



Sun Cluster 3.0 System Administration Guide

Sun Microsystems, Inc.
901 San Antonio Road
Palo Alto, CA 94303-4900
U.S.A. 650-960-1300

Part Number 806-1423
November 2000, Revision A

Copyright Copyright 2000 Sun Microsystems, Inc. 901 San Antonio Road, Palo Alto, California 94303-4900 U.S.A. All rights reserved.

This product or document is protected by copyright and distributed under licenses restricting its use, copying, distribution, and decompilation. No part of this product or document may be reproduced in any form by any means without prior written authorization of Sun and its licensors, if any. Third-party software, including font technology, is copyrighted and licensed from Sun suppliers.

Parts of the product may be derived from Berkeley BSD systems, licensed from the University of California. UNIX is a registered trademark in the U.S. and other countries, exclusively licensed through X/Open Company, Ltd. For Netscape Communicator™, the following notice applies: (c) Copyright 1995 Netscape Communications Corporation. All rights reserved.

Sun, Sun Microsystems, the Sun logo, AnswerBook2, docs.sun.com, Solstice DiskSuite, Sun Enterprise, Sun Enterprise SyMON, JumpStart, StorEdge, Sun Management Center, OpenBoot, and Solaris are trademarks, registered trademarks, or service marks of Sun Microsystems, Inc. in the U.S. and other countries. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc.

The OPEN LOOK and Sun™ Graphical User Interface was developed by Sun Microsystems, Inc. for its users and licensees. Sun acknowledges the pioneering efforts of Xerox in researching and developing the concept of visual or graphical user interfaces for the computer industry. Sun holds a non-exclusive license from Xerox to the Xerox Graphical User Interface, which license also covers Sun's licensees who implement OPEN LOOK GUIs and otherwise comply with Sun's written license agreements.

RESTRICTED RIGHTS: Use, duplication, or disclosure by the U.S. Government is subject to restrictions of FAR 52.227-14(g)(2)(6/87) and FAR 52.227-19(6/87), or DFAR 252.227-7015(b)(6/95) and DFAR 227.7202-3(a).

DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID.

Copyright 2000 Sun Microsystems, Inc., 901 San Antonio Road, Palo Alto, Californie 94303 Etats-Unis. Tous droits réservés.

Ce produit ou document est protégé par un copyright et distribué avec des licences qui en restreignent l'utilisation, la copie, la distribution, et la décompilation. Aucune partie de ce produit ou document ne peut être reproduite sous aucune forme, par quelque moyen que ce soit, sans l'autorisation préalable et écrite de Sun et de ses bailleurs de licence, s'il y en a. Le logiciel détenu par des tiers, et qui comprend la technologie relative aux polices de caractères, est protégé par un copyright et licencié par des fournisseurs de Sun.

Des parties de ce produit pourront être dérivées des systèmes Berkeley BSD licenciés par l'Université de Californie. UNIX est une marque déposée aux Etats-Unis et dans d'autres pays et licenciée exclusivement par X/Open Company, Ltd. La notice suivante est applicable à Netscape Communicator™: (c) Copyright 1995 Netscape Communications Corporation. Tous droits réservés.

Sun, Sun Microsystems, le logo Sun, AnswerBook2, docs.sun.com, Solstice DiskSuite, Sun Enterprise, Sun Enterprise SyMON, OpenBoot, et Solaris sont des marques de fabrique ou des marques déposées, ou marques de service, de Sun Microsystems, Inc. aux Etats-Unis et dans d'autres pays. Toutes les marques SPARC sont utilisées sous licence et sont des marques de fabrique ou des marques déposées de SPARC International, Inc. aux Etats-Unis et dans d'autres pays. Les produits portant les marques SPARC sont basés sur une architecture développée par Sun Microsystems, Inc.

L'interface d'utilisation graphique OPEN LOOK et Sun™ a été développée par Sun Microsystems, Inc. pour ses utilisateurs et licenciés. Sun reconnaît les efforts de pionniers de Xerox pour la recherche et le développement du concept des interfaces d'utilisation visuelle ou graphique pour l'industrie de l'informatique. Sun détient une licence non exclusive de Xerox sur l'interface d'utilisation graphique Xerox, cette licence couvrant également les licenciés de Sun qui mettent en place l'interface d'utilisation graphique OPEN LOOK et qui en outre se conforment aux licences écrites de Sun.

CETTE PUBLICATION EST FOURNIE "EN L'ETAT" ET AUCUNE GARANTIE, EXPRESSE OU IMPLICITE, N'EST ACCORDEE, Y COMPRIS DES GARANTIES CONCERNANT LA VALEUR MARCHANDE, L'APTITUDE DE LA PUBLICATION A REpondre A UNE UTILISATION PARTICULIERE, OU LE FAIT QU'ELLE NE SOIT PAS CONTREFAISANTE DE PRODUIT DE TIERS. CE DENI DE GARANTIE NE S'APPLIQUERAIT PAS, DANS LA MESURE OU IL SERAIT TENU JURIDIQUEMENT NUL ET NON AVENU.



Contents

Preface	9
1. Introduction to Administering Sun Cluster	13
1.1 Administering Sun Cluster Overview	13
1.2 Graphical User Interface	14
1.3 Command-Line Interface	15
1.4 Preparing to Administer the Cluster	16
1.4.1 Documenting a Sun Cluster Hardware Configuration	16
1.4.2 Using an Administrative Console	17
1.4.3 Backing Up the Cluster	17
1.5 Beginning to Administer the Cluster	17
▼ How to Remotely Log In to Sun Cluster	19
▼ How to Access the <code>scsetup</code> Utility	21
▼ How to Display Sun Cluster Release and Version Information	21
▼ How to Display Configured Resource Types, Resource Groups, and Resources	22
▼ How to Check the Status of Cluster Components	23
▼ How to View the Cluster Configuration	25
▼ How to Check the Global Mount Points	27
2. Shutting Down and Booting a Cluster	29
2.1 Shutting Down and Booting a Cluster Overview	29

▼	How to Shut Down a Cluster	31
▼	How to Boot a Cluster	32
▼	How to Reboot a Cluster	33
2.2	Shutting Down and Booting a Single Cluster Node	35
▼	How to Shut Down a Cluster Node	36
▼	How to Boot a Cluster Node	37
▼	How to Reboot a Cluster Node	38
▼	How to Boot a Cluster Node in Non-Cluster Mode	39
2.3	Troubleshooting Cluster and Cluster Node Problems	41
▼	How to Repair a Full <code>/var</code> File System	41
3.	Administering Global Devices and Cluster File Systems	43
3.1	Administering Global Devices and the Global Namespace Overview	44
3.1.1	Global Device Permissions for Solstice DiskSuite	45
3.1.2	VERITAS Volume Manager Administration Considerations	45
3.2	Administering Cluster File Systems Overview	46
3.3	Administering Disk Device Groups	46
▼	How to Update the Global Device Namespace	48
▼	How to Add and Register a Disk Device Group (Solstice DiskSuite)	49
▼	How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)	50
▼	How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)	50
▼	How to Add a New Volume to an Existing Disk Device Group (VERITAS Volume Manager)	52
▼	How to Make an Existing Disk Group Into a Disk Device Group (VERITAS Volume Manager)	53
▼	How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)	53
▼	How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)	54

▼	How to Register Disk Group Configuration Changes (VERITAS Volume Manager)	56
▼	How to Remove and Unregister a Disk Device Group (Solstice DiskSuite)	57
▼	How to Remove a Volume From a Disk Device Group (VERITAS Volume Manager)	57
▼	How to Remove and Unregister a Disk Device Group (VERITAS Volume Manager)	58
▼	How to Add a Node to a Disk Device Group (VERITAS Volume Manager)	59
▼	How to Change Disk Device Properties	60
▼	How to List a Disk Device Group Configuration	61
▼	How to Switch the Primary for a Device Group	62
3.4	Administering Cluster File Systems	63
▼	How to Add an Additional Cluster File System	63
▼	How to Remove a Cluster File System	66
▼	How to Check Global Mounts in a Cluster	69
3.4.1	How to Remove a Node From a Disk Device Group (Solstice DiskSuite)	69
3.4.2	How to Remove a Node From a Disk Device Group (VERITAS Volume Manager)	70
4.	Administering Quorum	73
4.1	Administering Quorum Overview	73
▼	How to Add a Quorum Device	75
4.1.1	How to Remove a Quorum Device	76
4.1.2	How to Remove the Last Quorum Device From a Cluster	78
▼	How to Replace a Quorum Device	80
▼	How to Put a Quorum Device Into Maintenance State	80
▼	How to Put a Cluster Node Into Maintenance State	81
▼	How to Bring a Node Out of Maintenance State and Reset Quorum	83
▼	How to List the Quorum Configuration	85
5.	Administering Cluster Interconnects and Public Networks	87

5.1	Administering the Cluster Interconnects	88
▼	How to Check the Status of the Cluster Interconnect	89
▼	How to Add Cluster Transport Cables and Transport Adapters	89
▼	How to Remove Cluster Transport Cables and Transport Adapters	92
▼	How to Enable a Cluster Transport Cable	96
▼	How to Disable a Cluster Transport Cable	97
5.2	Administering the Public Network	99
▼	How to Create a NAFO Group	100
▼	How to Delete a NAFO Group	102
▼	How to Add an Adapter to a NAFO Group	104
▼	How to Remove an Adapter From a NAFO Group	105
▼	How to Switch a NAFO Group's Active Adapter	106
▼	How to Check the Status of NAFO Groups	107
▼	How to Change Public Network Management Tunable Parameters	108
6.	Administering the Cluster	111
6.1	Administering the Cluster Overview	111
▼	How to Change the Cluster Name	112
▼	How to Map Node ID to Node Name	113
▼	How to Work With New Cluster Node Authentication	113
▼	How to Reset the Time of Day in a Cluster	115
▼	How to Enter the OpenBoot PROM (OBP) on a Node	116
6.2	Adding a Cluster Node	117
6.2.1	How to Add a Cluster Node to the Authorized Node List	118
6.3	Removing a Cluster Node	119
6.3.1	How to Remove a Node From the Cluster Software Configuration	121
7.	Patching Sun Cluster Software and Firmware	125
7.1	Patching Sun Cluster Overview	125

7.1.1	Sun Cluster Patch Tips	126
7.2	Patching Sun Cluster	127
▼	How to Apply a Rebooting Patch (Node)	127
▼	How to Apply a Rebooting Patch (Cluster and Firmware)	130
▼	How to Apply a Non-Rebooting Sun Cluster Patch	131
▼	How to Remove a Sun Cluster Patch	132
8.	Backing Up and Restoring a Cluster	135
8.1	Backing Up Cluster Files	135
▼	How to Find File System Names to Back Up	136
▼	How to Determine the Number of Tapes Needed for a Full Backup	137
▼	How to Back Up the root (/) File System	137
▼	How to Perform Online Backups for Mirrors (Solstice DiskSuite)	139
▼	How to Perform Online Backups for Volumes (VERITAS Volume Manager)	142
8.2	Restoring Cluster Files Overview	146
8.3	Restoring Cluster Files	147
▼	How to Restore Individual Files Interactively (Solstice DiskSuite)	148
▼	How to Restore the root (/) File System (Solstice DiskSuite)	148
▼	How to Restore a root (/) File System That Was on a Metadevice (Solstice DiskSuite)	151
▼	How to Restore a Non-Encapsulated root (/) File System (VERITAS Volume Manager)	156
▼	How to Restore an Encapsulated root (/) File System (VERITAS Volume Manager)	159

Preface

The *Sun Cluster 3.0 System Administration Guide* provides procedures for administering a Sun[™] Cluster 3.0 configuration.

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

The instructions in this book assume knowledge of the Solaris[™] operating environment and expertise with the volume manager software used with Sun Cluster.

Using UNIX Commands

This document contains information on commands specific to administering a Sun Cluster configuration. This document may not contain complete information on basic UNIX[®] commands and procedures.

See one or more of the following for this information:

- AnswerBook2[™] online documentation for the Solaris software environment
- Other software documentation that you received with your system
- Solaris operating environment man pages

Typographic Conventions

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

Shell Prompts

Shell	Prompt
C shell	<i>machine_name</i> %
C shell superuser	<i>machine_name</i> #
Bourne shell and Korn shell	\$
Bourne shell and Korn shell superuser	#

Related Documentation

Application	Title	Part Number
Hardware	<i>Sun Cluster 3.0 Hardware Guide</i>	806-1420
Software Installation	<i>Sun Cluster 3.0 Installation Guide</i>	806-1419
Data Services	<i>Sun Cluster 3.0 Data Services Installation and Configuration Guide</i>	806-1421
API Development	<i>Sun Cluster 3.0 Data Services Developers' Guide</i>	806-1422
Concepts	<i>Sun Cluster 3.0 Concepts</i>	806-1424
Error Messages	<i>Sun Cluster 3.0 Error Messages Manual</i>	806-1426
Release Notes	<i>Sun Cluster 3.0 Release Notes</i>	805-1428

Ordering Sun Documentation

Fatbrain.com, an Internet professional bookstore, stocks select product documentation from Sun Microsystems, Inc.

For a list of documents and how to order them, visit the Sun Documentation Center on Fatbrain.com at:

<http://www1.fatbrain.com/documentation/sun>

Accessing Sun Documentation Online

The `docs.sun.com`SM web site enables you to access Sun technical documentation on the Web. You can browse the `docs.sun.com` archive or search for a specific book title or subject at:

Getting Help

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

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

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

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

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

Introduction to Administering Sun Cluster

This chapter provides information on preparing to administer the cluster and the procedures for using Sun Cluster administration tools.

This is a list of the procedures in this chapter.

- “How to Remotely Log In to Sun Cluster” on page 19
- “How to Access the `scsetup` Utility” on page 21
- “How to Display Sun Cluster Release and Version Information” on page 21
- “How to Display Configured Resource Types, Resource Groups, and Resources” on page 22
- “How to Check the Status of Cluster Components” on page 23
- “How to View the Cluster Configuration” on page 25
- “How to Check the Global Mount Points” on page 27

1.1 Administering Sun Cluster Overview

Sun Cluster’s highly-available environment ensures that critical applications are available to end users. The system administrator’s job is to make sure that Sun Cluster is stable and operational.

Before undertaking an administrative task, you should have familiarized yourself with the planning information in the *Sun Cluster 3.0 Installation Guide* and the

glossary in the *Sun Cluster 3.0 Concepts* document. Sun Cluster administration is organized into tasks among the following manuals:

- Standard tasks, used to administer and maintain the cluster on a regular, perhaps daily basis. These tasks are described in this guide.
- Data service tasks, such as installation, configuration, and changing properties. These tasks are described in the *Sun Cluster 3.0 Data Services Installation and Configuration Guide*.
- Service tasks, such as adding or repairing storage or network hardware. These tasks are described in the *Sun Cluster 3.0 Hardware Guide*.

For the most part, you can perform Sun Cluster administration tasks while the cluster is operational, with the impact on cluster operation limited to a single node. For those procedures that require that the entire cluster be shut down, schedule downtime for off hours, when there is minimal impact on the system. If you plan to take down the cluster or a cluster node, notify users ahead of time.

1.2 Graphical User Interface

The Sun Cluster-supplied module for Sun Management Center™ (formerly Sun Enterprise SyMON™) GUI Console enables you to graphically display cluster resources, resource types, and resource groups. It also enables you to monitor configuration changes and check the status of cluster components. However, the Sun Cluster-supplied module for Sun Management Center currently cannot perform all Sun Cluster administration tasks. You must use the command-line interface for other operations. See Section 1.3 “Command-Line Interface” on page 15 for more information.

For information on installing and starting the Sun Cluster module for Sun Management Center, and for viewing the cluster-specific online help supplied with the Sun Cluster module, see the *Sun Cluster 3.0 Installation Guide*.

Sun Cluster has created the SNMP MIB for the Sun Cluster module of Sun Management Center. This MIB can be used as the data definition by third-party SNMP based management stations.

The Sun Cluster MIB file is located in `$BASEDIR/SUNWsymon/modules/cfg` on any cluster node.

The Sun Cluster MIB file is an ASN.1 specification of the Sun Cluster data that is modeled. This is the same specification used by all Sun Management Center MIBs. To use the Sun Cluster MIB, refer to the instructions for using other Sun Management Center MIBs in Sun MC 2.1 Software User's Guide, Appendix B, “SNMP MIBs for Sun Management Center Modules.”

1.3 Command-Line Interface

You can perform most Sun Cluster administration tasks interactively through the `scsetup(1M)` utility. Whenever possible, administration procedures in this guide are described using `scsetup`.

You can administer the following items through the `scsetup` utility:

- Quorum
- Cluster transport
- Private hostnames
- Device groups
- New node authentication
- Cluster name

Listed here are the other commands you use to administer Sun Cluster. See the man pages for more detailed information.

TABLE 1-1 Sun Cluster Command-Line Interface Commands

Command	Description
<code>ccp(1M)</code>	Starts remote console access to the cluster.
<code>pmfadm(1M)</code>	Provides administrative access to the process monitor facility.
<code>pnmset(1M)</code>	Configures Public Network Management (PNM).
<code>pnmstat(1M)</code>	Reports the status of Network Adapter Failover (NAFO) groups monitored by PNM.
<code>sccheck(1M)</code>	Checks and validates the global mount entries in the <code>/etc/vfstab</code> file.
<code>scconf(1M)</code>	Updates a Sun Cluster configuration. The <code>-p</code> option lists cluster configuration information.
<code>scdidadm(1M)</code>	Provides administrative access to the device ID configuration.
<code>scgdevs(1M)</code>	Runs the global device namespace administration script.
<code>scinstall(1M)</code>	Installs and configures Sun Cluster software; can be run interactively or non-interactively. The <code>-p</code> option displays release and package version information for the Sun Cluster software.

TABLE 1-1 Sun Cluster Command-Line Interface Commands *(continued)*

Command	Description
<code>scrgadm(1M)</code>	Manages the registration of resource types, the creation of resource groups, and the activation of resources within a resource group. The <code>-p</code> option displays information on installed resources, resource groups, and resource types.
<code>scsetup(1M)</code>	Runs the interactive cluster configuration utility.
<code>scshutdown(1M)</code>	Shuts down the entire cluster.
<code>scstat(1M)</code>	Provides a snapshot of the cluster status.
<code>scswitch(1M)</code>	Performs changes affecting node mastery and states for resource groups and disk device groups.

In addition, you use commands to administer the volume manager portion of Sun Cluster. These commands depend on the specific volume manager used in your cluster, either Solstice DiskSuite™ or VERITAS Volume Manager.

1.4 Preparing to Administer the Cluster

This section describes what to do to prepare for administering your cluster.

1.4.1 Documenting a Sun Cluster Hardware Configuration

As your Sun Cluster configuration grows and changes, you need to document the hardware aspects that are unique to your site. Labeling cables and connections between the various cluster components also saves administration time when it becomes necessary to change or upgrade the cluster.

Keeping such a record of your cluster configuration will also come in handy when servicing the cluster. If you use a third-party service provider, such record keeping will make a service provider's job easier.

1.4.2 Using an Administrative Console

If you choose, you can use a dedicated SPARC workstation, known as the *administrative console*, to administer the active cluster. Typically, you install and run the Cluster Control Panel (CCP) and Sun Management Center server and console software on the administrative console. For more information on the CCP, see “How to Remotely Log In to Sun Cluster” on page 19. For instructions on installing the Sun Management Center software, see the *Sun Cluster 3.0 Installation Guide*.

The administrative console is not a cluster node. The administrative console is used for remote access to the cluster nodes, either over the public network or through a network-based terminal concentrator. If your cluster consists of a Sun Enterprise™ 10000 server, you must have the ability to log in from the administrative console to the System Service Processor (SSP) and connect by using the `netcon` command.

Sun Cluster does not require a dedicated administrative console, but using one provides these benefits:

- Enables centralized cluster management by grouping console and management tools on the same machine
- Provides potentially quicker problem resolution by Enterprise Services or your service provider

1.4.3 Backing Up the Cluster

While not specifically discussed in this guide, it is important to back up your cluster on a regular basis. Even though Sun Cluster provides an HA environment, with mirrored copies of data on the storage devices, do not consider this a replacement for regular backups. Sun Cluster can survive multiple failures, but not a catastrophic failure in which all data is rendered unavailable. Therefore, you must have a backup procedure in place to protect against catastrophic data loss. Backups are also important to undo user or program errors.

The following information should be included as part of your backup.

- All file system partitions
- All database data if you are running DBMS data services
- Disk partition information for all cluster disks
- The `md.tab` file if you are using Solstice DiskSuite™ as your volume manager

1.5 Beginning to Administer the Cluster

Table 1–2 provides a starting point for administering your cluster.

TABLE 1-2 Sun Cluster 3.0 Administration Tools

If You Want To...	Then...	For More Information Go To...
Remotely Log in to the Cluster	Use the <code>ccp</code> command to launch the Cluster Control Panel (CCP). Then select one of the following icons: <code>cconsole</code> , <code>crlogin</code> , or <code>ctelnet</code> .	"How to Remotely Log In to Sun Cluster" on page 19
Interactively Configure the Cluster	Launch the <code>scsetup</code> utility.	"How to Access the <code>scsetup</code> Utility" on page 21
Display Sun Cluster Release Number and Version Information	Use the <code>scinstall</code> command with either the <code>-p</code> or <code>-pv</code> options.	"How to Display Sun Cluster Release and Version Information" on page 21
Display Installed Resources, Resource Groups, and Resource Types	Use the <code>scgradm -p</code> command.	"How to Display Configured Resource Types, Resource Groups, and Resources" on page 22
Graphically Monitor Cluster Components	Use the Sun Cluster module for Sun Management Center.	Sun Cluster module for Sun Management Center online help
Check the Status of Cluster Components	Use the <code>scstat</code> command.	"How to Check the Status of Cluster Components" on page 23

TABLE 1–2 Sun Cluster 3.0 Administration Tools *(continued)*

If You Want To...	Then...	For More Information Go To...
View the Cluster Configuration	Use the <code>seconf -p</code> command.	“How to View the Cluster Configuration” on page 25
Check Global Mount Points	Use the <code>sccheck</code> command.	“How to Check the Global Mount Points” on page 27
Look at Sun Cluster System Messages	Examine the <code>/var/adm/messages</code> file.	Solaris system administration documentation
Monitor the Status of Solstice DiskSuite	Use the <code>metastat</code> or <code>metatool</code> commands.	Solstice DiskSuite documentation
Monitor the Status of VERITAS Volume Manager	Use the <code>vxstat</code> or <code>vxva</code> commands.	VERITAS Volume Manager documentation

▼ How to Remotely Log In to Sun Cluster

The Cluster Control Panel (CCP) provides a launch pad for `cconsole`, `crlogin`, and `ctelnet` tools. All three tools start a multiple window connection to a set of specified nodes. The multiple-window connection consists of a host window for each of the specified nodes and a common window. Input directed into the common window is sent to each of these host windows. See the `ccp(1M)` and `cconsole(1M)` man pages for more information.

1. Verify that the following prerequisites are met. To start the Cluster Control Panel (CCP), you must:

- Install the appropriate Sun Cluster software (`SUNWcccon` package) on the administrative console.
- Make sure the `PATH` variable on the administrative console includes the Sun Cluster tools directory, `/opt/SUNWcluster/bin`, and `/usr/cluster/bin`.

You can specify an alternate location for the tools directory by setting the `$CLUSTER_HOME` environment variable.

- Configure the `clusters` file, the `serialports` file, and the `nsswitch.conf` file if using a terminal concentrator. These can be either `/etc` files or NIS/NIS+ databases. See `clusters(4)` and `serialports(4)` for more information.

2. Determine if you have a Sun Enterprise E10000 server platform.

- If no, proceed to Step 3 on page 20.
- If yes, log into the System Service Processor (SSP) and connect by using the `netcon` command. Once connected, enter `Shift~@` to unlock the console and gain write access.

3. Start the CCP launch pad.

From the administrative console, enter the following command.

```
# ccp clustername
```

The CCP launch pad appears.

- 4. To start a remote session with the cluster, click the appropriate icon (`cconsole`, `crlogin`, or `ctelnet`) in the CCP launch pad.**

1.5.0.1 Example

The following example shows the Cluster Control Panel.

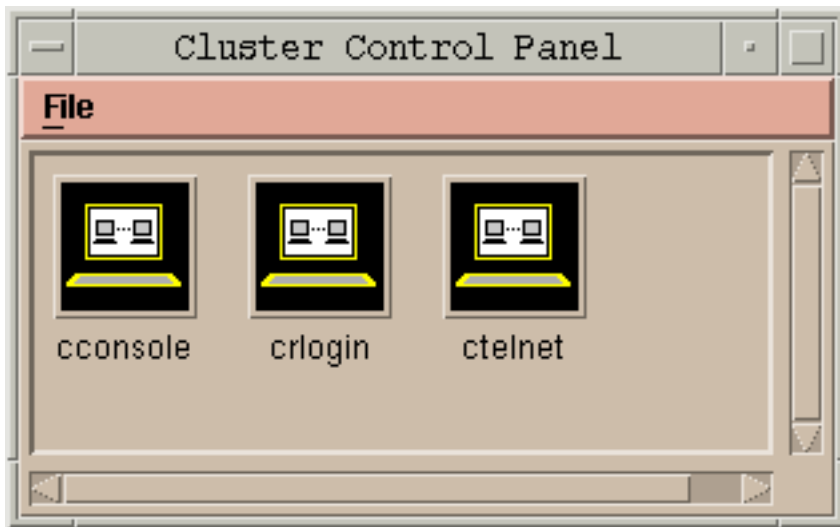


Figure 1-1 Cluster Control Panel

1.5.0.2 Where to Go From Here

You can also start `cconsole`, `crlogin`, or `ctelnet` sessions from the command line. See `cconsole(1M)` for more information.

▼ How to Access the `scsetup` Utility

The `scsetup(1M)` utility enables you to interactively configure quorum, cluster transport, private hostnames, device groups, and new node options for the cluster.

1. Become superuser on a node in the cluster.

2. Enter the `scsetup` utility.

```
# scsetup
```

The Main Menu appears.

3. Make your selection from the menu and follow the onscreen instructions.

See the `scsetup` online help for more information.

▼ How to Display Sun Cluster Release and Version Information

You do not need to be logged in as superuser to perform these procedures.

♦ **Display the Sun Cluster release number.**

```
% scinstall -p
```

♦ **Display the Sun Cluster release number and version strings for all Sun Cluster packages.**

```
% scinstall -pv
```

1.5.0.1 Examples—Displaying the Sun Cluster Release Number

The following example displays the cluster's release number.

```
% scinstall -p
3.0
```

1.5.0.2 Examples—Displaying Sun Cluster Release and Version Information

The following example displays the cluster's release information and version information for all packages.

```
% scinstall -pv
SunCluster 3.0
SUNWscr:      3.0.0,REV=1999.10.20.15.01
SUNWscdev:    3.0.0,REV=1999.10.20.15.01
SUNWscu:      3.0.0,REV=1999.10.20.15.01
SUNWscman:    3.0.0,REV=1999.10.20.15.01
SUNWscsal:    3.0.0,REV=1999.10.20.15.01
SUNWscsam:    3.0.0,REV=1999.10.20.15.01
SUNWrsmop:    3.0.0,REV=1999.10.20.15.01
SUNWsci:      3.0,REV=1999.09.08.17.43
SUNWscid:     3.0,REV=1999.09.08.17.43
SUNWscidx:    3.0,REV=1999.09.08.17.43
SUNWscvm:     3.0.0,REV=1999.10.20.15.01
```

▼ How to Display Configured Resource Types, Resource Groups, and Resources

You do not need to be logged in as superuser to perform this procedure.

- ◆ **Display the cluster's configured resource types, resource groups, and resources.**

```
% scrgadm -p
```

1.5.0.1 Example—Displaying Configured Resource Types, Resource Groups, and Resources

The following example shows the resource types (RT Name), resource groups (RG Name), and resources (RS Name) configured for the cluster `schost`.

```
% scgradm -p
RT Name: SUNW.SharedAddress
RT Description: HA Shared Address Resource Type
RT Name: SUNW.LogicalHostname
RT Description: Logical Hostname Resource Type
RG Name: schost-sa-1
RG Description:
  RS Name: schost-1
  RS Description:
  RS Type: SUNW.SharedAddress
  RS Resource Group: schost-sa-1
RG Name: schost-lh-1
RG Description:
  RS Name: schost-3
  RS Description:
  RS Type: SUNW.LogicalHostname
  RS Resource Group: schost-lh-1
```

▼ How to Check the Status of Cluster Components

You do not need to be logged in as superuser to perform this procedure.

- ◆ Check the status of cluster components.

```
% scstat -p
```

1.5.0.1 Example—Checking the Status of Cluster Components

The following example provides a sample of status information for cluster components returned by `scstat(1M)`.

```
% scstat -p
-- Cluster Nodes --

      Node name      Status
      -----
Cluster node:  phys-schost-1  Online
Cluster node:  phys-schost-2  Online
Cluster node:  phys-schost-3  Online
Cluster node:  phys-schost-4  Online
-----

-- Cluster Transport Paths --

      Endpoint      Endpoint      Status
      -----

```

(continued)

```

Transport path:  phys-schost-1:qfel phys-schost-4:qfel Path online
Transport path:  phys-schost-1:hmel phys-schost-4:hmel Path online
...

-----

-- Quorum Summary --

Quorum votes possible:      6
Quorum votes needed:        4
Quorum votes present:       6

-- Quorum Votes by Node --

      Node Name      Present Possible Status
      -----
Node votes:  phys-schost-1  1         1      Online
Node votes:  phys-schost-2  1         1      Online
...

-- Quorum Votes by Device --

      Device Name      Present Possible Status  Owner
      -----
Device votes:  /dev/did/rdisk/d2s2  1         1      Online  phys-schost-2
Device votes:  /dev/did/rdisk/d8s2  1         1      Online  phys-schost-4
...

-- Device Group Servers --

      Device Group      Primary      Secondary
      -----
Device group servers:  rmt/1          -          -
Device group servers:  rmt/2          -          -
Device group servers:  schost-1      phys-schost-2  phys-schost-1
Device group servers:  schost-3          -          -

-- Device Group Status --

      Device Group      Status
      -----
Device group status:    rmt/1          Offline
Device group status:    rmt/2          Offline
Device group status:    schost-1      Online
Device group status:    schost-3      Offline

-----

-- Resource Groups and Resources --

      Group Name      Resources
      -----
Resources: test-rg      test_1
Resources: real-property-rg  -

```

(continued)


```
Resources: failover-rg      -
Resources: descript-rg-1   -
...
-- Resource Groups --

      Group Name      Node Name      State
      -----
Group: test-rg        phys-schost-1  Offline
Group: test-rg        phys-schost-2  Offline
...
-- Resources --

      Resource Name      Node Name      State      Status Message
      -----
Resource: test_1         phys-schost-1  Offline    Offline
Resource: test_1         phys-schost-2  Offline    Offline
```

▼ How to View the Cluster Configuration

You do not need to be logged in as superuser to perform this procedure.

- ◆ View the cluster configuration.

```
% scconf -p
```

1.5.0.1 Example—Viewing the Cluster Configuration

The following example lists the cluster configuration.

```
% scconf -p
Cluster name:                cluster-1
Cluster ID:                  0x3908EE1C
Cluster install mode:        disabled
Cluster private net:         172.16.0.0
Cluster private netmask:     255.255.0.0
Cluster new node authentication: unix
Cluster new node list:       <NULL - Allow any node>
Cluster nodes:                phys-schost-1 phys-schost-2 phys-schost-3 phys-schost-4
4
Cluster node name:            phys-schost-1
Node ID:                      1
```

(continued)

```

Node enabled:                yes
Node private hostname:       clusternode1-priv
Node quorum vote count:      1
Node reservation key:        0x3908EE1C00000001
Node transport adapters:     hme1 qfe1 qfe2

Node transport adapter:      hme1
  Adapter enabled:           yes
  Adapter transport type:    dlpi
  Adapter property:          device_name=hme
  Adapter property:          device_instance=1
  Adapter property:          dlpi_heartbeat_timeout=10000
...
Cluster transport junctions:  hub0 hub1 hub2

Cluster transport junction:   hub0
  Junction enabled:           yes
  Junction type:              switch
  Junction port names:        1 2 3 4
...

```

```

Junction port:                1
  Port enabled:                yes

Junction port:                2
  Port enabled:                yes
...
Cluster transport cables
      Endpoint                Endpoint                State
      -----                -
Transport cable:  phys-schost-1:hme1@0  hub0@1          Enabled
Transport cable:  phys-schost-1:qfe1@0  hub1@1          Enabled
Transport cable:  phys-schost-1:qfe2@0  hub2@1          Enabled
Transport cable:  phys-schost-2:hme1@0  hub0@2          Enabled
...
Quorum devices:                d2 d8

Quorum device name:            d2
  Quorum device votes:         1
  Quorum device enabled:       yes
  Quorum device name:          /dev/did/rdisk/
d2s2
  Quorum device hosts (enabled): phys-schost-1
  phys-schost-2
  Quorum device hosts (disabled):
...
Device group name:             schost-3
  Device group type:           SDS
  Device group fallback enabled: no

```

(continued)

```

Device group node list:          phys-schost-3, phys-schost-
4
Diskset name:                   schost-3
...
```

▼ How to Check the Global Mount Points

The `sccheck(1M)` command checks the `/etc/vfstab` file for configuration errors with the cluster file system and its global mount points. The `sccheck` command only returns errors. If no errors are found, `sccheck` merely returns to the shell prompt.

Note - Run `sccheck` after making cluster configuration changes that have affected devices or volume management components.

1. Become superuser on a node in the cluster.
2. Verify the cluster configuration.

```
# sccheck
```

1.5.0.1 Example—Verifying the Cluster Configuration

The following example shows that the node `phys-schost-3` is missing the mount point `/global/schost-1`.

```
# sccheck
vfstab-check: WARNING - phys-schost-3 - Missing mount point /global/schost-1
```


Shutting Down and Booting a Cluster

This chapter provides the procedures for shutting down and booting a cluster and individual cluster nodes.

This is a list of the procedures in this chapter.

- “How to Shut Down a Cluster” on page 31
- “How to Boot a Cluster” on page 32
- “How to Reboot a Cluster” on page 33
- “How to Shut Down a Cluster Node” on page 36
- “How to Boot a Cluster Node” on page 37
- “How to Reboot a Cluster Node” on page 38
- “How to Boot a Cluster Node in Non-Cluster Mode” on page 39
- “How to Repair a Full `/var` File System” on page 41

For a high-level description of the related procedures in this chapter, see Table 2-1 and Table 2-2.

2.1 Shutting Down and Booting a Cluster Overview

The Sun Cluster `scshutdown(1M)` command stops cluster services in an orderly fashion and cleanly shuts down the cluster.

Note - Use `scshutdown` instead of the `shutdown` or `halt` commands to ensure proper shutdown of the entire cluster. The Solaris `shutdown` command is used to shut down individual nodes.

The `scshutdown` command stops a cluster by:

1. Taking all running resource groups offline
2. Unmounting all cluster file systems
3. Shutting down active device services
4. Running `init 0` and bringing all nodes to the `ok` PROM prompt

You might do this when moving a cluster from one location to another or if there was data corruption caused by an application error.

Note - If necessary, you can boot a node so that it does not participate in the cluster membership, that is, in non-cluster mode. This is useful when installing cluster software or for performing certain administrative procedures. See “How to Boot a Cluster Node in Non-Cluster Mode” on page 39 for more information.

TABLE 2-1 Task Map: Shutting Down and Booting a Cluster

Task	For Instructions, Go To...
Stop the cluster - Use <code>scshutdown</code>	“How to Shut Down a Cluster” on page 31
Start the cluster by booting all nodes. The nodes must have a working connection to the cluster interconnect to attain cluster membership.	“How to Boot a Cluster” on page 32
Shut down the cluster - Use <code>scshutdown</code> At the <code>ok</code> prompt, boot each node individually with the <code>boot</code> command. The nodes must have a working connection to the cluster interconnect to attain cluster membership.	“How to Reboot a Cluster” on page 33

▼ How to Shut Down a Cluster

1. **(Optional). For a cluster running Oracle Parallel Server (OPS), shut down all OPS database instances.**
Refer to the Oracle Parallel Server product documentation for shutdown procedures.
2. **Become superuser on a node in the cluster.**
3. **Shut down the cluster immediately by using the `scshutdownd(1M)` command.**
From a single node in the cluster, enter the following command.

```
# scshutdownd -g 0 -y
```

4. **Verify that all nodes have reached the `ok` PROM prompt.**
5. **If necessary, power off the nodes.**

2.1.0.1 Example—Shutting Down a Cluster

The following example shows the console output when stopping normal cluster operation and bringing down all nodes to the `ok` prompt. The `-g 0` option sets the shutdown grace period to zero, `-y` provides an automatic `yes` response to the confirmation question. Shutdown messages also appear on the consoles of the other nodes in the cluster.

```
# scshutdownd -g 0 -y
Sep  2 10:08:46 phys-schost-1 cl_runtime: WARNING: CMM monitoring disabled.
phys-schost-1#
INIT: New run level: 0
The system is coming down. Please wait.
System services are now being stopped.
/etc/rc0.d/K05initrgm: Calling scswitch -S (evacuate)
The system is down.
syncing file systems... done
Program terminated
ok
```

2.1.0.2 Where to Go From Here

See “How to Boot a Cluster” on page 32 to restart a cluster that has been shut down.

▼ How to Boot a Cluster

1. **To start a cluster whose nodes have been shut down and are at the `ok PROM` prompt, boot each node.**

The order in which the nodes are booted does not matter unless if you make configuration changes between shutdowns. In this case, you should start the nodes such that the one with the most current configuration boots first.

```
ok boot
```

Messages appear on the booted nodes' consoles as cluster components are activated.

Note - Cluster nodes must have a working connection to the cluster interconnect to attain cluster membership.

2. **Verify that the nodes booted without error and are online.**

The `scstat(1M)` command reports the nodes' status.

```
# scstat -n
```

2.1.0.1 Example—Booting a Cluster

The following example shows the console output when booting node `phys-schost-1` into the cluster. Similar messages appear on the consoles of the other nodes in the cluster.

```
ok boot
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: Node 1 with votecount = 1 added.
NOTICE: Node 2 with votecount = 1 added.
NOTICE: Node 3 with votecount = 1 added.
...
NOTICE: Node 1: attempting to join cluster
...
NOTICE: Node 2 (incarnation # 937690106) has become reachable.
NOTICE: Node 3 (incarnation # 937690290) has become reachable.
NOTICE: cluster has reached quorum.
NOTICE: node 1 is up; new incarnation number = 937846227.
NOTICE: node 2 is up; new incarnation number = 937690106.
NOTICE: node 3 is up; new incarnation number = 937690290.
NOTICE: Cluster members: 1 2 3
```

(continued)


```

...
NOTICE: Node 1: joined cluster
...
The system is coming up. Please wait.
checking ufs filesystems
...
reservation program successfully exiting
Print services started.
volume management starting.
The system is ready.
phys-schost-1 console login:

```

▼ How to Reboot a Cluster

Run the `scshutdownd(1M)` command to shut down the cluster, then boot the cluster with the `boot` command on each node.

1. **(Optional). For a cluster running Oracle Parallel Server (OPS), shut down all OPS database instances.**

Refer to the Oracle Parallel Server product documentation for shutdown procedures.

2. **Become superuser on a node in the cluster.**

3. **Shut down the cluster by using the `scshutdownd` command.**

From a single node in the cluster, enter the following command.

```
# scshutdownd -g 0 -y
```

This shuts down each node to the `ok` PROM prompt.

Note - Cluster nodes must have a working connection to the cluster interconnect to attain cluster membership.

4. **Boot each node.**

The order in which the nodes are booted does not matter unless if you make configuration changes between shutdowns. In this case, you should start the nodes such that the one with the most current configuration boots first.

```
ok boot
```

Messages appear on the booted nodes' consoles as cluster components are activated.

5. Verify that the nodes booted without error and are online.

The `scstat` command reports the nodes' status.

```
# scstat -n
```

2.1.0.1 Example—Rebooting a Cluster

The following example shows the console output when stopping normal cluster operation, bringing down all nodes to the `ok` prompt, then restarting the cluster. The `-g 0` option sets the grace period to zero, `-y` provides an automatic `yes` response to the confirmation question. Shutdown messages also appear on the consoles of other nodes in the cluster.

```
# scshutdown -g 0 -y
Sep  2 10:08:46 phys-schost-1 cl_runtime: WARNING: CMM monitoring disabled.
phys-schost-1#
INIT: New run level: 0
The system is coming down.  Please wait.
...
The system is down.
syncing file systems... done
Program terminated
ok boot
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
...
NOTICE: Node 1: attempting to join cluster
...
NOTICE: Node 2 (incarnation # 937690106) has become reachable.
NOTICE: Node 3 (incarnation # 937690290) has become reachable.
NOTICE: cluster has reached quorum.
...
NOTICE: Cluster members:   1  2  3
...
NOTICE: Node 1: joined cluster
...
The system is coming up.  Please wait.
checking ufs filesystems
...
reservation program successfully exiting
Print services started.
volume management starting.
The system is ready.
```

(continued)

```
phys-schost-1 console login:
```

2.2 Shutting Down and Booting a Single Cluster Node

Note - Use the `scswitch` command in conjunction with the Solaris `shutdown` command to shut down an individual node. Use the `scshutdown` command only when shutting down an entire cluster.

TABLE 2-2 Task Map: Shutting Down and Booting a Cluster Node

Task	For Instructions, Go To...
Stop a cluster node - Use <code>scswitch(1M)</code> and <code>shutdown(1M)</code>	“How to Shut Down a Cluster Node” on page 36
Start a node by booting it. The node must have a working connection to the cluster interconnect to attain cluster membership.	“How to Boot a Cluster Node” on page 37
Stop and restart (reboot) a cluster node - Use <code>scswitch</code> and <code>shutdown</code> The node must have a working connection to the cluster interconnect to attain cluster membership.	“How to Reboot a Cluster Node” on page 38
Boot a node so that it does not participate in cluster membership - Use <code>scswitch</code> and <code>shutdown</code> , then <code>boot -x</code>	“How to Boot a Cluster Node in Non-Cluster Mode” on page 39

▼ How to Shut Down a Cluster Node

1. **(Optional). For a cluster node running Oracle Parallel Server (OPS), shut down all OPS database instances.**

Refer to the Oracle Parallel Server product documentation for shutdown procedures.

2. **Become superuser on the cluster node to be shut down.**

3. **Shut down the cluster node by using the `scswitch` and `shutdown` commands.**

On the node to be shut down, enter the following command.

```
# scswitch -s -h node
# shutdown -g 0 -y
```

4. **Verify that the cluster node has reached the `ok` PROM prompt.**

5. **If necessary, power off the node.**

2.2.0.1 Example—Shutting Down a Cluster Node

The following example shows the console output when shutting down node `phys-schost-1`. The `-g 0` option sets the grace period to zero, `-y` provides an automatic `yes` response to the confirmation question. Shutdown messages for this node appear on the consoles of other nodes in the cluster.

```
# scswitch -s -h phys-schost-1
# shutdown -g 0 -y
Sep  2 10:08:46 phys-schost-1 cl_runtime: WARNING: CMM monitoring disabled.phys-schost-1#
INIT: New run level: 0
The system is coming down. Please wait.
Notice: rgmd is being stopped.
Notice: rpc.pmfd is being stopped.
Notice: rpc.fed is being stopped.
umount: /global/.devices/node@1 busy
umount: /global/phys-schost-1 busy
The system is down.
syncing file systems... done
```

(continued)

```
Program terminated
ok
```

2.2.0.2 Where to Go From Here

See “How to Boot a Cluster Node” on page 37 to restart a cluster node that has been shut down.

▼ How to Boot a Cluster Node

Note - Starting a cluster node can be affected by the quorum configuration. In a two-node cluster, you must have a quorum device configured such that the total quorum count for the cluster is three (one for each node and one for the quorum device). In this situation, if the first node is shut down, the second node continues to have quorum and runs as the sole cluster member. For the first node to come back in the cluster as a cluster node, the second node must be up and running and the required cluster quorum count (two) must be present.

1. **To start a cluster node that has been shut down, boot the node.**

```
ok boot
```

Messages appear on the booted node’s console, and on the member nodes’ consoles, as cluster components are activated.

Note - A cluster node must have a working connection to the cluster interconnect to attain cluster membership.

2. **Verify that the node has booted without error, and is online.**

The `scstat(1M)` command reports the status of a node.

```
# scstat -n
```

2.2.0.1 Example—Booting a Cluster Node

The following example shows the console output when booting node `phys-schost-1` into the cluster.

```
ok boot
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
...
NOTICE: Node 1: attempting to join cluster
...
NOTICE: Node 1: joined cluster
...
The system is coming up. Please wait.
checking ufs filesystems
...
reservation program successfully exiting
Print services started.
volume management starting.
The system is ready.
phys-schost-1 console login:
```

▼ How to Reboot a Cluster Node

1. **(Optional). For a cluster node running Oracle Parallel Server (OPS), shut down all OPS database instances.**

Refer to the Oracle Parallel Server product documentation for shutdown procedures.

2. **Become superuser on the cluster node to be shut down.**
3. **Shut down the cluster node by using the `scswitch` and `shutdown` commands.**

Enter these commands on the node to be shut down.

```
# scswitch -s -h node
# shutdown -g 0 -y -i 6
```

The `-i 6` option with the `shutdown` command causes the node to reboot after it shuts down to the `ok` PROM prompt.

Note - Cluster nodes must have a working connection to the cluster interconnect to attain cluster membership.

4. **Verify that the node has booted without error, and is online.**

The `scstat(1M)` command reports the status of a node.

```
# scstat -n
```

2.2.0.1 Example—Rebooting a Cluster Node

The following example shows the console output when shutting down and restarting node `phys-schost-1`. The `-g 0` option sets the grace period to zero, `-y` provides an automatic yes response to the confirmation question. Shutdown and startup messages for this node appear on the consoles of other nodes in the cluster.

```
# scswitch -S -h phys-schost-1
# shutdown -g 0 -y -i 6
Sep  2 10:08:46 phys-schost-1 cl_runtime: WARNING: CMM monitoring disabled.
phys-schost-1#
INIT: New run level: 6
The system is coming down. Please wait.
System services are now being stopped.
Notice: rgmd is being stopped.
Notice: rpc.pmfd is being stopped.
Notice: rpc.fed is being stopped.
umount: /global/.devices/node@1 busy
umount: /global/phys-schost-1 busy
The system is down.
syncing file systems... done
rebooting...
Resetting ...

'''
Sun Ultra 1 SBus (UltraSPARC 143MHz), No Keyboard
OpenBoot 3.11, 128 MB memory installed, Serial #7982421.
Ethernet address 8:0:20:79:cd:55, Host ID: 8079cd55.
...
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
...
NOTICE: Node 1: attempting to join cluster
...
NOTICE: Node 1: joined cluster
...
The system is coming up. Please wait.
The system is ready.
phys-schost-1 console login:
```

▼ How to Boot a Cluster Node in Non-Cluster Mode

You can boot a node so that it does not participate in the cluster membership, that is, in non-cluster mode. This is useful when installing the cluster software or for performing certain administrative procedures, such as patching a node.

1. Become superuser on the cluster node to be started in non-cluster mode.
2. Shut down the node by using the `scswitch` and `shutdown` commands.

```
# scswitch -s -h node
# shutdown -g 0 -y
```

3. Verify that the node is at the `ok` PROM prompt.
4. Boot the node in non-cluster mode by using the `boot(1M)` command with the `-x` option.

```
ok boot -x
```

Messages appear on the node's console stating that the node is not part of the cluster.

2.2.0.1 Example—Booting a Cluster Node in Non-Cluster Mode

The following example shows the console output when shutting down node `phys-schost-1` then restarting it in non-cluster mode. The `-g -0` option sets the grace period to zero, `-y` provides an automatic `yes` response to the confirmation question. Shutdown messages for this node appear on the consoles of other nodes in the cluster.

```
# scswitch -s -h phys-schost-1
# shutdown -g 0 -y
Sep  2 10:08:46 phys-schost-1 cl_runtime: WARNING: CMM monitoring disabled.
phys-schost-1#
...
rg_name = schost-sa-1 ...
offline node = phys-schost-2 ...
num of node = 0 ...
phys-schost-1#
INIT: New run level: 0
The system is coming down. Please wait.
System services are now being stopped.
Print services stopped.
syslogd: going down on signal 15
...
The system is down.
syncing file systems... done
WARNING: node 1 is being shut down.
```

(continued)


```
Program terminated
ok boot -x
...
Not booting as part of cluster
...
The system is ready.
phys-schost-1 console login:
```

2.3 Troubleshooting Cluster and Cluster Node Problems

This section describes solutions to problems that can arise during the day-to-day operation of a cluster and cluster nodes.

▼ How to Repair a Full `/var` File System

Both Solaris and Sun Cluster software write error messages to the `/var/adm/messages` file, which over time can fill the `/var` file system. If a cluster node's `/var` file system fills up, Sun Cluster might not be able to restart on that node. Additionally, you might not be able to log in to the node.

If a node reports a full `/var` file system and continues to run Sun Cluster services, use this procedure to clear the full file system.

1. **Become superuser on the cluster node with the full `/var` file system.**
2. **Clear the full file system.**

For example, delete nonessential files contained in the file system.

Administering Global Devices and Cluster File Systems

This chapter provides the procedures for administering global devices and cluster file systems.

This is a list of the procedures in this chapter.

- “How to Update the Global Device Namespace” on page 48
- “How to Add and Register a Disk Device Group (Solstice DiskSuite)” on page 49
- “How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)” on page 50
- “How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)” on page 50
- “How to Add a New Volume to an Existing Disk Device Group (VERITAS Volume Manager)” on page 52
- “How to Make an Existing Disk Group Into a Disk Device Group (VERITAS Volume Manager)” on page 53
- “How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)” on page 53
- “How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)” on page 54
- “How to Register Disk Group Configuration Changes (VERITAS Volume Manager)” on page 56
- “How to Remove and Unregister a Disk Device Group (Solstice DiskSuite)” on page 57
- “How to Remove a Volume From a Disk Device Group (VERITAS Volume Manager)” on page 57

- “How to Remove and Unregister a Disk Device Group (VERITAS Volume Manager)” on page 58
- “How to Add a Node to a Disk Device Group (VERITAS Volume Manager)” on page 59
- “How to Change Disk Device Properties ” on page 60
- “How to List a Disk Device Group Configuration” on page 61
- “How to Switch the Primary for a Device Group” on page 62
- “How to Add an Additional Cluster File System” on page 63
- “How to Remove a Cluster File System” on page 66
- “How to Check Global Mounts in a Cluster” on page 69
- Section 3.4.1 “How to Remove a Node From a Disk Device Group (Solstice DiskSuite)” on page 69
- Section 3.4.2 “How to Remove a Node From a Disk Device Group (VERITAS Volume Manager)” on page 70

For a high-level description of the related procedures in this chapter, see Table 3-1.

See the *Sun Cluster 3.0 Concepts* document for conceptual information related to global devices, the global namespace, disk device groups, and the cluster file system.

3.1 Administering Global Devices and the Global Namespace Overview

Administration of Sun Cluster disk device groups depends on the volume manager installed on the cluster. Solstice DiskSuite is “cluster-aware,” so you add, register, and remove disk device groups by using the Solstice DiskSuite `metaset(1M)` command. With VERITAS Volume Manager (VxVM), you create disk groups by using VxVM commands. Then you register the disk groups as Sun Cluster disk device groups through the `scsetup(1M)` utility. When removing VxVM disk device groups, you use both the `scsetup` utility and VxVM commands.

When administering disk device groups, or volume manager disk groups, you need to be on the cluster node that is the primary node for the group.

Normally, you do not need to administer the global device namespace because the global namespace is automatically set up during installation and automatically updated during Solaris operating environment reconfiguration reboots. However, if the global namespace needs to be regenerated or updated, you can run the `scgdevs(1M)` command from any cluster node. This causes the global namespace to be updated on all other cluster node members, as well as on nodes that might join the cluster in the future.

3.1.1 Global Device Permissions for Solstice DiskSuite

Changes made to global device permissions are not automatically propagated to all the nodes in the cluster for Solstice DiskSuite and disk devices. If you want to change permissions on global devices, you must manually change the permissions on all the nodes in the cluster. For example, if you want to change permissions on global device `/dev/global/dsk/d3s0` to 644, you must execute

```
# chmod 644 /dev/global/dsk/d3s0
```

on all nodes in the cluster.

VxVM does not support the `chmod` command. To change global device permissions in VxVM, consult the VxVM administrator's guide.

3.1.2 VERITAS Volume Manager Administration Considerations

For Sun Cluster to correctly maintain the VxVM namespace, when any configuration information for a disk group or volume is changed, you must register the Sun Cluster disk device group configuration changes. This ensures that the namespace on all cluster nodes is updated. Examples of configuration changes that impact the namespace include adding, removing, or renaming a volume; and changing the volume permissions, owner, or group ID.

Note - Never import or deport VxVM disk groups using VxVM commands once the disk group has been registered with the cluster as a Sun Cluster disk device group. The Sun Cluster software will handle all cases where disk groups need to be imported or deported.

Each VxVM disk group must have a cluster-wide unique minor number. By default, when a disk group is created, VxVM chooses a random number that is a multiple of 1000 as that disk group's base minor number. For most configurations with only a small number of disk groups, this is sufficient to guarantee uniqueness. However, it is possible that the minor number for a newly-created disk group will conflict with the minor number of a pre-existing disk group imported on a different cluster node. In this case, attempting to register the Sun Cluster disk device group will fail. To fix this problem, the new disk group should be given a new minor number that is a unique value and then registered as a Sun Cluster disk device group.

If you are setting up a mirrored volume, Dirty Region Logging (DRL) can be used to decrease volume recovery time in the event of a system crash. Use of DRL is strongly recommended.

3.2 Administering Cluster File Systems Overview

You use standard Solaris file system commands, such as `mount`, `newfs`, and so on, to administer the cluster file system. You mount cluster file systems by specifying the `-g` option to the `mount` command. Cluster file systems can also be automatically mounted at boot.

Note - No special Sun Cluster commands are necessary for cluster file system administration. You administer a cluster file system as you would any other Solaris file system.

3.3 Administering Disk Device Groups

Note - The `scsetup(1M)` utility is an interactive interface to the `scconf(1M)` command. When `scsetup` runs, it generates `scconf` commands. These generated commands are shown in the examples at the end of some procedures.

TABLE 3-1 Task Map: Administering Disk Device Groups

Task	For Instructions, Go To...
Update the global device namespace (without a reconfiguration reboot) - Use <code>scgdevs</code>	"How to Update the Global Device Namespace" on page 48
Add Solstice DiskSuite disksets and register them as disk device groups - Use <code>metaset</code>	"How to Add and Register a Disk Device Group (Solstice DiskSuite)" on page 49

TABLE 3-1 Task Map: Administering Disk Device Groups *(continued)*

Task	For Instructions, Go To...
<p>Add VERITAS Volume Manager disk groups as disk device groups</p> <ul style="list-style-type: none"> - Use VxVM commands and <code>scsetup</code> 	<p>“How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)” on page 50</p> <p>“How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)” on page 50</p> <p>“How to Add a New Volume to an Existing Disk Device Group (VERITAS Volume Manager)” on page 52</p> <p>“How to Make an Existing Disk Group Into a Disk Device Group (VERITAS Volume Manager)” on page 53</p> <p>“How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)” on page 53</p> <p>“How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)” on page 54</p> <p>“How to Register Disk Group Configuration Changes (VERITAS Volume Manager)” on page 56</p>
<p>Remove Solstice DiskSuite disk device groups from the configuration</p> <ul style="list-style-type: none"> - Use <code>metaset</code> and <code>metaclear</code> 	<p>“How to Remove and Unregister a Disk Device Group (Solstice DiskSuite)” on page 57</p>
<p>Remove VERITAS Volume Manager disk device groups from the configuration</p> <ul style="list-style-type: none"> - Use <code>scsetup</code> (to generate <code>scconf</code>) 	<p>“How to Remove a Volume From a Disk Device Group (VERITAS Volume Manager)” on page 57</p> <p>“How to Remove and Unregister a Disk Device Group (VERITAS Volume Manager)” on page 58</p>

TABLE 3-1 Task Map: Administering Disk Device Groups *(continued)*

Task	For Instructions, Go To...
Add a node to a VERITAS Volume Manager disk device group - Use <code>scsetup</code> (to generate <code>scconf</code>)	"How to Add a Node to a Disk Device Group (VERITAS Volume Manager)" on page 59
Change disk device group properties - Use <code>scsetup</code> (to generate <code>scconf</code>)	"How to Change Disk Device Properties " on page 60
Display disk device groups and properties - Use <code>scconf</code>	"How to List a Disk Device Group Configuration" on page 61
Switch the primary for a disk device group - Use <code>scswitch</code>	"How to Switch the Primary for a Device Group" on page 62

▼ How to Update the Global Device Namespace

Manually update the global device namespace when adding a new global device by running `scgdevs(1M)`.

Note - The `scgdevs` command does not have any effect if the node running the command is not currently a cluster member or if the `/global/.devices/node@nodeID` file system is not mounted.

1. Become superuser on a node of the cluster.
2. Use `scgdevs` to reconfigure the namespace.

```
# scgdevs
```

3.3.0.1 Example—Updating the Global Device Namespace

The following example shows output generated by a successful run of `scgdevs`.


```
# scgdevs
Configuring the /dev/global directory (global devices)...
obtaining access to all attached disks
reservation program successfully exiting
```

▼ How to Add and Register a Disk Device Group (Solstice DiskSuite)

Disk device groups map directly to Solstice DiskSuite disksets. When you create a diskset using `metaset(1M)`, you also create the disk device group with the same name and register it as a Sun Cluster disk device group.

1. **Become superuser on the node connected to the disks where you want to create the diskset.**
2. **Use `metaset` to add the Solstice DiskSuite diskset and register it as a disk device group with Sun Cluster.**

```
# metaset -s diskset -a -h node-list
```

- `-s diskset` Specifies the diskset to be created.
- `-a -h node-list` Adds the list of nodes that can master the diskset.

3. **Verify that the disk device group has been added.**

The disk device group name will match the diskset name specified with `metaset`.

```
# scconf -p | egrep disk-device-group
```

3.3.0.1 Example—Adding a Solstice DiskSuite Disk Device Group

The following example shows the creation of the diskset and disk device group and verifies that the disk device group has been created.

```
# metaset -s dg-schost-1
# scconf -p | egrep dg-schost-1
Device group name: dg-schost-1
```

▼ How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)

Note - This procedure is only for initializing disks. If you are encapsulating disks, use the procedure “How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)” on page 50.

After adding the VxVM disk group, you need to register the disk device group.

1. **Become superuser on a node of the cluster that is *physically connected* to the disks that make up the disk group being added.**

2. **Create the VxVM disk group and volume.**

Use your preferred method to create the disk group and volume.

Note - If you are setting up a mirrored volume, we strongly recommend that Dirty Region Logging (DRL) be used to decrease volume recovery time in the event of a system crash.

See the VERITAS Volume Manager documentation for the procedures to complete this step.

3.3.0.1 Where to Go From Here

The VxVM disk group must be registered as a Sun Cluster disk device group. See “How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)” on page 54.

▼ How to Create a New Disk Group When Encapsulating Disks (VERITAS Volume Manager)

Note - This procedure is only for encapsulating disks. If you are initializing disks, use the procedure “How to Create a New Disk Group When Initializing Disks (VERITAS Volume Manager)” on page 50.

You can make non-root disks into Sun Cluster disk device groups by first encapsulating them as VxVM disk groups, then registering them as Sun Cluster disk device groups.

Disk encapsulation is only supported during initial creation of VxVM disk groups. Once a disk group is created and registered, only disks which can be initialized should be added to the disk group.

1. Become superuser on a node of the cluster.

2. If the disk being encapsulated has file system entries in the `/etc/vfstab` file, make sure that the `mount at boot` option is set to `no`.

This can be set back to `yes` once the disk has been encapsulated and registered as a Sun Cluster disk device group.

3. Encapsulate the disks.

Use `vxdiskadm` menus or the graphical user interface to encapsulate the disks. VxVM requires two free partitions as well as unassigned cylinders at the beginning or the end of the disk. Slice 2 must also be set to the entire disk. See the `vxdiskadm(1M)` man page for more information.

4. Shut down and restart the node.

The `scswitch(1M)` command will switch over all resource groups and device groups from the node to the next preferred node. Then `shutdown(1M)` is used to shut down and restart the node.

```
# scswitch -s -h node
# shutdown -g 0 -i 6 -y
```

5. If necessary, switch all resource groups, and device groups back.

If the resource groups and device groups were initially configured to fail back to the primary node, this step is not necessary.

```
# scswitch -z -h node -D devgrp1 [ ,devgrp2,... ]
# scswitch -z -h node -g resgrp1 [ ,resgrp2,... ]
```

3.3.0.1 Where to Go From Here

The VxVM disk group must be registered as a Sun Cluster disk device group. See “How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)” on page 54.

▼ How to Add a New Volume to an Existing Disk Device Group (VERITAS Volume Manager)

Note - After adding the volume, you need to register the configuration change by using the procedure “How to Register Disk Group Configuration Changes (VERITAS Volume Manager)” on page 56.

When you add a new volume to an existing VxVM disk device group, you need to perform the procedure from the primary node for the disk device group, and the disk device group must be online.

1. **Become superuser on a node of the cluster.**
2. **Determine the primary node for the disk device group.**

```
# scstat -D
```

3. **Determine if the disk device group is offline.**

- If no, proceed to Step 4 on page 52.
- If yes, bring the disk group online.

```
# scswitch -z -D disk-device-group -h node
```

-z -D <i>disk-device-group</i>	Switches the specified device group.
-h <i>node</i>	Specifies the name of the node to switch the disk device group to.

4. **From the primary node (the node currently mastering the disk device group), create the VxVM volume in the disk group.**

Refer to your VERITAS Volume Manager documentation for the procedure used to create the VxVM volume.

3.3.0.1 Where to Go From Here

The change to the VxVM disk group must be registered to update the global namespace. See “How to Register Disk Group Configuration Changes (VERITAS Volume Manager)” on page 56.

▼ How to Make an Existing Disk Group Into a Disk Device Group (VERITAS Volume Manager)

You can make an existing VxVM disk group into a Sun Cluster disk device group by first importing the disk group onto the current node, then registering the disk group as a Sun Cluster disk device group.

1. **Become superuser on a node of the cluster.**
2. **Import the VxVM disk group onto the current node.**

```
# vxdg import diskgroup
```

3.3.0.1 Where to Go From Here

The VxVM disk group must be registered as a Sun Cluster disk device group. See “How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)” on page 54.

▼ How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)

If registering a VxVM disk device group fails due to a minor number conflict with another disk group, the new disk group must be assigned a new, unused minor number. After assigning the new minor number, you then rerun the procedure to register the disk group as a Sun Cluster disk device group.

1. **Become superuser on a node of the cluster.**
2. **Determine the minor numbers in use.**

```
# ls -l /dev/vx/dsk/*
```

3. **Choose any other multiple of 1000 that is not in use as the base minor number for the new disk group.**
4. **Assign the new minor number to the disk group.**

```
# vxdg reminor diskgroup base_minor_number
```

3.3.0.1 Example—How to Assign a New Minor Number to a Disk Device Group

This example shows the minor numbers 16000-16002 and 4000-4001 being used. The `vxvg remminor` command is used to remminor the new disk device group to use the base minor number 5000.

```
# ls -l /dev/vx/dsk/*
/dev/vx/dsk/dg1
brw----- 1 root    root    56,16000 Oct  7 11:32 dg1lv1
brw----- 1 root    root    56,16001 Oct  7 11:32 dg1lv2
brw----- 1 root    root    56,16002 Oct  7 11:32 dg1lv3

/dev/vx/dsk/dg2
brw----- 1 root    root    56,4000 Oct  7 11:32 dg2v1
brw----- 1 root    root    56,4001 Oct  7 11:32 dg2v2
# vxvg remminor dg3 5000
```

3.3.0.2 Where to Go From Here

The VxVM disk group must be registered as a Sun Cluster disk device group. See “How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)” on page 54.

▼ How to Register a Disk Group as a Disk Device Group (VERITAS Volume Manager)

This procedure uses the `scsetup(1M)` utility to register the associated VxVM disk group as a Sun Cluster disk device group.

Note - Once a disk device group has been registered, if you make a change to the VxVM disk group or volume, use the procedure “How to Register Disk Group Configuration Changes (VERITAS Volume Manager)” on page 56 to register the disk device group configuration changes. This will ensure that the global namespace is in the correct state.

The prerequisites to register a VxVM disk device group are:

- Superuser privilege on a node in the cluster.
- The name of the VxVM disk group to be registered as a disk device group.
- A preferred order of nodes to master the disk device group.

When you define the preference order, you also specify whether you want the disk device group to be switched back to the most preferred node in the event that the most preferred node goes down and later returns to the cluster.

See `scconf(1M)` for more information on node preference and failback options.

1. Become superuser on a node of the cluster.

2. Enter the `scsetup` utility.

```
# scsetup
```

The Main Menu appears.

3. To work with VxVM disk device groups, enter 3 (Device groups and volumes).

The Device Groups Menu appears.

4. To register a VxVM disk device group, enter 1 (Register a VxVM disk group as a device group).

Follow the instructions and enter the VxVM disk group to be registered as a Sun Cluster disk device group. If you encounter the following error while attempting to register the disk device group, use the procedure “How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)” on page 53. This procedure will enable you to assign a new minor number that does not conflict with a minor number used by existing disk device groups.

```
scconf: Failed to add device group - in use
```

5. Verify that the disk device group has been registered and brought online.

Look for the disk device information for the new disk displayed by the following command.

```
# scstat -D
```

3.3.0.1 Example—Registering a VERITAS Volume Manager Disk Device Group

The following example shows the `scconf` command generated by `scsetup` when it registers a VxVM disk device group (`dg1`), and the verification step. This example assumes that the VxVM disk group and volume were created previously.

```
# scconf -a -D type=vxvm,name=dg1,nodelist=phys-schost-1:phys-schost-2
# scstat -D
-- Device Group Servers --

Device group servers:  Device Group      Primary      Secondary
                       -----
                       dg1             phys-schost-1 phys-schost-2
```

(continued)

```
-- Device Group Status --
```

	Device Group	Status
	-----	-----
Device group status:	dgl	Online

3.3.0.2 Where to Go From Here

To create a cluster file system on the VxVM disk device group, see “How to Add an Additional Cluster File System” on page 63. If there are problems with the minor number, see “How to Assign a New Minor Number to a Disk Device Group (VERITAS Volume Manager)” on page 53.

▼ How to Register Disk Group Configuration Changes (VERITAS Volume Manager)

When you change any configuration information for a VxVM disk group or volume, you need to register the configuration changes for the Sun Cluster disk device group. This ensures that the global namespace is in the correct state.

1. **Become superuser on a node in the cluster.**
2. **Enter the `scsetup(1M)` utility.**

```
# scsetup
```

The Main Menu appears.

3. **To work with VxVM disk device groups, enter 3 (Device groups and volumes).**
The Device Groups Menu appears.
4. **To register configuration changes, enter 2 (Synchronize volume information for a VxVM device group).**
Follow the instructions and enter the VxVM disk group that has changed configuration.

3.3.0.1

Example—Registering VERITAS Volume Manager Disk Group Configuration Changes

The following example shows the `scconf` command generated by `scsetup` when it registers a changed VxVM disk device group (`dgl`). This example assumes that the VxVM disk group and volume were created previously.

```
# scconf -c -D name=dgl, sync
```

▼ How to Remove and Unregister a Disk Device Group (Solstice DiskSuite)

Disk device groups map directly to Solstice DiskSuite disksets. Thus, to remove a Solstice DiskSuite disk device group, you use the `metaclear(1M)` and `metaset(1M)` commands. These commands remove the disk device group with the same name and unregister the disk group as a Sun Cluster disk device group.

Refer to the Solstice DiskSuite documentation for the steps to remove a diskset.

▼ How to Remove a Volume From a Disk Device Group (VERITAS Volume Manager)

Note - After removing the volume from the disk device group, you must register the configuration changes to the disk device group using the procedure “How to Register Disk Group Configuration Changes (VERITAS Volume Manager)” on page 56.

1. Become superuser on a node of the cluster.
2. Determine the primary node for the disk device group.

```
# scstat -D
```

3. Determine if the disk device group is offline.

- If no, proceed to Step 4 on page 58.
- If yes, bring the disk group online.

```
# scswitch -z -D disk-device-group -h node
```

-z

Performs the switch.

-D *disk-device-group*

Specifies the device group to switch.

-h **node** Specifies the name of the node to become the new primary.

4. From the primary node (the node currently mastering the disk device group), remove the VxVM volume in the disk group.

```
# vxedit -g diskgroup -rf rm volume
```

-g **diskgroup** Specifies the VxVM disk group containing the volume.

-rf rm **volume** Removes the specified volume.

3.3.0.1 Where to Go From Here

After removing a volume, you must register the configuration changes to the disk device group. To register the configuration changes, see “How to Register Disk Group Configuration Changes (VERITAS Volume Manager)” on page 56.

▼ How to Remove and Unregister a Disk Device Group (VERITAS Volume Manager)

Removing a Sun Cluster disk device group will cause the corresponding VxVM disk group to be deported, not destroyed. However, even though the VxVM disk group still exists, it cannot be used in the cluster unless re-registered.

This procedure uses the `scsetup(1M)` utility to remove a VxVM disk group and unregister it as a Sun Cluster disk device group.

1. Become superuser on a node of the cluster.

2. Take the disk device group offline.

```
# scswitch -F -D disk-device-group
```

-F Places the disk device group offline.

-D **disk-device-group** Specifies the device group to take offline.

3. Enter the `scsetup` utility.

The Main Menu appears.

```
# scsetup
```

4. To work with VxVM device groups, enter 3 (Device groups and volumes).

The Device Groups Menu appears.

5. **To unregister a VxVM disk group, enter 3 (Unregister a VxVM device group).**
Follow the instructions and enter the VxVM disk group to be unregistered.

3.3.0.1

Example—Removing and Unregistering a VERITAS Volume Manager Disk Device Group

The following example shows the VxVM disk device group dg1 taken offline, and the `scconf(1M)` command generated by `scsetup` when it removes and unregisters the disk device group.

```
# scswitch -F -D dg1
# scconf -r -D name=dg1
```

▼ How to Add a Node to a Disk Device Group (VERITAS Volume Manager)

This procedure adds a node to a disk device group using the `scsetup(1M)` utility.

The prerequisites to add a node to a VxVM disk device group are:

- Superuser privilege on a node in the cluster
- The name of the VxVM device group to which the node will be added
- The name or node ID of the nodes to add

1. **Become superuser on a node of the cluster.**

2. **Enter the `scsetup(1M)` utility**

The Main Menu appears.

```
# scsetup
```

3. **To work with VxVM disk device groups, enter 3 (Device groups and volumes).**

The Device Groups Menu appears.

4. **To add a node to a VxVM disk device group, enter 4 (Add a node to a VxVM device group).**

Follow the instructions and enter the device group and node names.

5. **Verify that the node has been added.**

Look for the device group information for the new disk displayed by the following command.

```
# scconf -p
```

3.3.0.1 Example—Adding a Node to a VERITAS Volume Manager Disk Device Group

The following example shows the `scconf` command generated by `scsetup` when it adds a node (`phys-schost-3`) to a VxVM disk device group (`dg1`), and the verification step.

```
# scconf -a -D type=vxvm,name=dg1,nodelist=phys-schost-3
# scconf -p
...
Device group name:                dg1
Device type:                     VXVM
Failback enabled:                 yes
Node preference list:             phys-schost-1, phys-schost-3
```

▼ How to Change Disk Device Properties

The method for establishing the primary ownership of a disk device group is based on the setting of an ownership preference attribute called `preferenced`. If the attribute is not set, the primary owner of an otherwise unowned disk device group is the first node that attempts to access a disk in that group. However, if this attribute is set, you must specify the preferred order in which nodes attempt to establish ownership.

If you disable the `preferenced` attribute, then the `failback` attribute is also automatically disabled. However, if you attempt to enable or re-enable the `preferenced` attribute, you have the choice of enabling or disabling the `failback` attribute.

If the `preferenced` attribute is either enabled or re-enabled, you are required to re-establish the order of nodes in the primary ownership preference list.

This procedure uses `scsetup(1M)` to set or unset the `preferenced` attribute and the `failback` attribute for Solstice DiskSuite or VxVM disk device groups.

To run this procedure, you need the name of the disk device group for which you are changing attribute values.

1. **Become superuser on a node of the cluster.**
2. **Enter the `scsetup(1M)` utility**

The Main Menu appears.

```
# scsetup
```

3. To work with disk device groups, enter 3 (Device groups and volumes).

The Device Groups Menu appears.

4. To change a device group property, enter 6 (Change key properties of a VxVM or Solstice DiskSuite device group).

Follow the instructions to set the preferred and failback options for a device group.

5. Verify that the disk device group attributes have been changed.

Look for the device group information displayed by the following command.

```
# scconf -p
```

3.3.0.1 Example—Changing Disk Device Group Properties

The following example shows the `scconf` command generated by `scsetup` when it sets the attribute values for a disk device group (`dg-schost-1`).

```
# scconf -c -D name=dg-schost-1,nodelist=phys-schost-1:phys-schost-2,\
preferred=true,failback=enabled
# scconf -p
Device group name:          dg-schost-1
Device type:               SDS
Failback enabled:          yes
Node preference list:       phys-schost-1, phys-schost-2
Diskset name:              dg-schost-1
```

▼ How to List a Disk Device Group Configuration

You do not need to be superuser to list the configuration.

- ◆ Use `scconf(1M)` to list the disk device group configuration.

```
% scconf -p
```

3.3.0.1 Example—Listing the Disk Device Group Configuration

When using the `scconf` command, look for the information listed under device groups.

```
# scconf -p
...
Device group name: dg-schost-1
Device type: SDS
Failback enabled: yes
Node preference list: phys-schost-2, phys-schost-3
Diskset name: dg-schost-1
```

▼ How to Switch the Primary for a Device Group

This procedure can also be used to start (bring online) an inactive device group.

1. **Become superuser on a node of the cluster.**
2. **Use `scswitch(1M)` to switch the disk device group primary.**

```
# scswitch -z -D disk-device-group -h node
```

<code>-z</code>	Performs the switch.
<code>-D disk-device-group</code>	Specifies the device group to switch.
<code>-h node</code>	Specifies the name of the node to become the new primary.

3. **Verify that the disk device group has been switched to the new primary.**

Look for the disk device information for the device group displayed by the following command.

```
# scstat -D
```

3.3.0.1 Example—Switching the Primary for a Disk Device Group

The following example shows how to switch the primary for a disk device group and verify the change.

```
# scswitch -z -D dg-schost-1 -h phys-schost-1
# scstat -D
...
Device Group Name:          dg-schost-1
Status:                    Online
Primary:                   phys-schost-1
```

3.4 Administering Cluster File Systems

TABLE 3-2 Task Map: Administering Cluster File Systems

Task	For Instructions, Go To...
Add cluster file systems after the initial Sun Cluster installation - Use <code>newfs</code> and <code>makedir</code>	"How to Add an Additional Cluster File System" on page 63
Remove a cluster file system - Use <code>fuser</code> and <code>umount</code>	"How to Remove a Cluster File System" on page 66
Check global mount points in a cluster for consistency across nodes - Use <code>sccheck</code>	"How to Check Global Mounts in a Cluster" on page 69

▼ How to Add an Additional Cluster File System

Perform this task for each cluster file system you create after your initial Sun Cluster installation.



Caution - Be sure you have specified the correct disk device name. Creating a cluster file system destroys any data on the disks. If you specify the wrong device name, you will erase data that you may not intend to delete.

The prerequisites to add an additional cluster file system are:

- Superuser privilege on a node in the cluster.

- A device group (Solstice DiskSuite device group or a VxVM device group) or block disk slice upon which to create the cluster file system.

1. Become superuser on any node in the cluster.

Tip - For faster file system creation, become superuser on the current primary of the global device for which you are creating a file system.

2. Create a file system using the `newfs(1M)` command.

```
# newfs raw-disk-device
```

Table 3-3 shows examples of names for the *raw-disk-device* argument. Note that naming conventions differ for each volume manager.

TABLE 3-3 Sample Raw Disk Device Names

If Your Volume Manager Is ...	A Disk Device Name Might Be ...	Description
Solstice DiskSuite	/dev/md/oracle/rdisk/d1	Raw disk device d1 within the oracle metaset.
VERITAS Volume Manager	/dev/vx/rdisk/oradg/vol01	Raw disk device vol01 within the oradg disk group.
None	/dev/global/rdisk/d1s3	Raw disk device for block slice d1s3.

3. On each node in the cluster, create a mount point directory for the cluster file system.

A mount point is required *on each node*, even if the cluster file system will not be accessed on that node.

```
# mkdir -p /global/device-group/mount-point
```

device-group Name of the directory that corresponds to the name of the device group which contains the device.

mount-point	Name of the directory on which to mount the cluster file system.
--------------------	------------------------------------------------------------------

Tip - For ease of administration, create the mount point in the `/global/device-group` directory. This enables you to easily distinguish cluster file systems, which are globally available, from local file systems.

4. **On each node in the cluster, add an entry to the `/etc/vfstab` file for the mount point.**
 - a. **To automatically mount a cluster file system, set the `mount` at `boot` field to `yes`.**
 - b. **Use the following required mount options:**
 - The `global` mount option is required for all cluster file systems. This option identifies the file system as a cluster file system.
 - File system logging is required for all cluster file systems. UFS logging can be done either through the use of Solstice DiskSuite metatrans devices or directly through a Solaris UFS mount option. But, the two approaches should not be combined. If Solaris UFS logging is used directly, the `logging` mount option should be used. Otherwise, if metatrans file system logging is used, no additional mount option is needed.
 - c. **Ensure that, for each cluster file system, the information in its `/etc/vfstab` entry is identical on each node that has that entry.**
 - d. **Pay attention to boot order dependencies of the file systems.**
Normally, you should not nest the mount points for cluster file systems. For example, consider the scenario where `phys-schost-1` mounts disk device `d0` on `/global/oracle`, and `phys-schost-2` mounts disk device `d1` on `/global/oracle/logs`. With this configuration, `phys-schost-2` can boot up and mount `/global/oracle/logs` only after `phys-schost-1` boots and mounts `/global/oracle`.
 - e. **Make sure the entries in each node's `/etc/vfstab` file list common devices in the same order.**
For example, if `phys-schost-1` and `phys-schost-2` have a physical connection to devices `d0`, `d1`, and `d2`, the entries in their respective `/etc/vfstab` files should be listed as `d0`, `d1`, and `d2`.

Refer to the `vfstab(4)` man page for details.

5. **On any node in the cluster, verify that mount points exist and `/etc/vfstab` file entries are correct on all nodes of the cluster.**

```
# sccheck
```

If there are no errors, nothing is returned.

6. From any node in the cluster, mount the cluster file system.

```
# mount /global/device-group/mount-point
```

7. On each node of the cluster, verify that the cluster file system is mounted.

You can use either the `df(1M)` or `mount(1M)` command to list mounted file systems.

3.4.0.1 Example—Adding a Cluster File System

The following example creates a UFS cluster file system on the Solstice DiskSuite metadvice `/dev/md/oracle/rdsk/d1`.

```
# newfs /dev/md/oracle/rdsk/d1
...

[on each node:]
# mkdir -p /global/oracle/d1

# vi /etc/vfstab
#device          device      mount   FS      fsck    mount   mount
#to mount        to fsck    point   type    pass   at boot options
#
/dev/md/oracle/dsk/d1 /dev/md/oracle/rdsk/d1 /global/oracle/d1 ufs 2 yes global,logging
[save and exit]

[on one node:]
# sccheck

# mount /global/oracle/d1
# mount
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/
largefiles on Sun Oct 3 08:56:16 1999
```

▼ How to Remove a Cluster File System

You ‘remove’ a cluster file system by merely unmounting it. If you want to also remove or delete the data, remove the underlying disk device (or metadvice or volume) from the system.

Note - Cluster file systems are automatically unmounted as part of the system shutdown that occurs when you run `scshutdownd(1M)` to stop the entire cluster. A cluster file system is not unmounted when you run `shutdown` to stop a single node. However, if the node being shut down is the only node with a connection to the disk, any attempt to access the cluster file system on that disk results in an error.

The prerequisites to unmount cluster file systems are:

- Superuser privilege on a node in the cluster.
- The file system cannot be busy. A file system is considered busy if a user is in a directory in the file system, or if a program has a file open in that file system. The user or program could be running on any node in the cluster.

1. **Become superuser on a node in the cluster.**

2. **Determine which cluster file systems are mounted.**

```
# mount -v
```

3. **On each node, list all processes that are using the cluster file system, so you know which processes you are going to stop.**

```
# fuser -c [ -u ] mount-point
```

-c Reports on files that are mount points for file systems and any files within those mounted file systems.

-u (Optional) Displays the user login name for each process ID.

mount-point Specifies the name of the cluster file system for which you want to stop processes.

4. **On each node, stop all processes for the cluster file system.**

Use your preferred method for stopping processes. If necessary, use the following command to force termination of processes associated with the cluster file system.

```
# fuser -c -k mount-point
```

A `SIGKILL` is sent to each process using the cluster file system.

5. **On each node, verify that no processes are using the file system.**

```
# fuser -c mount-point
```

6. From just one node, umount the file system.

```
# umount mount-point
```

mount-point Specifies the name of the cluster file system you want to unmount. This can be either the directory name where the cluster file system is mounted, or the device name path of the file system.

7. (Optional) Edit the `/etc/vfstab` file to delete the entry for the cluster file system being removed.

Perform this step on each cluster node that has an entry for this cluster file system in its `/etc/vfstab` file.

8. (Optional) Remove the disk device group/metadevice/plex.

See your volume manager documentation for more information.

3.4.0.1 Example—Removing a Cluster File System

The following example removes a UFS cluster file system mounted on the Solstice DiskSuite metadevice `/dev/md/oracle/rdisk/d1`.

```
# mount -v
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/
largefiles on Sun Oct  3 08:56:16 1999
# fuser -c /global/oracle/d1
/global/oracle/d1: 4006c
# fuser -c -k /global/oracle/d1
/global/oracle/d1: 4006c
# fuser -c /global/oracle/d1
/global/oracle/d1:
# umount /global/oracle/d1

(on each node, remove the highlighted entry:)
# vi /etc/vfstab
#device          device          mount  FS      fsck    mount  mount
#to mount        to fsck        point  type    pass    at boot options
#
/dev/md/oracle/dsk/d1 /dev/md/oracle/rdisk/d1 /global/oracle/d1 ufs 2 yes global,logging
[Save and exit.]
```

Note - To remove the data on the cluster file system, remove the underlying device. See your volume manager documentation for more information.

▼ How to Check Global Mounts in a Cluster

The `sccheck(1M)` utility verifies the syntax of the entries for cluster file systems in the `/etc/vfstab` file. If there are no errors, nothing is returned.

Note - Run `sccheck` after making cluster configuration changes, such as removing a cluster file system, that have affected devices or volume management components.

1. Become superuser on a node in the cluster.
2. Check the cluster global mounts.

```
# sccheck
```

3.4.1 How to Remove a Node From a Disk Device Group (Solstice DiskSuite)

Use this procedure to remove a cluster node from disk device groups (diskset) running Solstice DiskSuite.

1. Determine the disk device group(s) of which the node to be removed is a member.

```
# scstat -D
```

2. Become superuser on the node that currently owns the disk device group from which you want to remove the node.
3. Delete from the disk device group the hostname of the node being removed.
Repeat this step for each disk device group from which the node is being removed.

```
# metaset -s setname -d -f -h node
```

<code>-s setname</code>	Specifies the disk device group (diskset) name
<code>-f</code>	Force
<code>-d</code>	Deletes from the disk device group
<code>-h nodelist</code>	Removes the node from the list of nodes that can master the disk device group

Note - The update can take several minutes to complete.

4. Verify that the node has been removed from the disk device group.

The disk device group name will match the diskset name specified with `metaset`.

```
# scstat -D
```

3.4.1.1 Example—Removing a Node From a Disk Device Group (SDS)

The following example shows the removal of the host name from a disk device group (`metaset`) and verifies that the node has been removed from the disk device group. Although the example shows the removal of a node from a single disk device group, a node can belong to more than one disk device group at a time. Repeat the `metaset` command for each disk device group from which you want to remove the node.

```
[Determine the disk device group(s) for the node:]
# scstat -D
  -- Device Group Servers --
                Device Group  Primary      Secondary
                -----
Device group servers: dg-schost-1  phys-schost-1  phys-schost-2
[Become superuser.]
[Remove the hostname from all disk device groups:]
# metaset -s dg-schost-1 -d -f -h phys-schost-2
[Verify removal of the node:]
# scstat -D
  -- Device Group Servers --
                Device Group  Primary      Secondary
                -----
Device group servers: dg-schost-1  phys-schost-1  -
```

3.4.2 How to Remove a Node From a Disk Device Group (VERITAS Volume Manager)

Use this procedure to remove a cluster node from an existing cluster disk device group (disk group) running VERITAS Volume Manager (VxVM).

1. Determine the disk device group of which the node to be removed is a member.

```
# scstat -D
```

2. Become superuser on a current cluster member node.

3. Execute the `scsetup` utility.

```
# scsetup
```

The Main Menu appears.

4. Reconfigure a disk device group by entering 3 (Device groups and volumes).

5. Remove the node from the VxVM disk device group by entering 5 (Remove a node from a VxVM device group).

Follow the prompts to remove the cluster node from the disk device group. You will be asked for information about the following:

VxVM device group

Node name

6. Verify that the node has been removed from the VxVM disk device group:

```
# scstat -D
...
Device group name: devicegroupname
Device group type: VxVM
Device group failback enabled: no
Device group node list: nodename
Diskgroup name: diskgroupname
...
```

3.4.2.1 Example—Removing a Node From a Disk Device Group (VxVM)

This example shows removal of the node named `phys-schost-4` from the `dg1` VxVM disk device group.

```
[Determine the disk device group for the node:]
# scstat -D
-- Device Group Servers --
      Device Group  Primary          Secondary
      -----
Device group servers: dg-schost-1  phys-schost-1  phys-schost-2
[Become superuser and execute the scsetup utility:]
# scsetup
[Select option 3:]
*** Main Menu ***
Please select from one of the following options:
```

(continued)

```

...
3) Device groups and volumes
...
Option: 3
[Select option 5:]
*** Device Groups Menu ***
Please select from one of the following options:
...
5) Remove a node from a VxVM device group
...
Option: 5
[Answer the questions to remove the node:]
>>> Remove a Node from a VxVM Device Group <<<
...
Is it okay to continue (yes/no) [yes]? yes
...
Name of the VxVM device group from which you want to remove a node? dg1
Name of the node to remove from this group? phys-schost-4
Is it okay to proceed with the update (yes/no) [yes]? yes

scconf -r -D name=dg1,nodelist=phys-schost-4

Command completed successfully.
Hit ENTER to continue:

```

```

[Quit the scsetup Device Groups Menu and Main Menu:]
...
Option: q
[Verify that the node was removed:]
# scstat -D
...
Device group name: dg1
Device group type: VxVM
Device group failback enabled: no
Device group node list: phys-schost-3
Diskgroup name: dg1
...

```


Administering Quorum

This chapter provides the procedures for administering quorum within Sun Cluster.

This is a list of the procedures in this chapter.

- “How to Add a Quorum Device” on page 75
- Section 4.1.1 “How to Remove a Quorum Device” on page 76
- “How to Replace a Quorum Device” on page 80
- “How to Put a Quorum Device Into Maintenance State” on page 80
- “How to Put a Cluster Node Into Maintenance State” on page 81
- “How to Bring a Node Out of Maintenance State and Reset Quorum” on page 83
- “How to List the Quorum Configuration” on page 85
- Section 4.1.1 “How to Remove a Quorum Device” on page 76
- Section 4.1.2 “How to Remove the Last Quorum Device From a Cluster” on page 78

Most examples shown in this chapter are from a three-node cluster.

See the *Sun Cluster 3.0 Concepts* document for a conceptual overview of quorum and quorum devices.

4.1 Administering Quorum Overview

You can use the `scconf(1M)` command to perform all quorum administrative procedures. In addition, some procedures can be accomplished by using the `scsetup(1M)` interactive utility. Whenever possible, quorum procedures are described in this chapter using `scsetup`.

If any quorum-related `scconf` command is interrupted or fails, the quorum configuration information can become inconsistent in the cluster configuration database. If this occurs, either rerun the command or run `scconf` with the `reset` option to reset the quorum configuration.

Note - The `scsetup(1M)` utility is an interactive interface to the `scconf(1M)` command. When `scsetup` runs, it generates `scconf` commands. These generated commands are shown in the examples at the end of the procedures.

There are two commands that can be used to view the quorum configuration: `scstat -q` option and `scconf -p`. Most of the verification steps in this chapter use `scconf`, but you can substitute `scstat -q` if you find its output more useful.

TABLE 4-1 Task Map: Administering Quorum

Task	For Instructions, Go To...
Add a quorum device to a cluster - Use <code>scsetup</code>	"How to Add a Quorum Device" on page 75
Remove a quorum device from a cluster - Use <code>scsetup</code> (to generate <code>scconf</code>)	Section 4.1.1 "How to Remove a Quorum Device" on page 76
Remove the last quorum device from a cluster - Use <code>scsetup</code> (to generate <code>scconf</code>)	Section 4.1.2 "How to Remove the Last Quorum Device From a Cluster" on page 78
Replace a quorum device in a cluster - Use the add and remove procedures	"How to Replace a Quorum Device" on page 80
Put a quorum device into maintenance state (While in maintenance state, the quorum device does not participate in voting to establish the quorum.) - Use <code>scsetup</code> (to generate <code>scconf</code>)	"How to Put a Quorum Device Into Maintenance State" on page 80

TABLE 4-1 Task Map: Administering Quorum *(continued)*

Task	For Instructions, Go To...
Put a cluster node into maintenance state (While in maintenance state, the node does not participate in voting to establish the quorum.) - Use <code>scsetup</code> (to generate <code>scconf</code>)	“How to Put a Cluster Node Into Maintenance State” on page 81
Reset the quorum configuration to its default state - Use <code>scsetup</code> (to generate <code>scconf</code>)	“How to Bring a Node Out of Maintenance State and Reset Quorum” on page 83
List the quorum devices and vote counts - Use <code>scconf</code>	“How to List the Quorum Configuration” on page 85

▼ How to Add a Quorum Device

To complete this procedure, identify a disk drive by its device ID (DID) that is shared by the nodes. Use the `scdidadm -L` command to see the list of DID names. Refer to `scdidadm(1M)` for additional information.

See `scsetup(1M)` and `scconf(1M)` for information on the commands used in this procedure.

- 1. Become superuser on a node of the cluster.**
- 2. Enter the `scsetup` utility.**

```
# scsetup
```

The Main Menu appears.

- 3. To work with quorum devices, enter 1 (Quorum).**

The Quorum Menu appears.

- 4. To add a quorum device, enter 1 (Add a quorum disk).**

Follow the instructions and enter the device to be used as the quorum device.

5. Verify that the quorum device has been added.

```
# scstat -q
```

6. Repeat Step 3 on page 75 through Step 5 on page 76 for each group of nodes that shares a storage enclosure.

4.1.0.1 Example—Adding a Quorum Device

The following example shows the `scconf` command generated by `scsetup` when it adds a quorum device and a verification step.

```
# scconf -a -q globaldev=d20
# scstat -q

-- Quorum Summary --

Quorum votes possible: 4
Quorum votes needed:   3
Quorum votes present:  4

-- Quorum Votes by Node --

      Node Name      Present Possible Status
      -----
Node votes: phys-schost-1    1         1   Online
Node votes: phys-schost-2    1         1   Online

-- Quorum Votes by Device --

      Device Name      Present Possible Status  Owner
      -----
Device votes: /dev/did/rdisk/d3s2  1         1   Online  phys-schost-1
Device votes: /dev/did/rdisk/d4s2  1         1   Online  phys-schost-1
```

4.1.1 How to Remove a Quorum Device

When a quorum disk is removed, it no longer participates in the voting to establish quorum. Note that all two-node clusters require that at least one quorum device be configured. If this is the last quorum device on a cluster, `scconf(1M)` will fail to remove the device from the configuration.

Note - If the device you intend to remove is the last quorum device in the cluster, see the procedure Section 4.1.2 “How to Remove the Last Quorum Device From a Cluster” on page 78.

1. Become superuser on a node in the cluster.

2. Determine the quorum device to be removed.

```
# scconf -pv | grep Quorum
```

3. Execute the `scsetup(1M)` utility.

```
# scsetup
```

The Main Menu appears.

4. Access the Quorum Menu by entering 1 (Quorum) to work with the quorum device.

5. Remove the quorum device by entering 2 (Remove a quorum disk).

Answer the questions displayed during the removal process.

6. Quit `scsetup`.

7. Verify that the quorum device is removed.

```
# scstat -q
```

4.1.1.1 Example—Removing a Quorum Device

This example shows how to remove a quorum device from a cluster with two or more quorum devices configured.

```
[Become superuser on a node and place the node to be removed in maintenance state.]
[Determine the quorum device to be removed:]
# scconf -pv | grep Quorum
[Execute the scsetup utility:]
# scsetup
[Select option 1:]
*** Main Menu ***
    Please select from one of the following options:
        1) Quorum
        ...
    Option: 1
[Select option 2:]
*** Quorum Menu ***
    Please select from one of the following options:
        ...
        2) Remove a quorum disk
        ...
    Option: 2
```

(continued)

```

[Answer the questions when prompted:]
>>> Remove a Quorum Disk <<<
...
Is it okay to continue (yes/no) [yes]? <Return>
Which quorum disk do you want to remove (d<N>)? d4
Is it okay to proceed with the update (yes/no) [yes]? <Return>

scconf -r -q globaldev=d4

Command completed successfully.
Hit ENTER to continue:
[Quit the scsetup Quorum Menu and Main Menu:]
...
Option: q
[Verify the quorum device is removed:]
# scstat -q

-- Quorum Summary --

Quorum votes possible:      3
Quorum votes needed:       2
Quorum votes present:      3

-- Quorum Votes by Node --

Node votes:      Node Name      Present Possible Status
Node votes:      phys-schost-1    1          1      Online
Node votes:      phys-schost-2    1          1      Online

-- Quorum Votes by Device --

Device votes:    Device Name      Present Possible Status Owner
Device votes:    /dev/did/rdisk/d3s2 1          1      Online phys-schost-1

```

4.1.2 How to Remove the Last Quorum Device From a Cluster

If the device you intend to remove is not the last quorum device in the cluster, use the previous procedure, Section 4.1.1 “How to Remove a Quorum Device” on page 76.

Note - All two-node clusters require at least one configured quorum device. If this is the last quorum device on a two-node cluster, the cluster must be placed into install mode before `scconf(1M)` allows you to remove the device from the configuration. This should only be done if a node is being removed from the cluster.

1. **Become superuser on a node in the cluster and place the node to be removed in maintenance state.**

See “How to Put a Cluster Node Into Maintenance State” on page 81.

2. Place the cluster in install mode.

```
# scconf -c -q installmode
```

3. Remove the quorum device using the `scconf` command.

```
# scconf -r -q globaldev=device
```

4. Verify that the quorum device has been removed.

```
# scstat -q
```

4.1.2.1 Example—Removing the Last Quorum Device

This example shows how to remove the last remaining quorum device in a cluster configuration.

```
[Become superuser on any node.]
[Place the cluster in install mode:]
# scconf -c -q installmode
[Remove the quorum device:]
# scconf -r -q globaldev=d3
[Verify the quorum device has been removed:]
# scstat -q

-- Quorum Summary --

Quorum votes possible:      2
Quorum votes needed:       2
Quorum votes present:      2

-- Quorum Votes by Node --

Node Name      Present Possible Status
-----
Node votes:    phys-schost-1    1         1    Online
Node votes:    phys-schost-2    1         1    Online

-- Quorum Votes by Device --

Device Name      Present Possible Status  Owner
-----
```

▼ How to Replace a Quorum Device

1. **Configure a new quorum device on the storage enclosure that currently includes the disk to be replaced.**

You need to first add a new quorum device to the configuration to take the place of the old device. See “How to Add a Quorum Device” on page 75 to add a new quorum device to the cluster.

2. **Remove the failed disk as a quorum device.**

See Section 4.1.1 “How to Remove a Quorum Device” on page 76 to remove the old quorum device from the configuration.

3. **Replace the failed disk.**

Refer to the hardware procedures in the *Sun Cluster 3.0 Hardware Guide* for your disk enclosure.

▼ How to Put a Quorum Device Into Maintenance State

Put a quorum device into maintenance state when taking the quorum device out of service for an extended period of time. This way, the device’s quorum vote count is set to zero and does not contribute to the quorum count while the device is being serviced.

You need to use the `scconf(1M)` command to put a quorum device into the maintenance state. The `scsetup` utility does not have the ability to put a quorum device into maintenance state.

Note - All two-node clusters require at least one configured quorum device. If this is the last quorum device on a two-node cluster, `scconf` will fail to put the device into maintenance state.

1. **Become superuser on a node of the cluster.**
2. **Put the quorum device into the maintenance state.**

```
# scconf -c -q globaldev=device,maintstate
```

-c Specifies the change form of the `scconf` command.

-q Manages the quorum options.

<code>globaldev=device</code>	Specifies the DID name of the disk device to change, for example, <code>d4</code> .
<code>maintstate</code>	Puts the node into maintenance state.

3. Verify that the quorum device is now in maintenance state.

The output for the device you placed in maintenance state should read zero for the Quorum Device Votes.

```
# scconf -p | grep -i quorum
```

4.1.2.1 Example—Putting a Quorum Device Into Maintenance State

The following example shows how to put a quorum device into maintenance state and how to verify the results.

```
# scconf -c -q globaldev=d20,maintstate
# scconf -p | grep -i quorum
Node quorum vote count:          1
Node quorum vote count:          1
Quorum devices:                  d20
Quorum device name:              d20
Quorum device votes:             0
Quorum device enabled:          no
Quorum device path:              /dev/did/rdisk/d20s2
Quorum device hosts (enabled):   phys-schost-2 phys-schost-3
Quorum device hosts (disabled):
```

4.1.2.2 Where to Go From Here

When you have completed the maintenance procedure on the quorum device, and want to re-enable the quorum device, see “How to Bring a Node Out of Maintenance State and Reset Quorum” on page 83 to reset the quorum device’s vote back to the default.

▼ How to Put a Cluster Node Into Maintenance State

Put a cluster node into maintenance state when taking the node out of service for an extended period of time. This way, the node does not contribute to the quorum count while it is being serviced. To put a cluster node into maintenance state, the node must be brought down using `scswitch(1M)` and `shutdown(1M)`.

Note - Use the Solaris `shutdown` command to shut down a single node. The `scshutdown` command should be used only when shutting down an entire cluster.

When a cluster node is brought down and put into maintenance state, all quorum devices that are configured with ports to the node have their quorum vote counts decremented by one. The node and quorum device vote counts are incremented by one when the node is taken out of maintenance mode and brought back online.

You need to use `scconf(1M)` to put a cluster node into the maintenance state. The `scsetup` utility does not have the ability to put a quorum device into maintenance state.

1. Become superuser on the node to be put into maintenance state.
2. On the node to be put into maintenance state, use `scswitch` to evacuate any resource groups and disk device groups from the node and then `shutdown` to bring it down and out of the cluster.

```
phys-schost-1# scswitch -s -h node
phys-schost-1# shutdown -g 0 -y
```

3. Become superuser on another node and put the node brought down in Step 2 on page 82 into the maintenance state.

```
phys-schost-2# scconf -c -q node=node,maintstate
```

<code>-c</code>	Specifies the change form of the <code>scconf</code> command.
<code>-q</code>	Manages the quorum options.
<code>node=</code> <i>node</i>	Specifies the node name or node ID of the node to change.
<code>maintstate</code>	Puts the node into maintenance state.

4. Verify that the cluster node is now in maintenance state.

```
phys-schost-2# scconf -p | grep -i vote
```

4.1.2.1 Example—Putting a Cluster Node Into Maintenance State

The following example moves a cluster node into maintenance state and verifies the results.

```

phys-schost-1# scswitch -S -h phys-schost-1
phys-schost-1# shutdown -g 0 -y
phys-schost-2# scconf -c -q node=phys-schost-1,maintstate
phys-schost-2# scconf -p | grep -i quorum
Node quorum vote count:                1
Quorum devices:                        d20
Quorum device name:                    d20
Quorum device votes:                    1
Quorum device enabled:                  yes
Quorum device path:                     /dev/did/rdisk/d5s2
Quorum device hosts (disabled):         phys-schost-1
Quorum device hosts (enabled):          phys-schost-2

```

4.1.2.2 Where to Go From Here

When you have completed the maintenance procedures, see “How to Bring a Node Out of Maintenance State and Reset Quorum” on page 83 to reset the quorum vote back to the default.

▼ How to Bring a Node Out of Maintenance State and Reset Quorum

After a quorum device or a node has been in maintenance state, and is being brought back online, use the following procedure to reset the quorum vote count to the default. For cluster nodes, the default quorum count is one. For quorum devices, the default quorum count is $N-1$, where N is the number of nodes with non-zero vote counts that have ports to the quorum device.

When a node has been put into maintenance state, the node’s quorum vote count is decremented by one. All quorum devices that are configured with ports to the node will also have their quorum vote counts decremented. When the quorum vote count is reset and a node is brought back out of maintenance state, both the node’s quorum vote count and the quorum device vote count are incremented by one.

You should run this procedure any time a node or a quorum device has been put into maintenance state and you are bringing it out of maintenance state.



Caution - If you do not specify either the `globaldev` or `node` options, the quorum count is reset for the entire cluster.

1. Become superuser on a node of the cluster.

2. Reset the quorum count.

This sample shows the quorum count being reset for a quorum device. Refer to Section 4.1.2.2 “Example—Resetting the Quorum Vote Count (Cluster Node and Quorum Devices)” on page 84 for an example using a cluster node rather than a quorum device.

For a cluster node, reset the quorum count from another node before rebooting the node or it might hang waiting for quorum.

```
# scconf -c -q globaldev=device,reset
```

-c	Specifies the change form of the <code>scconf</code> command.
-q	Manages the quorum options.
globaldev= device	Specifies the DID name of the quorum device to reset, for example, <code>d4</code> .
reset	The change flag that resets quorum.

3. If you are resetting the quorum count because a node was in maintenance state, reboot the node.

4. Verify the quorum vote count.

```
# scconf -p | grep -i quorum
```

4.1.2.1 Example—Resetting the Quorum Vote Count (Quorum Device)

The following example resets the quorum count for a quorum device back to the default and verifies the result.

```
# scconf -c -q globaldev=d20,reset
# scconf -p | grep -i quorum
Node quorum vote count:          1
Node quorum vote count:          1
Quorum devices:                  d20
Quorum device name:              d20
Quorum device votes:              1
Quorum device enabled:           yes
Quorum device path:              /dev/did/rdisk/d20s2
Quorum device hosts (enabled):    phys-schost-2 phys-schost-3
Quorum device hosts (disabled):
```

4.1.2.2 Example—Resetting the Quorum Vote Count (Cluster Node and Quorum Devices)

The following example resets the quorum count for a cluster node and its quorum devices back to their defaults and verifies the result.

```
# scconf -c -q node=phys-schost-1,reset
# scconf -pv | grep -i vote
Node quorum vote count:      1
Node quorum vote count:      1
Node quorum vote count:      1
(d20) Quorum device votes:    1
(d21) Quorum device votes:    1
```

▼ How to List the Quorum Configuration

You do not need to be superuser to list the quorum configuration.

- ◆ Use `scconf(1M)` to list the quorum configuration.

```
# scconf -p | grep -i quorum
```

4.1.2.1 Example—Listing the Quorum Configuration

```
# scconf -p | egrep ``Quorum | vote``
Node quorum vote count:      1
Node quorum vote count:      1
Quorum devices:              d20
Quorum device name:          d20
Quorum device votes:         1
Quorum device enabled:       yes
Quorum device path:          /dev/did/rdisk/d20s2
Quorum device hosts (enabled): phys-schost-2 phys-schost-3
Quorum device hosts (disabled):
```


Administering Cluster Interconnects and Public Networks

This chapter provides the software procedures for administering the Sun Cluster interconnects and public networks.

Administering the cluster interconnects and public networks consists of both hardware and software procedures. Typically, you configure the cluster interconnects and public networks, including NAFO groups, when you initially install and configure the cluster. If you later need to alter a cluster interconnect or public network configuration, you can use the software procedures in this chapter.

This is a list of the procedures in this chapter.

- “How to Check the Status of the Cluster Interconnect” on page 89
- “How to Add Cluster Transport Cables and Transport Adapters” on page 89
- “How to Remove Cluster Transport Cables and Transport Adapters” on page 92
- “How to Enable a Cluster Transport Cable” on page 96
- “How to Disable a Cluster Transport Cable ” on page 97
- “How to Create a NAFO Group” on page 100
- “How to Delete a NAFO Group” on page 102
- “How to Add an Adapter to a NAFO Group” on page 104
- “How to Remove an Adapter From a NAFO Group” on page 105
- “How to Switch a NAFO Group’s Active Adapter” on page 106
- “How to Check the Status of NAFO Groups” on page 107
- “How to Change Public Network Management Tunable Parameters” on page 108

For a high-level description of the related procedures in this chapter, see Table 5-1 and Table 5-2.

Refer to the *Sun Cluster 3.0 Concepts* document for background and overview information on the cluster interconnects and public networks.

5.1 Administering the Cluster Interconnects

This section provides the procedures for reconfiguring cluster interconnects, such as cluster transport adapters and cluster transport cables. These procedures require that you install Sun Cluster software.

Most of the time, you can use the `scsetup` utility to administer the cluster transport for the cluster interconnects. See the `scsetup(1M)` man page for more information.

For cluster software installation procedures, see the *Sun Cluster 3.0 Installation Guide*. For procedures about servicing cluster hardware components, see the *Sun Cluster 3.0 Hardware Guide*.

TABLE 5-1 Task Map: Administering the Cluster Interconnect

Task	For Instructions, Go To...
Administer the cluster transport - Use <code>scsetup</code>	"How to Access the <code>scsetup</code> Utility" on page 21
Check the status of the cluster interconnect - Use <code>scstat</code>	"How to Check the Status of the Cluster Interconnect" on page 89
Add a cluster transport cable or transport adapter - Use <code>scsetup</code>	"How to Add Cluster Transport Cables and Transport Adapters" on page 89
Remove a cluster transport cable or transport adapter - Use <code>scsetup</code>	"How to Remove Cluster Transport Cables and Transport Adapters" on page 92
Enable a cluster transport cable - Use <code>scsetup</code>	"How to Enable a Cluster Transport Cable" on page 96
Disable a cluster transport cable - Use <code>scsetup</code>	"How to Disable a Cluster Transport Cable " on page 97

▼ How to Check the Status of the Cluster Interconnect

You do not need to be logged in as superuser to perform this procedure.

1. Check the status of the cluster interconnect.

```
# scstat -W
```

2. Refer to the following for common status messages.

Status Message	Description and Possible Action
Path online	The path is currently functioning correctly. No action is necessary.
Path waiting	The path is currently being initialized. No action is necessary.
Path faulted	The path is not functioning. Take corrective action.

5.1.0.1 Example—Checking the Status of the Cluster Interconnect

The following example shows the status of a functioning cluster interconnect.

```
# scstat -W
-- Cluster Transport Paths --
      Endpoint                Endpoint                Status
      -----                -
Transport path:  phys-schost-1:qfe1  phys-schost-2:qfe1  Path online
Transport path:  phys-schost-1:qfe0  phys-schost-2:qfe0  Path online
Transport path:  phys-schost-1:qfe1  phys-schost-3:qfe1  Path online
Transport path:  phys-schost-1:qfe0  phys-schost-3:qfe0  Path online
Transport path:  phys-schost-2:qfe1  phys-schost-3:qfe1  Path online
Transport path:  phys-schost-2:qfe0  phys-schost-3:qfe0  Path online
```

▼ How to Add Cluster Transport Cables and Transport Adapters

1. Ensure that the physical cluster transport cables are installed.

For the procedure on installing a cluster transport cable, see the *Sun Cluster 3.0 Hardware Guide*.

2. Become superuser on any node in the cluster.

3. Execute the `scsetup` utility.

```
# scsetup
```

The Main Menu appears.

4. Access the Cluster Interconnect Menu by entering 2 (Cluster interconnect).

5. Add the cable by entering 1 (Add a transport cable).

Follow the instructions and enter the requested information. A junction is not needed for direct connections in a cluster with two nodes. However, you need to enter both the node and the adapter names of one of the endpoints of the cable you are trying to identify.

```
>>> Add a Cluster Transport Cable <<<
...
Is it okay to continue (yes/no) [yes]? <Return>
...
To which node do you want to add the cable? nodename
Name of the adapter to use on ``node``? adaptername
Name of the junction at the other end of the cable? junctionname
Okay to use the default for this cable connection [yes]? <Return>
Is it okay to proceed with the update (yes/no) [yes]? <Return>
```

6. Add the adapter by entering 2 (Add a transport adapter to a node).

Follow the instructions and enter the requested information. A junction is not needed for direct connections in a cluster with two nodes. However, you need to enter both the node and the adapter names of one of the endpoints of the cable you are trying to identify.

```
>>> Add a Cluster Transport Adapter <<<
...
Is it okay to continue (yes/no) [yes]? <Return>
...
To which node do you want to add the cable? nodename
Name of the adapter to use on ``node``? adaptername
Name of the junction at the other end of the cable? junctionname
Okay to use the default for this cable connection [yes]? <Return>
```

(continued)

```
Is it okay to proceed with the update (yes/no) [yes]? <Return>
```

7. Verify that the cluster transport cable or transport adapter is added.

```
# scconf -p | grep cable
# scconf -p | grep adapter
```

5.1.0.1 Example—Adding a Cluster Transport Cable and a Transport Adapter

The following example shows how to add a transport cable or transport adapter to a node using the `scsetup` command.

```
[Ensure the physical cable is installed.]
[Become superuser on a node in the cluster.]
[Execute the scsetup utility:]
# scsetup
[Enter option 2:]
*** Main Menu ***
Please select from one of the following options:
...
2) Cluster interconnect
... Option: 2
[Enter option 1 to add a cable:]
*** Cluster Interconnect Menu ***
Please select from one of the following options:
...
1) Add a transport cable
...
Option: 1
[Enter the requested information:]
>>> Add a Cluster Transport Cable <<<
...
Is it okay to continue (yes/no) [yes]? <Return> .
...
To which node is the cable attached? phys-schhost-2
Name of the adapter on 'phys-schost-2'? qfe-0
Is it okay to proceed with the update (yes/no) [yes]? <Return> .
scconf -a -m endpoint=phys-schost-2:hme1,endpoint=hub0
```

(continued)

```

Command completed successfully.      Hit ENTER to continue: <Return>
[Enter option 2 to add an adapter:]
*** Cluster Interconnect Menu ***
Please select from one of the following options:
...
2) Add a transport adapter to a node
...
Option: 2
[Enter the requested information:]
>>> Add a Cluster Transport Adapter <<<
...
Is it okay to continue (yes/no) [yes]? <Return> .
...
To which node is the cable attached? phys-schost-2
Name of the adapter on ``phys-schost-2``? qfe-0
Is it okay to proceed with the update (yes/no) [yes]? <Return> .
scsconf -a -m endpoint=phys-schost-2:hme1,endpoint=hub0
Command completed successfully.      Hit ENTER to continue: <Return>
[Verify that the cable and adapter are added:]
# scsconf -p | grep cable
Transport cable:    phys-schost-2:qfe0@1 ethernet-1@2    Enabled
Transport cable:    phys-schost-3:qfe0@1 ethernet-1@3    Enabled
Transport cable:    phys-schost-1:qfe0@0 ethernet-1@1    Enabled
# scsconf -p | grep adapter
Node transport adapters:                qfe2 hme1 qfe0
Node transport adapter:                  qfe0
Node transport adapters:                qfe0 qfe2 hme1
Node transport adapter:                  qfe0
Node transport adapters:                qfe0 qfe2 hme1
Node transport adapter:                  qfe0

```

▼ How to Remove Cluster Transport Cables and Transport Adapters

Use the following procedure to remove cluster transport cables and cluster transport adapters from a node. When a cable is disabled, the two endpoints of the cable remain configured. An adapter cannot be removed if it is still in use as an endpoint in a transport cable.



Caution - A cluster node needs at least one functioning cable attached to the cluster interconnect to be a member of the cluster. Always verify the status of a node's cluster interconnect before disabling a cable. Only disable a cable connection after you have verified that it is redundant; that is, that another connection is available. Disabling a node's last remaining working cable takes the node out of cluster membership.

1. Become superuser on any node in the cluster.
2. Check the status of the remaining cluster transport path.

```
# scstat -W
```



Caution - If you receive an error such as “path faulted” while attempting to remove one node of a two-node cluster, investigate the problem before continuing with this procedure. Such a problem could indicate that a node path is unavailable. Removing the remaining good path takes the node out of cluster membership and could result in a cluster reconfiguration.

3. Execute the `scsetup` utility.

```
# scsetup
```

The Main Menu appears.

4. Access the Cluster Interconnect Menu by entering 2 (Cluster interconnect).

5. Remove the cable by entering 4 (Remove a transport cable).

Follow the instructions and enter the requested information. You need to enter both the node and the adapter names of one of the endpoints of the cable you are trying to identify.

```
*** Cluster Interconnect Menu ***
Please select from one of the following options:
...
4) Remove a transport cable
...
Option: 4
...
>>> Remove a Cluster Transport Cable <<<
...
Is it okay to continue (yes/no) [yes]? <Return>
...
To which node is the cable attached? nodename
Name of the adapter on ``node``? adaptername
Is it okay to proceed with the update (yes/no) [yes]? <Return>
...
```

(continued)

```
Command completed successfully.      Hit ENTER to continue: <Return>
```

Note - If you are removing a physical cable, disconnect the cable between the port and the destination device.

6. To remove an adapter, enter 5 (Remove a transport adapter from a node).

Follow the instructions and enter the requested information. You need to enter both the node and the adapter names of one of the endpoints of the cable you are trying to identify.

```
*** Cluster Interconnect Menu **      Please select from one of the following options:
...
5) Remove a transport adapter
...
Option: 5
...
>>> Remove a Cluster Transport Adapter <<<
...
Is it okay to continue (yes/no) [yes]? <Return>
...
To which node is the cable attached? nodename
Name of the adapter on ``node``? adaptername
Is it okay to proceed with the update (yes/no) [yes]? <Return>
...
Command completed successfully.      Hit ENTER to continue: <Return>
```

Note - If you are removing a physical adapter from a node, see the *Sun Cluster 3.0 Hardware Guide* for hardware service procedures.

7. Verify that the cable or the adapter has been removed.

```
# scconf -p | grep cable
# scconf -p | grep adapter
```

The transport cable or adapter removed from the given node should not appear in the output from this command.

5.1.0.1 Example—Removing a Transport Cable or Transport Adapter

The following example shows how to remove a transport cable or transport adapter using the `scsetup` command.

```
[Become superuser on a node in the cluster.]
[Execute the scsetup utility:]
# scsetup
[Enter option 2:]
*** Main Menu ***
Please select from one of the following options:
...
2) Cluster interconnect
... Option: 2
[Enter option 4 to remove a cable:]
*** Cluster Interconnect Menu ***
Please select from one of the following options:
...
4) Remove a transport cable
...
Option: 4
[Enter the requested information:]
>>> Remove a Cluster Transport Cable <<<
...
Is it okay to continue (yes/no) [yes]? <Return> .
...
To which node is the cable attached? phys-schhost-2
Name of the adapter on ``phys-schost-2''? qfe-0
Is it okay to proceed with the update (yes/no) [yes]? <Return> .
scconf -a -m endpoint=phys-schost-2:hme1,endpoint=hub0
Command completed successfully. Hit ENTER to continue: <Return>
[Enter option 5 to remove an adapter:]
*** Cluster Interconnect Menu ***
Please select from one of the following options:
...
5) Remove a transport adapter from a node
...
Option: 5
[Enter the requested information:]
>>> Remove a Cluster Transport Adapter <<<
...
Is it okay to continue (yes/no) [yes]? <Return> .
...
To which node is the cable attached? phys-schhost-2
Name of the adapter on ``phys-schost-2''? qfe-0
Is it okay to proceed with the update (yes/no) [yes]? <Return> .
```

(continued)

```

scconf -r -m endpoint=phys-schost-2:hme1

Command completed successfully.      Hit ENTER to continue: <Return>
[Verify that the cable or adapter is removed:]
# scconf -p | grep cable
Transport cable:   phys-schost-2:qfe0@1 ethernet-1@2    Enabled
Transport cable:   phys-schost-3:qfe0@1 ethernet-1@3    Enabled
Transport cable:   phys-schost-1:qfe0@0 ethernet-1@1    Enabled
# scconf -p | grep adapter
Node transport adapters:   qfe2 hme1 qfe0
Node transport adapter:    qfe0
Node transport adapters:   qfe0 qfe2 hme1
Node transport adapter:    qfe0
Node transport adapters:   qfe0 qfe2 hme1
Node transport adapter:    qfe0

```

▼ How to Enable a Cluster Transport Cable

This option is used to enable an already existing cluster transport cable.

1. **Become superuser on any node in the cluster.**
2. **Enter the `scsetup` utility.**

```
# scsetup
```

The Main Menu appears.

3. **Access the Cluster Interconnect Menu by entering 2 (Cluster interconnect).**
4. **Enable the transport cable by entering 7 (Enable a transport cable).**

Follow the instructions when prompted. You need to enter both the node and the adapter names of one of the endpoints of the cable you are trying to identify.

5. **Verify that the cable is enabled.**

```
# scconf -p | grep cable
```


5.1.0.1 Example—Enabling a Cluster Transport Cable

This example shows how to enable a cluster transport cable on adapter `qfe-1` located on the node `phys-schost-2`.

```
[Become superuser on any node.]
[Enter the scsetup utility:]
# scsetup
[Enter option 7:]
*** Cluster Interconnect Menu ***
  Please select from one of the following options:
  ...
  7) Enable a transport cable
  ...
  Option: 7
[Enter the requested information:]
>>> Enable a Cluster Transport Cable <<<
...
Is it okay to continue (yes/no) [yes]? <Return>
...
To which node is the cable attached? phys-schost-2
Name of the adapter on ``phys-schost-2``? qfe-1
Is it okay to proceed with the update (yes/no) [yes]? <Return>

scconf -c -m endpoint=phys-schost-2:qfe1,state=enabled

  Command completed successfully.
  Hit ENTER to continue: <Return>
[Verify that the cable is enabled:]
# scconf -p | grep cable
Transport cable:  phys-schost-2:qfe1@0 ethernet-1@2    Enabled
Transport cable:  phys-schost-3:qfe0@1 ethernet-1@3    Enabled
Transport cable:  phys-schost-1:qfe0@0 ethernet-1@1    Enabled
```

▼ How to Disable a Cluster Transport Cable

You might need to disable a cluster transport cable to temporarily shut down a cluster interconnect path. This is useful when troubleshooting a cluster interconnect problem or when replacing cluster interconnect hardware.

When a cable is disabled, the two endpoints of the cable remain configured. An adapter cannot be removed if it is still in use as an endpoint in a transport cable.



Caution - A cluster node needs at least one functioning cable attached to the cluster interconnect to be a member of the cluster. Always verify the status of a node's cluster interconnect before disabling a cable. Only disable a cable connection after you have verified that it is redundant; that is, that another connection is available. Disabling a node's last remaining working cable takes the node out of cluster membership.

1. **Become superuser on any node in the cluster.**

2. Check the status of the cluster interconnect before disabling a cable.

```
# scstat -W
```



Caution - If you receive an error such as “path faulted” while attempting to remove one node of a two-node cluster, investigate the problem before continuing with this procedure. Such a problem could indicate that a node path is unavailable. Removing the remaining good path takes the node out of cluster membership and could result in a cluster reconfiguration.

3. Enter the `scsetup` utility.

```
# scsetup
```

The Main Menu appears.

4. Access the Cluster Interconnect Menu by entering 2 (Cluster interconnect).

5. Disable the cable by entering 8 (Disable a transport cable).

Follow the instructions and enter the requested information. All of the components on this cluster interconnect will be disabled. You need to enter both the node and the adapter names of one of the endpoints of the cable you are trying to identify.

6. Verify that the cable is disabled.

```
# sconfig -p | grep cable
```

5.1.0.1 Example—Disabling a Cluster Transport Cable

This example shows how to disable a cluster transport cable on adapter `qfe-1` located on the node `phys-schost-2`.

```
[Become superuser on any node.]
[Enter the scsetup utility:]
# scsetup
[Enter option 8:]
*** Cluster Interconnect Menu ***
    Please select from one of the following options:
    ...
```

(continued)

```

    8) Disable a transport cable
    ...
    Option: 8
[Enter the requested information:]
>>> Disable a Cluster Transport Cable <<<
    ...
    Is it okay to continue (yes/no) [yes]? <Return>
    ...
    To which node is the cable attached? phys-schhost-2
    Name of the adapter on ``phys-schost-2``? qfe-1
    Is it okay to proceed with the update (yes/no) [yes]? <Return>

scconf -c -m endpoint=phys-schost-2:qfe1,state=disabled

    Command completed successfully.
    Hit ENTER to continue: <Return>
[Verify that the cable is disabled:]
# scconf -p | grep cable
Transport cable:    phys-schost-2:qfe1@0 ethernet-1@2    Disabled
Transport cable:    phys-schost-3:qfe0@1 ethernet-1@3    Enabled
Transport cable:    phys-schost-1:qfe0@0 ethernet-1@1    Enabled

```

5.2 Administering the Public Network

If you need to alter a public network configuration, you can use the software procedures in this section.

When administering public network adapters pay attention to the following:

- Avoid unconfiguring (unplumbing) or bringing down the active adapter of a NAFO group without first switching over the active adapter to a backup adapter in the group. See “How to Switch a NAFO Group’s Active Adapter” on page 106.
- Avoid rewiring backup adapters to different subnets without first removing them from their respective NAFO groups.
- Logical adapter operations can be done on the active adapter even if monitoring is on for the group.
- You must maintain at least one public network connection for each node in the cluster. The cluster is inaccessible without a public network connection.

For cluster software installation procedures, see the *Sun Cluster 3.0 Installation Guide*. For procedures about servicing public networking hardware components, see the *Sun Cluster 3.0 Hardware Guide*.

TABLE 5-2 Task Map: Administering the Public Network

Task	For Instructions, Go To...
Create a NAFO group on a node.	"How to Create a NAFO Group" on page 100
Add more public network adapters to a node for increased availability of the public network.	"How to Add an Adapter to a NAFO Group" on page 104
Delete a NAFO group to not have any monitoring or failover for any adapter in the group.	"How to Delete a NAFO Group" on page 102
Remove backup adapters from an existing NAFO group to enable the adapter to be removed from the system, to be replaced, or to be reconnected to a different subnet and used as backup for another NAFO group.	"How to Remove an Adapter From a NAFO Group" on page 105
Switch the active adapter to a backup adapter so that the current active adapter can be removed from the NAFO group.	"How to Switch a NAFO Group's Active Adapter" on page 106
Check the status of NAFO groups.	"How to Check the Status of NAFO Groups" on page 107
Change parameters to tune the PNM fault detection and failover process.	"How to Change Public Network Management Tunable Parameters" on page 108

▼ How to Create a NAFO Group

1. Note the following requirements for creating a NAFO group:

- All public network adapters must be configured to belong to a NAFO group.
- For any given node, there can be at most one NAFO group on a given subnet.
- All adapters in a given NAFO group must be connected to the same subnet.

- Only one adapter in a given NAFO group can have a hostname association, that is, an `/etc/hostname.adapter` file.
 - A public network adapter can belong to only one NAFO group.
2. **Become superuser on the node being configured for a NAFO group.**
 3. **For this node, find out the public network adapters that are physically connected to the same subnet.**
These adapters form the backup adapters for the NAFO group.
 4. **Determine if you need to create an `/etc/hostname.adapter` file for one of the public network adapters.**
 - If no (the file already exists), proceed to Step 5 on page 101.
 - If yes (the file does not exist), create the file and add the hostname of the public network adapter's IP address to the file. For example:

```
# vi /etc/hostname.hme0
phys-schost-1
```

5. **Edit the `/etc/inet/hosts` file to add the IP address and corresponding hostname assigned to the public network adapter.**
For example, the following shows the IP address 192.29.75.101 and hostname `phys-schost-1` added to the `/etc/inet/hosts` file.

```
# vi /etc/inet/hosts
192.29.75.101    phys-schost-1
```

Note - If a naming service is used, this information should also exist in the naming service database.

6. **Create the NAFO group.**

```
# pnmset -c nafo-group -o create adapter [adapter ...]
```

- c **nafo-group** Performs a configuration subcommand for the specified NAFO group. NAFO groups must be named `nafoN`, where *N* is a nonnegative integer identifier for the group.

Group names are local to each node. Thus, the same NAFO group name can be used on multiple nodes.

`-o create` Creates the new NAFO group.

`adapter [adapter ...]` Specifies the public network adapter(s) that serves as the backup adapter. See Step 3 on page 101 above.

Note - If an adapter is already configured, it will be chosen as the active adapter and the `pnmset` command does not alter its state. Otherwise, one of the backup adapters will be configured and assigned the IP address found in the `/etc/hostname.adapter` file for the NAFO group.

7. Verify the status of the NAFO group.

```
# pnmstat -l
```

5.2.0.1 Example—Creating a NAFO Group

The following example shows the creation of a NAFO group (`nafo0`) configured with two network adapters (`qfe0` and `qfe1`).

```
# pnmstat -l
# pnmset -c nafo0 -o create qfe0 qfe1
# pnmstat -l
group  adapters      status  fo_time  act_adp
nafo0  qfe0:qfe1         OK      NEVER    qfe0
```

▼ How to Delete a NAFO Group

Delete a NAFO group when you do not want monitoring and failover for any adapter in the group. To be deleted, a NAFO group cannot be in use by logical host resource groups or shared address resource groups.

1. Become superuser on the node that contains the NAFO group that is being deleted.
2. Identify whether the NAFO group is being used by any logical host or shared address resources.

```
# scrgadm -pv
```

You can also use `scrgadm -pvv` (with two `v` flags) to locate the resources that are using the NAFO group you are going to delete.

3. Switch logical host resource groups and shared address resource groups using this NAFO group to another node.

```
# scswitch -z -g resource-group -h node
```

`-z -g resource-group` Switches the specified resource group.

`-h node` Specifies the name of the node to switch the resource group to.

4. Delete the NAFO group.

```
# pnmset -c nafo-group -o delete
```

`-c nafo-group` Specifies the NAFO group to be deleted.

`-o delete` Deletes the NAFO group.

5. Verify the status of the NAFO group.

The deleted NAFO group should not appear in the listing.

```
# pnmstat -l
```

5.2.0.1

Example—Deleting a NAFO Group

The following example shows the NAFO group named `nafo1` deleted from the system. Logical host resource group `lh-rg-1`, which uses this NAFO group, is first switched to a different node.

```
# scswitch -z -g lh-rg-1 -h phys-schost-2
# pnmstat -l
group    adapters      status  fo_time  act_adp
nafo0    qfe0:qfe1      OK      NEVER    qfe0
nafo1    qfe2           OK      NEVER    qfe2
# pnmset -c nafo1 -o delete
# pnmstat -l
group    adapters      status  fo_time  act_adp
nafo0    qfe0:qfe1      OK      NEVER    qfe0
```

▼ How to Add an Adapter to a NAFO Group

You can add adapters to an existing NAFO group to provide additional backup adapters for the NAFO group and thereby increase the availability of public network connectivity for the cluster node.

1. **Do you need to install the new public network adapter card(s) in the node(s)?**
 - If yes, see the *Sun Cluster 3.0 Hardware Guide* for instructions.
 - If no, proceed to Step 2 on page 104.
2. **Make sure the adapter to be added to the NAFO group is connected to the same subnet as the active adapter for the NAFO group.**
3. **Make sure the adapter is not plumbed, and that it does not have an associated `/etc/hostname.adapter` file.**
4. **Become superuser on the node that contains the NAFO group to which the new adapter is being added.**
5. **Add the adapter to the NAFO group.**

```
# pnmset -c nafo-group -o add adapter
```

`-c nafo-group` Specifies the NAFO group to which the new adapter is being added.

`-o add adapter` Specifies the public network adapter being added to the named NAFO group.

6. **Verify the status of the NAFO group.**

```
# pnmstat -l
```

5.2.0.1 Example—Adding an Additional Public Network Adapter to a NAFO Group

The following example shows the addition of adapter `qfe2` to NAFO group `nafo0` which already contained two adapters (`qfe0`, `qfe1`).

```
# pnmstat -l
group  adapters      status  fo_time  act_adp
nafo0  qfe0:qfe1        OK      NEVER    qfe0
# pnmset -c nafo0 -o add qfe2
```

(continued)


```
# pnmstat -l
group  adapters      status  fo_time  act_adp
nafo0  qfe0:qfe1:qfe2    OK      NEVER    qfe0
```

▼ How to Remove an Adapter From a NAFO Group

Remove backup adapters from an existing NAFO group to enable the adapter to be removed from the system, to be replaced, or to be reconnected to a different subnet and used as backup for another NAFO group.



Caution - Removing the last backup adapter from a NAFO group results in no protection against faults detected on the active adapter, reducing public network availability for the cluster node.

1. If you want to remove the active adapter, first switch to another adapter in the group.

See “How to Switch a NAFO Group’s Active Adapter” on page 106.

2. As superuser, remove the adapter from the NAFO group.

```
# pnmset -c nafo-group -o remove adapter
```

-c *nafo-group* Specifies the NAFO group from which to remove the adapter.

-o remove *adapter* Removes the adapter from the NAFO group.

3. Verify the status of the NAFO group.

The deleted adapter should not appear in the listing for the NAFO group.

```
# pnmstat -l
```

5.2.0.1 Example—Removing an Adapter From a NAFO Group

The following example removes adapter `qfe2` from NAFO group `nafo0`.

```
# pnmstat -l
group  adapters      status  fo_time  act_adp
nafo0  qfe0:qfe1:qfe2  OK      NEVER    qfe0
# pnmset -c nafo0 -o remove qfe2
# pnmstat -l
group  adapters      status  fo_time  act_adp
nafo0  qfe0:qfe1      OK      NEVER    qfe0
```

▼ How to Switch a NAFO Group's Active Adapter

Switch the active adapter to a backup adapter so that the current active adapter can be removed from the NAFO group. The `pnmmd(1M)` daemon moves all IP addresses hosted by the current active adapter to the new active adapter in a similar fashion as a fault-triggered adapter failover.

Note - Connections can experience a brief delay while the switchover is taking place. Otherwise, the operation is transparent to higher-level applications.

- 1. Ensure the physical connectivity of the new active adapter is identical to that of the current active adapter.**

If the new active adapter fails to host some of the IP addresses as the current active adapter, network and data services that depend on those IP addresses are interrupted until the physical connectivity is fixed or a subsequent successful failover occurs.

- 2. Become superuser on the node that contains the NAFO group whose active adapter you want to switch.**

- 3. Switch the active adapter.**

```
# pnmset -c nafo-group -o switch adapter
```

`-c nafo-group` Specifies the NAFO group containing the adapter to switch.

`-o switch adapter` Makes the specified adapter the active adapter in the NAFO group.

- 4. Verify the status of the NAFO group.**

The “switched-to” adapter should now appear as the active adapter.

```
# pnmstat -l
```

5.2.0.1

Example—Switching a NAFO Group’s Active Adapter

The following example switches the active adapter to `qfe1` from `qfe0`.

```
# pnmstat -l
group  adapters      status  fo_time  act_adp
nafo0  qfe0:qfe1        OK      NEVER    qfe0
# pnmset -c nafo0 -o switch qfe1
# pnmstat -l
group  adapters      status  fo_time  act_adp
nafo0  qfe0:qfe1        OK      11       qfe1
```

▼ How to Check the Status of NAFO Groups

- ♦ Run the `pnmstat(1M)` command to list information about the current setup and status of all NAFO groups on a node.

```
# pnmstat -l
```

You can also use the `pnmptor(1M)` and `pnmrtop(1M)` commands to get information on adapters.

5.2.0.1

Examples—Checking the Status of NAFO Groups

The following example shows the status of a node’s three NAFO groups.

```
# pnmstat -l
Group  adapters      status  fo_time  act_adp
nafo0  qfe5           OK      NEVER    qfe5
nafo1  qfe6           OK      NEVER    qfe6
nafo2  qfe7           OK      NEVER    qfe7
```

5.2.0.2

Examples—Finding the Active Adapter For a NAFO Group

The following example shows that the active adapter in NAFO group `nafo0` is adapter `qfe5`.

```
# pnmptor nafo0
qfe5
```

5.2.0.3 Examples—Finding an Adapter’s NAFO Group

The following example shows that adapter `qfe5` belongs to NAFO group `nafo0`.

```
# pnmrtop qfe5
nafo0
```

▼ How to Change Public Network Management Tunable Parameters

There are four tunable parameters in this algorithm: `inactive_time`, `ping_timeout`, `repeat_test`, and `slow_network`. These parameters provide an adjustable trade-off between speed and correctness of fault detection. See Table 5–3 for more information.

Use this procedure to change the default Public Network Management (PNM) values for the `pnmd(1M)` daemon.

1. Become superuser on a node in the cluster.
2. If it doesn’t already exist, create the `pnmparams` file.

```
# vi /etc/cluster/pnmparams
```

3. Use the following table to set PNM parameters.

Note - Settings in the `/etc/cluster/pnmparams` file apply to all NAFO groups on the node. Lines started with a pound sign (#) are ignored. Other lines in the file must be of the format: *variable=value*.

TABLE 5–3 Public Network Management Tunable Parameters

Parameter	Description
<code>inactive_time</code>	Number of seconds between successive probes of the packet counters of the current active adapter. Default is 5.
<code>ping_timeout</code>	Time-out value in seconds for the <code>ALL_HOST_MULTICAST</code> and subnet broadcast pings. Default is 4.
<code>repeat_test</code>	Number of times to do the ping sequence before declaring that the active adapter is faulty and failover is triggered. Default is 3.

TABLE 5–3 Public Network Management Tunable Parameters *(continued)*

Parameter	Description
<code>slow_network</code>	Number of seconds waited after each ping sequence before checking packet counters for any change. Default is 2.
<code>warmup_time</code>	Number of seconds waited after failover to a backupadapter before resuming fault monitoring. This allows extra time for any slow driver or port initialization. Default is 0.

Note - The changes don't take affect until the next time the `pnmd` daemon starts up.

5.2.0.1

Example—Changing Public Network Management Tunable Parameters

The following shows a sample `/etc/cluster/pnmparams` file, with two parameters changed from their default values.

```
inactive_time=3
repeat_test=5
```


Administering the Cluster

This chapter provides the procedures for administering items that affect the entire cluster.

This is a list of the procedures in this chapter.

- “How to Change the Cluster Name” on page 112
- “How to Map Node ID to Node Name” on page 113
- “How to Work With New Cluster Node Authentication” on page 113
- “How to Reset the Time of Day in a Cluster” on page 115
- “How to Enter the OpenBoot PROM (OBP) on a Node” on page 116
- Section 6.2.1 “How to Add a Cluster Node to the Authorized Node List” on page 118
- Section 6.3.1 “How to Remove a Node From the Cluster Software Configuration” on page 121

6.1 Administering the Cluster Overview

TABLE 6–1 Task Map: Administering the Cluster

Task	For Instructions, Go To...
Change the name of the cluster.	“How to Change the Cluster Name” on page 112
List node IDs and their corresponding node names.	“How to Change the Cluster Name” on page 112
Permit or deny new nodes to add themselves to the cluster.	“How to Work With New Cluster Node Authentication” on page 113
Change the time for a cluster using the Network Time Protocol (NTP).	“How to Reset the Time of Day in a Cluster” on page 115
Bring down a node and enter the OpenBoot™ PROM.	“How to Enter the OpenBoot PROM (OBP) on a Node” on page 116

▼ How to Change the Cluster Name

If necessary, you can change the cluster name after initial installation.

1. Become superuser on a node in the cluster.

2. Enter the `scsetup(1M)` utility.

```
# scsetup
```

The Main Menu appears.

3. To change the cluster name, enter 6 (Other cluster properties).

The Other Cluster Properties menu appears.

4. Make your selection from the menu and follow the onscreen instructions.

6.1.0.1 Example—Changing the Cluster Name

The following example shows the `scconf(1M)` command generated from the `scsetup` utility to change to the new cluster name, `dromedary`.

```
# scconf -c -C cluster=dromedary
```


▼ How to Map Node ID to Node Name

During Sun Cluster installation, each node is automatically assigned a unique node ID number. The node ID number is assigned to a node in the order in which it joins the cluster for the first time; once assigned, the number cannot be changed. The node ID number is often used in error messages to identify which cluster node the message concerns. Use this procedure to determine the mapping between node IDs and node names.

You do not need to be superuser to list configuration information.

1. Use `scconf(1M)` to list the cluster configuration information.

```
% scconf -pv | grep "Node ID"
```

6.1.0.1

Example—Mapping the Node ID to the Node Name

The following example shows the node ID assignments

```
% scconf -pv | grep ``Node ID``
(phys-schost-1) Node ID:          1
(phys-schost-2) Node ID:          2
(phys-schost-3) Node ID:          3
```

▼ How to Work With New Cluster Node Authentication

Sun Cluster enables you to determine if new nodes can add themselves to the cluster and with what type of authentication. You can permit any new node to join the cluster over the public network, deny new nodes from joining the cluster, or indicate a specific node that can join the cluster. New nodes can be authenticated by using either standard UNIX or Diffie-Hellman (DES) authentication. If you select DES authentication, you must also configure all necessary encryption keys before a node can join. See the `keyserv(1M)` and `publickey(4)` man pages for more information.

1. Become superuser on a node in the cluster.
2. Enter the `scsetup(1M)` utility.

```
# scsetup
```

The Main Menu appears.

3. To work with cluster authentication, enter 5 (New nodes).

The New Nodes menu appears.

4. Make your selection from the menu and follow the onscreen instructions.

6.1.0.1 Examples—Preventing New Machines From Being Added to the Cluster

The following example shows the `scconf(1M)` command generated from the `scsetup` utility that would prevent new machines from being added to the cluster.

```
# scconf -a -T node=.
```

6.1.0.2 Examples—Permitting All New Machines to Be Added to the Cluster

The following example shows the `scconf` command generated from the `scsetup` utility that would enable all new machines to be added to the cluster.

```
# scconf -r -T all
```

6.1.0.3 Examples—Specifying a New Machine to Be Added to the Cluster

The following example shows the `scconf` command generated from the `scsetup` utility to enable a single new machine to be added to the cluster.

```
# scconf -a -T node=phys-schost-4
```

6.1.0.4 Examples—Setting the Authentication to Standard UNIX

The following example shows the `scconf` command generated from the `scsetup` utility to reset to standard UNIX authentication for new nodes joining the cluster.

```
# scconf -c -T authtype=unix
```

6.1.0.5 Examples—Setting the Authentication to DES

The following example shows the `scconf` command generated from the `scsetup` utility to use DES authentication for new nodes joining the cluster.

```
# scconf -c -T authtype=des
```

Note - When using DES authentication, you need to also configure all necessary encryption keys before a node can join the cluster. See the `keyserv(1M)` and `publickey(4)` man pages for more information.

▼ How to Reset the Time of Day in a Cluster

Sun Cluster uses the Network Time Protocol (NTP) to maintain time synchronization between cluster nodes. Adjustments in the cluster occur automatically as needed when nodes synchronize their time. See the *Sun Cluster 3.0 Concepts* document and the *Network Time Protocol User's Guide* for more information.



Caution - When using NTP, do not attempt to adjust the cluster time while the cluster is up and running. This includes using the `date(1)`, `rdate(1M)`, or `xntpdate(1M)` commands interactively or within `cron(1M)` scripts.

1. Become superuser on a node in the cluster.

2. Shut down the cluster.

```
# scshutdown -g0 -y
```

3. Boot each node into non-cluster node.

```
ok boot -x
```

4. On a single node, set the time of day by running the `date(1)` command.

```
# date HHMMSS
```

5. On the other machines, synchronize the time to that node by running the `rdate(1M)` command.

```
# rdate hostname
```

6. Boot each node to restart the cluster.

```
# reboot
```

7. **Verify that the change took place on all cluster nodes.**

On each node, run the `date(1M)` command.

```
# date
```

▼ How to Enter the OpenBoot PROM (OBP) on a Node

Use this procedure if you need to configure or change OpenBoot PROM settings.

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

```
# 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 installed at a site.

2. **Shut down the cluster node gracefully by using the `scswitch(1M)` command to evacuate any resource or disk device groups and then `shutdown(1M)` to bring the node to the OBP prompt.**

```
# scswitch -s -h node
# shutdown -g 0 -y
```

3. **Send a break to the node.**

```
telnet> send brk
```

4. **Execute the OpenBoot PROM commands.**

6.2 Adding a Cluster Node

The following table lists the tasks to perform when adding a node to an existing cluster.

TABLE 6-2 Task Map: Adding a Node

Task	For Instructions, Go To...
Add the cluster interconnects to the new node. <ul style="list-style-type: none">- Install the host adapter, add the transport junction, cable the interconnect.	<i>Sun Cluster 3.0 Hardware Guide</i> <ul style="list-style-type: none">- Adding and Replacing Cluster Interconnect and Public Hardware
Add shared storage	<i>Sun Cluster 3.0 Hardware Guide</i> <ul style="list-style-type: none">- Installing and Replacing the StorEdge MultiPack Enclosure- Installing and Replacing the StorEdge D1000 Disk Array- Installing and Replacing the StorEdge A5x00 Disk Array
Add the node to the authorized node list <ul style="list-style-type: none">- Use <code>scsetup</code>.	<i>Sun Cluster 3.0 System Administration Guide</i> <ul style="list-style-type: none">- How to Add a Cluster Node
Install and configure the software on the new cluster node <ul style="list-style-type: none">- Install the Solaris Operating Environment and Sun Cluster software- Configure the node as part of the cluster	<i>Sun Cluster 3.0 Installation Guide</i> <ul style="list-style-type: none">- Installing and Configuring Sun Cluster Software

6.2.1 How to Add a Cluster Node to the Authorized Node List

Before adding a machine to an existing cluster, be sure the node has all of the necessary software and hardware correctly installed and configured, including a good physical connection to the private cluster interconnect, as indicated in the “Adding a Node” task map. Refer to the *Sun Cluster 3.0 Installation Guide* and the `scinstall(1M)` man page for more information regarding software installations. For hardware installations, refer to the *Sun Cluster 3.0 Hardware Guide* or the hardware documentation that shipped with your server.

1. **Become superuser on a current cluster member node.**
2. **Execute the `scsetup` utility.**

```
# scsetup
```

The Main Menu appears.

3. **Access the New Nodes Menu option by entering 5 at the Main Menu.**
4. **Modify the authorized list by entering 3 (Specify the name of the machine) at the New Nodes Menu.**
5. **Specify the name of a machine that can add itself.**
Follow the prompts to add the cluster node. You will be asked for the name of the node to be added.
6. **Verify that the node has been added to the authorized list.**

```
# scconf -p | grep ``Cluster new node``
```

6.2.1.1 Example—Adding a Cluster Node

The following example shows how to add a node named `phys-schost-3` to an existing cluster.

```
[Become a superuser.]
[Execute scsetup utility.]
# scsetup
*** Main Menu ***
    Please select from one of the following options:
    Option: 5
*** New Nodes Menu ***
```

(continued)

```

Please select from one of the following options:
...
3) Specify the name of a machine which may add itself
...
Option: 3
>>> Specify a Machine which may Install itself into the Cluster <<<
...
Is it okay to continue (yes/no) [yes]? <Return>
Name of the host to add to the list of recognized machines? phys-schost-3
Is it okay to proceed with the update (yes/no) [yes]? <Return>

scconf -a -T node=phys-schost-3

Command completed successfully.
[Quit the scsetup New Nodes Menu and Main Menu:]
...
Option: q
[Verify the node has been added.]
# scconf -p | grep ``Cluster new``
Cluster new node authentication:      unix
Cluster new node list:                phys-schost-3

```

6.2.1.2 Where to Go From Here

Sun Cluster 3.0 Installation Guide: Installing and Configuring Sun Cluster Software.

6.3 Removing a Cluster Node

The following table lists the tasks to perform when removing a node from an existing cluster.

TABLE 6-3 Task Map: Removing a Cluster Node

Task	For Instructions, Go To...
Place node being removed into maintenance state - Use <code>shutdown</code> and <code>scconf</code>	<i>Sun Cluster 3.0 System Administration Guide</i> : Chapter 4, Administering Quorum - How to Put a Cluster Node Into Maintenance State
Remove node from all resource groups - Use <code>scrgadm</code>	<i>Sun Cluster 3.0 Data Services Installation and Configuration Guide</i> : Chapter 9, Administering Data Service Resources - How to Remove a Node from an Existing Resource Group
Remove node from all device groups of which the node is a member - Use volume manager commands	<i>Sun Cluster 3.0 System Administration Guide</i> : Chapter 3, Administering Global Devices and Cluster File Systems - How to Remove a Node from a Disk Device Group (SDS) - How to Remove a Node from a Disk Device Group (VxVM)
Remove all logical transport connections to the node being removed - Use <code>scsetup</code>	<i>Sun Cluster 3.0 System Administration Guide</i> : Chapter 5, Administering Cluster Networks - How to Remove Cluster Transport Cables and Transport Adapters To remove the physical hardware from the node, see <i>Sun Cluster 3.0 Hardware Guide</i> : Chapter 3, Installing and Maintaining Cluster Interconnect and Public Network Hardware.
Remove all quorum devices shared with the node being removed - Use <code>scsetup</code>	<i>Sun Cluster 3.0 System Administration Guide</i> : Chapter 4, Administering Quorum - How to Remove a Quorum Device

TABLE 6–3 Task Map: Removing a Cluster Node *(continued)*

Task	For Instructions, Go To...
Remove node from the cluster software configuration - Use <code>scconf</code>	<i>Sun Cluster 3.0 System Administration Guide</i> : Chapter 6, Administering the Cluster - How to Remove a Cluster Node
Remove required shared storage from the node and cluster - Follow the procedures in your volume manager documentation and hardware guide	SDS or VxVM administration guide; <i>Sun Cluster 3.0 Hardware Guide</i> : - How to Remove a StorEdge MultiPack Enclosure - How to Remove a StorEdge D1000 Disk Array - How to Remove a StorEdge A5x00 Disk Array

6.3.1 How to Remove a Node From the Cluster Software Configuration

This is the last software configuration procedure that needs to be accomplished in the process for removing a node from a cluster. Before beginning this procedure, you must complete all the prerequisite tasks listed in the “Removing a Cluster Node” task map. When finished with this procedure, remove the hardware connections as described in the *Sun Cluster 3.0 Hardware Guide*.

1. Become superuser on a node in the cluster.

Note - Be sure you have placed the node in maintenance state and removed it from all resource groups, device groups, and quorum device configurations before continuing with this procedure.

2. Determine the boot disks in the cluster.

```
# scconf -pvv | grep Local_Disk
```

3. Identify which boot disks in the cluster are connected to the node being removed.

```
# scdidadm -L did-name
```

4. Remove the `localonly` flag from each boot disk.

```
# scconf -c -D name=devicename,localonly=false
```

5. Remove the node from all raw disk device groups, of which it is a member.

This step must be completed for each disk device group.

```
# scconf -pvv | grep nodename | grep Device
# scconf -r -D name=devicename,nodelist=nodename
```

6. Remove the node from the cluster.

```
# scconf -r -h node=nodename
```

7. Verify the node removal using `scstat`.

```
# scstat -n
```

Note - After the device has been removed from the cluster, you must reinstall the Solaris operating environment on the removed host before it can be placed back into service in any capacity.

6.3.1.1 Example—Removing a Cluster Node

This example shows how to remove a node (`phys-schost-2`) from a cluster.

```
[Become superuser on any node.]
[Determine the boot disks on the node:]
# scconf -pvv | grep Local_Disk
(dsk/d4) Device group type:          Local_Disk (dsk/
d3) Device group type:          Local_Disk
# scdidadm -L d4
...
4      phys-schost-2:/dev/rdsk/clt3d0 /dev/did/rdsk/d4
[Remove the localonly flag:]
```

(continued)

```
# scconf -c -D name=dsk/d4,localonly=false
[Remove the node from all raw disk device groups:]
# scconf -pvv | grep phys-schost-2 | grep Device
(dsk/d4) Device group node list:  phys-schost-2
(dsk/d2) Device group node list:  phys-schost-1, phys-schost-2 (dsk/
d1) Device group node list:  phys-schost-1, phys-schost-2
# scconf -r -D name=dsk/d4,nodelist=phys-schost-2
# scconf -r -D name=dsk/d2,nodelist=phys-schost-2
# scconf -r -D name=dsk/d1,nodelist=phys-schost-2
[Remove the node from the cluster:]
# scconf -r -h node=phys-schost-2
[Verify node removal:]# scstat -n

-- Cluster Nodes --

                Node name          Status
                -----          -
Cluster node:   phys-schost-1      Online
```

6.3.1.2 Where to Go From Here

- Sun Cluster 3.0 Hardware Guide:*
- How to Remove a StorEdge MultiPack Enclosure
 - How to Remove a StorEdge D1000 Disk Array
 - How to Remove a StorEdge A5x00 Disk Array

Patching Sun Cluster Software and Firmware

This chapter provides the procedures for adding and removing patches for a Sun Cluster configuration.

This is a list of the procedures in this chapter.

- “How to Apply a Non-Rebooting Sun Cluster Patch” on page 131
- “How to Apply a Rebooting Patch (Node)” on page 127
- “How to Apply a Rebooting Patch (Cluster and Firmware)” on page 130
- “How to Remove a Sun Cluster Patch” on page 132

7.1 Patching Sun Cluster Overview

Due to the nature of a cluster, all cluster member nodes must be at the same patch level for proper cluster operation. When patching a node with a Sun Cluster-specific patch, you might need to temporarily remove a node from cluster membership or stop the entire cluster before installing the patch. This section describes these steps.

Before applying a Sun Cluster patch, first check the Sun Cluster web page for any special instructions; for the current URL, see the *Sun Cluster 3.0 Release Notes* or contact Enterprise Services. If there aren't any special instructions, check the patch's README file.

Note - For Sun Cluster-specific patches, always defer to the patch's README file for instructions that supersede procedures in this chapter.

Patch installation on all cluster nodes falls into one of the following scenarios:

- **Rebooting patch (node)**—A node must be booted to single-user mode before the patch can be applied, then rebooted to join the cluster. In doing so, you need to put the node into a “quiet” state by first switching any resource groups or disk device groups from the node to be patched to another cluster member. Also, you apply the patch to one cluster node at a time to avoid bringing down the entire cluster.

The cluster itself remains available during this type of patch application, even though individual nodes are temporarily shut down. A patched node is able to rejoin a cluster as a member node even though other nodes are not yet at the same patch level.

- **Rebooting patch (cluster and firmware)**—The cluster must be shut down and each node must be booted to single-user mode to apply the software or firmware patch. Then you reboot the nodes to rejoin the cluster. For this type of patch, the cluster is unavailable during patch application.
- **Non-rebooting patch**—A node does not have to be in a “quiet” state (it can still be mastering resource groups or device groups), nor does it have to be shut down or rebooted when applying the patch. However, you should still apply the patch to one node at a time and verify that the patch works before patching another node.

Note - Underlying cluster protocols do not change due to a patch.

You use the `patchadd` command to apply a patch to the cluster, and `patchrm` to remove a patch (when possible).

7.1.1 Sun Cluster Patch Tips

Use the following tips to help you administer Sun Cluster patches more efficiently:

- Refer to the Sun Cluster website for any special instructions associated with the patch or firmware update. For the current URL, see the *Sun Cluster 3.0 Release Notes* or contact Enterprise Services.
- Always read the patch README file before applying the patch.
- Apply all patches (required and recommended) before running the cluster in a production environment.
- Check the hardware firmware levels and install any required firmware updates that may be needed.
- All nodes acting as cluster members must have the same patches.

- Keep cluster subsystem patches up to date. This includes volume management, storage device firmware, cluster transport, and so forth.
- Review patch reports regularly, such as once a quarter, and patch a Sun Cluster configuration using the recommended patch suite.
- Apply selective patches as recommended by Enterprise Services.
- Test failover after major patch updates; be prepared to back out the patch if cluster operation is degraded or impaired.

7.2 Patching Sun Cluster

TABLE 7-1 Task Map: Patching the Cluster

Task	For Instructions, Go To...
Apply a non-rebooting Sun Cluster patch to one node at a time without having to stop the node.	"How to Apply a Non-Rebooting Sun Cluster Patch" on page 131
Apply a rebooting Sun Cluster patch after taking the cluster member to non-cluster mode.	"How to Apply a Rebooting Patch (Node)" on page 127 "How to Apply a Rebooting Patch (Cluster and Firmware)" on page 130
Remove a Sun Cluster patch. - You can back out the patch if necessary.	"How to Remove a Sun Cluster Patch" on page 132

▼ How to Apply a Rebooting Patch (Node)

Apply the patch to one node in the cluster at a time to keep the cluster itself operational during the patch process. With this procedure, you must first shut down the node and boot it to single-user mode before applying the patch.

1. **Before applying the patch, check the Sun Cluster product web page for any special pre- or post-installation instructions.**
2. **Become superuser on the node to which you are applying the patch.**

3. List the resource groups and device groups on the node being patched.

```
# scrgadm -pv  
# scstat
```

4. Switch all resource groups, resources, and device groups from the node being patched to other cluster members.

```
# scswitch -s -h node
```

- h *node*** Specifies the node to switch resource and devices groups to.
- s** Evacuates all device services and resource groups from the specified node.

5. Shut down the node.

```
# shutdown -g 0 [-y]
```

- g 0** Specifies to shut down grace period to be zero.
- y** Specifies to answer *yes* to the confirmation prompt, so the command can be run without user intervention.

6. Boot the node in single-user mode.

```
ok boot -s
```

7. Apply the patch.

```
# patchadd patch-dir patch-id
```

patch-dir Specifies the directory location of the patch.

patch-id Specifies the patch number of a given patch.

8. Reboot the node into the cluster.

```
# reboot
```


9. Verify that the patch has been installed successfully.

```
# showrev -p | grep patch-id
```

10. Verify that the patch works, and that the node and cluster are operating normally.

11. Switch back all resource groups, resources, and device groups (optional).

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

12. Repeat Step 2 on page 127 through Step 11 on page 129 for all remaining cluster nodes.

7.2.0.1 Example—Applying a Rebooting Patch (Node)

The following example shows the application of a rebooting Sun Cluster patch to a node.

```
# scrgadm -pv
...
RG Name: schost-sa-1
...
# scstat
...
Device Group Name:                dg-schost-1
...
# scswitch -z -h phys-schost-2 -D dg-schost-1
# scswitch -z -h phys-schost-2 -g schost-sa-1
# shutdown -g 0 -y
...
ok boot -s
...
# patchadd 10-34567
...
# reboot
...
# showrev -p | grep 10-34567
# scswitch -z -h phys-schost-1 -D dg-schost-1
# scswitch -z -h phys-schost-1 -g schost-sa-1
```

7.2.0.2 Where to Go From Here

If you need to back out a patch, see “How to Remove a Sun Cluster Patch” on page 132.

▼ How to Apply a Rebooting Patch (Cluster and Firmware)

With this procedure, you must first shut down the cluster and boot each node to single-user mode before applying the patch.

1. Before applying the patch, check the Sun Cluster product web page for any special pre- or post-installation instructions.
2. Become superuser on a node in the cluster.
3. Shut down the cluster.

```
# scshutdown [-y] [-g grace-period] [``message``]
```

-y Specifies to answer *yes* to the confirmation prompt.

-g grace-period Specifies, in seconds, the amount of time to wait before shutting down. Default grace period is 60 seconds.

message Specifies the warning message to broadcast. Use quotes if *message* contains multiple words.

4. Boot each node into single-user mode.

On the console of each node, run the following command.

```
ok boot -s
```

5. Apply the software or firmware patch.

On one node at a time, run the following command.

```
# patchadd patch-dir patch-id
```

patch-dir Specifies the directory location of the patch.

patch-id Specifies the patch number of a given patch.

Note - Always defer to the instructions in the patch directory that supersede procedures in this chapter.

6. Verify that the patch has been installed successfully on each node.

```
# showrev -p | grep patch-id
```

7. After applying the patch to all nodes, reboot the nodes into the cluster.

On each node, run the following command.

```
# reboot
```

8. Verify that the patch works, and that the nodes and cluster are operating normally.

7.2.0.1 Example—Applying a Rebooting Patch (Cluster)

The following example shows the application of a rebooting Sun Cluster patch to a cluster.

```
# scshutdown -g 0 -y
...
ok boot -s
...
# patchadd 10-34567
(Apply patch to other cluster nodes)
...
# showrev -p | grep 10-34567
# reboot
```

7.2.0.2 Where to Go From Here

If you need to back out a patch, see “How to Remove a Sun Cluster Patch” on page 132.

▼ How to Apply a Non-Rebooting Sun Cluster Patch

Apply the patch to one node in the cluster at a time. When applying a non-rebooting patch, you do not need to first shut down the node receiving the patch.

1. Before applying the patch, check the Sun Cluster product web page for any special pre- or post-installation instructions.
2. Apply the patch on a single node.

```
# patchadd patch-dir patch-id
```

patch-dir Specifies the directory location of the patch.

patch-id Specifies the patch number of a given patch.

3. Verify that the patch has been installed successfully.

```
# showrev -p | grep patch-id
```

4. Verify that the patch works, and that the node and cluster are operating normally.
5. Repeat Step 2 on page 131 through Step 4 on page 132 for the remaining cluster nodes.

7.2.0.1 Example—Applying a Non-Rebooting Sun Cluster Patch

```
# patchadd 10-34567
...
# showrev -p | grep 10-34567
```

7.2.0.2 Where to Go From Here

If you need to back out a patch, see “How to Remove a Sun Cluster Patch” on page 132.

▼ How to Remove a Sun Cluster Patch

If necessary, you can back out (remove) a Sun Cluster patch.

1. Become superuser on the node from which you are removing the patch.
2. List the resource groups and device groups on the node having the patch removed.

```
# scrgadm -pv
# scstat
```

3. Switch all resource groups, resources, and device groups from the node having the patch removed to other cluster members.

```
# scswitch -s -h node
```

- h **node** Specifies the node to switch resource and devices groups to.
- S Evacuates all device services and resource groups from the specified node.

4. Shut down the node.

```
# shutdown [-y] [-g grace-period] [``message'']
```

- y Specifies to answer *yes* to the confirmation prompt.
- g **grace-period** Specifies, in seconds, the amount of time to wait before shutting down. Default grace period is 60 seconds.
- message** Specifies the warning message to broadcast. Use quotes if *message* contains multiple words.

5. Boot the node in non-cluster mode.

```
ok boot -x
```

6. Remove the patch.

```
# patchrm patch-id
```

- patch-id** Specifies the patch number of a given patch.

7. Reboot the node.

```
# reboot
```

8. Verify that the patch has been removed successfully.

```
# showrev -p | grep patch-id
```

9. Verify that the node and cluster are operating normally.

10. Switch back all resource groups, resources, and device groups.

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

11. Repeat Step 1 on page 132 through Step 10 on page 133 for the remaining cluster nodes.

7.2.0.1 Example—Removing a Sun Cluster Patch

The following example shows the removal of a Sun Cluster patch.

```
# scrgadm -pv
...
RG Name: schost-sa-1
...
# scstat
...
Device Group Name: dg-schost-1
...
# scswitch -S -h phys-schost-2
# shutdown -y -g 5 "Rebooting down node for maintenance"
...
ok boot -x
...
# patchrm 10-34567
...
# reboot
...
# pkgchk -v 10-34567
...
# scswitch -S -h phys-schost-1
```

Backing Up and Restoring a Cluster

This is a list of step-by-step instructions in this chapter.

- “How to Find File System Names to Back Up” on page 136
- “How to Determine the Number of Tapes Needed for a Full Backup” on page 137
- “How to Back Up the root (/) File System” on page 137
- “How to Perform Online Backups for Mirrors (Solstice DiskSuite)” on page 139
- “How to Perform Online Backups for Volumes (VERITAS Volume Manager)” on page 142
- “How to Restore Individual Files Interactively (Solstice DiskSuite)” on page 148
- “How to Restore the root (/) File System (Solstice DiskSuite)” on page 148
- “How to Restore a root (/) File System That Was on a Metadevice (Solstice DiskSuite)” on page 151
- “How to Restore a Non-Encapsulated root (/) File System (VERITAS Volume Manager)” on page 156
- “How to Restore an Encapsulated root (/) File System (VERITAS Volume Manager)” on page 159

8.1 Backing Up Cluster Files

TABLE 8-1 Task Map: Backing Up Cluster Files

Task	For Instructions, Go To...
Find the names of the file systems you want to back up.	"How to Find File System Names to Back Up" on page 136
Calculate how many tapes you will need to contain a full backup.	"How to Determine the Number of Tapes Needed for a Full Backup" on page 137
Back up the root file system.	"How to Back Up the root (/) File System" on page 137
Perform online backup for mirrored or plexed file systems.	"How to Perform Online Backups for Mirrors (Solstice DiskSuite)" on page 139
	"How to Perform Online Backups for Volumes (VERITAS Volume Manager)" on page 142

▼ How to Find File System Names to Back Up

Use this procedure to determine the names of the file systems you want to back up.

1. Display the contents of the `/etc/vfstab` file.

You do not need to be superuser to run this command.

```
% more /etc/vfstab
```

2. Look in the mount point column for the name of the file system you want to back up.

Use this name when you back up the file system.

```
% more /etc/vfstab
```

8.1.0.1 Example—Finding File System Names to Back Up

In the following example, the names of available file systems listed in the `/etc/vfstab` file are displayed.


```
% more /etc/vfstab
```

#device	device	mount	FS	fsck	mount	mount	
#to mount	to fsck	point	type	pass	at boot	options	
#							
#/dev/dsk/c1d0s2	/dev/rdisk/c1d0s2	/usr	ufs	1	yes	-	
f	-	/dev/fd	fd	-	no	-	
/proc	-	/proc	proc	-	no	-	
/dev/dsk/c1t6d0s1	-	-	swap	-	no	-	
/dev/dsk/c1t6d0s0	/dev/rdisk/c1t6d0s0	/	ufs	1	no	-	
/dev/dsk/c1t6d0s3	/dev/rdisk/c1t6d0s3	/cache	ufs	2	yes	-	
swap	-	/tmp	tmpfs	-	yes	-	

▼ How to Determine the Number of Tapes Needed for a Full Backup

Use this procedure to calculate the number of tapes you will need to back up a file system.

1. Become superuser on the cluster node you want to back up.
2. Estimate the size of the backup in bytes.

```
# ufsdump S filesystem
```

S Displays the estimated number of bytes needed to perform the backup.

filesystem Specifies the name of the file system you want to back up.

3. Divide the estimated size by the capacity of the tape to see how many tapes you need.

8.1.0.1 Example—Determining the Number of Tapes Needed

In the following example, the file system size of 905,881,620 bytes will easily fit on a 4 GB tape (905,881,620 ÷ 4,000,000,000).

```
# ufsdump S /global/phys-schost-1
905881620
```

▼ How to Back Up the root (/) File System

Use this procedure to back up the root (/) file system of a cluster node. Be sure the cluster is running problem-free before performing the backup procedure.

1. Become superuser on the cluster node you want to back up.
2. Switch each running data service from the node to be backed up to another node in the cluster.

```
# scswitch -z -D disk-device-group -h node
```

- z Performs the switch.
- D **disk-device-group** Name of the disk device group, which is the same as the diskset or disk group name.
- h **node** Name of the cluster node which serves as the primary of the disk device group.

3. Stop the node.

```
# shutdown -g0 -y
```

4. At the `ok` prompt, reboot in non-cluster mode.

```
ok boot -x
```

5. Back up the root (/) file system.

- If the root disk is not encapsulated, use the following command.

```
# ufsdump 0ucf dump-device /
```

- If the root disk is encapsulated, use the following command.

```
# ufsdump 0ucf dump-device /dev/vx/rdisk/rootvol
```

Refer to the `ufsdump(1M)` man page for more information.

6. Reboot the node in cluster mode.

```
# init 6
```

8.1.0.1 Example—Backing Up the root (/) File System

In the following example, the root (/) file system is backed up onto tape device `/dev/rmt/0`.

```
# ufsdump 0ucf /dev/rmt/0 /
DUMP: Writing 63 Kilobyte records
DUMP: Date of this level 0 dump: Tue Apr 18 18:06:15 2000
DUMP: Date of last level 0 dump: the epoch
DUMP: Dumping /dev/rdisk/c0t0d0s0 (phys-schost-1:/) to /dev/rmt/0
DUMP: Mapping (Pass I) [regular files]
DUMP: Mapping (Pass II) [directories]
DUMP: Estimated 859086 blocks (419.48MB).
DUMP: Dumping (Pass III) [directories]
DUMP: Dumping (Pass IV) [regular files]
DUMP: 859066 blocks (419.47MB) on 1 volume at 2495 KB/sec
DUMP: DUMP IS DONE
DUMP: Level 0 dump on Tue Apr 18 18:06:15 2000
```

▼ How to Perform Online Backups for Mirrors (Solstice DiskSuite)

A mirrored metadevice can be backed up without unmounting it or taking the entire mirror offline. One of the submirrors must be taken offline temporarily, thus losing mirroring, but it can be placed online and resynced as soon as the backup is complete, without halting the system or denying user access to the data. Using mirrors to perform online backups creates a backup that is a “snapshot” of an active file system.

A problem might occur if a program writes data onto the volume immediately before the `lockfs` command is run. To prevent this problem, temporarily stop all the services running on this node. Also, be sure the cluster is running problem-free before performing the backup procedure.

1. Become superuser on the cluster node you want to back up.
2. Use the `metaset(1M)` command to determine which node has the ownership on the backed up volume.

```
# metaset -s setname
```

`-s setname` Specifies the diskset name.

3. Use the `lockfs(1M)` command with the `-w` option to lock the file system from writes.

```
# lockfs -w mount-point
```

Note - You must lock the file system only if a UFS file system resides on the mirror. For example, if the metadevice is set up as a raw device for database management software or some other specific application, it would not be necessary to use the `lockfs` command. You may, however, want to run the appropriate vendor-dependent utility to flush any buffers and lock access.

4. Use the `metastat(1M)` command to determine the names of the submirrors.

```
# metastat -s setname -p
```

`-p` Displays the status in a format similar to the `md.tab` file.

5. Use the `metadetach(1M)` command to take one submirror offline from the mirror.

```
# metadetach -s setname mirror submirror
```

Note - Reads will continue to be made from the other submirrors. However, the offline submirror will be out of sync as soon as the first write is made to the mirror. This inconsistency is corrected when the offline submirror is brought back online. You don't need to run `fsck`.

6. Unlock the file systems and allow writes to continue, using the `lockfs` command with the `-u` option.

```
# lockfs -u mount-point
```

7. Perform a file system check.

```
# fsck /dev/md/diskset/rdisk/submirror
```

8. Back up the offline submirror to tape or another medium.

Use the `ufsdump(1M)` command or whatever other backup utility you normally use.

Note - Use the raw device (`/rdsk`) name for the submirror, rather than the block device (`/dsk`) name.

```
# ufsdump 0ucf dump-device submirror
```

9. Use the `metattach(1M)` command to place the metadvice back online.

```
# metattach -s setname mirror submirror
```

When the metadvice is placed online, it is automatically resynced with the mirror.

10. Use the `metastat` command to verify that the submirror is resyncing.

```
# metastat -s setname mirror
```

8.1.0.1

Example—Performing Online Backups for Mirrors (Solstice DiskSuite)

In the following example, the cluster node `phys-schost-1` is the owner of the metaset `schost-1`, therefore the backup procedure is performed from `phys-schost-1`. The mirror `/dev/md/schost-1/dsk/d0` consists of the submirrors `d10`, `d20`, and `d30`.

```
[Determine the owner of the metaset:]
# metaset -s schost-1
Set name = schost-1, Set number = 1
Host          Owner
phys-schost-1 Yes
...
[Lock the file system from writes:]
# lockfs -w /global/schost-1
[List the submirrors:]
# metastat -s schost-1 -p
schost-1/d0 -m schost-1/d10 schost-1/d20 schost-1/d30 1
schost-1/d10 1 1 d4s0
schost-1/d20 1 1 d6s0
schost-1/d30 1 1 d8s0
[Take a submirror offline:]
# metadetach -s schost-1 d0 d30
[Unlock the file system:]
# lockfs -u /
[Check the file system:]
# fsck /dev/md/schost-1/rdisk/d30
[Copy the submirror to the backup device:]
# ufsdump 0ucf /dev/rmt/0 /dev/md/schost-1/rdisk/d30
DUMP: Writing 63 Kilobyte records
DUMP: Date of this level 0 dump: Tue Apr 25 16:15:51 2000
DUMP: Date of last level 0 dump: the epoch
DUMP: Dumping /dev/md/schost-1/rdisk/d30 to /dev/rdisk/c1t9d0s0.
...
DUMP: DUMP IS DONE
```

(continued)

```
[Bring the submirror back online:]
# metattach -s schost-1 d0 d30
schost-1/d0: submirror schost-1/d30 is attached
[Resync the submirror:]
# metastat -s schost-1 d0
schost-1/d0: Mirror
    Submirror 0: schost-0/d10
        State: Okay
    Submirror 1: schost-0/d20
        State: Okay
    Submirror 2: schost-0/d30
        State: Resyncing
    Resync in progress: 42 % done
    Pass: 1
    Read option: roundrobin (default)
...
```

▼ How to Perform Online Backups for Volumes (VERITAS Volume Manager)

VERITAS Volume Manager identifies a mirrored volume as a plex. A plex can be backed up without unmounting it or taking the entire volume offline. This is done by creating a snapshot copy of the volume and backing up this temporary volume without halting the system or denying user access to the data.

Be sure the cluster is running problem-free before performing the backup procedure.

1. **Log on to any node in the cluster, and become superuser on the current primary node for the disk group on the cluster.**
2. **List the disk group information.**

```
# vxprint -g diskgroup
```

3. Run `scstat(1M)` to see which node has the disk group currently imported, indicating it is the primary node for the disk group.

```
# scstat -D
```

-D	Shows the status for all disk device groups.
----	----------------------------------------------

- 4. Create a snapshot of the volume using the `vxassist(1M)` command.**

```
# vxassist -g diskgroup snapstart volume
```

Note - Creating a snapshot can take a long time depending on the size of your volume.

5. Verify the new volume was created using the `vxprint(1M)` command.

```
# vxprint -g diskgroup
```

When the snapshot is complete, a status of Snapdone displays in the State field for the selected disk group.

6. Stop any data services that are accessing the file system using `scswitch(1M)`.

```
# scswitch -z -g nfs-rg -h ''''
```

Note - Stopping data services is recommended to ensure that the data file system is properly backed up. If no data services are running, you do not need to perform Step 6 on page 143 and Step 8 on page 143.

7. Create a backup volume named `bkup-vol` and attach the snapshot volume to it using the `vxassist` command.

```
# vxassist -g diskgroup snapshot volume bkup-vol
```

8. Restart any data services that were stopped in Step 6 on page 143, using the `scswitch` command.

```
# scswitch -z -g nfs-rg -h nodename
```

9. Verify the volume is now attached to the new volume `bkup-vol` using the `vxprint` command.

```
# vxprint -g diskgroup
```

10. Register the disk group configuration change using the `scconf(1M)` command.

```
# scconf -c -D name=diskgroup, sync
```

11. Check the backup volume using the `fsck` command.

```
# fsck -y /dev/vx/rdisk/diskgroup/bkup-vol
```

12. Perform a backup to copy the volume `bkup-vol` to tape or another medium.

Use the `ufsdump(1M)` command or the backup utility you normally use.

```
# ufsdump 0ucf dump-device /dev/vx/dsk/diskgroup/bkup-vol
```

13. Remove the temporary volume using `vxedit(1M)`.

```
# vxedit -rf rm bkup-vol
```

14. Register the disk group configuration changes using the `scconf` command.

```
# scconf -c -D name=disk-group, sync
```

8.1.0.1 Example—Performing Online Backups for Volumes (VERITAS Volume Manager)

In the following example, the cluster node `phys-schost-2` is the primary owner of the metaset disk group `schost-1`, therefore the backup procedure is performed from `phys-schost-2`. The volume `/vol101` is copied and then associated with a new volume, `bkup-vol`.

```
[Become superuser on the primary node.]
[Identify the current primary node for the disk group:]
# scstat -D
-- Device Group Servers --
      Device Group      Primary      Secondary
      -----
Device group servers:  rmt/1          -
Device group servers:  schost-1      phys-schost-2  phys-schost-1
-- Device Group Status --
      Device Group      Status
      -----
Device group status:   rmt/1          Offline
Device group status:   schost-1      Online
[List the disk group information:]
```

(continued)


```

# vxprint -g schost-1
TY NAME          ASSOC          KSTATE    LENGTH    PLOFFS    STATE    TUTILO    PUTILO
dg schost-1      schost-1      -          -          -          -          -          -

dm schost-101    c1t1d0s2    -          17678493    -          -          -          -
dm schost-102    c1t2d0s2    -          17678493    -          -          -          -
dm schost-103    c2t1d0s2    -          8378640     -          -          -          -
dm schost-104    c2t2d0s2    -          17678493    -          -          -          -
dm schost-105    c1t3d0s2    -          17678493    -          -          -          -
dm schost-106    c2t3d0s2    -          17678493    -          -          -          -

v  vol01         gen          ENABLED    204800     -          ACTIVE    -          -
pl vol01-01      vol01        ENABLED    208331     -          ACTIVE    -          -
sd schost-101-01 vol01-01     ENABLED    104139     0          -          -          -
sd schost-102-01 vol01-01     ENABLED    104139     0          -          -          -
pl vol01-02      vol01        ENABLED    208331     -          ACTIVE    -          -
sd schost-103-01 vol01-02     ENABLED    103680     0          -          -          -
sd schost-104-01 vol01-02     ENABLED    104139     0          -          -          -
pl vol01-03      vol01        ENABLED    LOGONLY    -          ACTIVE    -          -
sd schost-103-02 vol01-03     ENABLED    5          LOG        -          -          -
[Start the snapshot operation:]
# vxassist -g schost-1 snapstart vol01
[Verify the new volume was created:]
# vxprint -g schost-1
TY NAME          ASSOC          KSTATE    LENGTH    PLOFFS    STATE    TUTILO    PUTILO
dg schost-1      schost-1      -          -          -          -          -          -

dm schost-101    c1t1d0s2    -          17678493    -          -          -          -
dm schost-102    c1t2d0s2    -          17678493    -          -          -          -
dm schost-103    c2t1d0s2    -          8378640     -          -          -          -
dm schost-104    c2t2d0s2    -          17678493    -          -          -          -
dm schost-105    c1t3d0s2    -          17678493    -          -          -          -
dm schost-106    c2t3d0s2    -          17678493    -          -          -          -

v  vol01         gen          ENABLED    204800     -          ACTIVE    -          -
pl vol01-01      vol01        ENABLED    208331     -          ACTIVE    -          -
sd schost-101-01 vol01-01     ENABLED    104139     0          -          -          -
sd schost-102-01 vol01-01     ENABLED    104139     0          -          -          -
pl vol01-02      vol01        ENABLED    208331     -          ACTIVE    -          -
sd schost-103-01 vol01-02     ENABLED    103680     0          -          -          -
sd schost-104-01 vol01-02     ENABLED    104139     0          -          -          -
pl vol01-03      vol01        ENABLED    LOGONLY    -          ACTIVE    -          -
sd schost-103-02 vol01-03     ENABLED    5          LOG        -          -          -
pl vol01-04      vol01        ENABLED    208331     -          SNAPDONE  -          -
sd schost-105-01 vol01-04     ENABLED    104139     0          -          -          -
sd schost-106-01 vol01-04     ENABLED    104139     0          -          -          -
[Stop data services, if necessary:]
# scswitch -z -g nfs-rg -h ''''
[Create a copy of the volume:]
# vxassist -g schost-1 snapshot vol01 bkup-vol
[Restart data services, if necessary:]
# scswitch -z -g nfs-rg -h phys-schost-1
[Verify bkup-vol was created:]
# vxprint -g schost-1

```

(continued)

TY NAME	ASSOC	KSTATE	LENGTH	PLOFFS	STATE	TUTIL0	PUTIL0
dg schost-1	schost-1	-	-	-	-	-	-
dm schost-101	c1t1d0s2	-	17678493	-	-	-	-
dm schost-102	c1t2d0s2	-	17678493	-	-	-	-
dm schost-103	c2t1d0s2	-	8378640	-	-	-	-
dm schost-104	c2t2d0s2	-	17678493	-	-	-	-
dm schost-105	c1t3d0s2	-	17678493	-	-	-	-
dm schost-106	c2t3d0s2	-	17678493	-	-	-	-
v bkup-vol	gen	ENABLED	204800	-	ACTIVE	-	-
pl bkup-vol-01	bkup-vol	ENABLED	208331	-	ACTIVE	-	-
sd schost-105-01	bkup-vol-01	ENABLED	104139	0	-	-	-
sd schost-106-01	bkup-vol-01	ENABLED	104139	0	-	-	-
v vol01	gen	ENABLED	204800	-	ACTIVE	-	-
pl vol01-01	vol01	ENABLED	208331	-	ACTIVE	-	-
sd schost-101-01	vol01-01	ENABLED	104139	0	-	-	-
sd schost-102-01	vol01-01	ENABLED	104139	0	-	-	-
pl vol01-02	vol01	ENABLED	208331	-	ACTIVE	-	-
sd schost-103-01	vol01-02	ENABLED	103680	0	-	-	-
sd schost-104-01	vol01-02	ENABLED	104139	0	-	-	-
pl vol01-03	vol01	ENABLED	LOGONLY	-	ACTIVE	-	-
sd schost-103-02	vol01-03	ENABLED	5	LOG	-	-	-

```

[Synchronize the disk group with cluster framework:]
# scconf -c -D name=schost-1,sync
[Check the file systems:]
# fsck -y /dev/vx/rdisk/schost-1/bkup-vol
[Copy bkup-vol to the backup device:]
# ufsdump 0ucf /dev/rmt/0 /dev/vx/rdisk/schost-1/bkup-vol
DUMP: Writing 63 Kilobyte records
DUMP: Date of this level 0 dump: Tue Apr 25 16:15:51 2000
DUMP: Date of last level 0 dump: the epoch
DUMP: Dumping /dev/vx/dsk/schost-2/bkup-vol to /dev/rmt/0.
...
DUMP: DUMP IS DONE
[Remove the bkup-volume:]
# vxedit -rf rm bkup-vol
[Synchronize the disk group:]
# scconf -c -D name=schost-1,sync

```

8.2 Restoring Cluster Files Overview

The `ufsrestore` command copies files to disk, relative to the current working directory, from backups created using the `ufsdump` command. You can use

`ufsrestore` to reload an entire file system hierarchy from a level 0 dump and incremental dumps that follow it, or to restore one or more single files from any dump tape. If `ufsrestore` is run as superuser, files are restored with their original owner, last modification time, and mode (permissions).

Before you start to restore files or file systems, you need to know:

- which tapes you need
- the raw device name on which you want to restore the file system
- the type of tape drive you will use
- the device name (local or remote) for the tape drive
- the partition scheme of any failed disk, because the partitions and file systems must be exactly duplicated on the replacement disk

8.3 Restoring Cluster Files

TABLE 8-2 Task Map: Restoring Cluster Files

Task	For Instructions, Go To...
For Solstice DiskSuite, restore files interactively following Solaris restore procedures.	“How to Restore Individual Files Interactively (Solstice DiskSuite)” on page 148
For Solstice DiskSuite, restore the root (/) file system.	“How to Restore the root (/) File System (Solstice DiskSuite)” on page 148 “How to Restore a root (/) File System That Was on a Metadevice (Solstice DiskSuite)” on page 151
For VERITAS Volume Manager, restore a non-encapsulated root (/) file system.	“How to Restore a Non-Encapsulated root (/) File System (VERITAS Volume Manager)” on page 156
For VERITAS Volume Manager, restore an encapsulated root (/) file system.	“How to Restore an Encapsulated root (/) File System (VERITAS Volume Manager)” on page 159

▼ How to Restore Individual Files Interactively (Solstice DiskSuite)

Use this procedure to restore one or more individual files. Be sure the cluster is running problem-free before performing the restore procedure.

1. **Become superuser on the cluster node you want to restore.**
2. **Stop all the data services that are using the files to be restored.**

```
# scswitch -z -g rgname -h ""
```

3. **Restore the files using the `ufsrestore` command.**

▼ How to Restore the root (/) File System (Solstice DiskSuite)

Use this procedure to restore the root (/) file systems to a new disk, such as after replacing a bad root disk. The node being restored should not be booted. Be sure the cluster is running problem-free before performing the restore procedure.

Note - Since you must partition the new disk using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and recreate file systems as appropriate.

1. **Become superuser on a cluster node *other than* the node you want to restore.**
2. **Remove the hostname of the node being restored from all metaset(s) using the `metaset(1M)` command. Run this command from a node in the metaset other than the node you are removing.**

```
# metaset -s setname -f -d -h node
```

- | | |
|-------------------|------------------------------------------------------------|
| -s setname | Specifies the diskset name. |
| -f | Force. |
| -d | Deletes from the diskset. |
| -h node | Specifies the name of the node to delete from the diskset. |
3. **Replace the failed disk on the node on which the root (/) file system will be restored.**

Refer to disk replacement procedures in the documentation that came with your server.

4. Boot the node being restored.

- If using the Solaris CD-ROM, run the following command:

```
ok boot cdrom -s
```

- If using a JumpStart™ server, run the following command:

```
ok boot net -s
```

5. Create all the partitions and swap on the root disk using the `format(1M)` command.

Recreate the original partitioning scheme that was on the failed disk.

6. Create the root (`/`) file system and other file systems as appropriate, using the `newfs(1M)` command.

Recreate the original file systems that were on the failed disk.

Note - Be sure to create the `/global/.devices/node@nodeid` file system.

7. Mount the root (`/`) file system on a temporary mount point.

```
# mount device temp-mount-point
```

8. Use the following commands to restore the root (`/`) file system.

```
# cd temp-mount-point
# ufsrestore rvf dump-device
# rm restoresymtable
# cd /
# umount temp-mount-point
# fsck raw-disk-device
```

The file system is now restored.

9. Install a new boot block on the new disk.

```
# /usr/sbin/installboot /usr/platform/`uname -i`/lib/fs/ufs/bootblk raw-disk-device
```

10. Reboot the node in single-user mode.

```
# reboot -- "-s"
```

11. Replace the disk ID using the `scdidadm` command.

```
# sddidadm -R rootdisk
```

12. Use the `metadb(1M)` command to recreate the state database replicas.

```
# metadb -c copies -af raw-disk-device
```

-c *copies* Specifies the number of replicas to create.

-f *raw-disk-device* Raw disk device on which to create replicas.

-a Adds replicas.

13. Reboot the node in cluster mode.

a. Start the reboot.

```
# reboot
```

During this boot you might see an error or warning message, ending with the following instruction:

```
Type control-d to proceed with normal startup,  
(or give root password for system maintenance):
```

b. Press CTRL-d to boot into multiuser mode.

14. From a cluster node other than the restored node, use the `metaset(1M)` command to add the restored node to all metaset.

```
phys-schost-2# metaset -s setname -a -h node
```

-a	Creates and adds the host to the diskset.
----	-------------------------------------------

The node is rebooted into cluster mode. The cluster is ready to use.

8.3.0.1 Example—Restoring the root (/) File System (Solstice DiskSuite)

The following example shows the root (/) file system restored to the node `phys-schost-1` from the tape device `/dev/rmt/0`. The `metaset` command is run from another node in the cluster, `phys-schost-2`, to remove and later add back node `phys-schost-1` to the diskset `schost-1`. All other commands are run from `phys-schost-1`. A new boot block is created on `/dev/rdsd/c0t0d0s0`, and three state database replicas are recreated on `/dev/rdsd/c0t0d0s4`.

```
[Become superuser on a cluster node other than the node to be restored.]
[Remove the node from the metaset:]
phys-schost-2# metaset -s schost-1 -f -d -h phys-schost-1
[Replace the failed disk and boot the node:]
ok boot cdrom -s
[Use format and newfs to recreate partitions and file systems.]
[Mount the root file system on a temporary mount point:]
# mount /dev/dsk/c0t0d0s0 /a
[Restore the root file system:]
# cd /a
# ufsrestore rvf /dev/rmt/0
# rm restoresymtable
# cd /
# umount /a
# fsck /dev/rdisk/c0t0d0s0
[Install a new boot block:]
# /usr/sbin/installboot /usr/platform/'uname \
-i' /lib/fs/ufs/bootblk /dev/rdisk/c0t0d0s0
[Reboot in single-user mode:]
# reboot -- "-s"
[Replace the disk ID:]
# sccdidadm -R /dev/dsk/c0t0d0
[Recreate state database replicas:]
# metadb -c 3 -af /dev/rdisk/c0t0d0s4
# reboot
Press CTL-d to boot into multiuser mode.
[Add the node back to the metaset:]
phys-schost-2# metaset -s schost-1 -a -h phys-schost-1
```

▼ How to Restore a root (/) File System That Was on a Metadevice (Solstice DiskSuite)

Use this procedure to restore a root (/) file system that was on a metadvice when the backups were performed. Perform this procedure under circumstances such as

when a root disk is corrupted and replaced with a new disk. The node being restored should not be booted. Be sure the cluster is running problem-free before performing the restore procedure.

Note - Since you must partition the new disk using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and recreate file systems as appropriate.

1. **Become superuser on a cluster node with access to the metaset, *other than* the node you want to restore.**
2. **Use the `metaset(1M)` command to remove the hostname of the node being restored from all metaset.**

```
# metaset -s setname -f -d -h node
```

<code>-s setname</code>	Specifies the metaset name.
<code>-f</code>	Force.
<code>-d</code>	Deletes from the metaset.
<code>-h node</code>	Specifies the name of the node to delete from the metaset.

3. **Replace the failed disk on the node on which the root (/) file system will be restored.**

Refer to disk replacement procedures in the documentation that came with your server.

4. **Boot the node being restored.**

- If using the Solaris CD-ROM, run the following command:

```
ok boot cdrom -s
```

- If using a JumpStart server, run the following command:

```
ok boot net -s
```

5. **Create all the partitions and swap on the root disk using the `format(1M)` command.**

Recreate the original partitioning scheme that was on the failed disk.

6. **Create the root (/) file system and other file systems as appropriate, using the `newfs(1M)` command**

Recreate the original file systems that were on the failed disk.

Note - Be sure to create the `/global/.devices/node@nodeid` file system.

7. Mount the root (/) file system on a temporary mount point.

```
# mount device temp-mount-point
```

8. Use the following commands to restore the root (/) file system.

```
# cd temp-mount-point  
# ufsrestore rvf dump-device  
# rm restoresymtable
```

9. Install a new boot block on the new disk.

```
# /usr/sbin/installboot /usr/platform/`uname -i`/lib/fs/ufs/bootblk raw-disk-device
```

10. Remove the lines in the `/temp-mount-point/etc/system` file for MDD root information.

```
* Begin MDD root info (do not edit)  
forceload: misc/md_trans  
forceload: misc/md_raid  
forceload: misc/md_mirror  
forceload: misc/md_hotspares  
forceload: misc/md_stripe  
forceload: drv/pcipsy  
forceload: drv/glm  
forceload: drv/sd  
rootdev:/pseudo/md@0:0,10,blk  
* End MDD root info (do not edit)
```

11. Edit the `/temp-mount-point/etc/vfstab` file to change the root entry from a metadvice to a corresponding normal slice for each file system on the root disk that is part of the metadvice.

```
Example:
Change from---
/dev/md/dsk/d10  /dev/md/rdisk/d10  /      ufs  1    no    -

Change to---
/dev/dsk/c0t0d0s0 /dev/rdisk/c0t0d0s0 /usr  ufs  1    no    -
```

12. Unmount the temporary file system, and check the raw disk device.

```
# cd /
# umount  temp-mount-point
# fsck raw-disk-device
```

13. Reboot the node in single-user mode.

```
# reboot -- "-s"
```

14. Replace the disk ID using the `sddidadm` command.

```
# sddidadm -R rootdisk
```

15. Use the `metadb(1M)` command to recreate the state database replicas.

```
# metadb -c copies -af raw-disk-device
```

-c `copies` Specifies the number of replicas to create.

-af `raw-disk-device` Creates initial state database replicas on the named raw disk device.

16. Reboot the node in cluster mode.

a. Start the reboot.

```
# reboot
```

During this boot you will see error or warning messages, ending with the following instruction:

```
Type control-d to proceed with normal startup,  
(or give root password for system maintenance):
```

b. Press CTRL-d to boot into multiuser mode.

17. From a cluster node other than the restored node, use the `metaset(1M)` command to add the restored node to all metaset.

```
phys-schost-2# metaset -s setname -a -h node
```

-a Adds (creates) the metaset.

Set up the metadevice/mirror for root (/) according to the Solstice DiskSuite documentation.

The node is rebooted into cluster mode. The cluster is ready to use.

8.3.0.1 Example—Restoring a root (/) File System That Was on a Metadevice (Solstice DiskSuite)

The following example shows the root (/) file system restored to the node `phys-schost-1` from the tape device `/dev/rmt/0`. The `metaset` command is run from another node in the cluster, `phys-schost-2`, to remove and later add back node `phys-schost-1` to the metaset `schost-1`. All other commands are run from `phys-schost-1`. A new boot block is created on `/dev/rdsk/c0t0d0s0`, and three state database replicas are recreated on `/dev/rdsk/c0t0d0s4`.

```
[Become superuser on a cluster node with access to the metaset, other than the node to be restored.]  
[Remove the node from the metaset:]  
phys-schost-2# metaset -s schost-1 -f -d -h phys-schost-1  
[Replace the failed disk and boot the node:]  
ok boot cdrom -s  
[Use format and newfs to recreate partitions and file systems.]  
[Mount the root file system on a temporary mount point:]  
# mount /dev/dsk/c0t0d0s0 /a  
[Restore the root file system:]  
# cd /a  
# ufsrestore rvf /dev/rmt/0  
# rm restoresymtable  
[Install a new boot block:]  
# /usr/sbin/installboot /usr/platform/`uname \n  
-i`/lib/fs/ufs/bootblk /dev/rdsk/c0t0d0s0
```

(continued)

```

[Remove the lines in /temp-mount-point/etc/
system file for MDD root information:]
* Begin MDD root info (do not edit)
forceload: misc/md_trans
forceload: misc/md_raid
forceload: misc/md_mirror
forceload: misc/md_hotspares
forceload: misc/md_stripe
forceload: drv/pcipsy
forceload: drv/glm
forceload: drv/sd
rootdev:/pseudo/md@0:0,10,blk
* End MDD root info (do not edit)
[Edit the /temp-mount-point/etc/vfstab file]
Example:
Change from---
/dev/md/dsk/d10    /dev/md/rdisk/d10    /      ufs    1      no      -

Change to---
/dev/dsk/c0t0d0s0 /dev/rdisk/c0t0d0s0 /usr    ufs    1      no      -
[Unmount the temporary file system and check the raw disk device:]
# cd /
# umount /a
# fsck /dev/rdisk/c0t0d0s0
[Reboot in single-user mode:]
# reboot -- "-s"
[Replace the disk ID:]
# sddidadm -R /dev/dsk/c0t0d0
[Recreate state database replicas:]
# metadb -c 3 -af /dev/rdisk/c0t0d0s4
# reboot
Type CTRL-d to boot into multiuser mode.
[Add the node back to the metaset:]
phys-schost-2# metaset -s schost-1 -a -h phys-schost-1

```

▼ How to Restore a Non-Encapsulated root (/) File System (VERITAS Volume Manager)

Use this procedure to restore a non-encapsulated root (/) file system to a node. The node being restored should not be booted. Be sure the cluster is running problem-free before performing the restore procedure.

Note - Since you must partition the new disk using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and recreate file systems as appropriate.

1. Replace the failed disk on the node where the root file system will be restored.

Refer to disk replacement procedures in the documentation that came with your server.

2. Boot the node being restored.

- If using the Solaris CD-ROM, run the following command:

```
ok boot cdrom -s
```

- If using a JumpStart server, run the following command:

```
ok boot net -s
```

3. Create all the partitions and swap on the root disk using the `format(1M)` command.

Recreate the original partitioning scheme that was on the failed disk.

4. Create the root (/) file system and other file systems as appropriate, using the `newfs(1M)` command.

Recreate the original file systems that were on the failed disk.

Note - Be sure to create the `/global/.devices/node@nodeid` file system.

5. Mount the root (/) file system on a temporary mount point.

```
# mount device temp-mount-point
```

6. Restore the root (/) file system from backup, and unmount and check the file system.

```
# cd temp-mount-point
# ufsrestore rvf dump-device
# rm restoresymtable
# cd /
# umount temp-mount-point
# fsck raw-disk-device
```

The file system is now restored.

7. Install a new boot block on the new disk.

```
# /usr/sbin/installboot /usr/platform/`uname -i`/lib/fs/ufs/bootblk raw-disk-device
```

8. Reboot the node into single-user mode.

a. Start the reboot.

```
# reboot -- "-s"
```

During this boot you will see error or warning messages, ending with the following instruction:

```
Type control-d to proceed with normal startup,  
(or give root password for system maintenance):
```

b. Enter the root password.

9. Determine if the root disk group is on a single slice on the root disk.

- If yes, create and set up the root disk group:

```
# vxctl init  
# vxvg init rootdg  
# vxctl add disk diskslice type=simple  
# vxdisk -f init diskslice type=simple  
# vxvg adddisk diskslice  
# vxctl enable
```

- If no, proceed to Step 10 on page 158.

10. Update the disk ID using the `sddidadm` command.

```
# sddidadm -R /dev/rdisk/diskdevice
```

11. Press CTRL-d to resume in multiuser mode.

The node reboots into cluster mode. The cluster is ready to use.

8.3.0.1

Example—Restoring a Non-Encapsulated root (/) File System (VERITAS Volume Manager)

The following example shows a non-encapsulated root (/) file system restored to the node `phys-schost-1` from the tape device `/dev/rmt/0`.

```
[Replace the failed disk and boot the node:]
ok boot cdrom -s
[Use format and newfs to create partitions and file systems]
[Mount the root file system on a temporary mount point:]
# mount /dev/dsk/c0t0d0s0 /a
[Restore the root file system:]
# cd /a
# ufsrestore rvf /dev/rmt/0
# rm restoresymtable
# cd /
# umount /a
# fsck /dev/rdsk/c0t0d0s0
[Install a new boot block:]
# /usr/sbin/installboot /usr/platform/`uname \
-i`/lib/fs/ufs/bootblk /dev/rdsk/c0t0d0s0
[Reboot in single-user mode:]
# reboot -- "-s"
[If the root disk group is on a single slice on the root disk, create the new root disk group:]
# vxdctl init
# vxdg init rootdg
# vxdctl add disk c0t0d0s4 type=simple
# vxdisk -f init c0t0d0s4 type=simple
# vxdg adddisk c0t0d0s4
# vxdctl enable
[Update the disk ID:]
# scdidadm -R /dev/rdsk/c0t0d0
[Press CTRL-d to resume in multiuser mode]
```

▼ How to Restore an Encapsulated root (/) File System (VERITAS Volume Manager)

Use this procedure to restore an encapsulated root (/) file system to a node. The node being restored should not be booted. Be sure the cluster is running problem-free before performing the restore procedure.

Note - Since you must partition the new disk using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and recreate file systems as appropriate.

1. **Replace the failed disk on the node where the root file system will be restored.**
Refer to disk replacement procedures in the documentation that came with your server.

2. Boot the node being restored.

- If using the Solaris CD-ROM, run the following command:

```
ok boot cdrom -s
```

- If using a JumpStart server, run the following command:

```
ok boot net -s
```

3. Create all the partitions and swap on the root disk using the `format(1M)` command.

Recreate the original partitioning scheme that was on the failed disk.

4. Create the root (/) file system and other file systems as appropriate, using the `newfs(1M)` command.

Recreate the original file systems that were on the failed disk.

Note - Be sure to create the `/global/.devices/node@nodeid` file system.

5. Mount the root (/) file system on a temporary mount point.

```
# mount device temp-mount-point
```

6. Restore the root (/) file system from backup.

```
# cd temp-mount-point
# ufsrestore rvf dump-device
# rm restoresymtable
```

7. Create an empty `install-db` file.

This puts the node in VxVM install mode at the next reboot.

```
# touch /temp-mount-point/etc/vx/reconfig.d/state.d/install-db
```

8. Remove or comment out the following entries from the `/temp-mount-point/etc/system` file.


```
* rootdev:/pseudo/vxio@0:0
* set vxio:vol_rootdev_is_volume=1
```

- 9. Edit the `/temp-mount-point/etc/vfstab` file and replace all VxVM mount points with the standard disk devices for the root disk, such as `/dev/dsk/c0t0d0s0`.**

```
Example:
Change from---
/dev/vx/dsk/rootdg/rootvol /dev/md/rdisk/rootdg/rootvol /      ufs    1      no -

Change to---
/dev/dsk/c0t0d0s0 /dev/rdisk/c0t0d0s0 / ufs    1      no    -
```

- 10. Unmount the temporary file system and check the file system.**

```
# cd /
# umount temp-mount-point
# fsck raw-disk-device
```

- 11. Install the boot block on the new disk.**

```
# /usr/sbin/installboot /usr/platform/`uname -i`/lib/fs/ufs/bootblk raw-disk-device
```

- 12. Reboot the node in single-user mode.**

```
# reboot -- "-s"
```

- 13. Update the disk ID using `scdidadm(1M)`.**

```
# scdidadm -R /dev/rdisk/c0t0d0
```

- 14. Run `vxinstall`.**

```
# vxinstall
```

Choose to encapsulate the disk and reboot.

15. If there is a conflict in minor number with any other system, unmount the global devices and remirror the disk group.

- Unmount the global devices file system on the cluster node.

```
# umount /global/.devices/node@nodeid
```

- Remirror the rootdg disk group on the cluster node.

```
# vxvg remirror rootdg 100
```

16.

```
# shutdown -g0 -i 6 -y
```

Shut down and reboot the node in cluster mode.

8.3.0.1 Example—Restoring an Encapsulated root (/) File System (VERITAS Volume Manager)

The following example shows an encapsulated root (/) file system restored to the node `phys-schost-1` from the tape device `/dev/rmt/0`.

```
[Replace the failed disk and boot the node:]
ok boot cdrom -s
[Use format and newfs to create partitions and file systems]
[Mount the root file system on a temporary mount point:]
# mount /dev/dsk/c0t0d0s0 /a
[Restore the root file system:]
# cd /a
# ufsrestore rvf /dev/rmt/0
# rm restoresymtable
[Create an empty install-db file:]
# touch /a/etc/vx/reconfig.d/state.d/install-db
[Edit /etc/
system on the temporary file system and remove or comment out the following entries:]
# rootdev:/pseudo/vxio@0:0
# set vxio:vol_rootdev_is_volume=1
[Edit /etc/vfstab on the temporary file system:]
Example:
Change from---
/dev/vx/dsk/rootdg/rootvol /dev/md/rdsk/rootdg/rootvol / ufs 1 no-

Change to---
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0 / ufs 1 no -
[Unmount the temporary file system, then check the file system:]
```

(continued)

```
# cd /
# umount /a
# fsck /dev/rdisk/c0t0d0s0
[Install a new boot block:]
# /usr/sbin/installboot /usr/platform/`uname \
-i`/lib/fs/ufs/bootblk /dev/rdisk/c0t0d0s0
[Reboot in single-user mode:]
# reboot -- "-s"
[Update the disk ID:]
# sddidadm -R /dev/rdisk/c0t0d0
[Run vxinstall:]
# vxinstall
Choose to encapsulate the root disk.
[If there is a conflict in minor number, remminor the rootdg disk group:]
# umount /global/.devices/node@nodeid
# vxdg remminor rootdg 100
# shutdown -g0 -i 6 -y
```