

Oracle® Solaris Cluster System Administration Guide

Copyright © 2000, 2010, Oracle and/or its affiliates. All rights reserved.

This software and related documentation are provided under a license agreement containing restrictions on use and disclosure and are protected by intellectual property laws. Except as expressly permitted in your license agreement or allowed by law, you may not use, copy, reproduce, translate, broadcast, modify, license, transmit, distribute, exhibit, perform, publish, or display any part, in any form, or by any means. Reverse engineering, disassembly, or decompilation of this software, unless required by law for interoperability, is prohibited.

The information contained herein is subject to change without notice and is not warranted to be error-free. If you find any errors, please report them to us in writing.

If this is software or related software documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, the following notice is applicable:

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation shall be subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License (December 2007). Oracle America, Inc., 500 Oracle Parkway, Redwood City, CA 94065.

This software or hardware is developed for general use in a variety of information management applications. It is not developed or intended for use in any inherently dangerous applications, including applications which may create a risk of personal injury. If you use this software or hardware in dangerous applications, then you shall be responsible to take all appropriate fail-safe, backup, redundancy, and other measures to ensure its safe use. Oracle Corporation and its affiliates disclaim any liability for any damages caused by use of this software or hardware in dangerous applications.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. UNIX is a registered trademark licensed through X/Open Company, Ltd.

This software or hardware and documentation may provide access to or information on content, products, and services from third parties. Oracle Corporation and its affiliates are not responsible for and expressly disclaim all warranties of any kind with respect to third-party content, products, and services. Oracle Corporation and its affiliates will not be responsible for any loss, costs, or damages incurred due to your access to or use of third-party content, products, or services.

Contents

Preface	11
1 Introduction to Administering Oracle Solaris Cluster	17
Overview of Administering Oracle Solaris Cluster	17
Working With a Zone Cluster	18
Oracle Solaris OS Feature Restrictions	19
Administration Tools	20
Graphical User Interface	20
Command-Line Interface	20
Preparing to Administer the Cluster	22
Documenting an Oracle Solaris Cluster Hardware Configuration	22
Using an Administrative Console	22
Backing Up the Cluster	23
Beginning to Administer the Cluster	23
▼ How to Log Into the Cluster Remotely	25
▼ How to Connect Securely to Cluster Consoles	26
▼ How to Access the Cluster Configuration Utilities	27
▼ How to Display Oracle Solaris Cluster Patch Information	27
▼ How to Display Oracle Solaris Cluster Release and Version Information	28
▼ How to Display Configured Resource Types, Resource Groups, and Resources	30
▼ How to Check the Status of Cluster Components	31
▼ How to Check the Status of the Public Network	34
▼ How to View the Cluster Configuration	35
▼ How to Validate a Basic Cluster Configuration	43
▼ How to Check the Global Mount Points	45
▼ How to View the Contents of Oracle Solaris Cluster Command Logs	47

2	Oracle Solaris Cluster and RBAC	51
	Setting Up and Using RBAC With Oracle Solaris Cluster	51
	Oracle Solaris Cluster RBAC Rights Profiles	52
	Creating and Assigning an RBAC Role With an Oracle Solaris Cluster Management Rights Profile	53
	▼ How to Create a Role by Using the Administrative Roles Tool	53
	▼ How to Create a Role From the Command Line	55
	Modifying a User's RBAC Properties	57
	▼ How to Modify a User's RBAC Properties by Using the User Accounts Tool	57
	▼ How to Modify a User's RBAC Properties From the Command Line	58
3	Shutting Down and Booting a Cluster	59
	Overview of Shutting Down and Booting a Cluster	59
	▼ How to Shut Down a Cluster	60
	▼ How to Boot a Cluster	63
	▼ How to Reboot a Cluster	65
	Shutting Down and Booting a Single Node in a Cluster	68
	▼ How to Shut Down a Node	69
	▼ How to Boot a Node	72
	▼ How to Reboot a Node	74
	▼ How to Boot a Node in Noncluster Mode	78
	Repairing a Full /var File System	81
	▼ How to Repair a Full /var File System	81
4	Data Replication Approaches	83
	Understanding Data Replication	84
	Supported Data Replication Methods	84
	Using Storage-Based Data Replication Within a Cluster	85
	Requirements and Restrictions When Using Storage-Based Data Replication Within a Cluster	87
	Manual Recovery Concerns When Using Storage-Based Data Replication Within a Cluster	88
	Best Practices When Using Storage-Based Data Replication	89

5	Administering Global Devices, Disk-Path Monitoring, and Cluster File Systems	91
	Overview of Administering Global Devices and the Global Namespace	91
	Global Device Permissions for Solaris Volume Manager	92
	Dynamic Reconfiguration With Global Devices	92
	Veritas Volume Manager Administration Considerations	93
	Administering Storage-Based Replicated Devices	94
	Administering Hitachi TrueCopy Replicated Devices	95
	Administering EMC Symmetrix Remote Data Facility Replicated Devices	106
	Overview of Administering Cluster File Systems	117
	Cluster File System Restrictions	118
	Guidelines to Support VxFS	118
	Administering Device Groups	119
	▼ How to Update the Global-Devices Namespace	122
	▼ How to Change the Size of a <code>lofi</code> Device That Is Used for the Global-Devices Namespace	123
	Migrating the Global-Devices Namespace	124
	▼ How to Migrate the Global-Devices Namespace From a Dedicated Partition to a <code>lofi</code> Device	124
	▼ How to Migrate the Global-Devices Namespace From a <code>lofi</code> Device to a Dedicated Partition	125
	Adding and Registering Device Groups	127
	▼ How to Add and Register a Device Group (Solaris Volume Manager)	127
	▼ How to Add and Register a Device Group (Raw-Disk)	129
	▼ How to Add and Register a Replicated Device Group (ZFS)	130
	▼ How to Create a New Disk Group When Initializing Disks (Veritas Volume Manager)	131
	Maintaining Device Groups	132
	How to Remove and Unregister a Device Group (Solaris Volume Manager)	132
	▼ How to Remove a Node From All Device Groups	132
	▼ How to Remove a Node From a Device Group (Solaris Volume Manager)	133
	▼ How to Create a New Disk Group When Encapsulating Disks (Veritas Volume Manager)	135
	▼ How to Add a New Volume to an Existing Device Group (Veritas Volume Manager)	136
	▼ How to Convert an Existing Disk Group to a Device Group (Veritas Volume Manager)	137
	▼ How to Assign a New Minor Number to a Device Group (Veritas Volume Manager)	138
	▼ How to Register a Disk Group as a Device Group (Veritas Volume Manager)	139
	▼ How to Register Disk Group Configuration Changes (Veritas Volume Manager)	141
	▼ How to Convert a Local Disk Group to a Device Group (VxVM)	142

▼ How to Convert a Device Group to a Local Disk Group (VxVM)	143
▼ How to Remove a Volume From a Device Group (Veritas Volume Manager)	144
▼ How to Remove and Unregister a Device Group (Veritas Volume Manager)	145
▼ How to Add a Node to a Device Group (Veritas Volume Manager)	146
▼ How to Remove a Node From a Device Group (Veritas Volume Manager)	148
▼ How to Remove a Node From a Raw-Disk Device Group	149
▼ How to Change Device Group Properties	151
▼ How to Set the Desired Number of Secondaries for a Device Group	153
▼ How to List a Device Group Configuration	156
▼ How to Switch the Primary for a Device Group	157
▼ How to Put a Device Group in Maintenance State	158
Administering the SCSI Protocol Settings for Storage Devices	160
▼ How to Display the Default Global SCSI Protocol Settings for All Storage Devices	160
▼ How to Display the SCSI Protocol of a Single Storage Device	161
▼ How to Change the Default Global Fencing Protocol Settings for All Storage Devices	162
▼ How to Change the Fencing Protocol for a Single Storage Device	163
Administering Cluster File Systems	165
▼ How to Add a Cluster File System	165
▼ How to Remove a Cluster File System	168
▼ How to Check Global Mounts in a Cluster	170
Administering Disk-Path Monitoring	171
▼ How to Monitor a Disk Path	172
▼ How to Unmonitor a Disk Path	173
▼ How to Print Failed Disk Paths	174
▼ How to Resolve a Disk-Path Status Error	174
▼ How to Monitor Disk Paths From a File	175
▼ How to Enable the Automatic Rebooting of a Node When All Monitored Shared-Disk Paths Fail	177
▼ How to Disable the Automatic Rebooting of a Node When All Monitored Shared-Disk Paths Fail	178
6 Administering Quorum	179
Administering Quorum Devices	179
Dynamic Reconfiguration With Quorum Devices	181
Adding a Quorum Device	181
Removing or Replacing a Quorum Device	192

Maintaining Quorum Devices	196
Administering Oracle Solaris Cluster Quorum Servers	203
Overview of the Quorum Server Configuration File	203
Starting and Stopping the Quorum Server Software	204
▼ How to Start a Quorum Server	205
▼ How to Stop a Quorum Server	205
Displaying Information About the Quorum Server	206
Cleaning Up Stale Quorum Server Cluster Information	207
7 Administering Cluster Interconnects and Public Networks	211
Administering the Cluster Interconnects	211
Dynamic Reconfiguration With Cluster Interconnects	212
▼ How to Check the Status of the Cluster Interconnect	213
▼ How to Add Cluster Transport Cables, Transport Adapters, or Transport Switches	214
▼ How to Remove Cluster Transport Cables, Transport Adapters, and Transport Switches	216
▼ How to Enable a Cluster Transport Cable	219
▼ How to Disable a Cluster Transport Cable	220
▼ How to Determine a Transport Adapter's Instance Number	222
▼ How to Change the Private Network Address or Address Range of an Existing Cluster ..	223
Administering the Public Network	226
How to Administer IP Network Multipathing Groups in a Cluster	226
Dynamic Reconfiguration With Public Network Interfaces	227
8 Adding and Removing a Node	229
Adding a Node to a Cluster	229
▼ How to Add a Node to the Authorized Node List	230
Creating a Non-Voting Node (Zone) in a Global Cluster	232
Removing a Node From a Cluster	235
▼ How to Remove a Node From a Zone Cluster	236
▼ How to Remove a Node From the Cluster Software Configuration	236
▼ How to Remove a Non-Voting Node (Zone) From a Global Cluster	240
▼ How to Remove Connectivity Between an Array and a Single Node, in a Cluster With Greater Than Two-Node Connectivity	240
▼ How to Correct Error Messages	242

9	Administering the Cluster	245
	Overview of Administering the Cluster	245
	▼ How to Change the Cluster Name	246
	▼ How to Map Node ID to Node Name	247
	▼ How to Work With New Cluster Node Authentication	248
	▼ How to Reset the Time of Day in a Cluster	250
	▼ SPARC: How to Display the OpenBoot PROM (OBP) on a Node	252
	▼ How to Change the Node Private Hostname	253
	▼ How to Add a Private Hostname for a Non-Voting Node on a Global Cluster	255
	▼ How to Change the Private Hostname on a Non-Voting Node on a Global Cluster	256
	▼ How to Delete the Private Hostname for a Non-Voting Node on a Global Cluster	257
	▼ How to Rename a Node	258
	▼ How to Change the Logical Hostnames Used by Existing Oracle Solaris Cluster Logical Hostname Resources	259
	▼ How to Put a Node Into Maintenance State	260
	▼ How to Bring a Node Out of Maintenance State	261
	Configuring Load Limits	264
	Performing Zone Cluster Administrative Tasks	266
	▼ How to Remove a Zone Cluster	267
	▼ How to Remove a File System From a Zone Cluster	268
	▼ How to Remove a Storage Device From a Zone Cluster	270
	▼ How to Uninstall Oracle Solaris Cluster Software From a Cluster Node	271
	Troubleshooting a Node Uninstallation	274
	Creating, Setting Up, and Managing the Oracle Solaris Cluster SNMP Event MIB	275
	Troubleshooting	281
	Running an Application Outside the Global Cluster	281
	Restoring a Corrupted Diskset	283
10	Configuring Control of CPU Usage	287
	Introduction to CPU Control	287
	Choosing a Scenario	287
	Fair Share Scheduler	288
	Configuring CPU Control	289
	▼ How to Control CPU Usage in the Voting Node on a Global Cluster	289
	▼ How to Control CPU Usage in a Global-Cluster Non-Voting Node With the Default Processor Set	291

▼ How to Control CPU Usage in a Global-Cluster Non-Voting Node With a Dedicated Processor Set	293
11 Patching Oracle Solaris Cluster Software and Firmware	297
Overview of Patching Oracle Solaris Cluster	297
Oracle Solaris Cluster Patch Tips	298
Patching Oracle Solaris Cluster Software	299
▼ How to Apply a Rebooting Patch (Node)	299
▼ How to Apply a Rebooting Patch (Cluster)	303
▼ How to Apply a Nonrebooting Oracle Solaris Cluster Patch	307
▼ How to Apply Patches in Single-User Mode with Failover Nodes	308
Changing an Oracle Solaris Cluster Patch	311
12 Backing Up and Restoring a Cluster	313
Backing Up a Cluster	313
▼ How to Find File System Names to Back Up	314
▼ How to Determine the Number of Tapes Needed for a Full Backup	314
▼ How to Back Up the Root (/) File System	315
▼ How to Perform Online Backups for Mirrors (Solaris Volume Manager)	318
▼ How to Perform Online Backups for Volumes (Veritas Volume Manager)	320
▼ How to Back Up the Cluster Configuration	324
Restoring Cluster Files	325
▼ How to Restore Individual Files Interactively (Solaris Volume Manager)	325
▼ How to Restore the Root (/) File System (Solaris Volume Manager)	326
▼ How to Restore a Root (/) File System That Was on a Solaris Volume Manager Volume	328
▼ How to Restore a Nonencapsulated Root (/) File System (Veritas Volume Manager)	333
▼ How to Restore an Encapsulated Root (/) File System (Veritas Volume Manager)	335
13 Administering Oracle Solaris Cluster With the Graphical User Interfaces	339
Overview of Oracle Solaris Cluster Manager	339
SPARC: Overview of Sun Management Center	340
Configuring Oracle Solaris Cluster Manager	341
Setting up RBAC Roles	341
▼ How to Use the Common Agent Container to Change the Port Numbers for Services or Management Agents	341

▼ How to Change the Server Address for Oracle Solaris Cluster Manager	342
▼ How to Regenerate Common Agent Container Security Keys	343
Starting the Oracle Solaris Cluster Manager Software	344
▼ How to Start Oracle Solaris Cluster Manager	344
A Example	347
Configuring Host-Based Data Replication With Sun StorageTek Availability Suite Software	347
Understanding Sun StorageTek Availability Suite Software in a Cluster	348
Guidelines for Configuring Host-Based Data Replication Between Clusters	351
Task Map: Example of a Data Replication Configuration	356
Connecting and Installing the Clusters	357
Example of How to Configure Device Groups and Resource Groups	359
Example of How to Enable Data Replication	373
Example of How to Perform Data Replication	376
Example of How to Manage a Failover	381
Index	385

Preface

The *Oracle Solaris Cluster System Administration Guide* provides procedures for administering an Oracle Solaris Cluster configuration on both SPARC and x86 based systems.

Note – This Oracle Solaris Cluster release supports systems that use the SPARC and x86 families of processor architectures: UltraSPARC, SPARC64, AMD64, and Intel 64. In this document, x86 refers to the larger family of 64-bit x86 compatible products. Information in this document pertains to all platforms unless otherwise specified.

This document is intended for experienced system administrators with extensive knowledge of Oracle software and hardware. This document is not to be used as a planning or pre-sales guide.

The instructions in this book assume knowledge of the Oracle Solaris Operating System and expertise with the volume manager software used with Oracle Solaris Cluster.

Using UNIX Commands

This document contains information about commands specific to administering an Oracle Solaris Cluster configuration. This document might not contain complete information about basic UNIX commands and procedures.

See one or more of the following for this information:

- Online documentation for the Oracle Solaris software
- Other software documentation that you received with your system
- Oracle Solaris operating system (Oracle Solaris OS) man pages

Typographic Conventions

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

TABLE P-1 Typographic Conventions

Typeface	Meaning	Example
AaBbCc123	The names of commands, files, and directories, and onscreen computer output	Edit your <code>.login</code> file. Use <code>ls -a</code> to list all files. <code>machine_name% you have mail.</code>
AaBbCc123	What you type, contrasted with onscreen computer output	<code>machine_name% su</code> Password:
<i>aabbcc123</i>	Placeholder: replace with a real name or value	The command to remove a file is <code>rm filename</code> .
<i>AaBbCc123</i>	Book titles, new terms, and terms to be emphasized	Read Chapter 6 in the <i>User's Guide</i> . <i>A cache</i> is a copy that is stored locally. Do <i>not</i> save the file. Note: Some emphasized items appear bold online.

Shell Prompts in Command Examples

The following table shows the default UNIX system prompt and superuser prompt for shells that are included in the Oracle Solaris OS. Note that the default system prompt that is displayed in command examples varies, depending on the Oracle Solaris release.

TABLE P-2 Shell Prompts

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

Related Documentation

Information about related Oracle Solaris Cluster topics is available in the documentation that is listed in the following table. All Oracle Solaris Cluster documentation is available at <http://docs.sun.com>.

Topic	Documentation
Overview	<i>Oracle Solaris Cluster Overview</i>
Concepts	<i>Oracle Solaris Cluster Concepts Guide</i>
Hardware installation and administration	<i>Oracle Solaris Cluster 3.3 Hardware Administration Manual</i> Individual hardware administration guides
Software installation	<i>Oracle Solaris Cluster Software Installation Guide</i>
Data service installation and administration	<i>Oracle Solaris Cluster Data Services Planning and Administration Guide</i> Individual data service guides
Data service development	<i>Oracle Solaris Cluster Data Services Developer's Guide</i>
System administration	<i>Oracle Solaris Cluster System Administration Guide</i>
Error messages	<i>Oracle Solaris Cluster Error Messages Guide</i>
Command and function references	<i>Oracle Solaris Cluster Reference Manual</i>

For a complete list of Oracle Solaris Cluster documentation, see the release notes for your release of Oracle Solaris Cluster software at <http://docs.sun.com>.

Documentation, Support, and Training

See the following web sites for additional resources:

- [Documentation \(http://docs.sun.com\)](http://docs.sun.com)
- [Support \(http://www.oracle.com/us/support/systems/index.html\)](http://www.oracle.com/us/support/systems/index.html)
- [Training \(http://education.oracle.com\)](http://education.oracle.com) – Click the Sun link in the left navigation bar.

Oracle Welcomes Your Comments

Oracle welcomes your comments and suggestions on the quality and usefulness of its documentation. If you find any errors or have any other suggestions for improvement, go to <http://docs.sun.com> and click Feedback. Indicate the title and part number of the documentation along with the chapter, section, and page number, if available. Please let us know if you want a reply.

Oracle Technology Network (<http://www.oracle.com/technetwork/index.html>) offers a range of resources related to Oracle software:

- Discuss technical problems and solutions on the [Discussion Forums](http://forums.oracle.com) (<http://forums.oracle.com>).
- Get hands-on step-by-step tutorials with [Oracle By Example](http://www.oracle.com/technology/obe/start/index.html) (<http://www.oracle.com/technology/obe/start/index.html>).
- Download [Sample Code](http://www.oracle.com/technology/sample_code/index.html) (http://www.oracle.com/technology/sample_code/index.html).

Getting Help

Contact your service provider if you have problems installing or using Oracle Solaris Cluster. Provide the following information to your service provider.

- Your name and email address
- Your company name, address, and phone number
- The model and serial numbers of your systems
- The release number of the operating system, for example Oracle Solaris 10
- The release number of Oracle Solaris Cluster, for example, Oracle Solaris Cluster 3.3

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

Command	Function
<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>SPARC: prtdiag -v</code>	Displays system diagnostic information
<code>/usr/cluster/bin/clnode show-rev</code>	Displays Oracle Solaris Cluster release and package version information

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

Introduction to Administering Oracle Solaris Cluster

This chapter provides the following information about administering a global cluster and a zone cluster, and includes procedures for using Oracle Solaris Cluster administration tools:

- “Overview of Administering Oracle Solaris Cluster” on page 17
- “Oracle Solaris OS Feature Restrictions” on page 19
- “Administration Tools” on page 20
- “Preparing to Administer the Cluster” on page 22
- “Beginning to Administer the Cluster” on page 23

All procedures in this guide are for use on the Oracle Solaris 10 OS.

A global cluster is composed only of one or more global-cluster voting nodes and optionally, zero or more global-cluster non-voting nodes. A global cluster can optionally also include LINUX OS, or native brand, non-global zones that are not nodes, but high-availability containers (as resources). A zone cluster requires a global cluster. For general information about zone clusters, see the *Oracle Solaris Cluster Concepts Guide*.

A zone cluster is composed only of one or more cluster brand, voting nodes. A zone cluster depends on, and therefore requires, a global cluster. A global cluster does not contain a zone cluster. You cannot configure a zone cluster without a global cluster. A zone cluster has, at most, one zone cluster node on a machine. A zone-cluster node continues to operate only as long as the global-cluster voting node on the same machine continues to operate. If a global-cluster voting node on a machine fails, all zone-cluster nodes on that machine fail as well.

Overview of Administering Oracle Solaris Cluster

The Oracle Solaris Cluster highly available environment ensures that critical applications are available to end users. The system administrator's job is to make sure that the Oracle Solaris Cluster configuration is stable and operational.

Familiarize yourself with the planning information in the *Oracle Solaris Cluster Software Installation Guide* and the *Oracle Solaris Cluster Concepts Guide* before beginning

administration tasks. For instructions on creating a zone cluster, see “[Configuring a Zone Cluster](#)” in *Oracle Solaris Cluster Software Installation Guide*. Oracle Solaris Cluster administration is organized into tasks among the following manuals.

- Standard tasks, used to administer and maintain the global cluster or the zone cluster on a regular or even 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 *Oracle Solaris Cluster Data Services Planning and Administration Guide*.
- Service tasks, such as adding or repairing storage or network hardware. These tasks are described in the *Oracle Solaris Cluster 3.3 Hardware Administration Manual*.

In general, you can perform Oracle Solaris Cluster administration tasks while the cluster is operational. If you need take a node out of the cluster or even shut down the node, you can do so while the rest of the nodes continue cluster operations. Unless otherwise indicated, Oracle Solaris Cluster administrative tasks should be performed in the global-cluster voting node. For those procedures that require the entire cluster to be shut down, minimize impact on the system by scheduling downtime outside normal working hours. If you plan to shut down the cluster or a cluster node, notify users in advance.

Working With a Zone Cluster

Two Oracle Solaris Cluster administrative commands (`cluster` and `clnode`) can also run in a zone cluster. However, the scope of these commands is limited to the zone cluster where the command is issued. For example, using the `cluster` command in the global-cluster voting node retrieves all information about the voting global cluster and all the zone clusters. Using the `cluster` command in a zone cluster retrieves information about that specific zone cluster.

When you use the `clzonecluster` command in a voting node, the command affects all of the zone clusters in the global cluster. Zone cluster commands also affect all nodes on the zone cluster, even if the node is down when the command is issued.

Zone clusters support delegated administration of resources that are under Resource Group Manager (RGM) control. Therefore, zone cluster administrators can view, but not change, zone cluster dependencies that cross zone cluster boundaries. Only the administrator in a voting node can create, modify, or delete dependencies that cross zone cluster boundaries.

The following list contains the major administrative tasks performed on a zone cluster.

- Creating a zone cluster – Use the `clzonecluster configure` command to create a new zone cluster. See the instructions in “[Configuring a Zone Cluster](#)” in *Oracle Solaris Cluster Software Installation Guide*.
- Starting and rebooting a zone cluster – See [Chapter 3, “Shutting Down and Booting a Cluster.”](#)
- Adding a node to a zone cluster – See [Chapter 8, “Adding and Removing a Node.”](#)

- Removing a node from a zone cluster – See [“How to Remove a Node From a Zone Cluster”](#) on page 236.
- Viewing the configuration of a zone cluster – See [“How to View the Cluster Configuration”](#) on page 35.
- Validating the configuration of a zone cluster – See [“How to Validate a Basic Cluster Configuration”](#) on page 43.
- Stopping a zone cluster – See [Chapter 3, “Shutting Down and Booting a Cluster.”](#)

Oracle Solaris OS Feature Restrictions

Do not enable or disable the following Oracle Solaris Cluster services by using the Service Management Facility (SMF) management interface.

TABLE 1-1 Oracle Solaris ClusterServices

Oracle Solaris Cluster Services	FMRI
pnm	svc:/system/cluster/pnm:default
cl_event	svc:/system/cluster/cl_event:default
cl_eventlog	svc:/system/cluster/cl_eventlog:default
rpc_pmf	svc:/system/cluster/rpc_pmf:default
rpc_fed	svc:/system/cluster/rpc_fed:default
rgm	svc:/system/cluster/rgm:default
scdpm	svc:/system/cluster/scdpm:default
cl_ccra	svc:/system/cluster/cl_ccra:default
scsymon_srv	svc:/system/cluster/scsymon_srv:default
spm	svc:/system/cluster/spm:default
cl_svc_cluster_milestone	svc:/system/cluster/cl_svc_cluster_milestone:default
cl_svc_enable	svc:/system/cluster/cl_svc_enable:default
network-multipathing	svc:/system/cluster/network-multipathing

Administration Tools

You can perform administrative tasks on an Oracle Solaris Cluster configuration by using a graphical user interface (GUI) or by using the command line. The following section provides an overview of the GUI and command-line tools.

Graphical User Interface

Oracle Solaris Cluster software supports GUI tools that you can use to perform various administrative tasks on your cluster. These GUI tools are Oracle Solaris Cluster Manager and, if you are using Oracle Solaris Cluster software on a SPARC based system, Sun Management Center. See [Chapter 13, “Administering Oracle Solaris Cluster With the Graphical User Interfaces,”](#) for more information and for procedures about configuring Oracle Solaris Cluster Manager and Sun Management Center. For specific information about how to use Oracle Solaris Cluster Manager, see the online help for the GUI.

Command-Line Interface

You can perform most Oracle Solaris Cluster administration tasks interactively through the `clsetup(1CL)` utility. Whenever possible, administration procedures in this guide use the `clsetup` utility.

You can administer the following Main Menu items through the `clsetup` utility.

- Quorum
- Resource groups
- Data Services
- Cluster interconnect
- Device groups and volumes
- Private hostnames
- New nodes
- Other cluster tasks

Other commands that you use to administer an Oracle Solaris Cluster configuration are provided in the following list. See the man pages for more detailed information.

<code>ccp(1M)</code>	Starts remote console access to the cluster.
<code>if_mpadm(1M)</code>	Switches IP addresses from one adapter to another in an IP Network Multipathing group.
<code>claccess(1CL)</code>	Manages Oracle Solaris Cluster access policies for adding nodes.
<code>cldevice(1CL)</code>	Manages Oracle Solaris Cluster devices.

<code>cldevicegroup(1CL)</code>	Manages Oracle Solaris Cluster device groups.
<code>clinterconnect(1CL)</code>	Manages the Oracle Solaris Cluster interconnect.
<code>clnasdevice(1CL)</code>	Manages access to NAS devices for an Oracle Solaris Cluster configuration.
<code>clnode(1CL)</code>	Manages Oracle Solaris Cluster nodes.
<code>clquorum(1CL)</code>	Manages Oracle Solaris Cluster quorum.
<code>clreslogicalhostname(1CL)</code>	Manages Oracle Solaris Cluster resources for logical host names.
<code>clresource(1CL)</code>	Manages resources for Oracle Solaris Cluster data services.
<code>clresourcegroup(1CL)</code>	Manages resources for Oracle Solaris Cluster data services.
<code>clresourcetype(1CL)</code>	Manages resources for Oracle Solaris Cluster data services.
<code>clressharedaddress(1CL)</code>	Manages Oracle Solaris Cluster resources for shared addresses.
<code>clsetup(1CL)</code>	Configures an Oracle Solaris Cluster configuration interactively.
<code>clsnmphost(1CL)</code>	Administers Oracle Solaris Cluster SNMP hosts.
<code>clsnmpmib(1CL)</code>	Administers Oracle Solaris Cluster SNMP MIB.
<code>clsnmpuser(1CL)</code>	Administers Oracle Solaris Cluster SNMP users.
<code>cltelemetryattribute(1CL)</code>	Configures system resource monitoring.
<code>cluster(1CL)</code>	Manages the global configuration and the global status of the Oracle Solaris Cluster configuration.
<code>clvxvm(1CL)</code>	Initializes Veritas Volume Manager (VxVM) on an Oracle Solaris Cluster node and optionally performs root-disk encapsulation.
<code>clzonecluster(1CL)</code>	Creates and modifies a zone cluster.

In addition, you can use commands to administer the volume manager portion of an Oracle Solaris Cluster configuration. These commands depend on the specific volume manager that your cluster uses, either Veritas Volume Manager or Solaris Volume Manager.

Preparing to Administer the Cluster

This section describes how to prepare to administer your cluster.

Documenting an Oracle Solaris Cluster Hardware Configuration

Document the hardware aspects that are unique to your site as your Oracle Solaris Cluster configuration is scaled. To reduce administration, refer to your hardware documentation when you change or upgrade the cluster. Labeling cables and connections between the various cluster components can also make administration easier.

Reduce the time required by a third-party service provider when servicing your cluster by keeping records of your original cluster configuration, and subsequent changes.

Using an Administrative Console

You can use either a dedicated workstation or a workstation connected through a management network as the *administrative console*, to administer the active cluster. Typically, you install and run the Cluster Control Panel (CCP) and graphical user interface (GUI) tools on the administrative console. For more information about the CCP, see [“How to Log Into the Cluster Remotely” on page 25](#). For instructions on installing the Cluster Control Panel module for Sun Management Center and Oracle Solaris Cluster Manager GUI tools, see the [Oracle Solaris Cluster Software 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 SPARC cluster consists of Oracle's Sun Enterprise 10000 server, you must log in from the administrative console to the System Service Processor (SSP). Connect by using the `net con` command. The default method for `net con` to connect with a Sun Enterprise 10000 domain is through the network interface. If the network is inaccessible, you can use `net con` in “exclusive” mode by setting the `-f` option. You can also type `~*` during a normal `net con` session. Either of the previous solutions gives you the option of toggling to the serial interface if the network becomes unreachable.

Oracle Solaris Cluster does not require a dedicated administrative console, but using a console provides the following 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

Backing Up the Cluster

Back up your cluster on a regular basis. Even though Oracle Solaris Cluster software provides a highly available environment, with mirrored copies of data on the storage devices, Oracle Solaris Cluster software is not a replacement for regular backups. An Oracle Solaris Cluster configuration can survive multiple failures, but does not protect against user or program error, or catastrophic failure. Therefore, you must have a backup procedure in place to protect against data loss.

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

Beginning to Administer the Cluster

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

Note – The Oracle Solaris Cluster commands that you run only from the global-cluster voting node are not valid for use with zone clusters. See the appropriate Oracle Solaris Cluster man page for information about the valid use of a command in zones.

TABLE 1–2 Oracle Solaris Cluster Administration Tools

Task	Tool	Instructions
Log in to the cluster remotely	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> , <code>cssh</code> , or <code>ctelnet</code> .	“How to Log Into the Cluster Remotely” on page 25 “How to Connect Securely to Cluster Consoles” on page 26
Configure the cluster interactively	Start the <code>clzonecluster(1CL)</code> utility or the <code>clsetup(1CL)</code> utility.	“How to Access the Cluster Configuration Utilities” on page 27
Display Oracle Solaris Cluster release number and version information	Use the <code>clnode(1CL)</code> command with the <code>show - rev -v -node</code> subcommand and option.	“How to Display Oracle Solaris Cluster Release and Version Information” on page 28

TABLE 1-2 Oracle Solaris Cluster Administration Tools (Continued)

Task	Tool	Instructions
Display installed resources, resource groups, and resource types	Use the following commands to display the resource information: <ul style="list-style-type: none"> ■ <code>clresource(1CL)</code> ■ <code>clresourcegroup(1CL)</code> ■ <code>clresourcetype(1CL)</code> 	“How to Display Configured Resource Types, Resource Groups, and Resources” on page 30
Monitor cluster components graphically	Use Oracle Solaris Cluster Manager.	See online help
Administer some cluster components graphically	Use Oracle Solaris Cluster Manager or the Oracle Solaris Cluster module for Sun Management Center, which is available only with Oracle Solaris Cluster on SPARC based systems.	For Oracle Solaris Cluster Manager, see online help. For Sun Management Center, see Sun Management Center documentation.
Check the status of cluster components	Use the <code>cluster(1CL)</code> command with the <code>status</code> subcommand.	“How to Check the Status of Cluster Components” on page 31
Check the status of IP network multipathing groups on the public network	For a global cluster, use the <code>clnode(1CL)</code> status command with the <code>-m</code> option. For a zone cluster, use the <code>clzonecluster(1CL)</code> show command.	“How to Check the Status of the Public Network” on page 34
View the cluster configuration	For a global cluster, use the <code>cluster(1CL)</code> command with the <code>show</code> subcommand. For a zone cluster, use the <code>clzonecluster(1CL)</code> command with the <code>show</code> subcommand.	“How to View the Cluster Configuration” on page 35
View and display the configured NAS devices	For a global cluster or a zone cluster, use the <code>clzonecluster(1CL)</code> command with the <code>show</code> subcommand.	<code>clnasdevice(1CL)</code>
Check global mount points or verify the cluster configuration	For a global cluster, use the <code>cluster(1CL)cluster(1CL)</code> command with the <code>check</code> subcommand. For a zone cluster, use the <code>clzonecluster(1CL)</code> verify command.	“How to Validate a Basic Cluster Configuration” on page 43

TABLE 1-2 Oracle Solaris Cluster Administration Tools (Continued)

Task	Tool	Instructions
Look at the contents of Oracle Solaris Cluster command logs	Examine the <code>/var/cluster/logs/commandlog</code> file.	“How to View the Contents of Oracle Solaris Cluster Command Logs” on page 47
Look at Oracle Solaris Cluster system messages	Examine the <code>/var/adm/messages</code> file.	“Viewing System Messages” in System Administration Guide: Advanced Administration
Monitor the status of Solaris Volume Manager	Use the <code>metastat</code> command.	Solaris Volume Manager Administration Guide

▼ How to Log Into the Cluster Remotely

The Cluster Control Panel (CCP) provides a launchpad for the `cconsole`, `crlogin`, `cssh`, and `ctelnet` tools. All 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 to the common window is sent to each of the host windows, enabling you to run commands simultaneously on all nodes of the cluster.

You can also start `cconsole`, `crlogin`, `cssh`, or `ctelnet` sessions from the command line.

By default, the `cconsole` utility uses a `telnet` connection to the node consoles. To establish secure shell connections to the consoles instead, enable the Use SSH checkbox in the Options menu of the `cconsole` window. Or, specify the `-s` option when you issue the `ccp` or `cconsole` command.

See the [`ccp\(1M\)`](#) and [`cconsole\(1M\)`](#) man pages for more information.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

Before You Begin Verify that the following prerequisites are met before starting the CCP:

- Install the `SUNWccn` package on the administrative console.
- Ensure that the `PATH` variable on the administrative console includes the Oracle Solaris Cluster tools directories, `/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 you are using a terminal concentrator. The files can be either `/etc` files or NIS or NIS+ databases. See the [clusters\(4\)](#) and [serialports\(4\)](#) man pages for more information.
- 1 If you have a Sun Enterprise 10000 server platform, log in to the System Service Processor (SSP).
 - a. Connect by using the `netcon` command.
 - b. After the connection is made, type `Shift~@` to unlock the console and gain write access.
- 2 From the administrative console, start the CCP launchpad.


```
phys-schost# ccp clustername
```

 The CCP launchpad is displayed.
- 3 To start a remote session with the cluster, click the `cconsole` icon, `crlogin` icon, `cssh` icon, or `ctelnet` icon in the CCP launch pad.

▼ How to Connect Securely to Cluster Consoles

Perform this procedure to establish secure shell connections to the consoles of the cluster nodes.

Before You Begin Configure the `clusters` file, the `serialports` file, and the `nsswitch.conf` file if you are using a terminal concentrator. The files can be either `/etc` files or NIS or NIS+ databases.

Note – In the `serialports` file, assign the port number to use for secure connection to each console-access device. The default port number for secure shell connection is 22.

See the [clusters\(4\)](#) and [serialports\(4\)](#) man pages for more information.

- 1 Become superuser on the administrative console.
- 2 Start the `cconsole` utility in secure mode.


```
# cconsole -s [-l username] [-p ssh-port]
```

 - `-s` Enables secure shell connection.
 - `-l username` Specifies the user name for the remote connections. If the `-l` option is not specified, the user name that launched the `cconsole` utility is used.
 - `-p ssh-port` Specifies the secure shell port number to use. If the `-p` option is not specified, the default port number 22 is used for the secure connections.

▼ How to Access the Cluster Configuration Utilities

The `clsetup` utility enables you to interactively configure quorum, resource group, cluster transport, private hostname, device group, and new node options for the global cluster. The `clzonecluster` utility performs similar configuration tasks for a zone cluster. For more information, see the [clsetup\(1CL\)](#) and [clzonecluster\(1CL\)](#) man pages.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser on an active member node of a global cluster. Perform all steps of this procedure from a node of the global cluster.**

- 2 **Start the configuration utility.**

```
phys-schost# clsetup
```

- **For a global cluster, start the utility with the `clsetup` command.**

```
phys-schost# clsetup
```

The Main Menu is displayed.

- **For a zone cluster, start the utility with the `clzonecluster` command. The zone cluster in this example is `sczone`.**

```
phys-schost# clzonecluster configure sczone
```

You can view the available actions in the utility with the following option:

```
clzc:sczone> ?
```

- 3 **Choose your configuration from the menu. Follow the onscreen instructions to complete a task. For more detail, see the instructions in [“Configuring a Zone Cluster”](#) in *Oracle Solaris Cluster Software Installation Guide*.**

See Also See the `clsetup` or `clzonecluster` online help for more information.

▼ How to Display Oracle Solaris Cluster Patch Information

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

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- **Display the Oracle Solaris Cluster patch information:**

```
phys-schost# showrev -p
```

Oracle Solaris Cluster update releases are identified by the main product patch number plus the update version.

Example 1-1 Displaying Oracle Solaris Cluster Patch Information

The following example displays information about patch 110648-05.

```
phys-schost# showrev -p | grep 110648  
Patch: 110648-05 Obsoletes: Requires: Incompatibles: Packages:
```

▼ How to Display Oracle Solaris Cluster Release and Version Information

You do not need to be logged in as superuser to perform this procedure. Perform all steps of this procedure from a node of the global cluster.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- **Display Oracle Solaris Cluster release and version information:**

```
phys-schost# clnode show-rev -v -node
```

This command displays Oracle Solaris Cluster release number and version strings for all Oracle Solaris Cluster packages.

Example 1-2 Displaying Oracle Solaris Cluster Release and Version Information

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

```
phys-schost# clnode show-rev
3.2
```

```
phys-schost#% clnode show-rev -v
Oracle Solaris Cluster 3.3 for Solaris 10 sparc
SUNWscu: 3.3.0,REV=2010.06.14.03.44
SUNWsccomu: 3.3.0,REV=2010.06.14.03.44
SUNWsczr: 3.3.0,REV=2010.06.14.03.44
SUNWsccomzu: 3.3.0,REV=2010.06.14.03.44
SUNWsczu: 3.3.0,REV=2010.06.14.03.44
SUNWscsckr: 3.3.0,REV=2010.06.14.03.44
SUNWscscku: 3.3.0,REV=2010.06.14.03.44
SUNWscr: 3.3.0,REV=2010.06.14.03.44
SUNWscrtlh: 3.3.0,REV=2010.06.14.03.44
SUNWscnmr: 3.3.0,REV=2010.06.14.03.44
SUNWscnmu: 3.3.0,REV=2010.06.14.03.44
SUNWscdev: 3.3.0,REV=2010.06.14.03.44
SUNWscgds: 3.3.0,REV=2010.06.14.03.44
SUNWscsmf: 3.3.0,REV=2010.06.14.03.44
SUNWscman: 3.3.0,REV=2010.05.21.18.40
SUNWscsal: 3.3.0,REV=2010.06.14.03.44
SUNWscsam: 3.3.0,REV=2010.06.14.03.44
SUNWscvm: 3.3.0,REV=2010.06.14.03.44
SUNWmdmr: 3.3.0,REV=2010.06.14.03.44
SUNWmdmu: 3.3.0,REV=2010.06.14.03.44
SUNWscmasa: 3.3.0,REV=2010.06.14.03.44
SUNWscmasar: 3.3.0,REV=2010.06.14.03.44
SUNWscmasasen: 3.3.0,REV=2010.06.14.03.44
SUNWscmasazu: 3.3.0,REV=2010.06.14.03.44
SUNWscmasau: 3.3.0,REV=2010.06.14.03.44
SUNWscmautil: 3.3.0,REV=2010.06.14.03.44
SUNWscmautilr: 3.3.0,REV=2010.06.14.03.44
SUNWjfreechart: 3.3.0,REV=2010.06.14.03.44
SUNWscspmr: 3.3.0,REV=2010.06.14.03.44
SUNWscspmu: 3.3.0,REV=2010.06.14.03.44
SUNWscderby: 3.3.0,REV=2010.06.14.03.44
SUNWsc telemetry: 3.3.0,REV=2010.06.14.03.44
SUNWscgrepavs: 3.2.3,REV=2009.10.23.12.12
SUNWscgrepsrdf: 3.2.3,REV=2009.10.23.12.12
SUNWscgreptc: 3.2.3,REV=2009.10.23.12.12
SUNWscghb: 3.2.3,REV=2009.10.23.12.12
SUNWscgctl: 3.2.3,REV=2009.10.23.12.12
SUNWscims: 6.0,REV=2003.10.29
SUNWscics: 6.0,REV=2003.11.14
SUNWscapc: 3.2.0,REV=2006.12.06.18.32
SUNWscdns: 3.2.0,REV=2006.12.06.18.32
SUNWschadb: 3.2.0,REV=2006.12.06.18.32
SUNWschtt: 3.2.0,REV=2006.12.06.18.32
SUNWscslas: 3.2.0,REV=2006.12.06.18.32
SUNWscsckrb5: 3.2.0,REV=2006.12.06.18.32
SUNWscnfs: 3.2.0,REV=2006.12.06.18.32
SUNWscor: 3.2.0,REV=2006.12.06.18.32
SUNWscslmq: 3.2.0,REV=2006.12.06.18.32
SUNWscsap: 3.2.0,REV=2006.12.06.18.32
SUNWscslc: 3.2.0,REV=2006.12.06.18.32
SUNWscsapdb: 3.2.0,REV=2006.12.06.18.32
SUNWscsapenq: 3.2.0,REV=2006.12.06.18.32
SUNWscsaprepl: 3.2.0,REV=2006.12.06.18.32
```

```

SUNWscsapscs: 3.2.0,REV=2006.12.06.18.32
SUNWscsapwebas: 3.2.0,REV=2006.12.06.18.32
SUNWscsabl: 3.2.0,REV=2006.12.06.18.32
SUNWscsyb: 3.2.0,REV=2006.12.06.18.32
SUNWscwls: 3.2.0,REV=2006.12.06.18.32
SUNWsc9ias: 3.2.0,REV=2006.12.06.18.32
SUNWscPostgreSQL: 3.2.0,REV=2006.12.06.18.32
SUNWsczone: 3.2.0,REV=2006.12.06.18.32
SUNWscdhc: 3.2.0,REV=2006.12.06.18.32
SUNWsccebs: 3.2.0,REV=2006.12.06.18.32
SUNWscmqi: 3.2.0,REV=2006.12.06.18.32
SUNWscmqs: 3.2.0,REV=2006.12.06.18.32
SUNWscmys: 3.2.0,REV=2006.12.06.18.32
SUNWscsge: 3.2.0,REV=2006.12.06.18.32
SUNWscsaa: 3.2.0,REV=2006.12.06.18.32
SUNWscsag: 3.2.0,REV=2006.12.06.18.32
SUNWscsmb: 3.2.0,REV=2006.12.06.18.32
SUNWscsps: 3.2.0,REV=2006.12.06.18.32
SUNWscTomcat: 3.2.0,REV=2006.12.06.18.32

```

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

You can also accomplish this procedure by using the Oracle Solaris Cluster Manager GUI. Refer to [Chapter 13, “Administering Oracle Solaris Cluster With the Graphical User Interfaces,”](#) or see the Oracle Solaris Cluster Manager online help for more information.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

Before You Begin Users other than superuser require `solaris.cluster.read` RBAC authorization to use this subcommand.

- **Display the cluster's configured resource types, resource groups, and resources. Perform all steps of this procedure from a node of the global cluster.**

```
phys-schost# cluster show -t resource, resourcetype, resourcegroup
```

For information about individual resources, resource groups, and resource types, use the `show` subcommand with one of the following commands:

- `resource`
- `resource group`
- `resourcetype`

Example 1-3 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.

```
phys-schost# cluster show -t resource,resourcetype,resourcegroup

=== Registered Resource Types ===

Resource Type:                               SUNW.qfs
RT_description:                              SAM-QFS Agent on Oracle Solaris Cluster
RT_version:                                  3.1
API_version:                                  3
RT_basedir:                                   /opt/SUNWsamfs/sc/bin
Single_instance:                             False
Proxy:                                        False
Init_nodes:                                   All potential masters
Installed_nodes:                              <All>
Failover:                                     True
Pkglist:                                      <NULL>
RT_system:                                    False

=== Resource Groups and Resources ===

Resource Group:                               qfs-rg
RG_description:                              <NULL>
RG_mode:                                      Failover
RG_state:                                     Managed
Failback:                                     False
Nodelist:                                     phys-schost-2 phys-schost-1

--- Resources for Group qfs-rg ---

Resource:                                     qfs-res
Type:                                         SUNW.qfs
Type_version:                                3.1
Group:                                        qfs-rg
R_description:                               default
Resource_project_name:                       default
Enabled{phys-schost-2}:                       True
Enabled{phys-schost-1}:                       True
Monitored{phys-schost-2}:                     True
Monitored{phys-schost-1}:                     True
```

▼ How to Check the Status of Cluster Components

You can also accomplish this procedure by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

Note – The `cluster status` command also shows the status of a zone cluster.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

Before You Begin Users other than superuser require `solaris.cluster.read` RBAC authorization to use the `status` subcommand.

- **Check the status of cluster components. Perform all steps of this procedure from a node of the global cluster.**

```
phys-schost# cluster status
```

Example 1-4 Checking the Status of Cluster Components

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

```
phys-schost# cluster status
=== Cluster Nodes ===

--- Node Status ---

Node Name                Status
-----
phys-schost-1            Online
phys-schost-2            Online

=== Cluster Transport Paths ===

Endpoint1                Endpoint2                Status
-----
phys-schost-1:qfe1       phys-schost-4:qfe1       Path online
phys-schost-1:hme1       phys-schost-4:hme1       Path online

=== Cluster Quorum ===

--- Quorum Votes Summary ---

          Needed  Present  Possible
          -----
          3       3       4

--- Quorum Votes by Node ---

Node Name                Present  Possible  Status
-----
```

```

phys-schost-1  1          1          Online
phys-schost-2  1          1          Online

```

--- Quorum Votes by Device ---

Device Name	Present	Possible	Status
/dev/did/rdisk/d2s2	1	1	Online
/dev/did/rdisk/d8s2	0	1	Offline

=== Cluster Device Groups ===

--- Device Group Status ---

Device Group Name	Primary	Secondary	Status
schost-2	phys-schost-2	-	Degraded

--- Spare, Inactive, and In Transition Nodes ---

Device Group Name	Spare Nodes	Inactive Nodes	In Transition Nodes
schost-2	-	-	-

=== Cluster Resource Groups ===

Group Name	Node Name	Suspended	Status
test-rg	phys-schost-1	No	Offline
	phys-schost-2	No	Online
test-rg	phys-schost-1	No	Offline
	phys-schost-2	No	Error--stop failed
test-rg	phys-schost-1	No	Online
	phys-schost-2	No	Online

=== Cluster Resources ===

Resource Name	Node Name	Status	Message
test_1	phys-schost-1	Offline	Offline
	phys-schost-2	Online	Online
test_1	phys-schost-1	Offline	Offline
	phys-schost-2	Stop failed	Faulted
test_1	phys-schost-1	Online	Online
	phys-schost-2	Online	Online

Device Instance	Node	Status
-----	----	-----

```

/dev/did/rdisk/d2          phys-schost-1          Ok
/dev/did/rdisk/d3          phys-schost-1          Ok
                           phys-schost-2          Ok
/dev/did/rdisk/d4          phys-schost-1          Ok
                           phys-schost-2          Ok
/dev/did/rdisk/d6          phys-schost-2          Ok

```

```
=== Zone Clusters ===
```

```
--- Zone Cluster Status ---
```

Name	Node Name	Zone HostName	Status	Zone Status
sczone	schost-1	sczone-1	Online	Running
	schost-2	sczone-2	Online	Running

▼ How to Check the Status of the Public Network

You can also accomplish this procedure by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

To check the status of the IP Network Multipathing groups, use the `clnode(1CL)` command with the `status` subcommand.

Before You Begin Users other than superuser require `solaris.cluster.read` RBAC authorization to use this subcommand.

- **Check the status of cluster components. Perform all steps of this procedure from a node of the global cluster.**

```
phys-schost# clnode status -m
```

Example 1-5 Checking the Public Network Status

The following example provides a sample of status information for cluster components returned by the `clnode status` command.

```
% clnode status -m
--- Node IPMP Group Status ---

Node Name          Group Name      Status   Adapter   Status
-----
phys-schost-1     test-rg        Online   qfe1      Online
phys-schost-2     test-rg        Online   qfe1      Online
```

▼ How to View the Cluster Configuration

You can also perform this procedure by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

Before You Begin Users other than superuser require `solaris.cluster.read` RBAC authorization to use the `status` subcommand.

- **View the configuration of a global cluster or zone cluster. Perform all steps of this procedure from a node of the global cluster.**

```
% cluster show
```

Running the `cluster show` command from a global-cluster voting node shows detailed configuration information about the cluster and information for zone clusters, if you have configured them.

You can also use the `clzonecluster show` command to view the configuration information for just the zone cluster. Properties for a zone cluster include zone-cluster name, IP type, autoboot, and zone path. The `show` subcommand runs inside a zone cluster, and applies only to that particular zone cluster. Running the `clzonecluster show` command from a zone-cluster node retrieves status only about the objects visible to that specific zone cluster.

To display more information about the `cluster` command, use the verbose options. See the [cluster\(1CL\)](#) man page for details. See the [clzonecluster\(1CL\)](#) man page for more information about `clzonecluster`.

Example 1-6 Viewing the Global Cluster Configuration

The following example lists configuration information about the global cluster. If you have a zone cluster configured, it also lists that information.

```
phys-schost# cluster show
```

```
=== Cluster ===
```

```
Cluster Name:                cluster-1
installmode:                 disabled
heartbeat_timeout:          10000
heartbeat_quantum:          1000
private_netaddr:             172.16.0.0
private_netmask:             255.255.248.0
max_nodes:                   64
max_privatenets:             10
global_fencing:              Unknown
Node List:                   phys-schost-1
Node Zones:                   phys_schost-2:za
```

```
=== Host Access Control ===
```

```
Cluster name:                clustser-1
Allowed hosts:                phys-schost-1, phys-schost-2:za
Authentication Protocol:      sys
```

```
=== Cluster Nodes ===
```

```
Node Name:                   phys-schost-1
Node ID:                      1
Type:                         cluster
Enabled:                       yes
privatehostname:              clusternode1-priv
reboot_on_path_failure:       disabled
globalzoneshares:             3
defaulttpsetmin:              1
quorum_vote:                   1
quorum_defaultvote:           1
quorum_resv_key:               0x43CB1E1800000001
Transport Adapter List:        qfe3, hme0
```

```
--- Transport Adapters for phys-schost-1 ---
```

```
Transport Adapter:           qfe3
Adapter State:                Enabled
Adapter Transport Type:       dlpi
Adapter Property(device_name): qfe
Adapter Property(device_instance): 3
Adapter Property(lazy_free):  1
Adapter Property(dlpi_heartbeat_timeout): 10000
Adapter Property(dlpi_heartbeat_quantum): 1000
Adapter Property(nw_bandwidth): 80
Adapter Property(bandwidth):  10
Adapter Property(ip_address):  172.16.1.1
Adapter Property(netmask):      255.255.255.128
Adapter Port Names:            0
Adapter Port State(0):         Enabled
```

```
Transport Adapter:           hme0
Adapter State:                Enabled
Adapter Transport Type:       dlpi
Adapter Property(device_name): hme
```

```

Adapter Property(device_instance):      0
Adapter Property(lazy_free):            0
Adapter Property(dlpi_heartbeat_timeout): 10000
Adapter Property(dlpi_heartbeat_quantum): 1000
Adapter Property(nw_bandwidth):         80
Adapter Property(bandwidth):           10
Adapter Property(ip_address):           172.16.0.129
Adapter Property(netmask):              255.255.255.128
Adapter Port Names:                     0
Adapter Port State(0):                  Enabled

```

--- SNMP MIB Configuration on phys-schost-1 ---

```

SNMP MIB Name:      Event
State:              Disabled
Protocol:           SNMPv2

```

--- SNMP Host Configuration on phys-schost-1 ---

--- SNMP User Configuration on phys-schost-1 ---

```

SNMP User Name:      foo
Authentication Protocol: MD5
Default User:       No

```

```

Node Name:           phys-schost-2:za
Node ID:             2
Type:                cluster
Enabled:             yes
privatehostname:    clusternode2-priv
reboot_on_path_failure: disabled
globalzoneshares:  1
defaulttpsetmin:   2
quorum_vote:        1
quorum_defaultvote: 1
quorum_resv_key:    0x43CB1E1800000002
Transport Adapter List: hme0, qfe3

```

--- Transport Adapters for phys-schost-2 ---

```

Transport Adapter:      hme0
Adapter State:          Enabled
Adapter Transport Type: dlpi
Adapter Property(device_name): hme
Adapter Property(device_instance): 0
Adapter Property(lazy_free):      0
Adapter Property(dlpi_heartbeat_timeout): 10000
Adapter Property(dlpi_heartbeat_quantum): 1000
Adapter Property(nw_bandwidth):    80
Adapter Property(bandwidth):      10
Adapter Property(ip_address):      172.16.0.130
Adapter Property(netmask):         255.255.255.128
Adapter Port Names:          0
Adapter Port State(0):      Enabled

```

```

Transport Adapter:      qfe3
Adapter State:          Enabled
Adapter Transport Type: dlpi

```

```

Adapter Property(device_name):          qfe
Adapter Property(device_instance):      3
Adapter Property(lazy_free):            1
Adapter Property(dlpi_heartbeat_timeout): 10000
Adapter Property(dlpi_heartbeat_quantum): 1000
Adapter Property(nw_bandwidth):         80
Adapter Property(bandwidth):            10
Adapter Property(ip_address):           172.16.1.2
Adapter Property(netmask):              255.255.255.128
Adapter Port Names:                     0
Adapter Port State(0):                  Enabled

--- SNMP MIB Configuration on phys-schost-2 ---

SNMP MIB Name:                          Event
State:                                   Disabled
Protocol:                                SNMPv2

--- SNMP Host Configuration on phys-schost-2 ---

--- SNMP User Configuration on phys-schost-2 ---

=== Transport Cables ===

Transport Cable:                          phys-schost-1:qfe3,switch2@1
Cable Endpoint1:                         phys-schost-1:qfe3
Cable Endpoint2:                         switch2@1
Cable State:                             Enabled

Transport Cable:                          phys-schost-1:hme0,switch1@1
Cable Endpoint1:                         phys-schost-1:hme0
Cable Endpoint2:                         switch1@1
Cable State:                             Enabled

Transport Cable:                          phys-schost-2:hme0,switch1@2
Cable Endpoint1:                         phys-schost-2:hme0
Cable Endpoint2:                         switch1@2
Cable State:                             Enabled

Transport Cable:                          phys-schost-2:qfe3,switch2@2
Cable Endpoint1:                         phys-schost-2:qfe3
Cable Endpoint2:                         switch2@2
Cable State:                             Enabled

=== Transport Switches ===

Transport Switch:                          switch2
Switch State:                             Enabled
Switch Type:                              switch
Switch Port Names:                        1 2
Switch Port State(1):                     Enabled
Switch Port State(2):                     Enabled

Transport Switch:                          switch1
Switch State:                             Enabled
Switch Type:                              switch
Switch Port Names:                        1 2
Switch Port State(1):                     Enabled

```

Switch Port State(2): Enabled

=== Quorum Devices ===

Quorum Device Name: d3
 Enabled: yes
 Votes: 1
 Global Name: /dev/did/rdisk/d3s2
 Type: scsi
 Access Mode: scsi2
 Hosts (enabled): phys-schost-1, phys-schost-2

Quorum Device Name: qs1
 Enabled: yes
 Votes: 1
 Global Name: qs1
 Type: quorum_server
 Hosts (enabled): phys-schost-1, phys-schost-2
 Quorum Server Host: 10.11.114.83
 Port: 9000

=== Device Groups ===

Device Group Name: testdg3
 Type: SVM
 failback: no
 Node List: phys-schost-1, phys-schost-2
 preferred: yes
 numsecondaries: 1
 diskset name: testdg3

=== Registered Resource Types ===

Resource Type: SUNW.LogicalHostname:2
 RT_description: Logical Hostname Resource Type
 RT_version: 2
 API_version: 2
 RT_basedir: /usr/cluster/lib/rgm/rt/hafoip
 Single_instance: False
 Proxy: False
 Init_nodes: All potential masters
 Installed_nodes: <All>
 Failover: True
 Pkglist: SUNWscu
 RT_system: True

Resource Type: SUNW.SharedAddress:2
 RT_description: HA Shared Address Resource Type
 RT_version: 2
 API_version: 2
 RT_basedir: /usr/cluster/lib/rgm/rt/hascip
 Single_instance: False
 Proxy: False
 Init_nodes: <Unknown>
 Installed_nodes: <All>
 Failover: True

```

Pkglist:                SUNWscu
RT_system:              True

Resource Type:          SUNW.HASStoragePlus:4
RT_description:         HA Storage Plus
RT_version:             4
API_version:            2
RT_basedir:             /usr/cluster/lib/rgm/rt/hastorageplus
Single_instance:        False
Proxy:                  False
Init_nodes:             All potential masters
Installed_nodes:        <All>
Failover:               False
Pkglist:                SUNWscu
RT_system:              False

Resource Type:          SUNW.haderby
RT_description:         haderby server for Oracle Solaris Cluster
RT_version:             1
API_version:            7
RT_basedir:             /usr/cluster/lib/rgm/rt/haderby
Single_instance:        False
Proxy:                  False
Init_nodes:             All potential masters
Installed_nodes:        <All>
Failover:               False
Pkglist:                SUNWscderby
RT_system:              False

Resource Type:          SUNW.sctelemetry
RT_description:         sctelemetry service for Oracle Solaris Cluster
RT_version:             1
API_version:            7
RT_basedir:             /usr/cluster/lib/rgm/rt/sctelemetry
Single_instance:        True
Proxy:                  False
Init_nodes:             All potential masters
Installed_nodes:        <All>
Failover:               False
Pkglist:                SUNWsc telemetry
RT_system:              False

```

=== Resource Groups and Resources ===

```

Resource Group:         HA_RG
RG_description:         <Null>
RG_mode:                Failover
RG_state:               Managed
Failback:               False
Nodelist:               phys-schost-1 phys-schost-2

```

--- Resources for Group HA_RG ---

```

Resource:               HA_R
Type:                   SUNW.HASStoragePlus:4
Type_version:           4
Group:                  HA_RG
R_description:

```

```

Resource_project_name:          SCSLM_HA_RG
Enabled{phys-schost-1}:        True
Enabled{phys-schost-2}:        True
Monitored{phys-schost-1}:      True
Monitored{phys-schost-2}:      True

Resource Group:                 cl-db-rg
RG_description:                 <Null>
RG_mode:                        Failover
RG_state:                       Managed
Failback:                       False
Nodelist:                       phys-schost-1 phys-schost-2

--- Resources for Group cl-db-rg ---

Resource:                       cl-db-rs
Type:                           SUNW.haderby
Type_version:                   1
Group:                           cl-db-rg
R_description:
Resource_project_name:          default
Enabled{phys-schost-1}:        True
Enabled{phys-schost-2}:        True
Monitored{phys-schost-1}:      True
Monitored{phys-schost-2}:      True

Resource Group:                 cl-tlmtry-rg
RG_description:                 <Null>
RG_mode:                        Scalable
RG_state:                       Managed
Failback:                       False
Nodelist:                       phys-schost-1 phys-schost-2

--- Resources for Group cl-tlmtry-rg ---

Resource:                       cl-tlmtry-rs
Type:                           SUNW.sctelemetry
Type_version:                   1
Group:                           cl-tlmtry-rg
R_description:
Resource_project_name:          default
Enabled{phys-schost-1}:        True
Enabled{phys-schost-2}:        True
Monitored{phys-schost-1}:      True
Monitored{phys-schost-2}:      True

=== DID Device Instances ===

DID Device Name:                /dev/did/rdisk/d1
Full Device Path:               phys-schost-1:/dev/rdisk/c0t2d0
Replication:                    none
default_fencing:                global

DID Device Name:                /dev/did/rdisk/d2
Full Device Path:               phys-schost-1:/dev/rdisk/c1t0d0
Replication:                    none
default_fencing:                global

```

```

DID Device Name:                /dev/did/rdisk/d3
  Full Device Path:             phys-schost-2:/dev/rdisk/c2t1d0
  Full Device Path:             phys-schost-1:/dev/rdisk/c2t1d0
  Replication:                  none
  default_fencing:              global

DID Device Name:                /dev/did/rdisk/d4
  Full Device Path:             phys-schost-2:/dev/rdisk/c2t2d0
  Full Device Path:             phys-schost-1:/dev/rdisk/c2t2d0
  Replication:                  none
  default_fencing:              global

DID Device Name:                /dev/did/rdisk/d5
  Full Device Path:             phys-schost-2:/dev/rdisk/c0t2d0
  Replication:                  none
  default_fencing:              global

DID Device Name:                /dev/did/rdisk/d6
  Full Device Path:             phys-schost-2:/dev/rdisk/c1t0d0
  Replication:                  none
  default_fencing:              global

=== NAS Devices ===

Nas Device:                     nas_filer1
  Type:                          netapp
  User ID:                        root

Nas Device:                     nas2
  Type:                          netapp
  User ID:                        llai

```

Example 1-7 Viewing the Zone Cluster Configuration

The following example lists the properties of the zone cluster configuration.

```

% clzonecluster show
=== Zone Clusters ===

Zone Cluster Name:              sczone
  zonename:                      sczone
  zonename:                      /zones/sczone
  autoboot:                       TRUE
  ip-type:                        shared
  enable_priv_net:                TRUE

--- Solaris Resources for sczone ---

Resource Name:                  net
  address:                        172.16.0.1
  physical:                       auto

Resource Name:                  net
  address:                        172.16.0.2
  physical:                       auto

```

```
Resource Name: fs
  dir: /gz/db_qfs/CrsHome
  special: CrsHome
  raw:
  type: samfs
  options: []
```

```
Resource Name: fs
  dir: /gz/db_qfs/CrsData
  special: CrsData
  raw:
  type: samfs
  options: []
```

```
Resource Name: fs
  dir: /gz/db_qfs/OraHome
  special: OraHome
  raw:
  type: samfs
  options: []
```

```
Resource Name: fs
  dir: /gz/db_qfs/OraData
  special: OraData
  raw:
  type: samfs
  options: []
```

--- Zone Cluster Nodes for sczone ---

```
Node Name: sczone-1
  physical-host: sczone-1
  hostname: lzzone-1
```

```
Node Name: sczone-2
  physical-host: sczone-2
  hostname: lzzone-2
```

You can also view the NAS devices that are configured for global or zone clusters, by using the `clnasdevice show` subcommand or the Oracle Solaris Cluster Manager. See the [clnasdevice\(1CL\)](#) man page for more information.

▼ How to Validate a Basic Cluster Configuration

The `cluster(1CL)` command uses the `check` command to validate the basic configuration that is required for a global cluster to function properly. If no checks fail, `cluster check` returns to the shell prompt. If a check fails, `cluster check` produces reports in either the specified or the default output directory. If you run `cluster check` against more than one node, `cluster`

check produces a report for each node and a report for multinode checks. You can also use the `cluster list-checks` command to display a list of all available cluster checks.

You can run the `cluster check` command in verbose mode with the `-v` flag to display progress information.

Note – Run `cluster check` after performing an administration procedure that might result in changes to devices, volume management components, or the Oracle Solaris Cluster configuration.

Running the `clzonecluster(1CL)` command at the global—cluster voting node runs a set of checks to validate the configuration that is required for a zone cluster to function properly. If all checks pass, `clzonecluster verify` returns to the shell prompt and you can safely install the zone cluster. If a check fails, `clzonecluster verify` reports on the global-cluster nodes where the verification failed. If you run `clzonecluster verify` against more than one node, a report is produced for each node and a report for multinode checks. The `verify` subcommand is not allowed inside a zone cluster.

1 Become superuser on an active member node of a global cluster. Perform all steps of this procedure from a node of the global cluster.

```
phys-schost# su
```

2 Verify the cluster configuration.

- **Verify the configuration of the global cluster.**

```
phys-schost# cluster check
```

- **Verify the configuration of the zone cluster to see if a zone cluster can be installed.**

```
phys-schost# clzonecluster verify zoneclustername
```

Example 1–8 Checking the Global Cluster Configuration With All Checks Passing

The following example shows `cluster check` run in verbose mode against nodes `phys-schost-1` and `phys-schost-2` with all checks passing.

```
phys-schost# cluster check -v -h phys-schost-1,
phys-schost-2
```

```
cluster check: Requesting explorer data and node report from phys-schost-1.
cluster check: Requesting explorer data and node report from phys-schost-2.
cluster check: phys-schost-1: Explorer finished.
cluster check: phys-schost-1: Starting single-node checks.
cluster check: phys-schost-1: Single-node checks finished.
cluster check: phys-schost-2: Explorer finished.
cluster check: phys-schost-2: Starting single-node checks.
```

```

cluster check: phys-schost-2: Single-node checks finished.
cluster check: Starting multi-node checks.
cluster check: Multi-node checks finished
#

```

Example 1-9 Checking the Global Cluster Configuration With a Failed Check

The following example shows the node `phys-schost-2` in the cluster named `suncluster` minus the mount point `/global/phys-schost-1`. Reports are created in the output directory `/var/cluster/logs/cluster_check/<timestamp>`.

```

phys-schost# cluster check -v -h phys-schost-1,
phys-schost-2 -o
/var/cluster/logs/cluster_check/Dec5/

cluster check: Requesting explorer data and node report from phys-schost-1.
cluster check: Requesting explorer data and node report from phys-schost-2.
cluster check: phys-schost-1: Explorer finished.
cluster check: phys-schost-1: Starting single-node checks.
cluster check: phys-schost-1: Single-node checks finished.
cluster check: phys-schost-2: Explorer finished.
cluster check: phys-schost-2: Starting single-node checks.
cluster check: phys-schost-2: Single-node checks finished.
cluster check: Starting multi-node checks.
cluster check: Multi-node checks finished.
cluster check: One or more checks failed.
cluster check: The greatest severity of all check failures was 3 (HIGH).
cluster check: Reports are in /var/cluster/logs/cluster_check/<Dec5>.
#
# cat /var/cluster/logs/cluster_check/Dec5/cluster_check-results.suncluster.txt
...
=====
= ANALYSIS DETAILS =
=====
-----
CHECK ID : 3065
SEVERITY : HIGH
FAILURE : Global filesystem /etc/vfstab entries are not consistent across
all Oracle Solaris Cluster 3.x nodes.
ANALYSIS : The global filesystem /etc/vfstab entries are not consistent across
all nodes in this cluster.
Analysis indicates:
FileSystem '/global/phys-schost-1' is on 'phys-schost-1' but missing from 'phys-schost-2'.
RECOMMEND: Ensure each node has the correct /etc/vfstab entry for the
filesystem(s) in question.
...
#

```

▼ How to Check the Global Mount Points

The `cluster(ICL)` command includes checks that examine the `/etc/vfstab` file for configuration errors with the cluster file system and its global mount points.

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

- 1 **Become superuser on an active member node of a global cluster. Perform all steps of this procedure from a node of the global cluster.**

```
% su
```

- 2 **Verify the global cluster configuration.**

```
phys-schost# cluster check
```

Example 1-10 Checking the Global Mount Points

The following example shows the node `phys-schost-2` of the cluster named `suncluster` minus the mount point `/global/schost-1`. Reports are being sent to the output directory, `/var/cluster/logs/cluster_check/<timestamp>/`.

```
phys-schost# cluster check -v1 -h phys-schost-1,phys-schost-2 -o /var/cluster//logs/cluster_check/Dec5/
```

```
cluster check: Requesting explorer data and node report from phys-schost-1.
cluster check: Requesting explorer data and node report from phys-schost-2.
cluster check: phys-schost-1: Explorer finished.
cluster check: phys-schost-1: Starting single-node checks.
cluster check: phys-schost-1: Single-node checks finished.
cluster check: phys-schost-2: Explorer finished.
cluster check: phys-schost-2: Starting single-node checks.
cluster check: phys-schost-2: Single-node checks finished.
cluster check: Starting multi-node checks.
cluster check: Multi-node checks finished.
cluster check: One or more checks failed.
cluster check: The greatest severity of all check failures was 3 (HIGH).
cluster check: Reports are in /var/cluster/logs/cluster_check/Dec5.
#
# cat /var/cluster/logs/cluster_check/Dec5/cluster_check-results.suncluster.txt
...
=====
= ANALYSIS DETAILS =
=====
-----
CHECK ID : 3065
SEVERITY : HIGH
FAILURE : Global filesystem /etc/vfstab entries are not consistent across
all Oracle Solaris Cluster 3.x nodes.
ANALYSIS : The global filesystem /etc/vfstab entries are not consistent across
all nodes in this cluster.
Analysis indicates:
FileSystem '/global/phys-schost-1' is on 'phys-schost-1' but missing from 'phys-schost-2'.
RECOMMEND: Ensure each node has the correct /etc/vfstab entry for the
filesystem(s) in question.
...
#
```

```
# cat /var/cluster/logs/cluster_check/Dec5/cluster_check-results.phys-schost-1.txt
...
=====
= ANALYSIS DETAILS =
=====
-----
CHECK ID : 1398
SEVERITY : HIGH
FAILURE : An unsupported server is being used as an Oracle Solaris Cluster 3.x node.
ANALYSIS : This server may not been qualified to be used as an Oracle Solaris Cluster 3.x node.
Only servers that have been qualified with Oracle Solaris Cluster 3.x are supported as
Oracle Solaris Cluster 3.x nodes.
RECOMMEND: Because the list of supported servers is always being updated, check with
your Oracle representative to get the latest information on what servers
are currently supported and only use a server that is supported with Oracle Solaris Cluster 3.x.
...
#
```

▼ How to View the Contents of Oracle Solaris Cluster Command Logs

The `/var/cluster/logs/commandlog` ASCII text file contains records of selected Oracle Solaris Cluster commands that are executed in a cluster. The logging of commands starts automatically when you set up the cluster and ends when you shut down the cluster. Commands are logged on all nodes that are up and booted in cluster mode.

Commands that are not logged in this file include those commands that display the configuration and current state of the cluster.

Commands that are logged in this file include those commands that configure and change the current state of the cluster:

- `claccess`
- `cldevice`
- `cldevicegroup`
- `clinterconnect`
- `clnasdevice`
- `clnode`
- `clquorum`
- `clreslogicalhostname`
- `clresource`
- `clresourcegroup`
- `clresourcetype`
- `clressharedaddress`
- `clsetup`
- `clsnmphost`

- `clsnpmib`
- `clnmpuser`
- `cltelemetryattribute`
- `cluster`
- `clzonecluster`
- `scdidadm`

Records in the `commandlog` file can contain the following elements:

- Date and timestamp
- Name of the host from which the command was executed
- Process ID of the command
- Login name of the user who executed the command
- Command that the user executed, including all options and operands

Note – Command options are quoted in the `commandlog` file so that you can readily identify them and copy, paste, and execute them in the shell.

- Exit status of the executed command

Note – If a command aborts abnormally with unknown results, the Oracle Solaris Cluster software does *not* show an exit status in the `commandlog` file.

By default, the `commandlog` file is regularly archived once a week. To change the archiving policies for the `commandlog` file, on each node in the cluster, use the `crontab` command. See the [`crontab\(1\)`](#) man page for more information.

Oracle Solaris Cluster software maintains up to eight previously archived `commandlog` files on each cluster node at any given time. The `commandlog` file for the current week is named `commandlog`. The most recent complete week's file is named `commandlog.0`. The oldest complete week's file is named `commandlog.7`.

- **View the contents of the current week's commandlog file, one screen at a time.**

```
phys-schost# more /var/cluster/logs/commandlog
```

Example 1–11 Viewing the Contents of Oracle Solaris Cluster Command Logs

The following example shows the contents of the `commandlog` file that are displayed by the `more` command.

```
more -lines10 /var/cluster/logs/commandlog
11/11/2006 09:42:51 phys-schost-1 5222 root START - clsetup
11/11/2006 09:43:36 phys-schost-1 5758 root START - clrg add "app-sa-1"
11/11/2006 09:43:36 phys-schost-1 5758 root END 0
11/11/2006 09:43:36 phys-schost-1 5760 root START - clrg set -y
"RG_description=Department Shared Address RG" "app-sa-1"
11/11/2006 09:43:37 phys-schost-1 5760 root END 0
11/11/2006 09:44:15 phys-schost-1 5810 root START - clrg online "app-sa-1"
11/11/2006 09:44:15 phys-schost-1 5810 root END 0
11/11/2006 09:44:19 phys-schost-1 5222 root END -20988320
12/02/2006 14:37:21 phys-schost-1 5542 jbloggs START - clrg -c -g "app-sa-1"
-y "RG_description=Joe Bloggs Shared Address RG"
12/02/2006 14:37:22 phys-schost-1 5542 jbloggs END 0
```


Oracle Solaris Cluster and RBAC

This chapter describes role-based access control (RBAC) in relation to Oracle Solaris Cluster. Topics covered include:

- “Setting Up and Using RBAC With Oracle Solaris Cluster” on page 51
- “Oracle Solaris Cluster RBAC Rights Profiles” on page 52
- “Creating and Assigning an RBAC Role With an Oracle Solaris Cluster Management Rights Profile” on page 53
- “Modifying a User's RBAC Properties” on page 57

Setting Up and Using RBAC With Oracle Solaris Cluster

Use the following table to determine the documentation to consult about setting up and using RBAC. Specific steps that you follow to set up and use RBAC with Oracle Solaris Cluster software are provided later in this chapter.

Task	Instructions
Learn more about RBAC	Chapter 8, “Using Roles and Privileges (Overview),” in <i>System Administration Guide: Security Services</i>
Set up, manage elements, and use RBAC	Chapter 9, “Using Role-Based Access Control (Tasks),” in <i>System Administration Guide: Security Services</i>
Learn more about RBAC elements and tools	Chapter 10, “Role-Based Access Control (Reference),” in <i>System Administration Guide: Security Services</i>

Oracle Solaris Cluster RBAC Rights Profiles

Oracle Solaris Cluster Manager and selected Oracle Solaris Cluster commands and options that you issue at the command line use RBAC for authorization. Oracle Solaris Cluster commands and options that require RBAC authorization will require one or more of the following authorization levels. Oracle Solaris Cluster RBAC rights profiles apply to both voting and non-voting nodes in a global cluster.

`solaris.cluster.read` Authorization for list, show, and other read operations.

`solaris.cluster.admin` Authorization to change the state of a cluster object.

`solaris.cluster.modify` Authorization to change properties of a cluster object.

For more information about the RBAC authorization required by an Oracle Solaris Cluster command, see the command man page.

RBAC rights profiles include one or more RBAC authorizations. You can assign these rights profiles to users or to roles to give them different levels of access to Oracle Solaris Cluster. Oracle provides the following rights profiles with Oracle Solaris Cluster software.

Note – The RBAC rights profiles listed in the following table continue to support the old RBAC authorizations as defined in previous Oracle Solaris Cluster releases.

Rights Profile	Includes Authorizations	Role Identity Permission
Oracle Solaris Cluster Commands	None, but includes a list of Oracle Solaris Cluster commands that run with <code>euclid=0</code>	Execute selected Oracle Solaris Cluster commands that you use to configure and manage a cluster, including the following subcommands for all of the Oracle Solaris Cluster commands: <ul style="list-style-type: none"> ▪ list ▪ show ▪ status scha_control(1HA) scha_resource_get(1HA) scha_resource_setstatus(1HA) scha_resourcegroup_get(1HA) scha_resourcetype_get(1HA)
Basic Oracle Solaris User	This existing Oracle Solaris rights profile contains Oracle Solaris authorizations, as well as the following:	

Rights Profile	Includes Authorizations	Role Identity Permission
	<code>solaris.cluster.read</code>	Perform list, show, and other read operations for Oracle Solaris Cluster commands, as well as access the Oracle Solaris Cluster Manager GUI.
Cluster Operation	This rights profile is specific to Oracle Solaris Cluster software and contains the following authorizations:	
	<code>solaris.cluster.read</code>	Perform list, show, export, status, and other read operations, as well as access the Oracle Solaris Cluster Manager GUI.
	<code>solaris.cluster.admin</code>	Change the state of cluster objects.
System Administrator	This existing Oracle Solaris rights profile contains the same authorizations that the Cluster Management profile contains.	Perform the same operations that the Cluster Management role identity can perform, in addition to other system administration operations.
Cluster Management	This rights profile contains the same authorizations that the Cluster Operation profile contains, as well as the following authorization:	Perform the same operations that the Cluster Operation role identity can perform, as well as change properties of a cluster object.
	<code>solaris.cluster.modify</code>	

Creating and Assigning an RBAC Role With an Oracle Solaris Cluster Management Rights Profile

Use this task to create a new RBAC role with an Oracle Solaris Cluster Management Rights Profile and to assign users to this new role.

▼ How to Create a Role by Using the Administrative Roles Tool

Before You Begin To create a role, you must either assume a role that has the Primary Administrator rights profile assigned to it or run as root user.

1 Start the Administrative Roles tool.

To run the Administrative Roles tool, start the Solaris Management Console, as described in [“How to Assume a Role in the Solaris Management Console”](#) in *System Administration Guide: Security Services*. Open the User Tool Collection and click the Administrative Roles icon.

2 Start the Add Administrative Role wizard.

Select Add Administrative Role from the Action menu to start the Add Administrative Role wizard for configuring roles.

3 Set up a role to which the Cluster Management rights profile is assigned.

Use the Next and Back buttons to navigate between dialog boxes. Note that the Next button does not become active until you have filled in all required fields. The last dialog box enables you to review the entered data, at which point you can use the Back button to change entries or click Finish to save the new role. The following list summarizes the dialog box fields and buttons.

Role Name	Short name of the role.
Full Name	Long version of the name.
Description	Description of the role.
Role ID Number	UID for the role, automatically incremented.
Role Shell	The profile shells that are available to roles: Administrator's C, Administrator's Bourne, or Administrator's Korn shell.
Create a role mailing list	Makes a mailing list for users who are assigned to this role.
Available Rights / Granted Rights	Assigns or removes a role's rights profiles. Note that the system does not prevent you from typing multiple occurrences of the same command. The attributes that are assigned to the first occurrence of a command in a rights profile have precedence and all subsequent occurrences are ignored. Use the Up and Down arrows to change the order.
Server	Server for the home directory.
Path	Home directory path.
Add	Adds users who can assume this role. Must be in the same scope.
Delete	Deletes users who are assigned to this role.

Note – You need to place this profile first in the list of profiles that are assigned to the role.

4 Add users who need to use the Oracle Solaris Cluster Manager features or Oracle Solaris Cluster commands to the newly created role.

You use the `useradd(1M)` command to add a user account to the system. The `-P` option assigns a role to a user's account.

5 Click Finish.

6 Open a terminal window and become root.

7 Start and stop the name service cache daemon.

The new role does not take effect until the name service cache daemon is restarted. After becoming root, type the following text:

```
# /etc/init.d/nscd stop
# /etc/init.d/nscd start
```

▼ How to Create a Role From the Command Line

1 Become superuser or assume a role that provides `solaris.cluster.admin` RBAC authorization.

2 Select a method for creating a role:

- For roles in the local scope, use the `roleadd(1M)` command to specify a new local role and its attributes.
- Alternatively, for roles in the local scope, edit the `user_attr(4)` file to add a user with `type=role`.

Use this method only for emergencies.

- For roles in a name service, use the `smrole(1M)` command to specify the new role and its attributes.

This command requires authentication by superuser or a role that is capable of creating other roles. You can apply the `smrole` to all name services. This command runs as a client of the Solaris Management Console server.

3 Start and stop the name service cache daemon.

New roles do not take effect until the name service cache daemon is restarted. As root, type the following text:

```
# /etc/init.d/nscd stop
# /etc/init.d/nscd start
```

Example 2-1 Creating a Custom Operator Role by Using the `smrole` Command

The following sequence demonstrates how a role is created with the `smrole` command. In this example, a new version of the Operator role is created that has assigned to it the standard Operator rights profile and the Media Restore rights profile.

```
% su primaryadmin
# /usr/sadm/bin/smrole add -H myHost -- -c "Custom Operator" -n oper2 -a johnDoe \
-d /export/home/oper2 -F "Backup/Restore Operator" -p "Operator" -p "Media Restore"

Authenticating as user: primaryadmin

Type /? for help, pressing <enter> accepts the default denoted by [ ]
Please enter a string value for: password :: <type primaryadmin password>

Loading Tool: com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost
Login to myHost as user primaryadmin was successful.
Download of com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost was successful.

Type /? for help, pressing <enter> accepts the default denoted by [ ]
Please enter a string value for: password :: <type oper2 password>

# /etc/init.d/nscd stop
# /etc/init.d/nscd start
```

To view the newly created role (and any other roles), use `smrole` with the `list` option, as follows:

```
# /usr/sadm/bin/smrole list --
Authenticating as user: primaryadmin

Type /? for help, pressing <enter> accepts the default denoted by [ ]
Please enter a string value for: password :: <type primaryadmin password>

Loading Tool: com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost
Login to myHost as user primaryadmin was successful.
Download of com.sun.admin.usermgr.cli.role.UserMgrRoleCli from myHost was successful.
root                0                Super-User
primaryadmin        100              Most powerful role
sysadmin            101              Performs non-security admin tasks
oper2               102              Custom Operator
```

Modifying a User's RBAC Properties

You can modify a user's RBAC properties by using either the user accounts tool or the command line. To modify a user's RBAC properties, choose one of the following procedures.

- [“How to Modify a User's RBAC Properties by Using the User Accounts Tool” on page 57](#)
- [“How to Modify a User's RBAC Properties From the Command Line” on page 58](#)

▼ How to Modify a User's RBAC Properties by Using the User Accounts Tool

Before You Begin To modify a user's properties, you must run the User Tool Collection as root user or assume a role that has the primary administrator rights profile assigned to it.

1 Start the User Accounts tool.

To run the user accounts tool, start the Solaris Management Console, as described in [“How to Assume a Role in the Solaris Management Console” in *System Administration Guide: Security Services*](#). Open the User Tool Collection and click the User Accounts icon.

After the User Accounts tool starts, the icons for the existing user accounts are displayed in the view pane.

2 Click the User Account icon to be changed and select Properties from the Action menu (or double-click the user account icon).

3 Click the appropriate tab in the dialog box for the property to be changed, as follows:

- To change the roles that are assigned to the user, click the Roles tab and move the role assignment to be changed to the appropriate column: Available Roles or Assigned Roles.
- To change the rights profiles that are assigned to the user, click the Rights tab and move it to the appropriate column: Available Rights or Assigned Rights.

Note – Avoid assigning rights profiles directly to users. The preferred approach is to require users to assume roles in order to perform privileged applications. This strategy discourages users from abusing privileges.

▼ How to Modify a User's RBAC Properties From the Command Line

1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.

2 Choose the appropriate command:

- To change the authorizations, roles, or rights profiles that are assigned to a user who is defined in the local scope, use the `usermod(1M)` command.
- Alternatively, to change the authorizations, roles, or rights profiles that are assigned to a user who is defined in the local scope, edit the `user_attr` file.

Use this method for emergencies only.

- To change the authorizations, roles, or rights profiles that are assigned to a user who is defined in a name service, use the `smuser(1M)` command.

This command requires authentication as superuser or as a role that is capable of changing user files. You can apply `smuser` to all name services. `smuser` runs as a client of the Solaris Management Console server.

Shutting Down and Booting a Cluster

This chapter provides information about and procedures for shutting down and booting a global cluster, a zone cluster, and individual nodes. For information about booting a non-global zone, see Chapter 18, “Planning and Configuring Non-Global Zones (Tasks),” in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

- “Overview of Shutting Down and Booting a Cluster” on page 59
- “Shutting Down and Booting a Single Node in a Cluster” on page 68
- “Repairing a Full /var File System” on page 81

For a high-level description of the related procedures in this chapter, see “How to Boot a Node in Noncluster Mode” on page 78 and Table 3–2.

Overview of Shutting Down and Booting a Cluster

The Oracle Solaris Cluster `cluster(1CL)` shutdown command stops global cluster services in an orderly fashion and cleanly shuts down an entire global cluster. You can use the `cluster shutdown` command when moving the location of a global cluster, or to shut down the global cluster if an application error causes data corruption. The `clzonecluster halt` command stops a zone cluster that is running on a specific node or an entire zone cluster on all configured nodes. (You can also use the `cluster shutdown` command within a zone cluster.)

In the procedures in this chapter, `phys - schost#` reflects a global-cluster prompt. The `clzonecluster` interactive shell prompt is `clzc: schost>`.

Note – Use the `cluster shutdown` command to ensure proper shutdown of the entire global cluster. The Oracle Solaris shutdown command is used with the `clnode(1CL) evacuate` command to shut down individual nodes. See [“How to Shut Down a Cluster” on page 60](#) or [“Shutting Down and Booting a Single Node in a Cluster” on page 68](#) for more information.

The `cluster shutdown` and the `clzonecluster halt` commands stop all nodes in a global cluster or zone cluster, respectively, by performing the following actions:

1. Takes all running resource groups offline.
2. Unmounts all cluster file systems for a global cluster or a zone cluster.
3. The `cluster shutdown` command shuts down active device services on a global cluster or a zone cluster.
4. The `cluster shutdown` command runs `init 0` and brings all nodes on the cluster to the OpenBoot PROM `ok` prompt on a SPARC based system or the `Press any key to continue` message on the GRUB menu of an x86 based system. The GRUB menus are described in more detail in [“Booting an x86 Based System by Using GRUB \(Task Map\)” in *System Administration Guide: Basic Administration*](#). The `clzonecluster halt` command performs the `zoneadm -z zoneclustername halt` command to stop (but not shut down) the zones of the zone cluster.

Note – If necessary, you can boot a node in noncluster mode so that the node does not participate in cluster membership. Noncluster mode is useful when installing cluster software or for performing certain administrative procedures. See [“How to Boot a Node in Noncluster Mode” on page 78](#) for more information.

TABLE 3-1 Task List: Shutting Down and Booting a Cluster

Task	Instructions
Stop the cluster.	“How to Shut Down a Cluster” on page 60
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 63
Reboot the cluster.	“How to Reboot a Cluster” on page 65

▼ How to Shut Down a Cluster

You can shut down a global cluster, a zone cluster, or all zone clusters.



Caution – Do not use `send brk` on a cluster console to shut down a global-cluster node or a zone-cluster node. The command is not supported within a cluster.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **If your global cluster or zone cluster is running Oracle Real Application Clusters (RAC), shut down all instances of the database on the cluster you are shutting down.**
Refer to the Oracle RAC product documentation for shutdown procedures.
- 2 **Become superuser or assume a role that provides `solaris.cluster.admin` RBAC authorization on any node in the cluster. Perform all steps in this procedure from a node of the global cluster.**
- 3 **Shut down the global cluster, the zone cluster, or all zone clusters.**

- **Shut down the global cluster. This action also shuts down all zone clusters.**

```
phys-schost# cluster shutdown -g0 -y
```

- **Shut down a specific zone cluster.**

```
phys-schost# clzonecluster halt zoneclustername
```

- **Shut down all zone clusters.**

```
phys-schost# clzonecluster halt +
```

You can also use the `cluster shutdown` command within a zone cluster to shut down all zone clusters.

- 4 **Verify that all nodes on the global cluster or zone cluster are showing the `ok` prompt on a SPARC based system or a GRUB menu on an x86 based system.**

Do not power off any nodes until all nodes are at the `ok` prompt on a SPARC based system or in a boot subsystem on an x86 based system.

- **Verify that the global-cluster nodes are at the `ok` prompt on a SPARC based system or the `Press any key to continue` message on the GRUB based x86 systems.**

```
phys-schost# cluster status -t node
```

- **Use the `status` subcommand to verify that the zone cluster was shut down.**

```
phys-schost# clzonecluster status
```

5 If necessary, power off the nodes of the global cluster.

Example 3-1 Shutting Down a Zone Cluster

The following example shuts down a zone cluster called *sparse-sczone*.

```
phys-schost# clzonecluster halt sparse-sczone
Waiting for zone halt commands to complete on all the nodes of the zone cluster "sparse-sczone"...
Sep  5 19:06:01 schost-4 cl_runtime: NOTICE: Membership : Node 2 of cluster 'sparse-sczone' died.
Sep  5 19:06:01 schost-4 cl_runtime: NOTICE: Membership : Node 4 of cluster 'sparse-sczone' died.
Sep  5 19:06:01 schost-4 cl_runtime: NOTICE: Membership : Node 3 of cluster 'sparse-sczone' died.
Sep  5 19:06:01 schost-4 cl_runtime: NOTICE: Membership : Node 1 of cluster 'sparse-sczone' died.
phys-schost#
```

Example 3-2 SPARC: Shutting Down a Global Cluster

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

```
phys-schost# cluster shutdown -g0 -y
Wed Mar 10 13:47:32 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 clnode evacuate
The system is down.
syncing file systems... done
Program terminated
ok
```

Example 3-3 x86: Shutting Down a Global Cluster

The following example shows the console output when normal global-cluster operation is stopped and all nodes are shut down. In this example, the ok prompt is not displayed on all of the nodes. The `-g 0` option sets the shutdown grace period to zero, and the `-y` option provides an automatic yes response to the confirmation question. Shutdown messages also appear on the consoles of the other nodes in the global cluster.

```
phys-schost# cluster shutdown -g0 -y
May  2 10:32:57 phys-schost-1 cl_runtime:
WARNING: CMM: Monitoring disabled.
root@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 clnode evacuate
failfasts already disabled on node 1
Print services already stopped.
May  2 10:33:13 phys-schost-1 syslogd: going down on signal 15
The system is down.
syncing file systems... done
Type any key to continue

```

See Also See [“How to Boot a Cluster” on page 63](#) to restart a global cluster or a zone cluster that was shut down.

▼ How to Boot a Cluster

This procedure explains how to start a global cluster or zone cluster whose nodes have been shut down. For global-cluster nodes, the system displays the ok prompt on SPARC systems or the Press any key to continue message on the GRUB based x86 systems.

The phys-schost# prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

Note – To create a zone cluster, follow the instructions in [“Configuring a Zone Cluster” in Oracle Solaris Cluster Software Installation Guide](#).

1 Boot each node into cluster mode. Perform all steps in this procedure from a node of the global cluster.

- On SPARC based systems, run the following command.

```
ok boot
```

- On x86 based systems, run the following commands.

When the GRUB menu is displayed, select the appropriate Oracle Solaris entry and press Enter. The GRUB menu appears similar to the following:

```

GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-----+
| Solaris 10 /sol_10_x86                               |
| Solaris failsafe                                     |
|                                                       |
+-----+

```

Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.

For more information about GRUB based booting, see “[Booting an x86 Based System by Using GRUB \(Task Map\)](#)” in *System Administration Guide: Basic Administration*.

Note – Nodes must have a working connection to the cluster interconnect to attain cluster membership.

- **If you have a zone cluster, you can boot the entire zone cluster.**

```
phys-schost# clzonecluster boot zoneclustername
```

- **If you have more than one zone cluster, you can boot all zone clusters. Use + instead of the zoneclustername.**

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

The `cluster(1CL)` status command reports the global-cluster nodes' status.

```
phys-schost# cluster status -t node
```

When you run the `clzonecluster(1CL)` status command from a global-cluster node, the command reports the state of the zone-cluster node.

```
phys-schost# clzonecluster status
```

Note – If a node's `/var` file system fills up, Oracle Solaris Cluster might not be able to restart on that node. If this problem arises, see “[How to Repair a Full /var File System](#)” on page 81.

Example 3–4 SPARC: Booting a Global Cluster

The following example shows the console output when node `phys-schost-1` is booted into the global cluster. Similar messages appear on the consoles of the other nodes in the global cluster. When the `autoboot` property of a zone cluster is set to `true`, the system automatically boots the zone-cluster node after booting the global-cluster node on that machine.

When a global-cluster node reboots, all zone cluster nodes on that machine halt. Any zone-cluster node on that same machine with the `autoboot` property set to `true` boots after the global-cluster node restarts.

```
ok boot
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: Node phys-schost-1 with votecount = 1 added.
NOTICE: Node phys-schost-2 with votecount = 1 added.
NOTICE: Node phys-schost-3 with votecount = 1 added.
...
NOTICE: Node phys-schost-1: attempting to join cluster
...

```

```

NOTICE: Node phys-schost-2 (incarnation # 937690106) has become reachable.
NOTICE: Node phys-schost-3 (incarnation # 937690290) has become reachable.
NOTICE: cluster has reached quorum.
NOTICE: node phys-schost-1 is up; new incarnation number = 937846227.
NOTICE: node phys-schost-2 is up; new incarnation number = 937690106.
NOTICE: node phys-schost-3 is up; new incarnation number = 937690290.
NOTICE: Cluster members: phys-schost-1 phys-schost-2 phys-schost-3.
...

```

▼ How to Reboot a Cluster

To shut down a global cluster, run the `cluster shutdown` command and then boot the global cluster with the `boot` command on each node. To shut down a zone cluster, use the `clzonecluster halt` command and then use the `clzonecluster boot` command to boot the zone cluster. You can also use the `clzonecluster reboot` command. For more information, see the `cluster(1CL)`, `boot(1M)`, and `clzonecluster(1CL)` man pages.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **If your cluster is running Oracle RAC, shut down all instances of the database on the cluster you are shutting down.**
Refer to the Oracle RAC product documentation for shutdown procedures.
- 2 **Become superuser or assume a role that provides `soLarIS.cLuster.admin` RBAC authorization on any node in the cluster. Perform all steps in this procedure from a node of the global cluster.**
- 3 **Shut down the cluster.**

- **Shut down the global cluster.**

```
phys-schost# cluster shutdown -g0 -y
```

- **If you have a zone cluster, shut down the zone cluster from a global-cluster node.**

```
phys-schost# clzonecluster halt zoneclustername
```

Each node is shut down. You can also use the `cluster shutdown` command within a zone cluster to shut down the zone cluster.

Note – 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 is irrelevant unless you make configuration changes between shutdowns. If you make configuration changes between shutdowns, start the node with the most current configuration first.

- For a global-cluster node on a SPARC based system, run the following command.

```
ok boot
```

- For a global-cluster node on an x86 based system, run the following commands.

When the GRUB menu is displayed, select the appropriate Oracle Solaris OS entry and press Enter. The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-----+
| Solaris 10 /sol_10_x86                               |
| Solaris failsafe                                     |
|                                                     |
+-----+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

Note – Nodes must have a working connection to the cluster interconnect to attain cluster membership.

For more information about GRUB-based booting, see “[Booting an x86 Based System by Using GRUB \(Task Map\)](#)” in *System Administration Guide: Basic Administration*.

- For a zone cluster, type the following command on a single node of the global cluster to boot the zone cluster.

```
phys-schost# clzonecluster boot zoneclustername
```

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

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

- **The `clnode status` command reports the status of the nodes on the global cluster.**

```
phys-schost# clnode status
```

- **Running the `clzonecluster status` command on a global-cluster node reports the status of the zone-cluster nodes.**

```
phys-schost# clzonecluster status
```

You can also run the `cluster status` command within a zone cluster to see the status of the nodes.

Note – If a node's /var file system fills up, Oracle Solaris Cluster might not be able to restart on that node. If this problem arises, see [“How to Repair a Full /var File System” on page 81](#).

Example 3-5 Rebooting a Zone Cluster

The following example shows how to halt and boot a zone cluster called *sparse-sczone*. You can also use the `clzonecluster reboot` command.

```
phys-schost# clzonecluster halt sparse-sczone
Waiting for zone halt commands to complete on all the nodes of the zone cluster "sparse-sczone"...
Sep  5 19:17:46 schost-4 cl_runtime: NOTICE: Membership : Node 4 of cluster 'sparse-sczone' died.
Sep  5 19:17:46 schost-4 cl_runtime: NOTICE: Membership : Node 2 of cluster 'sparse-sczone' died.
Sep  5 19:17:46 schost-4 cl_runtime: NOTICE: Membership : Node 1 of cluster 'sparse-sczone' died.
Sep  5 19:17:46 schost-4 cl_runtime: NOTICE: Membership : Node 3 of cluster 'sparse-sczone' died.
phys-schost#
phys-schost# clzonecluster boot sparse-sczone
Waiting for zone boot commands to complete on all the nodes of the zone cluster "sparse-sczone"...
phys-schost# Sep  5 19:18:23 schost-4 cl_runtime: NOTICE: Membership : Node 1 of cluster
'sparse-sczone' joined.
Sep  5 19:18:23 schost-4 cl_runtime: NOTICE: Membership : Node 2 of cluster 'sparse-sczone' joined.
Sep  5 19:18:23 schost-4 cl_runtime: NOTICE: Membership : Node 3 of cluster 'sparse-sczone' joined.
Sep  5 19:18:23 schost-4 cl_runtime: NOTICE: Membership : Node 4 of cluster 'sparse-sczone' joined.

phys-schost#
phys-schost# clzonecluster status

=== Zone Clusters ===

--- Zone Cluster Status ---

Name           Node Name   Zone HostName   Status   Zone Status
-----
sparse-sczone  schost-1   sczone-1       Online   Running
                schost-2   sczone-2       Online   Running
                schost-3   sczone-3       Online   Running
                schost-4   sczone-4       Online   Running

phys-schost#
```

Example 3-6 SPARC: Rebooting a Global Cluster

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

```
phys-schost# cluster shutdown -g0 -y
Wed Mar 10 13:47:32 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 phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-2 (incarnation # 937690106) has become reachable.
NOTICE: Node phys-schost-3 (incarnation # 937690290) has become reachable.
NOTICE: cluster has reached quorum.
...
NOTICE: Cluster members: phys-schost-1 phys-schost-2 phys-schost-3.
...
NOTICE: Node phys-schost-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:
NOTICE: Node phys-schost-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:
```

Shutting Down and Booting a Single Node in a Cluster

You can shut down a global-cluster node, a zone-cluster node, or a non-global zone. This section provides instructions for shutting down a global-cluster node and a zone-cluster node.

To shut down a global-cluster node, use the `clnode evacuate` command with the Oracle Solaris shutdown command. Use the `cluster shutdown` command only when shutting down an entire global cluster.

On a zone-cluster node, use the `clzonecluster halt` command on a global cluster to shut down a single zone-cluster node or an entire zone cluster. You can also use the `clnode evacuate` and `shutdown` commands to shut down a zone-cluster node.

For information about shutting down and booting a non-global zone, see [Chapter 20](#), “Installing, Booting, Halting, Uninstalling, and Cloning Non-Global Zones (Tasks),” in *System*

Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones. See also `clnode(1CL)`, `shutdown(1M)`, and `clzonecluster(1CL)`.

In the procedures in this chapter, `phys - schost#` reflects a global-cluster prompt. The `clzonecluster` interactive shell prompt is `clzc : schost>`.

TABLE 3-2 Task Map: Shutting Down and Booting a Node

Task	Tool	Instructions
Stop a node.	For a global-cluster node, use <code>clnode(1CL) evacuate</code> and <code>shutdown</code> . For a zone-cluster node, use <code>clzonecluster(1CL) halt</code> .	“How to Shut Down a Node” on page 69
Start a node. The node must have a working connection to the cluster interconnect to attain cluster membership.	For a global-cluster node, use <code>boot</code> or <code>b</code> . For a zone-cluster node, use <code>clzonecluster(1CL) boot</code> .	“How to Boot a Node” on page 72
Stop and restart (reboot) a node on a cluster. The node must have a working connection to the cluster interconnect to attain cluster membership.	For a global-cluster node, use <code>clnode evacuate</code> and <code>shutdown</code> , followed by <code>boot</code> or <code>b</code> . For a zone-cluster node, use <code>clzonecluster(1CL) reboot</code> .	“How to Reboot a Node” on page 74
Boot a node so that the node does not participate in cluster membership.	For a global-cluster node, use <code>clnode evacuate</code> and <code>shutdown</code> commands, followed by <code>boot -x</code> on SPARC or GRUB menu entry editing on x86. If the underlying global cluster is booted in noncluster mode, the zone cluster node is automatically in noncluster mode.	“How to Boot a Node in Noncluster Mode” on page 78

▼ How to Shut Down a Node

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.



Caution – Do not use `send brk` on a cluster console to shut down a node on a global cluster or a zone cluster. The command is not supported within a cluster.

1 If your cluster is running Oracle RAC, shut down all instances of the database on the cluster you are shutting down.

Refer to the Oracle RAC product documentation for shutdown procedures.

2 Become superuser or assume a role that provides `soLaris.cLuster.admin` RBAC authorization on the cluster node to be shut down. Perform all steps in this procedure from a node of the global cluster.

3 If you want to halt a specific zone cluster member, skip Steps 4 - 6 and execute the following command from a global-cluster node:

```
phys-schost# clzonecluster halt -n physical-name zoneclustername
```

When you specify a particular zone-cluster node, you stop only that node. By default, the `halt` command stops the zone clusters on all nodes.

4 Switch all resource groups, resources, and device groups from the node being shut down to other global cluster members.

On the global-cluster node to shut down, type the following command. The `clnode evacuate` command switches over all resource groups and device groups, including all non-global zones, from the specified node to the next-preferred node. (You can also run `clnode evacuate` within a zone-cluster node.)

```
phys-schost# clnode evacuate node
```

node Specifies the node from which you are switching resource groups and device groups.

5 Shut down the node.

Specify the global-cluster node you want to shut down.

```
phys-schost# shutdown -g0 -y -i0
```

Verify that the global-cluster node is showing the `ok` prompt on a SPARC based system or the `Press any key to continue` message on the GRUB menu on an x86 based system.

6 If necessary, power off the node.

Example 3-7 SPARC: Shutting Down a Global-Cluster Node

The following example shows the console output when node `phys-schost-1` is shut down. The `-g0` option sets the grace period to zero, and the `-y` option provides an automatic yes response to the confirmation question. Shutdown messages for this node appear on the consoles of other nodes in the global cluster.

```
phys-schost# clnode evacuate nodename
phys-schost# shutdown -g0 -y
Wed Mar 10 13:47:32 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
Program terminated
ok
```

Example 3-8 x86: Shutting Down a Global-Cluster Node

The following example shows the console output when node `phys-schost-1` is shut down. The `-g0` option sets the grace period to zero, and the `-y` option provides an automatic yes response to the confirmation question. Shutdown messages for this node appear on the consoles of other nodes in the global cluster.

```
phys-schost# clnode evacuate phys-schost-1
phys-schost# shutdown -g0 -y
Shutdown started.   Wed Mar 10 13:47:32 PST 2004

Changing to init state 0 - please wait
Broadcast Message from root (console) on phys-schost-1 Wed Mar 10 13:47:32...
THE SYSTEM phys-schost-1 IS BEING SHUT DOWN NOW ! ! !
Log off now or risk your files being damaged

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 clnode evacuate
failfasts disabled on node 1
Print services already stopped.
Mar 10 13:47:44 phys-schost-1 syslogd: going down on signal 15
umount: /global/.devices/node@2 busy
umount: /global/.devices/node@1 busy
The system is down.
syncing file systems... done
```

```
WARNING: CMM: Node being shut down.
Type any key to continue
```

Example 3-9 Shutting Down a Zone-Cluster Node

The following example shows how use the `clzonecluster halt` to shut down a node on a zone cluster called *sparse-sczone*. (You can also run the `clnode evacuate` and `shutdown` commands in a zone-cluster node.)

```
phys-schost# clzonecluster status
```

```
=== Zone Clusters ===
```

```
--- Zone Cluster Status ---
```

Name	Node Name	Zone HostName	Status	Zone Status
sparse-sczone	schost-1	sczone-1	Online	Running
	schost-2	sczone-2	Online	Running
	schost-3	sczone-3	Online	Running
	schost-4	sczone-4	Online	Running

```
phys-schost#
```

```
phys-schost# clzonecluster halt -n schost-4 sparse-sczone
```

```
Waiting for zone halt commands to complete on all the nodes of the zone cluster "sparse-sczone"...
```

```
Sep 5 19:24:00 schost-4 cl_runtime: NOTICE: Membership : Node 3 of cluster 'sparse-sczone' died.
```

```
phys-host#
```

```
phys-host# clzonecluster status
```

```
=== Zone Clusters ===
```

```
--- Zone Cluster Status ---
```

Name	Node Name	Zone HostName	Status	Zone Status
sparse-sczone	schost-1	sczone-1	Online	Running
	schost-2	sczone-2	Online	Running
	schost-3	sczone-3	Offline	Installed
	schost-4	sczone-4	Online	Running

```
phys-schost#
```

See Also See [“How to Boot a Node”](#) on page 72 to restart a global-cluster node that was shut down.

▼ How to Boot a Node

If you intend to shut down or reboot other active nodes in the global cluster or zone cluster, wait until the `multiuser-server` milestone comes online for the node you are booting.

Otherwise, the node will not be available to take over services from other nodes in the cluster that you shut down or reboot. For information about booting a non-global zone, see [Chapter 20, “Installing, Booting, Halting, Uninstalling, and Cloning Non-Global Zones \(Tasks\)”](#), in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

Note – Starting a node can be affected by the quorum configuration. In a two-node cluster, you must have a quorum device configured so that the total quorum count for the cluster is three. You should have one quorum count for each node and one quorum count 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. The required cluster quorum count (two) must be present.

The `phys - s chost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

Note – Nodes must have a working connection to the cluster interconnect to attain cluster membership.

1 To start a global-cluster node or zone-cluster node that has been shut down, boot the node. Perform all steps in this procedure from a node of the global cluster.

- On SPARC based systems, run the following command.

```
ok boot
```

- On x86 based systems, run the following commands.

When the GRUB menu is displayed, select the appropriate Oracle Solaris entry and press Enter. The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-----+
| Solaris 10 /sol_10_x86                               |
| Solaris failsafe                                     |
|                                                       |
+-----+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

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

- If you have a zone cluster, you can specify a node to boot.

```
phys-schost# clzonecluster boot -n node zoneclustername
```

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

- **Running the `cluster status` command reports the status of a global-cluster node.**

```
phys-schost# cluster status -t node
```

- **Running the `clzonecluster status` command from a node on the global cluster reports the status of all zone-cluster nodes.**

```
phys-schost# clzonecluster status
```

A zone-cluster node can only be booted in cluster mode when the node hosting the node is booted in cluster mode.

Note – If a node's /var file system fills up, Oracle Solaris Cluster might not be able to restart on that node. If this problem arises, see “[How to Repair a Full /var File System](#)” on page 81.

Example 3–10 SPARC: Booting a Global-Cluster Node

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

```
ok boot
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
...
NOTICE: Node phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-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 Node

To shut down or reboot other active nodes in the global cluster or zone cluster, wait until the `multiuser-server` milestone comes online for the node that you are rebooting.

Otherwise, the node will not be available to take over services from other nodes in the cluster that you shut down or reboot. For information about rebooting a non-global zone, see [Chapter 20, “Installing, Booting, Halting, Uninstalling, and Cloning Non-Global Zones \(Tasks\)”](#) in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 If the global-cluster or zone-cluster node is running Oracle RAC, shut down all instances of the database on the node that you are shutting down.**
Refer to the Oracle RAC product documentation for shutdown procedures.
- 2 Become superuser or assume a role that provides `soLaris.cLuster.admin` RBAC authorization on the node to shut down. Perform all steps in this procedure from a node of the global cluster.**
- 3 Shut down the global-cluster node by using the `cLnode evacuate` and `shutdown` commands. Shut down the zone cluster with the `cLzonecLuster halt` command executed on a node of the global cluster. (The `cLnode evacuate` and `shutdown` commands also work in a zone cluster.)**

For a global cluster, type the following commands on the node to shut down. The `cLnode evacuate` command switches over all device groups from the specified node to the next-preferred node. The command also switches all resource groups from global or non-global zones on the specified node to the next-preferred global or non-global zones on other nodes.

- On a SPARC based system, run the following commands.

```
phys-schost# cLnode evacuate node
phys-schost# shutdown -g0 -y -i6
```

- On an x86 based system, run the following commands.

```
phys-schost# cLnode evacuate node
phys-schost# shutdown -g0 -y -i6
```

When the GRUB menu is displayed, select the appropriate Oracle Solaris entry and press Enter. The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-----+
| Solaris 10 /sol_10_x86                               |
| Solaris failsafe                                    |
|                                                     |
+-----+
Use the ^ and v keys to select which entry is highlighted.
```

Press enter to boot the selected OS, 'e' to edit the commands before booting, or 'c' for a command-line.

- Specify the zone-cluster node to shut down and reboot.

```
phys-schost# clzonecluster reboot - node zoneclustername
```

Note – Nodes must have a working connection to the cluster interconnect to attain cluster membership.

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

- **Verify that the global-cluster node is online.**

```
phys-schost# cluster status -t node
```

- **Verify that the zone-cluster node is online.**

```
phys-schost# clzonecluster status
```

Example 3–11 SPARC: Rebooting a Global-Cluster Node

The following example shows the console output when node `phys-schost-1` is rebooted. Messages for this node, such as shutdown and startup notification, appear on the consoles of other nodes in the global cluster.

```
phys-schost# clnode evacuate phys-schost-1
phys-schost# shutdown -g0 -y -i6
Shutdown started.   Wed Mar 10 13:47:32 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 #5932401.
Ethernet address 8:8:20:99:ab:77, Host ID: 8899ab77.
...
Rebooting with command: boot
...
Hostname: phys-schost-1
Booting as part of a cluster
```

```

...
NOTICE: Node phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-1: joined cluster
...
The system is coming up. Please wait.
The system is ready.
phys-schost-1 console login:

```

Example 3–12 x86: Rebooting a Global-Cluster Node

The following example shows the console output when rebooting node `phys-schost-1`. Messages for this node, such as shutdown and startup notification, appear on the consoles of other nodes in the global cluster.

```

phys-schost# clnode evacuate phys-schost-1
phys-schost # shutdown -y -g0 -i6

GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-----+
| Solaris 10 /sol_10_x86                               |
| Solaris failsafe                                     |
|                                                       |
+-----+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
Hostname: phys-schost-1
Booting as part of a cluster
...
NOTICE: Node phys-schost-1: attempting to join cluster
...
NOTICE: Node phys-schost-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:

```

Example 3–13 Rebooting a Zone-Cluster Node

The following example shows how to reboot a node on a zone cluster.

```

phys-schost# clzonecluster reboot -n schost-4 sparse-sczone
Waiting for zone reboot commands to complete on all the nodes of the zone cluster
"sparse-sczone"...
Sep  5 19:40:59 schost-4 cl_runtime: NOTICE: Membership : Node 3 of cluster
'sparse-sczone' died.
phys-schost# Sep  5 19:41:27 schost-4 cl_runtime: NOTICE: Membership : Node 3 of cluster

```

```
'sparse-sczone' joined.
phys-schost#
phys-schost# clzonecluster status

=== Zone Clusters ===

--- Zone Cluster Status ---
Name           Node Name   Zone HostName  Status   Zone Status
-----
sparse-sczone  schost-1   sczone-1       Online   Running
                schost-2   sczone-2       Online   Running
                schost-3   sczone-3       Online   Running
                schost-4   sczone-4       Online   Running

phys-schost#
```

▼ How to Boot a Node in Noncluster Mode

You can boot a global-cluster node in noncluster mode, where the node does not participate in the cluster membership. Noncluster mode is useful when installing the cluster software or performing certain administrative procedures, such as patching a node. A zone-cluster node cannot be in a boot state that is different from the state of the underlying global-cluster node. If the global-cluster node is booted in noncluster mode, the zone-cluster node is automatically in noncluster mode.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `soLaris.cLuster.admin` RBAC authorization on the cluster to be started in noncluster mode. Perform all steps in this procedure from a node of the global cluster.**

- 2 **Shut down the zone-cluster node by running the `clzonecluster halt` command on a node of the global cluster. Shut down the global-cluster node by using the `clnode evacuate` and `shutdown` commands.**

The `clnode evacuate` command switches over all device groups from the specified node to the next-preferred node. The command also switches all resource groups from global or non-global zones on the specified node to the next-preferred global or non-global zones on other nodes.

- **Shut down a specific global cluster.**

```
phys-schost# clnode evacuate node
```

```
phys-schost# shutdown -g0 -y
```

- **Shut down a specific zone-cluster node from a global-cluster node.**

```
phys-schost# clzonecluster halt -n node zoneclustername
```

You can also use the `clnode evacuate` and `shutdown` commands within a zone cluster.

- 3 **Verify that the global-cluster node is showing the `ok` prompt on an Oracle Solaris-based system or the `Press any key to continue` message on a GRUB menu on an x86 based system.**
- 4 **Boot the global-cluster node in noncluster mode.**

- On SPARC based systems, run the following command.

```
ok boot -xs
```

- On x86 based systems, run the following commands.

- a. **In the GRUB menu, use the arrow keys to select the appropriate Oracle Solaris entry and type `e` to edit its commands.**

The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
```

```
+-----+
| Solaris 10 /sol_10_x86                               |
| Solaris failsafe                                     |
|                                                       |
+-----+-----+
```

Use the `^` and `v` keys to select which entry is highlighted. Press enter to boot the selected OS, `'e'` to edit the commands before booting, or `'c'` for a command-line.

For more information about GRUB based booting, see [“Booting an x86 Based System by Using GRUB \(Task Map\)”](#) in *System Administration Guide: Basic Administration*.

- b. In the boot parameters screen, use the arrow keys to select the kernel entry and type e to edit the entry.**

The GRUB boot parameters screen appears similar to the following:

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+-----+
| root (hd0,0,a)                                     |
| kernel /platform/i86pc/multiboot                  |
| module /platform/i86pc/boot_archive                |
+-----+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.
```

- c. Add -x to the command to specify system boot in noncluster mode.**

[Minimal BASH-like line editing is supported. For the first word, TAB lists possible command completions. Anywhere else TAB lists the possible completions of a device/filename. ESC at any time exits.]

```
grub edit> kernel /platform/i86pc/multiboot -x
```

- d. Press the Enter key to accept the change and return to the boot parameters screen.**

The screen displays the edited command.

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+-----+
| root (hd0,0,a)                                     |
| kernel /platform/i86pc/multiboot -x              |
| module /platform/i86pc/boot_archive                |
+-----+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.-
```

- e. Type b to boot the node into noncluster mode.**

Note – This change to the kernel boot parameter command does not persist over the system boot. The next time you reboot the node, it will boot into cluster mode. To boot into noncluster mode instead, perform these steps again to add the -x option to the kernel boot parameter command.

Example 3–14 SPARC: Booting a Global-Cluster Node in Noncluster Mode

The following example shows the console output when node `phys - schost - 1` is shut down and restarted in noncluster mode. The `-g0` option sets the grace period to zero, the `-y` option

provides an automatic yes response to the confirmation question, and the `-i0` option invokes run level 0 (zero). Shutdown messages for this node appear on the consoles of other nodes in the global cluster.

```
phys-schost# clnode evacuate phys-schost-1
phys-schost# cluster shutdown -g0 -y
Shutdown started.    Wed Mar 10 13:47:32 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 phys-schost-1 is being shut down.
Program terminated

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

Repairing a Full /var File System

Both Oracle Solaris software and Oracle Solaris 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, Oracle Solaris Cluster might not be able to restart on that node. Additionally, you might not be able to log in to the node.

▼ How to Repair a Full /var File System

If a node reports a full `/var` file system and continues to run Oracle Solaris Cluster services, use this procedure to clear the full file system. Refer to “[Viewing System Messages](#)” in *System Administration Guide: Advanced Administration* in *System Administration Guide: Advanced Administration* for more information.

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

2 Clear the full file system.

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

Data Replication Approaches

This chapter describes data replication technologies you can use with Oracle Solaris Cluster software. *Data replication* is defined as copying data from a primary storage device to a backup or secondary device. If the primary device fails, your data is available from the secondary device. Data replication helps assure high availability and disaster tolerance for your cluster.

Oracle Solaris Cluster software supports the following types of data replication:

- Between clusters – Use Oracle Solaris Cluster Geographic Edition for disaster recovery
- Within a cluster – Use as a replacement for host-based mirroring within a campus cluster

To perform data replication, you must have a device group that has the same name as the object you are replicating. A device can belong to only one device group at a time, so if you already have an Oracle Solaris Cluster device group that contains the device, you must delete the group before you add that device to a new device group. For instructions on creating and managing Solaris Volume Manager, Veritas Volume Manager, ZFS, or raw-disk device groups, see [“Administering Device Groups” on page 119](#) in Chapter 5.

You must understand both host-based and storage-based data replication before you can select the replication approach that best serves your cluster. For more information about using Oracle Solaris Cluster Geographic Edition to manage your data replication for disaster recovery, see the [Oracle Solaris Cluster Geographic Edition Overview](#).

This chapter contains the following sections:

- [“Understanding Data Replication” on page 84](#)
- [“Using Storage-Based Data Replication Within a Cluster” on page 85](#)

Understanding Data Replication

Oracle Solaris Cluster supports the following approaches to data replication:

- *Host-based data replication* uses software to replicate disk volumes between geographically dispersed clusters in real time. Remote mirror replication enables data from the master volume of the primary cluster to be replicated to the master volume of the geographically dispersed secondary cluster. A remote mirror bitmap tracks differences between the master volume on the primary disk and the master volume on the secondary disk. An example of host-based replication software used for replication between clusters (and between a cluster and a host that is not in a cluster) is Sun StorageTek Availability Suite 4.

Host-based data replication is a less expensive data replication solution because it uses host resources, rather than special storage arrays. Databases, applications, or file systems that are configured to allow multiple hosts running the Oracle Solaris OS to write data to a shared volume are not supported (for example, Oracle 9iRAC and Oracle Parallel Server). For more information about using host-based data replication between two clusters, see *Oracle Solaris Cluster Geographic Edition Data Replication Guide for Sun StorageTek Availability Suite*. To see an example of host-based replication that does not use Oracle Solaris Cluster Geographic Edition, see Appendix A, “Configuring Host-Based Data Replication With Sun StorageTek Availability Suite Software” on page 347.

- *Storage-based data replication* uses software on the storage controller to move the work of data replication off the cluster nodes and onto the storage device. This software frees some node processing power to serve cluster requests. Examples of storage-based software that can replicate data inside a cluster or between clusters include Hitachi TrueCopy, Hitachi Universal Replicator, and EMC SRDF. Storage-based data replication can be especially important in campus cluster configurations and can simplify the infrastructure required. For more information about using storage-based data replication in a campus cluster environment, see “Using Storage-Based Data Replication Within a Cluster” on page 85.

For more information about using storage-based replication between two or more clusters and the Oracle Solaris Cluster Geographic Edition product that automates the process, see *Oracle Solaris Cluster Geographic Edition Data Replication Guide for Hitachi TrueCopy and Universal Replicator* and *Oracle Solaris Cluster Geographic Edition Data Replication Guide for EMC Symmetrix Remote Data Facility*. See also Appendix A, “Configuring Host-Based Data Replication With Sun StorageTek Availability Suite Software” on page 347 for a complete example of this type of cluster configuration.

Supported Data Replication Methods

Oracle Solaris Cluster software supports the following methods of data replication between clusters or within a cluster:

1. Replication Between Clusters – For disaster recovery, you can use host-based or storage-based replication to perform data replication between clusters. Generally, you would choose either host-based replication or storage-based replication, rather than a combination of the two. You can manage both types of replication with Oracle Solaris Cluster Geographic Edition software.
 - Host-Based Replication
 - Sun StorageTek Availability Suite, starting with the Oracle Solaris 10 OS

If you want to use host-based replication without Oracle Solaris Cluster Geographic Edition software, see the instructions in [Appendix A, “Example,” “Configuring Host-Based Data Replication With Sun StorageTek Availability Suite Software”](#) on page 347.
 - Storage-Based Replication
 - Hitachi TrueCopy and Hitachi Universal Replicator, through the Oracle Solaris Cluster Geographic Edition
 - EMC Symmetrix Remote Data Facility (SRDF), through the Oracle Solaris Cluster Geographic Edition

If you want to use storage-based replication without Oracle Solaris Cluster Geographic Edition software, see the documentation for your replication software.
2. Replication Within a Cluster – This method is used as a replacement for host-based mirroring.
 - Storage-Based Replication
 - Hitachi TrueCopy and Hitachi Universal Replicator
 - EMC Symmetrix Remote Data Facility (SRDF)
3. Application-Based Replication – Oracle Data Guard is an example of application-based replication software. This type of software is used only for disaster recovery. For more information, see the [Oracle Solaris Cluster Geographic Edition Data Replication Guide for Oracle Data Guard](#)

Using Storage-Based Data Replication Within a Cluster

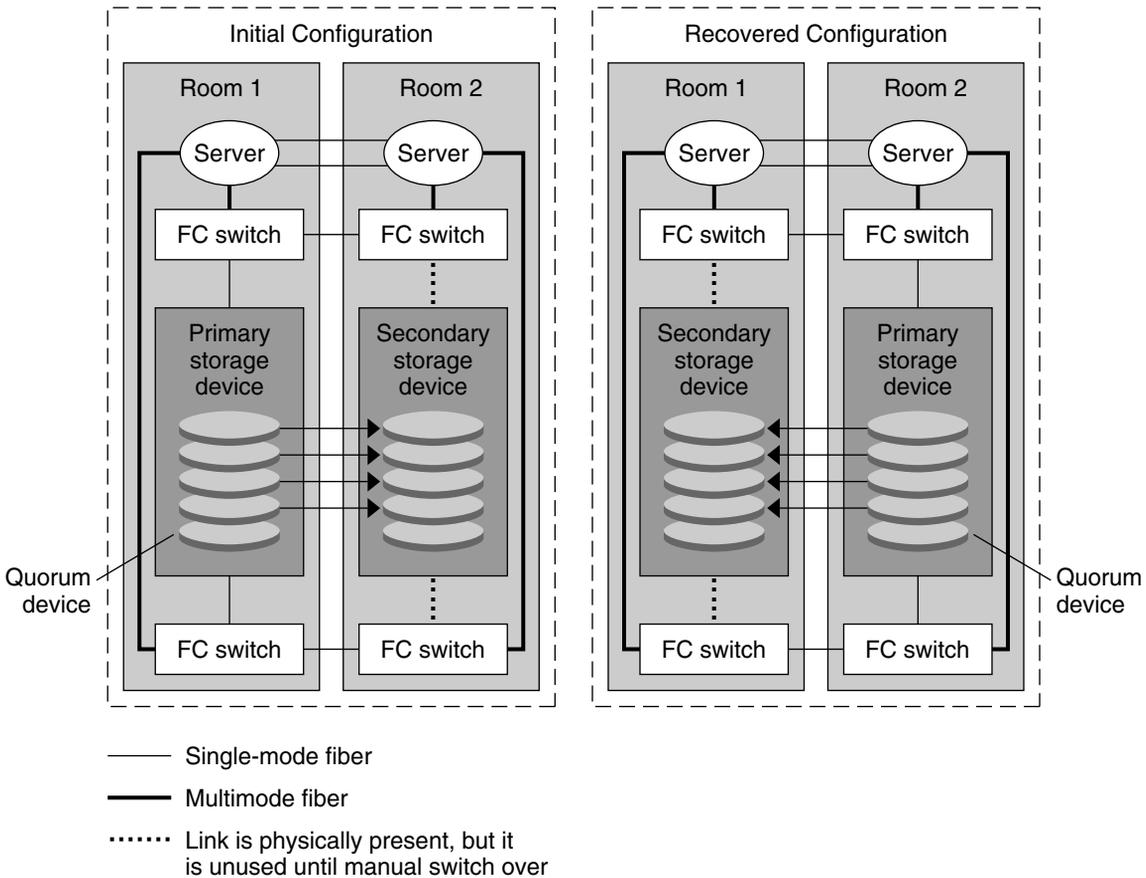
Storage-based data replication uses software installed on the storage device to manage the replication within a cluster or a campus cluster. Such software is specific to your particular storage device, and is not used for disaster recovery. Refer to the documentation that shipped with your storage device when configuring storage-based data replication.

Depending on the software you use, you can use either automatic or manual failover with storage-based data replication. Oracle Solaris Cluster supports both manual and automatic failover of the replicants with Hitachi TrueCopy, Hitachi Universal Replicator, and EMC SRDF software.

This section describes storage-based data replication as used in a campus cluster. [Figure 4-1](#) shows a sample two-room configuration where data is replicated between two storage arrays. In this configuration, the primary storage array is contained in the first room, where it provides data to the nodes in both rooms. The primary storage array also provides the secondary storage array with data to replicate.

Note – [Figure 4-1](#) illustrates that the quorum device is on an unreplicated volume. A replicated volume cannot be used as a quorum device.

FIGURE 4-1 Two-Room Configuration With Storage-Based Data Replication



Storage-based data replication with Hitachi TrueCopy or Hitachi Universal Replicator can be performed synchronously or asynchronously in the Oracle Solaris Cluster environment,

depending on the type of application you use. If you want to perform automatic failover in a campus cluster, use TrueCopy synchronously. Storage-based synchronous replication with EMC SRDF is supported with Oracle Solaris Cluster; asynchronous replication is not supported for EMC SRDF.

Do not use EMC SRDF's Domino mode or Adaptive Copy mode. Domino mode makes the local and target SRDF volumes unavailable to the host when the target is unavailable. Adaptive Copy mode is generally used for data migrations and data center moves and it not recommended for disaster recovery.

Do not use the Data or Status modes in Hitachi TrueCopy or Hitachi Universal Replicator. If the secondary storage device fails, you might experience problems writing to the primary storage device.

Requirements and Restrictions When Using Storage-Based Data Replication Within a Cluster

To ensure data integrity, use multipathing and the proper RAID package. The following list includes considerations for implementing a cluster configuration that uses storage-based data replication.

- Node-to-node distance is limited by the Oracle Solaris Cluster Fibre Channel and interconnect infrastructure. Contact your Oracle service provider for more information about current limitations and supported technologies.
- Do not configure a replicated volume as a quorum device. Locate any quorum devices on a shared, unreplicated volume or use the quorum server.
- Ensure that only the primary copy of the data is visible to cluster nodes. Otherwise, the volume manager might try to simultaneously access both primary and secondary copies of the data. Refer to the documentation that was shipped with your storage array for information about controlling the visibility of your data copies.
- EMC SRDF, Hitachi TrueCopy, and Hitachi Universal Replicator allow the user to define groups of replicated devices. Each replication device group requires an Oracle Solaris Cluster device group with the same name.
- Particular application-specific data might not be suitable for asynchronous data replication. Use your understanding of your application's behavior to determine how best to replicate application-specific data across the storage devices.
- If configuring the cluster for automatic failover, use synchronous replication.

For instructions on configuring the cluster for automatic failover of replicated volumes, see [“Administering Storage-Based Replicated Devices” on page 94](#).

- Oracle Real Application Clusters (RAC) is not supported with SRDF, Hitachi TrueCopy, and Hitachi Universal Replicator when replicating within a cluster. Nodes connected to replicas that are not currently the primary replica will not have write access. Any scalable application that requires direct write access from all nodes of the cluster cannot be supported with replicated devices.
- Veritas Cluster Volume Manager (CVM) and multi-owner Solaris Volume Manager for Oracle Solaris Cluster software are not supported.
- Do not use the Domino mode or Adaptive Copy mode in EMC SRDF. See [“Using Storage-Based Data Replication Within a Cluster” on page 85](#) for more information.
- Do not use the Data mode or Status mode in Hitachi TrueCopy or Hitachi Universal Replicator. See [“Using Storage-Based Data Replication Within a Cluster” on page 85](#) for more information.

Manual Recovery Concerns When Using Storage-Based Data Replication Within a Cluster

As with all campus clusters, those clusters that use storage-based data replication generally do not need intervention when they experience a single failure. However, if you are using manual failover and you lose the room that holds your primary storage device (as shown in [Figure 4–1](#)), problems arise in a two-node cluster. The remaining node cannot reserve the quorum device and cannot boot as a cluster member. In this situation, your cluster requires the following manual intervention:

1. Your Oracle service provider must reconfigure the remaining node to boot as a cluster member.
2. You or your Oracle service provider must configure an unreplicated volume of your secondary storage device as a quorum device.
3. You or your Oracle service provider must configure the remaining node to use the secondary storage device as primary storage. This reconfiguration might involve rebuilding volume manager volumes, restoring data, or changing application associations with storage volumes.

Best Practices When Using Storage-Based Data Replication

When setting up device groups that use the Hitachi TrueCopy or Hitachi Universal Replicator software for storage-based data replication, observe the following practices:

- Use synchronous replication to avoid the possibility of lost data if the primary site fails.
- A one-to-one relationship should exist between the Oracle Solaris Cluster global device group and the TrueCopy replication group defined in the `hor cm` configuration file. This allows both groups to move from node to node as a single unit.
- Global file system volumes and failover file system volumes cannot be mixed in the same replicated device group because they are controlled differently. Global file systems are controlled by a Device Configuration System (DCS), while failover file system volumes are controlled by HAS+. The primary for each could be a different node, causing conflicts on which node should be the replication primary.
- All RAID manager instances should be up and running at all times.

When using EMC SRDF software for storage-based data replication, use dynamic devices instead of static devices. Static devices require several minutes to change the replication primary and can impact failover time.

Administering Global Devices, Disk-Path Monitoring, and Cluster File Systems

This chapter provides information about and procedures for administering global devices, disk-path monitoring, and cluster file systems.

- “Overview of Administering Global Devices and the Global Namespace” on page 91
- “Administering Storage-Based Replicated Devices” on page 94
- “Overview of Administering Cluster File Systems” on page 117
- “Administering Device Groups” on page 119
- “Administering the SCSI Protocol Settings for Storage Devices” on page 160
- “Administering Cluster File Systems” on page 165
- “Administering Disk-Path Monitoring” on page 171

For a high-level description of the related procedures in this chapter, see [Table 5–4](#).

For conceptual information related to global devices, the global namespace, device groups, disk-path monitoring, and the cluster file system, see *Oracle Solaris Cluster Concepts Guide*.

Overview of Administering Global Devices and the Global Namespace

Administration of Oracle Solaris Cluster device groups depends on the volume manager that is installed on the cluster. Solaris Volume Manager is “cluster-aware,” so you add, register, and remove device groups by using the Solaris Volume Manager `metaset(1M)` command. If you are using Veritas Volume Manager (VxVM), you create disk groups by using VxVM commands. You register the disk groups as Oracle Solaris Cluster device groups with the `clsetup` utility. When removing VxVM device groups, you use both the `clsetup` command and VxVM commands.

Note – Global devices are not directly accessible from global-cluster non-voting nodes.

Oracle Solaris Cluster software automatically creates a raw-disk device group for each disk and tape device in the cluster. However, cluster device groups remain in an offline state until you access the groups as global devices. When administering 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. The global namespace is automatically set up during installation and automatically updated during Oracle Solaris OS reboots. However, if the global namespace needs to be updated, you can run the `cldevice populate` command from any cluster node. This command 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.

Global Device Permissions for Solaris Volume Manager

Changes made to global device permissions are not automatically propagated to all the nodes in the cluster for Solaris Volume Manager 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 issue the following command on all nodes in the cluster:

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

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

Dynamic Reconfiguration With Global Devices

You must consider the following issues when completing dynamic reconfiguration (DR) operations on disk and tape devices in a cluster.

- All of the requirements, procedures, and restrictions that are documented for the Oracle Solaris DR feature also apply to Oracle Solaris Cluster DR support. The only exception is for the operating system quiescence operation. Therefore, review the documentation for the Oracle Solaris DR feature *before* using the DR feature with Oracle Solaris Cluster software. You should review in particular the issues that affect non-network IO devices during a DR detach operation.
- Oracle Solaris Cluster rejects DR remove-board operations on active devices in the primary node. DR operations can be performed on inactive devices in the primary node and on any devices in the secondary nodes.
- After the DR operation, cluster data access continues as before.
- Oracle Solaris Cluster rejects DR operations that impact the availability of quorum devices. See [“Dynamic Reconfiguration With Quorum Devices” on page 181](#) for more information.



Caution – If the current primary node fails while you are performing the DR operation on a secondary node, cluster availability is impacted. The primary node will have no place to fail over until a new secondary node is provided.

To perform DR operations on global devices, complete the following steps in the order indicated.

TABLE 5-1 Task Map: Dynamic Reconfiguration With Disk and Tape Devices

Task	For Instructions
1. If a DR operation that affects an active device group must be performed on the current primary node, switch the primary and secondary nodes before performing the DR remove operation on the device	“How to Switch the Primary for a Device Group” on page 157
2. Perform the DR removal operation on the device being removed	<i>Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual</i> in the <i>Solaris 10 on Sun Hardware</i> collection.

Veritas Volume Manager Administration Considerations

- For Oracle Solaris Cluster software to maintain the VxVM namespace, you must register any VxVM disk group or volume changes as Oracle Solaris Cluster device group configuration changes. Registering these changes 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. Changing the volume permissions, owner, or group ID also impacts the namespace.

Note – Never import or deport VxVM disk groups by using VxVM commands after the disk group has been registered with the cluster as an Oracle Solaris Cluster device group. The Oracle Solaris Cluster software handles all cases where disk groups need to be imported or be 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, the minor number is sufficient to guarantee uniqueness. The minor number for a newly created disk group might conflict with the minor number of a preexisting disk group that was imported on a different node. In this case, attempting to register the Oracle Solaris

Cluster device group fails. To fix this problem, the new disk group should be given a new minor number that is a unique value and then registered as an Oracle Solaris Cluster device group.

- If you are setting up a mirrored volume, Dirty Region Logging (DRL) can be used to decrease volume recovery time after a node failure. Use of DRL is strongly recommended, although use of DRL could decrease I/O throughput.
- VxVM does not support the `chmod` command. To change global device permissions in VxVM, consult the VxVM administrator's guide.
- Oracle Solaris Cluster 3.3 software does not support the VxVM Dynamic Multipathing (DMP) management of multiple paths from the same node.
- If you use VxVM to set up shared disk groups for Oracle RAC, use the cluster functionality of VxVM as described in the *Veritas Volume Manager Administrator's Reference Guide*. Creating shared disk groups for Oracle RAC differs from creating other disk groups. You must import the Oracle RAC shared disk groups by using `vxvdg -s`. You do not register the Oracle RAC shared disk groups with the cluster framework. To create other VxVM disk groups, see “[How to Create a New Disk Group When Initializing Disks \(Veritas Volume Manager\)](#)” on page 131.

Administering Storage-Based Replicated Devices

You can configure an Oracle Solaris Cluster device group to contain devices that are replicated by using storage-based replication. Oracle Solaris Cluster software supports Hitachi TrueCopy and EMC Symmetrix Remote Data Facility software for storage-based replication.

Before you can replicate data with Hitachi TrueCopy or EMC Symmetrix Remote Data Facility software, you must be familiar with the storage-based replication documentation and have the storage-based replication product and the latest patches installed on your system. For information about installing the storage-based replication software, see the product documentation.

The storage-based replication software configures a pair of devices as replicas with one device as the primary replica and the other device as the secondary replica. At any given time, the device attached to one set of nodes will be the primary replicas. The device attached to the other set of nodes will be the secondary replica.

In an Oracle Solaris Cluster configuration, the primary replica is automatically moved whenever the Oracle Solaris Cluster device group to which the replica belongs is moved. Therefore, the replica primary should never be moved in an Oracle Solaris Cluster configuration directly. Rather, the takeover should be accomplished by moving the associated Oracle Solaris Cluster device group.



Caution – The name of the Oracle Solaris Cluster device group that you create (Solaris Volume Manager, Veritas Volume Manager, or raw-disk) must be the same as the name of the replicated device group.

This section contains the following procedures:

- “Administering Hitachi TrueCopy Replicated Devices” on page 95
- “Administering EMC Symmetrix Remote Data Facility Replicated Devices” on page 106

Administering Hitachi TrueCopy Replicated Devices

The following table lists the tasks you must perform to set up an Hitachi TrueCopy storage-based replicated device.

TABLE 5-2 Task Map: Administering an Hitachi TrueCopy Storage-Based Replicate Device

Task	Instructions
Install the TrueCopy software on your storage device and nodes	See the documentation that shipped with your Hitachi storage device.
Configure the Hitachi replication group	“How to Configure a Hitachi TrueCopy Replication Group” on page 95
Configure the DID device	“How to Configure DID Devices for Replication Using Hitachi TrueCopy” on page 97
Register the replicated group	“How to Add and Register a Device Group (Solaris Volume Manager)” on page 127 or “How to Register a Disk Group as a Device Group (Veritas Volume Manager)” on page 139
Verify the configuration	“How to Verify a Hitachi TrueCopy Replicated Global Device Group Configuration” on page 99

▼ How to Configure a Hitachi TrueCopy Replication Group

Before You Begin

First, configure the Hitachi TrueCopy device groups on shared disks in the primary cluster. This configuration information is specified in the `/etc/horcm.conf` file on each of the cluster's nodes that has access to the Hitachi array. For more information about how to configure the `/etc/horcm.conf` file, see the *Sun StorEdge SE 9900 V Series Command and Control Interface User and Reference Guide*.



Caution – The name of the Oracle Solaris Cluster device group that you create (Solaris Volume Manager, Veritas Volume Manager, ZFS, or raw-disk) must be the same as the name of the replicated device group.

- 1 **Become superuser or assume a role that provides `soLaris.cLuster.modifyRBAC` authorization on all nodes connected to the storage array.**

- 2 **Add the `horcm` entry to the `/etc/services` file.**

```
horcm 9970/udp
```

Specify a port number and protocol name for the new entry.

- 3 **Specify the Hitachi TrueCopy device group configuration information in the `/etc/horcm.conf` file.**

For instructions, refer to the documentation that shipped with your TrueCopy software.

- 4 **Start the TrueCopy CCI daemon by running the `horcmstart.sh` command on all nodes.**

```
# /usr/bin/horcmstart.sh
```

- 5 **If you have not already created the replica pairs, create them now.**

Use the `paircreate` command to create your replica pairs with the desired fence level. For instructions on creating the replica pairs, refer to your TrueCopy documentation.

- 6 **On each node configured with replicated devices, verify that data replication is set up correctly by using the `pairdisplay` command. A Hitachi TrueCopy or Hitachi Universal Replicator device group with a `fence_level` of `ASYNCR` can not share its `ctgid` with any other device groups on the system.**

```
# pairdisplay -g group-name
Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#,P/S,Status,Fence,Seq#,P-LDEV# M
group-name pair1(L) (CL1-C , 0, 9) 54321 58..P-VOL PAIR NEVER ,12345 29 -
group-name pair1(R) (CL1-A , 0, 29)12345 29..S-VOL PAIR NEVER ,----- 58 -
```

- 7 **Verify that all nodes can master the replication groups.**

- a. **Determine which node contains the primary replica and which node contains the secondary replica by using the `pairdisplay` command.**

```
# pairdisplay -g group-name
Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#,P/S,Status,Fence,Seq#,P-LDEV# M
group-name pair1(L) (CL1-C , 0, 9) 54321 58..P-VOL PAIR NEVER ,12345 29 -
group-name pair1(R) (CL1-A , 0, 29)12345 29..S-VOL PAIR NEVER ,----- 58 -
```

The node with the local (L) device in the P-VOL state contains the primary replica and the node with the local (L) device in the S-VOL state contains the secondary replica.

- b. **Make the secondary node the master by running the `horctakeover` command on the node that contains the secondary replica.**

```
# horctakeover -g group-name
```

Wait for the initial data copy to complete before proceeding to the next step.

- c. **Verify that the node that performed the `horctakeover` now has the local (L) device in the P-VOL state.**

```
# pairdisplay -g group-name
Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#,P/S,Status,Fence,Seq#,P-LDEV# M
group-name pair1(L) (CL1-C , 0, 9) 54321 58..S-VOL PAIR NEVER ,12345 29 -
group-name pair1(R) (CL1-A , 0, 29)12345 29..P-VOL PAIR NEVER ,----- 58 -
```

- d. **Run the `horctakeover` command on the node that originally contained the primary replica.**

```
# horctakeover -g group-name
```

- e. **Verify that the primary node has changed back to the original configuration by running the `pairdisplay` command.**

```
# pairdisplay -g group-name
Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#,P/S,Status,Fence,Seq#,P-LDEV# M
group-name pair1(L) (CL1-C , 0, 9) 54321 58..P-VOL PAIR NEVER ,12345 29 -
group-name pair1(R) (CL1-A , 0, 29)12345 29..S-VOL PAIR NEVER ,----- 58 -
```

Next Steps Continue the configuration of your replicated device by following the instructions in [“How to Configure DID Devices for Replication Using Hitachi TrueCopy”](#) on page 97.

▼ How to Configure DID Devices for Replication Using Hitachi TrueCopy

Before You Begin After you have configured a device group for your replicated device, you must configure the device identifier (DID) driver that the replicated device uses.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the cluster.**
- 2 Verify that the `horcmd` daemon is running on all nodes.**

The following command will start the daemon if it is not running. The system will display a message if the daemon is already running.

```
# /usr/bin/horcmstart.sh
```

3 Determine which node contains the secondary replica by running the `pairdisplay` command.

```
# pairdisplay -g group-name
Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#,P/S,Status,Fence,Seq#,P-LDEV# M
group-name pair1(L) (CL1-C , 0, 9) 54321 58..P-VOL PAIR NEVER ,12345 29 -
group-name pair1(R) (CL1-A , 0, 29)12345 29..S-VOL PAIR NEVER ,----- 58 -
```

The node with the local (L) device in the S-VOL state contains the secondary replica.

4 On the node with secondary replica (as determined by the previous step), configure the DID devices for use with storage-based replication.

This command combines the two separate DID instances for the device replica pairs into a single, logical DID instance. The single instance enables the device to be used by volume management software from both sides.



Caution – If multiple nodes are connected to the secondary replica, run this command on only one of these nodes.

```
# cldevice replicate -D primary-replica-nodename -S secondary-replica-nodename
```

primary-replica-nodename

Specifies the name of the remote node that contains the primary replica.

-S

Specifies a source node other than the current node.

secondary-replica-nodename

Specifies the name of the remote node that contains the secondary replica.

Note – By default, the current node is the source node. Use the -S option to specify a different source node.

5 Verify that the DID instances have been combined.

```
# cldevice list -v logical_DID_device
```

6 Verify that the TrueCopy replication is set.

```
# cldevice show logical_DID_device
```

The command output should indicate that TrueCopy is the replication type.

7 If the DID remapping did not successfully combine all replicated devices, combine the individual replicated devices manually.

Caution – Exercise extreme care when combining DID instances manually. Improper device remapping can cause data corruption.

- a. On all nodes that contains the secondary replica, run the `cldevice combine` command.

```
# cldevice combine -d destination-instance source-instance
```

-d The remote DID instance, which corresponds to the primary replica.
destination-instance

source-instance The local DID instance, which corresponds to the secondary replica.

- b. Verify that the DID remapping occurred successfully.

```
# cldevice list destination-instance source-instance
```

One of the DID instances should not be listed.

- 8 On all nodes, verify that the DID devices for all combined DID instances are accessible.

```
# cldevice list -v
```

Next Steps To complete the configuration of your replicated device group, perform the steps in the following procedures.

- “How to Add and Register a Device Group (Solaris Volume Manager)” on page 127 or “How to Register a Disk Group as a Device Group (Veritas Volume Manager)” on page 139

When registering the device group, make sure to give it the same name as the TrueCopy replication group.

- “How to Verify a Hitachi TrueCopy Replicated Global Device Group Configuration” on page 99

▼ How to Verify a Hitachi TrueCopy Replicated Global Device Group Configuration

Before You Begin Before you verify the global device group, you must first create it. You can use device groups from Solaris Volume Manager, Veritas Volume Manager, ZFS, or raw-disk. For more information, consult the following:

- “How to Add and Register a Device Group (Solaris Volume Manager)” on page 127
- “How to Add and Register a Device Group (Raw-Disk)” on page 129
- “How to Add and Register a Replicated Device Group (ZFS)” on page 130
- “How to Create a New Disk Group When Initializing Disks (Veritas Volume Manager)” on page 131



Caution – The name of the Oracle Solaris Cluster device group that you created (Solaris Volume Manager, Veritas Volume Manager, or raw-disk) must be the same as the name of the replicated device group.

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

1 Verify that the primary device group corresponds to the same node as the node that contains the primary replica.

```
# pairdisplay -g group-name
# cldevicegroup status -n nodename group-name
```

2 Verify that the replication property is set for the device group.

```
# cldevicegroup show -n nodename group-name
```

3 Verify that the replicated property is set for the device.

```
# usr/cluster/bin/cldevice status [-s state] [-n node[,?]] [+| [disk-device ]]
```

4 Perform a trial switchover to ensure that the device groups are configured correctly and the replicas can move between nodes.

If the device group is offline, bring it online.

```
# cldevicegroup switch -n nodename group-name
```

`-n nodename` The node to which the device group is switched. This node becomes the new primary

5 Verify that the switchover was successful by comparing the output of the following commands.

```
# pairdisplay -g group-name
# cldevicegroup status -n nodename group-name
```

Example: Configuring a TrueCopy Replication Group for Oracle Solaris Cluster

This example completes the Oracle Solaris Cluster specific steps necessary to set up TrueCopy replication in your cluster. The example assumes that you have already performed the following tasks:

- Set up your Hitachi LUNs
- Installed the TrueCopy software on your storage device and cluster nodes
- Configured the replication pairs on your cluster nodes

For instructions about configuring your replication pairs, see [“How to Configure a Hitachi TrueCopy Replication Group” on page 95](#).

This example involves a three-node cluster that uses TrueCopy. The cluster is spread across two remote sites, with two nodes at one site and one node at the other site. Each site has its own Hitachi storage device.

The following examples show the TrueCopy `/etc/horcm.conf` configuration file on each node.

EXAMPLE 5-1 TrueCopy Configuration File on Node 1

```
HORCM_DEV
#dev_group  dev_name  port#      TargetID   LU#        MU#
VG01        pair1     CL1-A      0           29
VG01        pair2     CL1-A      0           30
VG01        pair3     CL1-A      0           31
HORCM_INST
#dev_group  ip_address  service
VG01        node-3     horcm
```

EXAMPLE 5-2 TrueCopy Configuration File on Node 2

```
HORCM_DEV
#dev_group  dev_name  port#      TargetID   LU#        MU#
VG01        pair1     CL1-A      0           29
VG01        pair2     CL1-A      0           30
VG01        pair3     CL1-A      0           31
HORCM_INST
#dev_group  ip_address  service
VG01        node-3     horcm
```

EXAMPLE 5-3 TrueCopy Configuration File on Node 3

```
HORCM_DEV
#dev_group  dev_name  port#      TargetID   LU#        MU#
VG01        pair1     CL1-C      0           09
VG01        pair2     CL1-C      0           10
VG01        pair3     CL1-C      0           11
HORCM_INST
#dev_group  ip_address  service
VG01        node-1     horcm
VG01        node-2     horcm
```

In the preceding examples, three LUNs are replicated between the two sites. The LUNs are all in a replication group named VG01. The `pairdisplay` command verifies this information and shows that Node 3 has the primary replica.

EXAMPLE 5-4 pairdisplay Command Output on Node 1

```
# pairdisplay -g VG01
Group  PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV# M
VG01   pair1(L)     (CL1-A , 0, 29)61114 29..S-VOL PAIR NEVER ,----- 58 -
VG01   pair1(R)     (CL1-C , 0, 9)20064 58..P-VOL PAIR NEVER ,61114 29 -
VG01   pair2(L)     (CL1-A , 0, 30)61114 30..S-VOL PAIR NEVER ,----- 59 -
VG01   pair2(R)     (CL1-C , 0, 10)20064 59..P-VOL PAIR NEVER ,61114 30 -
VG01   pair3(L)     (CL1-A , 0, 31)61114 31..S-VOL PAIR NEVER ,----- 60 -
```

EXAMPLE 5-4 pairdisplay Command Output on Node 1 (Continued)

```
VG01 pair3(R) (CL1-C , 0, 11)20064 60..P-VOL PAIR NEVER ,61114 31 -
```

EXAMPLE 5-5 pairdisplay Command Output on Node 2

```
# pairdisplay -g VG01
Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV# M
VG01 pair1(L) (CL1-A , 0, 29)61114 29..S-VOL PAIR NEVER ,----- 58 -
VG01 pair1(R) (CL1-C , 0, 9)20064 58..P-VOL PAIR NEVER ,61114 29 -
VG01 pair2(L) (CL1-A , 0, 30)61114 30..S-VOL PAIR NEVER ,----- 59 -
VG01 pair2(R) (CL1-C , 0, 10)20064 59..P-VOL PAIR NEVER ,61114 30 -
VG01 pair3(L) (CL1-A , 0, 31)61114 31..S-VOL PAIR NEVER ,----- 60 -
VG01 pair3(R) (CL1-C , 0, 11)20064 60..P-VOL PAIR NEVER ,61114 31 -
```

EXAMPLE 5-6 pairdisplay Command Output on Node 3

```
# pairdisplay -g VG01
Group PairVol(L/R) (Port#,TID,LU),Seq#,LDEV#.P/S,Status,Fence, Seq#,P-LDEV# M
VG01 pair1(L) (CL1-C , 0, 9)20064 58..P-VOL PAIR NEVER ,61114 29 -
VG01 pair1(R) (CL1-A , 0, 29)61114 29..S-VOL PAIR NEVER ,----- 58 -
VG01 pair2(L) (CL1-C , 0, 10)20064 59..P-VOL PAIR NEVER ,61114 30 -
VG01 pair2(R) (CL1-A , 0, 30)61114 30..S-VOL PAIR NEVER ,----- 59 -
VG01 pair3(L) (CL1-C , 0, 11)20064 60..P-VOL PAIR NEVER ,61114 31 -
VG01 pair3(R) (CL1-A , 0, 31)61114 31..S-VOL PAIR NEVER ,----- 60 -
```

To see which disks are being used, use the `-fd` option of the `pairdisplay` command as shown in the following examples.

EXAMPLE 5-7 pairdisplay Command Output on Node 1, Showing Disks Used

```
# pairdisplay -fd -g VG01
Group PairVol(L/R) Device File ,Seq#,LDEV#.P/S,Status,Fence,Seq#,P-LDEV# M
VG01 pair1(L) c6t500060E800000000000000E0EBA0000001Dd0s2 61114 29..S-VOL PAIR NEVER ,----- 58 -
VG01 pair1(R) c5t500060E8000000000000004E600000003Ad0s2 20064 58..P-VOL PAIR NEVER ,61114 29 -
VG01 pair2(L) c6t500060E800000000000000E0EBA0000001Ed0s2 61114 30..S-VOL PAIR NEVER ,----- 59 -
VG01 pair2(R) c5t500060E8000000000000004E600000003Bd0s2 0064 59..P-VOL PAIR NEVER ,61114 30 -
VG01 pair3(L) c6t500060E800000000000000E0EBA0000001Fd0s2 61114 31..S-VOL PAIR NEVER ,----- 60 -
VG01 pair3(R) c5t500060E8000000000000004E600000003Cd0s2 20064 60..P-VOL PAIR NEVER ,61114 31 -
```

EXAMPLE 5-8 pairdisplay Command Output on Node 2, Showing Disks Used

```
# pairdisplay -fd -g VG01
Group PairVol(L/R) Device File ,Seq#,LDEV#.P/S,Status,Fence,Seq#,P-LDEV# M
VG01 pair1(L) c5t500060E800000000000000E0EBA0000001Dd0s2 61114 29..S-VOL PAIR NEVER ,----- 58 -
VG01 pair1(R) c5t500060E8000000000000004E600000003Ad0s2 20064 58..P-VOL PAIR NEVER ,61114 29 -
VG01 pair2(L) c5t500060E800000000000000E0EBA0000001Ed0s2 61114 30..S-VOL PAIR NEVER ,----- 59 -
VG01 pair2(R) c5t500060E8000000000000004E600000003Bd0s2 20064 59..P-VOL PAIR NEVER ,61114 30 -
VG01 pair3(L) c5t500060E800000000000000E0EBA0000001Fd0s2 61114 31..S-VOL PAIR NEVER ,----- 60 -
VG01 pair3(R) c5t500060E8000000000000004E600000003Cd0s2 20064 60..P-VOL PAIR NEVER ,61114 31 -
```

EXAMPLE 5-9 pairdisplay Command Output on Node 3, Showing Disks Used

```
# pairdisplay -fd -g VG01
Group PairVol(L/R) Device File                               ,Seq#,LDEV#.P/S,Status,Fence ,Seq#,P-LDEV# M
VG01 pair1(L) c5t50060E8000000000000004E600000003Ad0s2 20064 58..P-VOL PAIR NEVER ,61114 29 -
VG01 pair1(R) c6t50060E800000000000000EEBA0000001Dd0s2 61114 29..S-VOL PAIR NEVER ,----- 58 -
VG01 pair2(L) c5t50060E8000000000000004E600000003Bd0s2 20064 59..P-VOL PAIR NEVER ,61114 30 -
VG01 pair2(R) c6t50060E800000000000000EEBA0000001Ed0s2 61114 30..S-VOL PAIR NEVER ,----- 59 -
VG01 pair3(L) c5t50060E8000000000000004E600000003Cd0s2 20064 60..P-VOL PAIR NEVER ,61114 31 -
VG01 pair3(R) c6t50060E800000000000000EEBA0000001Fd0s2 61114 31..S-VOL PAIR NEVER ,----- 60 -
```

These examples show that the following disks are being used:

- On Node 1:
 - c6t50060E800000000000000EEBA0000001Dd0s2
 - c6t50060E800000000000000EEBA0000001Ed0s2
 - c6t50060E800000000000000EEBA0000001Fd0s2
- On Node 2:
 - c5t50060E800000000000000EEBA0000001Dd0s2
 - c5t50060E800000000000000EEBA0000001Ed0s2
 - c5t50060E800000000000000EEBA0000001Fd0s2
- On Node 3:
 - c5t50060E8000000000000004E600000003Ad0s2
 - c5t50060E8000000000000004E600000003Bd0s2
 - c5t50060E8000000000000004E600000003Cd0s2

To see the DID devices that corresponds to these disks, use the `cldevice list` command as shown in the following examples.

EXAMPLE 5-10 Displaying DIDs Corresponding to the Disks Used

```
# cldevice list -v
```

```
DID Device Full Device Path
-----
1 node-1:/dev/rdisk/c0t0d0 /dev/did/rdisk/d1
2 node-1:/dev/rdisk/c0t6d0 /dev/did/rdisk/d2
11 node-1:/dev/rdisk/c6t50060E800000000000000EEBA00000020d0 /dev/did/rdisk/d11
11 node-2:/dev/rdisk/c5t50060E800000000000000EEBA00000020d0 /dev/did/rdisk/d11
12 node-1:/dev/rdisk/c6t50060E800000000000000EEBA0000001Fd0 /dev/did/rdisk/d12
12 node-2:/dev/rdisk/c5t50060E800000000000000EEBA0000001Fd0 /dev/did/rdisk/d12
13 node-1:/dev/rdisk/c6t50060E800000000000000EEBA0000001Ed0 /dev/did/rdisk/d13
13 node-2:/dev/rdisk/c5t50060E800000000000000EEBA0000001Ed0 /dev/did/rdisk/d13
14 node-1:/dev/rdisk/c6t50060E800000000000000EEBA0000001Dd0 /dev/did/rdisk/d14
14 node-2:/dev/rdisk/c5t50060E800000000000000EEBA0000001Dd0 /dev/did/rdisk/d14
18 node-3:/dev/rdisk/c0t0d0 /dev/did/rdisk/d18
19 node-3:/dev/rdisk/c0t6d0 /dev/did/rdisk/d19
20 node-3:/dev/rdisk/c5t50060E8000000000000004E6000000013d0 /dev/did/rdisk/d20
21 node-3:/dev/rdisk/c5t50060E8000000000000004E600000003Dd0 /dev/did/rdisk/d21
22 node-3:/dev/rdisk/c5t50060E8000000000000004E600000003Cd0 /dev/did/rdisk/d2223
```

EXAMPLE 5-10 Displaying DIDs Corresponding to the Disks Used *(Continued)*

```

23 node-3:/dev/rdisk/c5t50060E800000000000004E600000003Bd0 /dev/did/rdsk/d23
24 node-3:/dev/rdisk/c5t50060E800000000000004E600000003Ad0 /dev/did/rdsk/d24

```

When combining the DID instances for each pair of replicated devices, `cldevice list` should combine DID instance 12 with 22, instance 13 with 23 and instance 14 with 24. Because Node 3 has the primary replica, run the `cldevice -T` command from either Node 1 or Node 2. Always combine the instances from a node that has the secondary replica. Run this command from a single node only, not on both nodes.

The following example shows the output when combining DID instances by running the command on Node 1.

EXAMPLE 5-11 Combining DID Instances

```

# cldevice replicate -D node-3
Remapping instances for devices replicated with node-3...
VG01 pair1 L node-1:/dev/rdsk/c6t50060E80000000000000EEBA0000001Dd0
VG01 pair1 R node-3:/dev/rdsk/c5t50060E800000000000004E600000003Ad0
Combining instance 14 with 24
VG01 pair2 L node-1:/dev/rdsk/c6t50060E80000000000000EEBA0000001Ed0
VG01 pair2 R node-3:/dev/rdsk/c5t50060E800000000000004E600000003Bd0
Combining instance 13 with 23
VG01 pair3 L node-1:/dev/rdsk/c6t50060E80000000000000EEBA0000001Fd0
VG01 pair3 R node-3:/dev/rdsk/c5t50060E800000000000004E600000003Cd0
Combining instance 12 with 22

```

Checking the `cldevice list` output, the LUNs from both sites now have the same DID instance. Having the same DID instance makes each replica pair look like a single DID device, as the following example shows.

EXAMPLE 5-12 Displaying the Combined DIDs

```

# cldevice list -v
DID Device Full Device Path
-----
1 node-1:/dev/rdsk/c0t0d0 /dev/did/rdsk/d1
2 node-1:/dev/rdsk/c0t6d0 /dev/did/rdsk/d2
11 node-1:/dev/rdsk/c6t50060E80000000000000EEBA00000020d0 /dev/did/rdsk/d11
11 node-2:/dev/rdsk/c5t50060E80000000000000EEBA00000020d0 /dev/did/rdsk/d11
18 node-3:/dev/rdsk/c0t0d0 /dev/did/rdsk/d18
19 node-3:/dev/rdsk/c0t6d0 /dev/did/rdsk/d19
20 node-3:/dev/rdsk/c5t50060E800000000000004E6000000013d0 /dev/did/rdsk/d20
21 node-3:/dev/rdsk/c5t50060E800000000000004E600000003Dd0 /dev/did/rdsk/d21
22 node-1:/dev/rdsk/c6t50060E80000000000000EEBA0000001Fd0 /dev/did/rdsk/d122
22 node-2:/dev/rdsk/c5t50060E80000000000000EEBA0000001Fd0 /dev/did/rdsk/d12
22 node-3:/dev/rdsk/c5t50060E800000000000004E600000003Cd0 /dev/did/rdsk/d22
23 node-1:/dev/rdsk/c6t50060E80000000000000EEBA0000001Ed0 /dev/did/rdsk/d13
23 node-2:/dev/rdsk/c5t50060E80000000000000EEBA0000001Ed0 /dev/did/rdsk/d13
23 node-3:/dev/rdsk/c5t50060E800000000000004E600000003Bd0 /dev/did/rdsk/d23

```


Configuration of the replication group is complete with this step. To verify that the configuration was successful, perform the steps in [“How to Verify a Hitachi TrueCopy Replicated Global Device Group Configuration”](#) on page 99.

Administering EMC Symmetrix Remote Data Facility Replicated Devices

The following table lists the tasks you must perform to set up and manage an EMC Symmetrix Remote Data Facility (SRDF) storage-based replicated device.

TABLE 5-3 Task Map: Administering an EMC SRDF Storage-Based Replicated Device

Task	Instructions
Install the SRDF software on your storage device and nodes	The documentation that shipped with your EMC storage device.
Configure the EMC replication group	“How to Configure an EMC SRDF Replication Group” on page 106
Configure the DID device	“How to Configure DID Devices for Replication Using EMC SRDF” on page 108
Register the replicated group	“How to Add and Register a Device Group (Solaris Volume Manager)” on page 127 or “How to Register a Disk Group as a Device Group (Veritas Volume Manager)” on page 139
Verify the configuration	“How to Verify EMC SRDF Replicated Global Device Group Configuration” on page 109
Manually recover data after a campus cluster's primary room completely fails	“How to Recover EMC SRDF Data after a Primary Room's Complete Failure” on page 115

▼ How to Configure an EMC SRDF Replication Group

Before You Begin EMC Solutions Enabler software must be installed on all cluster nodes before you configure an EMC Symmetrix Remote Data Facility (SRDF) replication group. First, configure the EMC SRDF device groups on shared disks in the cluster. For more information about how to configure the EMC SRDF device groups, see your EMC SRDF product documentation.

When using EMC SRDF, use dynamic devices instead of static devices. Static devices require several minutes to change the replication primary and can impact failover time.



Caution – The name of the Oracle Solaris Cluster device group that you create (Solaris Volume Manager, Veritas Volume Manager, or raw-disk) must be the same as the name of the replicated device group.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on all nodes connected to the storage array.**
- 2 **On each node configured with the replicated data, discover the symmetrix device configuration.**
This might take a few minutes.
`# /usr/symcli/bin/symcfg discover`
- 3 **If you have not already created the replica pairs, create them now.**
Use the `symrdf` command to create your replica pairs. For instructions on creating the replica pairs, refer to your SRDF documentation.
- 4 **On each node configured with replicated devices, verify that data replication is set up correctly.**
`# /usr/symcli/bin/symdg show group-name`
- 5 **Perform a swap of the device group.**
 - a. **Verify that the primary and secondary replicas are synchronized.**
`# /usr/symcli/bin/symrdf -g group-name verify -synchronized`
 - b. **Determine which node contains the primary replica and which node contains the secondary replica by using the `symdg show` command.**
`# /usr/symcli/bin/symdg show group-name`
The node with the RDF1 device contains the primary replica and the node with the RDF2 device state contains the secondary replica.
 - c. **Enable the secondary replica.**
`# /usr/symcli/bin/symrdf -g group-name failover`
 - d. **Swap the RDF1 and RDF2 devices.**
`# /usr/symcli/bin/symrdf -g group-name swap -refresh R1`
 - e. **Enable the replica pair.**
`# /usr/symcli/bin/symrdf -g group-name establish`
 - f. **Verify that the primary node and secondary replicas are synchronized.**
`# /usr/symcli/bin/symrdf -g group-name verify -synchronized`
- 6 **Repeat all of step 5 on the node which originally had the primary replica.**

Next Steps After you have configured a device group for your EMC SRDF replicated device, you must configure the device identifier (DID) driver that the replicated device uses.

▼ How to Configure DID Devices for Replication Using EMC SRDF

This procedure configures the device identifier (DID) driver that the replicated device uses.

Before You Begin The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the cluster.**
- 2 **Determine which DID devices correspond to the configured RDF1 and RDF2 devices.**

```
# /usr/symcli/bin/symsg show group-name
```

Note – If your system does not display the entire Oracle Solaris device patch, set the environment variable `SYMCLI_FULL_PDEVNAME` to 1 and retype the `symsg -show` command.

- 3 **Determine which DID devices correspond to the Oracle Solaris devices.**

```
# cldevice list -v
```

- 4 **For each pair of matched DID devices, combine the instances into a single replicated DID device. Run the following command from the RDF2/secondary side.**

```
# cldevice combine -t srdf -g replication-device-group \  
-d destination-instance source-instance
```

Note – The `-T` option is not supported for SRDF data replication devices.

<code>-t replication-type</code>	Specifies the replication type. For EMC SRDF, type SRDF .
<code>-g replication-device-group</code>	Specifies the name of the device group as shown in the <code>symsg show</code> command.
<code>-d destination-instance</code>	Specifies the DID instance that corresponds to the RDF1 device.
<code>source-instance</code>	Specifies the DID instance that corresponds to the RDF2 device.

Note – If you combine the wrong DID device, use the `-b` option for the `scdidadm` command to undo the combining of two DID devices.

```
# scdidadm -b device
```

```
-b device    The DID instance that corresponded to the destination_device when the instances
             were combined.
```

- 5 If the name of a replication device group changes, additional steps are required for Hitachi TrueCopy and SRDF. After you complete steps 1 through 4, perform the appropriate additional step.**

Item	Description
TrueCopy	If the name of the replication device group (and the corresponding global device group) changes, you must rerun the <code>cldevice replicate</code> command to update the replicated device information.
SRDF	If the name of the replication device group (and the corresponding global device group) changes, you must update the replicated device information by first using the <code>scdidadm -b</code> command to remove the existing information. The last step is to use the <code>cldevice combine</code> command to create a new, updated device.

- 6 Verify that the DID instances have been combined.**

```
# cldevice list -v device
```

- 7 Verify that the SRDF replication is set.**

```
# cldevice show device
```

- 8 On all nodes, verify that the DID devices for all combined DID instances are accessible.**

```
# cldevice list -v
```

Next Steps After you have configured the device identifier (DID) driver that the replicated device uses, you must verify the EMC SRDF replicated global device group configuration.

▼ How to Verify EMC SRDF Replicated Global Device Group Configuration

Before You Begin Before you verify the global device group, you must first create it. You can use device groups from Solaris Volume Manager, Veritas Volume Manager, ZFS, or raw-disk. For more information, consult the following:

- [“How to Add and Register a Device Group \(Solaris Volume Manager\)” on page 127](#)
- [“How to Add and Register a Device Group \(Raw-Disk\)” on page 129](#)
- [“How to Add and Register a Replicated Device Group \(ZFS\)” on page 130](#)

- “How to Create a New Disk Group When Initializing Disks (Veritas Volume Manager)” on page 131



Caution – The name of the Oracle Solaris Cluster device group that you created (Solaris Volume Manager, Veritas Volume Manager, or raw-disk) must be the same as the name of the replicated device group.

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Verify that the primary device group corresponds to the same node as the node that contains the primary replica.**

```
# symdbg -show group-name
# cldevicegroup status -n nodename group-name
```

- 2 **Perform a trial switchover to ensure that the device groups are configured correctly and the replicas can move between nodes.**

If the device group is offline, bring it online.

```
# cldevicegroup switch -n nodename group-name
```

-n *nodename* The node to which the device group is switched. This node becomes the new primary.

- 3 **Verify that the switchover was successful by comparing the output of the following commands.**

```
# symdbg -show group-name
# cldevicegroup status -n nodename group-name
```

Example: Configuring an SRDF Replication Group for Oracle Solaris Cluster

This example completes the Oracle Solaris Cluster specific steps necessary to set up SRDF replication in your cluster. The example assumes that you have already performed the following tasks:

- Completed pairing LUNS for replication between arrays.
- Installed the SRDF software on your storage device and cluster nodes.

This example involves a four-node cluster where two nodes are connected to one symmetrix and the other two nodes are connected to the second symmetrix. The SRDF device group is called `dg1`.

EXAMPLE 5-15 Creating Replica Pairs

Run the following command on all nodes.

```
# symcfg discover
! This operation might take up to a few minutes.
# symdev list pd
```

Symmetrix ID: 000187990182

Device Name		Directors			Device		
Sym	Physical	SA	:P DA	:IT Config	Attribute	Sts	Cap (MB)
0067	c5t600604800001879901*	16D:0	02A:C1	RDF2+Mir	N/Grp'd	RW	4315
0068	c5t600604800001879901*	16D:0	16B:C0	RDF1+Mir	N/Grp'd	RW	4315
0069	c5t600604800001879901*	16D:0	01A:C0	RDF1+Mir	N/Grp'd	RW	4315
...							

On all nodes on the RDF1 side, type:

```
# symdg -type RDF1 create dg1
# symld -g dg1 add dev 0067
```

On all nodes on the RDF2 side, type:

```
# symdg -type RDF2 create dg1
# symld -g dg1 add dev 0067
```

EXAMPLE 5-16 Verifying Data Replication Setup

From one node in the cluster, type:

```
# symdg show dg1
```

Group Name: dg1

```
Group Type                : RDF1      (RDFA)
Device Group in GNS       : No
Valid                      : Yes
Symmetrix ID              : 000187900023
Group Creation Time       : Thu Sep 13 13:21:15 2007
Vendor ID                  : EMC Corp
Application ID            : SYMCLI
```

```
Number of STD Devices in Group : 1
Number of Associated GK's      : 0
Number of Locally-associated BCV's : 0
Number of Locally-associated VDEV's : 0
Number of Remotely-associated BCV's (STD RDF): 0
Number of Remotely-associated BCV's (BCV RDF): 0
Number of Remotely-associ'd RBCV's (RBCV RDF) : 0
```

EXAMPLE 5-16 Verifying Data Replication Setup (Continued)

Standard (STD) Devices (1):

```

{
-----
LdevName          PdevName          Sym          Cap
Dev              Dev Att. Sts      (MB)
-----
DEV001            /dev/rdisk/c5t6006048000018790002353594D303637d0s2 0067      RW      4315
}
    
```

Device Group RDF Information

```

...
# symrdf -g dg1 establish
    
```

Execute an RDF 'Incremental Establish' operation for device group 'dg1' (y/[n]) ? y

An RDF 'Incremental Establish' operation execution is in progress for device group 'dg1'. Please wait...

```

Write Disable device(s) on RA at target (R2).....Done.
Suspend RDF link(s).....Done.
Mark target (R2) devices to refresh from source (R1).....Started.
Device: 0067 ..... Marked.
Mark target (R2) devices to refresh from source (R1).....Done.
Merge device track tables between source and target.....Started.
Device: 0067 ..... Merged.
Merge device track tables between source and target.....Done.
Resume RDF link(s).....Started.
Resume RDF link(s).....Done.
    
```

The RDF 'Incremental Establish' operation successfully initiated for device group 'dg1'.

```

#
# symrdf -g dg1 query
    
```

```

Device Group (DG) Name      : dg1
DG's Type                   : RDF2
DG's Symmetrix ID          : 000187990182
    
```

Target (R2) View					Source (R1) View				MODES	
Standard	Logical	Device	Dev	ST	LI	ST	Dev	Dev	MDA	STATE
A	T	E		A	N	T	E			
	R1 Inv	R2 Inv		R1 Inv	R2 Inv	R1 Inv	R2 Inv			
	Tracks	Tracks		Tracks	Tracks	Tracks	Tracks			
DEV001	0067	WD	0	0	RW	0067	RW	0	0	S.. Synchronized
Total	-----		-----		-----		-----			
MB(s)	0.0	0.0			0.0	0.0				

EXAMPLE 5-16 Verifying Data Replication Setup (Continued)

Legend for MODES:

M(ode of Operation): A = Async, S = Sync, E = Semi-sync, C = Adaptive Copy
 D(omino) : X = Enabled, . = Disabled
 A(daptive Copy) : D = Disk Mode, W = WP Mode, . = ACp off

#

EXAMPLE 5-17 Displaying DIDs Corresponding to the Disks Used

The same procedure applies to the RDF1 and RDF2 sides.

You can look under the PdevName field of output of the `dymdg show dg` command.

On the RDF1 side, type:

```
# symdg show dg1
Group Name: dg1
Group Type : RDF1 (RDFA)
...
Standard (STD) Devices (1):
{
-----
LdevName          PdevName          Sym          Cap
Dev Att. Sts      (MB)
-----
DEV001            /dev/rdisk/c5t6006048000018790002353594D303637d0s2 0067      RW      4315
}
Device Group RDF Information
...
```

To obtain the corresponding DID, type:

```
# scdidadm -L | grep c5t6006048000018790002353594D303637d0
217      pmoney1:/dev/rdisk/c5t6006048000018790002353594D303637d0 /dev/did/rdisk/d217
217      pmoney2:/dev/rdisk/c5t6006048000018790002353594D303637d0 /dev/did/rdisk/d217
#
```

To list the corresponding DID, type:

```
# cldevice show d217
=== DID Device Instances ===
DID Device Name: /dev/did/rdisk/d217
Full Device Path: pmoney2:/dev/rdisk/c5t6006048000018790002353594D303637d0
Full Device Path: pmoney1:/dev/rdisk/c5t6006048000018790002353594D303637d0
Replication: none
default_fencing: global
```

EXAMPLE 5-17 Displaying DIDs Corresponding to the Disks Used *(Continued)*

#

On the RDF2 side, type:

You can look under the PdevName field of output of `dymdg show dg` command.

`symdg show dg1`

Group Name: dg1

Group Type : RDF2 (RDFA)

...

Standard (STD) Devices (1):

{

LdevName	PdevName	Sym Dev	Att.	Sts	Cap (MB)
DEV001	/dev/rdisk/c5t6006048000018799018253594D303637d0	0067	WD		4315

}

Device Group RDF Information

...

To obtain the corresponding DID, type:

```
# scdidadm -L | grep c5t6006048000018799018253594D303637d0
108    pmoney4:/dev/rdisk/c5t6006048000018799018253594D303637d0 /dev/did/rdisk/d108
108    pmoney3:/dev/rdisk/c5t6006048000018799018253594D303637d0 /dev/did/rdisk/d108
#
```

To list the corresponding DID, type:

`cldevice show d108`

=== DID Device Instances ===

```
DID Device Name:      /dev/did/rdisk/d108
Full Device Path:    pmoney3:/dev/rdisk/c5t6006048000018799018253594D303637d0
Full Device Path:    pmoney4:/dev/rdisk/c5t6006048000018799018253594D303637d0
Replication:         none
default_fencing:     global
```

#

EXAMPLE 5-18 Combining DID instances

From the RDF2 side, type:

```
# cldevice combine -t srdf -g dg1 -d d217 d108
#
```

EXAMPLE 5-19 Displaying the Combined DIDs

From any node in the cluster, type:

```
# cldevice show d217 d108
cldevice: (C727402) Could not locate instance "108".

=== DID Device Instances ===

DID Device Name:                /dev/did/rdisk/d217
Full Device Path:                pmoney1:/dev/rdisk/c5t6006048000018790002353594D303637d0
Full Device Path:                pmoney2:/dev/rdisk/c5t6006048000018790002353594D303637d0
Full Device Path:                pmoney4:/dev/rdisk/c5t6006048000018799018253594D303637d0
Full Device Path:                pmoney3:/dev/rdisk/c5t6006048000018799018253594D303637d0
Replication:                    srdf
default_fencing:                 global

#
```

▼ How to Recover EMC SRDF Data after a Primary Room's Complete Failure

This procedure performs data recovery when a campus cluster's primary room fails completely, the primary room fails over to a secondary room, and then the primary room comes back online. The campus cluster's primary room is the primary node and storage site. The complete failure of a room includes the failure of both the host and the storage in that room. If the primary room fails, Oracle Solaris Cluster automatically fails over to the secondary room, makes the secondary room's storage device readable and writable, and enables the failover of the corresponding device groups and resource groups.

When the primary room returns online, you can manually recover the data from the SRDF device group that was written to the secondary room and resynchronize the data. This procedure recovers the SRDF device group by synchronizing the data from the original secondary room (this procedure uses *phys-campus-2* for the secondary room) to the original primary room (*phys-campus-1*). The procedure also changes the SRDF device group type to RDF1 on *phys-campus-2* and to RDF2 on *phys-campus-1*.

Before You Begin You must configure the EMC replication group and DID devices, as well as register the EMC replication group before you can perform a manual failover. For information about creating a Solaris Volume Manager device group, see [“How to Add and Register a Device Group \(Solaris Volume Manager\)” on page 127](#). For information about creating a Veritas Volume Manager device group, see [“How to Create a New Disk Group When Encapsulating Disks \(Veritas Volume Manager\)” on page 135](#).

Note – These instructions demonstrate one method you can use to manually recover SRDF data after the primary room fails over completely and then comes back online. Check the EMC documentation for additional methods.

Log into the campus cluster's primary room to perform these steps. In the procedure below, *dg1* is the SRDF device group name. At the time of the failure, the primary room in this procedure is *phys-campus-1* and the secondary room is *phys-campus-2*.

- 1 Log into the campus cluster's primary room and become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**
- 2 From the primary room, use the `symrdf` command to query the replication status of the RDF devices and view information about those devices.**

```
phys-campus-1# symrdf -g dg1 query
```

Tip – A device group that is in the `split` state is not synchronized.

- 3 If the RDF pair state is `split` and the device group type is `RDF1`, then force a failover of the SRDF device group.**

```
phys-campus-1# symrdf -g dg1 -force failover
```

- 4 View the status of the RDF devices.**

```
phys-campus-1# symrdf -g dg1 query
```

- 5 After the failover, you can swap the data on the RDF devices that failed over.**

```
phys-campus-1# symrdf -g dg1 swap
```

- 6 Verify the status and other information about the RDF devices.**

```
phys-campus-1# symrdf -g dg1 query
```

- 7 Establish the SRDF device group in the primary room.**

```
phys-campus-1# symrdf -g dg1 establish
```

- 8 Confirm that the device group is in a synchronized state and that the device group type is `RDF2`.**

```
phys-campus-1# symrdf -g dg1 query
```

Example 5–20 Manually Recovering EMC SRDF Data after a Primary Site Failover

This example provides the Oracle Solaris Cluster-specific steps necessary to manually recover EMC SRDF data after a campus cluster's primary room fails over, a secondary room takes over and records data, and then the primary room comes back online. In the example, the SRDF

device group is called *dg1* and the standard logical device is DEV001. The primary room is *phys-campus-1* at the time of the failure, and the secondary room is *phys-campus-2*. Perform the steps from the campus cluster's primary room, *phys-campus-1*.

```
phys-campus-1# symrdf -g dg1 query | grep DEV
DEV001 0012RW 0 0NR 0012RW 2031 0 S.. Split

phys-campus-1# symdg list | grep RDF
dg1 RDF1 Yes 00187990182 1 0 0 0 0

phys-campus-1# symrdf -g dg1 -force failover
...

phys-campus-1# symrdf -g dg1 query | grep DEV
DEV001 0012 WD 0 0 NR 0012 RW 2031 0 S.. Failed Over

phys-campus-1# symdg list | grep RDF
dg1 RDF1 Yes 00187990182 1 0 0 0 0

phys-campus-1# symrdf -g dg1 swap
...

phys-campus-1# symrdf -g dg1 query | grep DEV
DEV001 0012 WD 0 0 NR 0012 RW 0 2031 S.. Suspended

phys-campus-1# symdg list | grep RDF
dg1 RDF2 Yes 000187990182 1 0 0 0 0

phys-campus-1# symrdf -g dg1 establish
...

phys-campus-1# symrdf -g dg1 query | grep DEV
DEV001 0012 WD 0 0 RW 0012 RW 0 0 S.. Synchronized

phys-campus-1# symdg list | grep RDF
dg1 RDF2 Yes 000187990182 1 0 0 0 0
```

Overview of Administering Cluster File Systems

No special Oracle Solaris Cluster commands are necessary for cluster file system administration. Administer a cluster file system as you would any other Oracle Solaris file system, using standard Oracle Solaris file system commands, such as `mount` and `newfs`. Mount cluster file systems by specifying the `-g` option to the `mount` command. Cluster file systems can also be automatically mounted at boot. Cluster file systems are only visible from the voting node in a global cluster. If you require the cluster file system data to be accessible from a non-voting node, map the data to the non-voting node with `zoneadm(1M)` or `HASstoragePlus`.

Note – When the cluster file system reads files, the file system does not update the access time on those files.

Cluster File System Restrictions

The following restrictions apply to the cluster file system administration:

- The `unlink(1M)` command is not supported on directories that are not empty.
- The `lockfs -d` command is not supported. Use `lockfs -n` as a workaround.
- You cannot remount a cluster file system with the `directio` mount option added at remount time.
- ZFS for root file systems is supported, with one significant exception. If you use a dedicated partition of the boot disk for the global-devices file system, you must use only UFS as its file system. The global-devices namespace requires the proxy file system (PxFS) running on a UFS file system. However, a UFS file system for the global-devices namespace can coexist with a ZFS file system for the root (`/`) file system and other root file systems, for example, `/var` or `/home`. Alternatively, if you instead use a `lofi` device to host the global-devices namespace, there is no limitation on the use of ZFS for root file systems.

Guidelines to Support VxFS

The following VxFS features are not supported in an Oracle Solaris Cluster cluster file system. They are, however, supported in a local file system.

- Quick I/O
- Snapshots
- Storage checkpoints
- VxFS-specific mount options:
 - `convosync` (Convert `O_SYNC`)
 - `mincache`
 - `qlog`, `delaylog`, `tmplog`
- Veritas cluster file system (requires VxVM cluster feature & Veritas Cluster Server). The VxVM cluster feature is not supported on x86 based systems.

Cache advisories can be used, but the effect is observed on the given node only.

All other VxFS features and options that are supported in a cluster file system are supported by Oracle Solaris Cluster software. See VxFS documentation for details about VxFS options that are supported in a cluster configuration.

The following guidelines for using VxFS to create highly available cluster file systems are specific to an Oracle Solaris 3.3 configuration.

- Create a VxFS file system by the following procedures in the VxFS documentation.

- Mount and unmount a VxFS file system from the primary node. The primary node masters the disk on which the VxFS file system resides. A VxFS file system mount or unmount operation that is performed from a secondary node might fail.
- Perform all VxFS administration commands from the primary node of the VxFS cluster file system.

The following guidelines for administering VxFS cluster file systems are not specific to Oracle Solaris 3.3 software. However, the guidelines are different from the way you administer UFS cluster file systems.

- You can administer files on a VxFS cluster file system from any node in the cluster. The exception is `ioctls`, which you must issue only from the primary node. If you do not know whether an administration command involves `ioctls`, issue the command from the primary node.
- If a VxFS cluster file system fails over to a secondary node, all standard system-call operations that were in progress during failover are reissued transparently on the new primary. However, any `ioctl`-related operation in progress during the failover will fail. After a VxFS cluster file system failover, check the state of the cluster file system. Administrative commands that were issued on the old primary before failover might require corrective measures. See VxFS documentation for more information.

Administering Device Groups

As your cluster requirements change, you might need to add, remove, or modify the device groups on your cluster. Oracle Solaris Cluster provides an interactive interface called `clsetup` that you can use to make these changes. `clsetup` generates `cluster` commands. Generated commands are shown in the examples at the end of some procedures. The following table lists tasks for administering device groups and provides links to the appropriate procedures in this section.



Caution – Do not run `metaset -s setname -f -t` on a cluster node that is booted outside the cluster if other nodes are active cluster members and at least one of them owns the disk set.

Note – Oracle Solaris Cluster software automatically creates a raw-disk device group for each disk and tape device in the cluster. However, cluster device groups remain in an offline state until you access the groups as global devices.

TABLE 5-4 Task Map: Administering Device Groups

Task	Instructions
Update the global-devices namespace without a reconfiguration reboot by using the <code>cldevice populate</code> command	“How to Update the Global-Devices Namespace” on page 122
Change the size of a <code>lofi</code> device that is used for the global-devices namespace	“How to Change the Size of a <code>lofi</code> Device That Is Used for the Global-Devices Namespace” on page 123
Move an existing global-devices namespace	“How to Migrate the Global-Devices Namespace From a Dedicated Partition to a <code>lofi</code> Device” on page 124 “How to Migrate the Global-Devices Namespace From a <code>lofi</code> Device to a Dedicated Partition” on page 125
Add Solaris Volume Manager disksets and register them as device groups by using the <code>metaset</code> command	“How to Add and Register a Device Group (Solaris Volume Manager)” on page 127
Add and register a raw-disk device group by using the <code>cldevicegroup</code> command	“How to Add and Register a Device Group (Raw-Disk)” on page 129
Add a named device group for ZFS using the <code>cldevicegroup</code> command	“How to Add and Register a Replicated Device Group (ZFS)” on page 130
Add and register a new disk group as a device group using your preferred method	“How to Create a New Disk Group When Initializing Disks (Veritas Volume Manager)” on page 131
Remove Solaris Volume Manager device groups from the configuration by using the <code>metaset</code> and <code>metaclear</code> commands	“How to Remove and Unregister a Device Group (Solaris Volume Manager)” on page 132
Remove a node from all device groups by using the <code>cldevicegroup</code> , <code>metaset</code> , and <code>clsetup</code> commands	“How to Remove a Node From All Device Groups” on page 132
Remove a node from a Solaris Volume Manager device group by using the <code>metaset</code> command	“How to Remove a Node From a Device Group (Solaris Volume Manager)” on page 133

TABLE 5-4 Task Map: Administering Device Groups (Continued)

Task	Instructions
Add Veritas Volume Manager disk groups as device groups by using VxVM commands and <code>clsetup</code>	“How to Create a New Disk Group When Initializing Disks (Veritas Volume Manager)” on page 131
	“How to Create a New Disk Group When Encapsulating Disks (Veritas Volume Manager)” on page 135
	“How to Add a New Volume to an Existing Device Group (Veritas Volume Manager)” on page 136
	“How to Convert an Existing Disk Group to a Device Group (Veritas Volume Manager)” on page 137
	“How to Assign a New Minor Number to a Device Group (Veritas Volume Manager)” on page 138
	“How to Register a Disk Group as a Device Group (Veritas Volume Manager)” on page 139
	“How to Convert a Local Disk Group to a Device Group (VxVM)” on page 142
	“How to Convert a Device Group to a Local Disk Group (VxVM)” on page 143
	“How to Register Disk Group Configuration Changes (Veritas Volume Manager)” on page 141
Remove Veritas Volume Manager device groups from the configuration by using the <code>clsetup</code> (to generate <code>cldevicegroup</code>) commands	“How to Remove a Volume From a Device Group (Veritas Volume Manager)” on page 144
	“How to Remove and Unregister a Device Group (Veritas Volume Manager)” on page 145
Add a node to a Veritas Volume Manager device group by using <code>clsetup</code> to generate <code>cldevicegroup</code>	“How to Add a Node to a Device Group (Veritas Volume Manager)” on page 146
Remove a node from a Veritas Volume Manager device group by using <code>clsetup</code> to generate <code>cldevicegroup</code>	“How to Remove a Node From a Device Group (Veritas Volume Manager)” on page 148
Remove a node from a raw-disk device group by using the <code>cldevicegroup</code> command	“How to Remove a Node From a Raw-Disk Device Group” on page 149
Change device group properties by using <code>clsetup</code> to generate <code>cldevicegroup</code>	“How to Change Device Group Properties” on page 151
Display device groups and properties by using the <code>cldevicegroup show</code> command	“How to List a Device Group Configuration” on page 156

TABLE 5-4 Task Map: Administering Device Groups (Continued)

Task	Instructions
Change the desired number of secondaries for a device group by using <code>clsetup</code> to generate <code>cldevicegroup</code>	“How to Set the Desired Number of Secondaries for a Device Group” on page 153
Switch the primary for a device group by using the <code>cldevicegroup switch</code> command	“How to Switch the Primary for a Device Group” on page 157
Put a device group in maintenance state by using the <code>metaset</code> or <code>vx dg</code> command	“How to Put a Device Group in Maintenance State” on page 158

▼ How to Update the Global-Devices Namespace

When adding a new global device, manually update the global-devices namespace by running the `cldevice populate` command.

Note – The `cldevice populate` command does not have any effect if the node that is running the command is not currently a cluster member. The command also has no effect if the `/global/.devices/node@nodeID` file system is not mounted.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the cluster.**

- 2 **On each node in the cluster, run the `devfsadm(1M)` command.**

You can run this command on all nodes in the cluster at the same time.

- 3 **Reconfigure the namespace.**

```
# cldevice populate
```

- 4 **On each node, verify that the `cldevice populate` command has been completed before you attempt to create any disksets.**

The `cldevice` command calls itself remotely on all nodes, even when the command is run from just one node. To determine whether the `cldevice populate` command has completed processing, run the following command on each node of the cluster.

```
# ps -ef | grep cldevice populate
```

Example 5-21 Updating the Global-Devices Namespace

The following example shows the output generated by successfully running the `cldevice populate` command.

```
# devfsadm
cldevice populate
Configuring the /dev/global directory (global devices)...
obtaining access to all attached disks
reservation program successfully exiting
# ps -ef | grep cldevice populate
```

▼ How to Change the Size of a `lofi` Device That Is Used for the Global-Devices Namespace

If you use a `lofi` device for the global-devices namespace on one or more nodes of the global cluster, perform this procedure to change the size of the device.

- 1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on a node whose `lofi` device for the global-devices namespace you want to resize.**
- 2 Evacuate services off the node and reboot the node into noncluster mode**
Do this to ensure that global devices are not served from this node while you perform this procedure. For instructions, see [“How to Boot a Node in Noncluster Mode”](#) on page 78.
- 3 Unmount the global-device file system and detach its `lofi` device.**

The global-devices file system mounts locally.

```
phys-schost# umount /global/.devices/node\@'clinfo -n' > /dev/null 2>&1
```

Ensure that the `lofi` device is detached

```
phys-schost# lofiadm -d /.globaldevices
```

The command returns no output if the device is detached

Note – If the file system is mounted by using the `-m` option, no entry is added to the `mnttab` file. The `umount` command might report a warning similar to the following:

```
umount: warning: /global/.devices/node@2 not in mnttab  =====>>>
not mounted
```

This warning is safe to ignore.

- 4 Delete and recreate the `/.globaldevices` file with the required size.**

The following example shows the creation of a new `/.globaldevices` file that is 200 Mbytes in size.

```
phys-schost# rm /.globaldevices
phys-schost# mkfile 200M /.globaldevices
```

5 Create a new file system for the global-devices namespace.

```
phys-schost# lofiadm -a /.globaldevices  
phys-schost# newfs 'lofiadm /.globaldevices' < /dev/null
```

6 Boot the node into cluster mode.

The global devices are now populated on the new file system.

```
phys-schost# reboot
```

7 Migrate to the node any services that you want to run on that node.

Migrating the Global-Devices Namespace

You can create a namespace on a loopback file interface (lofi) device, rather than creating a global-devices namespace on a dedicated partition. This feature is useful if you are installing Oracle Solaris Cluster software on systems that are pre-installed with the Oracle Solaris 10 OS.

Note – ZFS for root file systems is supported, with one significant exception. If you use a dedicated partition of the boot disk for the global-devices file system, you must use only UFS as its file system. The global-devices namespace requires the proxy file system (PxFS) running on a UFS file system. However, a UFS file system for the global-devices namespace can coexist with a ZFS file system for the root (/) file system and other root file systems, for example, /var or /home. Alternatively, if you instead use a lofi device to host the global-devices namespace, there is no limitation on the use of ZFS for root file systems.

The following procedures describe how to move an existing global-devices namespace from a dedicated partition to a lofi device or the opposite:

- [“How to Migrate the Global-Devices Namespace From a Dedicated Partition to a lofi Device” on page 124](#)
- [“How to Migrate the Global-Devices Namespace From a lofi Device to a Dedicated Partition” on page 125](#)

▼ How to Migrate the Global-Devices Namespace From a Dedicated Partition to a lofi Device

- 1 Become superuser on the global-cluster voting node whose namespace location you want to change.**
- 2 Evacuate services off the node and reboot the node into noncluster mode.**

Do this to ensure that global devices are not served from this node while you perform this procedure. For instructions, see [“How to Boot a Node in Noncluster Mode” on page 78](#).

- 3 Ensure that a file named `/.globaldevices` does not exist on the node. If the file does exist, delete it.
- 4 Create the `lofi` device.


```
# mkfile 100m /.globaldevices# lofiadm -a /.globaldevices# \
LOFI_DEV='lofiadm /.globaldevices'# newfs 'echo ${LOFI_DEV} | \
sed -e 's/lofi/rlofi/g'' < /dev/null# lofiadm -d /.globaldevices
```
- 5 In the `/etc/vfstab` file, comment out the `global-devices` namespace entry. This entry has a mount path that begins with `/global/.devices/node@nodeID`.
- 6 Unmount the `global-devices` partition `/global/.devices/node@nodeID`.
- 7 Disable and re-enable the `globaldevices` and `scmountdev` SMF services.


```
# svcadm disable globaldevices# svcadm disable scmountdev# \
svcadm enable scmountdev# svcadm enable globaldevices
```

A `lofi` device is now created on `/.globaldevices` and mounted as the `global-devices` file system.
- 8 Repeat these steps on other nodes whose `global-devices` namespace you want to migrate from a partition to a `lofi` device.
- 9 From one node, populate the `global-device` namespaces.


```
# /usr/cluster/bin/cldevice populate
```

On each node, verify that the command has completed processing before you perform any further actions on the cluster.

```
# ps -ef \ grep cldevice populate
```

The `global-devices` namespace now resides on a `lofi` device.
- 10 Migrate to the node any services that you want to run on that node.

▼ How to Migrate the Global-Devices Namespace From a `lofi` Device to a Dedicated Partition

- 1 Become superuser on the `global-cluster` voting node whose namespace location you want to change.
- 2 Evacuate services off the node and reboot the node into noncluster mode

Do this to ensure that `global devices` are not served from this node while you perform this procedure. For instructions, see [“How to Boot a Node in Noncluster Mode”](#) on page 78.

- 3 **On a local disk of the node, create a new partition that meets the following requirements:**
 - Is at least 512 MByte in size
 - Uses the UFS file system
- 4 **Add an entry to the `/etc/vfstab` file for the new partition to be mounted as the `global-devices` file system.**
 - **Determine the current node's node ID.**

```
# /usr/sbin/clinfo -nnode ID
```
 - **Create the new entry in the `/etc/vfstab` file, using the following format:**

```
blockdevice rawdevice /global/.devices/node@nodeID ufs 2 no global
```

For example, if the partition that you choose to use is `/dev/did/rdisk/d5s3`, the new entry to add to the `/etc/vfstab` file would then be as follows: `/dev/did/dsk/d5s3 /dev/did/rdisk/d5s3 /global/.devices/node@3 ufs 2 no global`
- 5 **Unmount the global devices partition `/global/.devices/node@nodeID`.**
- 6 **Remove the `lofi` device that is associated with the `.globaldevices` file.**

```
# lofiadm -d /.globaldevices
```
- 7 **Delete the `.globaldevices` file.**

```
# rm /.globaldevices
```
- 8 **Disable and re-enable the `globaldevices` and `scmountdev` SMF services.**

```
# svcadm disable globaldevices# svcadm disable scmountdev# \
svcadm enable scmountdev# svcadm enable globaldevices
```

The partition is now mounted as the `global-devices` namespace file system.
- 9 **Repeat these steps on other nodes whose `global-devices` namespace you might want to migrate from a `lofi` device to a partition.**
- 10 **From one node in the cluster, run the `cldevice populate` command to populate the `global-devices` namespace.**

```
# /usr/cluster/bin/cldevice populate
```

Ensure that the process completes on all nodes of the cluster before you perform any further action on any of the nodes.

```
# ps -ef | grep cldevie populate
```

The `global-devices` namespace now resides on the dedicated partition.
- 11 **Migrate to the node any services that you want to run on that node.**

Adding and Registering Device Groups

You can add and register device groups for Solaris Volume Manager, ZFS, Veritas Volume Manager, or raw-disk.

▼ How to Add and Register a Device Group (Solaris Volume Manager)

Use the `metaset` command to create a Solaris Volume Manager disk set and register the disk set as an Oracle Solaris Cluster device group. When you register the disk set, the name that you assigned to the disk set is automatically assigned to the device group.

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.



Caution – The name of the Oracle Solaris Cluster device group that you create (Solaris Volume Manager, Veritas Volume Manager, or raw-disk) must be the same as the name of the replicated device group.

- 1 **Become superuser or assume a role that provides `soLaris.cLuster.modify` RBAC authorization on one of the nodes connected to the disks where you are creating the disk set.**
- 2 **Add the Solaris Volume Manager disk set and register it as a device group with Oracle Solaris Cluster. To create a multi-owner disk group, use the `-M` option.**

```
# metaset -s diskset -a -M -h nodelist
```

`-s diskset` Specifies the disk set to be created.

`-a -h nodelist` Adds the list of nodes that can master the disk set.

`-M` Designates the disk group as multi-owner.

Note – Running the `metaset` command to set up a Solaris Volume Manager device group on a cluster results in one secondary by default, regardless of the number of nodes that are included in that device group. You can change the desired number of secondary nodes by using the `clsetup` utility after the device group has been created. Refer to [“How to Set the Desired Number of Secondaries for a Device Group”](#) on page 153 for more information about disk failover.

3 If you are configuring a replicated device group, set the replication property for the device group.

```
# cldevicegroup sync devicegroup
```

4 Verify that the device group has been added.

The device group name matches the disk set name that is specified with `metaset`.

```
# cldevicegroup list
```

5 List the DID mappings.

```
# cldevice show | grep Device
```

- Choose drives that are shared by the cluster nodes that will master or potentially master the disk set.
- Use the full DID device name, which has the form `/dev/did/rdisk/dN`, when you add a drive to a disk set.

In the following example, the entries for DID device `/dev/did/rdisk/d3` indicate that the drive is shared by `phys-schost-1` and `phys-schost-2`.

```
=== DID Device Instances ===
DID Device Name:                /dev/did/rdisk/d1
  Full Device Path:              phys-schost-1:/dev/rdisk/c0t0d0
DID Device Name:                /dev/did/rdisk/d2
  Full Device Path:              phys-schost-1:/dev/rdisk/c0t6d0
DID Device Name:                /dev/did/rdisk/d3
  Full Device Path:              phys-schost-1:/dev/rdisk/c1t1d0
  Full Device Path:              phys-schost-2:/dev/rdisk/c1t1d0
...
```

6 Add the drives to the disk set.

Use the full DID path name.

```
# metaset -s setname -a /dev/did/rdisk/dN
```

`-s setname` Specifies the disk set name, which is the same as the device group name.

`-a` Adds the drive to the disk set.

Note – Do *not* use the lower-level device name (`cNtXdY`) when you add a drive to a disk set. Because the lower-level device name is a local name and not unique throughout the cluster, using this name might prevent the `metaset` from being able to switch over.

7 Verify the status of the disk set and drives.

```
# metaset -s setname
```

Example 5–22 Adding a Solaris Volume Manager Device Group

The following example shows the creation of the disk set and device group with the disk drives /dev/did/rdisk/d1 and /dev/did/rdisk/d2 and verifies that the device group has been created.

```
# metaset -s dg-schost-1 -a -h phys-schost-1

# cldevicegroup list
dg-schost-1
metaset -s dg-schost-1 -a /dev/did/rdisk/d1 /dev/did/rdisk/d2
```

▼ How to Add and Register a Device Group (Raw-Disk)

Oracle Solaris Cluster software supports the use of raw-disk device groups in addition to other volume managers. When you initially configure Oracle Solaris Cluster, device groups are automatically configured for each raw device in the cluster. Use this procedure to reconfigure these automatically created device groups for use with Oracle Solaris Cluster software.

Create a new device group of the raw-disk type for the following reasons:

- You want to add more than one DID to the device group
- You need to change the name of the device group
- You want to create a list of device groups without using the -v option of the cldg command



Caution – If you are creating a device group on replicated devices, the name of the device group that you create (Solaris Volume Manager, Veritas Volume Manager, or raw-disk) must be the same as the name of the replicated device group.

1 Identify the devices that you want to use and unconfigure any predefined device groups.

The following commands remove the predefined device groups for d7 and d8.

```
paris-1# cldevicegroup disable dsk/d7 dsk/d8
paris-1# cldevicegroup offline dsk/d7 dsk/d8
paris-1# cldevicegroup delete dsk/d7 dsk/d8
```

2 Create the new raw-disk device group, including the desired devices.

The following command creates a global device group, rawdg, which contains d7 and d8.

```
paris-1# cldevicegroup create -n phys-paris-1,phys-paris-2 -t rawdisk
-d d7,d8 rawdg
paris-1# /usr/cluster/lib/dcs/cldg show rawdg -d d7 rawdg
paris-1# /usr/cluster/lib/dcs/cldg show rawdg -d d8 rawdg
```

▼ How to Add and Register a Replicated Device Group (ZFS)

To replicate ZFS, you must create a named device group and list the disks that belong to the zpool. A device can belong to only one device group at a time, so if you already have an Oracle Solaris Cluster device group that contains the device, you must delete the group before you add that device to a new ZFS device group.

The name of the Oracle Solaris Cluster device group that you create (Solaris Volume Manager, Veritas Volume Manager, or raw-disk) must be the same as the name of the replicated device group.



Caution – Full support for ZFS with third-party data-replication technologies is pending. See the latest Oracle Solaris Cluster Release Notes for updates on ZFS support.

1 Delete the default device groups that correspond to the devices in the zpool.

For example, if you have a zpool called `mypool` that contains two devices `/dev/did/dsk/d2` and `/dev/did/dsk/d13`, you must delete the two default device groups called `d2` and `d13`.

```
# cldevicegroup offline dsk/d2 dsk/d13
# cldevicegroup remove dsk/d2 dsk/d13
```

2 Create a named device group with DIDs that correspond to those in the device group you removed in Step #1.

```
# cldevicegroup create -d d2,d13 -t rawdisk mypool
```

This action creates a device group called `mypool` (with the same name as the zpool), which manages the raw devices `/dev/did/dsk/d2` and `/dev/did/dsk/d13`.

3 Create a zpool that contains those devices.

```
# zpool create mypool mirror /dev/did/dsk/d2 /dev/did/dsk/d13
```

4 Create a resource group to manage migration of the replicated devices (in the device group) with only global zones in its nodelist.

```
# clrg create -n pnode1,pnode2 migrate_truecopydg-rg
```

5 Create a hasp-rs resource in the resource group you created in Step 4, setting the `globaldevicepaths` property to a device group of type `raw-disk`. You created this device group in Step #2.

```
# clr create -t HASStoragePlus -x globaldevicepaths=mypool -g \
migrate_truecopydg-rg hasp2migrate_mypool
```

6 If the application resource group will run in local zones, create a new resource group with the nodelist containing the appropriate local zones. The global zones corresponding to the local

zones must be in the nodelist of the resource group created in Step #4. Set the +++ value in the `rg_affinities` property from this resource group to the resource group you created in Step #4.

```
# clrg create -n pnode1:zone-1,pnode2:zone-2 -p \
RG_affinities=+++migrate_truecopydg-rg sybase-rg
```

- 7 Create an `HASStoragePlus` resource (`hasp-rs`) for the `zpool` you created in Step #3 in the resource group that you created in either Step #4 or #6. Set the `resource_dependencies` property to the `hasp-rs` resource that you created in Step #5.

```
# clrs create -g sybase-rg -t HASStoragePlus -p zpools=mypool \
-p resource_dependencies=hasp2migrate_my_pool \
-p ZpoolsSearchDir=/dev/did/dsk hasp2import_my_pool
```

- 8 Use the new resource group name where a device group name is required.

▼ 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 135.

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

If you use VxVM to set up shared disk groups for Oracle RAC, use the cluster functionality of VxVM as described in the *Veritas Volume Manager Administrator's Reference Guide*.

- 1 **Become superuser on any cluster node that is physically connected to 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, use Dirty Region Logging (DRL) to decrease volume recovery time after a node failure. However, DRL might decrease I/O throughput.

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

- 3 **Register the VxVM disk group as an Oracle Solaris Cluster device group.**

See “[How to Register a Disk Group as a Device Group \(Veritas Volume Manager\)](#)” on page 139.

Do not register the Oracle RAC shared disk groups with the cluster framework.

Maintaining Device Groups

You can perform a variety of administrative tasks for your device groups.

How to Remove and Unregister a Device Group (Solaris Volume Manager)

Device groups are Solaris Volume Manager disksets that have been registered with Oracle Solaris Cluster. To remove a Solaris Volume Manager device group, use the `metaclear` and `metaset` commands. These commands remove the device group with the same name and unregister the disk group as an Oracle Solaris Cluster device group.

Refer to the Solaris Volume Manager documentation for the steps to remove a disk set.

▼ How to Remove a Node From All Device Groups

Use this procedure to remove a cluster node from all device groups that list the node in their lists of potential primaries.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on the node that you are removing as a potential primary of all device groups.**
- 2 Determine the device group or groups of which the node to be removed is a member.**
Look for the node name in the Device group node list for each device group.

```
# cldevicegroup list -v
```
- 3 If any of the device groups identified in [Step 2](#) are of the device group type SVM, perform the steps in [“How to Remove a Node From a Device Group \(Solaris Volume Manager\)”](#) on page 133 for each device group of that type.**
- 4 If any of the device groups identified in [Step 2](#) are of the device group type VxVM, perform the steps in [“How to Remove a Node From a Device Group \(Veritas Volume Manager\)”](#) on page 148 for each device group of that type.**

- 5 **Determine the raw-device disk groups of which the node to be removed is a member.**

```
# cldevicegroup list -v
```
- 6 **If any of the device groups listed in Step 5 are of the device group types Disk or Local_Disk, perform the steps in “How to Remove a Node From a Raw-Disk Device Group” on page 149 for each of these device groups.**
- 7 **Verify that the node has been removed from the potential primaries list of all device groups.**
 The command returns nothing if the node is no longer listed as a potential primary of any device group.

```
# cldevicegroup list -v nodename
```

▼ How to Remove a Node From a Device Group (Solaris Volume Manager)

Use this procedure to remove a cluster node from the list of potential primaries of a Solaris Volume Manager device group. Repeat the `metaset` command for each device group from which you want to remove the node.



Caution – Do not run `metaset -s setname -f -t` on a cluster node that is booted outside the cluster if other nodes are active cluster members and at least one of them owns the disk set.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Verify that the node is still a member of the device group and that the device group is a Solaris Volume Manager device group.**
 Device group type SDS/SVM indicates a Solaris Volume Manager device group.

```
phys-schost-1% cldevicegroup show devicegroup
```
- 2 **Determine which node is the current primary for the device group.**

```
# cldevicegroup status devicegroup
```
- 3 **Become superuser on the node that currently owns the device group that you want to modify.**

4 Delete the node's hostname from the device group.

```
# metaset -s setname -d -h nodelist
```

-s *setname* Specifies the device group name.

-d Deletes from the device group the nodes identified with -h.

-h *nodelist* Specifies the node name of the node or nodes that will be removed.

Note – The update can take several minutes to complete.

If the command fails, add the -f (force) option to the command.

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

5 Repeat Step 4 for each device group from which the node is being removed as a potential primary.**6 Verify that the node has been removed from the device group.**

The device group name matches the disk set name that is specified with `metaset`.

```
phys-schost-1% cldevicegroup list -v devicegroup
```

Example 5–23 Removing a Node From a Device Group (Solaris Volume Manager)

The following example shows the removal of the hostname `phys-schost-2` from a device group configuration. This example eliminates `phys-schost-2` as a potential primary for the designated device group. Verify removal of the node by running the `cldevicegroup show` command. Check that the removed node is no longer displayed in the screen text.

```
[Determine the Solaris Volume Manager
 device group for the node:]
# cldevicegroup show dg-schost-1
=== Device Groups ===

Device Group Name:          dg-schost-1
Type:                      SVM
failback:                  no
Node List:                  phys-schost-1, phys-schost-2
preferenced:                yes
numsecondaries:            1
diskset name:              dg-schost-1
[Determine which node is the current primary for the device group:]
# cldevicegroup status dg-schost-1
=== Cluster Device Groups ===

--- Device Group Status ---

Device Group Name    Primary            Secondary        Status
-----
```

```

dg-schost-1          phys-schost-1  phys-schost-2  Online
[Become superuser on the node that currently owns the device group.]
[Remove the host name from the device group:]
# metaset -s dg-schost-1 -d -h phys-schost-2
[Verify removal of the node:]
phys-schost-1% cldevicegroup list -v dg-schost-1
=== Cluster Device Groups ===

--- Device Group Status ---

Device Group Name  Primary          Secondary        Status
-----
dg-schost-1       phys-schost-1   -                Online

```

▼ 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 131.

You can convert nonroot disks to Oracle Solaris Cluster device groups by encapsulating the disks as VxVM disk groups, then registering the disk groups as Oracle Solaris Cluster device groups.

Disk encapsulation is only supported during initial creation of a VxVM disk group. After a VxVM disk group is created and registered as an Oracle Solaris Cluster device group, only disks which can be initialized should be added to the disk group.

If you use VxVM to set up shared disk groups for Oracle RAC, use the cluster functionality of VxVM as described in the *Veritas Volume Manager Administrator's Reference Guide*.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any 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`.**

Set back to `yes` after the disk is encapsulated and registered as an Oracle Solaris Cluster 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 two must also be set to the entire disk. See the `vxdiskadm` man page for more information.

- 4 Shut down and restart the node.**

The `clnode evacuate` command switches over all resource groups and device groups including all non-voting nodes in a global cluster from the specified node to a next-preferred node. Use the `shutdown` command to shut down and restart the node.

```
# clnode evacuate node[...]  
# shutdown -g0 -y -i6
```

- 5 If necessary, switch all resource groups and device groups back to the original node.**

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

```
# cldevicegroup switch -n node devicegroup  
# clresourcegroup switch -z zone -n node resourcegroup
```

node The name of the node.

zone The name of the non-voting node, *node*, that can master the resource group. Specify *zone* only if you specified a non-voting node when you created the resource group.

- 6 Register the VxVM disk group as an Oracle Solaris Cluster device group.**

See “[How to Register a Disk Group as a Device Group \(Veritas Volume Manager\)](#)” on page 139.

Do not register the Oracle RAC shared disk groups with the cluster framework.

- 7 If you set the `mount at boot` option to `no` in [Step 2](#), set it back to `yes`.**

▼ **How to Add a New Volume to an Existing Device Group (Veritas Volume Manager)**

When you add a new volume to an existing VxVM device group, perform the procedure from the primary node of the online device group.

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 141.

The `phys - s chost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris . cluster . read` and `solaris . cluster . administer` RBAC authorization on any node of the cluster.**
- 2 **Determine the primary node for the device group to which you are adding the new volume.**
`# cldevicegroup status`
- 3 **If the device group is offline, bring the device group online.**
`# cldevicegroup switch -n nodename devicegroup`
nodename Specifies the name of the node to which to switch the device group. This node becomes the new primary.
devicegroup Specifies the device group to switch.
- 4 **From the primary node (the node currently mastering the 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.
- 5 **Synchronize the VxVM disk group changes to update the global namespace.**
`# cldevicegroup sync`
“[How to Register Disk Group Configuration Changes \(Veritas Volume Manager\)](#)” on page 141.

▼ How to Convert an Existing Disk Group to a Device Group (Veritas Volume Manager)

You can convert an existing VxVM disk group to an Oracle Solaris Cluster device group by importing the disk group onto the current node, then registering the disk group as an Oracle Solaris Cluster device group.

- 1 Become superuser on any node of the cluster.
- 2 Import the VxVM disk group to the current node.

```
# vxdg import diskgroup
```
- 3 Register the VxVM disk group as an Oracle Solaris Cluster device group.
 See [“How to Register a Disk Group as a Device Group \(Veritas Volume Manager\)”](#) on page 139.

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

If device group registration fails because of a minor number conflict with another disk group, you must assign the new disk group a new, unused minor number. After assigning the new minor number, rerun the procedure to register the disk group as an Oracle Solaris Cluster device group.

- 1 Become superuser on any node of the cluster.
- 2 Determine the minor numbers in use.

```
# ls -l /global/.devices/node@nodeid/dev/vx/dsk/*
```
- 3 Choose another multiple of 1000 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
```
- 5 Register the VxVM disk group as an Oracle Solaris Cluster device group.
 See [“How to Register a Disk Group as a Device Group \(Veritas Volume Manager\)”](#) on page 139.

Example 5–24 How to Assign a New Minor Number to a Device Group

This example uses the minor numbers 16000-16002 and 4000-4001. The `vxdg reminor` command is used to assign the base minor number 5000 to the new device group.

```
# ls -l /global/.devices/node@nodeid/dev/vx/dsk/*

/global/.devices/node@nodeid/dev/vx/dsk/dg1
brw----- 1 root   root    56,16000 Oct  7 11:32 dg1v1
brw----- 1 root   root    56,16001 Oct  7 11:32 dg1v2
brw----- 1 root   root    56,16002 Oct  7 11:32 dg1v3

/global/.devices/node@nodeid/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
# vxdg  remior  dg3 5000
```

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

This procedure uses the `clsetup` utility to register the associated VxVM disk group as an Oracle Solaris Cluster device group.

Note – After a device group has been registered with the cluster, never import or export a VxVM disk group by using VxVM commands. If you make a change to the VxVM disk group or volume, follow the procedure “[How to Register Disk Group Configuration Changes \(Veritas Volume Manager\)](#)” on page 141 to register the device group configuration changes. This procedure ensures that the global namespace is in the correct state.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

Before You Begin Ensure that the following prerequisites have been completed prior to registering a VxVM device group:

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

When you define the preference order, you also specify whether the device group should be switched back to the most preferred node if that node fails and later returns to the cluster.

See `cldevicegroup(1CL)` for more information about node preference and failback options.

Nonprimary cluster nodes (spares) transition to secondary according to the node preference order. The default number of secondaries for a device group is normally set to one. This default setting minimizes performance degradation that is caused by primary checkpointing of multiple secondary nodes during normal operation. For example, in a four-node cluster, the default behavior configures one primary, one secondary, and two spare nodes. See also “[How to Set the Desired Number of Secondaries for a Device Group](#)” on page 153.

- 1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the cluster.**

- 2 Start the `clsetup` utility.**

```
# clsetup
```

The Main Menu is displayed.

- 3 To work with VxVM device groups, type the number that corresponds to the option for device groups and volumes.**

The Device Groups Menu is displayed.

- 4 To register a VxVM device group, type the number that corresponds to the option for registering a VxVM disk group as a device group.**

Follow the instructions and type the name of the VxVM disk group to be registered as an Oracle Solaris Cluster device group.

If this device group is replicated by using storage-based replication, this name must match the replication group name.

If you use VxVM to set up shared disk groups for Oracle Parallel Server/Oracle RAC, you do not register the shared disk groups with the cluster framework. Use the cluster functionality of VxVM as described in the *Veritas Volume Manager Administrator's Reference Guide*.

- 5 If you encounter the following error while attempting to register the device group, remind the device group.**

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

To remind the device group, use the procedure “[How to Assign a New Minor Number to a Device Group \(Veritas Volume Manager\)](#)” on page 138. This procedure enables you to assign a new minor number that does not conflict with a minor number that an existing device group uses.

- 6 If you are configuring a replicated device group, set the replication property for the device group.**

```
# cldevicegroup sync devicegroup
```

- 7 Verify that the device group is registered and online.**

If the device group is properly registered, information for the new device group is displayed when you use the following command.

```
# cldevicegroup status devicegroup
```

Note – If you change any configuration information for a VxVM disk group or volume that is registered with the cluster, you must synchronize the device group by using `clsetup`. Such configuration changes include adding or removing volumes, as well as changing the group, owner, or permissions of existing volumes. Reregistration after configuration changes ensures that the global namespace is in the correct state. See [“How to Update the Global-Devices Namespace” on page 122](#).

Example 5–25 Registering a Veritas Volume Manager Device Group

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

```
# clsetup
# cldevicegroup create -t vxvm -n phys-schost-1,phys-schost-2 -p failback=true dg1

# cldevicegroup status dg1

=== Cluster Device Groups ===
--- Device Group Status ---

Device Group Name      Primary           Secondary         Status
-----
dg1                    phys-schost-1    phys-schost-2    Online
```

See Also To create a cluster file system on the VxVM device group, see [“How to Add a Cluster File System” on page 165](#).

If problems occur with the minor number, see [“How to Assign a New Minor Number to a Device Group \(Veritas Volume Manager\)” on page 138](#).

▼ 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 Oracle Solaris Cluster device group. Registration ensures that the global namespace is in the correct state.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node in the cluster.**

- 2 **Start the `clsetup` utility.**

```
# clsetup
```

The Main Menu is displayed.

- 3 **To work with VxVM device groups, type the number that corresponds to the option for device groups and volumes.**

The Device Groups Menu is displayed.

- 4 **To register configuration changes, type the number that corresponds to the option for synchronizing volume information for a VxVM device group.**

Follow the instructions and type the name of the VxVM disk group that has changed configuration.

Example 5–26 Registering Veritas Volume Manager Disk Group Configuration Changes

The following example shows the `cldevicegroup` command generated by `clsetup` a changed VxVM device group (`dg1`) is registered. This example assumes that the VxVM disk group and volume were created previously.

```
# clsetup
```

```
cldevicegroup sync dg1
```

▼ How to Convert a Local Disk Group to a Device Group (VxVM)

Perform this procedure to change a local VxVM disk group to a globally accessible VxVM device group.

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

- 2 **Start the `clsetup` utility**

```
# clsetup
```

- 3 Unset the `localonly` property.
 - a. Choose the Device Groups and Volumes menu item.
 - b. Choose the Reset a Local VxVM Disk Group to a VxVM Device Group menu item.
 - c. Follow the instructions to unset the `localonly` property.
- 4 Specify the nodes that can master the disk group.
 - a. Return to the main menu in the `clsetup` utility.
 - b. Choose the Device Groups and Volumes menu item.
 - c. Choose the Register a Diskgroup menu item.
 - d. Follow the instructions to specify the nodes that can master the disk group.
 - e. When finished, quit the `clsetup` utility.
- 5 Verify that the device group is configured.


```
phys-schost# cldevicegroup show
```

▼ How to Convert a Device Group to a Local Disk Group (VxVM)

Perform this procedure to change a VxVM device group to a local VxVM disk group that is not managed by Oracle Solaris Cluster software. The local disk group can have more than one node in its node list, but it can be mastered by only one node at a time.

- 1 Become superuser on a node of the cluster.
- 2 Take the device group offline.


```
phys-schost# cldevicegroup offline devicegroup
```
- 3 Unregister the device group.
 - a. Start the `clsetup` utility.


```
phys-schost# clsetup
```
 - b. Choose the Device Groups and Volumes menu item.

- c. Choose the Unregister a VxVM Disk Group menu item.
 - d. Follow the instructions to specify the VxVM disk group that you are unregistering from Oracle Solaris Cluster software.
 - e. Quit the `clsetup` utility.
- 4 Verify that the disk group is no longer registered with Oracle Solaris Cluster software.
`phys-schost# cldevicegroup status`
Command output should no longer show the device group that you unregistered.
- 5 Import the disk group.
`phys-schost# vxdg import diskgroup`
- 6 Set the `localonly` property of the disk group.
 - a. Start the `clsetup` utility.
`phys-schost# clsetup`
 - b. Choose the Device Groups and Volumes menu item.
 - c. Choose the Set a VxVM Disk Group as a Local Disk Group menu item.
 - d. Follow the instructions to set the `localonly` property and to specify the single node that is to exclusively master the disk group.
 - e. When finished, quit the `clsetup` utility.
- 7 Verify that the disk group is successfully configured as a local disk group.
`phys-schost# vxdg list diskgroup`

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

Note – After removing the volume from the device group, you must register the configuration changes to the device group by using the procedure [“How to Register Disk Group Configuration Changes \(Veritas Volume Manager\)”](#) on page 141.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.read` and `solaris.cluster.modify` RBAC authorization on any node of the cluster.**
- 2 **Determine the primary node and status for the device group.**

```
# cldevicegroup status devicegroup
```
- 3 **If the device group is offline, bring it online.**

```
# cldevicegroup online devicegroup
```
- 4 **From the primary node (the node currently mastering the device group), remove the VxVM volume in the disk group.**

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

`-g diskgroup` Specifies the VxVM disk group that contains the volume.

`-rf rm volume` Removes the specified volume. The `-r` option makes the operation recursive. The `-f` option is required to remove an enabled volume.
- 5 **Using the `clsetup` utility, register the device group configuration changes to update the global namespace.**
 See [“How to Register Disk Group Configuration Changes \(Veritas Volume Manager\)”](#) on page 141.

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

Removing an Oracle Solaris Cluster device group causes the corresponding VxVM disk group to be exported, not destroyed. However, even though the VxVM disk group still exists, it cannot be used in the cluster unless reregistered.

This procedure uses the `clsetup` utility to remove a VxVM disk group and unregister it as an Oracle Solaris Cluster device group.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the cluster.**

- 2 **Take the device group offline.**

```
# cldevicegroup offline devicegroup
```

- 3 **Start the `clsetup` utility.**

```
# clsetup
```

The Main Menu is displayed.

- 4 **To work with VxVM device groups, type the number that corresponds to the option for device groups and volumes.**

The Device Groups Menu is displayed.

- 5 **To unregister a VxVM disk group, type the number that corresponds to the option for unregistering a VxVM device group.**

Follow the instructions and type the name of the VxVM disk group to be unregistered.

Example 5–27 Removing and Unregistering a Veritas Volume Manager Device Group

The following example shows the VxVM device group `dg1` taken offline, and the `cldevicegroup` command generated by `clsetup` when it removes and unregisters the device group.

```
# cldevicegroup offline dg1
# clsetup
cldevicegroup delete dg1
```

▼ **How to Add a Node to a Device Group (Veritas Volume Manager)**

This procedure adds a node to a device group using the `clsetup` utility.

The prerequisites to add a node to a VxVM 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

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.read` and `solaris.cluster.modify` RBAC authorization on any node of the cluster.**

- 2 **Start the `clsetup` utility.**

```
# clsetup
```

The Main Menu is displayed.

- 3 **To work with VxVM device groups, type the number that corresponds to the option for device groups and volumes.**

The Device Groups Menu is displayed.

- 4 **To add a node to a VxVM device group, type the number that corresponds to the option for adding a node to a VxVM device group.**

Follow the instructions and type 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.

```
# cldevicegroup show devicegroup
```

Example 5–28 Adding a Node to a Veritas Volume Manager Device Group

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

```
# clsetup
cldevicegroup add-node -n phys-schost-3 dg1
# cldevicegroup show dg1

=== Device Groups ===

Device Group Name:          dg1
Type:                      VxVM
failback:                  yes
Node List:                  phys-schost-1, phys-schost-3
preferenced:                no
numsecondaries:             1
diskgroup names:           dg1
```

▼ How to Remove a Node From a Device Group (Veritas Volume Manager)

Use this procedure to remove a cluster node from the list of potential primaries of a Veritas Volume Manager (VxVM) device group (disk group).

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

1 Verify that the node is still a member of the group and that the group is an VxVM device group.

Device group type VxVM indicates a VxVM device group.

```
phys - schost - 1% cldevicegroup show devicegroup
```

2 Become superuser or assume a role that provides `solaris . cluster . read` and `solaris . cluster . modify` RBAC authorization on a current cluster member node.

3 Start the `clsetup` utility.

```
# clsetup
```

The Main Menu is displayed.

4 To reconfigure a device group, type the number that corresponds to the option for device groups and volumes.

5 To remove the node from the VxVM device group, type the number that corresponds to the option for removing a node from a VxVM device group.

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

- VxVM device group
- Node name

6 Verify that the node has been removed from the VxVM device group or groups.

```
# cldevicegroup show devicegroup
```

Example 5–29 Removing a Node From a Device Group (VxVM)

This example shows the removal of the node named `phys - schost - 1` from the `dg1` VxVM device group.

```

[Determine the VxVM device group for the node:]
# cldevicegroup show dg1

=== Device Groups ===

Device Group Name:          dg1
Type:                      VxVM
failback:                  no
Node List:                 phys-schost-1, phys-schost-2
preferenced:               no
numsecondaries:            1
diskgroup names:          dg1
[Become superuser and start the clsetup utility:]
# clsetup
Select Device groups and volumes>Remove a node from a VxVM device group.

Answer the questions when prompted.
You will need the following information.
Name:          Example:
VxVM device group name    dg1
node names              phys-schost-1

[Verify that the cldevicegroup command executed properly:]
cldevicegroup remove-node -n phys-schost-1 dg1

Command completed successfully.
Dismiss the clsetup Device Groups Menu and Main Menu.
[Verify that the node was removed:]
# cldevicegroup show dg1

=== Device Groups ===

Device Group Name:          dg1
Type:                      VxVM
failback:                  no
Node List:                 phys-schost-2
preferenced:               no
numsecondaries:            1
device names:              dg1

```

▼ How to Remove a Node From a Raw-Disk Device Group

Use this procedure to remove a cluster node from the list of potential primaries of a raw-disk device group.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.read` and `solaris.cluster.modify` RBAC authorization on a node in the cluster *other than the node to remove*.**
- 2 **Identify the device groups that are connected to the node being removed, and determine which are raw-disk device groups.**

```
# cldevicegroup show -n nodename -t rawdisk +
```

- 3 **Disable the `localonly` property of each `Local_Disk` raw-disk device group.**

```
# cldevicegroup set -p localonly=false devicegroup
```

See the `cldevicegroup(1CL)` man page for more information about the `localonly` property.

- 4 **Verify that you have disabled the `localonly` property of all raw-disk device groups that are connected to the node being removed.**

The `Disk` device group type indicates that the `localonly` property is disabled for that raw-disk device group.

```
# cldevicegroup show -n nodename -t rawdisk -v +
```

- 5 **Remove the node from all raw-disk device groups that are identified in [Step 2](#).**

You must complete this step for each raw-disk device group that is connected to the node being removed.

```
# cldevicegroup remove-node -n nodename devicegroup
```

Example 5–30 Removing a Node From a Raw Device Group

This example shows how to remove a node (`phys-schost-2`) from a raw-disk device group. All commands are run from another node of the cluster (`phys-schost-1`).

[Identify the device groups connected to the node being removed, and determine which are raw-disk device groups:]

```
phys-schost-1# cldevicegroup show -n phys-schost-2 -t rawdisk -v +
```

```
Device Group Name:      dsk/d4
Type:                   Disk
failback:               false
Node List:              phys-schost-2
preferenced:           false
localonly:              false
autogen                 true
numsecondaries:        1
device names:          phys-schost-2
```

```
Device Group Name:      dsk/d2
Type:                   VxVM
failback:               true
Node List:              pbrave2
preferenced:           false
localonly:              false
```

```

autogen                                true
numsecondaries:                        1
diskgroup name:                        vxdg1

Device Group Name:                     dsk/d1
Type:                                  SVM
failback:                              false
Node List:                             pbrave1, pbrave2
preferenced:                            true
localonly:                              false
autogen                                 true
numsecondaries:                         1
diskset name:                           ms1
(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
[Disable the localonly flag for each local disk on the node:]
phys-schost-1# cldevicegroup set -p localonly=false dsk/d4
[Verify that the localonly flag is disabled:]
phys-schost-1# cldevicegroup show -n phys-schost-2 -t rawdisk +
  (dsk/d4) Device group type:         Disk
  (dsk/d8) Device group type:         Local_Disk
[Remove the node from all raw-disk device groups:]

phys-schost-1# cldevicegroup remove-node -n phys-schost-2 dsk/d4
phys-schost-1# cldevicegroup remove-node -n phys-schost-2 dsk/d2
phys-schost-1# cldevicegroup remove-node -n phys-schost-2 dsk/d1

```

▼ How to Change Device Group Properties

The method for establishing the primary ownership of a 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 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 reestablish the order of nodes in the primary ownership preference list.

This procedure uses `clsetup` to set or unset the `preferenced` attribute and the `failback` attribute for Solaris Volume Manager or VxVM device groups.

Before You Begin To perform this procedure, you need the name of the device group for which you are changing attribute values.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Become superuser or assume a role that provides `solaris.cluster.read` and `solaris.cluster.modify` RBAC authorization on any node of the cluster.**
- 2 Start the `clsetup` utility.**

```
# clsetup
```

The Main Menu is displayed.
- 3 To work with device groups, type the number that corresponds to the option for device groups and volumes.**

The Device Groups Menu is displayed.
- 4 To change key properties of a device group, type the number that corresponds to the option for changing key properties of a VxVM or Solaris Volume Manager device group).**

The Change Key Properties Menu is displayed.
- 5 To change a device group property, type the number that corresponds to the option for changing the preferences and/or failback properties.**

Follow the instructions to set the `preferenced` and `failback` options for a device group.
- 6 Verify that the device group attributes have been changed.**

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

```
# cldevicegroup show -v devicegroup
```

Example 5-31 Changing Device Group Properties

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

```
# cldevicegroup set -p preferenced=true -p failback=true -p numsecondaries=1 \
-p nodelist=phys-schost-1,phys-schost-2 dg-schost-1
# cldevicegroup show dg-schost-1

=== Device Groups ===

Device Group Name:          dg-schost-1
Type:                      SVM
failback:                  yes
```

```

Node List:                phys-schost-1, phys-schost-2
preferenced:              yes
numsecondaries:           1
diskset names:            dg-schost-1

```

▼ How to Set the Desired Number of Secondaries for a Device Group

The `numsecondaries` property specifies the number of nodes within a device group that can master the group if the primary node fails. The default number of secondaries for device services is one. You can set the value to any integer between one and the number of operational nonprimary provider nodes in the device group.

This setting is an important factor in balancing cluster performance and availability. For example, increasing the desired number of secondaries increases the device group's opportunity to survive multiple failures that occur simultaneously within a cluster. Increasing the number of secondaries also decreases performance regularly during normal operation. A smaller number of secondaries typically results in better performance, but reduces availability. However, a larger number of secondaries does not always result in greater availability of the file system or device group in question. Refer to [Chapter 3, “Key Concepts for System Administrators and Application Developers,”](#) in *Oracle Solaris Cluster Concepts Guide* for more information.

If you change the `numsecondaries` property, secondary nodes are added or removed from the device group if the change causes a mismatch between the actual number of secondaries and the desired number.

This procedure uses the `clsetup` utility to set the `numsecondaries` property for all types of device groups. Refer to `cldevicegroup(1CL)` for information about device group options when configuring any device group.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.read` and `solaris.cluster.modify` RBAC authorization on any node of the cluster.**
- 2 **Start the `clsetup` utility.**

```
# clsetup
```

The Main Menu is displayed.

3 To work with device groups, select the Device Groups and Volumes menu item.

The Device Groups Menu is displayed.

4 To change key properties of a device group, select the Change Key Properties of a Device Group menu item.

The Change Key Properties Menu is displayed.

5 To change the desired number of secondaries, type the number that corresponds to the option for changing the numsecondaries property.

Follow the instructions and type the desired number of secondaries to be configured for the device group. The corresponding `cldevicegroup` command is then executed, a log is printed, and the utility returns to the previous menu.

6 Validate the device group configuration.

```
# cldevicegroup show dg-schost-1
=== Device Groups ===
```

```
Device Group Name:          dg-schost-1
Type:                      VxVm           This might also be SDS or Local_Disk.
failback:                  yes
Node List:                  phys-schost-1, phys-schost-2 phys-schost-3
preferenced:                yes
numsecondaries:             1
diskgroup names:           dg-schost-1
```

Note – If you change any configuration information for a VxVM disk group or volume that is registered with the cluster, you must reregister the device group by using `clsetup`. Such configuration changes include adding or removing volumes, as well as changing the group, owner, or permissions of existing volumes. Reregistration after configuration changes ensures that the global namespace is in the correct state. See [“How to Update the Global-Devices Namespace” on page 122.](#)

7 Verify that the device group attribute has been changed.

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

```
# cldevicegroup show -v devicegroup
```

Example 5–32 Changing the Desired Number of Secondaries (Solaris Volume Manager)

The following example shows the `cldevicegroup` command that is generated by `clsetup` when it configures the desired number of secondaries for a device group (`dg-schost-1`). This example assumes that the disk group and volume were created previously.

```
# cldevicegroup set -p numsecondaries=1 dg-schost-1
# cldevicegroup show -v dg-schost-1

=== Device Groups ===

Device Group Name:          dg-schost-1
Type:                      SVM
failback:                  yes
Node List:                  phys-schost-1, phys-schost-2
preferenced:                yes
numsecondaries:             1
diskset names:              dg-schost-1
```

Example 5-33 Setting the Desired Number of Secondaries (Veritas Volume Manager)

The following example shows the `cldevicegroup` command that is generated by `clsetup` when it sets the desired number of secondaries for a device group (`dg-schost-1`) to two. See [“How to Set the Desired Number of Secondaries for a Device Group” on page 153](#) for information about changing the desired number of secondaries after a device group is created.

```
# cldevicegroup set -p numsecondaries=2 dg-schost-1

# cldevicegroup show dg-schost-1
=== Device Groups ===

Device Group Name:          dg-schost-1
Type:                      VxVM
failback:                  yes
Node List:                  phys-schost-1, phys-schost-2
preferenced:                yes
numsecondaries:             1
diskgroup names:           dg-schost-1
```

Example 5-34 Setting the Desired Number of Secondaries to the Default Value

The following example shows use of a null string value to configure the default number of secondaries. The device group will be configured to use the default value, even if the default value changes.

```
# cldevicegroup set -p numsecondaries= dg-schost-1
# cldevicegroup show -v dg-schost-1

=== Device Groups ===

Device Group Name:          dg-schost-1
Type:                      SVM
failback:                  yes
Node List:                  phys-schost-1, phys-schost-2 phys-schost-3
preferenced:                yes
numsecondaries:             1
diskset names:              dg-schost-1
```

▼ How to List a Device Group Configuration

You do not need to be superuser to list the configuration. However, you do need `solaris.cluster.read` authorization.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- **Use one method from the following list.**

Oracle Solaris Cluster Manager GUI	See the Oracle Solaris Cluster Manager online help for more information.
<code>cldevicegroup show</code>	Use <code>cldevicegroup show</code> to list the configuration for all device groups in the cluster.
<code>cldevicegroup show <i>devicegroup</i></code>	Use <code>cldevicegroup show <i>devicegroup</i></code> to list the configuration of a single device group.
<code>cldevicegroup status <i>devicegroup</i></code>	Use <code>cldevicegroup status <i>devicegroup</i></code> to determine the status of a single device group.
<code>cldevicegroup status +</code>	Use <code>cldevicegroup status +</code> to determine the status of all device groups in the cluster.

Use the `-v` option with any of these commands to obtain more detailed information.

Example 5–35 Listing the Status of All Device Groups

```
# cldevicegroup status +
=== Cluster Device Groups ===
--- Device Group Status ---

Device Group Name   Primary           Secondary          Status
-----
dg-schost-1         phys-schost-2    phys-schost-1     Online
dg-schost-2         phys-schost-1    --                Offline
dg-schost-3         phys-schost-3    phy-shost-2       Online
```

Example 5–36 Listing the Configuration of a Particular Device Group

```
# cldevicegroup show dg-schost-1
=== Device Groups ===
```

```

Device Group Name:          dg-schost-1
Type:                      SVM
failback:                  yes
Node List:                 phys-schost-2, phys-schost-3
preferenced:               yes
numsecondaries:            1
diskset names:             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.

You can also bring an inactive device group online or switch the primary for a device group by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Become superuser or assume a profile that provides `solaris.cluster.modify` RBAC authorization on any node of the cluster.**

- 2 Use `cldevicegroup switch` to switch the device group primary.**

```
# cldevicegroup switch -n nodename devicegroup
```

`-n nodename` Specifies the name of the node to switch to. This node become the new primary.

`devicegroup` Specifies the device group to switch.

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

If the device group is properly registered, information for the new device group is displayed when you use the following command.

```
# cldevice status devicegroup
```

Example 5-37 Switching the Primary for a Device Group

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

```
# cldevicegroup switch -n phys-schost-1 dg-schost-1
```

```
# cldevicegroup status dg-schost-1

=== Cluster Device Groups ===

--- Device Group Status ---

Device Group Name   Primary       Secondary     Status
-----
dg-schost-1         phys-schost-1 phys-schost-2 OnLine
```

▼ How to Put a Device Group in Maintenance State

Putting a device group in maintenance state prevents that device group from automatically being brought online whenever one of its devices is accessed. You should put a device group in maintenance state when completing repair procedures that require that all I/O activity be acquiesced until completion of the repair. Putting a device group in maintenance state also helps prevent data loss by ensuring that a device group is not brought online on one node while the disk set or disk group is being repaired on another node.

For instructions on how to restore a corrupted diskset, see [“Restoring a Corrupted Diskset” on page 283](#).

Note – Before a device group can be placed in maintenance state, all access to its devices must be stopped, and all dependent file systems must be unmounted.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Place the device group in maintenance state.**
 - a. **If the device group is enabled, disable the device group.**

```
# cldevicegroup disable devicegroup
```
 - b. **Take the device group offline.**

```
# cldevicegroup offline devicegroup
```

- 2 If the repair procedure being performed requires ownership of a disk set or disk group, manually import that disk set or disk group.

For Solaris Volume Manager:

```
# metaset -C take -f -s diskset
```



Caution – If you are taking ownership of a Solaris Volume Manager disk set, you *must* use the `metaset -C take` command when the device group is in maintenance state. Using `metaset -t` brings the device group online as part of taking ownership. If you are importing a VxVM disk group, you must use the `-t` flag when importing the disk group. Using the `-t` flag prevents the disk group from automatically being imported if this node is rebooted.

For Veritas Volume Manager:

```
# vxvg -t import disk-group-name
```

- 3 Complete the repair procedure that you need to perform.
- 4 Release ownership of the disk set or disk group.



Caution – Before taking the device group out of maintenance state, you must release ownership of the disk set or disk group. Failure to release ownership can result in data loss.

- For Solaris Volume Manager:


```
# metaset -C release -s diskset
```
- For Veritas Volume Manager:


```
# vxvg deport diskgroupname
```

- 5 Bring the device group online.

```
# cldevicegroup online devicegroup
# cldevicegroup enable devicegroup
```

Example 5–38 Putting a Device Group in Maintenance State

This example shows how to put device group `dg-schost-1` in maintenance state, and remove the device group from maintenance state.

```
[Place the device group in maintenance state.]
# cldevicegroup disable dg-schost-1
# cldevicegroup offline dg-schost-1
[If needed, manually import the disk set or disk group.]
For Solaris Volume Manager:
# metaset -C take -f -s dg-schost-1
For Veritas Volume Manager:
# vxvg -t import dg1
```

```
[Complete all necessary repair procedures.]
```

```
[Release ownership.]
```

```
For Solaris Volume Manager:
```

```
# metaset -C release -s dg-schost-1
```

```
For Veritas Volume Manager:
```

```
# vxdg deport dg1
```

```
[Bring the device group online.]
```

```
# cldevicegroup online dg-schost-1
```

```
# cldevicegroup enable dg-schost-1
```

Administering the SCSI Protocol Settings for Storage Devices

Oracle Solaris Cluster software installation automatically assigns SCSI reservations to all storage devices. Use the following procedures to check the settings of devices and, if necessary, to override the setting for a device.

- “How to Display the Default Global SCSI Protocol Settings for All Storage Devices” on page 160
- “How to Display the SCSI Protocol of a Single Storage Device” on page 161
- “How to Change the Default Global Fencing Protocol Settings for All Storage Devices” on page 162
- “How to Change the Fencing Protocol for a Single Storage Device” on page 163

▼ How to Display the Default Global SCSI Protocol Settings for All Storage Devices

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `soLaris.cLuster.read` RBAC authorization.**
- 2 **From any node, display the current global default SCSI protocol setting.**

```
# cluster show -t global
```

For more information, see the [cluster\(1CL\)](#) man page.

Example 5–39 Displaying the Default Global SCSI Protocol Settings for All Storage Devices

The following example displays the SCSI protocol settings for all storage devices on the cluster.

```
# cluster show -t global

=== Cluster ===

Cluster Name:                racerxx
installmode:                 disabled
heartbeat_timeout:          10000
heartbeat_quantum:          1000
private_netaddr:             172.16.0.0
private_netmask:             255.255.248.0
max_nodes:                   64
max_privatenets:             10
global_fencing:              pathcount
Node List:                   phys-racerxx-1, phys-racerxx-2
```

▼ How to Display the SCSI Protocol of a Single Storage Device

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `soLaris.cLuster.read` RBAC authorization.**
- 2 **From any node, display the SCSI protocol setting of the storage device.**

```
# cldevice show device
device    The name of the device path or a device name.
```

For more information, see the `cldevice(1CL)` man page.

Example 5–40 Displaying the SCSI Protocol of a Single Device

The following example displays the SCSI protocol for the device `/dev/rdisk/c4t8d0`.

```
# cldevice show /dev/rdisk/c4t8d0

=== DID Device Instances ===

DID Device Name:                /dev/did/rdisk/d3
```

Full Device Path:	phappy1:/dev/rdsk/c4t8d0
Full Device Path:	phappy2:/dev/rdsk/c4t8d0
Replication:	none
default_fencing:	global

▼ How to Change the Default Global Fencing Protocol Settings for All Storage Devices

You can turn fencing on or off globally for all storage devices connected to a cluster. The default fencing setting of a single storage device overrides the global setting when the device's default fencing is set to `pathcount`, `prefer3`, or `nofencing`. If the default fencing setting of a storage device is set to `global`, the storage device will use the global setting. For example, if a storage device has the default setting `pathcount`, the setting will not change if you use this procedure to change the global SCSI protocol settings to `prefer3`. You must use the [“How to Change the Fencing Protocol for a Single Storage Device” on page 163](#) procedure to change the default setting of a single device.



Caution – If fencing is turned off under the wrong circumstances, your data can be vulnerable to corruption during application failover. Examine this data corruption possibility carefully when you are considering turning fencing off. Fencing can be turned off if the shared storage device does not support the SCSI protocol or if you want to allow access to the cluster's storage from hosts outside the cluster.

To change the default fencing setting for a quorum device, you must unconfigure the device, change the fencing setting, and reconfigure the quorum device. If you plan to turn fencing off and back on regularly for devices that include quorum devices, consider configuring quorum through a quorum server service to eliminate interruptions in quorum operation.

The `phys - s chost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**
- 2 Set the fencing protocol for all storage devices that are not quorum devices.**

```
cluster set -p global_fencing={pathcount | prefer3 | nofencing | nofencing-noscrub}
```

<code>-p global_fencing</code>	Sets the current global default fencing algorithm for all shared devices.
<code>prefer3</code>	Uses the SCSI-3 protocol for devices with more than two paths.
<code>pathcount</code>	Determines the fencing protocol by the number of DID paths that are attached to the shared device. The <code>pathcount</code> setting is used for quorum devices.
<code>nofencing</code>	Turns fencing off by setting the fencing status for all storage devices.
<code>nofencing-noscrub</code>	Scrubbing the device ensures that the device is cleared of all persistent SCSI reservation information and allows access to the storage from systems outside the cluster. Use the <code>nofencing-noscrub</code> option only for storage devices that have severe problems with SCSI reservations.

Example 5–41 Setting the Default Global Fencing Protocol Settings for All Storage Devices

The following example sets the fencing protocol for all storage devices on the cluster to the SCSI-3 protocol.

```
# cluster set -p global_fencing=prefer3
```

▼ How to Change the Fencing Protocol for a Single Storage Device

You can also set the fencing protocol for a single storage device.

Note – To change the default fencing setting for a quorum device, you must unconfigure the device, change the fencing setting, and reconfigure the quorum device. If you plan to turn fencing off and back on regularly for devices that include quorum devices, consider configuring quorum through a quorum server service to eliminate interruptions in quorum operation.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.



Caution – If fencing is turned off under the wrong circumstances, your data can be vulnerable to corruption during application failover. Examine this data corruption possibility carefully when you are considering turning fencing off. Fencing can be turned off if the shared storage device does not support the SCSI protocol or if you want to allow access to the cluster's storage from hosts outside the cluster.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**
- 2 **Set the fencing protocol of the storage device.**

```
cldevice set -p default_fencing ={pathcount | \
scsi3 | global | nofencing | nofencing-noscrub} device
```

<code>-p default_fencing</code>	Modifies the <code>default_fencing</code> property of the device.
<code>pathcount</code>	Determines the fencing protocol by the number of DID paths that are attached to the shared device.
<code>scsi3</code>	Uses the SCSI-3 protocol.
<code>global</code>	Uses the global default fencing setting. The global setting is used for non-quorum devices.
<code>nofencing</code>	Turns fencing off by setting the fencing status for the specified DID instance.
<code>nofencing-noscrub</code>	Scrubbing the device ensures that the device is cleared of all persistent SCSI reservation information and allows access to the storage device from systems outside the cluster. Use the <code>nofencing-noscrub</code> option only for storage devices that have severe problems with SCSI reservations.
<code>device</code>	Specifies the name of the device path or device name.

For more information, see the `cluster(1CL)` man page.

Example 5–42 Setting the Fencing Protocol of a Single Device

The following example sets the device `d5`, specified by device number, to the SCSI-3 protocol.

```
# cldevice set -p default_fencing=prefer3 d5
```

The following example turns default fencing off for the `d11` device.

```
#cldevice set -p default_fencing=nofencing d11
```

Administering Cluster File Systems

The cluster file system is a globally available file system that can be read and accessed from any node of the cluster.

TABLE 5-5 Task Map: Administering Cluster File Systems

Task	Instructions
Add cluster file systems after the initial Oracle Solaris Cluster installation	“How to Add a Cluster File System” on page 165
Remove a cluster file system	“How to Remove a Cluster File System” on page 168
Check global mount points in a cluster for consistency across nodes	“How to Check Global Mounts in a Cluster” on page 170

▼ How to Add a Cluster File System

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



Caution – Be sure you specify 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 might not intend to delete.

Ensure the following prerequisites have been completed prior to adding an additional cluster file system:

- Superuser privilege is established on a node in the cluster.
- Volume manager software be installed and configured on the cluster.
- A device group (Solaris Volume Manager device group or a VxVM device group) or block disk slice exists on which to create the cluster file system.

If you used Oracle Solaris Cluster Manager to install data services, one or more cluster file systems already exist if shared disks on which to create the cluster file systems were sufficient.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

1 Become superuser on any node in the cluster.

Perform this procedure from the global zone if non-global zones are configured in the cluster.

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

2 Create a file system.

Caution – Any data on the disks is destroyed when you create a file system. Be sure that you specify the correct disk device name. If you specify the wrong device name, you might erase data that you did not intend to delete.

- **For a UFS file system, use the `newfs(1M)` command.**

```
phys-schost# newfs raw-disk-device
```

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

Volume Manager	Sample Disk Device Name	Description
Solaris Volume Manager	/dev/md/nfs/rdisk/d1	Raw disk device d1 within the nfs disk set
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 d1s3

- **For a Veritas File System (VxFS) file system, follow the procedures that are provided in your VxFS documentation.**

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 is not accessed on that node.

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

```
phys-schost# mkdir -p /global/device-group/mountpoint/
```

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

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

4 On each node in the cluster, add an entry to the `/etc/vfstab` file for the mount point.

See the `vfstab(4)` man page for details.

Note – If non-global zones are configured in the cluster, ensure that you mount cluster file systems in the global zone on a path in the global zone's root directory.

- a. In each entry, specify the required mount options for the type of file system that you use.
- b. To automatically mount the cluster file system, set the `mount at boot` field to `yes`.
- c. Ensure that, for each cluster file system, the information in its `/etc/vfstab` entry is identical on each node.
- d. Ensure that the entries in each node's `/etc/vfstab` file list devices in the same order.
- e. Check the boot order dependencies of the 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 and mount `/global/oracle/logs/` only after `phys-schost-1` boots and mounts `/global/oracle/`.

5 On any node in the cluster, run the configuration check utility.

```
phys-schost# cluster check -k vfstab
```

The configuration check utility verifies that the mount points exist. The utility also verifies that `/etc/vfstab` file entries are correct on all nodes of the cluster. If no errors occur, nothing is returned.

For more information, see the `cluster(1CL)` man page.

6 Mount the cluster file system.

```
phys-schost# mount /global/device-group/mountpoint/
```

- For UFS, mount the cluster file system from any node in the cluster.
- For VxFS, mount the cluster file system from the current master of *device-group* to ensure that the file system mounts successfully.

In addition, unmount a VxFS file system from the current master of *device-group* to ensure that the file system unmounts successfully.

Note – To manage a VxFS cluster file system in an Oracle Solaris Cluster environment, run administrative commands only from the primary node on which the VxFS cluster file system is mounted.

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

You can use either the `df` command or `mount` command to list mounted file systems. For more information, see the [df\(1M\)](#) man page or [mount\(1M\)](#) man page.

Cluster file systems are accessible from both the global zone and the non-global zone.

Example 5–43 Creating a UFS Cluster File System

The following example creates a UFS cluster file system on the Solaris Volume Manager volume `/dev/md/oracle/rdisk/d1`. An entry for the cluster file system is added to the `vfstab` file on each node. Then from one node the `cluster check` command is run. After configuration check processing is completed successfully, the cluster file system is mounted from one node and verified on all nodes.

```
phys-schost# newfs /dev/md/oracle/rdisk/d1
...
phys-schost# mkdir -p /global/oracle/d1
phys-schost# 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
...
phys-schost# cluster check -k vfstab
phys-schost# mount /global/oracle/d1
phys-schost# mount
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/largefiles
on Sun Oct 3 08:56:16 2005
```

▼ How to Remove a Cluster File System

You *remove* a cluster file system by merely unmounting it. 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 `cluster shutdown` 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.

Ensure that the following prerequisites have been completed prior to unmounting cluster file systems:

- Superuser privilege is established on a node in the cluster.
- The file system is not busy. A file system is considered busy if a user is working 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 any 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 that you know which processes you are going to stop.

```
# fuser -c [ -u ] mountpoint
```

-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.

mountpoint 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 mountpoint
```

A SIGKILL is sent to each process that uses the cluster file system.

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

```
# fuser -c mountpoint
```

6 From just one node, unmount the file system.

```
# umount mountpoint
```

mountpoint 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/volume/plex.

See your volume manager documentation for more information.

Example 5-44 Removing a Cluster File System

The following example removes a UFS cluster file system that is mounted on the Solaris Volume Manager metadevice or volume/dev/md/oracle/rdisk/d1.

```
# mount -v
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/largefiles
# 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
#
/global/oracle/dsk/d1 /dev/md/oracle/rdisk/d1 /global/oracle/d1 ufs 2 yes global,logging
```

[Save and exit.]

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 `cluster(1CL)` utility verifies the syntax of the entries for cluster file systems in the `/etc/vfstab` file. If no errors occur, nothing is returned.

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

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

```
# cluster check -k vfstab
```

Administering Disk-Path Monitoring

Disk path monitoring (DPM) administration commands enable you to receive notification of secondary disk-path failure. Use the procedures in this section to perform administrative tasks that are associated with monitoring disk paths. Refer to [Chapter 3, “Key Concepts for System Administrators and Application Developers,”](#) in *Oracle Solaris Cluster Concepts Guide* for conceptual information about the disk-path monitoring daemon. Refer to the `cldevice(1CL)` man page for a description of the command options and related commands. For more information about tuning the `sdpmd` daemon, see the `sdpmd.conf(4)` man page. Also see the `syslogd(1M)` man page for logged errors that the daemon reports.

Note – Disk paths are automatically added to the monitoring list monitored when I/O devices are added to a node by using the `cldevice` command. Disk paths are also automatically unmonitored when devices are removed from a node by using Oracle Solaris Cluster commands.

TABLE 5-6 Task Map: Administering Disk-Path Monitoring

Task	Instructions
Monitor a disk path.	“How to Monitor a Disk Path” on page 172
Unmonitor a disk path.	“How to Unmonitor a Disk Path” on page 173
Print the status of faulted disk paths for a node.	“How to Print Failed Disk Paths” on page 174
Monitor disk paths from a file.	“How to Monitor Disk Paths From a File” on page 175
Enable or disable the automatic rebooting of a node when all monitored shared-disk paths fail.	“How to Enable the Automatic Rebooting of a Node When All Monitored Shared-Disk Paths Fail” on page 177 “How to Disable the Automatic Rebooting of a Node When All Monitored Shared-Disk Paths Fail” on page 178
Resolve an incorrect disk-path status. An incorrect disk-path status can be reported when the monitored DID device is unavailable at boot time, and the DID instance is not uploaded to the DID driver.	“How to Resolve a Disk-Path Status Error” on page 174

The procedures in the following section that issue the `cldevice` command include the disk-path argument. The disk-path argument consists of a node name and a disk name. The node name is not required and defaults to `all` if you do not specify it.

▼ How to Monitor a Disk Path

Perform this task to monitor disk paths in your cluster.



Caution – DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node in the cluster.**
- 2 **Monitor a disk path.**

```
# cldevice monitor -n node disk
```
- 3 **Verify that the disk path is monitored.**

```
# cldevice status device
```

Example 5–45 Monitoring a Disk Path on a Single Node

The following example monitors the `schost-1:/dev/did/rdisk/d1` disk path from a single node. Only the DPM daemon on the node `schost-1` monitors the path to the disk `/dev/did/dsk/d1`.

```
# cldevice monitor -n schost-1 /dev/did/dsk/d1
# cldevice status d1
```

Device Instance	Node	Status
/dev/did/rdisk/d1	phys-schost-1	Ok

Example 5–46 Monitoring a Disk Path on All Nodes

The following example monitors the `schost-1:/dev/did/dsk/d1` disk path from all nodes. DPM starts on all nodes for which `/dev/did/dsk/d1` is a valid path.

```
# cldevice monitor /dev/did/dsk/d1
# cldevice status /dev/did/dsk/d1
```

```

Device Instance  Node          Status
-----
/dev/did/rdsk/d1  phys-schost-1  Ok

```

Example 5–47 Rereading the Disk Configuration From the CCR

The following example forces the daemon to reread the disk configuration from the CCR and prints the monitored disk paths with status.

```

# cldevice monitor +
# cldevice status
Device Instance          Node          Status
-----
/dev/did/rdsk/d1         schost-1     Ok
/dev/did/rdsk/d2         schost-1     Ok
/dev/did/rdsk/d3         schost-1     Ok
                        schost-2     Ok
/dev/did/rdsk/d4         schost-1     Ok
                        schost-2     Ok
/dev/did/rdsk/d5         schost-1     Ok
                        schost-2     Ok
/dev/did/rdsk/d6         schost-1     Ok
                        schost-2     Ok
/dev/did/rdsk/d7         schost-2     Ok
/dev/did/rdsk/d8         schost-2     Ok

```

▼ How to Unmonitor a Disk Path

Use this procedure to unmonitor a disk path.



Caution – DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node in the cluster.**
- 2 **Determine the state of the disk path to unmonitor.**

```
# cldevice status device
```

3 On each node, unmonitor the appropriate disk paths.

```
# cldevice unmonitor -n node disk
```

Example 5-48 Unmonitoring a Disk Path

The following example unmonitors the `schost-2:/dev/did/rdsk/d1` disk path and prints disk paths with status for the entire cluster.

```
# cldevice unmonitor -n schost2 /dev/did/rdsk/d1
# cldevice status -n schost2 /dev/did/rdsk/d1
```

Device Instance	Node	Status
-----	----	-----
/dev/did/rdsk/d1	schost-2	Unmonitored

▼ How to Print Failed Disk Paths

Use the following procedure to print the faulted disk paths for a cluster.



Caution – DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

- 1 Become superuser on any node in the cluster.**
- 2 Print the faulted disk paths throughout the cluster.**

```
# cldevice status -s fail
```

Example 5-49 Printing Faulted Disk Paths

The following example prints faulted disk paths for the entire cluster.

```
# cldevice status -s fail
```

Device Instance	Node	Status
-----	----	-----
dev/did/dsk/d4	phys-schost-1	fail

▼ How to Resolve a Disk-Path Status Error

If the following events occur, DPM might not update the status of a failed path when it comes back online:

- A monitored-path failure causes a node reboot.

- The device under the monitored DID path does not come back online until after the rebooted node is back online.

The incorrect disk-path status is reported because the monitored DID device is unavailable at boot time, and therefore the DID instance is not uploaded to the DID driver. When this situation occurs, manually update the DID information.

1 From one node, update the global-devices namespace.

```
# cldevice populate
```

2 On each node, verify that command processing has completed before you proceed to the next step.

The command executes remotely on all nodes, even though the command is run from just one node. To determine whether the command has completed processing, run the following command on each node of the cluster.

```
# ps -ef | grep cldevice populate
```

3 Verify that, within the DPM polling time frame, the status of the faulted disk path is now Ok.

```
# cldevice status disk-device
```

Device Instance	Node	Status
-----	----	-----
dev/did/dsk/dN	phys-schost-1	Ok

▼ How to Monitor Disk Paths From a File

Use the following procedure to monitor or unmonitor disk paths from a file.

To change your cluster configuration by using a file, you must first export the current configuration. This export operation creates an XML file that you can then modify to set the configuration items you are changing. The instructions in this procedure describe this entire process.



Caution – DPM is not supported on nodes that run versions that were released prior to Sun Cluster 3.1 10/03 software. Do not use DPM commands while a rolling upgrade is in progress. After all nodes are upgraded, the nodes must be online to use DPM commands.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node in the cluster.**
- 2 **Export your device configuration to an XML file.**

```
# cldevice export -o configurationfile
```

`-o configurationfile` Specify the file name for your XML file.
- 3 **Modify the configuration file so that device paths are monitored.**

Find the device paths that you want to monitor, and set the `monitored` attribute to `true`.
- 4 **Monitor the device paths.**

```
# cldevice monitor -i configurationfile
```

`-i configurationfile` Specify the file name of the modified XML file.
- 5 **Verify that device path is now monitored.**

```
# cldevice status
```

Example 5–50 Monitor Disk Paths From a File

In the following example, the device path between the node `phys-schost-2` and device `d3` is monitored by using an XML file.

The first step is to export the current cluster configuration.

```
# cldevice export -o deviceconfig
```

The `deviceconfig` XML file shows that the path between `phys-schost-2` and `d3` is not currently monitored.

```
<?xml version="1.0"?>
<!DOCTYPE cluster SYSTEM "/usr/cluster/lib/xml/cluster.dtd">
<cluster name="brave_clus">
.
.
.
  <deviceList readonly="true">
    <device name="d3" ctd="c1t8d0">
      <devicePath nodeRef="phys-schost-1" monitored="true"/>
      <devicePath nodeRef="phys-schost-2" monitored="false"/>
    </device>
  </deviceList>
</cluster>
```

To monitor that path, set the `monitored` attribute to `true`, as follows.

```
<?xml version="1.0"?>
<!DOCTYPE cluster SYSTEM "/usr/cluster/lib/xml/cluster.dtd">
```

```

<cluster name="brave_clus">
.
.
.
  <deviceList readonly="true">
    <device name="d3" ctd="c1t8d0">
      <devicePath nodeRef="phys-schost-1" monitored="true"/>
      <devicePath nodeRef="phys-schost-2" monitored="true"/>
    </device>
  </deviceList>
</cluster>

```

Use the `cldevice` command to read the file and turn on monitoring.

```
# cldevice monitor -i deviceconfig
```

Use the `cldevice` command to verify that the device is now monitored.

```
# cldevice status
```

See Also For more detail about exporting cluster configuration and using the resulting XML file to set cluster configuration, see the [cluster\(1CL\)](#) and the [clconfiguration\(5CL\)](#) man pages.

▼ How to Enable the Automatic Rebooting of a Node When All Monitored Shared-Disk Paths Fail

When you enable this feature, a node automatically reboots, provided that the following conditions are met:

- All monitored shared-disk paths on the node fail.
- At least one of the disks is accessible from a different node in the cluster.

Rebooting the node restarts all resource groups and device groups that are mastered on that node on another node.

If all monitored shared-disk paths on a node remain inaccessible after the node automatically reboots, the node does not automatically reboot again. However, if any disk paths become available after the node reboots but then fail, the node automatically reboots again.

When you enable the `reboot_on_path_failure` property, the states of local-disk paths are not considered when determining if a node reboot is necessary. Only monitored shared disks are affected.

- 1 **On any node in the cluster, become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**

- 2 For *all* nodes in the cluster, enable the automatic rebooting of a node when all monitored shared-disk paths to it fail.

```
# clnode set -p reboot_on_path_failure=enabled +
```

▼ How to Disable the Automatic Rebooting of a Node When All Monitored Shared-Disk Paths Fail

When you disable this feature and all monitored shared-disk paths on a node fail, the node does *not* automatically reboot.

- 1 On any node in the cluster, become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.
- 2 For *all* nodes in the cluster, disable the automatic rebooting of a node when monitored all monitored shared-disk paths to it fail.

```
# clnode set -p reboot_on_path_failure=disabled +
```

Administering Quorum

This chapter provides the procedures for administering quorum devices within Oracle Solaris Cluster and Oracle Solaris Cluster quorum servers. For information about quorum concepts, see “Quorum and Quorum Devices” in *Oracle Solaris Cluster Concepts Guide*.

- “Administering Quorum Devices” on page 179
- “Administering Oracle Solaris Cluster Quorum Servers” on page 203

Administering Quorum Devices

A quorum device is a shared storage device or quorum server that is shared by two or more nodes and that contributes votes that are used to establish a quorum. This section provides the procedures for administering quorum devices.

You can use the `clquorum(1CL)` command to perform all quorum device administrative procedures. In addition, you can accomplish some procedures by using the `clsetup(1CL)` interactive utility or the Oracle Solaris Cluster Manager GUI. Whenever possible, quorum procedures are described in this section by using the `clsetup` utility. The Oracle Solaris Cluster Manager online help describes how to perform quorum procedures by using the GUI. When you work with quorum devices, keep in mind the following guidelines:

- All quorum commands must be run in the global-cluster voting node.
- If the `clquorum` command is interrupted or fails, the quorum configuration information can become inconsistent in the cluster configuration database. If this inconsistency occurs, either rerun the command or run the `clquorum reset` command to reset the quorum configuration.
- For highest availability of the cluster, ensure that the total number of votes that are contributed by quorum devices is less than the total number of votes that are contributed by nodes. Otherwise, the nodes cannot form a cluster if all quorum devices are unavailable, even if all nodes are functioning.

- Do not add a disk that is currently configured as a quorum device to a Solaris ZFS storage pool. If a configured quorum device is added to a Solaris ZFS storage pool, the disk is relabeled as an EFI disk and quorum configuration information is lost and the disk no longer provides a quorum vote to the cluster. Once a disk is in a storage pool, that disk can then be configured as a quorum device. Or, you can unconfigure the disk, add it to the storage pool, then reconfigure the disk as a quorum device.

Note – The `clsetup` command is an interactive interface to the other Oracle Solaris Cluster commands. When `clsetup` runs, the command generates the appropriate specific commands, in this case `clquorum` commands. These generated commands are shown in the examples at the end of the procedures.

To view the quorum configuration, use `clquorum show`. The `clquorum list` command displays the names of quorum devices in the cluster. The `clquorum status` command provides status and vote count information.

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

TABLE 6-1 Task List: Administering Quorum

Task	For Instructions
Add a quorum device to a cluster by using <code>clsetup(1CL)</code>	“Adding a Quorum Device” on page 181
Remove a quorum device from a cluster by using <code>clsetup</code> (to generate <code>clquorum</code>)	“How to Remove a Quorum Device” on page 192
Remove the last quorum device from a cluster by using <code>clsetup</code> (to generate <code>clquorum</code>)	“How to Remove the Last Quorum Device From a Cluster” on page 194
Replace a quorum device in a cluster by using the add and remove procedures	“How to Replace a Quorum Device” on page 195
Modify a quorum device list by using the add and remove procedures	“How to Modify a Quorum Device Node List” on page 196
Put a quorum device into maintenance state by using <code>clsetup</code> (to generate <code>clquorum</code>)	“How to Put a Quorum Device Into Maintenance State” on page 198
(While in maintenance state, the quorum device does not participate in voting to establish the quorum.)	
Reset the quorum configuration to its default state by using <code>clsetup</code> (to generate <code>clquorum</code>)	“How to Bring a Quorum Device Out of Maintenance State” on page 200

TABLE 6-1 Task List: Administering Quorum (Continued)

Task	For Instructions
List the quorum devices and vote counts by using the <code>c1quorum(1CL)</code> command	“How to List the Quorum Configuration” on page 201

Dynamic Reconfiguration With Quorum Devices

You must consider a few issues when completing dynamic reconfiguration (DR) operations on quorum devices in a cluster.

- All of the requirements, procedures, and restrictions that are documented for the Oracle Solaris DR feature also apply to Oracle Solaris Cluster DR support, except for the operating system quiescence operation. Therefore, review the documentation for the Oracle Solaris DR feature *before* using the DR feature with Oracle Solaris Cluster software. You should review in particular the issues that affect non-network IO devices during a DR detach operation.
- Oracle Solaris Cluster rejects DR remove-board operations that are performed when an interface is present that is configured for a quorum device.
- If the DR operation would pertain to an active device, Oracle Solaris Cluster rejects the operation and identifies the devices that would be affected by the operation.

To remove a quorum device, you must complete the following steps, in the order indicated.

TABLE 6-2 Task Map: Dynamic Reconfiguration With Quorum Devices

Task	For Instructions
1. Enable a new quorum device to replace the one being removed.	“Adding a Quorum Device” on page 181
2. Disable the quorum device to be removed.	“How to Remove a Quorum Device” on page 192
3. Perform the DR remove operation on the device being removed.	<i>Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual</i> (from the <i>Solaris 10 on Sun Hardware</i> collection.)

Adding a Quorum Device

This section provides procedures to add a quorum device. Ensure that all nodes in the cluster are online before adding a new quorum device. For information about determining the number of quorum vote counts necessary for your cluster, recommended quorum configurations, and failure fencing, see [“Quorum and Quorum Devices”](#) in *Oracle Solaris Cluster Concepts Guide*.



Caution – Do not add a disk that is currently configured as a quorum device to a Solaris ZFS storage pool. When a configured quorum device is added to a Solaris ZFS storage pool, the disk is relabeled as an EFI disk and quorum configuration information is lost and the disk no longer provides a quorum vote to the cluster. Once a disk is in a storage pool, that disk can then be configured as a quorum device. You can also unconfigure the disk, add it to the storage pool, and then reconfigure the disk as a quorum device.

The Oracle Solaris Cluster software supports the following types of quorum devices:

- Directly attached shared-disk (for SCSI or Serial Attached Technology Attachment (SATA) devices)
- Sun NAS
- Oracle's Sun Storage 7000 Unified Storage Systems
- Network Appliance (NetApp) NAS
- Oracle Solaris Cluster Quorum Server

Procedures for adding these devices are provided in the following sections:

- [“How to Add a Shared Disk Quorum Device” on page 183](#)
- [“How to Add a Network Appliance Network-Attached Storage \(NAS\) Quorum Device” on page 187](#)
- [“How to Add a Quorum Server Quorum Device” on page 189](#)

Note – You cannot configure replicated disks as quorum devices. If you try to add a replicated disk as a quorum device, you receive the following error message and the command exits with an error code.

Disk-name is a replicated device. Replicated devices cannot be configured as quorum devices.

A shared-disk quorum device is any attached storage device that is supported by Oracle Solaris Cluster software. The shared disk is connected to two or more nodes of your cluster. If you turn fencing on, a dual-ported disk can be configured as a quorum device that uses SCSI-2 or SCSI-3 (the default is SCSI-2). If fencing is turned on and your shared device is connected to more than two nodes, you can configure your shared disk as a quorum device that uses the SCSI-3 protocol (the default protocol for more than two nodes). You can use the SCSI override flag to make the Oracle Solaris Clustersoftware use the SCSI-3 protocol for dual-ported shared disks.

If you turn fencing off for a shared disk, you can then configure the disk as a quorum device that uses the software quorum protocol. This would be true regardless of whether the disk supports SCSI-2 or SCSI-3 protocols. Software quorum is a protocol from Oracle that emulates a form of SCSI Persistent Group Reservations (PGR).



Caution – If you are using disks that do not support SCSI (such as SATA), you should turn SCSI fencing off.

For quorum devices, you can use a disk that contains user data or is a member of a device group. View the protocol that is used by the quorum subsystem with a shared disk by looking at the access-mode value for the shared disk in the output from the `cluster show` command.

You can also accomplish these procedures by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

See the `clsetup(1CL)` and `clquorum(1CL)` man pages for information about the commands that are used in the following procedures.

▼ How to Add a Shared Disk Quorum Device

Oracle Solaris Cluster software supports shared-disk (both SCSI and SATA) devices as quorum devices. A SATA device does not support a SCSI reservation, and you must disable the SCSI reservation fencing flag and use the software quorum protocol to configure these disks as quorum devices.

To complete this procedure, identify a disk drive by its device ID (DID), which is shared by the nodes. Use the `cldevice show` command to see the list of DID names. Refer to the `cldevice(1CL)` man page for additional information. Ensure that all nodes in the cluster are online before adding a new quorum device.

Use this procedure to configure SCSI or SATA devices.

The `phys - s chost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the cluster.

2 Start the `clsetup` utility.

```
# clsetup
```

The `clsetup` Main Menu is displayed.

3 Type the number that corresponds to the option for Quorum.

The Quorum Menu is displayed.

- 4 **Type the number that corresponds to the option for adding a quorum device, then type yes when the `clsetup` utility asks you to confirm the quorum device that you are adding.**

The `clsetup` utility asks what type of quorum device you want to add.

- 5 **Type the number that corresponds to the option for a shared-disk quorum device.**

The `clsetup` utility asks which global device you want to use.

- 6 **Type the global device you are using.**

The `clsetup` utility asks you to confirm that the new quorum device should be added to the global device you specified.

- 7 **Type yes to continue adding the new quorum device.**

If the new quorum device is added successfully, the `clsetup` utility displays a message to that effect.

- 8 **Verify that the quorum device has been added.**

```
# clquorum list -v
```

Example 6-1 Adding a Shared Disk Quorum Device

The following example shows the `clquorum` command generated by `clsetup` when it adds a shared-disk quorum device and a verification step.

Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any cluster node.

```
[Start the clsetup utility:]
# clsetup
[Select Quorum>Add a quorum device]
[Answer the questions when prompted.]
[You will need the following information.]
  [Information:                Example:]
  [Directly attached shared disk  shared disk]
  [Global device                 d20]

[Verify that the clquorum command was completed successfully:]
clquorum add d20

  Command completed successfully.
[Quit the clsetup Quorum Menu and Main Menu.]
[Verify that the quorum device is added:]
# clquorum list -v
```

```
Quorum      Type
-----
d20         shared_disk
scphyshost-1 node
scphyshost-2 node
```

▼ How to Add a Sun NAS or Sun Storage 7000 Unified Storage Systems NAS Quorum Device

Ensure that all nodes in the cluster are online before adding a new quorum device.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

1 Use the Sun NAS GUI or Sun Unified Storage GUI to set up an iSCSI device on the Sun NAS filer.

If you have a Sun Unified Storage device, use the BUI to perform the following steps. If you have a Sun NAS device, use the following commands:

- a. Create a file volume that is approximately 50 Mbytes in size.
- b. For each node, create an iSCSI access list.
 - i. Use the name of the cluster as the iSCSI access list name.
 - ii. Add the initiator node name of each cluster node to the access list. CHAP and IQN are not needed.

c. Configure the iSCSI LUN

You can use the name of the backing file volume as the name of the LUN. Add the access list for each node to the LUN.

2 On each of the cluster nodes, discover the iSCSI LUN and set the iSCSI access list to static configuration.

```
# iscsiadm modify discovery -s enable

# iscsiadm list discovery
Discovery:
    Static: enabled
    Send Targets: disabled
    iSNS: disabled

# iscsiadm add static-config iqn.LUNName,IPAddress_of_NASDevice
# devfsadm -i iscsi
# cldevice refresh
```

3 From one cluster node, configure the DIDs for the iSCSI LUN.

```
# /usr/cluster/bin/cldevice populate
```

- 4 Identify the DID device that represents the NAS device LUN that has just been configured into the cluster using iSCSI. Use the `cldevice show` command to see the list of DID names. Refer to the [cldevice\(1CL\)](#) man page for additional information.
- 5 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the cluster.
- 6 Use the `clquorum` command to add the NAS device as a quorum device using the DID device identified in [Step 4](#).

```
# clquorum add d20
```

The cluster has default rules for deciding whether to use scsi-2, scsi-3, or software quorum protocols. See the [clquorum\(1CL\)](#) for more information.

Example 6–2 Adding a Sun NAS or Sun Storage 7000 Unified Storage Systems NAS Quorum Device

The following example shows the `clquorum` command generated by `clsetup` when it adds a Sun NAS quorum device and a verification step.

```
Add an iSCSI device on the Sun NAS filer.
Use the Sun NAS GUI to create a file volume that is approximately 50mb in size.
File Volume Operations -> Create File Volume
For each node, create an iSCSI access list.
iSCSI Configuration -> Configure Access List
Add the initiator node name of each cluster node to the access list.
*** Need GUI or command syntax for this step. ***
Configure the iSCSI LUN
iSCSI Configuration -> Configure iSCSI LUN
On each of the cluster nodes, discover the iSCSI LUN and set the iSCSI access list to static configuration.
iscsiadm modify discovery -s enable
iscsiadm list discovery
Discovery:
  Static: enabled
  Send Targets: enabled
  iSNS: disabled
iscsiadm add static-config
iqn.1986-03.com.sun0-1:000e0c66efe8.4604DE16.thinqorum,10.11.160.20
devsadm -i iscsi
From one cluster node, configure the DID devices for the iSCSI LUN.
/usr/cluster/bin/scldevice populate
/usr/cluster/bin/scldevice populate
Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on any cluster node.

[Add the NAS device as a quorum device
using the DID device:]
clquorum add d20
```

Command completed successfully.

▼ How to Add a Network Appliance Network-Attached Storage (NAS) Quorum Device

When you use a Network Appliance (NetApp) network-attached storage (NAS) device as a quorum device, the following are required:

- You must install the iSCSI license from NetApp.
- You must configure an iSCSI LUN on the clustered filer for use as the quorum device.
- You must configure the NetApp NAS unit to use NTP for synchronizing time.
- At least one of the NTP servers selected for the clustered filer must be an NTP server for the Oracle Solaris Cluster nodes.
- When booting the cluster, always boot the NAS device before you boot the cluster nodes.

If you boot devices in the wrong order, your nodes cannot find the quorum device. If a node should fail in this situation, your cluster might be unable to remain in service. If a service disruption occurs, you must either reboot the entire cluster or remove the NetApp NAS quorum device and add it again.

- A cluster can use a NAS device for only a single quorum device.

You can configure other shared storage if you need additional quorum devices. Additional clusters that use the same NAS device can use separate LUNs on that device as their quorum devices.

See the following Oracle Solaris Cluster documentation for information about installing a NetApp NAS storage device in an Oracle Solaris Cluster environment: [Oracle Solaris Cluster 3.3 With Network-Attached Storage Devices Manual](#).

The `phys - s chost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Make sure that all Oracle Solaris Cluster nodes are online and can communicate with the NetApp clustered filer.**
- 2 **Become superuser or assume a role that provides `solaris . cluster . modify` RBAC authorization on any node of the cluster.**
- 3 **Start the `clsetup` utility.**

```
# clsetup
```

The `clsetup` Main Menu is displayed.

4 Type the number that corresponds to the option for Quorum.

The Quorum Menu is displayed.

5 Type the number that corresponds to the option for adding a quorum device. Then type yes to confirm that you are adding a quorum device.

The `clsetup` utility asks what type of quorum device you want to add.

6 Type the number that corresponds to the option for a `netapp_nas` quorum device. Then type yes to confirm that you are adding a `netapp_nas` quorum device.

The `clsetup` utility asks you provide the name of the new quorum device.

7 Type the name of the quorum device you are adding.

The quorum device name can be any name you choose. The name is only used to process future administrative commands.

The `clsetup` utility asks you to provide the name of the filer for the new quorum device.

8 Type the name of the filer of the new quorum device.

This name is the network-accessible name or address of the filer.

The `clsetup` utility asks you to provide the LUN ID for the filer.

9 Type the ID of the quorum device LUN on the filer.

The `clsetup` utility asks if to the new quorum device should be added on the filer.

10 Type yes to continue adding the new quorum device.

If the new quorum device is added successfully, the `clsetup` utility displays a message to that effect.

11 Verify that the quorum device has been added.

```
# clquorum list -v
```

Example 6-3 Adding a NetApp NAS Quorum Device

The following example shows the `clquorum` command generated by `clsetup` when it adds a NetApp NAS quorum device. The example also shows a verification step.

```
Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on any cluster node.
[Start the clsetup utility:]
# clsetup
[Select Quorum>Add a quorum device]
[Answer the questions when prompted.]
[You will need the following information.]
  [Information:           Example:]
  [Quorum Device         Netapp_nas quorum device]
```

```

[Name:                qd1]
[Filer:               nas1.sun.com]
[LUN ID:              0]
[Verify that the clquorum command was completed successfully:]
clquorum add -t netapp_nas -p filer=nas1.sun.com,-p lun_id=0 qd1
    Command completed successfully.
[Quit the clsetup Quorum Menu and Main Menu.]
[Verify that the quorum device is added:]
# clquorum list -v
Quorum      Type
-----
qd1         netapp_nas
scphyshost-1  node
scphyshost-2  node

```

▼ How to Add a Quorum Server Quorum Device

Before You Begin Before you can add an Oracle Solaris Cluster Quorum Server as a quorum device, the Oracle Solaris Cluster Quorum Server software must be installed on the host machine and the quorum server must be started and running. For information about installing the quorum server, see the “How to Install and Configure Quorum Server Software” in *Oracle Solaris Cluster Software Installation Guide*.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the cluster.**
- 2 **Make sure that all Oracle Solaris Cluster nodes are online and can communicate with the Oracle Solaris Cluster Quorum Server.**
 - a. **Ensure that network switches that are directly connected to cluster nodes meet one of the following criteria:**
 - The switch supports Rapid Spanning Tree Protocol (RSTP).
 - Fast port mode is enabled on the switch.

One of these features is required to ensure immediate communication between cluster nodes and the quorum server. If this communication is significantly delayed by the switch, the cluster interprets this prevention of communication as loss of the quorum device.

- b. If the public network uses variable-length subnetting, also called Classless Inter-Domain Routing (CIDR), modify the following files on each node.**

If you use classful subnets, as defined in RFC 791, you do not need to perform these steps.

- i. Add to the `/etc/inet/netmasks` file an entry for each public subnet that the cluster uses.**

The following is an example entry which contains a public-network IP address and netmask:

```
10.11.30.0    255.255.255.0
```

- ii. Append `netmask + broadcast +` to the `hostname` entry in each `/etc/hostname.adapter` file.**

```
nodename netmask + broadcast +
```

- c. On each node in the cluster, add the quorum server hostname to the `/etc/inet/hosts` file or the `/etc/inet/ipnodes` file.**

Add a hostname-to-address mapping to the file, such as the following.

```
ipaddress qshost1
```

ipaddress The IP address of the computer where the quorum server is running.

qshost1 The hostname of the computer where the quorum server is running.

- d. If you use a naming service, add the quorum server host's name-to-address mapping to the name-service database.**

- 3 Start the `clsetup` utility.**

```
# clsetup
```

The `clsetup` Main Menu is displayed.

- 4 Type the number that corresponds to the option for Quorum.**

The Quorum Menu is displayed.

- 5 Type the number that corresponds to the option for adding a quorum device. Then type `yes` to confirm that you are adding a quorum device.**

The `clsetup` utility asks what type of quorum device you want to add.

- 6 Type the number that corresponds to the option for a quorum-server quorum device. Then type `yes` to confirm that you are adding a quorum-server quorum device.**

The `clsetup` utility asks you to provide the name of the new quorum device.

7 Type the name of the quorum device you are adding.

The quorum device name can be any name you choose. The name is only used to process future administrative commands.

The `clsetup` utility asks you to provide the name of the filer for the new quorum device.

8 Type the name of the host of the quorum server.

This name specifies the IP address of the machine where the quorum server runs or the hostname of the machine on the network.

Depending on the IPv4 or IPv6 configuration of the host, the IP address of the machine must be specified in the `/etc/hosts` file, the `/etc/inet/ipnodes` file, or both.

Note – The machine you specify must be reachable by all cluster nodes and must run the quorum server.

The `clsetup` utility asks you to provide the port number of the quorum server.

9 Type the port number that is used by the quorum server to communicate with the cluster nodes.

The `clsetup` utility asks you to confirm that the new quorum device should be added.

10 Type yes to continue adding the new quorum device.

If the new quorum device is added successfully, the `clsetup` utility displays a message to that effect.

11 Verify that the quorum device has been added.

```
# clquorum list -v
```

Example 6-4 Adding a Quorum Server Quorum Device

The following example shows the `clquorum` command that is generated by `clsetup` when it adds a quorum server quorum device. The example also shows a verification step.

Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any cluster node.

```
[Start the clsetup utility:]
# clsetup
[Select Quorum > Add a quorum device]
[Answer the questions when prompted.]
[You will need the following information.]
[Information:           Example:]
[Quorum Device         quorum_server quorum device]
[Name:                 qd1]
[Host Machine Name:    10.11.124.84]
[Port Number:         9001]
```

```
[Verify that the clquorum command was completed successfully:]
clquorum add -t quorum_server -p qshost=10.11.124.84,-p port=9001 qd1
```

Command completed successfully.

[Quit the clsetup Quorum Menu and Main Menu.]

[Verify that the quorum device is added:]

```
# clquorum list -v
```

```
Quorum      Type
-----
qd1         quorum_server
scphyshost-1 node
scphyshost-2 node
```

```
# clquorum status
```

```
=== Cluster Quorum ===
```

```
-- Quorum Votes Summary --
```

Needed	Present	Possible
-----	-----	-----
3	5	5

```
-- Quorum Votes by Node --
```

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

```
-- Quorum Votes by Device --
```

Device Name	Present	Possible	Status
-----	-----	-----	-----
qd1	1	1	Online
d3s2	1	1	Online
d4s2	1	1	Online

Removing or Replacing a Quorum Device

This section provides the following procedures for removing or replacing a quorum device:

- [“How to Remove a Quorum Device” on page 192](#)
- [“How to Remove the Last Quorum Device From a Cluster” on page 194](#)
- [“How to Replace a Quorum Device” on page 195](#)

▼ How to Remove a Quorum Device

You can also accomplish this procedure by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

When a quorum device 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, `clquorum(1CL)` will fail to remove the device from the configuration. If you are removing a node, remove all quorum devices connected to the node.

Note – If the device you intend to remove is the last quorum device in the cluster, see the procedure “[How to Remove the Last Quorum Device From a Cluster](#)” on page 194.

The `phys - s chost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node in the cluster.**
- 2 **Determine the quorum device to be removed.**
`# clquorum list -v`
- 3 **Execute the `clsetup(1CL)` utility.**
`# clsetup`
 The Main Menu is displayed.
- 4 **Type the number that corresponds to the option for Quorum.**
- 5 **Type the number that corresponds to the option to remove a quorum device.**
 Answer the questions displayed during the removal process.
- 6 **Quit `clsetup`.**
- 7 **Verify that the quorum device is removed.**
`# clquorum list -v`

Example 6–5 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 or assume a role that provides `solaris.cluster.modify` RBAC authorization on any cluster node.

```
[Determine the quorum device to be removed:]
# clquorum list -v
```

```
[Start the clsetup utility:]
# clsetup
[Select Quorum>Remove a quorum device]
[Answer the questions when prompted.]
Quit the clsetup Quorum Menu and Main Menu.]
[Verify that the quorum device is removed:]
# clquorum list -v

Quorum      Type
-----
scphyshost-1  node
scphyshost-2  node
scphyshost-3  node
```

Troubleshooting If you lose communications between the cluster and the quorum server host while removing a quorum server quorum device, you must clean up stale configuration information about the quorum server host. For instructions on performing this cleanup, see [“Cleaning Up Stale Quorum Server Cluster Information” on page 207](#).

▼ How to Remove the Last Quorum Device From a Cluster

This procedure removes the last quorum device from a two-node cluster by using the `clquorum` force option, `-F`. Generally, you should first remove the failed device and then add the replacement quorum device. If this is not the last quorum device in a two-node cluster, follow the steps in [“How to Remove a Quorum Device” on page 192](#).

Adding a quorum device involves a node reconfiguration, which touches the failed quorum device and panics the machine. The Force option lets you remove the failed quorum device without panicking the machine. The `clquorum(1CL)` command enables you to remove the device from the configuration. After you remove the failed quorum device, you can add a new device with the `clquorum add` command. See [“Adding a Quorum Device” on page 181](#).

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node in the cluster.**
- 2 **Remove the quorum device by using the `clquorum` command. If the quorum device failed, use the `-F` Force option to remove the failed device.**

```
# clquorum remove -F qd1
```

Note – You can also place the node to be removed in maintenance state and then remove the quorum device with the `clquorum remove quorum` command. The `clsetup(1CL)` cluster administration menu options are not available while the cluster is in install mode. See [“How to Put a Node Into Maintenance State” on page 260](#) for more information.

3 Verify that the quorum device has been removed.

```
# clquorum list -v
```

Example 6-6 Removing the Last Quorum Device

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

```
[Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization on any
 cluster node.]
[Place the cluster in install mode:]
# cluster set -p installmode=enabled
[Remove the quorum device:]
# clquorum remove d3
[Verify that the quorum device has been removed:]
# clquorum list -v
Quorum      Type
-----
scphyshost-1  node
scphyshost-2  node
scphyshost-3  node
```

▼ How to Replace a Quorum Device

Use this procedure to replace an existing quorum device with another quorum device. You can replace a quorum device with a similar device type, such as replacing a NAS device with another NAS device, or you can replace the device with a dissimilar device, such as replacing a NAS device with a shared disk.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

1 Configure a new quorum device.

You need to first add a new quorum device to the configuration to take the place of the old device. See [“Adding a Quorum Device” on page 181](#) to add a new quorum device to the cluster.

2 Remove the device that you are replacing as a quorum device.

See “[How to Remove a Quorum Device](#)” on page 192 to remove the old quorum device from the configuration.

3 If the quorum device is a failed disk, replace the disk.

Refer to the hardware procedures for your disk enclosure in the *Oracle Solaris Cluster 3.3 Hardware Administration Manual*.

Maintaining Quorum Devices

This section provides the following procedures for maintaining quorum devices:

- “[How to Modify a Quorum Device Node List](#)” on page 196
- “[How to Put a Quorum Device Into Maintenance State](#)” on page 198
- “[How to Bring a Quorum Device Out of Maintenance State](#)” on page 200
- “[How to List the Quorum Configuration](#)” on page 201
- “[How to Repair a Quorum Device](#)” on page 202

▼ How to Modify a Quorum Device Node List

You can use the `clsetup(1CL)` utility to add a node to or remove a node from the node list of an existing quorum device. To modify a quorum device's node list, you must remove the quorum device, modify the physical connections of nodes to the quorum device you removed, then add the quorum device to the cluster configuration again. When a quorum device is added, `clquorum(1CL)` automatically configures the node-to-disk paths for all nodes attached to the disk.

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the cluster.**2 Determine the name of the quorum device you are modifying.**

```
# clquorum list -v
```

3 Start the `clsetup` utility.

```
# clsetup
```

The Main Menu is displayed.

- 4 **Type the number that corresponds to the Quorum option.**
The Quorum Menu is displayed.
- 5 **Type the number that corresponds to the option to remove a quorum device.**
Follow the instructions. You will be asked the name of the disk to be removed.
- 6 **Add or delete the node connections to the quorum device.**
- 7 **Type the number that corresponds to the option to add a quorum device.**
Follow the instructions. You will be asked the name of the disk to be used as the quorum device.
- 8 **Verify that the quorum device has been added.**
`# clquorum list -v`

Example 6-7 Modifying a Quorum Device Node List

The following example shows how to use the `clsetup` utility to add nodes to or delete nodes from a quorum device node list. In this example, the quorum device name is `d2`, and the final result of the procedures adds another node to the quorum device node list.

[Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node in the cluster.]

[Determine the quorum device name:]

```
# clquorum list -v
Quorum          Type
-----          -
d2              shared_disk
sc-phys-schost-1 node
sc-phys-schost-2 node
sc-phys-schost-3 node
```

[Start the `clsetup` utility:]

```
# clsetup
```

[Type the number that corresponds with the quorum option.]

```
.
```

[Type the number that corresponds with the option to remove a quorum device.]

```
.
```

[Answer the questions when prompted.]

[You will need the following information:]

```
Information:      Example:
Quorum Device Name:  d2
```

[Verify that the `clquorum` command completed successfully:]

```
clquorum remove d2
Command completed successfully.
```

[Verify that the quorum device was removed.]

```
# clquorum list -v
Quorum      Type
-----
sc-phys-schost-1  node
sc-phys-schost-2  node
sc-phys-schost-3  node

[Type the number that corresponds with the Quorum option.]
.
[Type the number that corresponds with the option to add a quorum device.]
.
[Answer the questions when prompted.]
[You will need the following information:]

    Information      Example:
    quorum device name    d2

[Verify that the clquorum command was completed successfully:]
clquorum add d2
    Command completed successfully.

Quit the clsetup utility.

[Verify that the correct nodes have paths to the quorum device.
In this example, note that phys-schost-3 has been added to the
enabled hosts list.]
# clquorum show d2 | grep Hosts
=== Quorum Devices ===

Quorum Device Name:      d2
Hosts (enabled):        phys-schost-1, phys-schost-2, phys-schost-3

[Verify that the modified quorum device is online.]

# clquorum status d2
=== Cluster Quorum ===

--- Quorum Votes by Device ---

Device Name      Present      Possible      Status
-----
d2                1            1            Online
```

▼ How to Put a Quorum Device Into Maintenance State

Use the `clquorum(1CL)` command to put a quorum device into maintenance state. The `clsetup(1CL)` utility does not currently have this capability. You can also accomplish this procedure by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

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

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

To put a cluster node into maintenance state, see [“How to Put a Node Into Maintenance State” on page 260](#).

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

1 Become superuser or assume a role that provides `soLaris.cLuster.modify` RBAC authorization on any node of the cluster.

2 Put the quorum device into the maintenance state.

```
# clquorum disable device
```

device Specifies the DID name of the disk device to change, for example, `d4`.

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.

```
# clquorum status device
```

Example 6–8 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.

```
# clquorum disable d20
# clquorum status d20

=== Cluster Quorum ===

--- Quorum Votes by Device ---

Device Name      Present      Possible      Status
-----
d20              1            1            Offline
```

See Also To re-enable the quorum device, see [“How to Bring a Quorum Device Out of Maintenance State” on page 200.](#)

To put a node into maintenance state, see [“How to Put a Node Into Maintenance State” on page 260.](#)

▼ How to Bring a Quorum Device Out of Maintenance State

Run this procedure each time a quorum device is in maintenance state and you want to bring the quorum device out of maintenance state and reset the quorum vote count to the default.



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

When you configure a quorum device, Oracle Solaris Cluster software assigns the quorum device a vote count of $N-1$ where N is the number of connected votes to the quorum device. For example, a quorum device that is connected to two nodes with nonzero vote counts has a quorum count of one (two minus one).

- To bring a cluster node as well as its associated quorum devices out of maintenance state, see [“How to Bring a Node Out of Maintenance State” on page 261.](#)
- To learn more about quorum vote counts, see [“About Quorum Vote Counts” in Oracle Solaris Cluster Concepts Guide.](#)

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the cluster.**
- 2 Reset the quorum count.**

```
# clquorum enable device
```

device Specifies the DID name of the quorum device to reset, for example, `d4`.
- 3 If you are resetting the quorum count because a node was in maintenance state, reboot the node.**
- 4 Verify the quorum vote count.**

```
# clquorum show +
```

Example 6-9 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.

```
# clquorum enable d20
# clquorum show +

=== Cluster Nodes ===

Node Name:                phys-schost-2
Node ID:                   1
Quorum Vote Count:        1
Reservation Key:          0x43BAC41300000001

Node Name:                phys-schost-3
Node ID:                   2
Quorum Vote Count:        1
Reservation Key:          0x43BAC41300000002

=== Quorum Devices ===

Quorum Device Name:      d3
Enabled:                 yes
Votes:                   1
Global Name:             /dev/did/rdisk/d20s2
Type:                    shared_disk
Access Mode:             scsi2
Hosts (enabled):        phys-schost-2, phys-schost-3
```

▼ How to List the Quorum Configuration

You can also accomplish this procedure by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

You do not need to be superuser to list the quorum configuration. You can assume any role that provides `solaris.cluster.read` RBAC authorization.

Note – When you increase or decrease the number of node attachments to a quorum device, the quorum vote count is not automatically recalculated. You can reestablish the correct quorum vote if you remove all quorum devices and then add them back into the configuration. For a two-node cluster, temporarily add a new quorum device before you remove and add back the original quorum device. Then remove the temporary quorum device.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- Use **clquorum(1CL)** to list the quorum configuration.

```
% clquorum show +
```

Example 6–10 Listing the Quorum Configuration

```
% clquorum show +
```

```
=== Cluster Nodes ===
```

```
Node Name:                phys-schost-2
Node ID:                   1
Quorum Vote Count:        1
Reservation Key:          0x43BAC41300000001
```

```
Node Name:                phys-schost-3
Node ID:                   2
Quorum Vote Count:        1
Reservation Key:          0x43BAC41300000002
```

```
=== Quorum Devices ===
```

```
Quorum Device Name:      d3
Enabled:                  yes
Votes:                    1
Global Name:              /dev/did/rdisk/d20s2
Type:                     shared_disk
Access Mode:              scsi2
Hosts (enabled):         phys-schost-2, phys-schost-3
```

▼ How to Repair a Quorum Device

Use this procedure to replace a malfunctioning quorum device.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Remove the disk device that you are replacing as a quorum device.

Note – If the device you intend to remove is the last quorum device, you might want to first add another disk as a new quorum device. This step assures a valid quorum device if a failure occurs during the replacement procedure. See [“Adding a Quorum Device” on page 181](#) to add a new quorum device.

See [“How to Remove a Quorum Device” on page 192](#) to remove a disk device as a quorum device.

2 Replace the disk device.

To replace the disk device, see the hardware procedures for the disk enclosure in the *Oracle Solaris Cluster 3.3 Hardware Administration Manual*.

3 Add the replaced disk as a new quorum device.

See [“Adding a Quorum Device” on page 181](#) to add a disk as a new quorum device.

Note – If you added an additional quorum device in [Step 1](#), it is now safe to remove it. See [“How to Remove a Quorum Device” on page 192](#) to remove the quorum device.

Administering Oracle Solaris Cluster Quorum Servers

Oracle Solaris Cluster Quorum Server provides a quorum device that is not a shared storage device. This section provides procedure for administering Oracle Solaris Cluster quorum servers, including:

- [“Overview of the Quorum Server Configuration File” on page 203](#)
- [“Starting and Stopping the Quorum Server Software” on page 204](#)
- [“How to Start a Quorum Server” on page 205](#)
- [“How to Stop a Quorum Server” on page 205](#)
- [“Displaying Information About the Quorum Server” on page 206](#)
- [“Cleaning Up Stale Quorum Server Cluster Information” on page 207](#)

For information about installing and configuring Oracle Solaris Cluster quorum servers, see [“How to Install and Configure Quorum Server Software”](#) in *Oracle Solaris Cluster Software Installation Guide*.

Overview of the Quorum Server Configuration File

When you install the Oracle Solaris Cluster software, a default configuration file, `/etc/scqsd/scqsd.conf`, is created that contains information about a single default quorum server. Each line in the `/etc/scqsd/scqsd.conf` file has the following format:

<code>/usr/cluster/lib/sc/scqsd [-d <i>quorumdirectory</i>] [-i <i>instancename</i>] -p <i>port</i></code>	
<code>/usr/cluster/lib/sc/scqsd</code>	The full path to where you installed the Oracle Solaris Cluster software. This value must be <code>/usr/cluster/lib/sc/scqsd</code> .
<code>-d <i>quorumdirectory</i></code>	The path to the directory where the quorum server can store quorum data. The quorum server process creates one file per cluster in this directory to store cluster-specific quorum information. By default, the value of this option is <code>/var/scqsd</code> . This directory must be unique for each quorum server that you configure.
<code>-i <i>instancename</i></code>	A unique name that you choose for the quorum server instance.
<code>-p <i>port</i></code>	The port number on which the quorum server listens for requests from the cluster. The default port is 9000.

Instance names are optional. If you specify a name for the quorum server, that name must be unique among all quorum servers in your system. If you choose to omit the instance name option, you must refer to the quorum server by the port on which it listens.

Starting and Stopping the Quorum Server Software

These procedures describe how to start and stop the Oracle Solaris Cluster software.

By default, these procedures start and stop a single default quorum server unless you have customized the content of the quorum server configuration file, `/etc/scqsd/scqsd.conf`. The default quorum server is bound on port 9000 and uses the `/var/scqsd` directory for quorum information.

For information about customizing the quorum server configuration file, see [“Overview of the Quorum Server Configuration File” on page 203](#). For information about installing the Quorum Server software, see [“How to Install and Configure Quorum Server Software” in *Oracle Solaris Cluster Software Installation Guide*](#).

▼ How to Start a Quorum Server

- 1 Become superuser on the host where you want to start the Oracle Solaris Cluster software.
- 2 Use the `clquorumserver start` command to start the software.

```
# /usr/cluster/bin/clquorumserver start quorumserver
```

quorumserver Identifies the quorum server. You can use the port number on which the quorum server listens. If you provided an instance name in the configuration file, you can use the name instead.

To start a single quorum server, provide either the instance name or port number. To start all quorum servers, when you have multiple quorum servers configured, use the `+` operand.

Example 6–11 Starting All Configured Quorum Servers

The following example starts all the configured quorum servers.

```
# /usr/cluster/bin/clquorumserver start +
```

Example 6–12 Starting a Specific Quorum Server

The following example starts the quorum server that listens on port number 2000.

```
# /usr/cluster/bin/clquorumserver start 2000
```

▼ How to Stop a Quorum Server

- 1 Become superuser on the host where you want to start the Oracle Solaris Cluster software.
- 2 Use the `clquorumserver stop` command to stop the software.

```
# /usr/cluster/bin/clquorumserver stop [-d] quorumserver
```

`-d` Controls if the quorum server starts the next time you boot the machine. If you specify the `-d` option, the quorum server will not start the next time the machine boots.

quorumserver Identifies the quorum server. You can use the port number on which the quorum server listens. If you provided an instance name in the configuration file, you can use that name instead.

To stop a single quorum server, provide either the instance name or port number. To stop all quorum servers, when you have multiple quorum servers configured, use the + operand.

Example 6–13 Stopping All Configured Quorum Servers

The following example stops all the configured quorum servers.

```
# /usr/cluster/bin/clquorumserver stop +
```

Example 6–14 Stopping a Specific Quorum Server

The following example stops the quorum server that listens on port number 2000.

```
# /usr/cluster/bin/clquorumserver stop 2000
```

Displaying Information About the Quorum Server

You can display configuration information about the quorum server. For every cluster that configured the quorum server as a quorum device, this command shows the corresponding cluster name, cluster ID, list of reservation keys, and list of registration keys.

▼ How to Display Information About the Quorum Server

1 Become the superuser on the host where you want to display the quorum server information.

Users other than the superuser require `solaris.cluster.read` role-based access control (RBAC) authorization. For more information about RBAC rights profiles, see the `rbac(5)` man page.

2 Display the configuration information of the quorum server by using the `clquorumserver` command.

```
# /usr/cluster/bin/clquorumserver show quorumserver
```

quorumserver Identifies one or more quorum servers. You can specify the quorum server by instance name, or by port number. To display configuration information for all quorum servers, use the + operand.

Example 6–15 Displaying the Configuration of One Quorum Server

The following example displays the configuration information for the quorum server that uses port 9000. The command displays information for every cluster that has the quorum server

configured as a quorum device. This information includes the cluster name and ID, and the list of reservation and registration keys on the device.

In the following example, nodes with IDs 1, 2, 3, and 4 of cluster `bastille` have registered their keys on the quorum server. Also, because Node 4 owns the quorum device reservation, its key is displayed in the reservation list.

```
# /usr/cluster/bin/clquorumserver show 9000
=== Quorum Server on port 9000 ===
--- Cluster bastille (id 0x439A2EFB) Reservation ---
Node ID:                4
  Reservation key:      0x439a2efb00000004
--- Cluster bastille (id 0x439A2EFB) Registrations ---
Node ID:                1
  Registration key:     0x439a2efb00000001
Node ID:                2
  Registration key:     0x439a2efb00000002
Node ID:                3
  Registration key:     0x439a2efb00000003
Node ID:                4
  Registration key:     0x439a2efb00000004
```

Example 6–16 Displaying the Configuration of Several Quorum Servers

The following example displays the configuration information for three quorum servers, `qs1`, `qs2`, and `qs3`.

```
# /usr/cluster/bin/clquorumserver show qs1 qs2 qs3
```

Example 6–17 Displaying the Configuration of All Running Quorum Servers

The following example displays the configuration information for all running quorum servers:

```
# /usr/cluster/bin/clquorumserver show +
```

Cleaning Up Stale Quorum Server Cluster Information

To remove a quorum device of type `quorumserver`, use the `clquorum remove` command as described in [“How to Remove a Quorum Device” on page 192](#). Under normal operation, this command also removes the quorum server information about the quorum server host.

However, if the cluster loses communications with the quorum server host, removing the quorum device does not clean up this information.

The quorum server cluster information becomes invalid in the following circumstances:

- When a cluster is decommissioned without first removing the cluster quorum device by using the `clquorum remove` command
- When a `quorum_server` type quorum device is removed from a cluster while the quorum server host is down



Caution – If a quorum device of type `quorumserver` is not yet removed from the cluster, using this procedure to clean up a valid quorum server could compromise the cluster quorum.



How to Clean Up the Quorum Server Configuration Information

Before You Begin

Remove the quorum server quorum device from the cluster, as described in [“How to Remove a Quorum Device” on page 192](#).



Caution – If the cluster is still using this quorum server, performing this procedure will compromise cluster quorum.

- 1 **Become the superuser on the quorum server host.**
- 2 **Use the `clquorumserver clear` command to clean up the configuration file.**

```
# clquorumserver clear -c clustername -I clusterID quorumserver [-y]
```

`-c clustername` The name of the cluster that formerly used the quorum server as a quorum device.

You can obtain the cluster name by running `cluster show` on a cluster node.

`-I clusterID` The cluster ID.

The cluster ID is an 8-digit hexadecimal number. You can obtain the cluster ID by running `cluster show` on a cluster node.

`quorumserver` An identifier for one or more quorum servers.

The quorum server can be identified by a port number or an instance name. The port number is used by the cluster nodes to communicate with the quorum server. The instance name is specified in the quorum server configuration file, `/etc/scqsd/scqsd.conf`.

-y Force the `clquorumserver clear` command to clean up cluster information from the configuration file without first prompting for confirmation.

Use this option only if you are confident that you want outdated cluster information to be removed from the quorum server.

- 3 (Optional) If no other quorum devices are configured on this server instance, stop the quorum server.**

Example 6-18 Cleaning Up Outdated Cluster Information From the Quorum Server Configuration

This example removes information about the cluster named `sc-cluster` from the quorum server that uses port 9000.

```
# clquorumserver clear -c sc-cluster -I 0x4308D2CF 9000
```

The quorum server to be unconfigured must have been removed from the cluster. Unconfiguring a valid quorum server could compromise the cluster quorum. Do you want to continue? (yes or no) **y**

Administering Cluster Interconnects and Public Networks

This chapter provides the software procedures for administering the Oracle Solaris 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 Internet Protocol (IP) Network Multipathing groups, when you initially install and configure the cluster. Multipathing is installed automatically with the Oracle Solaris 10 OS, and you must enable it to use it. If you later need to alter a cluster interconnect network configuration, you can use the software procedures in this chapter. For information about configuring IP Network Multipathing groups in a cluster, see the section [“Administering the Public Network” on page 226](#).

This chapter provides information and procedures for the following topics.

- [“Administering the Cluster Interconnects” on page 211](#)
- [“Administering the Public Network” on page 226](#)

For a high-level description of the related procedures in this chapter, see [Table 7-1](#) and [Table 7-3](#).

Refer to the *Oracle Solaris Cluster Concepts Guide* document for background and overview information about the cluster interconnects and public networks.

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 Oracle Solaris Cluster software.

Most of the time, you can use the `clsetup` utility to administer the cluster transport for the cluster interconnects. See the `clsetup(1CL)` man page for more information. All cluster interconnect commands must be run in the global-cluster voting node.

For cluster software installation procedures, see the *Oracle Solaris Cluster Software Installation Guide*. For procedures about servicing cluster hardware components, see the *Oracle Solaris Cluster 3.3 Hardware Administration Manual*.

Note – You can usually choose to use the default port name, where appropriate, during cluster interconnect procedures. The default port name is the same as the internal node ID number of the node that hosts the adapter end of the cable.

TABLE 7-1 Task List: Administering the Cluster Interconnect

Task	Instructions
Administer the cluster transport by using <code>clsetup(1CL)</code>	“How to Access the Cluster Configuration Utilities” on page 27
Check the status of the cluster interconnect by using <code>clinterconnect status</code>	“How to Check the Status of the Cluster Interconnect” on page 213
Add a cluster transport cable, transport adapter, or switch by using <code>clsetup</code>	“How to Add Cluster Transport Cables, Transport Adapters, or Transport Switches” on page 214
Remove a cluster transport cable, transport adapter, or transport switch by using <code>clsetup</code>	“How to Remove Cluster Transport Cables, Transport Adapters, and Transport Switches” on page 216
Enable a cluster transport cable by using <code>clsetup</code>	“How to Enable a Cluster Transport Cable” on page 219
Disable a cluster transport cable by using <code>clsetup</code>	“How to Disable a Cluster Transport Cable” on page 220
Determining an transport adapter's instance number	“How to Determine a Transport Adapter's Instance Number” on page 222
Changing the IP address or the address range of an existing cluster	“How to Change the Private Network Address or Address Range of an Existing Cluster” on page 223

Dynamic Reconfiguration With Cluster Interconnects

You must consider a few issues when completing dynamic reconfiguration (DR) operations on cluster interconnects.

- All of the requirements, procedures, and restrictions that are documented for the Oracle Solaris DR feature also apply to Oracle Solaris Cluster DR support (except for the operating system quiescence operation). Therefore, review the documentation for the Oracle Solaris DR feature *before* using the DR feature with Oracle Solaris Cluster software. You should review in particular the issues that affect non-network IO devices during a DR detach operation.

- The Oracle Solaris Cluster software rejects DR remove-board operations performed on active private interconnect interfaces.
- You must completely remove an active adapter from the cluster in order to perform DR on an active cluster interconnect. Use the `clsetup` menu or the appropriate commands.



Caution – Oracle Solaris Cluster software requires that each cluster node has at least one functioning path to every other cluster node. Do not disable a private interconnect interface that supports the last path to any cluster node.

Complete the following procedures in the order indicated when performing DR operations on public network interfaces.

TABLE 7-2 Task Map: Dynamic Reconfiguration with Public Network Interfaces

Task	Instructions
1. Disable and remove the interface from the active interconnect	“Dynamic Reconfiguration With Public Network Interfaces” on page 227
2. Perform the DR operation on the public network interface.	<i>Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual</i> (from the <i>Solaris 10 on Sun Hardware</i> collection)

▼ How to Check the Status of the Cluster Interconnect

You can also accomplish this procedure by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

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

1 Check the status of the cluster interconnect.

```
% clinterconnect status
```

2 Refer to the following table for common status messages.

Status Message	Description and Possible Action
Path online	The path is currently functioning correctly. No action is necessary.

Status Message	Description and Possible Action
Path waiting	The path is currently being initialized. No action is necessary.
Faulted	The path is not functioning. This can be a transient state when paths are going between the waiting and online state. If the message persists when <code>clinterconnect status</code> is rerun, take corrective action.

Example 7-1 Checking the Status of the Cluster Interconnect

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

```
% clinterconnect status
-- 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, Transport Adapters, or Transport Switches

For information about the requirements for the cluster private transport, see “[Interconnect Requirements and Restrictions](#)” in *Oracle Solaris Cluster 3.3 Hardware Administration Manual*.

You can also accomplish this procedure by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

1 Ensure that the physical cluster transport cables are installed.

For the procedure on installing a cluster transport cable, see the *Oracle Solaris Cluster 3.3 Hardware Administration Manual*.

2 Become superuser on any node in the cluster.

3 Start the `clsetup` utility.

```
# clsetup
```

The Main Menu is displayed.

- 4 Type the number that corresponds to option for displaying the Cluster Interconnect Menu.**
- 5 Type the number that corresponds to the option for adding a transport cable.**
Follow the instructions and type the requested information.
- 6 Type the number that corresponds to the option for adding the transport adapter to a node.**
Follow the instructions and type the requested information.

If you plan to use any of the following adapters for the cluster interconnect, add the relevant entry to the `/etc/system` file on each cluster node. The entry becomes effective after the next system boot.

Adapter	Entry
ce	set ce:ce_taskq_disable=1
ipge	set ipge:ipge_taskq_disable=1
ixge	set ixge:ixge_taskq_disable=1

- 7 Type the number that corresponds to the option for adding the transport switch.**
Follow the instructions and type the requested information.
- 8 Verify that the cluster transport cable, transport adapter, or transport switch is added.**

```
# clinterconnect show node:adapter,adapternode
# clinterconnect show node:adapter
# clinterconnect show node:switch
```

Example 7-2 Adding a Cluster Transport Cable, Transport Adapter, or Transport Switch

The following example shows how to add a transport cable, transport adapter, or transport switch to a node by using the `clsetup` utility.

```
[Ensure that the physical cable is installed.]
[Start the clsetup utility:]
# clsetup
[Select Cluster interconnect]

[Select either Add a transport cable,
Add a transport adapter to a node,
or Add a transport switch.]
[Answer the questions when prompted.]
[You Will Need: ]
[Information:      Example:]
node names        phys-schost-1
adapter names     qfe2
```

```

    switch names          hub2
    transport type       dlpi
[Verify that the clinterconnect
command completed successfully:]Command completed successfully.
Quit the clsetup Cluster Interconnect Menu and Main Menu.
[Verify that the cable, adapter, and switch are added:]
# clinterconnect show phys-schost-1:qfe2,hub2
===Transport Cables ===
Transport Cable:                phys-schost-1:qfe2@0,hub2
Endpoint1:                      phys-schost-2:qfe0@0
Endpoint2:                      ethernet-1@2 ??? Should this be hub2?
State:                          Enabled

# clinterconnect show phys-schost-1:qfe2
=== Transport Adepters for qfe2
Transport Adapter:              qfe2
Adapter State:                 Enabled
Adapter Transport Type:        dlpi
Adapter Property (device_name): ce
Adapter Property (device_instance): 0
Adapter Property (lazy_free):  1
Adapter Property (dlpi_heartbeat_timeout): 10000
Adpater Property (dlpi_heartbeat_quantum): 1000
Adapter Property (nw_bandwidth): 80
Adapter Property (bandwidth):   70
Adapter Property (ip_address):  172.16.0.129
Adapter Property (netmask):     255.255.255.128
Adapter Port Names:            0
Adapter Port SState (0):       Enabled

# clinterconnect show phys-schost-1:hub2
=== Transport Switches ===
Transport Switch:              hub2
Switch State:                 Enabled
Switch Type:                  switch
Switch Port Names:            1 2
Switch Port State(1):         Enabled
Switch Port State(2):         Enabled

```

Next Steps To check the interconnect status of your cluster transport cable see [“How to Check the Status of the Cluster Interconnect”](#) on page 213.

▼ How to Remove Cluster Transport Cables, Transport Adapters, and Transport Switches

You can also accomplish this procedure by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

Use the following procedure to remove cluster transport cables, transport adapters, and transport switches from a node configuration. 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 on a transport cable.



Caution – Each cluster node needs at least one functioning transport path to every other node in the cluster. No two nodes should be isolated from one another. 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, ensure that another connection is available. Disabling a node's last remaining working cable takes the node out of cluster membership.

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

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

```
# clinterconnect status
```



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 operational path takes the node out of cluster membership and could result in a cluster reconfiguration.

- 3 **Start the `clsetup` utility.**

```
# clsetup
```

The Main Menu is displayed.
- 4 **Type the number that corresponds to the option for accessing the Cluster Interconnect menu.**
- 5 **Type the number that corresponds to the option for disabling the transport cable.**

Follow the instructions and type the requested information. You need to know the applicable node names, adapter names, and switch names.
- 6 **Type the number that corresponds to the option for removing the transport cable.**

Follow the instructions and type the requested information. You need to know the applicable node names, adapter names, and switch names.

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

7 Type the number that corresponds to the option for removing the transport adapter from a node.

Follow the instructions and type the requested information. You need to know the applicable node names, adapter names, and switch names.

Note – If you are removing a physical adapter from a node, see the *Oracle Solaris Cluster 3.3 Hardware Administration Manual* for hardware service procedures.

8 Type the number that corresponds to the option for removing a transport switch.

Follow the instructions and type the requested information. You need to know the applicable node names, adapter names, and switch names.

Note – A switch cannot be removed if any of the ports are still in use as endpoints on any transport cables.

9 Verify that the cable, adapter, or switch has been removed.

```
# clinterconnect show node:adapter,adapternode
# clinterconnect show node:adapter
# clinterconnect show node:switch
```

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

Example 7-3 Removing a Transport Cable, Transport Adapter, or Transport Switch

The following example shows how to remove a transport cable, transport adapter, or transport switch by using the `clsetup` command.

```
[Become superuser on any node in the cluster.]
[Start the utility:]
# clsetup
[Select Cluster interconnect.]
[Select either Remove a transport cable,
Remove a transport adapter to a node,
or Remove a transport switch.]
[Answer the questions when prompted.]
  You Will Need:
  Information           Example:
  node names            phys-schost-1
  adapter names         qfe1
  switch names          hub1
[Verify that the clinterconnect
command was completed successfully:]
Command completed successfully.
[Quit the clsetup utility Cluster Interconnect Menu and Main Menu.]
[Verify that the cable, adapter, or switch is removed:]
# clinterconnect show phys-schost-1:qfe2,hub2
===Transport Cables===
```

```

Transport Cable:                phys-schost-2:qfe2@0, hub2
  Cable Endpoint1:              phys-schost-2:qfe0@0
  Cable Endpoint2:              ethernet-1@2 ??? Should this be hub2???
  Cable State:                  Enabled

# clinterconnect show phys-schost-1:qfe2
=== Transport Adepters for qfe2
Transport Adapter:              qfe2
  Adapter State:                Enabled
  Adapter Transport Type:       dlpi
  Adapter Property (device_name): ce
  Adapter Property (device_instance): 0
  Adapter Property (lazy_free): 1
  Adapter Property (dlpi_heartbeat_timeout): 10000
  Adapter Property (dlpi_heartbeat_quantum): 1000
  Adapter Property (nw_bandwidth): 80
  Adapter Property (bandwidth): 70
  Adapter Property (ip_address): 172.16.0.129
  Adapter Property (netmask): 255.255.255.128
  Adapter Port Names:          0
  Adapter Port SState (0):     Enabled

# clinterconnect show phys-schost-1:hub2
=== Transport Switches ===
Transport Switch:              hub2
  Switch State:                 Enabled
  Switch Type:                  switch
  Switch Port Names:           1 2
  Switch Port State(1):        Enabled
  Switch Port State(2):        Enabled

```

▼ How to Enable a Cluster Transport Cable

You can also accomplish this procedure by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

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

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser on any node in the cluster.**
- 2 **Start the `clsetup` utility.**

```
# clsetup
```

The Main Menu is displayed.

- 3 Type the number that corresponds to the option for accessing the Cluster Interconnect menu and press the Return key.
- 4 Type the number that corresponds to the option for enabling the transport cable and press the Return key.

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

- 5 Verify that the cable is enabled.

```
# clinterconnect show node:adapter,adapternode
```

Example 7-4 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.]
[Start the clsetup utility:]
# clsetup
[Select Cluster interconnect>Enable a transport cable.]

[Answer the questions when prompted.]
[You will need the following information.]
  You Will Need:
Information:                               Example:
  node names                               phys-schost-2
  adapter names                             qfe1
  switch names                              hub1
[Verify that the scinterconnect
command was completed successfully:]

clinterconnect enable phys-schost-2:qfe1

Command completed successfully.
[Quit the clsetup Cluster Interconnect Menu and Main Menu.]
[Verify that the cable is enabled:]
# clinterconnect show phys-schost-1:qfe2,hub2
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 can also accomplish this procedure by using the Oracle Solaris Cluster Manager GUI. See the Oracle Solaris Cluster Manager online help for more information.

You might need to disable a cluster transport cable to temporarily shut down a cluster interconnect path. A temporary shutdown 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 – Each cluster node needs at least one functioning transport path to every other node in the cluster. No two nodes should be isolated from one another. 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, ensure that another connection is available. Disabling a node's last remaining working cable takes the node out of cluster membership.

The `phys - s chost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser on any node in the cluster.**
- 2 **Check the status of the cluster interconnect before disabling a cable.**

```
# clinterconnect status
```



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 operational path takes the node out of cluster membership and could result in a cluster reconfiguration.

- 3 **Start the `clsetup` utility.**
- 4 **Type the number that corresponds to the option for accessing the Cluster Interconnect Menu and press the Return key.**
- 5 **Type the number that corresponds to the option for disabling the transport cable and press the Return key.**

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

- 6 **Verify that the cable is disabled.**
- ```
clinterconnect show node:adapter,adapternode
```

**Example 7-5** 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.]
[Start the clsetup utility:]
clsetup
[Select Cluster interconnect>Disable a transport cable.]

[Answer the questions when prompted.]
[You will need the following information.]
 [You Will Need:]
Information: Example:
node names phys-schost-2
adapter names qfe1
switch names hub1
[Verify that the clinterconnect
command was completed successfully:]
Command completed successfully.
[Quit the clsetup Cluster Interconnect Menu and Main Menu.]
[Verify that the cable is disabled:]
clinterconnect show -p phys-schost-1:qfe2,hub2
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
```

## ▼ How to Determine a Transport Adapter's Instance Number

You need to determine a transport adapter's instance number to ensure that you add and remove the correct transport adapter through the clsetup command. The adapter name is a combination of the type of the adapter and the adapter's instance number.

### 1 Based on the slot number, find the adapter's name.

The following screen is an example and might not reflect your hardware.

```
prtdiag
...
===== IO Cards =====
 Bus Max
 Bus
IO Port Bus Freq Bus Dev,
Type ID Side Slot MHz Freq Func State Name Model

XYZ 8 B 2 33 33 2,0 ok xyz11c8,0-xyz11c8,d665.11c8.0.0
XYZ 8 B 3 33 33 3,0 ok xyz11c8,0-xyz11c8,d665.11c8.0.0
...

```

**2 Using the adapter's path, find the adapter's instance number.**

The following screen is an example and might not reflect your hardware.

```
grep sci /etc/path_to_inst
"/xyz@1f,400/pci11c8,0@2" 0 "ttt"
"/xyz@1f,4000.pci11c8,0@4 "ttt"
```

**3 Using the adapter's name and slot number, find the adapter's instance number.**

The following screen is an example and might not reflect your hardware.

```
prtconf
...
xyz, instance #0
 xyz11c8,0, instance #0
 xyz11c8,0, instance #1
...
```

## ▼ How to Change the Private Network Address or Address Range of an Existing Cluster

Use this procedure to change a private network address or the range of network addresses used or both.

**Before You Begin** Ensure that remote shell (rsh(1M)) or secure shell (ssh(1)) access for superuser is enabled to all cluster nodes.

**1 Reboot all cluster nodes into noncluster mode by performing the following substeps on each cluster node:**

a. **Become superuser or assume a role that provides `soLaris.cluster.admin` RBAC authorization on the cluster node to be started in noncluster mode.**

b. **Shut down the node by using the `clnode evacuate` and `cluster shutdown` commands.**

The `clnode evacuate` command switches over all device groups from the specified node to the next-preferred node. The command also switches all resource groups from voting or non-voting nodes on the specified node to the next-preferred voting or non-voting node.

```
clnode evacuate node
cluster shutdown -g0 -y
```

**2 From one node, start the `clsetup` utility.**

When run in noncluster mode, the `clsetup` utility displays the Main Menu for noncluster-mode operations.

- 3 Type the number that corresponds to the option for Change IP Address Range and press the Return key.**

The `clsetup` utility displays the current private-network configuration, then asks if you would like to change this configuration.

- 4 To change either the private-network IP address or the IP address range, type `yes` and press the Return key.**

The `clsetup` utility displays the default private-network IP address, `172.16.0.0`, and asks if it is okay to accept this default.

- 5 Change or accept the private-network IP address.**

- **To accept the default private-network IP address and proceed to changing the IP address range, type `yes` and press the Return key.**

The `clsetup` utility will ask if it is okay to accept the default netmask. Skip to the next step to enter your response.

- **To change the default private-network IP address, perform the following substeps.**

- a. Type `no` in response to the `clsetup` utility question about whether it is okay to accept the default address, then press the Return key.**

The `clsetup` utility will prompt for the new private-network IP address.

- b. Type the new IP address and press the Return key.**

The `clsetup` utility displays the default netmask and then asks if it is okay to accept the default netmask.

- 6 Change or accept the default private-network IP address range.**

The default netmask is `255.255.240.0`. This default IP address range supports up to 64 nodes, 12 zone clusters, and 10 private networks in the cluster.

- **To accept the default IP address range, type `yes` and press the Return key.**

Then skip to the next step.

- **To change the IP address range, perform the following substeps.**

- a. Type `no` in response to the `clsetup` utility's question about whether it is okay to accept the default address range, then press the Return key.**

When you decline the default netmask, the `clsetup` utility prompts you for the number of nodes and private networks, and zone clusters that you expect to configure in the cluster.

**b. Enter the number of nodes, private networks, and zone clusters that you expect to configure in the cluster.**

From these numbers, the `clsetup` utility calculates two proposed netmasks:

- The first netmask is the minimum netmask to support the number of nodes, private networks, and zone clusters that you specified.
- The second netmask supports twice the number of nodes, private networks, and zone clusters that you specified, to accommodate possible future growth.

**c. Specify either of the calculated netmasks, or specify a different netmask that supports the expected number of nodes, private networks, and zone clusters.**

**7 Type yes in response to the `clsetup` utility's question about proceeding with the update.**

**8 When finished, exit the `clsetup` utility.**

**9 Reboot each cluster node back into cluster mode by completing the following substeps for each cluster node:**

**a. Boot the node.**

- On SPARC based systems, run the following command.

```
ok boot
```

- On x86 based systems, run the following commands.

When the GRUB menu is displayed, select the appropriate Oracle Solaris entry and press Enter. The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
```

```
+-----+
| Solaris 10 /sol_10_x86 |
| Solaris failsafe |
| |
+-----+-----+
```

Use the `^` and `v` keys to select which entry is highlighted. Press enter to boot the selected OS, `'e'` to edit the commands before booting, or `'c'` for a command-line.

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

```
cluster status -t node
```

# Administering the Public Network

Oracle Solaris Cluster software supports the Oracle Solaris software implementation of Internet Protocol network Multipathing (IPMP) for public networks. Basic IPMP administration is the same for both cluster and noncluster environments. Multipathing is automatically installed when you install the Oracle Solaris 10 OS, and you must enable it to use it. Multipathing administration is covered in the appropriate Oracle Solaris OS documentation. However, review the guidelines that follow before administering IPMP in an Oracle Solaris Cluster environment.

## How to Administer IP Network Multipathing Groups in a Cluster

Before performing IPMP procedures on a cluster, consider the following guidelines.

- Each public network adapter must belong to an IPMP group.
- The `local-mac-address?` variable must have a value of `true` for Ethernet adapters.
- You can use probe-based IPMP groups or link-based IPMP groups in a cluster. A probe-based IPMP group tests the target IP address and provides the most protection by recognizing more conditions that might compromise availability.
- You must configure a test IP address for each adapter in the following kinds of multipathing groups:
  - All multiple-adapter multipathing groups require test IP addresses. Single-adapter multipathing groups do not require test IP addresses.
- Test IP addresses for all adapters in the same multipathing group must belong to a single IP subnet.
- Test IP addresses must not be used by normal applications because they are not highly available.
- No restrictions are placed on multipathing group naming. However, when configuring a resource group, the `netiflist` naming convention is any multipathing name followed by either the `nodeID` number or the node name. For example, given a multipathing group named `sc_ipmp0`, the `netiflist` naming could be either `sc_ipmp0@1` or `sc_ipmp0@phys-schost-1`, where the adapter is on the node `phys-schost-1`, which has the `nodeID` of 1.
- Avoid unconfiguring (unplumbing) or bringing down an adapter of an IP Network Multipathing group without first switching over the IP addresses from the adapter to be removed to an alternate adapter in the group, using the `if_mpadm(1M)` command.
- Avoid rewiring adapters to different subnets without first removing them from their respective multipathing groups.

- Logical adapter operations can be done on an adapter even if monitoring is on for the multipathing 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.
- To view the status of IP Network Multipathing groups on a cluster, use the `command.clinternconnect status` command

For more information about IP Network Multipathing, see the appropriate documentation in the Oracle Solaris OS system administration documentation set.

TABLE 7-3 Task Map: Administering the Public Network

| Oracle Solaris OS Release | Instructions                                                                        |
|---------------------------|-------------------------------------------------------------------------------------|
| Oracle Solaris 10 OS      | “IP Network Multipathing Topics” in <i>System Administration Guide: IP Services</i> |

For cluster software installation procedures, see the *Oracle Solaris Cluster Software Installation Guide*. For procedures about servicing public networking hardware components, see the *Oracle Solaris Cluster 3.3 Hardware Administration Manual*.

## Dynamic Reconfiguration With Public Network Interfaces

You must consider a few issues when completing dynamic reconfiguration (DR) operations on public network interfaces in a cluster.

- All of the requirements, procedures, and restrictions that are documented for the Oracle Solaris DR feature also apply to Oracle Solaris Cluster DR support (except for the operating system quiescence operation). Therefore, review the documentation for the Oracle Solaris DR feature *before* using the DR feature with Oracle Solaris Cluster software. You should review in particular the issues that affect non-network IO devices during a DR detach operation.
- DR remove-board operations can succeed only when public network interfaces are not active. Before removing an active public network interface, switch the IP addresses from the adapter to be removed to another adapter in the multipathing group, using the `if_mpadm(1M)` command.
- If you try to remove a public network interface card without having properly disabled it as an active network interface, Oracle Solaris Cluster rejects the operation and identifies the interface that would be affected by the operation.



**Caution** – For multipathing groups with two adapters, if the remaining network adapter fails while you are performing the DR remove operation on the disabled network adapter, availability is impacted. The remaining adapter has no place to fail over for the duration of the DR operation.

Complete the following procedures in the order indicated when performing DR operations on public network interfaces.

TABLE 7-4 Task Map: Dynamic Reconfiguration With Public Network Interfaces

| Task                                                                                                                                    | Instructions                                                                                                                                                                                                              |
|-----------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1. Switch the IP addresses from the adapter to be removed to another adapter in the multipathing group, using the <code>if_mpadm</code> | <code>if_mpadm(1M)</code> man page.<br><i>Part VI, “IPMP,” in <a href="#">System Administration Guide: IP Services</a></i>                                                                                                |
| 2. Remove the adapter from the multipathing group by using the <code>ifconfig</code> command                                            | <code>ifconfig(1M)</code> man page<br><i>Part VI, “IPMP,” in <a href="#">System Administration Guide: IP Services</a></i>                                                                                                 |
| 3. Perform the DR operation on the public network interface                                                                             | <i><a href="#">Sun Enterprise 10000 DR Configuration Guide</a> and the <a href="#">Sun Enterprise 10000 Dynamic Reconfiguration Reference Manual</a> (from the <a href="#">Solaris 10 on Sun Hardware</a> collection)</i> |

# Adding and Removing a Node

---

This chapter provides instructions on how to add a node to a cluster and how to remove a node:

- “Adding a Node to a Cluster” on page 229
- “Removing a Node From a Cluster” on page 235

For information about cluster maintenance tasks, see [Chapter 9, “Administering the Cluster.”](#)

## Adding a Node to a Cluster

This section describes how to add a node to a global cluster or a zone cluster. You can create a new zone-cluster node on a node of the global cluster that hosts the zone cluster, as long as that global-cluster node does not already host a node of that particular zone cluster. You cannot convert an existing non-voting node on a global cluster into a zone-cluster node.

In this chapter, `phys - s chost#` reflects a global-cluster prompt. The `clzonecluster` interactive shell prompt is `clzc : schost>`.

The following table lists the tasks to perform to add a node to an existing cluster. Perform the tasks in the order shown.

**TABLE 8-1** Task Map: Adding a Node to an Existing Global or Zone Cluster

| Task                                                                                                             | Instructions                                                     |
|------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------|
| Install the host adapter on the node and verify that the existing cluster interconnects can support the new node | <i>Oracle Solaris Cluster 3.3 Hardware Administration Manual</i> |
| Add shared storage                                                                                               | <i>Oracle Solaris Cluster 3.3 Hardware Administration Manual</i> |

TABLE 8-1 Task Map: Adding a Node to an Existing Global or Zone Cluster (Continued)

| Task                                                                                                                                                           | Instructions                                                                                                                          |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|
| Add the node to the authorized node list by using <code>clsetup</code>                                                                                         | “How to Add a Node to the Authorized Node List” on page 230                                                                           |
| Install and configure the software on the new cluster node                                                                                                     | Chapter 2, “Installing Software on Global-Cluster Nodes,” in <i>Oracle Solaris Cluster Software Installation Guide</i>                |
| Add the new node to an existing cluster                                                                                                                        | “Adding a Node to a Cluster” on page 229                                                                                              |
| If the cluster is configured in an Oracle Solaris Cluster Geographic Edition partnership, configure the new node as an active participant in the configuration | “How to Add a New Node to a Cluster in a Partnership” in <i>Oracle Solaris Cluster Geographic Edition System Administration Guide</i> |

## ▼ How to Add a Node to the Authorized Node List

Before adding an Oracle Solaris host or a virtual machine to an existing global cluster or a zone cluster, ensure that the node has all of the necessary hardware correctly installed and configured, including an operational physical connection to the private cluster interconnect.

For hardware installation information, refer to the *Oracle Solaris Cluster 3.3 Hardware Administration Manual* or the hardware documentation that shipped with your server.

This procedure enables a machine to install itself into a cluster by adding its node name to the list of authorized nodes for that cluster.

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 On a current global-cluster member, become superuser on the current cluster member. Perform these steps from a node of a global cluster.**
- 2 Ensure that you have correctly completed all prerequisite hardware installation and configuration tasks that are listed in the task map for Table 8-1.**
- 3 Start the `clsetup` utility.**

```
phys - schost# clsetup
```

The Main Menu is displayed.

---

**Note** – To add a node to a zone cluster, use the `clzonecluster` utility. See Step 9 for instructions to manually add a zone to a zone cluster.

---

- 4 **Type the number that corresponds to the option for displaying the New Nodes Menu and press the Return key.**
- 5 **Type the number that corresponds to the option to modify the authorized list and press the Return key. Specify the name of a machine that can add itself.**  
Follow the prompts to add the node's name to the cluster. You are asked for the name of the node to be added.
- 6 **Verify that the task has been performed successfully.**  
The `clsetup` utility prints a “Command completed successfully” message if it completes the task without error.
- 7 **To prevent any new machines from being added to the cluster, type the number that corresponds to option to instruct the cluster to ignore requests to add new machines. Press the Return key.**  
Follow the `clsetup` prompts. This option tells the cluster to ignore all requests over the public network from any new machine that is trying to add itself to the cluster.
- 8 **Quit the `clsetup` utility.**
- 9 **To manually add a node to a zone cluster, you must specify the Oracle Solaris host and the virtual node name. You must also specify a network resource to be used for public network communication on each node. In the following example, the zone name is `sczone`, and `bge0` is the public network adapter on both machines.**

```
clzc: sczone>add node
clzc: sczone: node>set physical-host=phys-cluster-1
clzc: sczone: node>set hostname=hostname1
clzc: sczone: node>add net
clzc: sczone: node: net>set address=hostname1
clzc: sczone: node: net>set physical=bge0
clzc: sczone: node: net>end
clzc: sczone: node>end
clzc: sczone>add node
clzc: sczone: node>set physical-host=phys-cluster-2
clzc: sczone: node>set hostname=hostname2
clzc: sczone: node>add net
clzc: sczone: node: net>set address=hostname2
clzc: sczone: node: net>set physical=bge0
clzc: sczone: node: net>end
clzc: sczone: node>end
```

For detailed instructions on configuring the node, see “Configuring a Zone Cluster” in *Oracle Solaris Cluster Software Installation Guide*.

**10 Install and configure the software on the new cluster node.**

Use either `cluster create` or JumpStart software to complete the installation and configuration of the new node, as described in the *Oracle Solaris Cluster Software Installation Guide*.

**Example 8–1 Adding a Global-Cluster Node to the Authorized Node List**

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

```
[Become superuser and execute the clsetup utility.]
phys-schost# clsetup
[Select New nodes>Specify the name of a machine which may add itself.]
[Answer the questions when prompted.]
[Verify that the command completed successfully.]

claccess allow -h phys-schost-3

 Command completed successfully.
[Select Prevent any new machines from being added to the cluster.]
[Quit the clsetup New Nodes Menu and Main Menu.]
[Install the cluster software.]
```

**See Also** [clsetup\(1CL\)](#)

For a complete list of tasks for adding a cluster node, see [Table 8–1](#), “Task Map: Adding a Cluster Node.”

To add a node to an existing resource group, see the *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

## Creating a Non-Voting Node (Zone) in a Global Cluster

This section provides the following information and procedures to create a non-voting node, simply referred to as a *zone*, on a global-cluster node.

**▼ How to Create a Non-Voting Node in a Global Cluster****1 Become superuser on the global-cluster node where you are creating the non-voting node.**

You must be working in the global zone.

## 2 Verify on each node that multiuser services for the Service Management Facility (SMF) are online.

If services are not yet online for a node, wait until the state changes to online before you proceed to the next step.

```
phys-schost# svcs multi-user-server node
STATE STIME FMRI
online 17:52:55 svc:/milestone/multi-user-server:default
```

## 3 Configure, install, and boot the new zone.

---

**Note** – You must set the autoboot property to true to support resource-group functionality in the non-voting node on the global cluster.

---

Follow procedures in the Solaris documentation:

- a. Perform procedures in [Chapter 18, “Planning and Configuring Non-Global Zones \(Tasks\),”](#) in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.
- b. Perform procedures in [“Installing and Booting Zones”](#) in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.
- c. Perform procedures in [“How to Boot a Zone”](#) in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

## 4 Verify that the zone is in the ready state.

```
phys-schost# zoneadm list -v
ID NAME STATUS PATH
0 global running /
1 my-zone ready /zone-path
```

## 5 (Optional) For a shared-IP zone, assign a private IP address and a private hostname to the zone.

The following command chooses and assigns an available IP address from the cluster's private IP-address range. The command also assigns the specified private hostname, or host alias, to the zone and maps it to the assigned private IP address.

```
phys-schost# clnode set -p zprivatehostname=hostalias node:zone
```

|                            |                                                     |
|----------------------------|-----------------------------------------------------|
| -p                         | Specifies a property.                               |
| zprivatehostname=hostalias | Specifies the zone private hostname, or host alias. |
| node                       | The name of the node.                               |
| zone                       | The name of the global-cluster non-voting node.     |

**6 Perform the initial internal zone configuration.**

Follow the procedures in “Performing the Initial Internal Zone Configuration” in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*. Choose either of the following methods:

- Log in to the zone.
- Use an `/etc/sysidcfg` file.

**7 In the non-voting node, modify the `nsswitch.conf` file.**

These changes enable the zone to resolve searches for cluster-specific hostnames and IP addresses.

**a. Log in to the zone.**

```
phys-schost# zlogin -c zonename
```

**b. Open the `/etc/nsswitch.conf` file for editing.**

```
sczone# vi /etc/nsswitch.conf
```

**c. Add the `cluster` switch to the beginning of the `lookups` for the `hosts` and `netmasks` entries, followed by the `files` switch.**

The modified entries should appear similar to the following:

```
...
hosts: cluster files nis [NOTFOUND=return]
...
netmasks: cluster files nis [NOTFOUND=return]
...
```

**d. For all other entries, ensure that the `files` switch is the first switch that is listed in the entry.****e. Exit the zone.****8 If you created an exclusive-IP zone, configure IPMP groups in each `/etc/hostname.interface` file that is on the zone.**

You must configure an IPMP group for each public-network adapter that is used for data-service traffic in the zone. This information is not inherited from the global zone. See “Public Networks” in *Oracle Solaris Cluster Software Installation Guide* for more information about configuring IPMP groups in a cluster.

**9 Set up name-to-address mappings for all logical hostname resources that are used by the zone.****a. Add name-to-address mappings to the `/etc/inet/hosts` file on the zone.**

This information is not inherited from the global zone.

**b. If you use a name server, add the name-to-address mappings.**

## Removing a Node From a Cluster

This section provides instructions on how to remove a node on a global cluster or a zone cluster. You can also remove a specific zone cluster from a global cluster. The following table lists the tasks to perform to remove a node from an existing cluster. Perform the tasks in the order shown.



**Caution** – If you remove a node using only this procedure for a RAC configuration, the removal might cause the node to panic during a reboot. For instructions on how to remove a node from a RAC configuration, see [“How to Remove Support for Oracle RAC From Selected Nodes”](#) in *Oracle Solaris Cluster Data Service for Oracle Real Application Clusters Guide*. After you complete that process, follow the appropriate steps below.

TABLE 8-2 Task Map: Removing a Node

| Task                                                                             | Instructions                                                                                                                                                                                                                                                                                                                                                                                                              |
|----------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Move all resource groups and device groups off the node to be removed            | <code>clnode evacuate node</code>                                                                                                                                                                                                                                                                                                                                                                                         |
| Verify that the node can be removed by checking the allowed hosts                | <code>claccess show node</code><br><code>claccess allow -h node-to-remove</code>                                                                                                                                                                                                                                                                                                                                          |
| If the node cannot be removed, give the node access to the cluster configuration |                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Remove the node from all device groups                                           | <a href="#">“How to Remove a Node From a Device Group (Solaris Volume Manager)”</a> on page 133                                                                                                                                                                                                                                                                                                                           |
| Remove all quorum devices connected to the node being removed                    | <b>This step is optional if you are removing a node from a two-node cluster.</b><br><br><a href="#">“How to Remove a Quorum Device”</a> on page 192<br><br>Note that although you must remove the quorum device before you remove the storage device in the next step, you can add the quorum device back immediately afterward.<br><br><a href="#">“How to Remove the Last Quorum Device From a Cluster”</a> on page 194 |
| Put the node being removed into noncluster mode                                  | <a href="#">“How to Put a Node Into Maintenance State”</a> on page 260                                                                                                                                                                                                                                                                                                                                                    |
| Remove a node from a zone cluster                                                | <a href="#">“How to Remove a Node From a Zone Cluster”</a> on page 236                                                                                                                                                                                                                                                                                                                                                    |
| Remove a node from the cluster software configuration                            | <a href="#">“How to Remove a Node From the Cluster Software Configuration”</a> on page 236                                                                                                                                                                                                                                                                                                                                |

TABLE 8-2 Task Map: Removing a Node (Continued)

| Task                                                                     | Instructions                                                                       |
|--------------------------------------------------------------------------|------------------------------------------------------------------------------------|
| (Optional) Uninstall Oracle Solaris Cluster software from a cluster node | “How to Uninstall Oracle Solaris Cluster Software From a Cluster Node” on page 271 |
| Remove an entire zone cluster                                            | “How to Remove a Zone Cluster” on page 267                                         |

## ▼ How to Remove a Node From a Zone Cluster

You can remove a node from a zone cluster by halting the node, uninstalling it, and removing the node from the configuration. If you decide later to add the node back into the zone cluster, follow the instructions in [Table 8-1](#). Most of these steps are performed from the global-cluster node.

- 1 **Become superuser on a node of the global cluster.**
- 2 **Shut down the zone-cluster node you want to remove by specifying the node and its zone cluster.**

```
phys-schost# clzonecluster halt -n node zoneclustername
```

You can also use the `clnode evacuate` and `shutdown` commands within a zone cluster.

- 3 **Uninstall the zone-cluster node.**

```
phys-schost# clzonecluster uninstall -n node zoneclustername
```

- 4 **Remove the zone-cluster node from the configuration.**

Use the following commands:

```
phys-schost# clzonecluster configure zoneclustername
```

```
clzc:sczone> remove node physical-host=zoneclusternodename
```

- 5 **Verify that the node was removed from the zone cluster.**

```
phys-schost# clzonecluster status
```

## ▼ How to Remove a Node From the Cluster Software Configuration

Perform this procedure to remove a node from the global cluster.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Ensure that you have removed the node from all resource groups, device groups, and quorum device configurations and put it into maintenance state before you continue with this procedure.**
- 2 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on the node that you want to remove. Perform all steps in this procedure from a node of the global cluster.**
- 3 **Boot the global-cluster node that you want to remove into noncluster mode. For a zone-cluster node, follow the instructions in “[How to Remove a Node From a Zone Cluster](#)” on page 236 before you perform this step.**

- On SPARC based systems, run the following command.

```
ok boot -x
```

- On x86 based systems, run the following commands.

```
shutdown -g -y -i0
```

Press any key to continue

- a. **In the GRUB menu, use the arrow keys to select the appropriate Oracle Solaris entry and type `e` to edit its commands.**

The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
```

```
+-----+
| Solaris 10 /sol_10_x86 |
| Solaris failsafe |
| |
+-----+
+-----+
```

Use the `^` and `v` keys to select which entry is highlighted.

Press `enter` to boot the selected OS, `'e'` to edit the commands before booting, or `'c'` for a command-line.

For more information about GRUB based booting, see “[Booting an x86 Based System by Using GRUB \(Task Map\)](#)” in *System Administration Guide: Basic Administration*.

- b. **In the boot parameters screen, use the arrow keys to select the kernel entry and type `e` to edit the entry.**

The GRUB boot parameters screen appears similar to the following:

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
```

```
+-----+
| root (hd0,0,a) |
| kernel /platform/i86pc/multiboot |
| module /platform/i86pc/boot_archive |
+-----+
```

```
+-----+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.
```

**c. Add -x to the command to specify system boot into noncluster mode.**

[ Minimal BASH-like line editing is supported. For the first word, TAB lists possible command completions. Anywhere else TAB lists the possible completions of a device/filename. ESC at any time exits. ]

```
grub edit> kernel /platform/i86pc/multiboot -x
```

**d. Press the Enter key to accept the change and return to the boot parameters screen.**

The screen displays the edited command.

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
```

```
+-----+
| root (hd0,0,a) |
| kernel /platform/i86pc/multiboot -x |
| module /platform/i86pc/boot_archive |
+-----+
```

```
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.-
```

**e. Type b to boot the node into noncluster mode.**

This change to the kernel boot parameter command does not persist over the system boot. The next time you reboot the node, it will boot into cluster mode. To boot into noncluster mode instead, perform these steps again to add the -x option to the kernel boot parameter command.

---

**Note** – If the node to be removed is not available or can no longer be booted, run the following command on any active cluster node: **clnode clear -F <node-to-be-removed>**. Verify the node removal by running **clnode status <nodename>**.

---

**4 From the node you want to remove, delete the node from the cluster.**

```
phys-schost# clnode remove -F
```

If the **clnode remove** command fails and a stale node reference exists, run **clnode clear -F nodename** on an active node.

---

**Note** – If you are removing the last node in the cluster, the node must be in noncluster mode with no active nodes left in the cluster.

---

**5 From another cluster node, verify the node removal.**

```
phys-schost# clnode status nodename
```

**6 Complete the node removal.**

- If you intend to uninstall the Oracle Solaris Cluster software from the removed node, proceed to [“How to Uninstall Oracle Solaris Cluster Software From a Cluster Node” on page 271](#).
- If you do not intend to uninstall the Oracle Solaris Cluster software from the removed node, you can physically remove the node from the cluster by removing the hardware connections as described in *Oracle Solaris Cluster 3.3 Hardware Administration Manual*.

**Example 8–2 Removing a Node From the Cluster Software Configuration**

This example shows how to remove a node (phys-schost-2) from a cluster. The `clnode remove` command is run in noncluster mode from the node you want to remove from the cluster (phys-schost-2).

```
[Remove the node from the cluster:]
phys-schost-2# clnode remove
phys-schost-1# clnode clear -F phys-schost-2
[Verify node removal:]
phys-schost-1# clnode status
-- Cluster Nodes --
 Node name Status

Cluster node: phys-schost-1 Online
```

**See Also** To uninstall Oracle Solaris Cluster software from the removed node, see [“How to Uninstall Oracle Solaris Cluster Software From a Cluster Node” on page 271](#).

For hardware procedures, see the *Oracle Solaris Cluster 3.3 Hardware Administration Manual*.

For a comprehensive list of tasks for removing a cluster node, see [Table 8–2](#).

To add a node to an existing cluster, see [“How to Add a Node to the Authorized Node List” on page 230](#).

## ▼ How to Remove a Non-Voting Node (Zone) From a Global Cluster

- 1 Become superuser on the global-cluster node where you created the non-voting node.
- 2 Delete the non-voting node from the system.

Follow the procedures in “Deleting a Non-Global Zone From the System” in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

## ▼ How to Remove Connectivity Between an Array and a Single Node, in a Cluster With Greater Than Two-Node Connectivity

Use this procedure to detach a storage array from a single cluster node, in a cluster that has three-node or four-node connectivity.

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Back up all database tables, data services, and volumes that are associated with the storage array that you are removing.**
- 2 **Determine the resource groups and device groups that are running on the node to be disconnected.**

```
phys - schost# clresourcegroup status
phys - schost# cldevicegroup status
```

- 3 **If necessary, move all resource groups and device groups off the node to be disconnected.**




---

**Caution (SPARC only)** – If your cluster is running Oracle RAC software, shut down the Oracle RAC database instance that is running on the node before you move the groups off the node. For instructions, see the *Oracle Database Administration Guide*.

---

```
phys - schost# clnode evacuate node
```

The `clnode evacuate` command switches over all device groups from the specified node to the next-preferred node. The command also switches all resource groups from voting or non-voting nodes on the specified node to the next-preferred voting or non-voting node.

**4 Put the device groups into maintenance state.**

For the procedure on acquiescing I/O activity to Veritas shared disk groups, see your VxVM documentation.

For the procedure on putting a device group in maintenance state, see [“How to Put a Node Into Maintenance State” on page 260](#).

**5 Remove the node from the device groups.**

If you use VxVM or a raw disk, use the `cldevicegroup(1CL)` command to remove the device groups.

**6 For each resource group that contains an HAS`storagePlus` resource, remove the node from the resource group's node list.**

```
phys-schost# clresourcegroup remove-node -z zone -n node + | resourcegroup
```

*node*     The name of the node.

*zone*     The name of the non-voting node that can master the resource group. Specify *zone* only if you specified a non-voting node when you created the resource group.

See the [Oracle Solaris Cluster Data Services Planning and Administration Guide](#) for more information about changing a resource group's node list.

---

**Note** – Resource type, resource group, and resource property names are case sensitive when `clresourcegroup` is executed.

---

**7 If the storage array that you are removing is the last storage array that is connected to the node, disconnect the fiber-optic cable between the node and the hub or switch that is connected to this storage array (otherwise, skip this step).**

**8 If you are removing the host adapter from the node that you are disconnecting, and power off the node. If you are removing the host adapter from the node that you are disconnecting, skip to [Step 11](#).**

**9 Remove the host adapter from the node.**

For the procedure on removing host adapters, see the documentation for the node.

**10 Without booting the node, power on the node.**

- 11 If Oracle RAC software has been installed, remove the Oracle RAC software package from the node that you are disconnecting.**

```
phys-schost# pkgrm SUNWscum
```



**Caution (SPARC only)** – If you do not remove the Oracle RAC software from the node that you disconnected, the node panics when the node is reintroduced to the cluster and potentially causes a loss of data availability.

- 12 Boot the node in cluster mode.**

- On SPARC based systems, run the following command.

```
ok boot
```

- On x86 based systems, run the following commands.

When the GRUB menu is displayed, select the appropriate Oracle Solaris entry and press Enter. The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-----+
| Solaris 10 /sol_10_x86 |
| Solaris failsafe |
| |
+-----+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

- 13 On the node, update the device namespace by updating the /devices and /dev entries.**

```
phys-schost# devfsadm -C
cldevice refresh
```

- 14 Bring the device groups back online.**

For procedures about bringing a Veritas shared disk group online, see your Veritas Volume Manager documentation.

For information about bringing a device group online, see [“How to Bring a Node Out of Maintenance State”](#) on page 261.

## ▼ How to Correct Error Messages

To correct any error messages that occurred while attempting to perform any of the cluster node removal procedures, perform the following procedure.

- 1 Attempt to rejoin the node to the global cluster. Perform this procedure only on a global cluster.**

```
phys-schost# boot
```

**2 Did the node successfully rejoin the cluster?**

- If no, proceed to [Step b](#).
- If yes, perform the following steps to remove the node from device groups.

**a. If the node successfully rejoins the cluster, remove the node from the remaining device group or groups.**

Follow procedures in [“How to Remove a Node From All Device Groups”](#) on page 132.

**b. After you remove the node from all device groups, return to [“How to Uninstall Oracle Solaris Cluster Software From a Cluster Node”](#) on page 271 and repeat the procedure.****3 If the node could not rejoin the cluster, rename the node's `/etc/cluster/ccr` file to any other name you choose, for example, `ccr.old`.**

```
mv /etc/cluster/ccr /etc/cluster/ccr.old
```

**4 Return to [“How to Uninstall Oracle Solaris Cluster Software From a Cluster Node”](#) on page 271 and repeat the procedure.**



# Administering the Cluster

---

This chapter provides administrative procedures that affect an entire global cluster or a zone cluster:

- “Overview of Administering the Cluster” on page 245
- “Performing Zone Cluster Administrative Tasks” on page 266
- “Troubleshooting” on page 281

For information about adding or removing a node from the cluster, see [Chapter 8, “Adding and Removing a Node.”](#)

## Overview of Administering the Cluster

This section describes how to perform administrative tasks for the entire global cluster or zone cluster. The following table lists these administrative tasks and the associated procedures. You generally perform cluster administrative tasks in the global zone. To administer a zone cluster, at least one machine that will host the zone cluster must be up in cluster mode. All zone-cluster nodes are not required to be up and running; Oracle Solaris Cluster replays any configuration changes when the node that is currently out of the cluster rejoins the cluster.

---

**Note** – By default, power management is disabled so that it does not interfere with the cluster. If you enable power management for a single-node cluster, the cluster is still running but it can become unavailable for a few seconds. The power management feature attempts to shut down the node, but it does not succeed.

---

In this chapter, `phys - schost#` reflects a global-cluster prompt. The `clzonecluster` interactive shell prompt is `clzc : schost>`.

TABLE 9-1 Task List: Administering the Cluster

| Task                                                                                                                                                      | Instructions                                                                                                                                   |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------|
| Add or remove a node from a cluster                                                                                                                       | Chapter 8, “Adding and Removing a Node”                                                                                                        |
| Change the name of the cluster                                                                                                                            | “How to Change the Cluster Name” on page 246                                                                                                   |
| List node IDs and their corresponding node names                                                                                                          | “How to Map Node ID to Node Name” on page 247                                                                                                  |
| Permit or deny new nodes to add themselves to the cluster                                                                                                 | “How to Work With New Cluster Node Authentication” on page 248                                                                                 |
| Change the time for a cluster by using the Network Time Protocol (NTP)                                                                                    | “How to Reset the Time of Day in a Cluster” on page 250                                                                                        |
| Shut down a node to the OpenBoot PROM ok prompt on a SPARC based system or to the Press any key to continue message in a GRUB menu on an x86 based system | “SPARC: How to Display the OpenBoot PROM (OBP) on a Node” on page 252                                                                          |
| Add or change the private hostname                                                                                                                        | “How to Add a Private Hostname for a Non-Voting Node on a Global Cluster” on page 255<br>“How to Change the Node Private Hostname” on page 253 |
| Put a cluster node in maintenance state                                                                                                                   | “How to Put a Node Into Maintenance State” on page 260                                                                                         |
| Rename a Node                                                                                                                                             | “How to Rename a Node” on page 258                                                                                                             |
| Bring a cluster node out of maintenance state                                                                                                             | “How to Bring a Node Out of Maintenance State” on page 261                                                                                     |
| Configure load limits for each node                                                                                                                       | “How to Configure Load Limits on a Node” on page 264                                                                                           |
| Move a zone cluster; prepare a zone cluster for applications, remove a zone cluster                                                                       | “Performing Zone Cluster Administrative Tasks” on page 266                                                                                     |
| Uninstall Oracle Solaris Cluster software from a node                                                                                                     | “How to Uninstall Oracle Solaris Cluster Software From a Cluster Node” on page 271                                                             |
| Add and manage an SNMP Event MIB                                                                                                                          | “How to Enable an SNMP Event MIB” on page 276<br>“How to Add an SNMP User on a Node” on page 279                                               |

## ▼ How to Change the Cluster Name

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

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser on any node in the global cluster.**
- 2 **Start the `clsetup` utility.**  

```
phys-schost# clsetup
```

The Main Menu is displayed.
- 3 **To change the cluster name, type the number that corresponds to the option for Other Cluster Properties.**  

The Other Cluster Properties menu is displayed.
- 4 **Make your selection from the menu and follow the onscreen instructions.**
- 5 **If you want the service tag for Oracle Solaris Cluster to reflect the new cluster name, delete the existing Oracle Solaris Cluster tag and restart the cluster. To delete the Oracle Solaris Cluster service tag instance, complete the following substeps on all nodes in the cluster.**
  - a. **List all of the service tags.**  

```
phys-schost# stclient -x
```
  - b. **Find the Oracle Solaris Cluster service tag instance number, then run the following command.**  

```
phys-schost# stclient -d -i service_tag_instance_number
```
  - c. **Reboot all the nodes in the cluster.**  

```
phys-schost# reboot
```

### Example 9-1 Changing the Cluster Name

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

```
phys-schost# cluster -c dromedary
```

## ▼ How to Map Node ID to Node Name

During Oracle Solaris 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. After the node ID number is 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 for a global cluster or a zone cluster. One step in this procedure is performed from a node of the global cluster. The other step is performed from a zone-cluster node.

- 1 Use the `clnode(1CL)` command to list the cluster configuration information for the global cluster.

```
phys-schost# clnode show | grep Node
```

- 2 You can also list the Node IDs for a zone cluster. The zone-cluster node has the same Node ID as the global cluster-node where it is running.

```
phys-schost# zlogin szone clnode -v | grep Node
```

### Example 9-2 Mapping the Node ID to the Node Name

The following example shows the node ID assignments for a global cluster.

```
phys-schost# clnode show | grep Node
=== Cluster Nodes ===
Node Name: phys-schost1
Node ID: 1
Node Name: phys-schost2
Node ID: 2
Node Name: phys-schost3
Node ID: 3
```

## ▼ How to Work With New Cluster Node Authentication

Oracle Solaris Cluster enables you to determine if new nodes can add themselves to the global cluster and the type of authentication to use. 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.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Become superuser on any node in the global cluster.

**2 Start the `clsetup(1CL)` utility.**

```
phys-schost# clsetup
```

The Main Menu is displayed.

**3 To work with cluster authentication, type the number that corresponds to the option for new nodes.**

The New Nodes menu is displayed.

**4 Make your selection from the menu and follow the onscreen instructions.****Example 9–3 Preventing a New Machine From Being Added to the Global Cluster**

The `clsetup` utility generates the `claccess` command. The following example shows the `claccess` command that prevents new machines from being added to the cluster.

```
phys-schost# claccess deny -h hostname
```

**Example 9–4 Permitting All New Machines to Be Added to the Global Cluster**

The `clsetup` utility generates the `claccess` command. The following example shows the `claccess` command that enables all new machines to be added to the cluster.

```
phys-schost# claccess allow-all
```

**Example 9–5 Specifying a New Machine to Be Added to the Global Cluster**

The `clsetup` utility generates the `claccess` command. The following example shows the `claccess` command that enables a single new machine to be added to the cluster.

```
phys-schost# claccess allow -h hostname
```

**Example 9–6 Setting the Authentication to Standard UNIX**

The `clsetup` utility generates the `claccess` command. The following example shows the `claccess` command that resets to standard UNIX authentication for new nodes that are joining the cluster.

```
phys-schost# claccess set -p protocol=sys
```

**Example 9–7 Setting the Authentication to DES**

The `clsetup` utility generates the `claccess` command. The following example shows the `claccess` command that uses DES authentication for new nodes that are joining the cluster.

```
phys-schost# claccess set -p protocol=des
```

When using DES authentication, you must also configure all necessary encryption keys before a node can join the cluster. For more information, see the [keyserv\(1M\)](#) and [publickey\(4\)](#) man pages.

## ▼ How to Reset the Time of Day in a Cluster

Oracle Solaris Cluster software uses the Network Time Protocol (NTP) to maintain time synchronization between cluster nodes. Adjustments in the global cluster occur automatically as needed when nodes synchronize their time. For more information, see the [Oracle Solaris Cluster Concepts Guide](#) and the [Network Time Protocol User's Guide](#).



**Caution** – When using NTP, do not attempt to adjust the cluster time while the cluster is up and running. Do not adjust the time by using the [date\(1\)](#), [rdate\(1M\)](#), [xntpd\(1M\)](#), or [svcadm\(1M\)](#) commands interactively or within [cron\(1M\)](#) scripts.

---

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser on any node in the global cluster.**
- 2 **Shut down the global cluster.**  

```
phys-schost# cluster shutdown -g0 -y -i 0
```
- 3 **Verify that the node is showing the ok prompt on a SPARC based system or the Press any key to continue message on the GRUB menu on an x86 based system.**
- 4 **Boot the node in noncluster mode.**
  - On SPARC based systems, run the following command.  

```
ok boot -x
```
  - On x86 based systems, run the following commands.

```
shutdown -g -y -i0
```

Press any key to continue

- a. In the GRUB menu, use the arrow keys to select the appropriate Oracle Solaris entry and type **e** to edit its commands.

The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
```

```
+-----+
| Solaris 10 /sol_10_x86 |
| Solaris failsafe |
+-----+
```

Use the ^ and v keys to select which entry is highlighted. Press enter to boot the selected OS, 'e' to edit the commands before booting, or 'c' for a command-line.

For more information about GRUB based booting, see [“Booting an x86 Based System by Using GRUB \(Task Map\)”](#) in *System Administration Guide: Basic Administration*.

- b. In the boot parameters screen, use the arrow keys to select the kernel entry and type **e** to edit the entry.

The GRUB boot parameters screen appears similar to the following:

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
```

```
+-----+
| root (hd0,0,a) |
| kernel /platform/i86pc/multiboot |
| module /platform/i86pc/boot_archive |
+-----+
```

Use the ^ and v keys to select which entry is highlighted. Press 'b' to boot, 'e' to edit the selected command in the boot sequence, 'c' for a command-line, 'o' to open a new line after ('O' for before) the selected line, 'd' to remove the selected line, or escape to go back to the main menu.

- c. Add **-x** to the command to specify system boot into noncluster mode.

```
[Minimal BASH-like line editing is supported. For the first word, TAB
lists possible command completions. Anywhere else TAB lists the possible
completions of a device/filename. ESC at any time exits.]
```

```
grub edit> kernel /platform/i86pc/multiboot -x
```

- d. Press the **Enter** key to accept the change and return to the boot parameters screen.

The screen displays the edited command.

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
```

```
+-----+
| root (hd0,0,a) |
| kernel /platform/i86pc/multiboot -x |
| module /platform/i86pc/boot_archive |
+-----+
```

Use the ^ and v keys to select which entry is highlighted. Press 'b' to boot, 'e' to edit the selected command in the

boot sequence, 'c' for a command-line, 'o' to open a new line after ('O' for before) the selected line, 'd' to remove the selected line, or escape to go back to the main menu.-

**e. Type b to boot the node into noncluster mode.**

---

**Note** – This change to the kernel boot parameter command does not persist over the system boot. The next time you reboot the node, it will boot into cluster mode. To boot into noncluster mode instead, perform these steps again to add the -x option to the kernel boot parameter command.

---

**5 On a single node, set the time of day by running the date command.**

```
phys-schost# date HHMM.SS
```

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

```
phys-schost# rdate hostname
```

**7 Boot each node to restart the cluster.**

```
phys-schost# reboot
```

**8 Verify that the change occurred on all cluster nodes.**

On each node, run the date command.

```
phys-schost# date
```

## ▼ SPARC: How to Display the OpenBoot PROM (OBP) on a Node

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

The phys-schost# prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

**1 Connect to the console on the node to be shut down.**

```
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 `clnode evacuate` command, then the `shutdown` command. The `clnode evacuate` command switches over all device groups from the specified node to the next-preferred node. The command also switches all resource groups from the global cluster's specified voting or non-voting node to the next-preferred voting or non-voting node.**

```
phys-schost# clnode evacuate node
shutdown -g0 -y
```




---

**Caution** – Do not use `send brk` on a cluster console to shut down a cluster node.

---

- 3 **Execute the OBP commands.**

## ▼ How to Change the Node Private Hostname

Use this procedure to change the private hostname of a cluster node after installation has been completed.

Default private host names are assigned during initial cluster installation. The default private hostname takes the form `clusternode<nodeid>-priv`, for example: `clusternode3-priv`. Change a private hostname only if the name is already in use in the domain.




---

**Caution** – Do not attempt to assign IP addresses to new private host names. The clustering software assigns them.

---

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Disable, on all nodes in the cluster, any data service resources or other applications that might cache private host names.**

```
phys-schost# clresource disable resource[,...]
```

Include the following in the applications you disable.

- HA-DNS and HA-NFS services, if configured
- Any application that has been custom-configured to use the private hostname
- Any application that is being used by clients over the private interconnect

For information about using the `clresource` command, see the `clresource(1CL)` man page and the *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

- 2 If your NTP configuration file refers to the private hostname that you are changing, bring down the Network Time Protocol (NTP) daemon on each node of the cluster.**

Use the `svcadm` command to shut down the Network Time Protocol (NTP) daemon. See the `svcadm(1M)` man page for more information about the NTP daemon.

```
phys-schost# svcadm disable ntp
```

- 3 Run the `clsetup(1CL)` utility to change the private hostname of the appropriate node.**

Run the utility from only one of the nodes in the cluster.

---

**Note** – When selecting a new private hostname, ensure that the name is unique to the cluster node.

---

- 4 Type the number that corresponds to the option for the private hostname.**

- 5 Type the number that corresponds to the option for changing a private hostname.**

Answer the questions when prompted. You are asked the name of the node whose private hostname you are changing (`clusternode< nodeid> -priv`), and the new private hostname.

- 6 Flush the name service cache.**

Perform this step on each node in the cluster. Flushing prevents the cluster applications and data services from trying to access the old private hostname.

```
phys-schost# nscd -i hosts
```

- 7 If you changed a private hostname in your NTP configuration file, update your NTP configuration file (`ntp.conf` or `ntp.conf.cluster`) on each node.**

- a. Use the editing tool of your choice.**

If you perform this step at installation, also remember to remove names for nodes that are configured. The default template is preconfigured with 16 nodes. Typically, the `ntp.conf.cluster` file is identical on each cluster node.

- b. Verify that you can successfully ping the new private hostname from all cluster nodes.**

- c. Restart the NTP daemon.**

Perform this step on each node of the cluster.

Use the `svcadm` command to restart the NTP daemon.

```
svcadm enable ntp
```

**8 Enable all data service resources and other applications that were disabled in Step 1.**

```
phys-schost# clresource enable resource[,...]
```

For information about using the `clresource` command, see the `clresource(1CL)` man page and the *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

**Example 9–8 Changing the Private Hostname**

The following example changes the private hostname from `clusternode2-priv` to `clusternode4-priv`, on node `phys-schost-2`.

```
[Disable all applications and data services as necessary.]
phys-schost-1# /etc/init.d/xntpd stop
phys-schost-1# clnode show | grep node
...
private hostname: clusternode1-priv
private hostname: clusternode2-priv
private hostname: clusternode3-priv
...
phys-schost-1# clsetup
phys-schost-1# nscd -i hosts
phys-schost-1# vi /etc/inet/ntp.conf
...
peer clusternode1-priv
peer clusternode4-priv
peer clusternode3-priv
phys-schost-1# ping clusternode4-priv
phys-schost-1# /etc/init.d/xntpd start
[Enable all applications and data services disabled at the beginning of the procedure.]
```

## ▼ How to Add a Private Hostname for a Non-Voting Node on a Global Cluster

Use this procedure to add a private hostname for a non-voting node on a global cluster after installation has been completed. In the procedures in this chapter, `phys-schost#` reflects a global-cluster prompt. Perform this procedure only on a global cluster.

**1 Run the `clsetup(1CL)` utility to add a private hostname on the appropriate zone.**

```
phys-schost# clsetup
```

**2 Type the number that corresponds to the option for private host names and press the Return key.**

- 3 **Type the number that corresponds to the option for adding a zone private hostname and press the Return key.**

Answer the questions when prompted. There is no default for a global-cluster non-voting node private hostname. You will need to provide a hostname.

## ▼ How to Change the Private Hostname on a Non-Voting Node on a Global Cluster

Use this procedure to change the private hostname of a non-voting node after installation has been completed.

Private host names are assigned during initial cluster installation. The private hostname takes the form `clusternode<nodeid>-priv`, for example: `clusternode3-priv`. Change a private hostname only if the name is already in use in the domain.




---

**Caution** – Do not attempt to assign IP addresses to new private hostnames. The clustering software assigns them.

---

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **On all nodes in the global cluster, disable any data service resources or other applications that might cache private host names.**

```
phys-schost# clresource disable resource1, resource2
```

Include the following in the applications you disable.

- HA-DNS and HA-NFS services, if configured
- Any application that has been custom-configured to use the private hostname
- Any application that is being used by clients over the private interconnect

For information about using the `clresource` command, see the `clresource(1CL)` man page and the *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

- 2 **Run the `clsetup(1CL)` utility to change the private hostname of the appropriate non-voting node on the global cluster.**

```
phys-schost# clsetup
```

You need to perform this step only from one of the nodes in the cluster.

---

**Note** – When selecting a new private hostname, ensure that the name is unique to the cluster.

---

- 3 Type the number that corresponds to the option for private hostnames and press the Return key.**
- 4 Type the number that corresponds to the option for adding a zone private hostname and press the Return key.**  
No default exists for a non-voting node of a global cluster's private hostname. You need to provide a hostname.
- 5 Type the number that corresponds to the option for changing a zone private hostname.**  
Answer the questions when prompted. You are asked for the name of the non-voting node whose private hostname is being changed (`clusternode< nodeid> -priv`), and the new private hostname.
- 6 Flush the name service cache.**  
Perform this step on each node in the cluster. Flushing prevents the cluster applications and data services from trying to access the old private hostname.  

```
phys-schost# nscd -i hosts
```
- 7 Enable all data service resources and other applications that were disabled in [Step 1](#).**

## ▼ How to Delete the Private Hostname for a Non-Voting Node on a Global Cluster

Use this procedure to delete a private hostname for a non-voting node on a global cluster. Perform this procedure only on a global cluster.

- 1 Run the `clsetup(1CL)` utility to delete a private hostname on the appropriate zone.**
- 2 Type the number that corresponds to the option for zone private hostname.**
- 3 Type the number that corresponds to the option for deleting a zone private hostname.**
- 4 Type the name of the non-voting node's private hostname that you are deleting.**

## ▼ How to Rename a Node

You can change the name of a node that is part of an Oracle Solaris Cluster configuration. You must rename the Oracle Solaris hostname before you can rename the node. Use the `clnode rename` command to rename the node.

The following instructions apply to any application that is running in a global cluster.

- 1 On the global cluster, become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.
- 2 If you are renaming a node in an Oracle Solaris Cluster Geographic Edition cluster that is in a partnership of an Oracle Solaris configuration, you must perform additional steps. If the cluster where you are performing the rename procedure is primary for the protection group, and you want to have the application in the protection group online, you can switch the protection group to the secondary cluster during the rename procedure. For more information on Geographic Edition clusters and nodes, see [Chapter 5, "Administering Cluster Partnerships," in \*Oracle Solaris Cluster Geographic Edition System Administration Guide\*](#).
- 3 Rename the Oracle Solaris host names by completing the steps in "[How to Change a System's Host Name](#)" in *System Administration Guide: Advanced Administration*, except do *not* perform a reboot at the end of the procedure. Instead, perform a cluster shutdown after you complete these steps.
- 4 Boot all cluster nodes into noncluster mode.
 

```
ok> boot -x
```
- 5 In noncluster mode on the node where you renamed the Oracle Solaris hostname, rename the node and run the `cmd` command on each renamed host. Rename one node at a time.
 

```
clnode rename -n newnodename oldnodename
```
- 6 Update any existing references to the previous hostname in the applications that run on the cluster.
- 7 Confirm that the node was renamed by checking the command messages and log files.
- 8 Reboot all nodes into cluster mode.
 

```
sync;sync;sync;/etc/reboot
```
- 9 Verify the node displays the new name.
 

```
clnode status -v
```
- 10 If you are renaming a node on a Geographic Edition cluster node and the partner cluster of the cluster that contains the renamed node still references the previous nodename, the protection

group's synchronization status will appear as an *Error*. You must update the protection group from one node of the partner cluster that contains the renamed node by using the `geopg update <pg>`. After you complete that step, run the `geopg start -e global <pg>` command. At a later time, you can switch the protection group back to the cluster with the renamed node.

- 11 You can choose to change the logical hostname resources' `hostnameList` property. See [“How to Change the Logical Hostnames Used by Existing Oracle Solaris Cluster Logical Hostname Resources” on page 259](#) for instructions on this optional step.

## ▼ How to Change the Logical Hostnames Used by Existing Oracle Solaris Cluster Logical Hostname Resources

You can choose to change the logical hostname resource's `hostnameList` property either before or after you rename the node by following the steps in [“How to Rename a Node” on page 258](#). This step is optional.

- 1 On the global cluster, become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.
- 2 Optionally, you can change the logical hostnames used by any of the existing Oracle Solaris Cluster Logical Hostname resources.

The following steps show how to configure the `apache-lh-res` resource to work with the new logical hostname, and must be executed in cluster mode.

- a. In cluster mode, take the Apache resource groups that contain the logical hostnames offline.

```
clrg offline apache-rg
```

- b. Disable the Apache logical hostname resources.

```
clr disable apache-lh-res
```

- c. Provide the new hostname list.

```
clr set -p HostnameList=test-2 apache-lh-res
```

- d. Change the application's references for previous entries in the `hostnameList` property to reference the new entries.

- e. Enable the new Apache logical hostname resources

```
clr enable apache-lh-res
```

- f. Bring the Apache resource groups online.

```
clrg online apache-rg
```

- g. Confirm that the application started correctly by running the following command checking a client.

```
clrs status apache-rs
```

## ▼ How to Put a Node Into Maintenance State

Put a global-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 node into maintenance state, the node must be shut down with `clnode(1CL)` evacuate and `cluster(1CL)` shutdown commands.

---

**Note** – Use the Oracle Solaris shutdown command to shut down a single node. Use the `cluster shutdown` command only when shutting down an entire cluster.

---

When a cluster node is shut down and put in 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 removed from maintenance mode and brought back online.

Use the `clquorum(1CL)` disable command to put a cluster node into maintenance state.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on the global-cluster node that you are putting into maintenance state.**
- 2 **Evacuate any resource groups and device groups from the node. The `clnode evacuate` command switches over all resource groups and device groups, including all non-voting nodes from the specified node to the next-preferred node.**

```
phys-schost# clnode evacuate node
```

- 3 **Shut down the node that you evacuated.**

```
phys-schost# shutdown -g0 -y-i 0
```

- 4 **Become superuser or assume a role that provides `soLaris.cLuster.modify` RBAC authorization on another node in the cluster and put the node that you shut down in [Step 3](#) in maintenance state.**

```
phys-schost# clquorum disable node
```

node Specifies the name of a node that you want to put into maintenance mode.

- 5 **Verify that the global-cluster node is now in maintenance state.**

```
phys-schost# clquorum status node
```

The node that you put into maintenance state should have a Status of `offLine` and `0` (zero) for Present and Possible quorum votes.

### Example 9–9 Putting a Global-Cluster Node Into Maintenance State

The following example puts a cluster node into maintenance state and verifies the results. The `clnode status` output shows the Node votes for `phys-schost-1` to be `0` (zero) and the status to be `Offline`. The Quorum Summary should also show reduced vote counts. Depending on your configuration, the Quorum Votes by Device output might indicate that some quorum disk devices are offline as well.

[On the node to be put into maintenance state:]

```
phys-schost-1# clnode evacuate phys-schost-1
phys-schost-1# shutdown -g0 -y -i0
```

[On another node in the cluster:]

```
phys-schost-2# clquorum disable phys-schost-1
phys-schost-2# clquorum status phys-schost-1
```

```
-- Quorum Votes by Node --
```

| Node Name     | Present | Possible | Status  |
|---------------|---------|----------|---------|
| -----         | -----   | -----    | -----   |
| phys-schost-1 | 0       | 0        | Offline |
| phys-schost-2 | 1       | 1        | Online  |
| phys-schost-3 | 1       | 1        | Online  |

**See Also** To bring a node back online, see [“How to Bring a Node Out of Maintenance State”](#) on page 261.

## ▼ How to Bring a Node Out of Maintenance State

Use the following procedure to bring a global-cluster node back online and 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 nonzero vote counts that have ports to the quorum device.

When a node has been put in 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 removed from maintenance state, both the node's quorum vote count and the quorum device vote count are incremented by one.

Run this procedure any time a global-cluster node has been put in maintenance state and you are removing it from maintenance state.



---

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

---

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the global cluster other than the one in maintenance state.**
- 2 Depending on the number of nodes that you have in your global cluster configuration, perform one of the following steps:**
  - If you have two nodes in your cluster configuration, go to [Step 4](#).
  - If you have more than two nodes in your cluster configuration, go to [Step 3](#).

- 3 If the node that you are removing from maintenance state will have quorum devices, reset the cluster quorum count from a node other than the one in maintenance state.**

You must reset the quorum count from a node other than the node in maintenance state before rebooting the node, or the node might hang while waiting for quorum.

```
phys-schost# clquorum reset
```

reset                    The change flag that resets quorum.

- 4 Boot the node that you are removing from maintenance state.**

- 5 Verify the quorum vote count.**

```
phys-schost# clquorum status
```

The node that you removed from maintenance state should have a status of `online` and show the appropriate vote count for `Present` and `Possible` quorum votes.

## Example 9–10 Removing a Cluster Node From Maintenance State and Resetting the Quorum Vote Count

The following example resets the quorum count for a cluster node and its quorum devices to their defaults and verifies the result. The `cluster status` output shows the Node votes for `phys-schost-1` to be 1 and the status to be `online`. The Quorum Summary should also show an increase in vote counts.

```
phys-schost-2# clquorum reset
```

- On SPARC based systems, run the following command.

```
ok boot
```

- On x86 based systems, run the following commands.

When the GRUB menu is displayed, select the appropriate Oracle Solaris entry and press Enter. The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-----+
| Solaris 10 /sol_10_x86 |
| Solaris failsafe |
| |
+-----+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

```
phys-schost-1# clquorum status
```

```
--- Quorum Votes Summary ---
```

| Needed | Present | Possible |
|--------|---------|----------|
| -----  | -----   | -----    |
| 4      | 6       | 6        |

```
--- Quorum Votes by Node ---
```

| Node Name     | Present | Possible | Status |
|---------------|---------|----------|--------|
| -----         | -----   | -----    | -----  |
| phys-schost-2 | 1       | 1        | Online |
| phys-schost-3 | 1       | 1        | Online |

```
--- Quorum Votes by Device ---
```

| Device Name          | Present | Possible | Status |
|----------------------|---------|----------|--------|
| -----                | -----   | -----    | -----  |
| /dev/did/rdisk/d3s2  | 1       | 1        | Online |
| /dev/did/rdisk/d17s2 | 0       | 1        | Online |
| /dev/did/rdisk/d31s2 | 1       | 1        | Online |

## Configuring Load Limits

You can enable the automatic distribution of resource group load across nodes or zones by setting load limits. You can configure a set of load limits for each cluster node. You assign load factors to resource groups, and the load factors correspond to the defined load limits of the nodes. The default behavior is to distribute resource group load evenly across all the available nodes in the resource group's node list.

The resource groups are started on a node from the resource group's node list by the RGM so that the node's load limits are not exceeded. As resource groups are assigned to nodes by the RGM, the resource groups' load factors on each node are summed up to provide a total load. The total load is then compared against that node's load limits.

A load limit consists of the following items:

- A user-assigned name.
- A soft limit value – You can temporarily exceed a soft load limit.
- A hard limit value – Hard load limits can never be exceeded and are strictly enforced.

You can set both the hard limit and the soft limit in a single command. If one of the limits is not explicitly set, the default value is used. Hard and soft load limits for each node are created and modified with the `clnode create-loadlimit`, `clnode set-loadlimit`, and `clnode delete-loadlimit` commands. See the `clnode(1CL)` man page for more information.

You can configure a resource group to have a higher priority so that it is less likely to be displaced from a specific node. You can also set a `preemption_mode` property to determine if a resource group will be preempted from a node by a higher-priority resource group because of node overload. A `concentrate_load` property also allows you to concentrate the resource group load onto as few nodes as possible. The default value of the `concentrate_load` property is `FALSE` by default.

---

**Note** – You can configure load limits on nodes in a global cluster or a zone cluster. You can use the command line, the `clsetup` utility, or the Oracle Solaris Cluster Manager interface to configure load limits. The following procedure illustrates how to configure load limits using the command line.

---

### ▼ How to Configure Load Limits on a Node

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on any node of the global cluster.**

**2 Create and set a load limit for the nodes that you want to use load balancing.**

```
clnode create-loadlimit -p limitname=mem_load -Z zc1 -p
softlimit=11 -p hardlimit=20 node1 node2 node3
```

In this example, the zone cluster name is `zc1`. The sample property is called `mem_load` and has a soft limit of 11 and a hard load limit of 20. Hard and soft limits are optional arguments and default to unlimited if you do not specifically define them. See the [clnode\(1CL\)](#) man page for more information.

**3 Assign load factor values to each resource group.**

```
clresourcegroup set -p load_factors=mem_load@50,factor2@1 rg1 rg2
```

In this example, the load factors are set on the two resource groups, `rg1` and `rg2`. The load factor settings correspond to the defined load limits of the nodes. You can also perform this step during the creation of the resource group with the `clresourcegroup create` command. See the [clresourcegroup\(1CL\)](#) man page for more information.

**4 If desired, you can redistribute the existing load (clrg remaster).**

```
clresourcegroup remaster rg1 rg2
```

This command can move resource groups off their current master to other nodes to achieve uniform load distribution.

**5 If desired, you can give some resource groups a higher priority than others.**

```
clresourcegroup set -p priority=600 rg1
```

The default priority is 500. Resource groups with higher priority values get precedence in node assignment over resource groups with lower priorities.

**6 If desired, you can set the Preemption\_mode property.**

```
clresourcegroup set -p Preemption_mode=No_cost rg1
```

See the [clresourcegroup\(1CL\)](#) man page for more information on the `HAS_COST`, `NO_COST`, and `NEVER` options.

**7 If desired, you can also set the Concentrate\_Load flag.**

```
cluster set -p Concentrate_Load=TRUE
```

**8 If desired, you can specify an affinity between resource groups.**

A strong positive or negative affinity takes precedence over load distribution. A strong affinity can never be violated, nor can a hard load limit. If you set both strong affinities and hard load limits, some resource groups might be forced to remain offline if both constraints cannot be satisfied.

The following example specifies a strong positive affinity between resource group `rg1` in zone cluster `zc1` and resource group `rg2` in zone cluster `zc2`.

```
clresourcegroup set -p RG_affinities=++zc2:rg2 zc1:rg1
```

## 9 Verify the status of all global-cluster nodes and zone-cluster nodes in the cluster.

```
clnode status -Z all -v
```

The output includes any load limit settings that are defined on the node or on its non-global zones.

# Performing Zone Cluster Administrative Tasks

You can perform other administrative tasks on a zone cluster, such as moving the zone path, preparing a zone cluster to run applications, and cloning a zone cluster. All of these commands must be performed from the voting node of the global cluster.

---

**Note** – The Oracle Solaris Cluster commands that you run only from the voting node in the global cluster are not valid for use with zone clusters. See the appropriate Oracle Solaris Cluster man page for information about the valid use of a command in zones.

---

TABLE 9-2 Other Zone-Cluster Tasks

| Task                                         | Instructions                                                                                                                                                                                                                      |
|----------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Move the zone path to a new zone path        | <code>clzonecluster move -f zonepath zoneclustername</code>                                                                                                                                                                       |
| Prepare the zone cluster to run applications | <code>clzonecluster ready -n nodename zoneclustername</code>                                                                                                                                                                      |
| Clone a zone cluster                         | <code>clzonecluster clone -Z source- zoneclustername [-m copymethod] zoneclustername</code><br>Halt the source zone cluster before you use the <code>clone</code> subcommand. The target zone cluster must already be configured. |
| Remove a zone cluster                        | <a href="#">“How to Remove a Zone Cluster” on page 267</a>                                                                                                                                                                        |
| Remove a file system from a zone cluster     | <a href="#">“How to Remove a File System From a Zone Cluster” on page 268</a>                                                                                                                                                     |
| Remove a storage device from a zone cluster  | <a href="#">“How to Remove a Storage Device From a Zone Cluster” on page 270</a>                                                                                                                                                  |
| Troubleshoot a node uninstallation           | <a href="#">“Troubleshooting a Node Uninstallation” on page 274</a>                                                                                                                                                               |

TABLE 9-2 Other Zone-Cluster Tasks (Continued)

| Task                                                                 | Instructions                                                                                               |
|----------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------|
| Create, set up, and manage the Oracle Solaris Cluster SNMP Event MIB | <a href="#">“Creating, Setting Up, and Managing the Oracle Solaris Cluster SNMP Event MIB” on page 275</a> |

## ▼ How to Remove a Zone Cluster

You can delete a specific zone cluster or use a wildcard to remove all zone clusters that are configured on the global cluster. The zone cluster must be configured before you remove it.

- 1 Become a superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on the node of the global cluster. Perform all steps in this procedure from a node of the global cluster.**
- 2 Delete all resource groups and their resources from the zone cluster.**

```
phys-schost# clresourcegroup delete -F -Z zoneclustername +
```

---

**Note** – This step is performed from a global-cluster node. To perform this step from a node of the zone cluster instead, log into the zone-cluster node and omit `-Z zonecluster` from the command.

---

- 3 Halt the zone cluster.**

```
phys-schost# clzonecluster halt zoneclustername
```

- 4 Uninstall the zone cluster.**

```
phys-schost# clzonecluster uninstall zoneclustername
```

- 5 Unconfigure the zone cluster.**

```
phys-schost# clzonecluster delete zoneclustername
```

### Example 9-11 Removing a Zone Cluster From a Global Cluster

```
phys-schost# clresourcegroup delete -F -Z sczone +
```

```
phys-schost# clzonecluster halt sczone
```

```
phys-schost# clzonecluster uninstall sczone
```

```
phys-schost# clzonecluster delete sczone
```

## ▼ How to Remove a File System From a Zone Cluster

A file system can be exported to a zone cluster using either a direct mount or a loopback mount.

Zone clusters support direct mounts for the following:

- UFS
- Vxfs
- QFS standalone shared file system
- ZFS (exported as a data set)

Zone clusters can manage loopback mounts for the following:

- UFS
- Vxfs
- QFS standalone shared file system
- QFS shared file system
- PxFs on UFS
- PxFs on Vxfs

For instructions on adding a file system to a zone cluster, see “[Adding File Systems to a Zone Cluster](#)” in *Oracle Solaris Cluster Software Installation Guide*.

The `phys - schost#` prompt reflects a global-cluster prompt. This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser on a node of the global cluster that hosts the zone cluster. Some steps in this procedure are performed from a node of the global cluster. Other steps are performed from a node of the zone cluster.**
- 2 **Delete the resources related to the file system being removed.**

- a. **Identify and remove the Oracle Solaris Cluster resource types, such as `HASstoragePlus` and `SUNW.ScalMountPoint`, that are configured for the zone cluster's file system that you are removing.**

```
phys - schost# clresource delete -F -Z zoneclustername fs_zone_resources
```

- b. **If applicable, identify and remove the Oracle Solaris Cluster resources of type `SUNW.qfs` that are configured in the global cluster for the file system that you are removing.**

```
phys - schost# clresource delete -F fs_global_resources
```

Use the `-F` option carefully because it forces the deletion of all the resources you specify, even if you did not disable them first. All the resources you specified are removed from the resource-dependency settings of other resources, which can cause a loss of service in the cluster. Dependent resources that are not deleted can be left in an invalid state or in an error state. For more information, see the `clresource(1CL)` man page.

---

**Tip** – If the resource group for the removed resource later becomes empty, you can safely delete the resource group.

---

**3 Determine the path to the file-system mount point directory. For example:**

```
phys-schost# clzonecluster configure zoneclustername
```

**4 Remove the file system from the zone-cluster configuration.**

```
phys-schost# clzonecluster configure zoneclustername
```

```
clzc:zoneclustername> remove fs dir=filesystemdirectory
```

```
clzc:zoneclustername> commit
```

The file system mount point is specified by **dir=**.

**5 Verify the removal of the file system.**

```
phys-schost# clzonecluster show -v zoneclustername
```

### Example 9–12 Removing a Highly Available File System in a Zone Cluster

This example shows how to remove a file system with a mount-point directory (`/local/ufs-1`) that is configured in a zone cluster called `sczone`. The resource is `hasp-rs` and is of the type `HASStoragePlus`.

```
phys-schost# clzonecluster show -v sczone
...
Resource Name: fs
 dir: /local/ufs-1
 special: /dev/md/ds1/dsk/d0
 raw: /dev/md/ds1/rdisk/d0
 type: ufs
 options: [logging]
...
phys-schost# clresource delete -F -Z sczone hasp-rs
phys-schost# clzonecluster configure sczone
clzc:sczone> remove fs dir=/local/ufs-1
clzc:sczone> commit
phys-schost# clzonecluster show -v sczone
```

### Example 9–13 Removing a Highly Available ZFS File System in a Zone Cluster

This example shows to remove a ZFS file systems in a ZFS pool called `HAzpool`, which is configured in the `sczone` zone cluster in resource `hasp-rs` of type `SUNW.HASStoragePlus`.

```
phys-schost# clzonecluster show -v sczone
...
Resource Name: dataset
 name: HAzpool
```

```

...
phys-schost# clresource delete -F -Z sczone hasp-rs
phys-schost# clzonecluster configure sczone
clzc:sczone> remove dataset name=HAzpool
clzc:sczone> commit
phys-schost# clzonecluster show -v sczone

```

## ▼ How to Remove a Storage Device From a Zone Cluster

You can remove storage devices, such as SVM disksets and DID devices, from a zone cluster. Perform this procedure to remove a storage device from a zone cluster.

- 1 **Become superuser on a node of the global cluster that hosts the zone cluster. Some steps in this procedure are performed from a node of the global cluster. Other steps can be performed from a node of the zone cluster.**

- 2 **Delete the resources related to the devices being removed. Identify and remove the Oracle Solaris Cluster resource types, such as SUNW.HASStoragePlus and SUNW.ScalDeviceGroup, that are configured for the zone cluster's devices that you are removing.**

```
phys-schost# clresource delete -F -Z zoneclustername dev_zone_resources
```

- 3 **Determine the match entry for the devices to be removed.**

```
phys-schost# clzonecluster show -v zoneclustername
...
Resource Name: device
match: <device_match>
...

```

- 4 **Remove the devices from the zone-cluster configuration.**

```
phys-schost# clzonecluster configure zoneclustername
clzc:zoneclustername> remove device match=<devices_match>
clzc:zoneclustername> commit
clzc:zoneclustername> end

```

- 5 **Reboot the zone cluster.**

```
phys-schost# clzonecluster reboot zoneclustername
```

- 6 **Verify the removal of the devices.**

```
phys-schost# clzonecluster show -v zoneclustername
```

### Example 9–14 Removing an SVM Disk Set From a Zone Cluster

This example shows how to remove an SVM disk set called `apachedg` configured in a zone cluster called `sczone`. The set number of the `apachedg` disk set is 3. The devices are used by the `zc_rs` resource that is configured in the cluster.

```

phys-schost# clzonecluster show -v sczone
...
Resource Name: device
match: /dev/md/apachedg/*dsk/*
Resource Name: device
match: /dev/md/shared/3/*dsk/*
...
phys-schost# clresource delete -F -Z sczone zc_rs

phys-schost# ls -l /dev/md/apachedg
lrwxrwxrwx 1 root root 8 Jul 22 23:11 /dev/md/apachedg -> shared/3
phys-schost# clzonecluster configure sczone
clzc:sczone> remove device match=/dev/md/apachedg/*dsk/*
clzc:sczone> remove device match=/dev/md/shared/3/*dsk/*
clzc:sczone> commit
clzc:sczone> end
phys-schost# clzonecluster reboot sczone
phys-schost# clzonecluster show -v sczone

```

### Example 9–15 Removing a DID Device From a Zone Cluster

This example shows how to remove DID devices d10 and d11, which are configured in a zone cluster called sczone. The devices are used by the zc\_rs resource that is configured in the cluster.

```

phys-schost# clzonecluster show -v sczone
...
Resource Name: device
match: /dev/did/*dsk/d10*
Resource Name: device
match: /dev/did/*dsk/d11*
...
phys-schost# clresource delete -F -Z sczone zc_rs
phys-schost# clzonecluster configure sczone
clzc:sczone> remove device match=/dev/did/*dsk/d10*
clzc:sczone> remove device match=/dev/did/*dsk/d11*
clzc:sczone> commit
clzc:sczone> end
phys-schost# clzonecluster reboot sczone
phys-schost# clzonecluster show -v sczone

```

## ▼ How to Uninstall Oracle Solaris Cluster Software From a Cluster Node

Perform this procedure to uninstall Oracle Solaris Cluster software from a global-cluster node before you disconnect it from a fully established cluster configuration. You can use this procedure to uninstall software from the last remaining node of a cluster.

---

**Note** – To uninstall Oracle Solaris Cluster software from a node that has not yet joined the cluster or is still in installation mode, do not perform this procedure. Instead, go to “How to Uninstall Oracle Solaris Cluster Software to Correct Installation Problems” in the *Oracle Solaris Cluster Software Installation Guide*.

---

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Ensure that you have correctly completed all prerequisite tasks in the task map to remove a cluster node.**

See [Table 8–2](#).

---

**Note** – Ensure that you have removed the node from the cluster configuration by using `clnode remove` before you continue with this procedure.

---

- 2 Become superuser on an active member of the global cluster *other than* the global-cluster node that you are uninstalling. Perform this procedure from a global-cluster node.**
- 3 From the active cluster member, add the node that you intend to uninstall to the cluster's node authentication list.**

```
phys-schost# claccess allow -h hostname
```

-h                      Specifies the name of the node to be added to the node's authentication list.

Alternately, you can use the `clsetup(1CL)` utility. See “How to Add a Node to the Authorized Node List” on page 230 for procedures.

- 4 Become superuser on the node to uninstall.**
  - 5 If you have a zone cluster, uninstall it.**
- ```
phys-schost# clzonecluster uninstall -F zoneclustername
```
- For specific steps, “How to Remove a Zone Cluster” on page 267.
- 6 If your node has a dedicated partition for the global devices namespace, reboot the global-cluster node into noncluster mode.**

- On a SPARC based system, run the following command.

```
# shutdown -g0 -y -i0ok boot -x
```

- On an x86 based system, run the following commands.

```
# shutdown -g0 -y -i0
...
<<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type    b [file-name] [boot-flags] <ENTER>  to boot with options
or      i <ENTER>                             to enter boot interpreter
or      <ENTER>                               to boot with defaults

<<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -x
```

- 7 In the `/etc/vfstab` file, remove all globally mounted file-system entries *except* the `/global/.devices` global mounts.**
- 8 If you intend to reinstall Oracle Solaris Cluster software on this node, remove the Oracle Solaris Cluster entry from the Sun Java Enterprise System (Java ES) product registry.**

If the Java ES product registry contains a record that Oracle Solaris Cluster software was installed, the Java ES installer shows the Oracle Solaris Cluster component grayed out and does not permit reinstallation.

 - a. Start the Java ES uninstaller.**

Run the following command, where *ver* is the version of the Java ES distribution from which you installed Oracle Solaris Cluster software.

```
# /var/sadm/prod/SUNWentsysver/uninstall
```
 - b. Follow the prompts to select Oracle Solaris Cluster to uninstall.**

For more information about using the `uninstall` command, see [Chapter 8, “Uninstalling,” in *Sun Java Enterprise System 5 Update 1 Installation Guide for UNIX*](#).
- 9 If you do not intend to reinstall the Oracle Solaris Cluster software on this cluster, disconnect the transport cables and the transport switch, if any, from the other cluster devices.**
 - a. If the uninstalled node is connected to a storage device that uses a parallel SCSI interface, install a SCSI terminator to the open SCSI connector of the storage device after you disconnect the transport cables.**

If the uninstalled node is connected to a storage device that uses Fibre Channel interfaces, no termination is necessary.
 - b. Follow the documentation that shipped with your host adapter and server for disconnection procedures.**

Tip – If you use a loopback file interface (lofi) device, the Java ES uninstaller automatically removes the lofi file, which is called `.globaldevices`. For more information about migrating a `global-devices` namespace to a lofi, see [“Migrating the Global-Devices Namespace” on page 124](#).

Troubleshooting a Node Uninstallation

This section describes error messages that you might receive when you run the `clnode remove` command and the corrective actions to take.

Unremoved Cluster File System Entries

The following error messages indicate that the global-cluster node you removed still has cluster file systems referenced in its `vfstab` file.

```
Verifying that no unexpected global mounts remain in /etc/vfstab ... failed
clnode: global-mount1 is still configured as a global mount.
clnode: global-mount1 is still configured as a global mount.
clnode: /global/dg1 is still configured as a global mount.
```

```
clnode: It is not safe to uninstall with these outstanding errors.
clnode: Refer to the documentation for complete uninstall instructions.
clnode: Uninstall failed.
```

To correct this error, return to [“How to Uninstall Oracle Solaris Cluster Software From a Cluster Node” on page 271](#) and repeat the procedure. Ensure that you successfully complete [Step 7](#) in the procedure before you rerun the `clnode remove` command.

Unremoved Listing in Device Groups

The following error messages indicate that the node you removed is still listed with a device group.

```
Verifying that no device services still reference this node ... failed
clnode: This node is still configured to host device service "
service".
```

```
clnode: This node is still configured to host device service "
service2".
```

```
clnode: This node is still configured to host device service "
service3".
```

```
clnode: This node is still configured to host device service "
dg1".
```

```
clnode: It is not safe to uninstall with these outstanding errors.
clnode: Refer to the documentation for complete uninstall instructions.
clnode: Uninstall failed.
```

Missing Uninstall Script

If you did not use the `installer` program to install or upgrade Sun Cluster 3.1 or 3.2 software that you now want to remove, there is no uninstall script to use for that software version. Instead, perform the following steps to uninstall the software.

Note – The Oracle Solaris Cluster 3.3 software is installed with the installer program, so these steps do not apply to that version of the software.

▼ How to Uninstall Sun Cluster 3.1 and 3.2 Software Without an Uninstall Script

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**
- 2 **Change to a directory that is not associated with any Sun Cluster package.**

```
# cd /directory
```
- 3 **Uninstall the Sun Cluster software from the node.**

```
# scinstall -r
```
- 4 **Rename the product registry file, to allow possible reinstallation of the software in the future.**

```
# mv /var/sadm/install/productregistry /var/sadm/install/productregistry.sav
```

Creating, Setting Up, and Managing the Oracle Solaris Cluster SNMP Event MIB

This section describes how to create, set up, and manage the Simple Network Management Protocol (SNMP) event Management Information Base (MIB). This section also describes how to enable, disable, and change the Oracle Solaris Cluster SNMP event MIB.

The Oracle Solaris Cluster software currently supports one MIB, the event MIB. The SNMP manager software traps cluster events in real time. When enabled, the SNMP manager automatically sends trap notifications to all hosts that are defined by the `clsnmphost` command. The MIB maintains a read-only table of the most current 50 events. Because clusters generate numerous notifications, only events with a severity of `warning` or greater are sent as trap notifications. This information does not persist across reboots.

The SNMP event MIB is defined in the `sun-cluster-event-mib.mib` file and is located in the `/usr/cluster/lib/mib` directory. You can use this definition to interpret the SNMP trap information.

The default port number for the event SNMP module is 11161, and the default port for the SNMP traps is 11162. These port numbers can be changed by modifying the Common Agent Container property file, which is `/etc/cacao/instances/default/private/cacao.properties`.

Creating, setting up, and managing an Oracle Solaris Cluster SNMP event MIB can involve the following tasks.

TABLE 9-3 Task Map: Creating, Setting Up, and Managing the Oracle Solaris Cluster SNMP Event MIB

Task	Instructions
Enable an SNMP event MIB	“How to Enable an SNMP Event MIB” on page 276
Disable an SNMP event MIB	“How to Disable an SNMP Event MIB” on page 277
Change an SNMP event MIB	“How to Change an SNMP Event MIB” on page 277
Add an SNMP host to the list of hosts that will receive trap notifications for the MIBs	“How to Enable an SNMP Host to Receive SNMP Traps on a Node” on page 278
Remove an SNMP host	“How to Disable an SNMP Host From Receiving SNMP Traps on a Node” on page 279
Add an SNMP user	“How to Add an SNMP User on a Node” on page 279
Remove an SNMP user	“How to Remove an SNMP User From a Node” on page 280

▼ How to Enable an SNMP Event MIB

This procedure shows how to enable an SNMP event MIB.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**
- 2 Enable the SNMP event MIB.**

```
phys-schost-1# clsnmpmib enable [-n node] MIB
```

`[-n node]` Specifies the *node* on which the event MIB that you want to enable is located. You can specify a node ID or a node name. If you do not specify this option, the current node is used by default.

MIB Specifies the name of the MIB that you want to enable. In this case, the MIB name must be event.

▼ How to Disable an SNMP Event MIB

This procedure shows how to disable an SNMP event MIB.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**
- 2 **Disable the SNMP event MIB.**

```
phys-schost-1# clnmpmib disable -n node MIB
```

-n node Specifies the *node* on which the event MIB that you want to disable is located. You can specify a node ID or a node name. If you do not specify this option, the current node is used by default.

MIB Specifies the type of the MIB that you want to disable. In this case, you must specify event.

▼ How to Change an SNMP Event MIB

This procedure shows how to change the protocol for an SNMP event MIB.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**
- 2 **Change the protocol of the SNMP event MIB.**

```
phys-schost-1# clnmpmib set -n node -p version=value MIB
```

-n *node*

Specifies the *node* on which the event MIB that you want to change is located. You can specify a node ID or a node name. If you do not specify this option, the current node is used by default.

-p version=*value*

Specifies the version of SNMP protocol to use with the MIBs. You specify *value* as follows:

- version=SNMPv2
- version=snmpv2
- version=2
- version=SNMPv3
- version=snmpv3
- version=3

MIB

Specifies the name of the MIB or MIBs to which to apply the subcommand. In this case, you must specify event.

▼ How to Enable an SNMP Host to Receive SNMP Traps on a Node

This procedure shows how to add an SNMP host on a node to the list of hosts that will receive trap notifications for the MIBs.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**
- 2 **Add the host to the SNMP host list of a community on another node.**

```
phys-schost-1# clnmhost add -c SNMPcommunity [-n node] host
```

-c *SNMPcommunity*

Specifies the SNMP community name that is used in conjunction with the hostname.

You must specify the SNMP community name *SNMPcommunity* when you add a host to a community other than `public`. If you use the `add` subcommand without the `-c` option, the subcommand uses `public` as the default community name.

If the specified community name does not exist, this command creates the community.

-n *node*

Specifies the name of the *node* of the SNMP host that is provided access to the SNMP MIBs in the cluster. You can specify a node name or a node ID. If you do not specify this option, the current node is used by default.

host

Specifies the name, IP address, or IPv6 address of a host that is provided access to the SNMP MIBs in the cluster.

▼ How to Disable an SNMP Host From Receiving SNMP Traps on a Node

This procedure shows how to remove an SNMP host on a node from the list of hosts that will receive trap notifications for the MIBs.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `soLaris.cLuster.mOdiFy` RBAC authorization.**
- 2 **Remove the host from the SNMP host list of a community on the specified node.**

```
phys-schost-1# cLsnmPhost remove -c SNMPcommunity -n node host
```

`remove`

Removes the specified SNMP host from the specified node.

-c *SNMPcommunity*

Specifies the name of the SNMP community from which the SNMP host is removed.

-n *node*

Specifies the name of the *node* on which the SNMP host is removed from the configuration. You can specify a node name or a node ID. If you do not specify this option, the current node is used by default.

host

Specifies the name, IP address, or IPv6 address of the host that is removed from the configuration.

To remove all hosts in the specified SNMP community, use a plus sign (+) for *host* with the -c option. To remove all hosts, use the plus sign (+) for *host*.

▼ How to Add an SNMP User on a Node

This procedure shows how to add an SNMP user to the SNMP user configuration on a node.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**
- 2 **Add the SNMP user.**

```
phys-schost-1# clsnmpuser create -n node -a authentication \
                -f password user
```

- n *node* Specifies the node on which the SNMP user is added. You can specify a node ID or a node name. If you do not specify this option, the current node is used by default.
- a *authentication* Specifies the authentication protocol that is used to authorize the user. The value of the authentication protocol can be SHA or MD5.
- f *password* Specifies a file that contains the SNMP user passwords. If you do not specify this option when you create a new user, the command prompts for a password. This option is valid only with the `add` subcommand.

You must specify user passwords on separate lines in the following format:

```
user:password
```

Passwords cannot contain the following characters or a space:

- ; (semicolon)
- : (colon)
- \ (backslash)
- \n (newline)

user Specifies the name of the SNMP user that you want to add.

▼ How to Remove an SNMP User From a Node

This procedure shows how to remove an SNMP user from the SNMP user configuration on a node.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**

- 2 **Remove the SNMP user.**

```
phys-schost-1# clsnmpuser delete -n node user
```

-n node Specifies the node from which the SNMP user is removed. You can specify a node ID or a node name. If you do not specify this option, the current node is used by default.

user Specifies the name of the SNMP user that you want to remove.

Troubleshooting

This section contains troubleshooting procedures that you can use for testing purposes.

Running an Application Outside the Global Cluster

▼ How to Take a Solaris Volume Manager Metaset From Nodes Booted in Noncluster Mode

Use this procedure to run an application outside the global cluster for testing purposes.

- 1 **Determine if the quorum device is used in the Solaris Volume Manager metaset, and determine if the quorum device uses SCSI2 or SCSI3 reservations.**

```
phys-schost# clquorum show
```

- a. **If the quorum device is in the Solaris Volume Manager metaset, add a new quorum device which is not part of the metaset to be taken later in noncluster mode.**

```
phys-schost# clquorum add did
```

- b. **Remove the old quorum device.**

```
phys-schost# clquorum remove did
```

- c. **If the quorum device uses a SCSI2 reservation, scrub the SCSI2 reservation from the old quorum and verify that no SCSI2 reservations remain.**

To learn how to run the `pgre` commands, you need to install and use the Diagnostic Toolkit (SUNWscdtk) package provided by Oracle Support.

- 2 **Evacuate the global-cluster node that you want to boot in noncluster mode.**

```
phys-schost# clresourcegroup evacuate -n targetnode
```

- 3 **Take offline any resource group or resource groups that contain HAStorage or HAStoragePlus resources and contain devices or file systems affected by the metaset that you want to later take in noncluster mode.**
`phys-schost# clresourcegroup offline resourcegroupname`
- 4 **Disable all the resources in the resource groups that you took offline.**
`phys-schost# clresource disable resourcename`
- 5 **Unmanage the resource groups.**
`phys-schost# clresourcegroup unmanage resourcegroupname`
- 6 **Take offline the corresponding device group or device groups.**
`phys-schost# cldevicegroup offline devicegroupname`
- 7 **Disable the device group or device groups.**
`phys-schost# cldevicegroup disable devicegroupname`
- 8 **Boot the passive node into noncluster mode.**
`phys-schost# reboot -x`
- 9 **Verify that the boot process has been completed on the passive node before proceeding.**
`phys-schost# svcs -x`
- 10 **Determine if any SCSI3 reservations exist on the disks in the metasets. Run the following command on all disks in the metasets.**
`phys-schost# /usr/cluster/lib/sc/scsi -c inkeys -d /dev/did/rdisk/dids2`
- 11 **If any SCSI3 reservations exist on the disks, scrub them.**
`phys-schost# /usr/cluster/lib/sc/scsi -c scrub -d /dev/did/rdisk/dids2`
- 12 **Take the metaset on the evacuated node.**
`phys-schost# metaset -s name -C take -f`
- 13 **Mount the file system or file systems that contain the defined device on the metaset.**
`phys-schost# mount device mountpoint`
- 14 **Start the application and perform the desired test. After finishing the test, stop the application.**
- 15 **Reboot the node and wait until the boot process has ended.**
`phys-schost# reboot`
- 16 **Bring online the device group or device groups.**
`phys-schost# cldevicegroup online -e devicegroupname`

17 Start the resource group or resource groups.

```
phys-schost# clresourcegroup online -eM resourcegroupname
```

Restoring a Corrupted Diskset

Use this procedure if a diskset is corrupted or in a state that the nodes in the cluster are unable to take ownership of the diskset. If your attempts to clear the state have failed, use this procedure as a last attempt to fix the diskset.

These procedures work for Solaris Volume Manager metaset and multi-owner Solaris Volume Manager metaset.

▼ How to Save the Solaris Volume Manager Software Configuration

Restoring a disk set from scratch can be time-consuming and error prone. A better alternative is to use the `metastat` command to regularly back up replicas or use Explorer (SUNWexplo) to create a backup. You can then use the saved configuration to recreate the diskset. You should save the current configuration into files (using the `prtvtoc` and the `metastat` commands), and then recreate the disk set and its components. See [“How to Recreate the Solaris Volume Manager Software Configuration”](#) on page 284.

1 Save the partition table for each disk in the disk set.

```
# /usr/sbin/prtvtoc /dev/global/rdisk/diskname > /etc/lvm/diskname.vtoc
```

2 Save the Solaris Volume Manager software configuration.

```
# /bin/cp /etc/lvm/md.tab /etc/lvm/md.tab_ORIGINAL
```

```
# /usr/sbin/metastat -p -s setname >> /etc/lvm/md.tab
```

Note – Other configuration files, such as the `/etc/vfstab` file, might reference the Solaris Volume Manager software. This procedure assumes that an identical Solaris Volume Manager software configuration is rebuilt and therefore, the mount information is the same. If Explorer (SUNWexplo) is run on a node that owns the set, it retrieves the `prtvtoc` and `metaset -p` information.

▼ How to Purge the Corrupted Diskset

Purging a set from a node or all nodes removes the configuration. To purge a diskset from a node, the node must not have ownership of the diskset.

1 Run the purge command on all nodes.

```
# /usr/sbin/metaset -s setname -P
```

Running this command removes the diskset information from the database replicas, as well as the Oracle Solaris Cluster repository. The `-P` and `-C` options allow a diskset to be purged without the need to completely rebuild the Solaris Volume Manager environment.

Note – If a multi-owner diskset is purged while nodes were booted out of cluster mode, you might need to install and use the Diagnostic Toolkit (SUNWscdtk) package provided by Oracle Support. The toolkit removes the information from the dcs configuration files. See [Step 2](#).

2 If you want to remove only the diskset information from the database replicas, use the following command.

```
# /usr/sbin/metaset -s setname -C purge
```

You should generally use the `-P` option, rather than the `-C` option. Using the `-C` option can cause a problem recreating the diskset because the Oracle Solaris Cluster software still recognizes the diskset.

- a. If you used the `-C` option with the `metaset` command, first create the diskset to see if a problem occurs.
- b. If a problem exists, use the Diagnostic Toolkit (SUNWscdtk) package to remove the information from the dcs configuration files.

If the purge options fail, verify that you have installed the latest kernel and metadvice patches and contact Oracle Solaris Cluster.

▼ How to Recreate the Solaris Volume Manager Software Configuration

Use this procedure only if you experience a complete loss of your Solaris Volume Manager software configuration. The steps assume that you have saved your current Solaris Volume Manager configuration and its components and purged the corrupted diskset.

Note – Mediators should be used only on two-node clusters.

1 Create a new diskset.

```
# /usr/sbin/metaset -s setname -a -h nodename1 nodename2
```

If this is a multi-owner diskset, use the following command to create a new diskset.

```
/usr/sbin/metaset -s setname -aM -h nodename1 nodename2
```

2 On the same host where the set was created, add mediator hosts if required (two nodes only).

```
/usr/sbin/metaset -s setname -a -m nodename1 nodename2
```

- 3 **Add the same disks back into the diskset from this same host.**

```
/usr/sbin/metaset -s setname -a /dev/did/rdisk/diskname /dev/did/rdisk/diskname
```

- 4 **If you purged the diskset and are recreating it, the Volume Table of Contents (VTOC) should remain on the disks, so you can skip this step. However, if you are recreating a set to recover, you should format the disks according to a saved configuration in the `/etc/lvm/diskname.vtoc` file. For example:**

```
# /usr/sbin/fmthard -s /etc/lvm/d4.vtoc /dev/global/rdsk/d4s2
```

```
# /usr/sbin/fmthard -s /etc/lvm/d8.vtoc /dev/global/rdsk/d8s2
```

You can run this command on any node.

- 5 **Check the syntax in the existing `/etc/lvm/md.tab` file for each metadvice.**

```
# /usr/sbin/metainit -s setname -n -a metadvice
```

- 6 **Create each metadvice from a saved configuration.**

```
# /usr/sbin/metainit -s setname -a metadvice
```

- 7 **If a filesystem exists on the metadvice, run the `fsck` command.**

```
# /usr/sbin/fsck -n /dev/md/setname/rdsk/metadvice
```

If the `fsck` command displays only a few errors, such as superblock count, then the device was probably reconstructed correctly. You can then run the `fsck` command without the `-n` option. If multiple errors appear, verify that you reconstructed the metadvice correctly. If you have, review the `fsck` errors to determine if the filesystem can be recovered. If it cannot, you should restore the data from a backup.

- 8 **Concatenate all other metaset on all cluster nodes to the `/etc/lvm/md.tab` file and then concatenate the local diskset.**

```
# /usr/sbin/metastat -p >> /etc/lvm/md.tab
```


Configuring Control of CPU Usage

If you want to control the usage of CPU, configure the CPU control facility. For more information about configuring the CPU control facility, see the [rg_properties\(5\)](#) man page. This chapter provides information about the following topics:

- “Introduction to CPU Control” on page 287
- “Configuring CPU Control” on page 289

Introduction to CPU Control

Oracle Solaris Cluster software enables you to control the usage of CPU.

The CPU control facility builds on the functionality available in the Oracle Solaris OS. For information about zones, projects, resource pools, processor sets, and scheduling classes, see *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

On the Oracle Solaris OS, you can do the following:

- Assign CPU shares to resource groups
- Assign processors to resource groups

Choosing a Scenario

Depending on the configuration choices you make and version of the operating system you choose, you can have different levels of CPU control. All aspects of CPU control described in this chapter are dependent on the resource group property `RG_SLM_TYPE` being set to automated.

Table 10–1 provides a description of the different configuration scenarios available.

TABLE 10-1 CPU Control Scenarios

Description	Instructions
<p>Resource group runs in the global-cluster voting node.</p> <p>Assign CPU shares to resource groups and zones, providing values for <code>project.cpu-shares</code> and <code>zone.cpu-shares</code>.</p> <p>You can perform this procedure whether or not global-cluster non-voting nodes are configured.</p>	<p>“How to Control CPU Usage in the Voting Node on a Global Cluster” on page 289</p>
<p>Resource group runs in a global-cluster non-voting zone by using the default processor set.</p> <p>Assign CPU shares to resource groups and zones, providing values for <code>project.cpu-shares</code> and <code>zone.cpu-shares</code>.</p> <p>Perform this procedure if you do not need to control the size of the processor set.</p>	<p>“How to Control CPU Usage in a Global-Cluster Non-Voting Node With the Default Processor Set” on page 291</p>
<p>Resource group runs in a global-cluster non-voting node with a dedicated processor set.</p> <p>Assign CPU shares to resource groups, providing values for <code>project.cpu-shares</code>, <code>zone.cpu-shares</code>, and maximum number of processors in a dedicated processor set.</p> <p>Set the minimum number of processor sets in a dedicated processor set.</p> <p>Perform this procedure if you want to control CPU shares and the size of a processor set. You can exercise this control only in a global-cluster non-voting node by using a dedicated processor set.</p>	<p>“How to Control CPU Usage in a Global-Cluster Non-Voting Node With a Dedicated Processor Set” on page 293</p>

Fair Share Scheduler

The first step in the procedures to assign CPU shares to resource groups is to set the scheduler for the system to be the fair share scheduler (FSS). By default, the scheduling class for the Oracle Solaris OS is timesharing schedule (TS). Set the scheduler to be FSS to have the shares configuration take effect.

You can create a dedicated processor set regardless of the scheduler class you choose.

Configuring CPU Control

This section includes the following procedures:

- “How to Control CPU Usage in the Voting Node on a Global Cluster” on page 289
- “How to Control CPU Usage in a Global-Cluster Non-Voting Node With the Default Processor Set” on page 291
- “How to Control CPU Usage in a Global-Cluster Non-Voting Node With a Dedicated Processor Set” on page 293

▼ How to Control CPU Usage in the Voting Node on a Global Cluster

Perform this procedure to assign CPU shares to a resource group that will be executed in a global-cluster voting node.

If a resource group is assigned CPU shares, Oracle Solaris Cluster software performs the following tasks when it starts a resource of the resource group in a global-cluster voting node:

- Augments the number of CPU shares assigned to the voting node (`zone.cpu-shares`) with the specified number of CPU shares, if this has not already been done.
- Creates a project named `SCSLM_resourcegroup_name` in the voting node, if this has not already been done. This project is specific to the resource group, and is assigned the specified number of CPU shares (`project.cpu-shares`).
- Starts the resource in the `SCSLM_resourcegroup_name` project.

For more information about configuring the CPU control facility, see the [rg_properties\(5\)](#) man page.

1 Set the default scheduler for the system to be fair share scheduler (FSS).

```
# dispadmin -d FSS
```

FSS becomes the default scheduler on next reboot. To make this configuration take effect immediately, use the `pricontrl` command.

```
# pricontrl -s -C FSS
```

Using the combination of the `pricontrl` and `dispadmin` commands ensures that FSS becomes the default scheduler immediately and remains so after reboot. For more information about setting a scheduling class, see the [dispadmin\(1M\)](#) and [pricontrl\(1\)](#) man pages.

Note – If the FSS is not the default scheduler, your CPU shares assignment will not take effect.

2 On each node to use CPU control, configure the number of shares for the global-cluster voting nodes and the minimum number of CPUs available in the default processor set.

Setting these parameters helps protect processes running in the voting nodes from competing for CPUs with processes running in non-voting nodes. If you do not assign a value to the `globalzoneshares` and `defaultpsetmin` properties, these properties take their default values.

```
# clnode set [-p globalzoneshares=integer] \  
[-p defaultpsetmin=integer] \  
node
```

`-p defaultpsetmin=defaultpsetmininteger` Sets the minimum number of CPU shares available in the default processor set. The default value is 1.

`-p globalzoneshares=integer` Sets the number of shares assigned to the voting node. The default value is 1.

`node` Specifies nodes on which properties are to be set.

In setting these properties, you are setting properties for the voting node. If you do not set these properties, you cannot benefit from the `RG_SLM_PSET_TYPE` property in non-voting nodes.

3 Verify that you correctly set these properties.

```
# clnode show node
```

For the node you specify, the `clnode` command prints the properties set and the values that are set for these properties. If you do not set the CPU control properties with `clnode`, they take the default value.

4 Configure the CPU control facility.

```
# clresourcegroup create -p RG_SLM_TYPE=automated \  
[-p RG_SLM_CPU_SHARES=value] resource_group_name
```

`-p RG_SLM_TYPE=automated` Enables you to control CPU usage and automates some steps to configure the Oracle Solaris OS for system resource management.

`-p RG_SLM_CPU_SHARES=value` Specifies the number of CPU shares that are assigned to the resource group-specific project, `project.cpu-shares` and determines the number of CPU shares that are assigned to the voting node zone `.cpu-shares`.

`resource_group_name` Specifies the name of the resource group.

In this procedure, you do not set the `RG_SLM_PSET_TYPE` property. In the voting node, this property takes the value `default`.

This step creates a resource group. You could alternatively use the `clresourcegroup set` command to modify an existing resource group.

5 Activate the configuration change.

```
# clresourcegroup online -M resource_group_name
```

resource_group_name Specifies the name of the resource group.

Note – Do not remove or modify the `SCSLM_resource_group_name` project. You can add more resource control manually to the project, for example, by configuring the project `.max-lwps` property. For more information, see the [projmod\(1M\)](#) man page.

▼ How to Control CPU Usage in a Global-Cluster Non-Voting Node With the Default Processor Set

Perform this procedure if you want to assign CPU shares for resource groups in a global-cluster non-voting node, but do not need to create a dedicated processor set.

If a resource group is assigned CPU shares, Oracle Solaris Cluster software performs the following tasks when starting a resource of that resource group in a non-voting node:

- Creates a pool named `SCSLM_resource_group_name`, if this has not already been done.
- Associates the `SCSLM_pool_zone_name` pool to the default processor set.
- Dynamically binds the non-voting node to the `SCSLM_poolzone_name` pool.
- Augments the number of CPU shares assigned to the non-voting node (`zone.cpu-shares`) with the specified number of CPU shares, if this has not already been done.
- Creates a project named `SCSLM_resourcegroup_name` in the non-voting node, if this has not already been done. This project is specific to the resource group, and is assigned the specified number of CPU shares (`project.cpu-shares`).
- Starts the resource in the `SCSLM_resourcegroup_name` project.

For more information about configuring the CPU control facility, see the [rg_properties\(5\)](#) man page.

1 Set the default scheduler for the system to be fair share scheduler (FSS).

```
# dispadmin -d FSS
```

FSS becomes the default scheduler on next reboot. To make this configuration take effect immediately, use the `priocntl` command:

```
# priocntl -s -C FSS
```

Using the combination of the `priocntl` and `dispadmin` commands ensures that FSS becomes the default schedule immediately and remains so after reboot. For more information about setting a scheduling class, see the [dispadmin\(1M\)](#) and [priocntl\(1\)](#) man pages.

Note – If the FSS is not the default scheduler, your CPU shares assignment will not take effect.

2 On each node to use CPU control, configure the number of shares for the global-cluster voting node and the minimum number of CPUs available in the default processor set.

Setting these parameters helps protect processes running in the voting node from competing for CPUs with processes running in global-cluster non-voting nodes. If you do not assign a value to the `globalzoneshares` and `defaultpsetmin` properties, these properties take their default values.

```
# clnode set [-p globalzoneshares=integer] \  
[-p defaultpsetmin=integer] \  
node
```

<code>-p globalzoneshares=<i>integer</i></code>	Sets the number of shares assigned to the voting node. The default value is 1.
<code>-p defaultpsetmin=<i>defaultpsetmininteger</i></code>	Sets the minimum number of CPUs available in the default processor set. The default value is 1.
<code><i>node</i></code>	Identifies nodes on which properties are to be set.

In setting these properties, you are setting properties for the voting node.

3 Verify that you correctly set these properties:

```
# clnode show node
```

For the node you specify, the `clnode` command prints the properties set and the values that are set for these properties. If you do not set the CPU control properties with `clnode`, they take the default value.

4 Configure the CPU control facility.

```
# clresourcegroup create -p RG_SLM_TYPE=automated \  
[-p RG_SLM_CPU_SHARES=value] resource_group_name
```

<code>-p RG_SLM_TYPE=automated</code>	Enables you to control CPU usage and automates some steps to configure the Oracle Solaris OS for system resource management.
<code>-p RG_SLM_CPU_SHARES=<i>value</i></code>	Specifies the number of CPU shares assigned to the resource group-specific project (<code>project.cpu-shares</code>) and determines the number of CPU shares assigned to the global-cluster non-voting node (<code>zone.cpu-shares</code>).
<code><i>resource_group_name</i></code>	Specifies the name of the resource group.

This step creates a resource group. You could alternatively use the `clresourcegroup set` command to modify an existing resource group.

You cannot set `RG_SLM_TYPE` to `automated` in a non-voting node if a pool other than the default pool is in the zone configuration or if the zone is dynamically bound to a pool other than the default pool. See the [zonecfg\(1M\)](#) and [poolbind\(1M\)](#) man pages for information about zone configuration and pool binding respectively. View your zone configuration as follows:

```
# zonecfg -z zone_name info pool
```

Note – A resource such as an `HAStoragePlus` or a `LogicalHostname` resource was configured to start in a non-voting node but with the `GLOBAL_ZONE` property set to `TRUE` is started in the voting node. Even if you set the `RG_SLM_TYPE` property to `automated`, this resource does not benefit from the CPU shares configuration and is treated as in a resource group with `RG_SLM_TYPE` set to `manual`.

In this procedure, you do not set the `RG_SLM_PSET_TYPE` property. Oracle Solaris Cluster uses the default processor set.

5 Activate the configuration change.

```
# clresourcegroup online -M resource_group_name
```

resource_group_name Specifies the name of the resource group.

If you set `RG_SLM_PSET_TYPE` to `default`, Oracle Solaris Cluster creates a pool, `SCSLM_pool_zone_name`, but does not create a processor set. In this case, `SCSLM_pool_zone_name` is associated with the default processor set.

If online resource groups are no longer configured for CPU control in a non-voting node, the CPU share value for the non-voting node takes the value of `zone.cpu-shares` in the zone configuration. This parameter has a value of 1 by default. For more information about zone configuration, see the [zonecfg\(1M\)](#) man page.

Note – Do not remove or modify the `SCSLM_resource_group_name` project. You can add more resource control manually to the project, for example by configuring the `project.max-lwps` property. For more information, see the [projmod\(1M\)](#) man page.

▼ How to Control CPU Usage in a Global-Cluster Non-Voting Node With a Dedicated Processor Set

Perform this procedure if you want your resource group to execute in a dedicated processor set.

If a resource group is configured to execute in a dedicated processor set, Oracle Solaris Cluster software performs the following tasks when it starts a resource of the resource group in a global-cluster non-voting node:

- Creates a pool named `SCSLM_pool_zone_name`, if this has not already been done.
- Creates a dedicated processor set. The size of the processor set is determined by using the `RG_SLM_CPU_SHARES` and `RG_SLM_PSET_MIN` properties.
- Associates the `SCSLM_pool_zone_name` pool to the created processor set.
- Dynamically binds the non-voting node to the `SCSLM_pool_zone_name` pool.
- Augments the number of CPU shares assigned to the non-voting node with the specified number of CPU shares, if this has not already been done.
- Creates a project named `SCSLM_resourcegroup_name` in the non-voting node, if this has not already been done. This project is specific to the resource group, and is assigned the specified number of CPU shares (`project.cpu-shares`).
- Starts the resource in the `SCSLM_resourcegroup_name` project.

1 Set the scheduler for the system to be fair share scheduler (FSS).

```
# dispadmin -d FSS
```

FSS becomes the default scheduler on next reboot. To make this configuration take effect immediately, use the `priocntl` command.

```
# priocntl -s -C FSS
```

Using the combination of the `priocntl` and `dispadmin` commands ensures that FSS becomes the default schedule immediately and remains so after reboot. For more information about setting a scheduling class, see the [dispadmin\(1M\)](#) and [priocntl\(1\)](#) man pages.

Note – If the FSS is not the default scheduler, your CPU shares assignment will not take effect.

2 On each node to use CPU control, configure the number of shares for the global-cluster voting node and the minimum number of CPUs available in the default processor set.

Setting these parameters helps protect processes running in the voting node from competing for CPUs with processes running in non-voting nodes. If you do not assign a value to the `globalzonesthreshares` and `defaultpsetmin` properties, these properties take their default values.

```
# clnode set [-p globalzonesthreshares=integer] \
[-p defaultpsetmin=integer] \
node
```

`-p defaultpsetmin=defaultpsetmininteger` Sets the minimum number of CPUs available in the default processor set. The default is 1.

`-p globalzonesthreshares=integer` Sets the number of shares assigned to the voting node. The default is 1.

`node` Identifies nodes on which properties are to be set.

In setting these properties, you are setting properties for the voting node.

3 Verify that you correctly set these properties:

```
# clnode show node
```

For the node you specify, the `clnode` command prints the properties set and the values that are set for these properties. If you do not set the CPU control properties with `clnode`, they take the default value.

4 Configure the CPU control facility.

```
# clresourcegroup create -p RG_SLM_TYPE=automated \
  [-p RG_SLM_CPU_SHARES=value] \
  -p -y RG_SLM_PSET_TYPE=value \
  [-p RG_SLM_PSET_MIN=value] resource_group_name
```

-p RG_SLM_TYPE=automated	Enables you to control CPU control usage and automates some steps to configure the Oracle Solaris OS for system resource management.
-p RG_SLM_CPU_SHARES=value	Specifies the number of CPU shares assigned to the resource group-specific project (<code>project.cpu-shares</code>) and determines the number of CPU shares assigned to the non-voting node (<code>zone.cpu-shares</code>) and the maximum number of processors in a processor set.
-p RG_SLM_PSET_TYPE=value	Enables the creation of a dedicated processor set. To have a dedicated processor set, you can set this property to <code>strong</code> or <code>weak</code> . The values <code>strong</code> and <code>weak</code> are mutually exclusive. That is, you cannot configure resource groups in the same zone so that some are <code>strong</code> and others <code>weak</code> .
-p RG_SLM_PSET_MIN=value	Determines the minimum number of processors in the processor set.
<i>resource_group_name</i>	Specifies the name of the resource group.

This step creates a resource group. You can alternatively use the `clresourcegroup set` command to modify an existing resource group.

You cannot set `RG_SLM_TYPE` to `automated` in a non-voting node if a pool other than the default pool is in the zone configuration or if the zone is dynamically bound to a pool other than the default pool. See the [zoncfg\(1M\)](#) and [poolbind\(1M\)](#) man pages for information about zone configuration and pool binding respectively. View your zone configuration as follows:

```
# zoncfg -z zone_name info pool
```

Note – A resource such as an `HASStoragePlus` or a `LogicalHostname` resource configured to start in a non-voting node but with the `GLOBAL_ZONE` property set to `TRUE` is started in the voting node. Even if you set the `RG_SLM_TYPE` property to `automated`, this resource does not benefit from the CPU shares and dedicated processor set configuration and is treated as in a resource group with `RG_SLM_TYPE` set to `manual`.

5 Activate the configuration change.

resource_group_name Specifies the name of the resource group.

Note – Do not remove or modify the `SCSLM_resource_group_name` project. You can add more resource control manually to the project, for example by configuring the `project.max-lwps` property. For more information, see the [projmod\(1M\)](#) man page.

Changes made to `RG_SLM_CPU_SHARES` and `RG_SLM_PSET_MIN` while the resource group is online are taken into account dynamically. However, if `RG_SLM_PSET_TYPE` is set to `strong`, and if not enough CPUs are available to accommodate the change, the change requested for `RG_SLM_PSET_MIN` is not applied. In this case, a warning message is displayed. On next switchover, errors because insufficient CPUs can occur if not enough CPUs are available to acknowledge the values that you configured for `RG_SLM_PSET_MIN`.

If an online resource group is no longer configured for CPU control in the non-voting node, the CPU share value for the non-voting node takes the value of `zone.cpu-shares`. This parameter has a value of 1 by default.

Patching Oracle Solaris Cluster Software and Firmware

This chapter provides the procedures for adding and removing patches for an Oracle Solaris Cluster configuration that are contained in the following sections.

- [“Overview of Patching Oracle Solaris Cluster” on page 297](#)
- [“Patching Oracle Solaris Cluster Software” on page 299](#)

Overview of Patching Oracle Solaris Cluster

Due to the nature of a cluster, all cluster member nodes must be at the same patch level for proper cluster operation. Occasionally, when patching a node with an Oracle Solaris Cluster 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 an Oracle Solaris Cluster patch, check the patch's README file. Also, check the upgrade requirements for your storage devices to determine which patch method they require.

Note – For Oracle Solaris Cluster patches, always defer to the patch's README file and to SunSolve for instructions that supersede procedures in this chapter.

Patch installation on all cluster nodes can be described by one of the following scenarios:

Rebooting patch (node)

A node must be booted to single-user mode, using the command `boot -sx` or `shutdown -g -y -i0`, before the patch or firmware can be applied, then rebooted to join the cluster. First you need to put the node into a “quiet” state by switching any resource groups or device groups from the node to be patched to another cluster member. Also, apply the patch or firmware to one cluster node at a time to avoid shutting down the entire cluster.

	The cluster itself remains available during this type of patch application, even though individual nodes are temporarily unavailable. 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)	The cluster must be stopped and each node must be booted to single-user mode, using the command <code>boot -sx</code> or <code>shutdown -g -y -i0</code> , to apply the software or firmware patch. Then, reboot the nodes to rejoin the cluster. For this type of patch, the cluster is unavailable during patch application.
Nonrebooting 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 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 because of a patch.

Use the `patchadd` command to apply a patch to the cluster, and `patchrm` to remove a patch (when possible).

Oracle Solaris Cluster Patch Tips

Use the following tips to help you administer Oracle Solaris Cluster patches more efficiently:

- Always read the patch README file before applying the patch.
- Check the upgrade requirements of your storage devices to determine which patch method they require.
- 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 might be needed.
- All nodes acting as cluster members must have the same patches.
- Keep cluster subsystem patches up to date. These patches include, for example, volume management, storage device firmware, and cluster transport.
- Review patch reports regularly, such as once a quarter, and patch an Oracle Solaris Cluster configuration by 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.

Patching Oracle Solaris Cluster Software

TABLE 11-1 Task Map: Patching the Cluster

Task	Instructions
Apply a nonrebooting Oracle Solaris Cluster patch to one node at a time without stopping the node	“How to Apply a Nonrebooting Oracle Solaris Cluster Patch” on page 307
Apply a rebooting Oracle Solaris Cluster patch after taking the cluster member to noncluster mode	“How to Apply a Rebooting Patch (Node)” on page 299 “How to Apply a Rebooting Patch (Cluster)” on page 303
Apply a patch in single-user mode when your cluster contains failover nodes	“How to Apply Patches in Single-User Mode with Failover Nodes” on page 308
Remove an Oracle Solaris Cluster patch	“Changing an Oracle Solaris Cluster Patch” on page 311

▼ 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 stop the node in the cluster and boot it to single-user mode by using the `boot -sx` or `shutdown -g -y -i0` command, before applying the patch.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Before applying the patch, check the Oracle Solaris Cluster product web site for any special preinstallation or postinstallation instructions.**
- 2 **Become superuser or assume a role that provides `solaris.cluster.admin` RBAC authorization on the node to which you are applying the patch.**
- 3 **List the resource groups and device groups on the node being patched.**

```
# clresourcegroup status -n node
# cldevicegroup status -n node
```

4 Switch all resource groups, resources, and device groups from the node being patched to other cluster members.

```
# clnode evacuate -n node
```

evacuate Evacuates all device groups and resource groups, including all global-cluster non-voting nodes.

-n node Specifies the node from which you are switching the resource groups and device groups.

5 Shut down the node.

```
# shutdown -g0 [-y]
[-i0]
```

6 Boot the node in noncluster, single-user mode.

- On SPARC based systems, run the following command.

```
ok boot -sx
```

- On x86 based systems, run the following commands.

```
phys-schost# shutdown -g -y -i0
```

Press any key to continue

- a. In the GRUB menu, use the arrow keys to select the appropriate Oracle Solaris entry and type **e** to edit its commands.

The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
```

```
+-----+
| Solaris 10 /sol_10_x86
| Solaris failsafe
|
+-----+
```

Use the ^ and v keys to select which entry is highlighted.

Press enter to boot the selected OS, 'e' to edit the commands before booting, or 'c' for a command-line.

For more information about GRUB based booting, see [“Booting an x86 Based System by Using GRUB \(Task Map\)”](#) in *System Administration Guide: Basic Administration*.

- b. In the boot parameters screen, use the arrow keys to select the kernel entry and type **e** to edit the entry.

The GRUB boot parameters screen appears similar to the following:

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
```

```
+-----+
| root (hd0,0,a)
| kernel /platform/i86pc/multiboot
| module /platform/i86pc/boot_archive
|
+-----+
```

Use the ^ and v keys to select which entry is highlighted. Press 'b' to boot, 'e' to edit the selected command in the boot sequence, 'c' for a command-line, 'o' to open a new line after ('O' for before) the selected line, 'd' to remove the selected line, or escape to go back to the main menu.

c. Add -sx to the command to specify that the system boot into noncluster mode.

[Minimal BASH-like line editing is supported. For the first word, TAB lists possible command completions. Anywhere else TAB lists the possible completions of a device/filename. ESC at any time exits.]

```
grub edit> kernel /platform/i86pc/multiboot -sx
```

d. Press the Enter key to accept the change and return to the boot parameters screen.

The screen displays the edited command.

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+-----+
| root (hd0,0,a)                                     |
| kernel /platform/i86pc/multiboot -sx              |
| module /platform/i86pc/boot_archive               |
+-----+
```

Use the ^ and v keys to select which entry is highlighted. Press 'b' to boot, 'e' to edit the selected command in the boot sequence, 'c' for a command-line, 'o' to open a new line after ('O' for before) the selected line, 'd' to remove the selected line, or escape to go back to the main menu.-

e. Type b to boot the node into noncluster mode.

Note – This change to the kernel boot parameter command does not persist over the system boot. The next time you reboot the node, it will boot into cluster mode. To boot into noncluster mode instead, perform these steps to again to add the -sx option to the kernel boot parameter command.

7 Apply the software or firmware patch.

```
# patchadd -M 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, which supersede procedures in this chapter.

8 Verify that the patch has been installed successfully.

```
# showrev -p | grep patch-id
```

9 Reboot the node into the cluster.

```
# reboot
```

10 Verify that the patch works, and that the node and cluster are operating normally.**11 Repeat Step 2 through Step 10 for all remaining cluster nodes.****12 Switch resource groups and device groups as needed.**

After you reboot all the nodes, the last node rebooted will not have the resource groups and device groups online.

```
# cldevicegroup switch -n node + | devicegroup ...
# clresourcegroup switch -n node[:zone][,...] + | resource-group ...
```

node The name of the node to which you are switching the resource groups and device groups.

zone The name of the global cluster non-voting node (node) that can master the resource group. Specify zone only if you specified a non-voting node when you created the resource group.

13 Check to see if you need to commit the patch software by using the `scversions` command.

```
# /usr/cluster/bin/scversions
```

You will see one of the following results:

```
Upgrade commit is needed.
```

```
Upgrade commit is NOT needed. All versions match.
```

14 If a commit is needed, commit the patch software.

```
# scversions -c
```

Note – Running `scversions` will cause one or more CMM reconfigurations, depending on the situation.

Example 11-1 Applying a Rebooting Patch (Node)

The following example shows the application of a rebooting Oracle Solaris Cluster patch to a node.

```
# clresourcegroup status -n rg1
...Resource Group      Resource
-----
rg1                      rs-2
rg1                      rs-3
...
```

```
# cldevicegroup status -n nodedg-schost-1
...
Device Group Name:                dg-schost-1
...
# clnode evacuate phys-schost-2
# shutdown -g0 -y -i0
...
```

Boot the node in noncluster, single-user mode.

- SPARC: Type:
 - ok boot -sx
- x86: Boot the node in noncluster, single-user mode. See boot steps in the proceeding procedure.

```
# patchadd -M /var/tmp/patches 234567-05
...
# showrev -p | grep 234567-05

...
# reboot
...
# cldevicegroup switch -n phys-schost-1 dg-schost-1
# clresourcegroup switch -n phys-schost-1 schost-sa-1
# scversions
Upgrade commit is needed.
# scversions -c
```

See Also If you need to back out a patch, see [“Changing an Oracle Solaris Cluster Patch”](#) on page 311.

▼ How to Apply a Rebooting Patch (Cluster)

With this procedure, you must first stop the cluster and boot each node to single-user mode by using the `boot -sx` or `shutdown -g -y -i0` command, before applying the patch.

- 1 Before applying the patch, check the Oracle Solaris Cluster product web site for any special preinstallation or postinstallation instructions.
- 2 Become superuser on any node in the cluster.
- 3 Shut down the cluster.

```
# cluster 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.

4 Boot each node into noncluster, single-user mode.

On the console of each node, run the following commands.

- On SPARC based systems, run the following command.

```
ok boot -sx
```

- On x86 based systems, run the following commands.

```
phys-schost# shutdown -g -y -i0
```

Press any key to continue

- a. In the GRUB menu, use the arrow keys to select the appropriate Oracle Solaris entry and type **e** to edit its commands.

The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-----+
| Solaris 10 /sol_10_x86          |
| Solaris failsafe                |
|                                 |
+-----+

```

Use the ^ and v keys to select which entry is highlighted. Press enter to boot the selected OS, 'e' to edit the commands before booting, or 'c' for a command-line.

For more information about GRUB based booting, see [“Booting an x86 Based System by Using GRUB \(Task Map\)”](#) in *System Administration Guide: Basic Administration*.

- b. In the boot parameters screen, use the arrow keys to select the kernel entry and type **e** to edit the entry.

The GRUB boot parameters screen appears similar to the following:

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+-----+
| root (hd0,0,a)                  |
| kernel /platform/i86pc/multiboot |
| module /platform/i86pc/boot_archive |
+-----+

```

Use the ^ and v keys to select which entry is highlighted. Press 'b' to boot, 'e' to edit the selected command in the boot sequence, 'c' for a command-line, 'o' to open a new line after ('O' for before) the selected line, 'd' to remove the selected line, or escape to go back to the main menu.

- c. Add -sx to the command to specify that the system boot into noncluster mode.

```
[ Minimal BASH-like line editing is supported. For the first word, TAB
lists possible command completions. Anywhere else TAB lists the possible
completions of a device/filename. ESC at any time exits. ]
```

```
grub edit> kernel /platform/i86pc/multiboot -sx
```

d. Press the Enter key to accept the change and return to the boot parameters screen.

The screen displays the edited command.

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+-----+
| root (hd0,0,a)                                     |
| kernel /platform/i86pc/multiboot -sx              |
| module /platform/i86pc/boot_archive               |
+-----+
```

Use the ^ and v keys to select which entry is highlighted. Press 'b' to boot, 'e' to edit the selected command in the boot sequence, 'c' for a command-line, 'o' to open a new line after ('O' for before) the selected line, 'd' to remove the selected line, or escape to go back to the main menu.-

e. Type b to boot the node into noncluster mode.

Note – This change to the kernel boot parameter command does not persist over the system boot. The next time you reboot the node, it will boot into cluster mode. To boot into noncluster mode instead, perform these steps to again to add the -sx option to the kernel boot parameter command.

5 Apply the software or firmware patch.

On one node at a time, run the following command.

```
# patchadd -M 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 Check to see if you need to commit the patch software by using the scversions command.

```
# /usr/cluster/bin/scversions
```

You will see one of the following results:

```
Upgrade commit is needed.
```

```
Upgrade commit is NOT needed. All versions match.
```

9 If a commit is needed, commit the patch software.

```
# scversions -c
```

Note – Running `scversions` will cause one or more CMM reconfigurations, depending on the situation.

10 Verify that the patch works, and that the nodes and cluster are operating normally.

Example 11–2 Applying a Rebooting Patch (Cluster)

The following example shows the application of a rebooting Oracle Solaris Cluster patch to a cluster.

```
# cluster shutdown -g0 -y
...
```

Boot the cluster in noncluster, single-user mode.

- SPARC: Type:

```
ok boot -sx
```

- x86: Boot each node into noncluster, single-user mode. See the preceding procedure for steps.

```
...
# patchadd -M /var/tmp/patches 234567-05
(Apply patch to other cluster nodes)
...
# showrev -p | grep 234567-05
# reboot
# scversions
Upgrade commit is needed.
# scversions -c
```

See Also If you need to back out a patch, see [“Changing an Oracle Solaris Cluster Patch”](#) on page 311.

▼ How to Apply a Nonrebooting Oracle Solaris Cluster Patch

Apply the patch to one node in the cluster at a time. When applying a nonrebooting patch, you do not need to first the node that is receiving the patch.

- 1 Before applying the patch, check the Oracle Solaris Cluster product web page for any special preinstallation or postinstallation instructions.

- 2 Apply the patch on a single node.

```
# patchadd -M 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](#) through [Step 4](#) for the remaining cluster nodes.

- 6 Check to see if you need to commit the patch software by using the `scversions` command.

```
# /usr/cluster/bin/scversions
```

You will see one of the following results:

```
Upgrade commit is needed.
```

```
Upgrade commit is NOT needed. All versions match.
```

- 7 If a commit is needed, commit the patch software.

```
# scversions -c
```

Note – Running `scversions` will cause one or more CMM reconfigurations, depending on the situation.

Example 11–3 Applying a Nonrebooting Oracle Solaris Cluster Patch

```
# patchadd -M /tmp/patches 234567-05
...
# showrev -p | grep 234567-05
# scversions
```

```
Upgrade commit is needed.
# scversions -c
```

See Also If you need to back out a patch, see [“Changing an Oracle Solaris Cluster Patch”](#) on page 311.

▼ How to Apply Patches in Single-User Mode with Failover Nodes

Perform this task to apply patches in single-user mode with failover nodes. This patch method is required if you use the Oracle Solaris Cluster Data Service for Solaris Containers in a failover configuration with Oracle Solaris Cluster software.

1 Verify that the quorum device is not configured for one of the LUNs used as shared storage that is part of the disksets that contain the zone path that is manually taken in this procedure.

- a. Determine if the quorum device is used in the disksets containing the zonpaths, and determine if the quorum device uses SCSI2 or SCSI3 reservations.

```
# clquorum show
```

- b. If the quorum device is within a LUN of the disksets, add a new LUN as a quorum device that is not part of any disk set containing the zone path.

```
# clquorum add new-didname
```

- c. Remove the old quorum device.

```
# clquorum remove old-didname
```

- d. If SCSI2 reservations are used for the old quorum device, scrub SCSI2 reservations from the old quorum and verify that there are no SCSI2 reservations left.

To learn how to run the `pgre` commands, you need to install and use the Diagnostic Toolkit (SUNWscdtk) package provided by Oracle Support.

Note – If you accidentally scrub reservation keys on your active quorum device, you must remove and re-add the quorum device to put new reservations keys on your quorum device.

2 Evacuate the node you want to patch.

```
# clresourcegroup evacuate -n node1
```

3 Take offline the resource group or resource groups that contain HA Solaris Container resources.

```
# clresourcegroup offline resourcegroupname
```

- 4 **Disable all the resources in the resource groups that you took offline.**
`# clresource disable resourcename`
- 5 **Unmanage the resource groups you took offline.**
`# clresourcegroup unmanage resourcegroupname`
- 6 **Take offline the corresponding device group or device groups.**
`# cldevicegroup offline cldevicegroupname`
- 7 **Disable the device groups that you took offline**
`# cldevicegroup disable devicegroupname`
- 8 **Boot the passive node out of the cluster.**
`# reboot -- -x`
- 9 **Verify that the SMF start methods are completed on the passive node before proceeding.**
`# svcs -x`
- 10 **Verify that any reconfiguration process on the active node has completed.**
`# cluster status`
- 11 **Determine if SCSI-2 reservations exist on the disk in the disk set and release the keys. Follow these instructions to determine if SCSI-2 reservations exist and then release them.**
 - For all disks in the disk set, run the following command: `/usr/cluster/lib/sc/scsi -c disfailfast -d /dev/did/rdisk/d#s2.`
 - If keys are listed, release them by running the following command:
`/usr/cluster/lib/sc/scsi -c release -d /dev/did/rdisk/d#s2.`

When you finish releasing the reservation keys, skip Step #12 and proceed to Step #13.
- 12 **Determine if there are any SCSI-3 reservations on the disks in the disksets.**
 - a. **Run the following command on all disks in the disksets.**
`# /usr/cluster/lib/sc/scsi -c inkeys -d /dev/did/rdisk/didnames2`
 - b. **If keys are listed, scrub them.**
`# /usr/cluster/lib/sc/scsi -c scrub -d /dev/did/rdisk/didnames2`
- 13 **Take ownership of the metaset on the passive node**
`# metaset -s disksetname -C take -f`
- 14 **Mount the file system or file systems that contains the zone path on the passive node.**
`# mount device mountpoint`

- 15 **Switch to single-user mode on the passive node.**
`# init s`
- 16 **Halt all possible booted zones that are not under the Oracle Solaris Cluster Data Service for Solaris Container control.**
`# zoneadm -z zonename halt`
- 17 **(Optional) If you install multiple patches, for performance reasons you can choose to boot all the configured zones in single-user mode.**
`# zoneadm -z zonename boot -s`
- 18 **Apply the patches.**
- 19 **Reboot the node and wait until all its SMF start methods are finished. Perform the `svcs -a` command only after the node has been rebooted.**
`# reboot`

`# svcs -a`
The first node is now ready.
- 20 **Evacuate the second node you want to patch.**
`# clresourcegroup evacuate -n node2`
- 21 **Repeat steps 8 through 13 for the second node.**
- 22 **Detach the zones you patched already to speed up the patch process.**
`# zoneadm -z zonename detach`
- 23 **Switch to single-user mode on the passive node.**
`# init s`
- 24 **Halt all possible booted zones that are not under the Oracle Solaris Cluster Data Service for Solaris Container control.**
`# zoneadm -z zonename halt`
- 25 **(Optional) If you install multiple patches, for performance reasons you can choose to boot all the configured zones in single-user mode.**
`# zoneadm -z zonename boot -s`
- 26 **Apply the patches.**
- 27 **Attach the zones you detached.**
`# zoneadm -z zonename attach -F`

28 Reboot the node into cluster mode.

```
# reboot
```

29 Bring online the device group or device groups.**30 Start the resource groups.****31 Check to see if you need to commit the patch software by using the `scversions` command.**

```
# /usr/cluster/bin/scversions
```

You will see one of the following results:

```
Upgrade commit is needed.
```

```
Upgrade commit is NOT needed. All versions match.
```

32 If a commit is needed, commit the patch software.

```
# scversions -c
```

Note – Running `scversions` will cause one or more CMM reconfigurations, depending on the situation.

Changing an Oracle Solaris Cluster Patch

To remove an Oracle Solaris Cluster patch that you've applied to your cluster, you must first remove the new Oracle Solaris Cluster patch, and then reapply the previous patch or update release. To remove the new Oracle Solaris Cluster patch, see the following procedures. To reapply a previous Oracle Solaris Cluster patch, see one of the following procedures:

- [“How to Apply a Rebooting Patch \(Node\)” on page 299](#)
- [“How to Apply a Rebooting Patch \(Cluster\)” on page 303](#)
- [“How to Apply a Nonrebooting Oracle Solaris Cluster Patch” on page 307](#)

Note – Before applying an Oracle Solaris Cluster patch, check the patch's README file.

▼ How to Remove a Non-Rebooting Oracle Solaris Cluster Patch

1 Become superuser on any node in the cluster.**2 Remove the non-rebooting patch.**

```
# patchrm patchid
```

▼ How to Remove a Rebooting Oracle Solaris Cluster Patch

- 1 Become superuser on any node in the cluster.
- 2 Boot the cluster node into noncluster mode. For information about booting a node into noncluster mode, see [“How to Boot a Node in Noncluster Mode” on page 78](#).
- 3 Remove the rebooting patch.

```
# patchrm patchid
```
- 4 Reboot the cluster node back into cluster mode.

```
# reboot
```
- 5 Repeat steps 2 through 4 for each cluster node.

Backing Up and Restoring a Cluster

This chapter provides the following sections:

- “Backing Up a Cluster” on page 313
- “Restoring Cluster Files” on page 325

Backing Up a Cluster

TABLE 12-1 Task Map: Backing Up Cluster Files

Task	Instructions
Find the names of the file systems you want to back up	“How to Find File System Names to Back Up” on page 314
Calculate how many tapes you need to contain a full backup	“How to Determine the Number of Tapes Needed for a Full Backup” on page 314
Back up the root file system	“How to Back Up the Root (/) File System” on page 315
Perform online backup for mirrored or plexed file systems	“How to Perform Online Backups for Mirrors (Solaris Volume Manager)” on page 318 “How to Perform Online Backups for Volumes (Veritas Volume Manager)” on page 320
Back up the cluster configuration	“How to Back Up the Cluster Configuration” on page 324
Back up disk partitioning configuration for storage disk	See the documentation for your storage disk

▼ How to Find File System Names to Back Up

Use this procedure to determine the names of the file systems that you want to back up.

1 Display the contents of the `/etc/vfstab` file.

You do not need to be superuser or assume an equivalent role to run this command.

```
# more /etc/vfstab
```

2 Look in the mount-point column for the name of the file system that you are backing up.

Use this name when you back up the file system.

```
# more /etc/vfstab
```

Example 12-1 Finding File System Names to Back Up

The following example displays the names of available file systems that are listed in the `/etc/vfstab` file.

```
# 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 that you need to back up a file system.

1 Become superuser or assume an equivalent role on the cluster node that you are backing 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.

Example 12-2 Determining the Number of Tapes Needed

In the following example, the file system size of 905,881,620 bytes easily fits on a 4-Gbyte 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. Ensure that the cluster is running without errors before performing the backup procedure.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on the cluster node that you are backing up.**
- 2 **Switch each running data service from the node to be backed up to another node in the cluster.**

```
# clnode evacuate node
```

node Specifies the node from which you are switching resource groups and device groups.
- 3 **Shut down the node.**

```
# shutdown -g0 -y -i0
```
- 4 **Reboot the node in noncluster mode.**
 - On SPARC based systems, run the following command.

```
ok boot -xs
```
 - On x86 based systems, run the following commands.

```
phys-schost# shutdown -g -y -i0
```

Press any key to continue

- a. In the GRUB menu, use the arrow keys to select the appropriate Oracle Solaris entry and type **e** to edit its commands.

The GRUB menu appears similar to the following:

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-----+
| Solaris 10 /sol_10_x86                               |
| Solaris failsafe                                     |
|                                                       |
+-----+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

For more information about GRUB based booting, see [“Booting an x86 Based System by Using GRUB \(Task Map\)”](#) in *System Administration Guide: Basic Administration*.

- b. In the boot parameters screen, use the arrow keys to select the kernel entry and type **e** to edit the entry.

The GRUB boot parameters screen appears similar to the following:

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+-----+
| root (hd0,0,a)                                       |
| kernel /platform/i86pc/multiboot                    |
| module /platform/i86pc/boot_archive                 |
+-----+
Use the ^ and v keys to select which entry is highlighted.
Press 'b' to boot, 'e' to edit the selected command in the
boot sequence, 'c' for a command-line, 'o' to open a new line
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.
```

- c. Add **-x** to the command to specify that the system boot into noncluster mode.

```
[ Minimal BASH-like line editing is supported. For the first word, TAB
lists possible command completions. Anywhere else TAB lists the possible
completions of a device/filename. ESC at any time exits. ]
```

```
grub edit> kernel /platform/i86pc/multiboot -x
```

- d. Press the **Enter** key to accept the change and return to the boot parameters screen.

The screen displays the edited command.

```
GNU GRUB version 0.95 (615K lower / 2095552K upper memory)
+-----+
| root (hd0,0,a)                                       |
| kernel /platform/i86pc/multiboot -x                 |
| module /platform/i86pc/boot_archive                 |
+-----+
Use the ^ and v keys to select which entry is highlighted.
```

Press 'b' to boot, 'e' to edit the selected command in the boot sequence, 'c' for a command-line, 'o' to open a new line after ('O' for before) the selected line, 'd' to remove the selected line, or escape to go back to the main menu.-

e. Type b to boot the node into noncluster mode.

Note – This change to the kernel boot parameter command does not persist over the system boot. The next time you reboot the node, it will boot into cluster mode. To boot into noncluster mode instead, perform these steps to again to add the -x option to the kernel boot parameter command.

5 Back up the root (/) file system by creating a UFS snapshot.

a. Ensure that the file system has enough disk space for the backing-store file.

```
# df -k
```

b. Ensure that a backing-store file of the same name and location does not already exist.

```
# ls /backing-store-file
```

c. Create the UFS snapshot.

```
# fssnap -F ufs -o bs=/backing-store-file /file-system
```

d. Verify that the snapshot has been created.

```
# /usr/lib/fs/ufs/fssnap -i /file-system
```

6 Back up the file-system snapshot.

```
# ufsdump 0ucf /dev/rmt/0 snapshot-name
```

For example:

```
# ufsdump 0ucf /dev/rmt/0 /dev/rfssnap/1
```

7 Verify that the snapshot is backed up.

```
# ufsrestore ta /dev/rmt/0
```

8 Reboot the node in cluster mode.

```
# init 6
```

Example 12-3 Backing Up the Root (/) File System

In the following example, a snapshot of the root (/) file system is saved to /scratch/usr.back.file in the /usr directory. ⁶

```
# fssnap -F ufs -o bs=/scratch/usr.back.file /usr  
/dev/fssnap/1
```

▼ How to Perform Online Backups for Mirrors (Solaris Volume Manager)

A mirrored Solaris Volume Manager volume 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 resynchronized 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, ensure the cluster is running without errors before performing the backup procedure.

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume an equivalent role on the cluster node that you are backing 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 disk set name.

- 3 **Use the `lockfs(1M)` command with the `-w` option to lock the file system from writes.**

```
# lockfs -w mountpoint
```

Note – You must lock the file system only if a UFS file system resides on the mirror. For example, if the Solaris Volume Manager volume is set up as a raw device for database management software or some other specific application, you do not need to use the `lockfs` command. You might, however, 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 continue to be made from the other submirrors. However, the offline submirror is unsynchronized as soon as the first write is made to the mirror. This inconsistency is corrected when the offline submirror is brought back online. You do not need to run `fsck`.

6 Unlock the file systems and allow writes to continue, using the `lockfs` command with the `-u` option.

```
# lockfs -u mountpoint
```

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 the backup utility that you usually use.

```
# ufsdump 0ucf dump-device submirror
```

Note – Use the raw device (`/rdsk`) name for the submirror, rather than the block device (`/dsk`) name.

9 Use the `metattach(1M)` command to place the metadvice or volume back online.

```
# metattach -s setname mirror submirror
```

When the metadvice or volume is placed online, it is automatically resynchronized with the mirror.

10 Use the `metastat` command to verify that the submirror is resynchronizing.

```
# metastat -s setname mirror
```

Example 12–4 Performing Online Backups for Mirrors (Solaris Volume Manager)

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
[Bring the submirror back online:]
# metattach -s schost-1 d0 d30
schost-1/d0: submirror schost-1/d30 is attached
[Resynchronize 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 result is accomplished 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.

Ensure that the cluster is running without errors before performing the backup procedure.

The `phys -schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Log on to any node in the cluster, and become superuser or assume a role that provides `soLaris.cluster.admin` RBAC authorization on the current primary node for the disk group on the cluster.**

- 2 **List the disk group information.**

```
# vxprint -g diskgroup
```

- 3 **Determine which node has the disk group currently imported, indicating it is the primary node for the disk group.**

```
# cldevicegroup status
```

- 4 **Create a snapshot of the volume.**

```
# vxassist -g diskgroup snapstart volume
```

Note – Creating a snapshot can take a long time, depending on the size of your volume.

- 5 **Verify that the new volume was created.**

```
# 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.**

```
# clresourcegroup offline resource-group
```

Note – Stop all data services 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](#) and [Step 8](#).

- 7 **Create a backup volume named `bkup-vol` and attach the snapshot volume to it.**

```
# vxassist -g diskgroup snapshot volume bkup-vol
```

- 8 **Restart any data services that were stopped in [Step 6](#), using the `clresourcegroup` command.**

```
# clresourcegroup online -zone -n node resourcegroup
```

node The name of the node.

zone The name of the global-cluster non-voting node (*node*) that can master the resource group. Specify *zone* only if you specified a non-voting node when you created the resource group.

- 9 **Verify the volume is now attached to the new volume bkup-vol.**
`# vxprint -g diskgroup`
- 10 **Register the device group configuration change.**
`# cldevicegroup sync diskgroup`
- 11 **Check the backup volume.**
`# 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 that you normally use.
`# ufsdump 0ucf dump-device /dev/vx/dsk/diskgroup/bkup-vol`
- 13 **Remove the temporary volume.**
`# vxedit -rf rm bkup-vol`
- 14 **Register the disk group configuration changes.**
`# cldevicegroup sync diskgroup`

Example 12-5 Performing Online Backups for Volumes (Veritas Volume Manager)

In the following example, the cluster node `phys-schost-2` is the primary owner of the device 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 or assume a role that provides `solaris.cluster.admin` RBAC authorization on the primary node.]

[Identify the current primary node for the device group:]

```
# cldevicegroup status
-- 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 device group information:]

```
# vxprint -g schost-1
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  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     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  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:]
# clresourcegroup offline nfs-rg
[Create a copy of the volume:]
# vxassist -g schost-1 snapshot vol01 bkup-vol
[Restart data services, if necessary:]
# clresourcegroup online -n phys-schost-1 nfs-rg
[Verify bkup-vol was created:]
# vxprint -g schost-1
TY NAME          ASSOC      KSTATE    LENGTH    PLOFFS    STATE     TUTIL0    PUTIL0
dg schost-1      schost-1   -         -         -         -         -         -

dm schost-101    c1t1d0s2  -          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:]
# cldevicegroup sync schost-1
[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:]
# cldevicegroup sync schost-1

```

▼ How to Back Up the Cluster Configuration

To ensure that your cluster configuration is archived and to facilitate easy recovery of the your cluster configuration, periodically back up your cluster configuration. Oracle Solaris Cluster provides the ability to export your cluster configuration to an eXtensible Markup Language (XML) file.

- 1 **Log on to any node in the cluster, and become superuser or assume a role that provides `soLaris.cluster.read` RBAC authorization.**

- 2 **Export the cluster configuration information to a file.**

```
# /usr/cluster/bin/cluster export -o configfile
```

configfile The name of the XML configuration file that the cluster command is exporting the cluster configuration information to. For information about the XML configuration file, see [clconfiguration\(5CL\)](#)

- 3 **Verify that the cluster configuration information was successfully exported to the XML file.**

```
# vi configfile
```

Restoring Cluster Files

The `ufsrestore(1M)` command copies files to disk, relative to the current working directory, from backups created by using the `ufsdump(1M)` 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 or assumed an equivalent role, 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 the following information.

- Which tapes you need
- The raw device name on which you are restoring the file system
- The type of tape drive you are using
- 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

TABLE 12-2 Task Map: Restoring Cluster Files

Task	Instructions
For Solaris Volume Manager, restore files interactively	“How to Restore Individual Files Interactively (Solaris Volume Manager)” on page 325
For Solaris Volume Manager, restore the root (/) file system	“How to Restore the Root (/) File System (Solaris Volume Manager)” on page 326 “How to Restore a Root (/) File System That Was on a Solaris Volume Manager Volume” on page 328
For Veritas Volume Manager, restore a root (/) file system	“How to Restore a Nonencapsulated Root (/) File System (Veritas Volume Manager)” on page 333
For Veritas Volume Manager, restore an encapsulated root (/) file system	“How to Restore an Encapsulated Root (/) File System (Veritas Volume Manager)” on page 335

▼ How to Restore Individual Files Interactively (Solaris Volume Manager)

Use this procedure to restore one or more individual files. Ensure that the cluster is running without errors before performing the restore procedure.

- 1 **Become superuser or assume a role that provides `solaris.cluster.admin` RBAC authorization on the cluster node you are restoring.**

- 2 Stop all the data services that are using the files to be restored.

```
# clresourcegroup offline resource-group
```

- 3 Restore the files.

```
# ufsrestore
```

▼ How to Restore the Root (/) File System (Solaris Volume Manager)

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. Ensure that the cluster is running without errors before performing the restore procedure.

Note – Because you must partition the new disk by using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and recreate file systems as appropriate.

The `phys - schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 **Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on a cluster node with access to the disksets to which the node to be restored is also attached.**

Use a node *other than* the node that you are restoring.

- 2 **Remove the hostname of the node being restored from all metaset.**

Run this command from a node in the metaset other than the node that you are removing. Because the recovering node is offline, the system will display an `RPC: Rpcbnd failure - RPC: Timed out error`. Ignore this error and continue to the next step.

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

-s setname	Specifies the disk set name.
-f	Deletes the last host from the disk set.
-d	Deletes from the disk set.
-h nodelist	Specifies the name of the node to delete from the disk set.

3 Restore the root (/) and /usr file systems.

To restore the root and /usr file systems, follow the procedure in [Chapter 26, “Restoring UFS Files and File Systems \(Tasks\)”](#), in *System Administration Guide: Devices and File Systems*. Omit the step in the Oracle Solaris OS procedure to reboot the system.

Note – Ensure that you create the /global/.devices/node@nodeid file system.

4 Reboot the node in multiuser mode.

```
# reboot
```

5 Replace the device ID.

```
# cldevice repair rootdisk
```

6 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.

7 From a cluster node other than the restored node add the restored node to all disksets.

```
phys-schost-2# metaset -s setname -a -h nodelist
```

-a Creates and adds the host to the disk set.

The node is rebooted into cluster mode. The cluster is ready to use.

Example 12–6 Restoring the Root (/) File System (Solaris Volume Manager)

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 disk set `schost-1`. All other commands are run from `phys-schost-1`. A new boot block is created on `/dev/rdisk/c0t0d0s0`, and three state database replicas are recreated on `/dev/rdisk/c0t0d0s4`.

```
[Become superuser or assume a role that provides solaris.cluster.modify RBAC authorization 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:]
```

```
Restore the root (/) and /usr file system using the procedure in the Solaris system administration documentation
```

```
[Reboot:]
```

```
# reboot
```

```
[Replace the disk ID:]
```

```
# cldevice repair /dev/dsk/c0t0d0
```

```
[Re-create state database replicas:]
# metadb -c 3 -af /dev/rdisk/c0t0d0s4
[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 Solaris Volume Manager Volume

Use this procedure to restore a root (/) file system that was on a Solaris Volume Manager volume 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. Ensure that the cluster is running without errors before performing the restore procedure.

Note – Because you must partition the new disk by using the same format as the failed disk, identify the partitioning scheme before you begin this procedure, and recreate file systems as appropriate.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 1 Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization on a cluster node with access to the disk set, *other than* the node you restoring.**
Use a node *other than* the node that you are restoring.
- 2 Remove the hostname of the node being restored from all disksets to which it is attached. Execute the following command once for each diskset.**

```
# metaset -s setname -d -h hostname
```

<code>-s setname</code>	Specifies the metaset name.
<code>-f</code>	Deletes the last host from the disk set.
<code>-d</code>	Deletes from the metaset.
<code>-h nodelist</code>	Specifies the name of the node to delete from the metaset.
<code>-h hostname</code>	Specifies the name of the host.

`-m mediator_host_list` Specifies the name of the mediator host to add or delete from the disk set.

3 If the node is a dual-string mediator host, remove the mediator. Execute the following command once for each diskset to which the node is attached.

```
# metaset -ssetname-d -m hostname
```

4 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 shipped with your server.

5 Boot the node that you are restoring. The repaired node is booted into single user mode from the CD-ROM, so Solaris Volume Manager is not running on the node.

- If you are using the Oracle Solaris OS CD, note the following:

- SPARC: Type:

```
ok boot cdrom -s
```

- x86: Insert the CD into the system's CD drive and boot the system by shutting it down and then turning it off and on. In the Current Boot Parameters screen, type b or i.

```
<<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@
7,1/sd@0,0:a
Boot args:

Type b [file-name] [boot-flags] <ENTER> to boot with options
or i <ENTER> to enter boot interpreter
or <ENTER> to boot with defaults
```

```
<<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -s
```

- If you are using a Solaris JumpStart server, note the following:

- SPARC: Type:

```
ok boot net -s
```

- x86: Insert the CD into the system's CD drive and boot the system by shutting it down and then turning it off and on. In the Current Boot Parameters screen, type b or i.

```
<<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@
7,1/sd@0,0:a
Boot args:

Type b [file-name] [boot-flags] <ENTER> to boot with options
or i <ENTER> to enter boot interpreter
or <ENTER> to boot with defaults
```

```
<<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -s
```

6 Create all the partitions and swap space on the root disk by using the format command.

Re-create the original partitioning scheme that was on the failed disk.

7 Create the root (/) file system and other file systems as appropriate, by using the newfs command

Re-create the original file systems that were on the failed disk.

Note – Ensure that you create the `/global/.devices/node@nodeid` file system.

8 Mount the root (/) file system on a temporary mount point.

```
# mount device temp-mountpoint
```

9 Use the following commands to restore the root (/) file system.

```
# cd temp-mountpoint
# ufsrestore rvf dump-device
# rm restoresymtable
```

10 Install a new boot block on the new disk.

```
# /usr/sbin/installboot /usr/platform/'uname -i'/lib/fs/ufs/bootblk
raw-disk-device
```

11 Remove the lines in the `/temp-mountpoint/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)
```

12 Edit the `/temp-mountpoint/etc/vfstab` file to change the root entry from a Solaris Volume Manager volume to a corresponding normal slice for each file system on the root disk that is part of the metadvice or volume.

Example:

Change from–

```
/dev/md/dsk/d10 /dev/md/rdisk/d10 / ufs 1 no -
```

Change to–

```
/dev/dsk/c0t0d0s0 /dev/rdisk/c0t0d0s0 / ufs 1 no -
```

13 Unmount the temporary file system, and check the raw disk device.

```
# cd /
# umount temp-mountpoint
# fsck raw-disk-device
```

14 Reboot the node in multiuser mode.

```
# reboot
```

15 Replace the device ID.

```
# cldevice repair rootdisk
```

16 Use the `metadb` 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.

17 From a cluster node other than the restored node, add the restored node to all disksets.

```
phys-schost-2# metaset -s setname -a -h nodelist
```

-a Adds (creates) the metaset.

Set up the volume/mirror for root (/) according to the documentation.

The node is rebooted into cluster mode.

18 If the node was a dual-string mediator host, re-add the mediator.

```
phys-schost-2# metaset -s setname -a -m hostname
```

Example 12-7 Restoring a Root (/) File System That Was on a Solaris Volume Manager Volume

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/rdisk/c0t0d0s0`, and three state database replicas are recreated on `/dev/rdisk/c0t0d0s4`.

[Become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization 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 -d -h phys-schost-1
```

[Replace the failed disk and boot the node:]

Boot the node from the Oracle Solaris OS CD:

- SPARC: Type:

```
ok boot cdrom -s
```

- x86: Insert the CD into the system's CD drive and boot the system by shutting it down and then turning it off and on. In the Current Boot Parameters screen, type `b` or `i`.

```

<<< Current Boot Parameters >>>
Boot path: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/pci8086,341a@7,1/
sd@0,0:a
Boot args:

Type  b [file-name] [boot-flags] <ENTER> to boot with options
     or  i <ENTER>                        to enter boot interpreter
     or  <ENTER>                          to boot with defaults

<<< timeout in 5 seconds >>>
Select (b)oot or (i)nterpreter: b -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 \
-i'/lib/fs/ufs/bootblk /dev/rdisk/c0t0d0s0

[Remove the lines in /temp-mountpoint/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-mountpoint/etc/vfstab file]
Example:
Change from-
/dev/md/dsk/d10 /dev/md/rdsk/d10 / ufs 1 no -
Change to-
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0 /usr ufs 1 no -
[Unmount the temporary file system and check the raw disk device:]
# cd /
# umount /a
# fsck /dev/rdsk/c0t0d0s0
[Reboot:]
# reboot
[Replace the disk ID:]
# cldevice repair /dev/rdsk/c0t0d0
[Re-create state database replicas:]
# metadb -c 3 -af /dev/rdsk/c0t0d0s4

```

```
[Add the node back to the metaset:]  
phys-schost-2# metaset -s schost-1 -a -h phys-schost-1
```

▼ How to Restore a Nonencapsulated Root (/) File System (Veritas Volume Manager)

Use this procedure to restore a nonencapsulated root (/) file system to a node. The node being restored should not be booted. Ensure the cluster is running without errors before performing the restore procedure.

Note – Because 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.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

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 shipped with your server.

2 Boot the node that you are restoring.

- If you are using the Oracle Solaris OS CD, at the OpenBoot PROM ok prompt, type the following command:

```
ok boot cdrom -s
```

- If you are using a Solaris JumpStart server, at the OpenBoot PROM ok prompt, type the following command:

```
ok boot net -s
```

3 Create all the partitions and swap on the root disk by using the format command.

Re-create 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 command.

Re-create the original file systems that were on the failed disk.

Note – Ensure that you create the `/global/.devices/node@nodeid` file system.

5 Mount the root (/) file system on a temporary mount point.

```
# mount device temp-mountpoint
```

6 Restore the root (/) file system from backup, and unmount and check the file system.

```
# cd temp-mountpoint
# ufsrestore rvf dump-device
# rm restoresymtable
# cd /
# umount temp-mountpoint
# 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 in multiuser mode.

```
# reboot
```

9 Update the device ID.

```
# cldevice repair /dev/rdisk/disk-device
```

10 Press Control-d to resume in multiuser mode.

The node reboots into cluster mode. The cluster is ready to use.

Example 12–8 Restoring a Nonencapsulated Root (/) File System (Veritas Volume Manager)

The following example shows a nonencapsulated root (/) file system that is restored to the node `phys-schost-1` from the tape device `/dev/rmt/0`.

[Replace the failed disk and boot the node:]

Boot the node from the Oracle Solaris OS CD. At the OpenBoot PROM ok prompt, type the following command:

```
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/rdisk/c0t0d0s0
[Install a new boot block:]
# /usr/sbin/installboot /usr/platform/'uname \
-i'/lib/fs/ufs/bootblk /dev/rdisk/c0t0d0s0

[Reboot:]
# reboot
[Update the disk ID:]
# cldevice repair /dev/rdisk/c0t0d0
```

▼ 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. Ensure the cluster is running with errors before performing the restore procedure.

Note – Because 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.

The `phys-schost#` prompt reflects a global-cluster prompt. Perform this procedure on a global cluster.

This procedure provides the long forms of the Oracle Solaris Cluster commands. Most commands also have short forms. Except for the long and short forms of the command names, the commands are identical.

- 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 shipped with your server.
- 2 **Boot the node that you are restoring.**
 - If you are using the Oracle Solaris OS CD, at the OpenBoot PROM `ok` prompt, type the following command:
`ok boot cdrom -s`
 - If you are using a Solaris JumpStart server, at the OpenBoot PROM `ok` prompt, type the following command:
`ok boot net -s`

3 Create all the partitions and swap space on the root disk by using the format command.

Re-create the original partitioning scheme that was on the failed disk.

4 Create the root (/) file system and other file systems as appropriate, by using the newfs command.

Re-create the original file systems that were on the failed disk.

Note – Ensure that you create the `/global/.devices/node@nodeid` file system.

5 Mount the root (/) file system on a temporary mount point.

```
# mount device temp-mountpoint
```

6 Restore the root (/) file system from backup.

```
# cd temp-mountpoint
# ufsrestore rvf dump-device
# rm restoresymtable
```

7 Create an empty install-db file.

This file puts the node in VxVM installation mode at the next reboot.

```
# touch \
/temp-mountpoint/etc/vx/reconfig.d/state.d/install-db
```

8 Remove the following entries from the /temp-mountpoint/etc/system file.

```
* rootdev:/pseudo/vxio@0:0
* set vxio:vol_rootdev_is_volume=1
```

9 Edit the /temp-mountpoint/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/vx/rdisk/rootdg/rootvol / ufs 1 no -
```

Change to–

```
/dev/dsk/c0t0d0s0 /dev/rdsk/c0t0d0s0 / ufs 1 no -
```

10 Unmount the temporary file system and check the file system.

```
# cd /
# umount temp-mountpoint
# 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 multiuser mode.

```
# reboot
```

- 13 Update the device ID by using `sccidadm(1M)`.

```
# cldevice repair /dev/rdisk/c0t0d0
```
- 14 Run the `clvxvm encapsulate` command to encapsulate the disk and reboot.
- 15 If a conflict in minor number occurs with any other system, unmount the global devices and re-minor the disk group.
 - Unmount the global devices file system on the cluster node.

```
# umount /global/.devices/node@nodeid
```
 - Re-minor the rootdg disk group on the cluster node.

```
# vxdg reminor rootdg 100
```
- 16 Shut down and reboot the node in cluster mode.

```
# shutdown -g0 -i6 -y
```

Example 12-9 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:]

Boot the node from the Oracle Solaris OS CD. At the OpenBoot PROM `ok` prompt, type the following command:

```
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/vx/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:]
```

```
# 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:]
# reboot
[Update the disk ID:]
# cldevice repair /dev/rdisk/c0t0d0
[Encapsulate the disk:~:]
# vxinstall
Choose to encapsulate the root disk.
[If a conflict in minor number occurs, remirror the rootdg disk group:]
# umount /global/.devices/node@nodeid
# vxdg remirror rootdg 100
# shutdown -g0 -i6 -y
```

See Also For instructions about how to mirror the encapsulated root disk, see the [Oracle Solaris Cluster Software Installation Guide](#).

Administering Oracle Solaris Cluster With the Graphical User Interfaces

This chapter provides descriptions of Oracle Solaris Cluster Manager and Sun Management Center graphical user interface (GUI) tools, which you can use to administer many aspects of a cluster. It also contains procedures to configure and start Oracle Solaris Cluster Manager. The online help that is included with the Oracle Solaris Cluster Manager GUI provides instructions for accomplishing various Oracle Solaris Cluster administrative tasks.

This chapter includes the following:

- “Overview of Oracle Solaris Cluster Manager” on page 339
- “SPARC: Overview of Sun Management Center” on page 340
- “Configuring Oracle Solaris Cluster Manager” on page 341
- “Starting the Oracle Solaris Cluster Manager Software” on page 344

Overview of Oracle Solaris Cluster Manager

Oracle Solaris Cluster Manager is a GUI that enables you to graphically display cluster information, check the status of cluster components, and monitor configuration changes. Oracle Solaris Cluster Manager also enables you to perform many administrative tasks for the following Oracle Solaris Cluster components.

- Adapters
- Cables
- Data services
- Global devices
- Interconnects
- Junctions
- Node load limits
- NAS devices
- Nodes
- Quorum devices
- Resource groups

- Resources

Information about installing and using Oracle Solaris Cluster Manager can be found in the following locations.

- **Installing Oracle Solaris Cluster Manager:** See the *Oracle Solaris Cluster Software Installation Guide*.
- **Starting Oracle Solaris Cluster Manager:** See “Starting the Oracle Solaris Cluster Manager Software” on page 344.
- **Configuring port numbers, server addresses, security certificates, and users:** See “Configuring Oracle Solaris Cluster Manager” on page 341.
- **Installing and administering aspects of your cluster using Oracle Solaris Cluster Manager:** See the online help supplied with Oracle Solaris Cluster Manager.
- **Regenerating Oracle Solaris Cluster Manager security keys:** See “How to Regenerate Common Agent Container Security Keys” on page 343.

Note – However, Oracle Solaris Cluster Manager currently cannot perform all Oracle Solaris Cluster administrative tasks. You must use the command-line interface for some operations.

SPARC: Overview of Sun Management Center

The Oracle Solaris Cluster 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 Oracle Solaris Cluster module for Sun Management Center cannot perform Oracle Solaris Cluster configuration tasks. You must use the command-line interface for configuration operations. See “Command-Line Interface” in Chapter 1 for more information.

For information about installing and starting the Oracle Solaris Cluster module for Sun Management Center, see the [Chapter 8, “Installing the Oracle Solaris Cluster Module to Sun Management Center,”](#) in *Oracle Solaris Cluster Software Installation Guide*.

The Oracle Solaris Cluster module of Sun Management Center is Simple Network Management Protocol (SNMP) compliant. Oracle Solaris Cluster has created a Management Information Base (MIB) that can be used as the data definition by third-party management stations based on SNMP.

The Oracle Solaris Cluster MIB file is located at `/opt/SUNWsymon/modules/cfg/sun-cluster-mib.mib` on any cluster node.

The Oracle Solaris Cluster MIB file is an ASN.1 specification of the Oracle Solaris Cluster data that is modeled. This is the same specification used by all Sun Management Center MIBs. To

use the Oracle Solaris Cluster MIB, refer to the instructions for using other Sun Management Center MIBs in the in “SNMP MIBs for Sun Management Centre Modules” in *Sun Management Center 3.6 User's Guide*.

Configuring Oracle Solaris Cluster Manager

Oracle Solaris Cluster Manager is a GUI that you can use to administer and view the status of all aspects of quorum devices, IPMP groups, interconnect components, and global devices. You can use the GUI in place of many of the Oracle Solaris Cluster CLI commands.

The procedure for installing Oracle Solaris Cluster Manager on your cluster is included in the *Oracle Solaris Cluster Software Installation Guide*. The Oracle Solaris Cluster Manager online help contains instructions for completing various tasks by using the GUI.

This section contains the following procedures for reconfiguring Oracle Solaris Cluster Manager after initial installation.

- “Setting up RBAC Roles” on page 341
- “How to Change the Server Address for Oracle Solaris Cluster Manager” on page 342
- “How to Regenerate Common Agent Container Security Keys” on page 343

Setting up RBAC Roles

The Oracle Solaris Cluster Manager uses RBAC to determine who has rights to administer the cluster. Several RBAC rights profiles are included in the Oracle Solaris Cluster software. You can assign these rights profiles to users or to roles to give users different levels of access to Oracle Solaris Cluster. For more information about how to set up and manage RBAC for Oracle Solaris Cluster software, see [Chapter 2, “Oracle Solaris Cluster and RBAC.”](#)

▼ How to Use the Common Agent Container to Change the Port Numbers for Services or Management Agents

If the default port numbers for your common agent container services conflict with other running processes, you can use the `cacaoadm` command to change the port number of the conflicting service or management agent on each node of the cluster.

- 1 On all cluster nodes, stop the common agent container management daemon.

```
# /opt/bin/cacaoadm stop
```

- 2 Stop Sun Java Web Console.

```
# /usr/sbin/smcwebserver stop
```

- 3 Retrieve the port number currently used by the common agent container service with the `get-param` subcommand.

```
# /opt/bin/cacaoadm get-param parameterName
```

You can use the `cacaoadm` command to change the port numbers for the following common agent container services. The following list provides some examples of services and agents that can be managed by the common agent container, along with corresponding parameter names.

JMX connector port	<code>jmxmp-connector-port</code>
SNMP port	<code>snmp-adapter-port</code>
SNMP trap port	<code>snmp-adapter-trap-port</code>
Command stream port	<code>commandstream-adapter-port</code>

- 4 Change a port number.

```
# /opt/bin/cacaoadm set-param parameterName=parameterValue
=parameterValue
```

- 5 Repeat [Step 4](#) on each node of the cluster.

- 6 Restart Sun Java Web Console.

```
# /usr/sbin/smcwebserver start
```

- 7 Restart the common agent container management daemon on all cluster nodes.

```
# /opt/bin/cacaoadm start
```

▼ How to Change the Server Address for Oracle Solaris Cluster Manager

If you change the hostname of a cluster node, you must change the address from which Oracle Solaris Cluster Manager runs. The default security certificate is generated based on the node's hostname at the time Oracle Solaris Cluster Manager is installed. To reset the node's hostname, delete the certificate file, keystore and restart Oracle Solaris Cluster Manager. Oracle Solaris Cluster Manager automatically creates a new certificate file with the new hostname. You must complete this procedure on any node that has had its hostname changed.

- 1 Remove the certificate file, keystore, located in `/etc/opt/webconsole`.

```
# cd /etc/opt/webconsole
# pkgrm keystore
```

- 2 Restart Oracle Solaris Cluster Manager.

```
# /usr/sbin/smcwebserver restart
```

▼ How to Regenerate Common Agent Container Security Keys

Oracle Solaris Cluster Manager uses strong encryption techniques to ensure secure communication between the Oracle Solaris Cluster Manager web server and each cluster node.

The keys that Oracle Solaris Cluster Manager uses are stored under the `/etc/opt/SUNWcacao/security` directory on each node. They should be identical across all cluster nodes.

Under normal operation, these keys can be left in their default configuration. If you change the hostname of a cluster node, you must regenerate the common agent container security keys. You might also need to regenerate the keys because a possible key compromise (for example, root compromise on the machine). To regenerate the security keys, use the following procedure.

- 1 **On all cluster nodes, stop the common agent container management daemon.**

```
# /opt/bin/cacaoadm stop
```

- 2 **On one node of the cluster, regenerate the security keys.**

```
phys-schost-1# /opt/bin/cacaoadm create-keys --force
```

- 3 **Restart the common agent container management daemon on the node on which you regenerated the security keys.**

```
phys-schost-1# /opt/bin/cacaoadm start
```

- 4 **Create a tar file of the `/etc/cacao/instances/default` directory.**

```
phys-schost-1# cd /etc/cacao/instances/default
phys-schost-1# tar cf /tmp/SECURITY.tar security
```

- 5 **Copy the `/tmp/Security.tar` file to each of the cluster nodes.**

- 6 **On each node to which you copied the `/tmp/SECURITY.tar` file, extract the security files.**

Any security files that already exist in the `/etc/opt/SUNWcacao/` directory are overwritten.

```
phys-schost-2# cd /etc/cacao/instances/default
phys-schost-2# tar xf /tmp/SECURITY.tar
```

- 7 **Delete the `/tmp/SECURITY.tar` file from each node in the cluster.**

You must delete each copy of the tar file to avoid security risks.

```
phys-schost-1# rm /tmp/SECURITY.tar
```

```
phys-schost-2# rm /tmp/SECURITY.tar
```

- 8 On all nodes, restart the common agent container management daemon.**

```
phys-schost-1# /opt/bin/cacaoadm start
```

- 9 Restart Oracle Solaris Cluster Manager.**

```
# /usr/sbin/smcwebserver restart
```

Starting the Oracle Solaris Cluster Manager Software

The Oracle Solaris Cluster Manager graphical user interface (GUI) provides an easy way to administer some aspects of the Oracle Solaris Cluster software. See the Oracle Solaris Cluster Manager online help for more information.

Both Sun Java Web Console and the common agent container are started automatically when you boot the cluster. If you need to verify that Sun Java Web Console and the common agent container are running, see the Troubleshooting section immediately following this procedure.

▼ How to Start Oracle Solaris Cluster Manager

This procedure shows how to start Oracle Solaris Cluster Manager on your cluster.

- 1 Determine if you intend to access Oracle Solaris Cluster Manager by using the cluster node root user name and password or set up a different user name and password?**

- If you will access Oracle Solaris Cluster Manager by using the cluster node root user name, go to [Step 5](#).
- If you intend to set up a different user name and password, go to [Step 3](#) to set up Oracle Solaris Cluster Manager user accounts.

- 2 Become superuser on a cluster node.**

- 3 Create a user account to access the cluster through Oracle Solaris Cluster Manager.**

You use the `useradd(1M)` command to add a user account to the system. You must set up at least one user account to access Oracle Solaris Cluster Manager if you do not use the root system account. Oracle Solaris Cluster Manager user accounts are used only by Oracle Solaris Cluster Manager. These accounts do not correspond to any Oracle Solaris OS system user accounts. Creating and assigning an RBAC role to a user account is described in more detail in “[Creating and Assigning an RBAC Role With an Oracle Solaris Cluster Management Rights Profile](#)” on page 53.

Note – Users who do not have a user account set up on a particular node cannot access the cluster through Oracle Solaris Cluster Manager from that node, nor can users manage that node through another cluster node to which the users do have access.

- 4 (Optional) Repeat **Step 3** to set up additional user accounts.
- 5 From the administrative console or any other machine outside the cluster, start a browser.
- 6 Ensure that the browser's disk and memory cache sizes are set to a value that is greater than 0.
- 7 Ensure that the Java and Javascript are enabled in the browser.
- 8 From the browser, connect to the Oracle Solaris Cluster Manager port on one node of the cluster.
The default port number is 6789.
`https://node:6789/`
- 9 Accept any certificates that are presented by the web browser.
The Java Web Console login page is displayed.
- 10 Enter the user name and password for the user you want to access Oracle Solaris Cluster Manager.
- 11 Click the **Log In** button.
The Java Web Console application launch page is displayed.
- 12 Click the **Oracle Solaris Cluster Manager** link under the **Systems** category.
- 13 Accept any additional certificates that are presented by the web browser.
- 14 If you cannot connect to the Oracle Solaris Cluster Manager, perform the following substeps to determine if a restricted network profile was chosen during Solaris installation and to restore external access to the Java Web Console service.
If you choose a restricted network profile during Oracle Solaris installation, external access for the Sun Java Web Console service is restricted. This network is required to use the Oracle Solaris Cluster Manager GUI.
 - a. Determine whether the Java Web Console service is restricted.

```
# svcprop /system/webconsole:console | grep tcp_listen
```

 If the value of the `tcp_listen` property is not true, the web console service is restricted.

b. Restore external access to the Java Web Console service.

```
# svccfg
svc:> select system/webconsole
svc:/system/webconsole> setprop options/tcp_listen=true
svc:/system/webconsole> quit
# /usr/sbin/smcwebserver restart
```

c. Verify that the service is available.

```
# netstat -a | grep 6789
```

If the service is available, the command output returns an entry for 6789, which is the port number used to connect to Java Web Console.

Troubleshooting If after performing this procedure you cannot connect to Oracle Solaris Cluster Manager, determine if the Sun Java Web Console is running by entering `/usr/sbin/smcwebserver status`. If the Sun Java Web Console is not running, manually start it by entering `/usr/sbin/smcwebserver start`. If you still cannot connect to Oracle Solaris Cluster Manager, determine if the common agent container is running by entering `usr/sbin/cacaoadm status`. If the common agent container is not running, manually start it by entering `/usr/sbin/cacaoadm start`.

Example

Configuring Host-Based Data Replication With Sun StorageTek Availability Suite Software

This appendix provides an alternative to host-based replication that does not use Oracle Solaris Cluster Cluster Geographic Edition. Oracle recommends that you use Oracle Solaris Cluster Geographic Edition for host-based replication to simplify the configuration and operation of host-based replication within a cluster. See [“Understanding Data Replication” on page 84](#).

The example in this appendix shows how to configure host-based data replication between clusters using Sun StorageTek Availability Suite 4.0 software. The example illustrates a complete cluster configuration for an NFS application that provides detailed information about how individual tasks can be performed. All tasks should be performed in the global-cluster voting node. The example does not include all of the steps that are required by other applications or other cluster configurations.

If you use role-based access control (RBAC) instead of superuser to access the cluster nodes, ensure that you can assume an RBAC role that provides authorization for all Oracle Solaris Cluster commands. This series of data replication procedures requires the following Oracle Solaris Cluster RBAC authorizations if the user is not superuser:

- `solaris.cluster.modify`
- `solaris.cluster.admin`
- `solaris.cluster.read`

See the *System Administration Guide: Security Services* for more information about using RBAC roles. See the Oracle Solaris Cluster man pages for the RBAC authorization that each Oracle Solaris Cluster subcommand requires.

Understanding Sun StorageTek Availability Suite Software in a Cluster

This section introduces disaster tolerance and describes the data replication methods that Sun StorageTek Availability Suite software uses.

Disaster tolerance is the ability of a system to restore an application on an alternate cluster when the primary cluster fails. Disaster tolerance is based on *data replication* and *failover*. Failover is the automatic relocation of a resource group or device group from a primary cluster to a secondary cluster. If the primary cluster fails, the application and the data are immediately available on the secondary cluster.

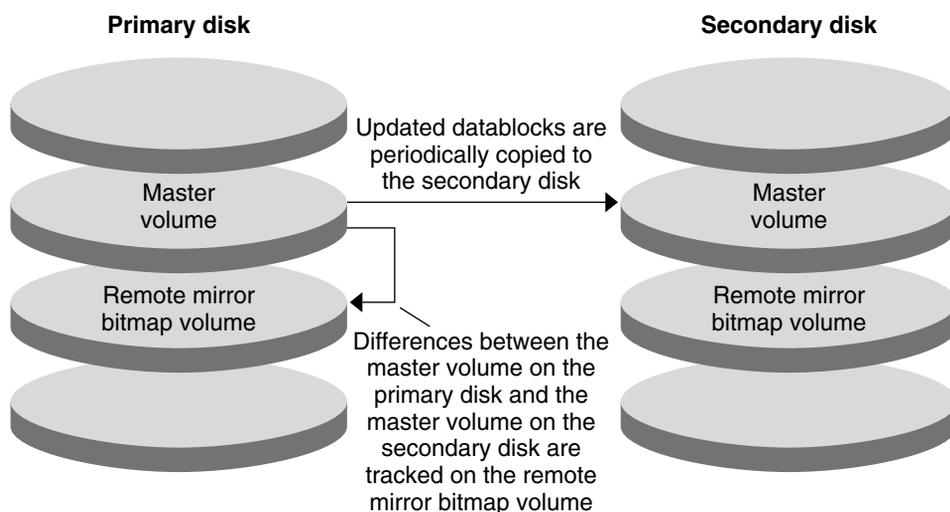
Data Replication Methods Used by Sun StorageTek Availability Suite Software

This section describes the remote mirror replication method and the point-in-time snapshot method used by Sun StorageTek Availability Suite software. This software uses the `sndradm(1RPC)` and `iiadm(1II)` commands to replicate data.

Remote Mirror Replication

[Figure A-1](#) shows remote mirror replication. Data from the master volume of the primary disk is replicated to the master volume of the secondary disk through a TCP/IP connection. A remote mirror bitmap tracks differences between the master volume on the primary disk and the master volume on the secondary disk.

FIGURE A-1 Remote Mirror Replication



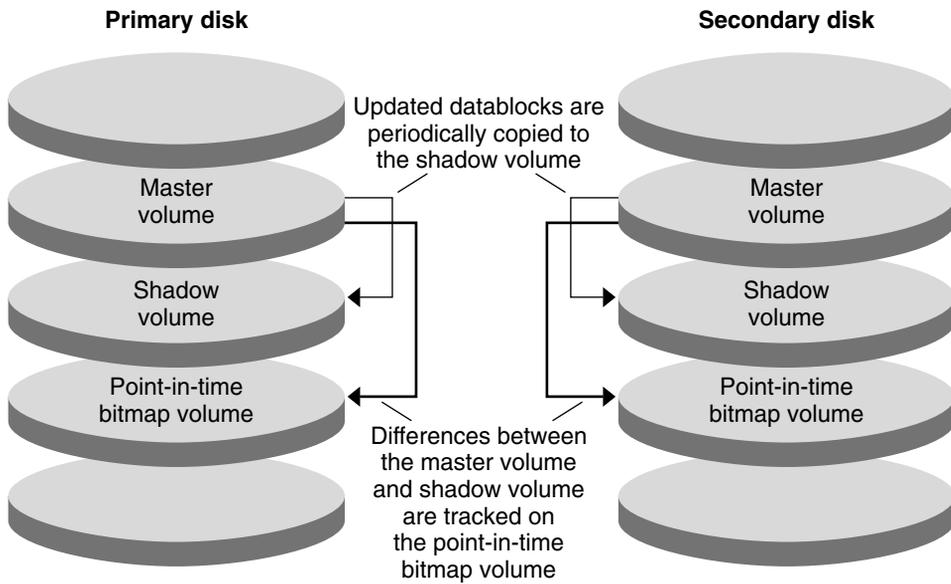
Remote mirror replication can be performed synchronously in real time, or asynchronously. Each volume set in each cluster can be configured individually, for synchronous replication or asynchronous replication.

- In synchronous data replication, a write operation is not confirmed as complete until the remote volume has been updated.
- In asynchronous data replication, a write operation is confirmed as complete before the remote volume is updated. Asynchronous data replication provides greater flexibility over long distances and low bandwidth.

Point-in-Time Snapshot

Figure A-2 shows point-in-time snapshot. Data from the master volume of each disk is copied to the shadow volume on the same disk. The point-in-time bitmap tracks differences between the master volume and the shadow volume. When data is copied to the shadow volume, the point-in-time bitmap is reset.

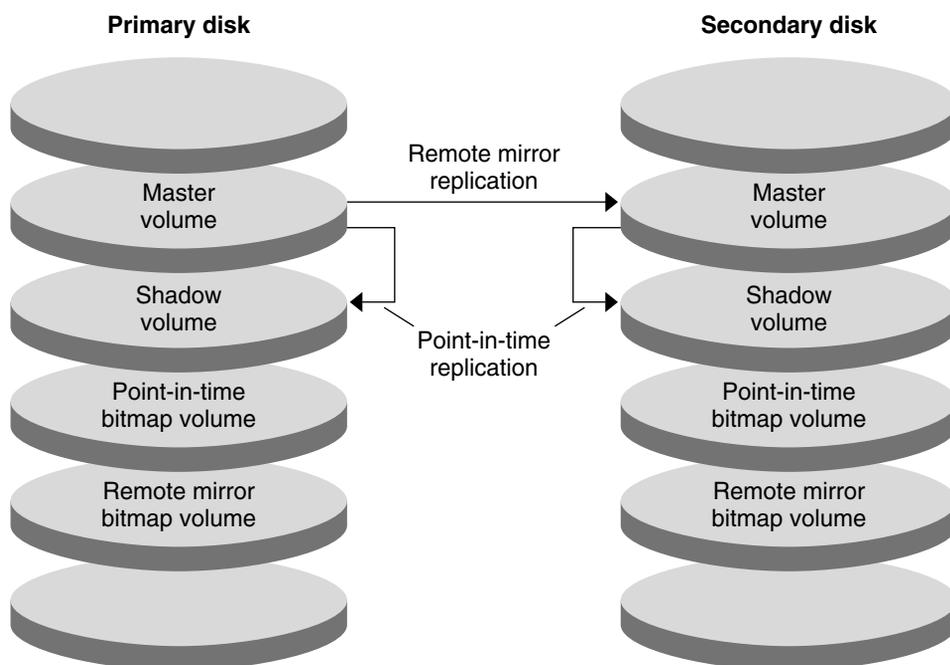
FIGURE A-2 Point-in-Time Snapshot



Replication in the Example Configuration

Figure A-3 illustrates how remote mirror replication and point-in-time snapshot are used in this example configuration.

FIGURE A-3 Replication in the Example Configuration



Guidelines for Configuring Host-Based Data Replication Between Clusters

This section provides guidelines for configuring data replication between clusters. This section also contains tips for configuring replication resource groups and application resource groups. Use these guidelines when you are configuring data replication for your cluster.

This section discusses the following topics:

- “Configuring Replication Resource Groups” on page 351
- “Configuring Application Resource Groups” on page 352
 - “Configuring Resource Groups for a Failover Application” on page 353
 - “Configuring Resource Groups for a Scalable Application” on page 354
- “Guidelines for Managing a Failover” on page 355

Configuring Replication Resource Groups

Replication resource groups collocate the device group under Sun StorageTek Availability Suite software control with the logical hostname resource. A replication resource group must have the following characteristics:

- Be a failover resource group

A failover resource can run on only one node at a time. When a failover occurs, failover resources take part in the failover.

- Have a logical hostname resource

The logical hostname must be hosted by the primary cluster. After a failover, the logical hostname must be hosted by the secondary cluster. The Domain Name System (DNS) is used to associate the logical hostname with a cluster.

- Have an HAStoragePlus resource

The HAStoragePlus resource enforces the failover of the device group when the replication resource group is switched over or failed over. Oracle Solaris Cluster software also enforces the failover of the replication resource group when the device group is switched over. In this way, the replication resource group and the device group are always colocated, or mastered by the same node.

The following extension properties must be defined in the HAStoragePlus resource:

- *GlobalDevicePaths*. This extension property defines the device group to which a volume belongs.
- *AffinityOn property = True*. This extension property causes the device group to switch over or fail over when the replication resource group switches over or fails over. This feature is called an *affinity switchover*.
- *ZPoolsSearchDir*. This extension property is required for using ZFS file system.

For more information about HAStoragePlus, see the [SUNW.HAStoragePlus\(5\)](#) man page.

- Be named after the device group with which it is colocated, followed by `-stor-rg`
For example, `devgrp-stor-rg`.
- Be online on both the primary cluster and the secondary cluster

Configuring Application Resource Groups

To be highly available, an application must be managed as a resource in an application resource group. An application resource group can be configured for a failover application or a scalable application.

Application resources and application resource groups configured on the primary cluster must also be configured on the secondary cluster. Also, the data accessed by the application resource must be replicated to the secondary cluster.

This section provides guidelines for configuring the following application resource groups:

- “Configuring Resource Groups for a Failover Application” on page 353
- “Configuring Resource Groups for a Scalable Application” on page 354

Configuring Resource Groups for a Failover Application

In a failover application, an application runs on one node at a time. If that node fails, the application fails over to another node in the same cluster. A resource group for a failover application must have the following characteristics:

- Have an HAStoragePlus resource to enforce the failover of the device group when the application resource group is switched over or failed over

The device group is colocated with the replication resource group and the application resource group. Therefore, the failover of the application resource group enforces the failover of the device group and replication resource group. The application resource group, the replication resource group, and the device group are mastered by the same node.

Note, however, that a failover of the device group or the replication resource group does not cause a failover of the application resource group.

- If the application data is globally mounted, the presence of an HAStoragePlus resource in the application resource group is not required but is advised.
- If the application data is mounted locally, the presence of an HAStoragePlus resource in the application resource group is required.

Without an HAStoragePlus resource, the failover of the application resource group would not trigger the failover of the replication resource group and device group. After a failover, the application resource group, replication resource group, and device group would not be mastered by the same node.

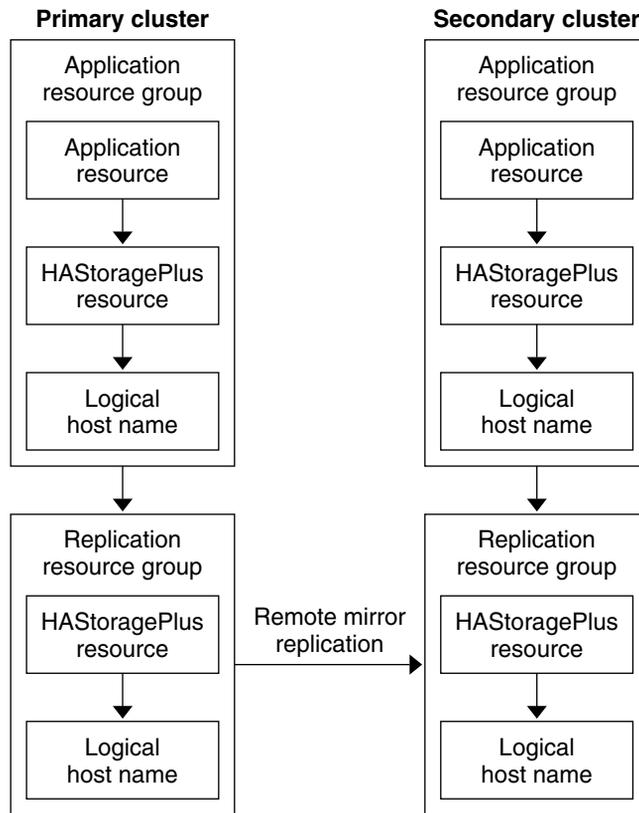
For more information about HAStoragePlus, see the [SUNW.HAStoragePlus\(5\)](#) man page.

- Must be online on the primary cluster and offline on the secondary cluster

The application resource group must be brought online on the secondary cluster when the secondary cluster takes over as the primary cluster.

[Figure A-4](#) illustrates the configuration of an application resource group and a replication resource group in a failover application.

FIGURE A-4 Configuration of Resource Groups in a Failover Application



Configuring Resource Groups for a Scalable Application

In a scalable application, an application runs on several nodes to create a single, logical service. If a node that is running a scalable application fails, failover does not occur. The application continues to run on the other nodes.

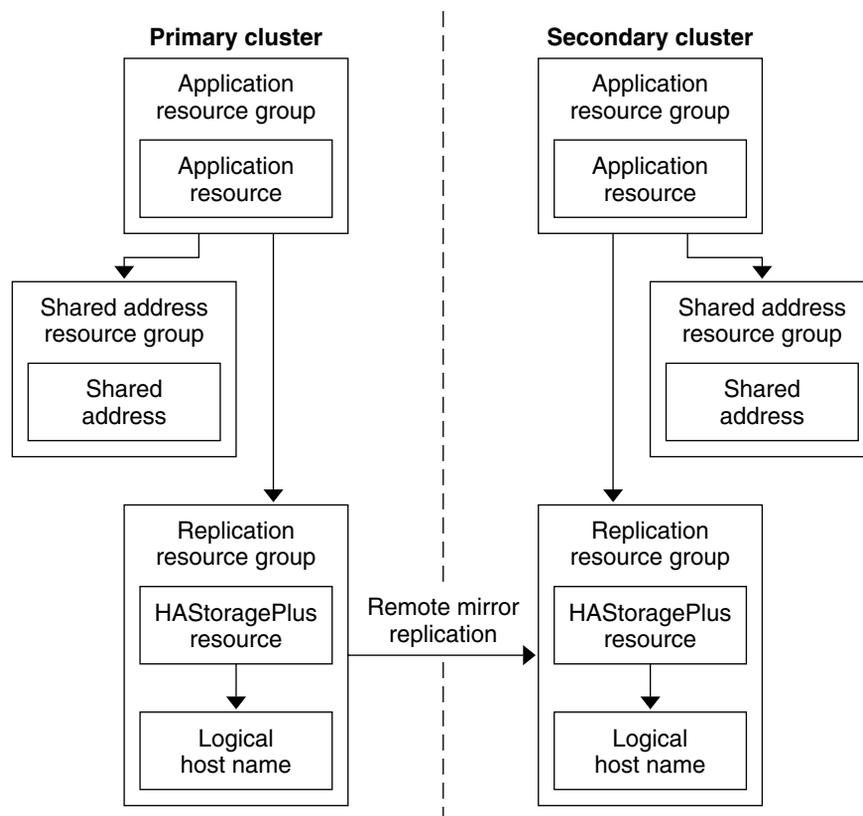
When a scalable application is managed as a resource in an application resource group, it is not necessary to collocate the application resource group with the device group. Therefore, it is not necessary to create an HAStoragePlus resource for the application resource group.

A resource group for a scalable application must have the following characteristics:

- Have a dependency on the shared address resource group
 - The nodes that are running the scalable application use the shared address to distribute incoming data.
- Be online on the primary cluster and offline on the secondary cluster

Figure A-5 illustrates the configuration of resource groups in a scalable application.

FIGURE A-5 Configuration of Resource Groups in a Scalable Application

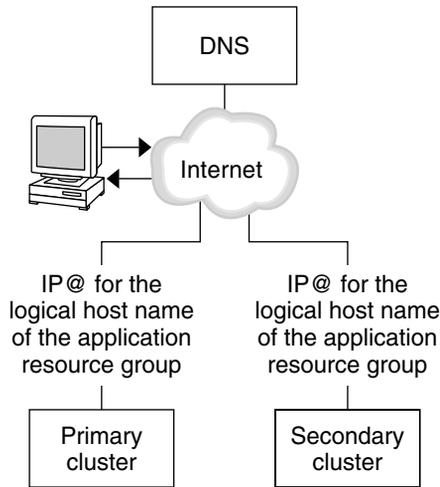


Guidelines for Managing a Failover

If the primary cluster fails, the application must be switched over to the secondary cluster as soon as possible. To enable the secondary cluster to take over, the DNS must be updated.

The DNS associates a client with the logical hostname of an application. After a failover, the DNS mapping to the primary cluster must be removed, and a DNS mapping to the secondary cluster must be created. [Figure A-6](#) shows how the DNS maps a client to a cluster.

FIGURE A-6 DNS Mapping of a Client to a Cluster



To update the DNS, use the `nsupdate` command. For information, see the `nsupdate(1M)` man page. For an example of how to manage a failover, see [“Example of How to Manage a Failover” on page 381](#).

After repair, the primary cluster can be brought back online. To switch back to the original primary cluster, perform the following tasks:

1. Synchronize the primary cluster with the secondary cluster to ensure that the primary volume is up-to-date.
2. Update the DNS so that clients can access the application on the primary cluster.

Task Map: Example of a Data Replication Configuration

[Table A-1](#) lists the tasks in this example of how data replication was configured for an NFS application by using Sun StorageTek Availability Suite software.

TABLE A-1 Task Map: Example of a Data Replication Configuration

Task	Instructions
1. Connect and install the clusters	“Connecting and Installing the Clusters” on page 357
2. Configure device groups, file systems for the NFS application, and resource groups on the primary cluster and on the secondary cluster	“Example of How to Configure Device Groups and Resource Groups” on page 359

TABLE A-1 Task Map: Example of a Data Replication Configuration (Continued)

Task	Instructions
3. Enable data replication on the primary cluster and on the secondary cluster	“How to Enable Replication on the Primary Cluster” on page 373 “How to Enable Replication on the Secondary Cluster” on page 375
4. Perform data replication	“How to Perform a Remote Mirror Replication” on page 376 “How to Perform a Point-in-Time Snapshot” on page 377
5. Verify the data replication configuration	“How to Verify That Replication Is Configured Correctly” on page 378

Connecting and Installing the Clusters

[Figure A-7](#) illustrates the cluster configuration the example configuration uses. The secondary cluster in the example configuration contains one node, but other cluster configurations can be used.

FIGURE A-7 Example Cluster Configuration

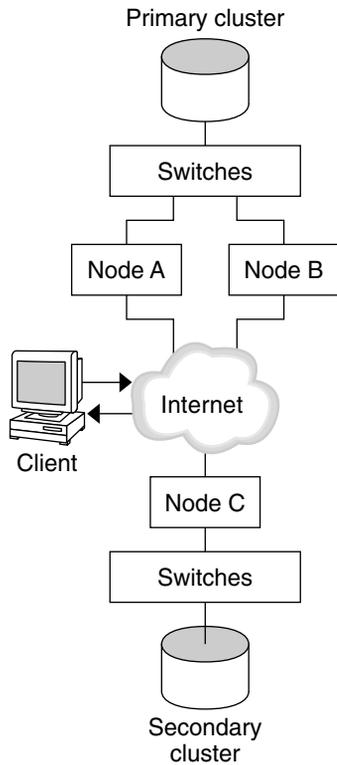


Table A-2 summarizes the hardware and software that the example configuration requires. The Oracle Solaris OS, Oracle Solaris Cluster software, and volume manager software must be installed on the cluster nodes *before* Sun StorageTek Availability Suite software and patches are installed.

TABLE A-2 Required Hardware and Software

Hardware or Software	Requirement
Node hardware	Sun StorageTek Availability Suite software is supported on all servers that use Oracle Solaris OS. For information about which hardware to use, see the <i>Oracle Solaris Cluster 3.3 Hardware Administration Manual</i> .
Disk space	Approximately 15 Mbytes.

TABLE A-2 Required Hardware and Software (Continued)

Hardware or Software	Requirement
Oracle Solaris OS	<p>Oracle Solaris OS releases that are supported by Oracle Solaris Cluster software.</p> <p>All nodes must use the same version of the Oracle Solaris OS.</p> <p>For information about installation, see the <i>Oracle Solaris Cluster Software Installation Guide</i>.</p>
Oracle Solaris Cluster software	<p>Oracle Solaris Cluster 3.3 software.</p> <p>For information about installation, see the <i>Oracle Solaris Cluster Software Installation Guide</i>.</p>
Volume manager software	<p>Solaris Volume Manager software or Veritas Volume Manager (VxVM) software.</p> <p>All nodes must use the same version of volume manager software.</p> <p>For information about installation see Chapter 4, “Configuring Solaris Volume Manager Software,” in <i>Oracle Solaris Cluster Software Installation Guide</i> and Chapter 5, “Installing and Configuring Veritas Volume Manager,” in <i>Oracle Solaris Cluster Software Installation Guide</i>.</p>
Sun StorageTek Availability Suite software	<p>For information about how to install the software, see the installation manuals for your release of Sun StorageTek Availability Suite software:</p> <ul style="list-style-type: none"> ▪ Sun StorageTek Availability Suite 4.0 – Sun StorageTek Availability documentation
Sun StorageTek Availability Suite software patches	<p>For information about the latest patches, see http://www.sunsolve.com.</p>

Example of How to Configure Device Groups and Resource Groups

This section describes how device groups and resource groups are configured for an NFS application. For additional information, see “Configuring Replication Resource Groups” on page 351 and “Configuring Application Resource Groups” on page 352.

This section contains the following procedures:

- “How to Configure a Device Group on the Primary Cluster” on page 361
- “How to Configure a Device Group on the Secondary Cluster” on page 362
- “How to Configure the File System on the Primary Cluster for the NFS Application” on page 363
- “How to Configure the File System on the Secondary Cluster for the NFS Application” on page 364
- “How to Create a Replication Resource Group on the Primary Cluster” on page 365

- [“How to Create a Replication Resource Group on the Secondary Cluster” on page 367](#)
- [“How to Create an NFS Application Resource Group on the Primary Cluster” on page 368](#)
- [“How to Create an NFS Application Resource Group on the Secondary Cluster” on page 370](#)
- [“How to Verify That Replication Is Configured Correctly” on page 378](#)

The following table lists the names of the groups and resources that are created for the example configuration.

TABLE A-3 Summary of the Groups and Resources in the Example Configuration

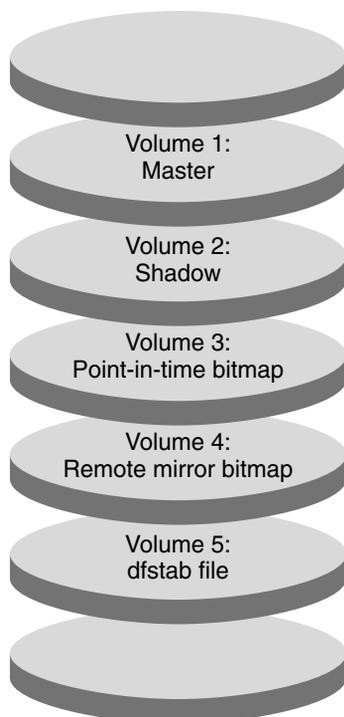
Group or Resource	Name	Description
Device group	devgrp	The device group
Replication resource group and resources	devgrp-stor-rg	The replication resource group
	lhost-reprg-prim, lhost-reprg-sec	The logical host names for the replication resource group on the primary cluster and the secondary cluster
	devgrp-stor	The HAStoragePlus resource for the replication resource group
Application resource group and resources	nfs-rg	The application resource group
	lhost-nfsrg-prim, lhost-nfsrg-sec	The logical host names for the application resource group on the primary cluster and the secondary cluster
	nfs-dg-rs	The HAStoragePlus resource for the application
	nfs-rs	The NFS resource

With the exception of `devgrp-stor-rg`, the names of the groups and resources are example names that can be changed as required. The replication resource group must have a name with the format *devicegroupname-stor-rg*.

This example configuration uses VxVM software. For information about Solaris Volume Manager software, see [Chapter 4, “Configuring Solaris Volume Manager Software,” in *Oracle Solaris Cluster Software Installation Guide*](#).

The following figure illustrates the volumes that are created in the device group.

FIGURE A-8 Volumes for the Device Group



Note – The volumes that are defined in this procedure must not include disk-label private areas, for example, cylinder 0. The VxVM software manages this constraint automatically.

▼ How to Configure a Device Group on the Primary Cluster

Before You Begin Ensure that you have completed the following tasks:

- Read the guidelines and requirements in the following sections:
 - “Understanding Sun StorageTek Availability Suite Software in a Cluster” on page 348
 - “Guidelines for Configuring Host-Based Data Replication Between Clusters” on page 351
- Set up the primary and secondary clusters as described in “Connecting and Installing the Clusters” on page 357.

1 Access nodeA as superuser or assume a role that provides `soLaris.cLuster.modify RBAC` authorization.

The node nodeA is the first node of the primary cluster. For a reminder of which node is nodeA, see [Figure A-7](#).

2 Create a disk group on nodeA that contains volume 1, vol01 through volume 4, vol04.

For information about configuring a disk group by using the VxVM software, see the [Chapter 5, “Installing and Configuring Veritas Volume Manager,”](#) in *Oracle Solaris Cluster Software Installation Guide*.

3 Configure the disk group to create a device group.

```
nodeA# cldevicegroup create -t vxvm -n nodeA nodeB devgrp
```

The device group is called devgrp.

4 Create the file system for the device group.

```
nodeA# newfs /dev/vx/rdisk/devgrp/vol01 < /dev/null
```

```
nodeA# newfs /dev/vx/rdisk/devgrp/vol02 < /dev/null
```

No file system is needed for vol03 or vol04, which are instead used as raw volumes.

Next Steps Go to “[How to Configure a Device Group on the Secondary Cluster](#)” on page 362.

▼ How to Configure a Device Group on the Secondary Cluster

Before You Begin Complete the procedure “[How to Configure a Device Group on the Primary Cluster](#)” on page 361.

1 Access nodeC as superuser or assume a role that provides solaris.cluster.modify RBAC authorization.**2 Create a disk group on nodeC that contains four volumes: volume 1, vol01, through volume 4, vol04.****3 Configure the disk group to create a device group.**

```
nodeC# cldevicegroup create -t vxvm -n nodeC devgrp
```

The device group is named devgrp.

4 Create the file system for the device group.

```
nodeC# newfs /dev/vx/rdisk/devgrp/vol01 < /dev/null
```

```
nodeC# newfs /dev/vx/rdisk/devgrp/vol02 < /dev/null
```

No file system is needed for vol03 or vol04, which are instead used as raw volumes.

Next Steps Go to “[How to Configure the File System on the Primary Cluster for the NFS Application](#)” on page 363.

▼ How to Configure the File System on the Primary Cluster for the NFS Application

Before You Begin Complete the procedure “[How to Configure a Device Group on the Secondary Cluster](#)” on page 362.

- 1 **On nodeA and nodeB, become superuser or assume a role that provides `solaris.cluster.admin` RBAC authorization.**

- 2 **On nodeA and nodeB, create a mount-point directory for the NFS file system.**

For example:

```
nodeA# mkdir /global/mountpoint
```

- 3 **On nodeA and nodeB, configure the master volume to be mounted automatically on the mount point.**

Add or replace the following text in the `/etc/vfstab` file on nodeA and nodeB. The text must be on a single line.

```
/dev/vx/dsk/devgrp/vol01 /dev/vx/rdisk/devgrp/vol01 \  
/global/mountpoint ufs 3 no global,logging
```

For a reminder of the volumes names and volume numbers that are used in the device group, see [Figure A-8](#).

- 4 **On nodeA, create a volume for the file system information that is used by the Oracle Solaris Cluster HA for NFS data service.**

```
nodeA# vxassist -g devgrp make vol05 120m disk1
```

Volume 5, `vol05`, contains the file system information that is used by the Oracle Solaris Cluster HA for NFS data service.

- 5 **On nodeA, resynchronize the device group with the Oracle Solaris Cluster software.**

```
nodeA# cldevicegroup sync devgrp
```

- 6 **On nodeA, create the file system for `vol05`.**

```
nodeA# newfs /dev/vx/rdisk/devgrp/vol05
```

- 7 **On nodeA and nodeB, create a mount point for `vol05`.**

The following example creates the mount point `/global/etc`.

```
nodeA# mkdir /global/etc
```

- 8 **On nodeA and nodeB, configure `vol05` to be mounted automatically on the mount point.**

Add or replace the following text in the `/etc/vfstab` file on nodeA and nodeB. The text must be on a single line.

```
/dev/vx/dsk/devgrp/vol05 /dev/vx/rdisk/devgrp/vol05 \  
/global/etc ufs 3 yes global,logging
```

9 Mount vol05 on nodeA.

```
nodeA# mount /global/etc
```

10 Make vol05 accessible to remote systems.**a. Create a directory called /global/etc/SUNW.nfs on nodeA.**

```
nodeA# mkdir -p /global/etc/SUNW.nfs
```

b. Create the file /global/etc/SUNW.nfs/dfstab.nfs-rs on nodeA.

```
nodeA# touch /global/etc/SUNW.nfs/dfstab.nfs-rs
```

c. Add the following line to the /global/etc/SUNW.nfs/dfstab.nfs-rs file on nodeA.

```
share -F nfs -o rw -d "HA NFS" /global/mountpoint
```

Next Steps Go to [“How to Configure the File System on the Secondary Cluster for the NFS Application” on page 364.](#)

▼ How to Configure the File System on the Secondary Cluster for the NFS Application

Before You Begin Complete the procedure [“How to Configure the File System on the Primary Cluster for the NFS Application” on page 363.](#)

1 On nodeC, become superuser or assume a role that provides solaris.cluster.admin RBAC authorization.**2 On nodeC, create a mount-point directory for the NFS file system.**

For example:

```
nodeC# mkdir /global/mountpoint
```

3 On nodeC, configure the master volume to be mounted automatically on the mount point.

Add or replace the following text in the `/etc/vfstab` file on nodeC. The text must be on a single line.

```
/dev/vx/dsk/devgrp/vol01 /dev/vx/rdisk/devgrp/vol01 \  
/global/mountpoint ufs 3 no global,logging
```

4 On nodeC, create a volume for the file system information that is used by the Oracle Solaris Cluster HA for NFS data service.

```
nodeC# vxassist -g devgrp make vol05 120m disk1
```

Volume 5, vol05, contains the file system information that is used by the Oracle Solaris Cluster HA for NFS data service.

- 5 On nodeC, resynchronize the device group with the Oracle Solaris Cluster software.

```
nodeC# cldevicegroup sync devgrp
```

- 6 On nodeC, create the file system for vol05.

```
nodeC# newfs /dev/vx/rdisk/devgrp/vol05
```

- 7 On nodeC, create a mount point for vol05.

The following example creates the mount point /global/etc.

```
nodeC# mkdir /global/etc
```

- 8 On nodeC, configure vol05 to be mounted automatically on the mount point.

Add or replace the following text in the /etc/vfstab file on nodeC. The text must be on a single line.

```
/dev/vx/dsk/devgrp/vol05 /dev/vx/rdsk/devgrp/vol05 \  
/global/etc ufs 3 yes global,logging
```

- 9 Mount vol05 on nodeC.

```
nodeC# mount /global/etc
```

- 10 Make vol05 accessible to remote systems.

- a. Create a directory called /global/etc/SUNW.nfs on nodeC.

```
nodeC# mkdir -p /global/etc/SUNW.nfs
```

- b. Create the file /global/etc/SUNW.nfs/dfstab.nfs-rs on nodeC.

```
nodeC# touch /global/etc/SUNW.nfs/dfstab.nfs-rs
```

- c. Add the following line to the /global/etc/SUNW.nfs/dfstab.nfs-rs file on nodeC:

```
share -F nfs -o rw -d "HA NFS" /global/mountpoint
```

Next Steps Go to [“How to Create a Replication Resource Group on the Primary Cluster”](#) on page 365.

▼ How to Create a Replication Resource Group on the Primary Cluster

Before You Begin Complete the procedure [“How to Configure the File System on the Secondary Cluster for the NFS Application”](#) on page 364.

- 1 Access nodeA as superuser or assume a role that provides `solaris.cluster.modify`, `solaris.cluster.admin`, and `solaris.cluster.read` RBAC authorization.

- 2 Register the SUNW.HAStoragePlus resource type.

```
nodeA# clresourcetype register SUNW.HAStoragePlus
```

3 Create a replication resource group for the device group.

```
nodeA# clresourcegroup create -n nodeA,nodeB devgrp-stor-rg
```

-n nodeA,nodeB Specifies that cluster nodes nodeA and nodeB can master the replication resource group.

devgrp-stor-rg The name of the replication resource group. In this name, devgrp specifies the name of the device group.

4 Add a SUNW.HASStoragePlus resource to the replication resource group.

```
nodeA# clresource create -g devgrp-stor-rg -t SUNW.HASStoragePlus \
-p GlobalDevicePaths=devgrp \
-p AffinityOn=True \
devgrp-stor
```

-g Specifies the resource group to which resource is added.

-p GlobalDevicePaths= Specifies the extension property that Sun StorageTek Availability Suite software relies on.

-p AffinityOn=True Specifies that the SUNW.HASStoragePlus resource must perform an affinity switchover for the global devices and cluster file systems defined by -x GlobalDevicePaths=. Therefore, when the replication resource group fails over or is switched over, the associated device group is switched over.

For more information about these extension properties, see the [SUNW.HASStoragePlus\(5\)](#) man page.

5 Add a logical hostname resource to the replication resource group.

```
nodeA# clreslogicalhostname create -g devgrp-stor-rg lhost-reprg-prim
```

The logical hostname for the replication resource group on the primary cluster is named lhost-reprg-prim.

6 Enable the resources, manage the resource group, and bring the resource group online.

```
nodeA# clresourcegroup online -e -M -n nodeA devgrp-stor-rg
```

-e Enables associated resources.

-M Manages the resource group.

-n Specifies the node on which to bring the resource group online.

7 Verify that the resource group is online.

```
nodeA# clresourcegroup status devgrp-stor-rg
```

Examine the resource group state field to confirm that the replication resource group is online on nodeA.

Next Steps Go to “How to Create a Replication Resource Group on the Secondary Cluster” on page 367.

▼ How to Create a Replication Resource Group on the Secondary Cluster

Before You Begin Complete the procedure “How to Create a Replication Resource Group on the Primary Cluster” on page 365.

1 Access nodeC as superuser or assume a role that provides `solaris.cluster.modify`, `solaris.cluster.admin`, and `solaris.cluster.read` RBAC authorization.

2 Register `SUNW.HASStoragePlus` as a resource type.

```
nodeC# clresourcetype register SUNW.HASStoragePlus
```

3 Create a replication resource group for the device group.

```
nodeC# clresourcegroup create -n nodeC devgrp-stor-rg
```

`create` Creates the resource group.

`-n` Specifies the node list for the resource group.

`devgrp` The name of the device group.

`devgrp-stor-rg` The name of the replication resource group.

4 Add a `SUNW.HASStoragePlus` resource to the replication resource group.

```
nodeC# clresource create \
-t SUNW.HASStoragePlus \
-p GlobalDevicePaths=devgrp \
-p AffinityOn=True \
devgrp-stor
```

`create` Creates the resource.

`-t` Specifies the resource type.

-p `GlobalDevicePaths=` Specifies the extension property that Sun StorageTek Availability Suite software relies on.

-p `AffinityOn=True` Specifies that the `SUNW.HASStoragePlus` resource must perform an affinity switchover for the global devices and cluster file systems defined by `-x GlobalDevicePaths=`. Therefore, when the replication resource group fails over or is switched over, the associated device group is switched over.

`devgrp-stor` The `HASStoragePlus` resource for the replication resource group.

For more information about these extension properties, see the [SUNW.HASStoragePlus\(5\)](#) man page.

5 Add a logical hostname resource to the replication resource group.

```
nodeC# clreslogicalhostname create -g devgrp-stor-rg lhost-reprg-sec
```

The logical hostname for the replication resource group on the primary cluster is named `lhost-reprg-sec`.

6 Enable the resources, manage the resource group, and bring the resource group online.

```
nodeC# clresourcegroup online -e -M -n nodeC devgrp-stor-rg
```

`online` Brings online.

`-e` Enables associated resources.

`-M` Manages the resource group.

`-n` Specifies the node on which to bring the resource group online.

7 Verify that the resource group is online.

```
nodeC# clresourcegroup status devgrp-stor-rg
```

Examine the resource group state field to confirm that the replication resource group is online on `nodeC`.

Next Steps Go to [“How to Create an NFS Application Resource Group on the Primary Cluster”](#) on page 368.

▼ How to Create an NFS Application Resource Group on the Primary Cluster

This procedure describes how application resource groups are created for NFS. This procedure is specific to this application and cannot be used for another type of application.

Before You Begin Complete the procedure [“How to Create a Replication Resource Group on the Secondary Cluster”](#) on page 367.

1 Access nodeA as superuser or assume a role that provides `solaris.cluster.modify`, `solaris.cluster.admin`, and `solaris.cluster.read` RBAC authorization.**2 Register `SUNW.nfs` as a resource type.**

```
nodeA# clresourcetype register SUNW.nfs
```

3 If `SUNW.HASStoragePlus` has not been registered as a resource type, register it.

```
nodeA# clresourcetype register SUNW.HASStoragePlus
```

4 Create an application resource group for the device group devgrp.

```
nodeA# clresourcegroup create \
-p Pathprefix=/global/etc \
-p Auto_start_on_new_cluster=False \
-p RG_dependencies=devgrp-stor-rg \
nfs-rg
```

Pathprefix=/global/etc

Specifies the directory into which the resources in the group can write administrative files.

Auto_start_on_new_cluster=False

Specifies that the application resource group is not started automatically.

RG_dependencies=devgrp-stor-rg

Specifies the resource group that the application resource group depends on. In this example, the application resource group depends on the replication resource group devgrp-stor-rg.

If the application resource group is switched over to a new primary node, the replication resource group is automatically switched over. However, if the replication resource group is switched over to a new primary node, the application resource group must be manually switched over.

nfs-rg

The name of the application resource group.

5 Add a SUNW.HASStoragePlus resource to the application resource group.

```
nodeA# clresource create -g nfs-rg \
-t SUNW.HASStoragePlus \
-p FileSystemMountPoints=/global/mountpoint \
-p AffinityOn=True \
nfs-dg-rs
```

create

Creates the resource.

-g

Specifies the resource group to which the resource is added.

-t SUNW.HASStoragePlus

Specifies that the resource is of the type SUNW.HASStoragePlus.

-p FileSystemMountPoints=/global/

Specifies that the mount point for the file system is global.

-p AffinityOn=True

Specifies that the application resource must perform an affinity switchover for the global devices and cluster file systems defined by -p GlobalDevicePaths=. Therefore, when the application resource group fails over or is switched over, the associated device group is switched over.

```
nfs - dg - rs
```

The name of the HAStoragePlus resource for the NFS application.

For more information about these extension properties, see the [SUNW.HAStoragePlus\(5\)](#) man page.

6 Add a logical hostname resource to the application resource group.

```
nodeA# clreslogicalhostname create -g nfs-rg \
lhost-nfsrg-prim
```

The logical hostname of the application resource group on the primary cluster is named `lhost-nfsrg-prim`.

7 Enable the resources, manage the application resource group, and bring the application resource group online.

a. Enable the HAStoragePlus resource for the NFS application.

```
nodeA# clresource enable nfs-rs
```

b. Bring the application resource group online on nodeA .

```
nodeA# clresourcegroup online -e -M -n nodeA nfs-rg
```

`online` Brings the resource group online.

`-e` Enables the associated resources.

`-M` Manages the resource group.

`-n` Specifies the node on which to bring the resource group online.

`nfs - rg` The name of the resource group.

8 Verify that the application resource group is online.

```
nodeA# clresourcegroup status
```

Examine the resource group state field to determine whether the application resource group is online for nodeA and nodeB.

Next Steps Go to [“How to Create an NFS Application Resource Group on the Secondary Cluster”](#) on page 370.

▼ How to Create an NFS Application Resource Group on the Secondary Cluster

Before You Begin Complete the procedure [“How to Create an NFS Application Resource Group on the Primary Cluster”](#) on page 368.

- 1 **Access nodeC as superuser or assume a role that provides `solaris.cluster.modify`, `solaris.cluster.admin`, and `solaris.cluster.read` RBAC authorization.**

- 2 **Register `SUNW.nfs` as a resource type.**

```
nodeC# clresourcetype register SUNW.nfs
```

- 3 **If `SUNW.HASStoragePlus` has not been registered as a resource type, register it.**

```
nodeC# clresourcetype register SUNW.HASStoragePlus
```

- 4 **Create an application resource group for the device group.**

```
nodeC# clresourcegroup create \
-p Pathprefix=/global/etc \
-p Auto_start_on_new_cluster=False \
-p RG_dependencies=devgrp-stor-rg \
nfs-rg
```

```
create
```

Creates the resource group.

```
-p
```

Specifies a property of the resource group.

```
Pathprefix=/global/etc
```

Specifies a directory into which the resources in the group can write administrative files.

```
Auto_start_on_new_cluster=False
```

Specifies that the application resource group is not started automatically.

```
RG_dependencies=devgrp-stor-rg
```

Specifies the resource groups that the application resource group depends on. In this example, the application resource group depends on the replication resource group.

If the application resource group is switched over to a new primary node, the replication resource group is automatically switched over. However, if the replication resource group is switched over to a new primary node, the application resource group must be manually switched over.

```
nfs-rg
```

The name of the application resource group.

- 5 **Add a `SUNW.HASStoragePlus` resource to the application resource group.**

```
nodeC# clresource create -g nfs-rg \
-t SUNW.HASStoragePlus \
-p FileSystemMountPoints=/global/mountpoint \
-p AffinityOn=True \
nfs-dg-rs
```

```
create
```

Creates the resource.

- g
Specifies the resource group to which the resource is added.
 - t SUNW.HASStoragePlus
Specifies that the resource is of the type SUNW.HASStoragePlus.
 - p
Specifies a property of the resource.
- FileSystemMountPoints=/global/
Specifies that the mount point for the file system is global.

AffinityOn=True
Specifies that the application resource must perform an affinity switchover for the global devices and cluster file systems defined by -x GlobalDevicePaths=. Therefore, when the application resource group fails over or is switched over, the associated device group is switched over.

nfs-dg-rs
The name of the HASStoragePlus resource for the NFS application.

For more information about these extension properties, see the [SUNW.HASStoragePlus\(5\)](#) man page.

6 Add a logical hostname resource to the application resource group.

```
nodeC# clreslogicalhostname create -g nfs-rg \  
lhost-nfsrg-sec
```

The logical hostname of the application resource group on the secondary cluster is named lhost-nfsrg-sec.

7 Add an NFS resource to the application resource group.

```
nodeC# clresource create -g nfs-rg \  
-t SUNW.nfs -p Resource_dependencies=nfs-dg-rs nfs-rg
```

8 Ensure that the application resource group does not come online on nodeC.

```
nodeC# clresource disable -n nodeC nfs-rs  
nodeC# clresource disable -n nodeC nfs-dg-rs  
nodeC# clresource disable -n nodeC lhost-nfsrg-sec  
nodeC# clresourcegroup online -n "" nfs-rg
```

The resource group remains offline after a reboot, because *Auto_start_on_new_cluster=False*.

9 If the global volume is mounted on the primary cluster, unmount the global volume from the secondary cluster.

```
nodeC# umount /global/mountpoint
```

If the volume is mounted on a secondary cluster, the synchronization fails.

Next Steps Go to [“Example of How to Enable Data Replication”](#) on page 373.

Example of How to Enable Data Replication

This section describes how data replication is enabled for the example configuration. This section uses the Sun StorageTek Availability Suite software commands `sndradm` and `iiadm`. For more information about these commands, see the Sun StorageTek Availability documentation.

This section contains the following procedures:

- “How to Enable Replication on the Primary Cluster” on page 373
- “How to Enable Replication on the Secondary Cluster” on page 375

▼ How to Enable Replication on the Primary Cluster

- 1 Access `nodeA` as superuser or assume a role that provides `solaris.cluster.read` RBAC authorization.

- 2 Flush all transactions.

```
nodeA# lockfs -a -f
```

- 3 Confirm that the logical host names `lhost-reprg-prim` and `lhost-reprg-sec` are online.

```
nodeA# clresourcegroup status
nodeC# clresourcegroup status
```

Examine the state field of the resource group.

- 4 Enable remote mirror replication from the primary cluster to the secondary cluster.

This step enables replication from the master volume on the primary cluster to the master volume on the secondary cluster. In addition, this step enables replication to the remote mirror bitmap on `vol04`.

- If the primary cluster and secondary cluster are unsynchronized, run this command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -e lhost-reprg-prim \
/dev/vx/rdisk/devgrp/vol01 \
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdisk/devgrp/vol01 \
/dev/vx/rdisk/devgrp/vol04 ip sync
```

- If the primary cluster and secondary cluster are synchronized, run this command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -E lhost-reprg-prim \
/dev/vx/rdisk/devgrp/vol01 \
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdisk/devgrp/vol01 \
/dev/vx/rdisk/devgrp/vol04 ip sync
```

5 Enable autosynchronization.

Run this command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -a on lhost-reprg-prim \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 ip sync
```

This step enables autosynchronization. When the active state of autosynchronization is set to on, the volume sets are resynchronized if the system reboots or a failure occurs.

6 Verify that the cluster is in logging mode.

Use the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -P
```

The output should resemble the following:

```
/dev/vx/rdisk/devgrp/vol01 ->  
lhost-reprg-sec:/dev/vx/rdisk/devgrp/vol01  
autosync: off, max q writes:4194304, max q fbas:16384, mode:sync,ctag:  
devgrp, state: logging
```

In logging mode, the state is logging, and the active state of autosynchronization is off. When the data volume on the disk is written to, the bitmap file on the same disk is updated.

7 Enable point-in-time snapshot.

Use the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/iadm -e ind \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol02 \  
/dev/vx/rdisk/devgrp/vol03  
nodeA# /usr/sbin/iadm -w \  
/dev/vx/rdisk/devgrp/vol02
```

This step enables the master volume on the primary cluster to be copied to the shadow volume on the same cluster. The master volume, shadow volume, and point-in-time bitmap volume must be in the same device group. In this example, the master volume is vol01, the shadow volume is vol02, and the point-in-time bitmap volume is vol03.

8 Attach the point-in-time snapshot to the remote mirror set.

Use the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -I a \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol02 \  
/dev/vx/rdisk/devgrp/vol03
```

This step associates the point-in-time snapshot with the remote mirror volume set. Sun StorageTek Availability Suite software ensures that a point-in-time snapshot is taken before remote mirror replication can occur.

Next Steps Go to [“How to Enable Replication on the Secondary Cluster”](#) on page 375.

▼ **How to Enable Replication on the Secondary Cluster**

Before You Begin Complete the procedure [“How to Enable Replication on the Primary Cluster”](#) on page 373.

1 Access nodeC as superuser.

2 Flush all transactions.

```
nodeC# lockfs -a -f
```

3 Enable remote mirror replication from the primary cluster to the secondary cluster.

Use the following command for Sun StorageTek Availability Suite software:

```
nodeC# /usr/sbin/sndradm -n -e lhost-reprg-prim \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 ip sync
```

The primary cluster detects the presence of the secondary cluster and starts synchronization. Refer to the system log file `/var/adm` for Sun StorageTek Availability Suite for information about the status of the clusters.

4 Enable independent point-in-time snapshot.

Use the following command for Sun StorageTek Availability Suite software:

```
nodeC# /usr/sbin/iiadm -e ind \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol02 \  
/dev/vx/rdisk/devgrp/vol03  
nodeC# /usr/sbin/iiadm -w \  
/dev/vx/rdisk/devgrp/vol02
```

5 Attach the point-in-time snapshot to the remote mirror set.

Use the following command for Sun StorageTek Availability Suite software:

```
nodeC# /usr/sbin/sndradm -I a \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol02 \  
/dev/vx/rdisk/devgrp/vol03
```

Next Steps Go to [“Example of How to Perform Data Replication”](#) on page 376.

Example of How to Perform Data Replication

This section describes how data replication is performed for the example configuration. This section uses the Sun StorageTek Availability Suite software commands `sndradm` and `iiadm`. For more information about these commands, see the Sun StorageTek Availability Suite documentation.

This section contains the following procedures:

- “How to Perform a Remote Mirror Replication” on page 376
- “How to Perform a Point-in-Time Snapshot” on page 377
- “How to Verify That Replication Is Configured Correctly” on page 378

▼ How to Perform a Remote Mirror Replication

In this procedure, the master volume of the primary disk is replicated to the master volume on the secondary disk. The master volume is `vol01` and the remote mirror bitmap volume is `vol04`.

1 Access nodeA as superuser.

2 Verify that the cluster is in logging mode.

Run the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -P
```

The output should resemble the following:

```
/dev/vx/rdisk/devgrp/vol01 ->
lhost-reprg-sec:/dev/vx/rdisk/devgrp/vol01
autosync: off, max q writes:4194304, max q fbas:16384, mode:sync,ctag:
devgrp, state: logging
```

In logging mode, the state is `logging`, and the active state of autosynchronization is `off`. When the data volume on the disk is written to, the bitmap file on the same disk is updated.

3 Flush all transactions.

```
nodeA# lockfs -a -f
```

4 Repeat Step 1 through Step 3 on nodeC.

5 Copy the master volume of nodeA to the master volume of nodeC.

Run the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -m lhost-reprg-prim \
/dev/vx/rdisk/devgrp/vol01 \
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdisk/devgrp/vol01 \
/dev/vx/rdisk/devgrp/vol04 ip sync
```

6 Wait until the replication is complete and the volumes are synchronized.

Run the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -w lhost-reprg-prim \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 ip sync
```

7 Confirm that the cluster is in replicating mode.

Run the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -P
```

The output should resemble the following:

```
/dev/vx/rdisk/devgrp/vol01 ->  
lhost-reprg-sec:/dev/vx/rdisk/devgrp/vol01  
autosync: on, max q writes:4194304, max q fbas:16384, mode:sync,ctag:  
devgrp, state: replicating
```

In replicating mode, the state is `replicating`, and the active state of autosynchronization is on. When the primary volume is written to, the secondary volume is updated by Sun StorageTek Availability Suite software.

Next Steps Go to [“How to Perform a Point-in-Time Snapshot” on page 377](#).

▼ How to Perform a Point-in-Time Snapshot

In this procedure, point-in-time snapshot is used to synchronize the shadow volume of the primary cluster to the master volume of the primary cluster. The master volume is `vol01`, the bitmap volume is `vol04`, and the shadow volume is `vol02`.

Before You Begin Complete the procedure [“How to Perform a Remote Mirror Replication” on page 376](#).

1 Access nodeA as superuser or assume a role that provides `solaris.cluster.modify` and `solaris.cluster.admin` RBAC authorization.**2 Disable the resource that is running on nodeA.**

```
nodeA# clresource disable -n nodeA nfs-rs
```

3 Change the primary cluster to logging mode.

Run the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -l lhost-reprg-prim \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 ip sync
```

When the data volume on the disk is written to, the bitmap file on the same disk is updated. No replication occurs.

4 Synchronize the shadow volume of the primary cluster to the master volume of the primary cluster.

Run the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/iiadm -u s /dev/vx/rdisk/devgrp/vol02
nodeA# /usr/sbin/iiadm -w /dev/vx/rdisk/devgrp/vol02
```

5 Synchronize the shadow volume of the secondary cluster to the master volume of the secondary cluster.

Run the following command for Sun StorageTek Availability Suite software:

```
nodeC# /usr/sbin/iiadm -u s /dev/vx/rdisk/devgrp/vol02
nodeC# /usr/sbin/iiadm -w /dev/vx/rdisk/devgrp/vol02
```

6 Restart the application on nodeA.

```
nodeA# clresource enable -n nodeA nfs-rs
```

7 Resynchronize the secondary volume with the primary volume.

Run the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -u lhost-reprg-prim \
/dev/vx/rdisk/devgrp/vol01 \
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \
/dev/vx/rdisk/devgrp/vol01 \
/dev/vx/rdisk/devgrp/vol04 ip sync
```

Next Steps Go to [“How to Verify That Replication Is Configured Correctly”](#) on page 378.

▼ How to Verify That Replication Is Configured Correctly

Before You Begin Complete the procedure [“How to Perform a Point-in-Time Snapshot”](#) on page 377.

- 1 Access nodeA and nodeC as superuser or assume a role that provides `solaris.cluster.admin` RBAC authorization.**
- 2 Verify that the primary cluster is in replicating mode, with autosynchronization on.**

Use the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -P
```

The output should resemble the following:

```
/dev/vx/rdisk/devgrp/vol01 ->
lhost-reprg-sec:/dev/vx/rdisk/devgrp/vol01
autosync: on, max q writes:4194304, max q fbas:16384, mode:sync,ctag:
devgrp, state: replicating
```

In replicating mode, the state is `replicating`, and the active state of autosynchronization is on. When the primary volume is written to, the secondary volume is updated by Sun StorageTek Availability Suite software.

3 If the primary cluster is not in replicating mode, put it into replicating mode.

Use the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -u lhost-reprg-prim \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 ip sync
```

4 Create a directory on a client machine.

a. Log in to a client machine as superuser.

You see a prompt that resembles the following:

```
client-machine#
```

b. Create a directory on the client machine.

```
client-machine# mkdir /dir
```

5 Mount the directory to the application on the primary cluster, and display the mounted directory.

a. Mount the directory to the application on the primary cluster.

```
client-machine# mount -o rw lhost-nfsrg-prim:/global/mountpoint /dir
```

b. Display the mounted directory.

```
client-machine# ls /dir
```

6 Mount the directory to the application on the secondary cluster, and display the mounted directory.

a. Unmount the directory from the application on the primary cluster.

```
client-machine# umount /dir
```

b. Take the application resource group offline on the primary cluster.

```
nodeA# clresource disable -n nodeA nfs-rs  
nodeA# clresource disable -n nodeA nfs-dg-rs  
nodeA# clresource disable -n nodeA lhost-nfsrg-prim  
nodeA# clresourcegroup online -n "" nfs-rg
```

c. Change the primary cluster to logging mode.

Run the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -l lhost-reprg-prim \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 ip sync
```

```
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 ip sync
```

When the data volume on the disk is written to, the bitmap file on the same disk is updated. No replication occurs.

d. Ensure that the PathPrefix directory is available.

```
nodeC# mount | grep /global/etc
```

e. Bring the application resource group online on the secondary cluster.

```
nodeC# clresourcegroup online -n nodeC nfs-rg
```

f. Access the client machine as superuser.

You see a prompt that resembles the following:

```
client-machine#
```

g. Mount the directory that was created in [Step 4](#) to the application on the secondary cluster.

```
client-machine# mount -o rw lhost-nfsrg-sec:/global/mountpoint /dir
```

h. Display the mounted directory.

```
client-machine# ls /dir
```

7 Ensure that the directory displayed in [Step 5](#) is the same as the directory displayed in [Step 6](#).

8 Return the application on the primary cluster to the mounted directory.

a. Take the application resource group offline on the secondary cluster.

```
nodeC# clresource disable -n nodeC nfs-rs  
nodeC# clresource disable -n nodeC nfs-dg-rs  
nodeC# clresource disable -n nodeC lhost-nfsrg-sec  
nodeC# clresourcegroup online -n "" nfs-rg
```

b. Ensure that the global volume is unmounted from the secondary cluster.

```
nodeC# umount /global/mountpoint
```

c. Bring the application resource group online on the primary cluster.

```
nodeA# clresourcegroup online -n nodeA nfs-rg
```

d. Change the primary cluster to replicating mode.

Run the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -u lhost-reprg-prim \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 ip sync
```

When the primary volume is written to, the secondary volume is updated by Sun StorageTek Availability Suite software.

See Also [“Example of How to Manage a Failover” on page 381](#)

Example of How to Manage a Failover

This section describes how to provoke a failover and how the application is transferred to the secondary cluster. After a failover, update the DNS entries. For additional information, see [“Guidelines for Managing a Failover” on page 355](#).

This section contains the following procedures:

- [“How to Provoke a Switchover” on page 381](#)
- [“How to Update the DNS Entry” on page 382](#)

▼ How to Provoke a Switchover

- 1 **Access nodeA and nodeC as superuser or assume a role that provides `solaris.cluster.admin` RBAC authorization.**

- 2 **Change the primary cluster to logging mode.**

Run the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -l lhost-reprg-prim \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 lhost-reprg-sec \  
/dev/vx/rdisk/devgrp/vol01 \  
/dev/vx/rdisk/devgrp/vol04 ip sync
```

When the data volume on the disk is written to, the bitmap volume on the same device group is updated. No replication occurs.

- 3 **Confirm that the primary cluster and the secondary cluster are in logging mode, with autosynchronization off.**

- a. **On nodeA, confirm the mode and setting:**

Run the following command for Sun StorageTek Availability Suite software:

```
nodeA# /usr/sbin/sndradm -P
```

The output should resemble the following:

```
/dev/vx/rdisk/devgrp/vol01 ->  
lhost-reprg-sec:/dev/vx/rdisk/devgrp/vol01  
autosync:off, max q writes:4194304,max q fbas:16384,mode:sync,ctag:  
devgrp, state: logging
```

b. On nodeC, confirm the mode and setting:

Run the following command for Sun StorageTek Availability Suite software:

```
nodeC# /usr/sbin/sndradm -P
```

The output should resemble the following:

```
/dev/vx/rdisk/devgrp/vol01 <-
lhost-reprg-prim:/dev/vx/rdisk/devgrp/vol01
autosync:off, max q writes:4194304,max q fbas:16384,mode:sync,ctag:
devgrp, state: logging
```

For nodeA and nodeC, the state should be logging, and the active state of autosynchronization should be off.

4 Confirm that the secondary cluster is ready to take over from the primary cluster.

```
nodeC# fsck -y /dev/vx/rdisk/devgrp/vol01
```

5 Switch over to the secondary cluster.

```
nodeC# clresourcegroup switch -n nodeC nfs-rg
```

Next Steps Go to “[How to Update the DNS Entry](#)” on page 382.

▼ How to Update the DNS Entry

For an illustration of how DNS maps a client to a cluster, see [Figure A–6](#).

Before You Begin Complete the procedure “[How to Provoke a Switchover](#)” on page 381.

1 Start the nsupdate command.

For information, see the [nsupdate\(1M\)](#) man page.

2 Remove the current DNS mapping between the logical hostname of the application resource group and the cluster IP address, for both clusters.

```
> update delete lhost-nfsrg-prim A
> update delete lhost-nfsrg-sec A
> update delete ipaddress1rev.in-addr.arpa ttl PTR lhost-nfsrg-prim
> update delete ipaddress2rev.in-addr.arpa ttl PTR lhost-nfsrg-sec
```

ipaddress1rev The IP address of the primary cluster, in reverse order.

ipaddress2rev The IP address of the secondary cluster, in reverse order.

ttl The time to live, in seconds. A typical value is 3600.

3 Create a new DNS mapping between the logical hostname of the application resource group and the cluster IP address, for both clusters.

Map the primary logical hostname to the IP address of the secondary cluster and map the secondary logical hostname to the IP address of the primary cluster.

```
> update add lhost-nfsrg-prim ttl A ipaddress2fwd
> update add lhost-nfsrg-sec ttl A ipaddress1fwd
> update add ipaddress2rev.in-addr.arpa ttl PTR lhost-nfsrg-prim
> update add ipaddress1rev.in-addr.arpa ttl PTR lhost-nfsrg-sec
```

ipaddress2fwd The IP address of the secondary cluster, in forward order.

ipaddress1fwd The IP address of the primary cluster, in forward order.

Index

A

- adapters, transport, 216
- Add Administrative Role wizard, description, 53
- adding
 - cluster file system, 165–168
 - custom roles (RBAC), 56
 - device group, 127, 129
 - directly attached shared-disk quorum devices, 183
 - Network-Attached Storage quorum devices, 187
 - new volumes to device groups, 136
 - nodes, 229–234
 - nodes to a device group, 146
 - nodes to a global cluster, 230
 - nodes to a zone cluster, 230
 - quorum devices, 183
 - quorum server quorum devices, 189
 - roles (RBAC), 53
 - SNMP hosts, 278
 - SNMP users, 279
 - Solaris Volume Manager device groups, 129
 - Sun NAS quorum devices, 185
 - Sun Storage 7000 Unified Storage Systems NAS
 - quorum devices, 185
 - transport cables, adapters, and switches, 214
 - ZFS device groups, 130
- administering
 - cluster file system, 117
 - cluster interconnects and public networks, 211–228
 - cluster with graphical user interface (GUI)
 - tool, 339–346
 - EMC SRDF replicated devices, 106–117
 - global-cluster non-voting nodes, 18

administering (*Continued*)

- global cluster settings, 245–285
- global clusters, 18
- Hitachi TrueCopy replicated devices, 95–106
- IPMP, 211–228
- quorum, 179–203
- storage-based replicated devices, 94–117
- zone clusters, 18, 266

- administrative console, 22
- affinity failover, extension property for data replication, 352
- affinity switchover, configuring for data replication, 366
- application resource groups
 - configuring for data replication, 368–370
 - guidelines, 352
- applying
 - a nonrebooting patch, 307
 - patches, 299
- asynchronous data replication, 86, 349
- attributes, *See* properties
- autoboot property, 233
- Availability Suite, using for data replication, 347

B

- backing up
 - cluster, 23, 313–324
 - file system, 314
 - mirrors online, 318
 - root file systems, 315

- backing up (*Continued*)
 - volumes online, 320
- best practices
 - EMC SRDF, 89
 - Hitachi TrueCopy, 89
 - storage-based data replication, 89
- bitmap
 - point-in-time snapshot, 349
 - remote mirror replication, 348
- boot command, 63–65
- booting
 - global cluster, 59–82
 - global-cluster nodes, 68–81
 - nodes, 68–81
 - non-global zones, 68
 - noncluster mode, 78
 - zone cluster, 59–82
 - zone-cluster nodes, 68–81
- C**
- cables, transport, 216
- campus clusters
 - recovery with storage-based data replication, 88
 - storage-based data replication, 85–89
- cconsole command, 22, 25
- ccp command, 20, 25
- changing
 - cluster name, 246–247
 - numsecondaries property, 153
 - Oracle Solaris Cluster Manager server address, 342
 - port number, using Common Agent Container, 341
 - primary nodes, 157–158
 - private hostnames, 253
 - properties, 151
 - SNMP event MIB protocol, 277
- checking
 - cluster interconnect status, 213
 - global mount points, 45, 170
- claccess command, 20
- cldevice command, 20
- cldevicegroup command, 20
- clinterconnect command, 20
- clnasdevice command, 20
- clnode check command, 20
- clnode command, 264
- clquorum command, 20
- clreslogicalhostname command, 20
- clresource command, 20
 - deleting resources and resource groups, 267
- clresourcegroup command, 20, 264–266
- clresourcetype command, 20
- clressharedaddress command, 20
- clsetup utility, 20, 27
- clsnmphost command, 20
- clsnmpmib command, 20
- clsnmpuser command, 20
- cltelemattribute command, 20
- cluster
 - applying a rebooting patch, 303
 - backing up, 23, 313–324
 - changing name, 246–247
 - node authentication, 248
 - restoring files, 325
 - setting time of day, 250
- cluster check command, 20
 - vfstab file check, 167
- Cluster Control Panel (CCP), 22
- cluster file system, 91–178
 - adding, 165–168
 - administration, 117
 - global-cluster non-voting nodes, 117
 - global-cluster voting nodes, 117
 - mount options, 167
 - removing, 168–170
 - verifying the configuration, 167
- cluster interconnects
 - administering, 211–228
 - dynamic reconfiguration, 212
 - status checking, 213
- cluster shutdown command, 59–68
- clzonecluster
 - boot, 63–65
 - description, 27
 - halt, 59–68
- clzonecluster command, 20
- command-line administration tool, 20

- commands
 - boot, 63–65
 - cconsole, 22, 25
 - ccp, 20, 25
 - claccess, 20
 - cldevice, 20
 - cldevicegroup, 20
 - clinterconnect, 20
 - clnasdevice, 20
 - clnode check, 20
 - clquorum, 20
 - clreslogicalhostname, 20
 - clresource, 20
 - clresourcegroup, 20
 - clresourcetype, 20
 - clressharedaddress, 20
 - clsetup, 20
 - clsnmp host, 20
 - clsnmp mib, 20
 - clsnmp user, 20
 - cltelemetryattribute, 20
 - cluster check, 20, 23, 43, 45
 - cluster shutdown, 59–68
 - clzonecluster, 20, 59–68
 - clzonecluster boot, 63–65
 - clzonecluster verify, 43
 - crlogin, 25
 - cssh, 25
 - ctelnet, 25
 - metaset, 91–94
 - netcon, 22
 - Common Agent Container
 - changing port number, 341
 - regenerating security keys, 343
 - configuration examples (campus clustering), two-room, storage-based data replication, 85–89
 - configuring
 - data replication, 347–383
 - device group minor number, 138
 - Hitachi TrueCopy replicated devices, 95–97
 - configuring load limits, on nodes, 264–266
 - consoles
 - connecting to, 25
 - secure connections, 26
 - CPU shares
 - configuring, 287
 - controlling, 287
 - global-cluster non-voting nodes, 291
 - global-cluster non-voting nodes, dedicated processor set, 293
 - global-cluster voting nodes, 289
 - creating, new disk groups, 131
 - crlogin command, 25
 - cssh command, 25
 - ctelnet command, 25
- D**
- data replication, 83–89
 - asynchronous, 349
 - configuring
 - affinity switchover, 352, 366
 - device groups, 360
 - file systems for an NFS application, 363–364
 - NFS application resource groups, 368–370
 - definition, 84–85
 - enabling, 373–375
 - example, 376–381
 - example configuration, 356
 - guidelines
 - configuring resource groups, 351
 - managing failover, 355
 - managing switchover, 355
 - host-based, 84
 - introduction, 348
 - managing a failover, 381–383
 - point-in-time snapshot, 349, 377–378
 - remote mirror, 348, 376–377
 - required hardware and software, 358
 - resource groups
 - application, 352
 - configuring, 351
 - creating, 365–367
 - failover applications, 353
 - naming convention, 352
 - scalable applications, 354–355
 - shared address, 354
 - storage-based, 84, 85–89

- data replication (*Continued*)
 - synchronous, 349
 - updating a DNS entry, 382–383
 - verifying the configuration, 378–381
 - dedicated processor set, configuring, 293
 - device groups
 - adding, 129
 - administration overview, 119
 - changing properties, 151
 - configuring for data replication, 360
 - listing configuration, 156
 - maintenance state, 158
 - new minor number assignment, 138
 - primary ownership, 151
 - raw-disk
 - adding, 129
 - removing and unregistering, 132, 145
 - SVM
 - adding, 127
 - verifying registration, 144
 - devices, global, 91–178
 - DID information, manually updating, 174–175
 - directly attached shared-disk quorum devices,
 - adding, 183
 - disabling transport cables, 220
 - disaster tolerance, definition, 348
 - disk groups
 - creating, 131
 - modifying, 137
 - registering, 139
 - registering configuration changes, 141
 - disk path
 - monitoring, 91–178
 - print faulted disk paths, 174
 - resolving status error, 174–175
 - unmonitoring, 173–174
 - displaying configured resources, 30
 - Domain Name System (DNS)
 - guidelines for updating, 355
 - updating in data replication, 382–383
 - DR, *See* dynamic reconfiguration
 - dynamic reconfiguration, 92–93
 - cluster interconnects, 212
 - public network interfaces, 227
 - dynamic reconfiguration (*Continued*)
 - quorum devices, 181
- E**
- EMC SRDF
 - Adaptive Copy, 87
 - administering, 106–117
 - best practices, 89
 - configuration example, 110–117
 - configuring DID devices, 108–109
 - configuring replication group, 106–108
 - Domino mode, 87
 - recovering after a campus cluster's primary room completely fails over, 115–117
 - requirements, 87
 - restrictions, 87
 - verifying configuration, 109–110
 - enabling and disabling SNMP event MIB, 276, 277
 - enabling transport cables, 219
 - encapsulating disks, 135
 - error messages
 - removing nodes, 242–243
 - `/var/adm/messages` file, 81
 - `/etc/inet/hosts` file, configuring on exclusive-IP zones, 234
 - `/etc/nsswitch.conf` file, non-global zone modifications, 234
 - `/etc/vfstab` file, 45
 - adding mount points, 166
 - verifying the configuration, 167
 - event MIB
 - changing SNMP protocol, 277
 - enabling and disabling SNMP, 276, 277
 - example configurations (campus clustering), two-room, storage-based replication, 85–89
 - examples, creating a cluster file system, 168
 - exclusive-IP zones, *See* Oracle Solaris Containers
 - extension properties for data replication
 - application resource, 369, 371
 - replication resource, 366, 367

F

- failback property, 151
- failover applications for data replication
 - affinity switchover, 352
 - guidelines
 - managing failover, 355
 - resource groups, 353
 - managing, 381–383
- fair share scheduler, CPU shares configuration, 288
- file system
 - backing up, 314
 - finding names, 314
 - NFS application
 - configuring for data replication, 363–364
 - removing it in a zone cluster, 266
 - restoring encapsulated root, 335
 - restoring nonencapsulated root, 333
 - restoring root
 - description, 326
 - from metadvice, 328
 - from volume, 328
- files
 - /etc/vfstab, 45
 - md.conf, 127
 - md.tab, 23
 - ntp.conf.cluster, 254
 - restoring interactively, 325
- finding
 - file system names, 314
 - node IDs for a global cluster, 247
 - node IDs for a zone cluster, 247

G

- global
 - devices, 91–178
 - dynamic reconfiguration, 92–93
 - setting permissions, 92
 - mount points, checking, 45, 170
 - namespace, 91–94, 122
- global cluster
 - administration, 245–285
 - booting, 59–82
 - component status, 31

global cluster (*Continued*)

- definition, 17
- rebooting, 65
- removing nodes, 236
- shutting down, 59–82
- validating configuration, 43
- viewing configuration, 35
- global-cluster nodes
 - booting, 68–81
 - rebooting, 74–78
 - shutting down, 68–81
- global-cluster non-voting nodes
 - adding a private hostname, 255
 - administering cluster file system, 117
 - administration, 18
 - applying patches, 302
 - changing the private hostname, 256
 - CPU shares, 291, 293
 - private hostname, deleting, 257
 - shutting down and rebooting, 68
- global-cluster voting nodes
 - administering cluster file system, 117
 - CPU shares, 289
- global-devices namespace, migrating, 124
- global file systems, *See* cluster file systems
- GUI administration tool, 20, 339–346
 - Oracle Solaris Cluster Manager, 339
 - Sun Management Center, 340

H

- Hitachi TrueCopy
 - administering, 95–106
 - best practices, 89
 - configuration example, 100–106
 - configuring DID devices, 97–99
 - configuring replication group, 95–97
 - Data or Status modes, 87
 - requirements, 87
 - restrictions, 87
 - verifying configuration, 99–100
- host-based data replication
 - definition, 84
 - example, 347–383

hosts
 adding and removing SNMP, 278, 279
hosts file, configuring on exclusive-IP zones, 234

I

IP addresses, adding to a naming service for
 exclusive-IP zones, 234

IPMP

 administration, 226
 groups on exclusive-IP zones
 configuring, 234
 status, 34

K

/kernel/drv/,md.conf file, 127

L

last quorum device, removing, 194
listing
 device group configuration, 156
 quorum configuration, 201
load limits
 concent rate_load property, 264
 configuring on nodes, 264
 preempt ion_mode property, 264
local mirroring, *See* storage-based replication
lofi file, uninstalling, 274
logical hostname resource, role in data replication
 failover, 352
login, remote, 25

M

maintaining, quorum device, 198
maintenance state
 bringing a quorum device out of, 200
 nodes, 260
 putting a quorum device into, 198

manually update DID information, 174–175
md.tab file, 23
metaset command, 91–94
MIB
 changing SNMP event protocol, 277
 enabling and disabling SNMP event, 276, 277
migrating, global-devices namespace, 124
mirrors, online backup, 318
modifying
 disk groups, 137
 quorum device node lists, 196
 users (RBAC), 57
monitoring
 disk paths, 172–173
 shared-disk paths, 177–178
mount options for cluster file systems,
 requirements, 167
mount points
 global, 45
 modifying the /etc/vfstab file, 166
multiuser services, verifying, 233

N

namespace
 global, 91–94
 migrating, 124
naming convention
 raw-disk devices, 166
 replication resource groups, 352
naming service, adding IP-address mappings for
 exclusive-IP zones, 234
NAS, *See* Network-Attached Storage quorum devices
NetApp, *See* Network-Attached Storage quorum devices
net con command, 22
Network Appliance, *See* Network-Attached Storage
 quorum devices
Network-Attached Storage quorum devices, adding and
 installing, 187
Network File System (NFS), configuring application file
 systems for data replication, 363–364
nodes
 adding, 229–234
 adding to device groups, 146

nodes (*Continued*)

- applying a rebooting patch, 299
- authentication, 248
- booting, 68–81
- configuring load limits, 264–266
- connecting to, 25
- finding ID, 247
- primary, 92–93, 151
- putting into maintenance state, 260
- removing
 - error messages, 242–243
- removing from a zone cluster, 236
- removing from device groups, 132, 148
- removing nodes from a global cluster, 236
- removing non-voting nodes from a global cluster, 240
- renaming in a global cluster, 258
- renaming in a zone cluster, 258
- secondary, 151
- shutting down, 68–81
- noncluster mode boot, 78
- nsswitch.conf file, non-global zone modifications, 234
- ntp.conf.cluster file, 254
- numsecondaries property, 153

O

- OpenBoot PROM (OBP), 252
- Oracle Solaris Cluster Manager, 20, 339
 - changing server address, 342
 - RBAC roles, setting, 341
 - starting, 344
- Oracle Solaris Containers
 - autoboot property, 233
 - exclusive-IP zones
 - configuring IPMP groups, 234
 - configuring the hosts file, 234
 - nsswitch.conf file modifications, 234
 - shared-IP zones, 233
- Oracle Solaris OS
 - administrative tasks for a global cluster, 18
 - CPU control, 287
 - global cluster definition, 17

Oracle Solaris OS (*Continued*)

- host-based replication, 84–85
- SMF, 233
- special instructions for booting nodes, 72–74
- special instructions for rebooting a node, 74–78
- svcadm command, 253
- zone cluster definition, 17
- out of service, quorum device, 198
- overview, quorum, 179–203

P

- patches
 - applying a nonrebooting patch, 307
 - applying a rebooting patch, 299
 - applying to cluster and firmware, 303
 - tips, 298
- patching, in global-cluster non-voting nodes, 302
- patching Oracle Solaris Cluster software, 297–299
- permissions, global device, 92
- point-in-time snapshot
 - definition, 349
 - performing, 377–378
- port number, changing using Common Agent Container, 341
- power management, 245
- primary node switching for a device group, 157–158
- primary ownership of device groups, 151
- printing, faulted disk paths, 174
- private hostnames
 - assigning to zones, 233
 - changing, 253
 - changing on global-cluster non-voting nodes, 256
 - deleting on global-cluster non-voting nodes, 257
 - global-cluster non-voting nodes, 255
- profiles, RBAC rights, 52–53
- properties
 - fallback, 151
 - numsecondaries, 153
 - preferenced, 151
- public network
 - administration, 211–228
 - dynamic reconfiguration, 227

Q

quorum

- administration, 179–203
- overview, 179–203

quorum devices

- adding, 183
 - directly attached shared-disk quorum devices, 183
 - Network-Attached Storage quorum devices, 187
 - quorum server quorum devices, 189
 - Sun NAS quorum devices, 185
 - Sun Storage 7000 Unified Storage Systems NAS quorum devices, 185
 - and storage-based replication, 88
 - dynamic reconfiguration of devices, 181
 - listing configuration, 201
 - maintenance state, bringing a device out of, 200
 - maintenance state, putting a device into, 198
 - modifying node lists, 196
 - removing, 181, 192
 - removing last quorum device, 194
 - repairing, 202
 - replacing, 195–196
- quorum server quorum devices
- adding, 189
 - requirements for installing, 189
 - troubleshooting removals, 194
- quorum servers, *See* quorum server quorum devices

R

raw-disk device group, adding, 129

raw-disk devices, naming conventions, 166

RBAC, 51–58

- for global-cluster voting nodes, 52

- for non-voting nodes, 52

Oracle Solaris Cluster Manager, 341

rights profiles (description), 52–53

tasks

- adding custom roles, 56

- adding roles, 53

- modifying users, 57

- setting up, 51

- using, 51

rebooting

- global cluster, 65

- global-cluster nodes, 74–78

- zone cluster, 65

- zone-cluster nodes, 74–78

recovery, clusters with storage-based data replication, 88

regenerating, security keys, 343

registering

- disk group configuration changes, 141

- disk groups as device groups, 139

release information, 27, 28

remote login, 25

remote mirror replication

- definition, 348

- performing, 376–377

remote mirroring, *See* storage-based replication

remote replication, *See* storage-based replication

removing

- cluster file system, 168–170

- device groups, 145

- from a zone cluster, 236

- last quorum device, 194

- nodes, 235, 236

- nodes from all device groups, 132

- nodes from device groups, 148

- non-voting nodes on a global cluster, 240

- quorum devices, 181, 192

- resources and resource groups from a zone cluster, 267

- SNMP hosts, 279

- SNMP users, 280

- Solaris Volume Manager device groups, 132

- storage arrays, 240

- transport cables, adapters, and switches, 216

- volumes from a device group, 144–145

renaming nodes

- in a global cluster, 258

- in a zone cluster, 258

repairing, quorum device, 202

repairing full `/var/adm/messages` file, 81

replacing quorum devices, 195–196

replication, *See* data replication

replication, storage-based, 85–89

- resource groups
 - data replication
 - configuring, 351
 - guidelines for configuring, 351
 - role in failover, 351
 - resources
 - deleting, 267
 - displaying configured types, 30
 - restarting
 - global-cluster nodes, 74–78
 - zone-cluster nodes, 74–78
 - restoring
 - cluster files, 325
 - encapsulated root file system, 335
 - files interactively, 325
 - nonencapsulated root file system, 333
 - root file system, 326
 - from metadvice, 328
 - from volume, 328
 - rights profiles, RBAC, 52–53
 - role
 - adding custom roles, 56
 - adding roles, 53
 - setting up, 51
 - Role-Based Access Control, *See* RBAC
- S**
- SATA, 182, 183
 - scalable applications for data replication, 354–355
 - secondaries
 - default number, 151
 - setting desired number, 153
 - secure connections to cluster consoles, 26
 - secure shell, 25, 26
 - security keys, regenerating, 343
 - setting cluster time, 250
 - setting up, roles (RBAC), 51
 - shared address resource groups for data replication, 354
 - shared-disk paths
 - disabling automatic reboot, 178
 - enabling automatic reboot, 177–178
 - monitoring, 171–178
 - shared-IP zones, *See* Oracle Solaris Containers
 - showrev -p command, 27, 28
 - shutting down
 - global cluster, 59–82
 - global-cluster nodes, 68–81
 - nodes, 68–81
 - non-global zones, 68
 - zone cluster, 59–82
 - zone-cluster nodes, 68–81
 - SMF, verifying online services, 233
 - snapshot, point-in-time, 349
 - snapshots, *See* storage-based replication
 - SNMP
 - adding users, 279
 - changing protocol, 277
 - disabling hosts, 279
 - enabling and disabling event MIB, 276, 277
 - enabling hosts, 278
 - removing users, 280
 - Solaris OS
 - See also* Oracle Solaris OS
 - Solaris Volume Manager, raw-disk device names, 166
 - SRDF
 - See* EMC SRDF
 - ssh, 26
 - starting
 - global cluster, 63–65
 - global-cluster nodes, 68–81
 - nodes, 68–81
 - Oracle Solaris Cluster Manager, 344
 - zone cluster, 63–65
 - zone-cluster nodes, 68–81
 - status
 - global cluster component, 31
 - zone cluster component, 31
 - stopping
 - global cluster, 65
 - global-cluster nodes, 68–81
 - nodes, 68–81
 - zone cluster, 65
 - zone-cluster nodes, 68–81
 - storage arrays, removing, 240
 - storage-based data replication, 85–89
 - and quorum devices, 88

storage-based data replication (*Continued*)

- best practices, 89
 - definition, 84
 - recovery, 88
 - requirements, 87
 - restrictions, 87
- storage-based replicated devices,
administering, 94–117
- Sun Management Center
description, 20
installation, 22
overview, 340
- Sun NAS quorum devices, adding, 185
- Sun Storage 7000 Unified Storage Systems NAS quorum
devices, adding, 185
- Sun StorageTek Availability Suite, using for data
replication, 347
- SunMC, *See* Sun Management Center
- supported features, VxFS, 117
- switchback, guidelines for performing in data
replication, 356
- switches, transport, 216
- switching, primary node for a device group, 157–158
- switchover for data replication
affinity switchover, 352
performing, 381–383
- synchronous data replication, 86, 349
- System Service Processor (SSP), 22

T

- transport adapters, adding, 214, 216
- transport cables
 - adding, 214, 216
 - disabling, 220
 - enabling, 219
- transport switches, adding, 214, 216
- TrueCopy
 - See* Hitachi TrueCopy

U

- uninstalling
 - lofi device file, 274
 - Oracle Solaris Cluster software, 271
- unmonitoring, disk paths, 173–174
- unregistering
 - device groups, 145
 - Solaris Volume Manager device groups, 132
- updating global namespace, 122
- User Accounts tool, description, 57
- users
 - adding SNMP, 279
 - modifying properties, 57
 - removing SNMP, 280
- using, roles (RBAC), 51
- `/usr/cluster/bin/clresource`, deleting resource
groups, 267
- `/usr/cluster/bin/cluster check` command, `vfstab`
file check, 167

V

- validating
 - global-cluster configuration, 43
 - zone-cluster configuration, 43
- `/var/adm/messages` file, 81
- verifying
 - data replication configuration, 378–381
 - SMF, 233
 - `vfstab` configuration, 167
- Veritas
 - administration, 93–94
 - online backups, 320
 - restoring encapsulated root file system, 335
 - restoring nonencapsulated root file system, 333
- Veritas File System (VxFS)
 - administering, 168
 - mounting cluster file systems, 167
- Veritas Volume Manager (VxVM), raw-disk device
names, 166
- `vfstab` file
 - adding mount points, 166
 - verifying the configuration, 167

- viewing
 - global-cluster configuration, 35
 - zone-cluster configuration, 35
- volume, *See* storage-based replication
- volume managers, Veritas, 93–94
- volumes
 - adding to device groups, 136
 - backing up online, 320
 - removing from device groups, 144–145
- VxFS supported features, 117
- VxVM, 93–94

Z

- ZFS
 - adding device groups, 130
 - removing file system, 268–270
 - replication, 130
 - restrictions for root file systems, 118
- zone cluster
 - administration, 245–285
 - booting, 59–82
 - cloning, 266
 - component status, 31
 - definition, 17
 - moving a zonepath, 266
 - preparing it for applications, 266
 - rebooting, 65
 - removing a file system, 266
 - shutting down, 59–82
 - validating configuration, 43
 - viewing configuration, 35
- zone-cluster nodes
 - booting, 68–81
 - rebooting, 74–78
 - shutting down, 68–81
- zone path, moving, 266

