

If your network uses jumbo frames, review the requirements in "Public Network Hardware: Requirements When Using Jumbo Frames" on page 51 and see the documentation that shipped with your network interface card for information about how to configure jumbo frames.

Adding Public Network Adapters: Where to Go From Here

To add a new public network adapter to an IPMP group, see *Oracle Solaris Administration: IP Services*.

Replacing Public Network Adapters

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

For procedures about how to administer public network connections, see the *Oracle Solaris Administration: IP Services*.

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

Replacing Public Network Adapters: Where to Go From Here

To add the new public network adapter to a IPMP group, see the *Oracle Solaris Cluster System Administration Guide*.

Removing Public Network Adapters

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

For procedures about how to administer public network connections, see Chapter 27, "Introducing IPMP (Overview)," in *Oracle Solaris Administration: IP Services*.

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

SPARC: 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.

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"
```

OnOracle Solaris 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 software driver for the Sun GigaSwift Ethernet adapter is known as the Cassini Ethernet (`ce`) driver. The Oracle Solaris Cluster software supports the `ce` driver for cluster interconnect and public network applications. Consult your Oracle service representative for details about the network interface products that are supported.

When you use the `ce` Sun Ethernet driver for the private cluster interconnect, add the following kernel parameters to the `/etc/system` file on all the nodes in the cluster to avoid communication problems over the private cluster interconnect.

```
set ce:ce_taskq_disable=1
set ce:ce_ring_size=1024
set ce:ce_comp_ring_size=4096
```

If you do not set these three kernel parameters when using the `ce` driver for the private cluster interconnect, one or more of the cluster nodes might panic due to a loss of communication between the nodes of the cluster. In these cases, check for the following panic messages.

```
Reservation conflict
CMM: Cluster lost operational quorum; aborting
CMM: Halting to prevent split brain with node name
```

If you are using the ce driver and your cluster interconnect uses a back-to-back connection, do not disable auto-negotiation. If you must disable auto-negotiation, when you want to force 1000 Mbit operation for example, manually specify the link master, or clock master, for the connection.

When manually specifying the link master, you must set one side of the back-to-back connection to provide the clock signal and the other side to use this clock signal. Use the ndd(1M) command to manually specify the link master and follow the guidelines listed below.

- Set the link_master or master_cfg_value parameter to 1 (clock master) on one side of the back-to-back connection and to 0 on the other side.
- Specify the link_master parameter for ce driver versions up to and including 1.118.
- Specify the master_cfg_value parameter for ce driver versions that are released after 1.118.
- Set the master_cfg_value parameter to 1.

To determine the version of the ce driver, use the modinfo command, as shown in the following example.

```
# modinfo | grep ce
84 78068000  4e016 222   1  ce (CE Ethernet Driver v1.148)
```

EXAMPLE 5–1 Using the ndd Command When You Want to Force 1000 Mbit Operation

This example shows how to use the ndd command when you want to force 1000 Mbit operation with a back-to-back connection and the version of the ce driver is lower than or equal to 1.118.

```
# ndd -set /dev/ce link_master 0
```

This example shows how to use the ndd command when you want to force 1000 Mbit operation with a back-to-back connection and the version of the ce driver is greater than or equal to 1.119.

```
# ndd -set /dev/ce master_cfg_enable 1
# ndd -set /dev/ce master_cfg_value 0
```

SPARC: GigaSwift Ethernet Driver and Jumbo Frames

If you are using jumbo frames, you must edit the ce.conf file to configure them, as explained in the Sun GigaSwift documentation.

The driver documentation instructs you to grep certain entries from the /etc/path_to_inst file to determine your entries for the ce.conf file. An entry modified for an Oracle Solaris Cluster node resembles the following:

```
# grep ce /etc/path_to_inst
"/node@1/pci@8,600000/network@1" 0 "ce"
```

When editing the `ce.conf` file, remove the */node@nodeID* identifier prefix from the entries that you put into the driver configuration file. For the example above, the entry to put into the configuration file is:

```
"/pci@8,600000/network@1" 0 "ce"
```

◆ ◆ ◆ CHAPTER 6

Maintaining Platform Hardware

This chapter contains information about node hardware in a cluster environment. It contains the following topics:

- "Mirroring Internal Disks on Servers that Use Internal Hardware Disk Mirroring or Integrated Mirroring" on page 57
- "Configuring Cluster Nodes With a Single, Dual-Port HBA" on page 61

Mirroring Internal Disks on Servers that Use Internal Hardware Disk Mirroring or Integrated Mirroring

Some servers support the mirroring of internal hard drives (internal hardware disk mirroring or integrated mirroring) to provide redundancy for node data. To use this feature in a cluster environment, follow the steps in this section.

The best way to set up hardware disk mirroring is to perform RAID configuration during cluster installation, before you configure multipathing. For instructions on performing this configuration, see the *Oracle Solaris Cluster Software Installation Guide*. If you need to change your mirroring configuration after you have established the cluster, you must perform some cluster-specific steps to clean up the device IDs, as described in the procedure that follows.

Note – Specific servers might have additional restrictions. See the documentation that shipped with your server hardware.

For specifics about how to configure your server's internal disk mirroring, refer to the documents that shipped with your server and the `raidctl(1M)` man page.

How to Configure Internal Disk Mirroring After the Cluster Is Established

Before You Begin This procedure assumes that you have already installed your hardware and software and have established the cluster. To configure an internal disk mirror during cluster installation, see the *Oracle Solaris Cluster Software Installation Guide*.

The Oracle Enterprise Manager Ops Center software helps you patch and monitor your data center assets. Oracle Enterprise Manager Ops Center helps improve operational efficiency and ensures that you have the latest software patches for your software. Contact your Oracle representative to purchase Oracle Enterprise Manager Ops Center.

Additional information for using the Oracle patch management tools is provided in *Oracle Solaris Administration Guide: Basic Administration* on the Oracle Technology Network. Refer to the version of this manual for the Oracle Solaris OS release that you have installed.

If you must apply a patch when a node is in noncluster mode, you can apply it in a rolling fashion, one node at a time, unless instructions for a patch require that you shut down the entire cluster. Follow the procedures in "How to Apply a Rebooting Patch (Node)" in *Oracle Solaris Cluster System Administration Guide* to prepare the node and to boot it in noncluster mode. For ease of installation, consider applying all patches at the same time. That is, apply all patches to the node that you place in noncluster mode.

For required firmware, see the *Oracle Technology Network*.

Caution – If there are state database replicas on the disk that you are mirroring, you must recreate them during this procedure.

1. **If necessary, prepare the node for establishing the mirror.**

   a. **Determine the resource groups and device groups that are running on the node.**

   Record this information because you use it later in this procedure to return resource groups and device groups to the node.

   Use the following command:

   ```
   # clresourcegroup status -n nodename
   # cldevicegroup status -n nodename
   ```

   b. **If necessary, move all resource groups and device groups off the node.**

   ```
   # clnode evacuate fromnode
   ```

2. **Configure the internal mirror.**

   ```
   # raidctl -c clt0d0 clt1d0
   ```

-c *clt0d0 clt1d0* Creates the mirror of primary disk to the mirror disk. Enter the name of your primary disk as the first argument. Enter the name of the mirror disk as the second argument.

3 **Boot the node into single user mode.**

```
# reboot -- -S
```

4 **Clean up the device IDs.**

Use the following command:

```
# cldevice repair /dev/rdsk/clt0d0
```

/dev/rdsk/clt0d0 Updates the cluster's record of the device IDs for the primary disk. Enter the name of your primary disk as the argument.

5 **Confirm that the mirror has been created and only the primary disk is visible to the cluster.**

```
# cldevice list
```

The command lists only the primary disk, and not the mirror disk, as visible to the cluster.

6 **Boot the node back into cluster mode.**

```
# reboot
```

7 **If you are using Solaris Volume Manager and if the state database replicas are on the primary disk, recreate the state database replicas.**

```
# metadb -a /dev/rdsk/clt0d0s4
```

8 **If you moved device groups off the node in Step 1, restore device groups to the original node.**

Perform the following step for each device group you want to return to the original node.

```
# cldevicegroup switch -n nodename devicegroup1[ devicegroup2 ...]
```

-n *nodename* The node to which you are restoring device groups.

devicegroup1[*devicegroup2* …] The device group or groups that you are restoring to the node.

9 **If you moved resource groups off the node in Step 1, move all resource groups back to the node.**

Perform the following step for each resource group you want to return to the original node.

```
# clresourcegroup switch -n nodename resourcegroup1[ resourcegroup2 ...]
```

nodename For failover resource groups, the node to which the groups are returned. For scalable resource groups, the node list to which the groups are returned.

resourcegroup1[*resourcegroup2* …] The resource group or groups that you are returning to the node or nodes.

▼ How to Remove an Internal Disk Mirror

1 If necessary, prepare the node for removing the mirror.

a. Determine the resource groups and device groups that are running on the node.

Record this information because you use this information later in this procedure to return resource groups and device groups to the node.

Use the following command:

```
# clresourcegroup status -n nodename
# cldevicegroup status -n nodename
```

b. If necessary, move all resource groups and device groups off the node.

```
# clnode evacuate fromnode
```

2 Remove the internal mirror.

```
# raidctl -d clt0d0
```

`-d` *clt0d0* Deletes the mirror of primary disk to the mirror disk. Enter the name of your primary disk as the argument.

3 Boot the node into single user mode.

```
# reboot -- -S
```

4 Clean up the device IDs.

Use the following command:

```
# cldevice repair /dev/rdsk/clt0d0 /dev/rdsk/clt1d0
```

/dev/rdsk/clt0d0 /dev/rdsk/clt1d0 Updates the cluster's record of the device IDs. Enter the names of your disks separated by spaces.

5 Confirm that the mirror has been deleted and that both disks are visible.

```
# cldevice list
```

The command lists both disks as visible to the cluster.

6 Boot the node back into cluster mode.

```
# reboot
```

7 If you are using Solaris Volume Manager and if the state database replicas are on the primary disk, recreate the state database replicas.

```
# metadb -c 3 -ag /dev/rdsk/clt0d0s4
```

8 **If you moved device groups off the node in Step 1, restore the device groups to the original node.**

```
# cldevicegroup switch -n nodename devicegroup1 devicegroup2 ...
```

-n *nodename*
The node to which you are restoring device groups.

devicegroup1[*devicegroup2* …]
The device group or groups that you are restoring to the node.

9 **If you moved resource groups off the node in Step 1, restore the resource groups and device groups to the original node.**

Perform the following step for each resource group you want to return to the original node.

```
# clresourcegroup switch -n nodename resourcegroup[ resourcegroup2 ...]
```

nodename
For failover resource groups, the node to which the groups are restored. For scalable resource groups, the node list to which the groups are restored.

resourcegroup[*resourcegroup2* …]
The resource group or groups that you are restoring to the node or nodes.

Configuring Cluster Nodes With a Single, Dual-Port HBA

This section explains the use of dual-port host bus adapters (HBAs) to provide both connections to shared storage in the cluster. While Oracle Solaris Cluster supports this configuration, it is less redundant than the recommended configuration. You *must* understand the risks that a dual-port HBA configuration poses to the availability of your application, if you choose to use this configuration.

This section contains the following topics:

- "Risks and Trade-offs When Using One Dual-Port HBA" on page 62
- "Supported Configurations When Using a Single, Dual-Port HBA" on page 62
- "Cluster Configuration When Using Solaris Volume Manager and a Single Dual-Port HBA" on page 63
- "Cluster Configuration When Using Solaris Volume Manager for Oracle Solaris Cluster and a Single Dual-Port HBA" on page 63

Risks and Trade-offs When Using One Dual-Port HBA

You should strive for as much separation and hardware redundancy as possible when connecting each cluster node to shared data storage. This approach provides the following advantages to your cluster:

- The best assurance of high availability for your clustered application
- Good failure isolation
- Good maintenance robustness

Oracle Solaris Cluster is usually layered on top of a volume manager, mirrored data with independent I/O paths, or a multipathed I/O link to a hardware RAID arrangement. Therefore, the cluster software does not expect a node ever to ever lose access to shared data. These redundant paths to storage ensure that the cluster can survive any single failure.

Oracle Solaris Cluster does support certain configurations that use a single, dual-port HBA to provide the required two paths to the shared data. However, using a single, dual-port HBA for connecting to shared data increases the vulnerability of your cluster. If this single HBA fails and takes down both ports connected to the storage device, the node is unable to reach the stored data. How the cluster handles such a dual-port failure depends on several factors:

- The cluster configuration
- The volume manager configuration
- The node on which the failure occurs
- The state of the cluster when the failure occurs

If you choose one of these configurations for your cluster, you must understand that the supported configurations mitigate the risks to high availability and the other advantages. The supported configurations do not eliminate these previously mentioned risks.

Supported Configurations When Using a Single, Dual-Port HBA

Oracle Solaris Cluster supports the following volume manager configurations when you use a single, dual-port HBA for connecting to shared data:

- Solaris Volume Manager with more than one disk in each diskset and no dual-string mediators configured. For details about this configuration, see “Cluster Configuration When Using Solaris Volume Manager and a Single Dual-Port HBA” on page 63.
- Solaris Volume Manager for Oracle Solaris Cluster. For details about this configuration, see “Cluster Configuration When Using Solaris Volume Manager for Oracle Solaris Cluster and a Single Dual-Port HBA” on page 63.

Cluster Configuration When Using Solaris Volume Manager and a Single Dual-Port HBA

If the Solaris Volume Manager metadbs lose replica quorum for a diskset on a cluster node, the volume manager panics the cluster node. Oracle Solaris Cluster then takes over the diskset on a surviving node and your application fails over to a secondary node.

To ensure that the node panics and is fenced off if it loses its connection to shared storage, configure each metaset with at least two disks. In this configuration, the metadbs stored on the disks create their own replica quorum for each diskset.

Dual-string mediators are not supported in Solaris Volume Manager configurations that use a single dual-port HBA. Using dual-string mediators prevents the service from failing over to a new node.

Configuration Requirements

When configuring Solaris Volume Manager metasets, ensure that each metaset contains at least two disks. Do not configure dual-string mediators.

Expected Failure Behavior with Solaris Volume Manager

When a dual-port HBA fails with both ports in this configuration, the cluster behavior depends on whether the affected node is primary for the diskset.

- If the affected node is primary for the diskset, Solaris Volume Manager panics that node because it lacks required state database replicas. Your cluster reforms with the nodes that achieve quorum and brings the diskset online on a new primary node.
- If the affected node is not primary for the diskset, your cluster remains in a degraded state.

Failure Recovery with Solaris Volume Manager

Follow the instructions for replacing an HBA in your storage device documentation.

Cluster Configuration When Using Solaris Volume Manager for Oracle Solaris Cluster and a Single Dual-Port HBA

Because Solaris Volume Manager for Oracle Solaris Cluster uses raw disks only and is specific to Oracle Real Application Clusters (RAC), no special configuration is required.

Expected Failure Behavior with Solaris Volume Manager for Oracle Solaris Cluster

When a dual-port HBA fails and takes down both ports in this configuration, the cluster behavior depends on whether the affected node is the current master for the multi-owner diskset.

- If the affected node is the current master for the multi-owner diskset, the node does not panic. If any other node fails or is rebooted, the affected node will panic when it tries to update the replicas. The volume manager chooses a new master for the diskset if the surviving nodes can achieve quorum.
- If the affected node is not the current master for the multi-owner diskset, the node remains up but the device group is in a degraded state. If an additional failure affects the master node and Solaris Volume Manager for Oracle Solaris Cluster attempts to remaster the diskset on the node with the failed paths, that node will also panic. A new master will be chosen if any surviving nodes can achieve quorum.

Failure Recovery with Solaris Volume Manager for Oracle Solaris Cluster

Follow the instructions for replacing an HBA in your storage device documentation.

Kernel Cage DR Recovery

When you perform a Dynamic Reconfiguration (DR) remove operation on a memory board with kernel cage memory, the affected node becomes unresponsive so heartbeat monitoring for that node is suspended on all other nodes and the node's quorum vote count is set to 0. After DR is completed, the heartbeat monitoring of the affected node is automatically re-enabled and the quorum vote count is reset to 1. If the DR operation does not complete, you might need to manually recover. For general information about DR, see “Dynamic Reconfiguration Support” in *Oracle Solaris Cluster Concepts Guide*.

The `monitor-heartbeat` subcommand is not supported in an exclusive-IP zone cluster. For more information about this command, see the `cluster`(1CL) man page.

Preparing the Cluster for Kernel Cage DR

When you use a DR operation to remove a system board containing kernel cage memory (memory used by the Oracle Solaris OS), the system must be quiesced in order to allow the memory contents to be copied to another system board. In a clustered system, the tight coupling between cluster nodes means that the quiescing of one node for repair can cause operations on non-quiesced nodes to be delayed until the repair operation is complete and the node is unquiesced. For this reason, using DR to remove a system board containing kernel cage memory from a cluster node requires careful planning and preparation.

Use the following information to reduce the impact of the DR quiesce on the rest of the cluster:

- I/O operations for file systems or global device groups with their primary or secondary on the quiesced node will hang until the node is unquiesced. If possible, ensure that the node being repaired is not the primary for any global file systems or device groups.
- I/O to SVM multi-owner disksets that include the quiesced node will hang until the node is unquiesced.
- Updates to the CCR require communication between all cluster members. Any operations that result in CCR updates should not be performed while the DR operation is ongoing. Configuration changes are the most common cause of CCR updates.
- Many cluster commands result in communication among cluster nodes. Refrain from running cluster commands during the DR operation.
- Applications and cluster resources on the node being quiesced will be unavailable for the duration of the DR event. The time required to move applications and resources to another node should be weighed against the expected outage time of the DR event.
- Scalable applications such as Oracle RAC often have a different membership standard, and have communication and synchronization actions among members. Scalable application instances on the node to be repaired should be brought offline before you initiate the DR operation.

▼ How to Recover From an Interrupted Kernel Cage DR Operation

If the DR operation does not complete, perform the following steps to re-enable heartbeat timeout monitoring for that node and to reset the quorum vote count.

1 **If DR does not complete successfully, manually re-enable heartbeat timeout monitoring.**

From a single cluster node (which is not the node where the DR operation was performed), run the following command.

```
# cluster monitor-heartbeat
```

Use this command only in the global zone. Messages display indicating that monitoring has been enabled.

2 **If the node that was dynamically reconfigured paused during boot, allow it to finish booting and join the cluster membership.**

If the node is at the `ok` prompt, boot it now.

3 **Verify that the node is now part of the cluster membership and check the quorum vote count of the cluster nodes by running the following command on a single node in the cluster.**

```
# clquorum status
--- Quorum Votes by Node (current status) ---

Node Name       Present       Possible       Status
---------       -------       --------       ------
pnode1          1             1              Online
pnode2          1             1              Online
pnode3          0             0              Online
```

4 **If one of the nodes has a vote count of 0, reset its vote count to 1 by running the following command on a single node in the cluster.**

```
# clquorum votecount -n nodename 1
```

nodename The hostname of the node that has a quorum vote count of 0.

5 **Verify that all nodes now have a quorum vote count of 1.**

```
# clquorum status
--- Quorum Votes by Node (current status) ---

Node Name       Present       Possible       Status
---------       -------       --------       ------
pnode1          1             1              Online
pnode2          1             1              Online
pnode3          1             1              Online
```

CHAPTER 7

Campus Clustering With Oracle Solaris Cluster Software

In campus clustering, nodes or groups of nodes are located in separate rooms, sometimes several kilometers apart. In addition to providing the usual benefits of using an Oracle Solaris Cluster, correctly designed campus clusters can generally survive the loss of any single room and continue to provide their services.

This chapter introduces the basic concepts of campus clustering and provides some configuration and setup examples. The following topics are covered:

- "Requirements for Designing a Campus Cluster" on page 67
- "Guidelines for Designing a Campus Cluster" on page 70
- "Determining Campus Cluster Connection Technologies" on page 77
- "Installing and Configuring Interconnect, Storage, and Fibre Channel Hardware" on page 78
- "Additional Campus Cluster Configuration Examples" on page 79

This chapter does not explain clustering, provide information about clustering administration, or furnish details about hardware installation and configuration. For conceptual and administrative information, see the *Oracle Solaris Cluster Concepts Guide* and the *Oracle Solaris Cluster System Administration Guide*.

Requirements for Designing a Campus Cluster

When designing your campus cluster, all of the requirements for a standard cluster still apply. Plan your cluster to eliminate any single point of failure in nodes, cluster interconnect, data storage, and public network. Just as in the standard cluster, a campus cluster requires redundant connections and switches. Disk multipathing helps ensure that each node can access each shared storage device. These concerns are universal for Oracle Solaris Cluster.

After you have a valid cluster plan, follow the requirements in this section to ensure a correct campus cluster. To achieve maximum benefits from your campus cluster, consider implementing the "Guidelines for Designing a Campus Cluster" on page 70.

Note – This chapter describes ways to design your campus cluster using fully tested and supported hardware components and transport technologies. You can also design your campus cluster according to Oracle Solaris Cluster's specification, regardless of the components used.

To build a specifications-based campus cluster, contact your Oracle representative, who will assist you with the design and implementation of your specific configuration. This process ensures that the configuration that you implement complies with the specification guidelines, is interoperable, and is supportable.

Selecting Networking Technologies

Your campus cluster must observe all requirements and limitations of the technologies that you choose to use. “Determining Campus Cluster Connection Technologies” on page 77 provides a list of tested technologies and their known limitations.

When planning your cluster interconnect, remember that campus clustering requires redundant network connections.

Connecting to Storage

A campus cluster must include at least two rooms using two independent SANs to connect to the shared storage. See Figure 7–1 for an illustration of this configuration.

If you are using Oracle Real Application Clusters (RAC), all nodes that support Oracle RAC must be fully connected to the shared storage devices. Also, all rooms of a specifications-based campus cluster must be fully connected to the shared storage devices.

See “Quorum in Clusters With Four Rooms or More” on page 75 for a description of a campus cluster with both direct and indirect storage connections.

Sharing Data Storage

Your campus cluster must use SAN-supported storage devices for shared storage. When planning the cluster, ensure that it adheres to the SAN requirements for all storage connections. See the SAN Solutions documentation site (`http://www.oracle.com/technetwork/indexes/documentation/index.html`) for information about SAN requirements.

Oracle Solaris Cluster software supports two methods of data replication: host-based replication and storage-based replication. Host-based data replication can mirror a campus cluster's shared data. If one room of the cluster is lost, another room must be able to provide access to the data. Therefore, mirroring between shared disks must always be performed across rooms, rather than within rooms. Both copies of the data should never be located in a single room. Host-based data replication can be a less expensive solution because it uses locally-attached disks and does not require special storage arrays.

An alternative to host-based replication is storage-based replication, which moves the work of data replication off the cluster nodes and onto the storage device. Storage-based data replication can simplify the infrastructure required, which can be useful in campus cluster configurations.

For more information on both types of data replication and supported software, see Chapter 4, “Data Replication Approaches,” in *Oracle Solaris Cluster System Administration Guide*.

Complying With Quorum Device Requirements

You must use a quorum device for a two-node cluster. For larger clusters, a quorum device is optional. These are standard cluster requirements.

Note – In Oracle Solaris Cluster 3.3, a quorum device can be a storage device or a quorum server.

In addition, you can configure quorum devices to ensure that specific rooms can form a cluster in the event of a failure. For guidelines about where to locate your quorum device, see “Deciding How to Use Quorum Devices” on page 74.

Replicating Solaris Volume Manager Disksets

If you use Solaris Volume Manager as your volume manager for shared device groups, carefully plan the distribution of your replicas. In two-room configurations, all disksets should be configured with an additional replica in the room that houses the cluster quorum device.

For example, in three-room two-node configurations, a single room houses both the quorum device and at least one extra disk that is configured in each of the disksets. Each diskset should have extra replicas in the third room.

Note – You can use a quorum disk for these replicas.

Refer to your Solaris Volume Manager documentation for details about configuring diskset replicas.

Guidelines for Designing a Campus Cluster

In planning a campus cluster, your goal is to build a cluster that can at least survive the loss of a room and continue to provide services. The concept of a room must shape your planning of redundant connectivity, storage replication, and quorum. Use the following guidelines to assist in managing these design considerations.

Determining the Number of Rooms in Your Cluster

The concept of a room, or location, adds a layer of complexity to the task of designing a campus cluster. Think of a *room* as a functionally independent hardware grouping, such as a node and its attendant storage, or a quorum device that is physically separated from any nodes. Each room is 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 failure to safeguard against, as described in the following table.

TABLE 7–1 Definitions of "Room"

Failure Scenario	Sample Definitions of "Room"
Power-line failure	Isolated and independent power supplies
Minor accidents, furniture collapse, water seepage	Different parts of a physical room
Small fire, fire sprinklers starting	Different physical areas (for example, sprinkler zone)
Structural failure, building-wide fire	Different buildings
Large-scale natural disaster (for example, earthquake or flood)	Different corporate campuses up to several kilometers apart

Oracle Solaris Cluster does support two-room campus clusters. These clusters are valid and might offer nominal insurance against disasters. However, consider adding a small third room, possibly even a secure closet or vault (with a separate power supply and correct cabling), to contain the quorum device or a third server.

Whenever a two-room campus cluster loses a room, it has only a 50 percent chance of remaining available. If the room with fewest quorum votes is the surviving room, the surviving nodes cannot form a cluster. In this case, your cluster requires manual intervention from your Oracle service provider before it can become available.

The advantage of a three-room or larger cluster is that, if any one of the three rooms is lost, automatic failover can be achieved. Only a correctly configured three-room or larger campus cluster can guarantee system availability if an entire room is lost (assuming no other failures).

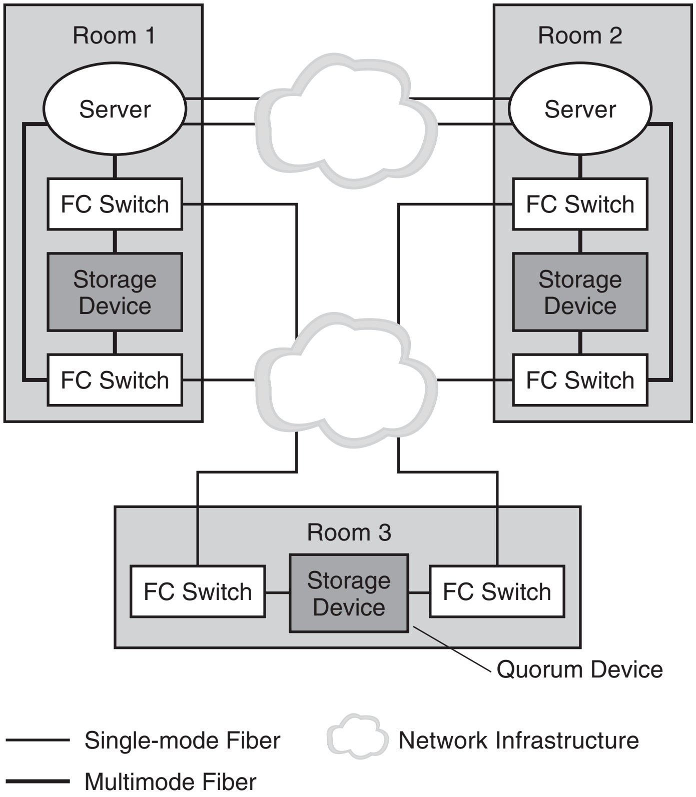
Three-Room Campus Cluster Examples

A three-room campus cluster configuration supports up to eight nodes. Three rooms enable you to arrange your nodes and quorum device so that your campus cluster can reliably survive the loss of a single room and still provide cluster services. The following example configurations all follow the campus cluster requirements and the design guidelines described in this chapter.

- Figure 7–1 shows a three-room, two-node campus cluster. In this arrangement, two rooms each contain a single node and an equal number of disk arrays to mirror shared data. The third room contains at least one disk subsystem, attached to both nodes and configured with a quorum device.
- Figure 7–2 shows an alternative three-room, two-node campus cluster.
- Figure 7–3 shows a three-room, three-node cluster. In this arrangement, two rooms each contain one node and an equal number of disk arrays. The third room contains a small server, which eliminates the need for a storage array to be configured as a quorum device.

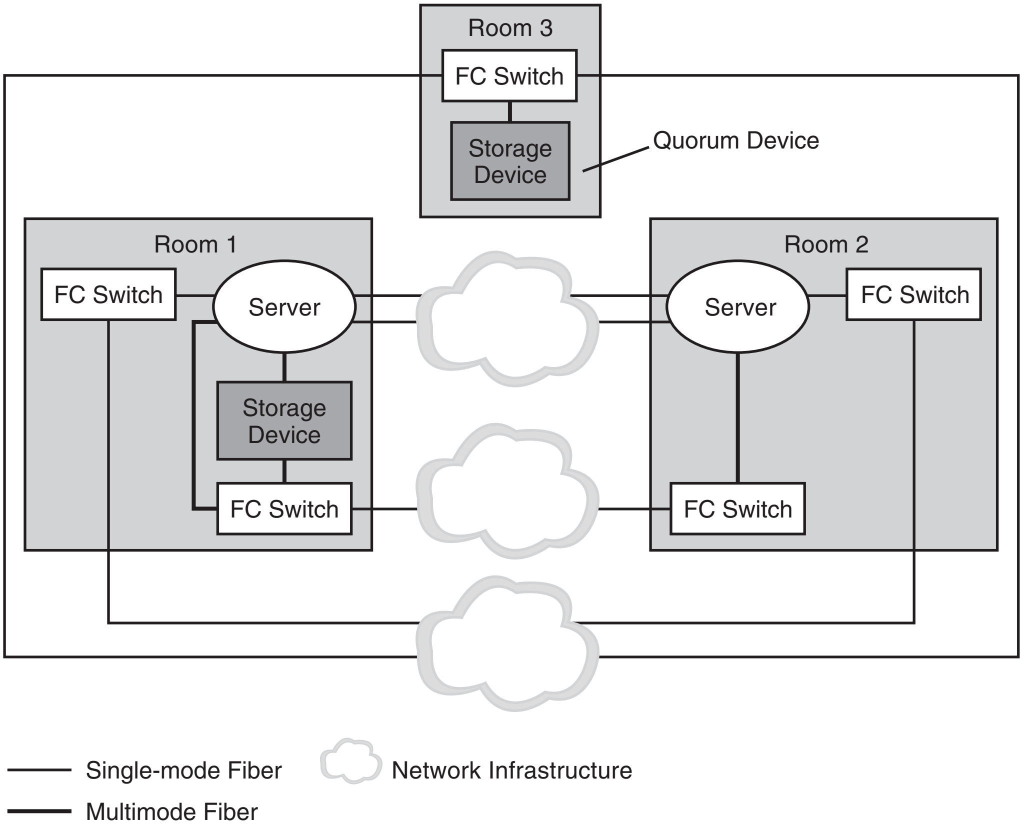
Note – These examples illustrate general configurations and are not intended to indicate required or recommended setups. For simplicity, the diagrams and explanations concentrate only on features that are unique to understanding campus clustering. For example, public-network Ethernet connections are not shown.

FIGURE 7–1 Basic Three-Room, Two-Node Campus Cluster Configuration With Multipathing



In the configuration that is shown in the following figure, if at least two rooms are up and communicating, recovery is automatic. Only three-room or larger configurations can guarantee that the loss of any one room can be handled automatically.

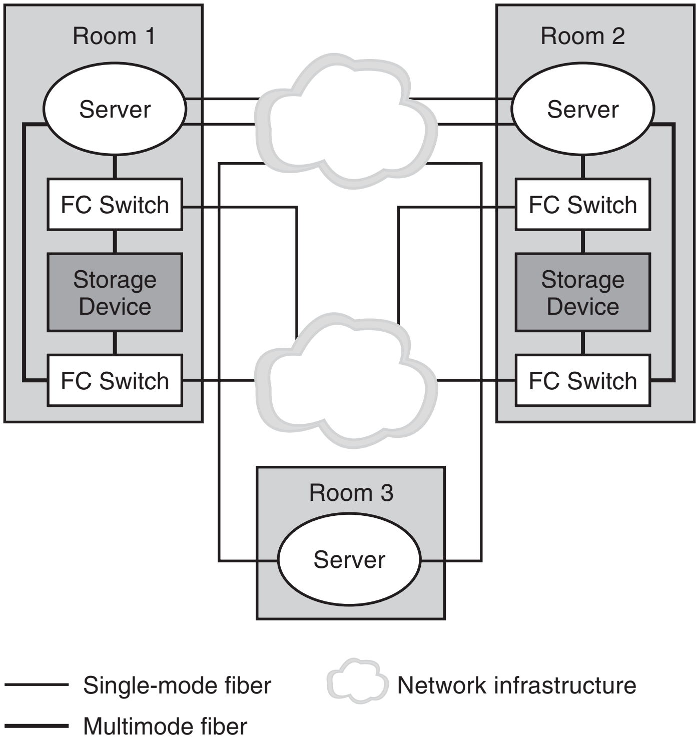
FIGURE 7–2 Minimum Three-Room, Two-Node Campus Cluster Configuration Without Multipathing



In the configuration shown in the following figure, one room contains one node and shared storage. A second room contains a cluster node only. The third room contains shared storage only. A LUN or disk of the storage device in the third room is configured as a quorum device.

This configuration provides the reliability of a three-room cluster with minimum hardware requirements. This campus cluster can survive the loss of any single room without requiring manual intervention.

FIGURE 7–3 Three-Room, Three-Node Campus Cluster Configuration



In the configuration that is shown in the preceding figure, a server acts as the quorum vote in the third room. This server does not necessarily support data services. Instead, it replaces a storage device as the quorum device.

Deciding How to Use Quorum Devices

When adding quorum devices to your campus cluster, your goal should be to balance the number of quorum votes in each room. No single room should have a much larger number of votes than the other rooms because loss of that room can bring the entire cluster down.

For campus clusters with more than three rooms and three nodes, quorum devices are optional. Whether you use quorum devices in such a cluster, and where you place them, depends on your assessment of the following:

- Your particular cluster topology
- The specific characteristics of the rooms involved
- Resiliency requirements for your cluster

As with two-room clusters, locate the quorum device in a room you determine is more likely to survive any failure scenario. Alternatively, you can locate the quorum device in a room that you *want* to form a cluster, in the event of a failure. Use your understanding of your particular cluster requirements to balance these two criteria.

Refer to the *Oracle Solaris Cluster Concepts Guide* for general information about quorum devices and how they affect clusters that experience failures. If you decide to use one or more quorum devices, consider the following recommended approach:

1. For each room, total the quorum votes (nodes) for that room.
2. Define a quorum device in the room that contains the lowest number of votes and that contains a fully connected shared storage device.

When your campus cluster contains more than two nodes, *do not* define a quorum device if each room contains the same number of nodes.

The following sections discuss quorum devices in various sizes of campus clusters.

- "Quorum in Clusters With Four Rooms or More" on page 75
- "Quorum in Three-Room Configurations" on page 77
- "Quorum in Two-Room Configurations" on page 77

Quorum in Clusters With Four Rooms or More

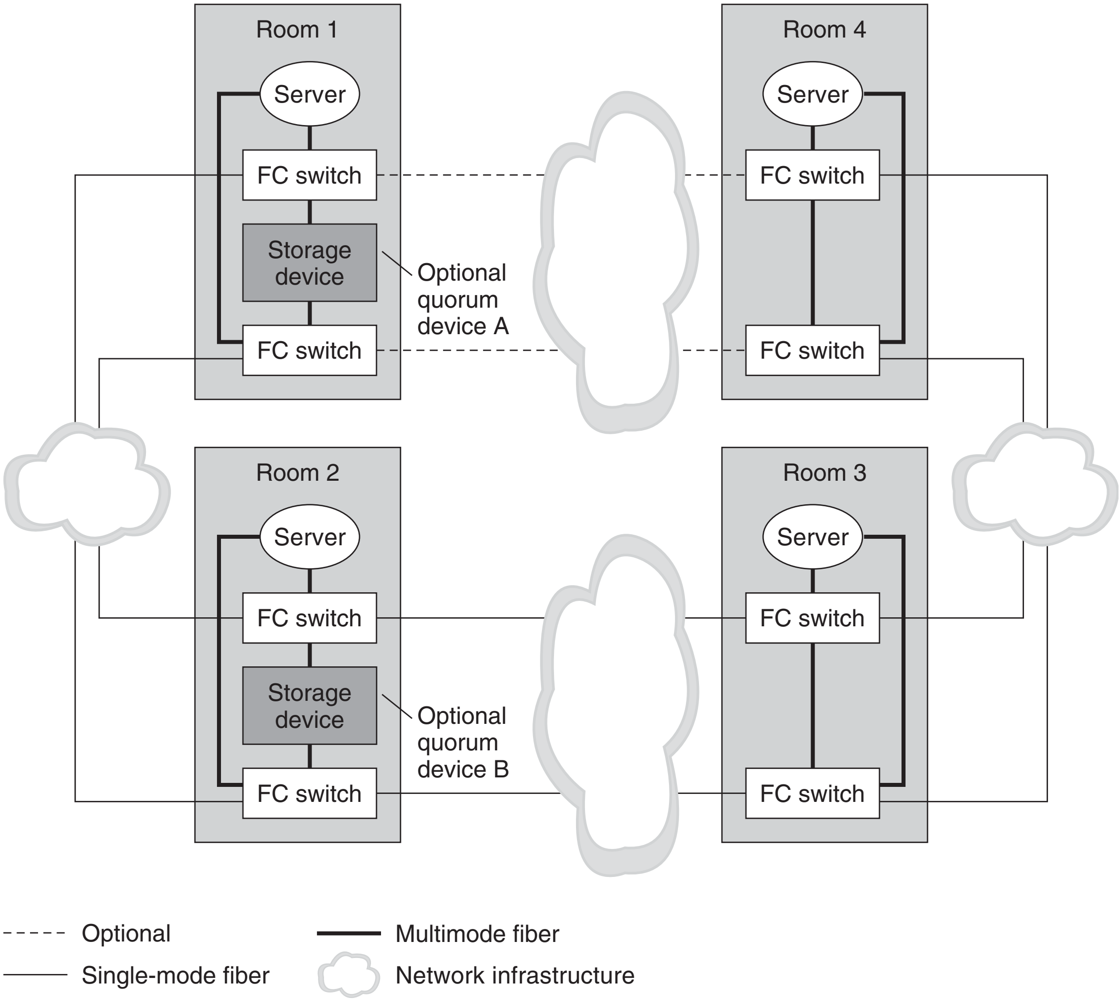
The following figure illustrates a four-node campus cluster with fully connected storage. Each node is in a separate room. Two rooms also contain the shared storage devices, with data mirrored between them.

Note that the quorum devices are marked *optional* in the illustration. This cluster does not require a quorum device. With no quorum devices, the cluster can still survive the loss of any single room.

Consider the effect of adding *Quorum Device A*. Because the cluster contains four nodes, each with a single quorum vote, the quorum device receives three votes. Four votes (one node and the quorum device, or all four nodes) are required to form the cluster. This configuration is not optimal, because the loss of *Room 1* brings down the cluster. The cluster is not available after the loss of that single room.

If you then add *Quorum Device B*, both *Room 1* and *Room 2* have four votes. Six votes are required to form the cluster. This configuration is clearly better, as the cluster can survive the random loss of any single room.

FIGURE 7–4 Four-Room, Four-Node Campus Cluster



Note – In Figure 7–4, the cluster interconnect is not shown.

Consider the optional I/O connection between *Room 1* and *Room 4*. Although fully connected storage is preferable for reasons of redundancy and reliability, fully redundant connections might not always be possible in campus clusters. Geography might not accommodate a particular connection, or the project's budget might not cover the additional fiber.

In such a case, you can design a campus cluster with indirect access between some nodes and the storage. In Figure 7–4, if the optional I/O connection is omitted, *Node 4* must access the storage indirectly.

Quorum in Three-Room Configurations

In three-room, two-node campus clusters, you should use the third room for the quorum device (Figure 7–1) or a server (Figure 7–3). Isolating the quorum device gives your cluster a better chance to maintain availability after the loss of one room. If at least one node and the quorum device remain operational, the cluster can continue to operate.

Quorum in Two-Room Configurations

In two-room configurations, the quorum device occupies the same room as one or more nodes. Place the quorum device in the room that is more likely to survive a failure scenario if all cluster transport and disk connectivity are lost between rooms. If *only* cluster transport is lost, the node that shares a room with the quorum device is not necessarily the node that reserves the quorum device first. For more information about quorum and quorum devices, see the *Oracle Solaris Cluster Concepts Guide*.

Determining Campus Cluster Connection Technologies

This section lists example technologies for the private cluster interconnect and for the data paths and their various distance limits. In some cases, it is possible to extend these limits. For more information, ask your Oracle representative.

Cluster Interconnect Technologies

The following table lists example node-to-node link technologies and their limitations.

TABLE 7–2 Campus Cluster Interconnect Technologies and Distance Limits

Link Technology	Maximum Distance	Comments
100 Mbps Ethernet	100 meters per segment	unshielded twisted pair (UTP)
1000 Mbps Ethernet	100 meters per segment	UTP
1000 Mbps Ethernet	260 meters per segment	62.5/125 micron multimode fiber (MMF)
1000 Mbps Ethernet	550 meters per segment	50/125 micron MMF
1000 Mbps Ethernet (FC)	10 kilometers at 1 Gbps	9/125 micron single-mode fiber (SMF)

TABLE 7–2 Campus Cluster Interconnect Technologies and Distance Limits *(Continued)*

Link Technology	Maximum Distance	Comments
DWDM	200 kilometers and up	
Other		Consult your Oracle representative

Always check your vendor documentation for technology-specific requirements and limitations.

Storage Area Network Technologies

The following table lists example link technologies for the cluster data paths and the distance limits for a single interswitch link (ISL).

TABLE 7–3 ISL Limits

Link Technology	Maximum Distance	Comments
FC short-wave gigabit interface converter (GBIC)	500 meters at 1 Gbps	50/125 micron MMF
FC long-wave GBIC	10 kilometers at 1 Gbps	9/125 micron SMF
FC short-wave small form-factor pluggable (SFP)	300 meters at 2 Gbps	62.5/125 micron MMF
FC short-wave SFP	500 meters at 2 Gbps	62.5/125 micron MMF
FC long-wave SFP	10 kilometers at 2 Gbps	9/125 micron SMF
DWDM	200 kilometers and up	
Other		Consult your Oracle representative

Installing and Configuring Interconnect, Storage, and Fibre Channel Hardware

Generally, using interconnect, storage, and Fibre Channel (FC) hardware does not differ markedly from standard cluster configurations.

The steps for installing Ethernet-based campus cluster interconnect hardware are the same as the steps for standard clusters. Refer to “Installing Ethernet or InfiniBand Cluster Interconnect Hardware” on page 37. When installing the media converters, consult the accompanying documentation, including requirements for fiber connections.

The guidelines for installing virtual local area networks interconnect networks are the same as the guidelines for standard clusters. See "Configuring VLANs as Private Interconnect Networks" on page 39.

The steps for installing shared storage are the same as the steps for standard clusters.

Campus clusters 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 the steps for standard clusters.

If your switch supports flexibility in the buffer allocation mechanism, (for example the QLogic switch with donor ports), make certain you allocate a sufficient number of buffers to the ports that are dedicated to interswitch links (ISLs). If your switch has a fixed number of frame buffers (or buffer credits) per port, you do not have this flexibility.

Calculating Buffer Credits

The following rules determine the number of buffers that you might need:

- For 1 Gbps, calculate buffer credits as:

  `(`*length-in-km*`) x (0.6)`

  Round the result up to the next whole number. For example, a 10 km connection requires 6 buffer credits, and a 7 km connection requires 5 buffer credits.
- For 2 Gbps, calculate buffer credits as:

  `(`*length-in-km*`) x (1.2)`

  Round the result up to the next whole number. For example, a 10 km connection requires 12 buffer credits, while a 7 km connection requires 9 buffer credits.

For greater speeds or for more details, refer to your switch documentation for information about computing buffer credits.

Additional Campus Cluster Configuration Examples

While detailing all of the configurations that are possible in campus clustering is beyond the scope of this document, the following illustrations depict variations on the configurations that were previously shown.

- Three-room campus cluster with a multipathing solution implemented (Figure 7–5)
- Two-room campus cluster with a multipathing solution implemented (Figure 7–6)
- Two-room campus cluster without a multipathing solution implemented (Figure 7–7)

FIGURE 7–5 Three-Room Campus Cluster With a Multipathing Solution Implemented

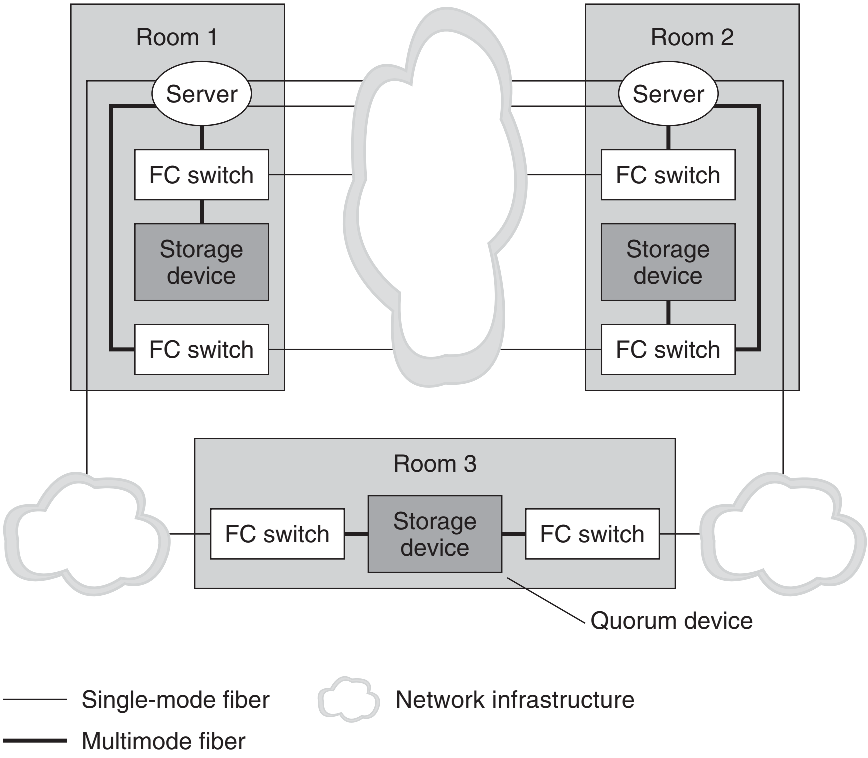
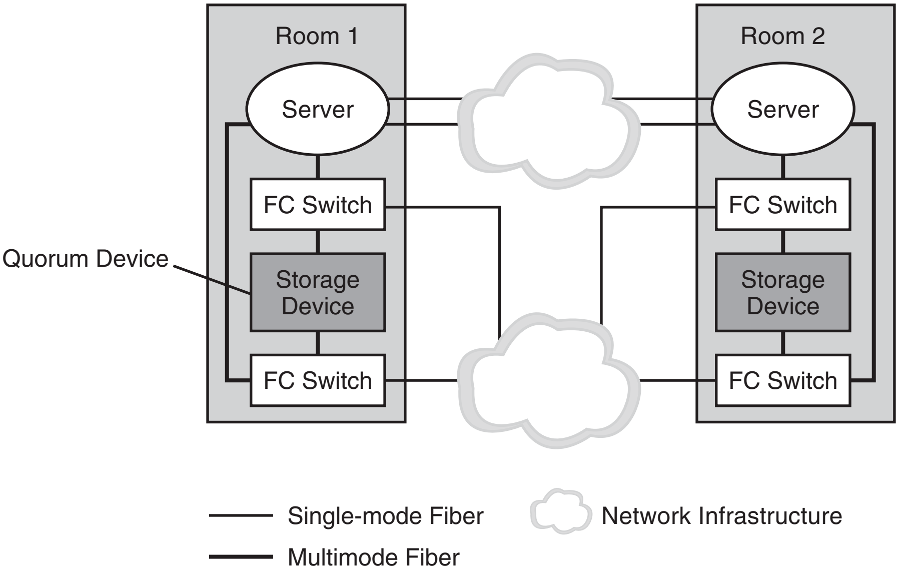


Figure 7–6 shows a two-room campus cluster that uses partner pairs of storage devices and four FC switches, with a multipathing solution implemented. The four switches are added to the cluster for greater redundancy and potentially better I/O throughput. Other possible configurations that you could implement include using Oracle's Sun StorEdge T3 partner groups or Oracle's Sun StorEdge 9910/9960 arrays with Solaris I/O multipathing software installed.

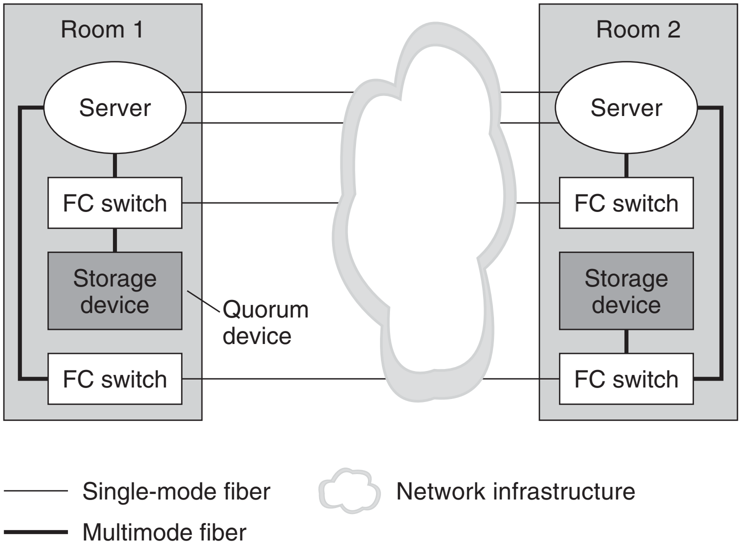
For information about Solaris I/O multipathing software for the Oracle Solaris 10 OS, see the *Solaris Fibre Channel Storage Configuration and Multipathing Support Guide*.

FIGURE 7–6 Two-Room Campus Cluster With a Multipathing Solution Implemented



The configuration in the following figure could be implemented by using Oracle's Sun StorEdge T3 or T3+ arrays in single-controller configurations, rather than partner groups.

FIGURE 7–7 Two-Room Campus Cluster Without a Multipathing Solution Implemented



CHAPTER 8

Verifying Oracle Solaris Cluster Hardware Redundancy

This chapter describes the tests for verifying and demonstrating the high availability (HA) of your Oracle Solaris Cluster configuration. The tests in this chapter assume that you installed Oracle Solaris Cluster hardware, the Oracle Solaris Operating System, and Oracle Solaris Cluster software. All nodes should be booted as cluster members.

This chapter 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 IPMP groups are not single points of failure.

To perform the tests in "How to Test Device Group Redundancy Using Resource Group Failover" on page 84 and "How to Test Cluster Interconnects" on page 85, 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 node and a secondary node for a particular device group.

Use the following command to determine the initial primary and secondary:`cldevicegroup status` with the `-n` option.

For conceptual information about primary nodes, secondary nodes, failover, device groups, or cluster hardware, see your Oracle Solaris Cluster concepts documentation.

Testing Node Redundancy

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

▼ How to Test Device Group Redundancy Using Resource Group Failover

Before You Begin To perform this procedure, become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.

1 **Create an HAStoragePlus resource group with which to test.**

Use the following command:

```
# clresourcegroup create testgroup
# clresourcetype register SUNW.HAStoragePlus
# clresource create -t HAStoragePlus -g testgroup \
  -p GlobalDevicePaths=/dev/md/red/dsk/d0 \
  -p Affinityon=true testresource
```

`clresourcetype register` If the `HAStoragePlus` resource type is not already registered, register it.

/dev/md/red/dsk/d0 Replace this path with your device path.

2 **Identify the node that masters the `testgroup`.**

```
# clresourcegroup status testgroup
```

3 **Power off the primary node for the `testgroup`.**

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

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

Use the following command to check the output for the resource group ownership:

```
# clresourcegroup status testgroup
```

5 **Power on the initial primary node. 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 cluster.

6 **From the initial primary node, return ownership of the resource group to the initial primary node.**

```
# clresourcegroup switch -n nodename testgroup
```

In these commands, *nodename* is the name of the primary node.

7. **Verify that the initial primary node has ownership of the resource group.**

   Use the following command to look for the output that shows the device group ownership.

   ```
   # clresourcegroup status testgroup
   ```

Testing Cluster Interconnect Redundancy

This section provides the procedure for testing cluster interconnect redundancy.

▼ How to Test Cluster Interconnects

Before You Begin To perform this procedure, become superuser or assume a role that provides `solaris.cluster.read` and `solaris.cluster.modify` RBAC authorization.

1. **Disconnect one of the cluster transport cables from a node in the cluster.**

   Messages similar to the following appear on the consoles of each node and are logged in the `/var/adm/messages` file.

```
Nov  4 08:27:21 node1 genunix: WARNING: ce1: fault detected external to device; service degraded
Nov  4 08:27:21 node1 genunix: WARNING: ce1: xcvr addr:0x01 - link down
Nov  4 08:27:31 node1 genunix: NOTICE: clcomm: Path node1:ce1 - node1:ce0 being cleaned up
Nov  4 08:27:31 node1 genunix: NOTICE: clcomm: Path node1:ce1 - node1:ce0 being drained
Nov  4 08:27:31 node1 genunix: NOTICE: clcomm: Path node1:ce1 - node1:ce0 being constructed
Nov  4 08:28:31 node1 genunix: NOTICE: clcomm: Path node1:ce1 - node1:ce0 errors during initiation
Nov  4 08:28:31 node1 genunix: WARNING: Path node1:ce1 - node1:ce0 initiation
encountered errors, errno = 62.
  Remote node may be down or unreachable through this path.
```

2. **Verify that Oracle Solaris Cluster has registered that the interconnect is down.**

   Use the following command to verify that the interconnect path displays as `Faulted`.

   ```
   # clinterconnect status
   ```

3. **Reconnect the cluster transport cable**

   Messages similar to the following appear on the consoles of each node and are logged in the `/var/adm/messages` file.

```
Nov  4 08:30:26 node1 genunix: NOTICE: ce1: fault cleared external to device; service available
Nov  4 08:30:26 node1 genunix: NOTICE: ce1: xcvr addr:0x01 - link up 1000 Mbps full duplex
Nov  4 08:30:26 node1 genunix: NOTICE: clcomm: Path node1:ce1 - node1:ce0 being initiated
Nov  4 08:30:26 node1 genunix: NOTICE: clcomm: Path node1:ce1 - node1:ce0 online
```

4 **Verify that Oracle Solaris Cluster has registered that the interconnect is up.**

Use the following command to verify that the interconnect path displays as `Online`.

```
# clinterconnect status
```

5 **Repeat Step 1 through Step 4 on each cluster transport cable in the node.**

6 **Repeat Step 1 through Step 5 on each node in the cluster.**

Testing Public Network Redundancy

This section provides the procedure for testing public network redundancy.

▼ How to Test Public Network Redundancy

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

Before You Begin To perform this procedure, become superuser or assume a role that provides `solaris.cluster.read` RBAC authorization.

1 **Create a logical hostname resource group which is the failover hostname to use the IPMP groups on the system.**

Use the following command:

```
# clresourcegroup create lhtestgroup
# clreslogicalhostname create -g lhtestgroup  logicalhostname
# clresourcegroup online lhtestgroup
```

logicalhostname The IP address that is hosted on the device on which an IPMP group is configured.

2 **Determine the adapter on which the *logicalhostname* exists.**

```
# ifconfig -a
```

3 **Disconnect one public network cable from the adapter you identified in Step 2.**

4 **If there are no more adapters in the group, skip to Step 7.**

5 **If there is another adapter in the group, verify that the logical hostname failed over to that adapter.**

```
# ifconfig -a
```

6 **Continue to disconnect adapters in the group, until you have disconnected the last adapter.**

The resource group (`lhtestgroup`) should fail over to the secondary node.

7 **Verify that the resource group failed over to the secondary node.**

Use the following command:

```
# clnode status lhtestgroup
```

8 **Reconnect all adapters in the group.**

9 **From the initial primary node, return ownership of the resource group to the initial primary node.**

```
# clresourcegroup switch -n nodename lhtestgroup
```

In these commands, *nodename* is the name of the original primary node.

10 **Verify that the resource group is running on the original primary node.**

Use the following command:

```
# clnode status lhtestgroup
```

Index