Oracle® Solaris Cluster 4.3 System Administration Guide

Part No: E56680
June 2017

Oracle Solaris Cluster 4.3 System Administration Guide

Part No: E56680

Copyright © 2000, 2017, 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 documentation that is delivered to the U.S. Government or anyone licensing it on behalf of the U.S. Government, then the following notice is applicable:

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

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 that 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.

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. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.

This software or hardware and documentation may provide access to or information about 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 unless otherwise set forth in an applicable agreement between you and Oracle. 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, except as set forth in an applicable agreement between you and Oracle.

Access to Oracle Support

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

Référence: E56680

Copyright © 2000, 2017, Oracle et/ou ses affiliés. Tous droits réservés.

Ce logiciel et la documentation qui l'accompagne sont protégés par les lois sur la propriété intellectuelle. Ils sont concédés sous licence et soumis à des restrictions d'utilisation et de divulgation. Sauf stipulation expresse de votre contrat de licence ou de la loi, vous ne pouvez pas copier, reproduire, traduire, diffuser, modifier, accorder de licence, transmettre, distribuer, exposer, exécuter, publier ou afficher le logiciel, même partiellement, sous quelque forme et par quelque procédé que ce soit. Par ailleurs, il est interdit de procéder à toute ingénierie inverse du logiciel, de le désassembler ou de le décompiler, excepté à des fins d'interopérabilité avec des logiciels tiers ou tel que prescrit par la loi.

Les informations fournies dans ce document sont susceptibles de modification sans préavis. Par ailleurs, Oracle Corporation ne garantit pas qu'elles soient exemptes d'erreurs et vous invite, le cas échéant, à lui en faire part par écrit.

Si ce logiciel, ou la documentation qui l'accompagne, est livré sous licence au Gouvernement des Etats-Unis, ou à quiconque qui aurait souscrit la licence de ce logiciel pour le compte du Gouvernement des Etats-Unis, la notice suivante s'applique :

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

Ce logiciel ou matériel a été développé pour un usage général dans le cadre d'applications de gestion des informations. Ce logiciel ou matériel n'est pas conçu ni n'est destiné à être utilisé dans des applications à risque, notamment dans des applications pouvant causer un risque de dommages corporels. Si vous utilisez ce logiciel ou ce matériel dans le cadre d'applications dangereuses, il est de votre responsabilité de prendre toutes les mesures de secours, de sauvegarde, de redondance et autres mesures nécessaires à son utilisation dans des conditions optimales de sécurité. Oracle Corporation et ses affiliés déclinent toute responsabilité quant aux dommages causés par l'utilisation de ce logiciel ou matériel pour des applications dangereuses.

Oracle et Java sont des marques déposées d'Oracle Corporation et/ou de ses affiliés. Tout autre nom mentionné peut correspondre à des marques appartenant à d'autres propriétaires qu'Oracle.

Intel et Intel Xeon sont des marques ou des marques déposées d'Intel Corporation. Toutes les marques SPARC sont utilisées sous licence et sont des marques ou des marques déposées de SPARC International, Inc. AMD, Opteron, le logo AMD et le logo AMD Opteron sont des marques ou des marques déposées d'Advanced Micro Devices. UNIX est une marque déposée de The Open Group.

Ce logiciel ou matériel et la documentation qui l'accompagne peuvent fournir des informations ou des liens donnant accès à des contenus, des produits et des services émanant de tiers. Oracle Corporation et ses affiliés déclinent toute responsabilité ou garantie expresse quant aux contenus, produits ou services émanant de tiers, sauf mention contraire stipulée dans un contrat entre vous et Oracle. En aucun cas, Oracle Corporation et ses affiliés ne sauraient être tenus pour responsables des pertes subies, des coûts occasionnés ou des dommages causés par l'accès à des contenus, produits ou services tiers, ou à leur utilisation, sauf mention contraire stipulée dans un contrat entre vous et Oracle.

Accès aux services de support Oracle

Les clients Oracle qui ont souscrit un contrat de support ont accès au support électronique via My Oracle Support. Pour plus d'informations, visitez le site `http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info` ou le site `http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs` si vous êtes malentendant.

Contents

Figures

Tables

Examples

Using This Documentation

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

- **Overview** – Describes how to configure an Oracle Solaris Custer configuration
- **Audience** – Technicians, system administrators, and authorized service providers
- **Required knowledge** – Advanced experience troubleshooting and replacing hardware

Product Documentation Library

Documentation and resources for this product and related products are available at http://www.oracle.com/pls/topic/lookup?ctx=E56676-01.

Feedback

Provide feedback about this documentation at http://www.oracle.com/goto/docfeedback.

♦ ♦ ♦ CHAPTER 1

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:

All procedures in this guide are for use on the Oracle Solaris 11 Operating System.

A global cluster is composed of one or more global-cluster nodes. A global cluster can also include `solaris` or `solaris10` branded non-global zones that are not nodes but rather are configured with the HA for Zones data service.

A zone cluster is composed of one or more non-global zones of the `solaris`, `solaris10`, or `labeled` brand that are set with the `cluster` attribute. No other brand type is permitted in a zone cluster. A `labeled` branded zone cluster is only for use with the Trusted Extensions feature of Oracle Solaris software. You create a zone cluster by using the `clzonecluster` command, the `clsetup` utility, or the Oracle Solaris Cluster Manager browser interface.

You can run supported services on the zone cluster similar to a global cluster, with the isolation that is provided by Oracle Solaris zones. A zone cluster depends on, and therefore requires, a global cluster. A global cluster does not contain a zone 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 node on the same machine continues to operate. If a global-cluster node on a machine fails, all zone-cluster nodes on that machine fail as well. For general information about zone clusters, see *Oracle Solaris Cluster 4.3 Concepts Guide*.

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 following manuals before beginning administration tasks.

- Chapter 1, "Planning the Oracle Solaris Cluster Configuration" in *Oracle Solaris Cluster 4.3 Software Installation Guide*
- *Oracle Solaris Cluster 4.3 Concepts Guide*

Oracle Solaris Cluster administration is organized into tasks among the following manuals.

Note - Some of these tasks can be done using the Oracle Solaris Cluster Manager browser interface. This is noted in the individual task procedures. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

- 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 4.3 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 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 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 node retrieves all information about the 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 global-cluster 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 a zone-cluster 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 global-cluster 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.

- **Starting and rebooting a zone cluster** – See Chapter 3, "Shutting Down and Booting a Cluster". You can also boot and reboot a zone cluster by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.
- **Adding a node to a zone cluster** – See Chapter 8, "Administering Cluster Nodes".
- **Removing a node from a zone cluster** – See "How to Remove a Node From a Zone Cluster" on page 232. You can also uninstall the software from a zone cluster node by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.
- **Viewing the configuration of a zone cluster** – See "How to View the Cluster Configuration" on page 42. You can also view the configuration of a zone cluster by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.
- **Validating the configuration of a zone cluster** – See "How to Validate a Basic Cluster Configuration" on page 53.
- Stopping a zone cluster – See Chapter 3, "Shutting Down and Booting a Cluster". You can also shut down a zone cluster by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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 Oracle Solaris Cluster Services

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 for an Oracle Solaris Cluster configuration by using the command line or the Oracle Solaris Cluster Manager browser interface. The following section provides an overview of Oracle Solaris Cluster Manager and the command-line tools.

Oracle Solaris Cluster Manager Browser Interface

Oracle Solaris Cluster software supports a browser interface, Oracle Solaris Cluster Manager, that you can use to perform various administrative tasks on your cluster. See Chapter 13, "Using the Oracle Solaris Cluster Manager Browser Interface" for more information. You can also get Oracle Solaris Cluster Manager log-in instructions at "How to Access Oracle Solaris Cluster Manager" on page 313.

Following are some of the tasks that you can perform in Oracle Solaris Cluster Manager:

- Create and update a zone cluster
- Create resources and resource groups
- Add a file system, logical host, or shared storage to a zone cluster
- Create an Oracle Database data service
- Manage nodes in a global cluster or a zone cluster

- Add and manage quorum devices and servers
- Add and manage NAS storage devices, and manage disks and device groups
- Manage Geographic Edition partnerships

Command-Line Interface

You can perform most Oracle Solaris Cluster administration tasks interactively through the `clsetup` 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
- Zone cluster
- 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.

`if_mpadm`(1M)

Switches IP addresses from one adapter to another in an IP Network Multipathing group.

`claccess`(1CL)

Manages Oracle Solaris Cluster access policies for adding nodes.

`cldevice`(1CL)

Manages Oracle Solaris Cluster devices.

`cldevicegroup`(1CL)

Manages Oracle Solaris Cluster device groups.

`clinterconnect`(1CL)

Manages the Oracle Solaris Cluster interconnect.

`clnasdevice`(1CL)

Manages access to NAS devices for an Oracle Solaris Cluster configuration.

`clnode`(1CL)

Manages Oracle Solaris Cluster nodes.

`clquorum`(1CL)

Manages Oracle Solaris Cluster quorum.

`clreslogicalhostname`(1CL)

Manages Oracle Solaris Cluster resources for logical host names.

`clresource`(1CL)

Manages resources for Oracle Solaris Cluster data services.

`clresourcegroup`(1CL)

Manages resources for Oracle Solaris Cluster data services.

`clresourcetype`(1CL)

Manages resources for Oracle Solaris Cluster data services.

`clressharedaddress`(1CL)

Manages Oracle Solaris Cluster resources for shared addresses.

`clsetup`(1CL)

Creates a zone cluster and interactively configures an Oracle Solaris Cluster configuration.

`clsnmphost`(1CL)

Administers Oracle Solaris Cluster SNMP hosts.

`clsnmpmib`(1CL)

Administers Oracle Solaris Cluster SNMP MIB.

`clsnmpuser`(1CL)

Administers Oracle Solaris Cluster SNMP users.

`cltelemetryattribute`(1CL)

Configures system resource monitoring.

`cluster`(1CL)

Manages the global configuration and the global status of the Oracle Solaris Cluster configuration.

`clzonecluster`(1CL)

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.

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.

You can use the `pconsole` utility to run terminal windows for each cluster node plus a master window that issues the commands you type there to all nodes at the same time. For information about installing `pconsole` software on the administrative console, see "How to Install pconsole Software on an Administrative Console" in *Oracle Solaris Cluster 4.3 Software Installation Guide*.

You can use the Oracle Solaris Cluster Manager browser interface to configure, monitor, and administer the cluster and cluster components. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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.

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

Administering the Cluster

Table 2, "Oracle Solaris Cluster Administration Tools," on page 30 provides a starting point for administering your cluster.

TABLE 2 Oracle Solaris Cluster Administration Tools

Task	Tool	Instructions
Log in to the cluster remotely	Use the Oracle Solaris `pconsole` utility from the command line to log into the cluster remotely.	"Logging Into the Cluster Remotely" on page 32 "How to Connect Securely to Cluster Consoles" on page 32
Configure the cluster interactively	Use the `clzonecluster` command or the `clsetup` utility.	"How to Access the Cluster Configuration Utilities" on page 33

Task	Tool	Instructions
Display Oracle Solaris Cluster release number and version information	Use the `clnode` command with the `show-rev -v` *-node* subcommand and option.	"How to Display Oracle Solaris Cluster Release and Version Information" on page 34
Display installed resources, resource groups, and resource types	Use the following commands to display the resource information: ■ `clresource` ■ `clresourcegroup` ■ `clresourcetype`	"How to Display Configured Resource Types, Resource Groups, and Resources" on page 36
Monitor cluster components graphically	Use Oracle Solaris Cluster Manager.	See the online help.
Administer some cluster components graphically	Use Oracle Solaris Cluster Manager.	See the online help.
Check the status of cluster components	Use the `cluster` command with the `status` subcommand.	"How to Check the Status of Cluster Components" on page 39
Check the status of IPMP groups on the public network	For a global cluster, use the `clnode status` command with the `-m` option. For a zone cluster, use the `clzonecluster` command with the `show` subcommand.	"How to Check the Status of the Public Network" on page 42
View the cluster configuration	For a global cluster, use the `cluster` command with the `show` subcommand. For a zone cluster, use the `clzonecluster` command with the `show` subcommand.	"How to View the Cluster Configuration" on page 42
View and display the configured NAS devices	For a global cluster or a zone cluster, use the `clzonecluster` command with the `show` subcommand.	`clnasdevice`(1CL)
Check global mount points or verify the cluster configuration	For a global cluster, use the `cluster` command with the `check` subcommand. For a zone cluster, use the `clzonecluster verify` command.	"How to Validate a Basic Cluster Configuration" on page 53
Look at the contents of Oracle Solaris Cluster command logs	Examine the `/var/cluster/logs/ commandlog` file.	"How to View the Contents of Oracle Solaris Cluster Command Logs" on page 61
Look at Oracle Solaris Cluster system messages	Examine the `/var/adm/messages` file.	"System Message Formats" in *Troubleshooting System Administration Issues in Oracle Solaris 11.3*

Task	Tool	Instructions
		You can also see a node's system messages in the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.
Monitor the status of Solaris Volume Manager	Use the metastat command.	*Solaris Volume Manager Administration Guide*

Logging Into the Cluster Remotely

You can use the Parallel Console Access (pconsole) utility from the command line to log into the cluster remotely. The pconsole utility is part of the Oracle Solaris terminal/pconsole package. Install the package by executing pkg install terminal/pconsole. The pconsole utility creates a host terminal window for each remote host that you specify on the command line. The utility also opens a central, or master, console window that propagates what you input there to each of the connections that you open.

The pconsole utility can be run from within X Windows or in console mode. Install pconsole on the machine that you will use as the administrative console for the cluster. If you have a terminal server that allows you to connect to specific port numbers on the IP address of the server, you can specify the port number in addition to the hostname or IP address as *terminal-server*:*portnumber*.

See the pconsole(1) man page for more information.

How to Connect Securely to Cluster Consoles

If your terminal concentrator or system controller supports ssh, you can use the pconsole utility to connect to the consoles of those systems. The pconsole utility is part of the Oracle Solaris terminal/pconsole package and is installed when you install that package. The pconsole utility creates a host terminal window for each remote host that you specify on the command line. The utility also opens a central, or master, console window that propagates what you input there to each of the connections that you open. See the pconsole(1) man page for more information.

▼ How to Access the Cluster Configuration Utilities

The `clsetup` utility enables you to interactively create a zone cluster, and configure quorum, resource groups, cluster transports, private hostnames, device groups, 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.

Note - You can also accomplish this procedure by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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. **Assume the `root` role 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 q 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> ?
     ```

   You can also use the interactive `clsetup` utility or the Oracle Solaris Cluster Manager browser interface to create a zone cluster or add a file system or storage device in the cluster scope. All other zone cluster configuration tasks are performed with the **`clzonecluster configure`** command. See the *Oracle Solaris Cluster 4.3 Software Installation Guide* for instructions about using the `clsetup` utility or Oracle Solaris Cluster Manager to configure a zone cluster.

3. **Choose your configuration from the menu.**
   Follow the onscreen instructions to complete a task. For more detail, see the instructions in "Creating and Configuring a Zone Cluster" in *Oracle Solaris Cluster 4.3 Software Installation Guide*.

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

▼ How to Display Oracle Solaris Cluster Release and Version Information

You do not need to be logged in as the `root` role 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 Displaying Oracle Solaris Cluster Release and Version Information

The following example displays the cluster's release information and version information for the packages that shipped with Oracle Solaris Cluster 4.2.

```
phys-schost# clnode show-rev
4.2

phys-schost#% clnode show-rev -v

Oracle Solaris Cluster 4.2 for Oracle Solaris 11 sparc
ha-cluster/data-service/apache                          :4.2-0.30
ha-cluster/data-service/dhcp                            :4.2-0.30
```

```
ha-cluster/data-service/dns                            :4.2-0.30
ha-cluster/data-service/goldengate                     :4.2-0.30
ha-cluster/data-service/glassfish-message-queue        :4.2-0.30
ha-cluster/data-service/ha-ldom                        :4.2-0.30
ha-cluster/data-service/ha-zones                       :4.2-0.30
ha-cluster/data-service/iplanet-web-server             :4.2-0.30
ha-cluster/data-service/jd-edwards-enterpriseone       :4.2-0.30
ha-cluster/data-service/mysql                          :4.2-0.30
ha-cluster/data-service/nfs                            :4.2-0.30
ha-cluster/data-service/obiee                          :4.2-0.30
ha-cluster/data-service/oracle-database                :4.2-0.30
ha-cluster/data-service/oracle-ebs                     :4.2-0.30
ha-cluster/data-service/oracle-external-proxy          :4.2-0.30
ha-cluster/data-service/oracle-http-server             :4.2-0.30
ha-cluster/data-service/oracle-pmn-server              :4.2-0.30
ha-cluster/data-service/oracle-traffic-director        :4.2-0.30
ha-cluster/data-service/peoplesoft                     :4.2-0.30
ha-cluster/data-service/postgresql                     :4.2-0.30
ha-cluster/data-service/samba                          :4.2-0.30
ha-cluster/data-service/sap-livecache                  :4.2-0.30
ha-cluster/data-service/sapdb                          :4.2-0.30
ha-cluster/data-service/sapnetweaver                   :4.2-0.30
ha-cluster/data-service/siebel                         :4.2-0.30
ha-cluster/data-service/sybase                         :4.2-0.30
ha-cluster/data-service/timesten                       :4.2-0.30
ha-cluster/data-service/tomcat                         :4.2-0.30
ha-cluster/data-service/weblogic                       :4.2-0.30
ha-cluster/developer/agent-builder                     :4.2-0.30
ha-cluster/developer/api                               :4.2-0.30
ha-cluster/geo/geo-framework                           :4.2-0.30
ha-cluster/geo/manual                                  :4.2-0.30
ha-cluster/geo/replication/availability-suite          :4.2-0.30
ha-cluster/geo/replication/data-guard                  :4.2-0.30
ha-cluster/geo/replication/sbp                         :4.2-0.30
ha-cluster/geo/replication/srdf                        :4.2-0.30
ha-cluster/geo/replication/zfs-sa                      :4.2-0.30
ha-cluster/group-package/ha-cluster-data-services-full     :4.2-0.30
ha-cluster/group-package/ha-cluster-framework-full     :4.2-0.30
ha-cluster/group-package/ha-cluster-framework-l10n     :4.2-0.30
ha-cluster/group-package/ha-cluster-framework-minimal  :4.2-0.30
ha-cluster/group-package/ha-cluster-framework-scm      :4.2-0.30
ha-cluster/group-package/ha-cluster-framework-slm      :4.2-0.30
ha-cluster/group-package/ha-cluster-full               :4.2-0.30
ha-cluster/group-package/ha-cluster-geo-full           :4.2-0.30
ha-cluster/group-package/ha-cluster-geo-incorporation  :4.2-0.30
ha-cluster/group-package/ha-cluster-incorporation      :4.2-0.30
ha-cluster/group-package/ha-cluster-minimal            :4.2-0.30
ha-cluster/group-package/ha-cluster-quorum-server-full     :4.2-0.30
```

```
ha-cluster/group-package/ha-cluster-quorum-server-l10n     :4.2-0.30
ha-cluster/ha-service/derby                                :4.2-0.30
ha-cluster/ha-service/gds                                  :4.2-0.30
ha-cluster/ha-service/gds2                                 :4.2-0.30
ha-cluster/ha-service/logical-hostname                     :4.2-0.30
ha-cluster/ha-service/smf-proxy                            :4.2-0.30
ha-cluster/ha-service/telemetry                            :4.2-0.30
ha-cluster/library/cacao                                   :4.2-0.30
ha-cluster/library/ucmm                                    :4.2-0.30
ha-cluster/locale                                          :4.2-0.30
ha-cluster/release/name                                    :4.2-0.30
ha-cluster/service/management                              :4.2-0.30
ha-cluster/service/management/slm                          :4.2-0.30
ha-cluster/service/quorum-server                           :4.2-0.30
ha-cluster/service/quorum-server/locale                    :4.2-0.30
ha-cluster/service/quorum-server/manual/locale             :4.2-0.30
ha-cluster/storage/svm-mediator                            :4.2-0.30
ha-cluster/system/cfgchk                                   :4.2-0.30
ha-cluster/system/core                                     :4.2-0.30
ha-cluster/system/dsconfig-wizard                          :4.2-0.30
ha-cluster/system/install                                  :4.2-0.30
ha-cluster/system/manual                                   :4.2-0.30
ha-cluster/system/manual/data-services                     :4.2-0.30
ha-cluster/system/manual/locale                            :4.2-0.30
ha-cluster/system/manual/manager                           :4.2-0.30
ha-cluster/system/manual/manager-glassfish3:4.2-0.30
```

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

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 - You can also view resources and resource groups through the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

Before You Begin Users other than the `root` role require `solaris.cluster.read` authorization to use this subcommand.

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

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

Perform all steps of this procedure from a node of the global cluster. 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 2 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.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:                                         <NULL>
RT_system:                                       True
Global_zone:                                     True

=== Resource Groups and Resources ===

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

```
--- Resources for Group tel-rg ---

Resource:                                    tel-res
Type:                                         SUNW.sctelemetry
Type_version:                                 4.0
Group:                                        tel-rg
R_description:
Resource_project_name:                        default
Enabled{phys-schost-2}:                       True
Enabled{phys-schost-1}:                       True
Monitored{phys-schost-2}:                     True
Monitored{phys-schost-1}:                     True

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
Global_zone:                                  True

=== 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:
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

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 - You can also check the status of cluster components by using the Oracle Solaris Cluster Manager browser interface. See the Oracle Solaris Cluster Manager online help for more information.

Oracle Solaris Cluster Manager and the `cluster status` command also show the status of a zone cluster.

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

- **Check the status of cluster components.**

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

Perform all steps of this procedure from a node of the global cluster.

Example 3 Checking the Status of Cluster Components

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

```
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:nge1     phys-schost-4:nge1     Path online
phys-schost-1:e1000g1  phys-schost-4:e1000g1  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/rdsk/d2s2    1            1             Online
/dev/did/rdsk/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/rdsk/d2            phys-schost-1             Ok

/dev/did/rdsk/d3            phys-schost-1             Ok
                            phys-schost-2             Ok

/dev/did/rdsk/d4            phys-schost-1             Ok
                            phys-schost-2             Ok

/dev/did/rdsk/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-2sczone-2OnlineRunning
```

How to Check the Status of the Public Network

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 command with the `clnode status` command.

Before You Begin Users other than the `root` role require `solaris.cluster.read` authorization to use this subcommand.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to check the status of the node. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

- **Check the status of cluster components.**

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

Perform all steps of this procedure from a node of the global cluster.

Example 4 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    nge2       Online
phys-schost-2 test-rg Onlinenge3 Online
```

How to View the Cluster Configuration

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 - You can also view a cluster's configuration through the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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

- **View the configuration of a global cluster or zone cluster.**

```
% cluster show
```

Perform all steps of this procedure from a node of the global cluster.

Running the `cluster show` command from a global-cluster 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 5 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
clusterid:                                       0x4DA2C888
installmode:                                     disabled
heartbeat_timeout:                               10000
heartbeat_quantum:                               1000
private_netaddr:                                 172.11.0.0
```

```
  private_netmask:                                 255.255.248.0
  max_nodes:                                       64
  max_privatenets:                                 10
  num_zoneclusters:                                12
  udp_session_timeout:                             480
  concentrate_load:                                False
  global_fencing:                                  prefer3
  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
    Enabled:                                     yes
    privatehostname:                             clusternode1-priv
    reboot_on_path_failure:                      disabled
    globalzoneshares:                            3
    defaultpsetmin:                              1
    quorum_vote:                                 1
    quorum_defaultvote:                          1
    quorum_resv_key:                             0x43CB1E1800000001
    Transport Adapter List:                      net1, net3

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

    Transport Adapter:                          net1
      Adapter State:                             Enabled
      Adapter Transport Type:                    dlpi
      Adapter Property(device_name):             net
      Adapter Property(device_instance):         1
      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:                          net3
```

```
Adapter State:                              Enabled
Adapter Transport Type:                     dlpi
Adapter Property(device_name):              net
Adapter Property(device_instance):          3
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
defaultpsetmin:                                2
quorum_vote:                                   1
quorum_defaultvote:                            1
quorum_resv_key:                               0x43CB1E1800000002
Transport Adapter List:                        e1000g1, nge1

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

Transport Adapter:                         e1000g1
Adapter State:                              Enabled
Adapter Transport Type:                     dlpi
Adapter Property(device_name):              e1000g
Adapter Property(device_instance):          2
```

```
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:                         nge1
Adapter State:                              Enabled
Adapter Transport Type:                     dlpi
Adapter Property(device_name):              nge
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:e1000g1,switch2@1
Cable Endpoint1:                              phys-schost-1:e1000g1
Cable Endpoint2:                              switch2@1
Cable State:                                  Enabled

Transport Cable:                             phys-schost-1:nge1,switch1@1
Cable Endpoint1:                              phys-schost-1:nge1
Cable Endpoint2:                              switch1@1
Cable State:                                  Enabled

Transport Cable:                             phys-schost-2:nge1,switch1@2
```

```
Cable Endpoint1:                              phys-schost-2:nge1
Cable Endpoint2:                              switch1@2
Cable State:                                  Enabled

Transport Cable:                             phys-schost-2:e1000g1,switch2@2
Cable Endpoint1:                              phys-schost-2:e1000g1
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/rdsk/d3s2
Type:                                         shared_disk
Access Mode:                                  scsi3
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
  preferenced:                                   yes
  numsecondaries:                                1
  diskset name:                                  testdg3

=== Registered Resource Types ===

Resource Type:                                  SUNW.LogicalHostname:2
  RT_description:                                Logical Hostname Resource Type
  RT_version:                                    4
  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:                                       <NULL>
  RT_system:                                     True
  Global_zone:                                   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:                                       <NULL>
  RT_system:                                     True
  Global_zone:                                   True
Resource Type:                                  SUNW.HAStoragePlus: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:                                       <NULL>
```

```
RT_system:                                       True
Global_zone:                                     True
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:                                         <NULL>
RT_system:                                       True
Global_zone:                                     True
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:                                         <NULL>
RT_system:                                       True
Global_zone:                                     True
=== 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.HAStoragePlus: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/rdsk/d1
Full Device Path:                              phys-schost-1:/dev/rdsk/c0t2d0
Replication:                                   none
```

```
default_fencing:                              global

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

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

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

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

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

=== NAS Devices ===

Nas Device:                                  nas_filer1
Type:                                         sun_uss
nodeIPs{phys-schost-2}:                       10.134.112.112
nodeIPs{phys-schost-1                         10.134.112.113
User ID: root
```

Example 6 Viewing the Zone Cluster Configuration

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

```
% clzonecluster show
=== Zone Clusters ===

Zone Cluster Name:                             sczone
zonename:                                       sczone
zonepath:                                       /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:                                             /local/ufs-1
special:                                         /dev/md/ds1/dsk/d0
raw:                                             /dev/md/ds1/rdsk/d0
type:                                            ufs
options:                                         [logging]

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` command uses the `check` subcommand 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.

In addition to basic checks, which run without user interaction, the command can also run interactive checks and functional checks. Basic checks are run when the `-k` *keyword* option is not specified.

- Interactive checks require information from the user that the checks cannot determine. The check prompts the user for the needed information, for example, the firmware version number. Use the `-k interactive` keyword to specify one or more interactive checks.
- Functional checks exercise a specific function or behavior of the cluster. The check prompts for user input, such as which node to fail over to, as well as confirmation to begin or continue the check. Use the `-k functional` *check-id* keyword to specify a functional check. Perform only one functional check at a time.

Note - Because some functional checks involve interrupting cluster service, do not start any functional check until you have read the detailed description of the check and determined whether you need to first take the cluster out of production. To display this information, use the following command:

```
% cluster list-checks -v -C checkID
```

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 from the global-cluster 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. **Assume the `root` role on an active member node of a global cluster.**

   ```
   phys-schost# su
   ```

   Perform all steps of this procedure from a node of the global cluster.

2. **Ensure that you have the most current checks.**

   a. **Go to the Patches & Updates tab of My Oracle Support.**

   b. **In the Advanced Search, select `Solaris Cluster` as the Product and type `check` in the Description field.**

      The search locates Oracle Solaris Cluster software updates that contain checks.

   c. **Apply any software updates that are not already installed on your cluster.**

3. **Run basic validation checks.**

```
phys-schost# cluster check -v -o outputdir
```

`-v` Verbose mode.

`-o` *outputdir* Redirects output to the *outputdir* subdirectory.

This command runs all available basic checks. No cluster functionality is affected.

4. **Run interactive validation checks.**

   ```
   phys-schost# cluster check -v -k interactive -o outputdir
   ```

   `-k interactive` Specifies running interactive validation checks.

   The command runs all available interactive checks and prompts you for needed information about the cluster. No cluster functionality is affected.

5. **Run functional validation checks.**

   a. **List all available functional checks in nonverbose mode.**

      ```
      phys-schost# cluster list-checks -k functional
      ```

   b. **Determine which functional checks perform actions that would interfere with cluster availability or services in a production environment.**

      For example, a functional check might trigger a node panic or a failover to another node.

      ```
      phys-schost# cluster list-checks -v -C check-ID
      ```

      `-C` *check-ID* Specifies a specific check.

   c. **If the functional check that you want to perform might interrupt cluster functioning, ensure that the cluster is not in production.**

   d. **Start the functional check.**

      ```
      phys-schost# cluster check -v -k functional -C check-ID -o outputdir
      ```

      `-k functional` Specifies running functional validation checks.

      Respond to prompts from the check to confirm that the check should run, and for any information or actions you must perform.

   e. **Repeat Step c and Step d for each remaining functional check to run.**

Note - For record-keeping purposes, specify a unique *outputdir* subdirectory name for each check you run. If you reuse an *outputdir* name, output for the new check overwrites the existing contents of the reused *outputdir* subdirectory.

---

6. **If you have a zone cluster configured, verify the configuration of the zone cluster to see if a zone cluster can be installed.**

   ```
   phys-schost# clzonecluster verify zone-cluster-name
   ```

7. **Make a recording of the cluster configuration for future diagnostic purposes.**

   See "How to Record Diagnostic Data of the Cluster Configuration" in *Oracle Solaris Cluster 4.3 Software Installation Guide*.

Example 7 Checking the Global Cluster Configuration With All Basic 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 -n 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 8 Listing Interactive Validation Checks

The following example lists all interactive checks that are available to run on the cluster. Example output shows a sampling of possible checks; actual available checks vary for each configuration.

```
# cluster list-checks -k interactive
 Some checks might take a few moments to run (use -v to see progress)...
 I6994574:(Moderate)Fix for GLDv3 interfaces on cluster transport vulnerability applied?
```

Example 9 Running a Functional Validation Check

The following example first shows the verbose listing of functional checks. The verbose description is then listed for the check F6968101, which indicates that the check would disrupt cluster services. The cluster is taken out of production. The functional check is then run with verbose output logged to the funct.test.F6968101.12Jan2011 subdirectory. Example output shows a sampling of possible checks; actual available checks vary for each configuration.

```
# cluster list-checks -k functional
 F6968101: (Critical) Perform resource group switchover
 F6984120: (Critical) Induce cluster transport network failure - single adapter.
 F6984121: (Critical) Perform cluster shutdown
 F6984140: (Critical) Induce node panic
# cluster list-checks -v -C F6968101
 F6968101: (Critical) Perform resource group switchover
Keywords: SolarisCluster3.x, functional
Applicability: Applicable if multi-node cluster running live.
Check Logic: Select a resource group and destination node. Perform
'clresourcegroup switch' on specified resource group
either to specified node or to all nodes in succession.
Version: 1.2
Revision Date: 12/10/10
```

Take the cluster out of production

```
# cluster list-checks -k functional -C F6968101 -o funct.test.F6968101.12Jan2011
F6968101
initializing...
initializing xml output...
loading auxiliary data...
starting check run...
 pschost1, pschost2, pschost3, pschost4: F6968101.... starting:
Perform resource group switchover


  ============================================================

   >>> Functional Check

    'Functional' checks exercise cluster behavior. It is recommended that you
    do not run this check on a cluster in production mode.' It is recommended
    that you have access to the system console for each cluster node and
    observe any output on the consoles while the check is executed.

    If the node running this check is brought down during execution the check
    must be rerun from this same node after it is rebooted into the cluster in
    order for the check to be completed.
```

```
    Select 'continue' for more details on this check.

          1) continue
          2) exit

          choice: 1

  ============================================================

   >>> Check Description <<<
```

Follow onscreen directions

Example 10 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 -n 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 SolarisCluster 4.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` command includes checks that examine the `/etc/vfstab` file for configuration errors with the cluster file system and its global mount points. See the `cluster`(1CL) man page for more information.

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

---

1. **Assume the `root` role 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 11 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 -n 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 4.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 4.x node.
ANALYSIS : This server may not been qualified to be used as an Oracle Solaris Cluster
 4.x node.
Only servers that have been qualified with Oracle Solaris Cluster 4.0 are supported as
Oracle Solaris Cluster 4.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 4.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`
- `clsnmpmib`
- `clnsmpuser`
- `cltelemetryattribute`
- `cluster`
- `clzonecluster`

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 12 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
```

♦ ♦ ♦ C H A P T E R 2

Oracle Solaris Cluster and User Rights

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

Setting Up and Assigning Rights Profiles With Oracle Solaris Cluster

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

Task	Instructions
Learn more about user rights	Chapter 1, "About Using Rights to Control Users and Processes" in *Securing Users and Processes in Oracle Solaris 11.3*
Set up, manage elements, and use rights	Chapter 3, "Assigning Rights in Oracle Solaris" in *Securing Users and Processes in Oracle Solaris 11.3*
Learn more about rights elements and tools	Chapter 8, "Reference for Oracle Solaris Rights" in *Securing Users and Processes in Oracle Solaris 11.3*

Oracle Solaris Cluster Authorizations

Selected Oracle Solaris Cluster commands and options that you issue at the command line use one or more authorizations. Oracle Solaris Cluster commands and options that require authorization will require an Oracle Solaris Cluster management rights profile that includes one or more of the following authorization levels. Oracle Solaris Cluster management rights profiles apply to nodes in a global cluster and in a zone 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 create, delete, and change properties of a cluster object.

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

Oracle Solaris Cluster Management Rights Profiles

Oracle Solaris Cluster provides rights profiles that collect just the rights that are required for a particular Oracle Solaris Cluster administrative task. You can assign these rights profiles to users or to roles, to give them specific Oracle Solaris Cluster administrator rights.

Rights Profile	Authorizations and Security Attributes	Rights Granted

Oracle Solaris Cluster Commands	A list of Oracle Solaris Cluster commands that run with the `euid=0` security attribute.	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:
		■ `list`
		■ `show`
		■ `status`
		`scha_control`
		`scha_resource_get`
		`scha_resource_setstatus`

Rights Profile	Authorizations and Security Attributes	Rights Granted
		scha_resourcegroup_get scha_resourcetype_get
Basic Oracle Solaris User	This existing Oracle Solaris rights profile contains Oracle Solaris authorizations, as well as the following authorization:	solaris.cluster.read – Perform list, show, and other read operations for Oracle Solaris Cluster commands, as well as access the Oracle Solaris Cluster Manager browser interface.
Cluster Operation	This rights profile is specific to Oracle Solaris Cluster software and contains the following authorizations:	solaris.cluster.read – Perform list, show, export, status, and other read operations., as well as access the Oracle Solaris Cluster Manager browser interface. solaris.cluster.admin – 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. **Note -** To enable private network communication when you use Oracle Solaris Cluster Manager to create a zone cluster, you must run the wizard as a role that also includes the Network Management rights profile.
Cluster Management	This rights profile contains the same authorizations that the Cluster Operation profile contains, as well as the solaris.cluster.modify authorization.	Perform the same operations that the Cluster Operation role identity can perform, as well as change properties of a cluster object.

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

Use this task to create a new role with an Oracle Solaris Cluster Management rights profile and to assign users to this new role.

▼ How to Create and Assign a Role From the Command Line

1. **Select a method for creating a role:**

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

     Use this method only for emergencies.

- For roles in a name service, use the `roleadd` and `rolemod` commands to specify the new role and its attributes. For more information, see the `roleadd`(1M) and `rolemod`(1M) man pages.

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

2. **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 13 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 Rights Properties

You can modify a user's rights properties by using either the user accounts tool or the command line. Choose one of the following procedures.

- “How to Modify a User's Rights Properties by Using the User Accounts Tool” on page 69
- “How to Modify a User's Rights Properties From the Command Line” on page 70

▼ How to Modify a User's Rights 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 System 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 “Using Your Assigned Administrative Rights” in *Securing Users and Processes in Oracle Solaris 11.3*. 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 Rights Properties From the Command Line

1. **Assume a role that provides `solaris.cluster.modify` authorization.**

2. **Choose the appropriate command:**

   - To change user properties that are assigned to a user who is defined in the local scope or in an LDAP repository, use the `usermod` command. For more information, see the `usermod`(1M) man page.
   - 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 manage roles locally or in a name service such as an LDAP repository, use the `roleadd` or `rolemod` commands. For more information, see the `roleadd`(1M) or `rolemod`(1M) man pages.

     These commands require authentication as the `root` role that is capable of changing user files. You can apply these commands to all name services. See "Commands That Are Used for Managing Users, Roles, and Groups" in *Managing User Accounts and User Environments in Oracle Solaris 11.3*.

   The Forced Privilege and Stop Rights profiles that ship with Oracle Solaris 11 cannot be modified.

CHAPTER 3

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.

- "Overview of Shutting Down and Booting a Cluster" on page 71
- "Shutting Down and Booting a Single Node in a Cluster" on page 87
- "Repairing a Full `/var` File System" on page 102

For a high-level description of the related procedures in this chapter, see "How to Boot a Node in Noncluster Mode" on page 99 and Table 4, "Task Map: Shutting Down and Booting a Node," on page 88.

Overview of Shutting Down and Booting a Cluster

The Oracle Solaris Cluster `cluster 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.) For more information, see the `cluster`(1CL) man page.

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 evacuate` command to shut down individual nodes. For more information, see "How to Shut Down a Cluster" on page 73, "Shutting Down and Booting a Single Node in a Cluster" on page 87, or the `clnode(1CL)` man page.

You can also evacuate a node by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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. For more information about GRUB based booting, see "Booting a System" in *Booting and Shutting Down Oracle Solaris 11.3 Systems*. The `clzonecluster halt` command performs the `zoneadm -z` *zone-cluster-name* `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 99 for more information.

TABLE 3 Task List: Shutting Down and Booting a Cluster

Task	Instructions
Stop the cluster.	"How to Shut Down a Cluster" on page 73
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 75
Reboot the cluster.	"How to Reboot a Cluster" on page 80

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. **(x86 only) 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. **Assume a role that provides `solaris.cluster.admin` 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 zone-cluster-name
     ```

   - **Shut down all zone clusters.**

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

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

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.

- **Check the status of one or more global-cluster nodes from another global-cluster node which is still up and running in the cluster.**

```
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 14 Shutting Down a Zone Cluster

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

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

Example 15 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 16 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 75 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 "Creating and Configuring a Zone Cluster" in *Oracle Solaris Cluster 4.3 Software Installation Guide* or use the Oracle Solaris Cluster Manager browser interface to create the zone cluster.

---

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.

     For more information about GRUB based booting, see "Booting a System" in *Booting and Shutting Down Oracle Solaris 11.3 Systems*.

   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 zone-cluster-name
     ```

   - **If you have more than one zone cluster, you can boot all zone clusters. Use the plus sign (+) instead of the *zone-cluster-name*.**

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

   The `cluster status` command reports the global-cluster nodes' status.

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

   When you run the `clzonecluster status` 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 102. For more information, see the `clzonecluster(1CL)` man page.

Example 17 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.
...
```

Example 18 x86: Booting a Cluster

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

```
ATI RAGE SDRAM BIOS P/N GR-xlint.007-4.330
*   BIOS Lan-Console 2.0
Copyright (C) 1999-2001  Intel Corporation
MAC ADDR: 00 02 47 31 38 3C
AMIBIOS (C)1985-2002 American Megatrends Inc.,
```

```
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064
2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

Press <F2> to enter SETUP, <F12> Network

Adaptec AIC-7899 SCSI BIOS v2.57S4
(c) 2000 Adaptec, Inc. All Rights Reserved.
Press <Ctrl><A> for SCSISelect(TM) Utility!

Ch B,  SCSI ID: 0 SEAGATE  ST336605LC        160
SCSI ID: 1 SEAGATE  ST336605LC        160
SCSI ID: 6 ESG-SHV  SCA HSBP M18      ASYN
Ch A,  SCSI ID: 2 SUN      StorEdge 3310     160
SCSI ID: 3 SUN      StorEdge 3310     160

AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064

2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

SunOS - Intel Platform Edition             Primary Boot Subsystem, vsn 2.0

Current Disk Partition Information

Part#   Status    Type      Start       Length
================================================
1     Active   X86 BOOT     2428       21852
2              SOLARIS     24280     71662420
3              <unused>
4              <unused>
Please select the partition you wish to boot: *       *

Solaris DCB
```

```
loading /solaris/boot.bin

SunOS Secondary Boot version 3.00

Solaris Intel Platform Edition Booting System

Autobooting from bootpath: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/
pci8086,341a@7,1/sd@0,0:a

If the system hardware has changed, or to boot from a different
device, interrupt the autoboot process by pressing ESC.
Press ESCape to interrupt autoboot in 2 seconds.
Initializing system
Please wait...
Warning: Resource Conflict - both devices are added

NON-ACPI device: ISY0050
Port: 3F0-3F5, 3F7; IRQ: 6; DMA: 2
ACPI device: ISY0050
Port: 3F2-3F3, 3F4-3F5, 3F7; IRQ: 6; DMA: 2

<<< 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:
Size: 275683 + 22092 + 150244 Bytes
/platform/i86pc/kernel/unix loaded - 0xac000 bytes used
SunOS Release 5.9 Version Generic_112234-07 32-bit
Copyright 1983-2003 Sun Microsystems, Inc.  All rights reserved.
Use is subject to license terms.
configuring IPv4 interfaces: e1000g2.
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) with votecount = 1 added.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) with votecount = 1 added.
NOTICE: CMM: Quorum device 1 (/dev/did/rdsk/d1s2) added; votecount = 1, bitmask
of nodes with configured paths = 0x3.
NOTICE: clcomm: Adapter e1000g3 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being initiated
```

```
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 online
NOTICE: clcomm: Adapter e1000g0 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being constructed
NOTICE: CMM: Node phys-schost-1: attempting to join cluster.
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being initiated
NOTICE: CMM: Quorum device /dev/did/rdsk/d1s2: owner set to node 1.
NOTICE: CMM: Cluster has reached quorum.
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) is up; new incarnation number = 1068496374.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) is up; new incarnation number = 1068496374.
NOTICE: CMM: Cluster members: phys-schost-1 phys-schost-2.
NOTICE: CMM: node reconfiguration #1 completed.
NOTICE: CMM: Node phys-schost-1: joined cluster.
```

▼ 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. **Assume a role that provides `solaris.cluster.admin` 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 zone-cluster-name
```

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.

  For more information about GRUB based booting, see "Booting a System" in *Booting and Shutting Down Oracle Solaris 11.3 Systems*.

- 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 zone-cluster-name
```

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

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 102.

Example 19 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 20 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:
```

Example 21 x86: Rebooting a Cluster

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

```
# 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

ATI RAGE SDRAM BIOS P/N GR-xlint.007-4.330
*                                         BIOS Lan-Console 2.0
Copyright (C) 1999-2001  Intel Corporation
MAC ADDR: 00 02 47 31 38 3C
AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064
```

```
2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

Press <F2> to enter SETUP, <F12> Network

Adaptec AIC-7899 SCSI BIOS v2.57S4
(c) 2000 Adaptec, Inc. All Rights Reserved.
Press <Ctrl><A> for SCSISelect(TM) Utility!

Ch B,  SCSI ID: 0 SEAGATE  ST336605LC        160
SCSI ID: 1 SEAGATE  ST336605LC        160
SCSI ID: 6 ESG-SHV  SCA HSBP M18      ASYN
Ch A,  SCSI ID: 2 SUN      StorEdge 3310     160
SCSI ID: 3 SUN      StorEdge 3310     160

AMIBIOS (C)1985-2002 American Megatrends Inc.,
Copyright 1996-2002 Intel Corporation
SCB20.86B.1064.P18.0208191106
SCB2 Production BIOS Version 2.08
BIOS Build 1064

2 X Intel(R) Pentium(R) III CPU family      1400MHz
Testing system memory, memory size=2048MB
2048MB Extended Memory Passed
512K L2 Cache SRAM Passed
ATAPI CD-ROM SAMSUNG CD-ROM SN-124

SunOS - Intel Platform Edition             Primary Boot Subsystem, vsn 2.0

Current Disk Partition Information

Part#   Status    Type      Start       Length
================================================
1     Active   X86 BOOT     2428       21852
2              SOLARIS     24280     71662420
3              <unused>
4              <unused>
Please select the partition you wish to boot: *       *

Solaris DCB

loading /solaris/boot.bin

SunOS Secondary Boot version 3.00
```

```
Solaris Intel Platform Edition Booting System

Autobooting from bootpath: /pci@0,0/pci8086,2545@3/pci8086,1460@1d/
pci8086,341a@7,1/sd@0,0:a

If the system hardware has changed, or to boot from a different
device, interrupt the autoboot process by pressing ESC.
Press ESCape to interrupt autoboot in 2 seconds.
Initializing system
Please wait...
Warning: Resource Conflict - both devices are added

NON-ACPI device: ISY0050
Port: 3F0-3F5, 3F7; IRQ: 6; DMA: 2
ACPI device: ISY0050
Port: 3F2-3F3, 3F4-3F5, 3F7; IRQ: 6; DMA: 2

<<< 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
Size: 275683 + 22092 + 150244 Bytes
/platform/i86pc/kernel/unix loaded - 0xac000 bytes used
SunOS Release 5.9 Version Generic_112234-07 32-bit
Copyright 1983-2003 Sun Microsystems, Inc.  All rights reserved.
Use is subject to license terms.
configuring IPv4 interfaces: e1000g2.
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) with votecount = 1 added.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) with votecount = 1 added.
NOTICE: CMM: Quorum device 1 (/dev/did/rdsk/d1s2) added; votecount = 1, bitmask
of nodes with configured paths = 0x3.
NOTICE: clcomm: Adapter e1000g3 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being initiated
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 online
NOTICE: clcomm: Adapter e1000g0 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being constructed
NOTICE: CMM: Node phys-schost-1: attempting to join cluster.
```

```
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being initiated
NOTICE: CMM: Quorum device /dev/did/rdsk/d1s2: owner set to node 1.
NOTICE: CMM: Cluster has reached quorum.
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) is up; new incarnation number = 1068496374.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) is up; new incarnation number = 1068496374.
NOTICE: CMM: Cluster members: phys-schost-1 phys-schost-2.
NOTICE: CMM: node reconfiguration #1 completed.
NOTICE: CMM: Node phys-schost-1: joined cluster.
WARNING: mod_installdrv: no major number for rsmrdt
ip: joining multicasts failed (18) on clprivnet0 - will use link layer
broadcasts for multicast
The system is coming up.  Please wait.
checking ufs filesystems
/dev/rdsk/c1t0d0s5: is clean.
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 online
NIS domain name is dev.eng.mycompany.com
starting rpc services: rpcbind keyserv ypbind done.
Setting netmask of e1000g2 to 192.168.255.0
Setting netmask of e1000g3 to 192.168.255.128
Setting netmask of e1000g0 to 192.168.255.128
Setting netmask of clprivnet0 to 192.168.255.0
Setting default IPv4 interface for multicast: add net 224.0/4: gateway phys-schost-1
syslog service starting.
obtaining access to all attached disks


*****************************************************************************
*
* The X-server can not be started on display :0...
*
*****************************************************************************
volume management starting.
Starting Fault Injection Server...
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 or a zone-cluster node. 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 more information, see the `clnode`(1CL), `shutdown`(1M), and `clzonecluster`(1CL) man pages.

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

TABLE 4 Task Map: Shutting Down and Booting a Node

Task	Tool	Instructions
Stop a node.	For a global-cluster node, use the `clnode evacuate` and `shutdown` commands. For a zone-cluster node, use the `clzonecluster halt` command.	"How to Shut Down a Node" on page 88
Start a node. The node must have a working connection to the cluster interconnect to attain cluster membership.	For a global-cluster node, use the `boot` or `b` command. For a zone-cluster node, use the `clzonecluster boot` command.	"How to Boot a Node" on page 92
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 the `clnode evacuate` and `shutdown` commands, followed by `boot` or `b`. For a zone-cluster node, use the `clzonecluster reboot` command.	"How to Reboot a Node" on page 96
Boot a node so that the node does not participate in cluster membership.	For a global-cluster node, use `clnode evacuate` and `shutdown` commands, followed by `boot -x` 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 99

▼ 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.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to evacuate a global-cluster node and switch all resource groups and device groups to the next-preferred node. You can also shut down a zone-cluster node. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

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. **Assume a role that provides `solaris.cluster.admin` 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 zone-cluster-name
   ```

   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 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.**

   Execute the shutdown command on 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 22 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 phys-schost-1
phys-schost# shutdown -g0 -y -i0
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 23 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 -i0
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 24 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 92 to restart a global-cluster node that was shut down.

How to Boot 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.

Note - You can also boot a zone cluster node by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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.

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.

If you are running Oracle Solaris Cluster in a guest domain, rebooting the control or I/O domain can have an impact on the running guest domain, including the domain going down. You should rebalance the workload to other nodes and stop the guest domain running Oracle Solaris Cluster before you reboot the control or I/O domain.

When a control or I/O domain is rebooted, heartbeats are not received or sent by the guest domain. This causes split brain and a cluster reconfiguration to occur. Since the control or I/O domain is rebooting, the guest domain cannot access any shared devices. The other cluster nodes will fence this guest domain from the shared devices. When the control or I/O domain finishes its reboot, I/O resumes on the guest domain and any I/O to shared storage causes the guest domain to panic because it has been fenced off the shared disks as part of the cluster reconfiguration. You can mitigate this issue if a guest is employing two I/O domains for redundancy and you reboot the I/O domains one at a time.

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.

     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 zone-cluster-name
     ```

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 102.

Example 25 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:
```

Example 26 x86: Booting a Cluster Node

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

```
                     <<< 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: Size: 276915 + 22156 + 150372 Bytes
/platform/i86pc/kernel/unix loaded - 0xac000 bytes used
SunOS Release 5.9 Version on81-feature-patch:08/30/2003 32-bit
```

```
Copyright 1983-2003 Sun Microsystems, Inc.  All rights reserved.
Use is subject to license terms.
configuring IPv4 interfaces: e1000g2.
Hostname: phys-schost-1
Booting as part of a cluster
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) with votecount = 1 added.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) with votecount = 1 added.
NOTICE: CMM: Quorum device 1 (/dev/did/rdsk/d1s2) added; votecount = 1, bitmask
of nodes with configured paths = 0x3.
WARNING: CMM: Initialization for quorum device /dev/did/rdsk/d1s2 failed with
error EACCES. Will retry later.
NOTICE: clcomm: Adapter e1000g3 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being constructed
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 being initiated
NOTICE: clcomm: Path phys-schost-1:e1000g3 - phys-schost-2:e1000g3 online
NOTICE: clcomm: Adapter e1000g0 constructed
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being constructed
NOTICE: CMM: Node phys-schost-1: attempting to join cluster.
WARNING: CMM: Reading reservation keys from quorum device /dev/did/rdsk/d1s2
failed with error 2.
NOTICE: CMM: Cluster has reached quorum.
NOTICE: CMM: Node phys-schost-1 (nodeid = 1) is up; new incarnation number =
1068503958.
NOTICE: CMM: Node phys-schost-2 (nodeid = 2) is up; new incarnation number =
1068496374.
NOTICE: CMM: Cluster members: phys-schost-1 phys-schost-2.
NOTICE: CMM: node reconfiguration #3 completed.
NOTICE: CMM: Node phys-schost-1: joined cluster.
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 being initiated
NOTICE: clcomm: Path phys-schost-1:e1000g0 - phys-schost-2:e1000g0 online
NOTICE: CMM: Retry of initialization for quorum device /dev/did/rdsk/d1s2 was
successful.
WARNING: mod_installdrv: no major number for rsmrdt
ip: joining multicasts failed (18) on clprivnet0 - will use link layer
broadcasts for multicast
The system is coming up.  Please wait.
checking ufs filesystems
/dev/rdsk/c1t0d0s5: is clean.
NIS domain name is dev.eng.mycompany.com
starting rpc services: rpcbind keyserv ypbind done.
Setting netmask of e1000g2 to 192.168.255.0
Setting netmask of e1000g3 to 192.168.255.128
Setting netmask of e1000g0 to 192.168.255.128
Setting netmask of clprivnet0 to 192.168.255.0
Setting default IPv4 interface for multicast: add net 224.0/4: gateway phys-schost-1
syslog service starting.
obtaining access to all attached disks
```

```
****************************************************************************
*
* The X-server can not be started on display :0...
*
****************************************************************************
volume management starting.
Starting Fault Injection Server...
The system is ready.

phys-schost-1 console login:
```

How to Reboot a Node

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

Note - You can also reboot a zone-cluster node by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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 a method for any resource times out and cannot be killed, the node will be rebooted only if the resource's `Failover_mode` property is set to `HARD`. If the `Failover_mode` property is set to any other value, the node will not be rebooted.

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.

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. **Assume a role that provides `solaris.cluster.admin` 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 zones on the specified node to the next-preferred global zone on other nodes.

   Note - To shut down a single node, use the `shutdown -g0 -y -i6` command. To shut down multiple nodes at the same time, use the `shutdown -g0 -y -i0` command to halt the nodes. After all the nodes have halted, use the `boot` command on all nodes to boot them back in to the cluster.

   - On a SPARC based system, run the following commands to reboot a single node.

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

     ```
     phys-schost# shutdown -g0 -y -i6
     ```

   - On an x86 based system, run the following commands to reboot a single node.

     ```
     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.

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

     ```
     phys-schost# clzonecluster reboot - node zone-cluster-name
     ```

   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 27 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 28 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 updating 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. **Assume a role that provides `solaris.cluster.admin` 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 or the global-cluster 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 global zones on the specified node to the next-preferred global zones on other nodes.

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

     ```
     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 zone-cluster-name
     ```

     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.

      For more information about GRUB based booting, see "Booting a System" in *Booting and Shutting Down Oracle Solaris 11.3 Systems*.

   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.

   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/kernel/$ISADIR/unix -B $ZFS-BOOTFS -x
```

d. **Press the Enter key to accept the change and return to the boot parameters screen.**
The screen displays the edited command.

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 29 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 start on that node at the next boot up. 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 "System Message Formats" in *Troubleshooting System Administration Issues in Oracle Solaris 11.3* for more information.

1. **Assume the `root` role 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.

♦ ♦ ♦ CHAPTER 4

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*

You must understand host-based, application-based, and storage-based data replication before you can select the replication approach that best serves your cluster. For more information about using the Oracle Solaris Cluster Geographic Edition framework to manage your data replication for disaster recovery, see the *Oracle Solaris Cluster 4.3 Geographic Edition Overview*.

This chapter contains the following section:

- "Understanding Data Replication" on page 103
- "Using Storage-Based Data Replication Within a Campus Cluster" on page 106

Understanding Data Replication

Oracle Solaris Cluster software supports host-based and storage-based 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, or between a cluster and a host that is not in a cluster, is the Availability Suite feature of Oracle Solaris.

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, such as Oracle RAC.

  - For more information about using host-based data replication between two clusters, see *Oracle Solaris Cluster Geographic Edition Data Replication Guide for Oracle Solaris Availability Suite*.
  - To see an example of host-based replication that does *not* use the Oracle Solaris Cluster Geographic Edition framework, see Appendix A, Appendix A, "Deployment Example: Configuring Host-Based Data Replication Between Clusters With Availability Suite Software".

- *Application-based data replication* uses an application's built-in mechanisms for replicating application data to maintain synchronized copies of the application on the Geographic Edition partner clusters. Different applications offer a variety of configurations for this setup, which allows an array of options on how to replicate that data and how to use the data on the secondary system.
  - For more information about using application-based data replication between two or more clusters and the Oracle Solaris Cluster Geographic Edition product that automates the process, see *Oracle Solaris Cluster 4.3 Geographic Edition Overview*.
  - Application-based replication is not supported for a campus cluster environment.
- *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. An example of storage-based software that can replicate data inside a cluster or between clusters is 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 Campus Cluster" on page 106.
  - 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 4.3 Geographic Edition Overview*.

Supported Data Replication Methods

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

- **Replication Between Clusters** – For disaster recovery using Oracle Solaris Cluster Geographic Edition software, you can use host-based, application-based, or storage-based replication to perform data replication between clusters. You can manage all supported types of replication with Geographic Edition software.

  Alternatively, you can use the Availability Suite feature of Oracle Solaris to perform data replication between clusters without also using Geographic Edition software. For more information, see Appendix A, "Deployment Example: Configuring Host-Based Data Replication Between Clusters With Availability Suite Software".

Type of Replication	Supported Data Replication Products
Host-Based Replication	Availability Suite feature of Oracle Solaris. ZFS snapshots
Application-Based Replication	MySQL Oracle Data Guard Oracle GoldenGate
Storage-Based Replication	EMC Symmetrix Remote Data Facility (SRDF). Hitachi TrueCopy and Hitachi Universal Replicator. Oracle ZFS Storage Appliance.

  For more information, see "Data Replication" in *Oracle Solaris Cluster 4.3 Geographic Edition Overview*.

  If you want to use host-based replication without Oracle Solaris Cluster Geographic Edition software, see the instructions in Appendix A, "Deployment Example: Configuring Host-Based Data Replication Between Clusters With Availability Suite Software".

- **Replication Within a Campus Cluster** – This method is used as a replacement for host-based mirroring using Solaris Volume Manager or ZFS.

Type of Replication	Supported Data Replication Products
Storage-Based Replication	EMC Symmetrix Remote Data Facility (SRDF).

Using Storage-Based Data Replication Within a Campus Cluster

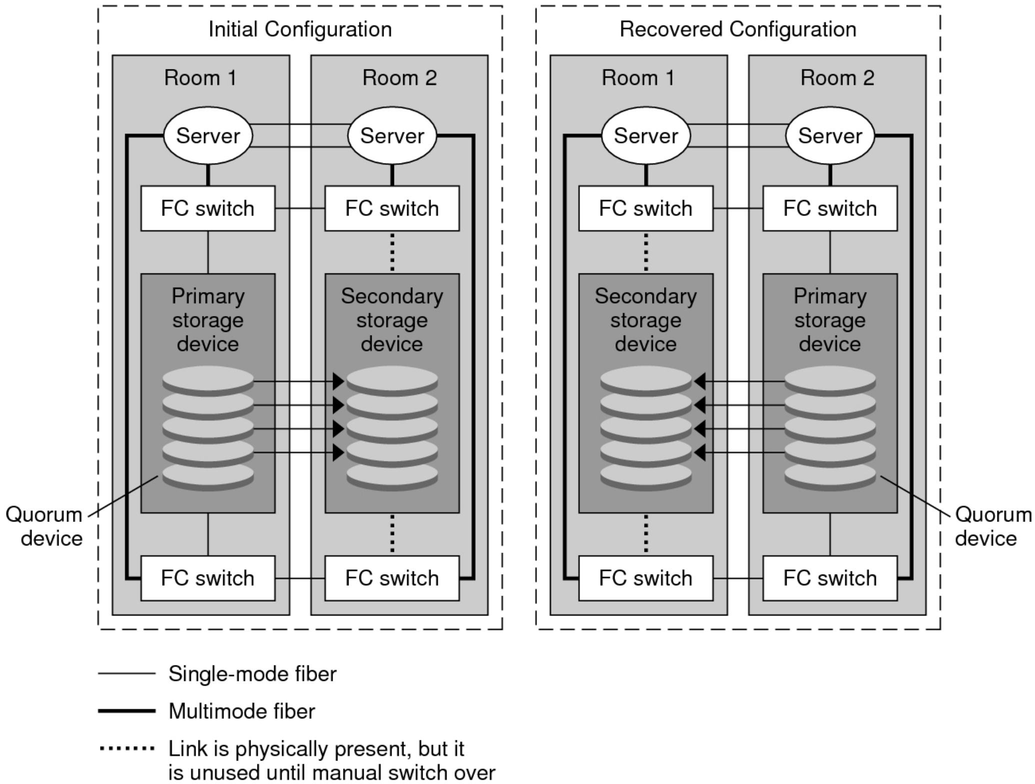
Storage-based data replication uses software installed on the storage device to manage the replication within a cluster, called a campus cluster. Such software is specific to your particular storage device, and is not used for disaster recovery. Refer to the documentation that shipped with your storage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

This section describes storage-based data replication as used in a campus cluster. Figure 1, “Two-Room Configuration With Storage-Based Data Replication,” on page 107 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 1, “Two-Room Configuration With Storage-Based Data Replication,” on page 107 illustrates that the quorum device is on an unreplicated volume. A replicated volume cannot be used as a quorum device.

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



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.

If contact with the remote storage device is lost, ensure that an application that is running on the primary cluster is not blocked by specifying a `Fence_level` of `never` or `async`. If you specify a `Fence_level` of `data` or `status`, the primary storage device refuses updates if the updates cannot be copied to the remote storage device.

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

- If configuring the cluster for automatic failover, use synchronous replication.

  For instructions on configuring the cluster for automatic failover of replicated volumes, see "Configuring and Administering Storage-Based Replicated Devices" on page 113. For details on the requirements to design a campus cluster, see "Shared Data Storage" in *Oracle Solaris Cluster Hardware Administration Manual*.
- 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.
- 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 allows the user to define groups of replicated devices. Each replication device group requires an Oracle Solaris Cluster device group with the same name.
- For a three-site, or three-data-center, configuration using EMC SRDF with concurrent or cascaded RDF devices, you must add the following entry in the Solutions Enabler SYMCLI options file on all participating cluster nodes:

  ```
  SYMAPI_2SITE_CLUSTER_DG=device-group:rdf-group-number
  ```

  This entry enables the cluster software to automate the movement of the application between the two SRDF synchronous sites. The *rdf-group-number* in the entry represents the RDF group that connects the host's local symmetrix to the second site's symmetrix.

  For more information about three-data-center configurations, see "Three-Data-Center (3DC) Topologies" in *Oracle Solaris Cluster 4.3 Geographic Edition Overview*.
- Oracle Real Application Clusters (Oracle RAC) is not supported with SRDF 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.

- Solaris Volume Manager for Sun Cluster multi-owner disksets are not supported.
- Do not use the Domino mode or Adaptive Copy mode in EMC SRDF. See "Using Storage-Based Data Replication Within a Campus Cluster" on page 106 for more information.

Manual Recovery Concerns When Using Storage-Based Data Replication Within a Campus 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 1, "Two-Room Configuration With Storage-Based Data Replication," on page 107), 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 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.

♦ ♦ ♦ CHAPTER 5

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.

For a high-level description of the related procedures in this chapter, see Table 7, "Task Map: Administering Device Groups," on page 128.

For conceptual information related to global devices, the global namespace, device groups, disk-path monitoring, and the cluster file system, see *Oracle Solaris Cluster 4.3 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` command. For more information, see the `metaset(1M)` man page.

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
```

Dynamic Reconfiguration With Global Devices

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

- All of the requirements, procedures, and restrictions that are documented for the Oracle Solaris dynamic reconfiguration feature also apply to Oracle Solaris Cluster dynamic reconfiguration support. The only exception is for the operating system quiescence operation. Therefore, review the documentation for the Oracle Solaris dynamic reconfiguration feature *before* using the dynamic reconfiguration feature with Oracle Solaris Cluster software. You should review in particular the issues that affect non-network IO devices during a dynamic reconfiguration detach operation.
- Oracle Solaris Cluster rejects dynamic reconfiguration remove-board operations on active devices in the primary node. Dynamic reconfiguration operations can be performed on inactive devices in the primary node and on any devices in the secondary nodes.
- After the dynamic reconfiguration operation, cluster data access continues as before.

- Oracle Solaris Cluster rejects dynamic reconfiguration operations that impact the availability of quorum devices. See "Dynamic Reconfiguration With Quorum Devices" on page 177 for more information.

Caution - If the current primary node fails while you are performing the dynamic reconfiguration 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 dynamic reconfiguration operations on global devices, complete the following steps in the order indicated.

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

Task	For Instructions
1. If a dynamic reconfiguration operation that affects an active device group must be performed on the current primary node, switch the primary and secondary nodes before performing the dynamic reconfiguration remove operation on the device	"How to Switch the Primary for a Device Group" on page 152
2. Perform the dynamic reconfiguration removal operation on the device being removed	Check the documentation that came with your system.

Configuring and 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 EMC Symmetrix Remote Data Facility software for storage-based replication.

Before you can replicate data with 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 updates 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 or raw-disk) must be the same as the name of the replicated device group.

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 6 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 114
Configure the DID device	"How to Configure DID Devices for Replication Using EMC SRDF" on page 116
Register the replicated group	"How to Add and Register a Device Group (Solaris Volume Manager)" on page 135
Verify the configuration	"How to Verify EMC SRDF Replicated Global Device Group Configuration" on page 118
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 124

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 or raw-disk) must be the same as the name of the replicated device group.

1. **Assume a role that provides `solaris.cluster.modify` authorization on all nodes connected to the storage array.**

2. **For a three-site or three-data-center implementation using concurrent SRDF or cascaded devices, set the `SYMAPI_2SITE_CLUSTER_DG` parameter.**

   Add the following entry to the Solutions Enabler options file on all participating cluster nodes:

   `SYMAPI_2SITE_CLUSTER_DG=`*device-group*`:`*rdf-group-number*

device-group	Specifies the name of the device group.
rdf-group-number	Specifies the RDF group that connects the host's local symmetrix to the second site's symmetrix.

   This entry enables the cluster software to automate the movement of the application between the two SRDF synchronous sites.

   For more information about three-data-center configurations, see "Three-Data-Center (3DC) Topologies" in *Oracle Solaris Cluster 4.3 Geographic Edition Overview*.

3. **On each node configured with the replicated data, discover the symmetrix device configuration.**

   This might take a few minutes.

   ```
   # /usr/symcli/bin/symcfg discover
   ```

4. **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.

   Note - If using concurrent RDF devices for a three-site or three-data-center implementation, add the following parameter to all `symrdf` commands:

   `-rdfg` *rdf-group-number*

   Specifying the RDF group number to the `symrdf` command ensures that the `symrdf` operation is directed to the correct RDF group.

   ---

5. **On each node configured with replicated devices, verify that data replication is set up correctly.**

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

6. **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
```

7. **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. Ensure that the specified DID device instances are replicas of each other and that they belong to the specified replication group.

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. **Assume a role that provides `solaris.cluster.modify` authorization on any node of the cluster.**

2. **Determine which DID devices correspond to the configured RDF1 and RDF2 devices.**

   ```
   # /usr/symcli/bin/symdg 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 `symdg -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.

   -t *replication-type*

   Specifies the replication type. For EMC SRDF, type **`SRDF`**.

   -g *replication-device-group*

   Specifies the name of the device group as shown in the `symdg show` command.

   -d *destination-instance*

   Specifies the DID instance that corresponds to the RDF1 device.

   source-instance

   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, perform the following additional steps.**

   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 `scdidadm -b` command to remove the existing information. The last step is to use the `cldevice combine` 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, ZFS, or raw-disk. For more information, consult the following:

- “How to Add and Register a Device Group (Solaris Volume Manager)” on page 135
- “How to Add and Register a Device Group (Raw-Disk)” on page 137
- “How to Add and Register a Replicated Device Group (ZFS)” on page 138

Caution - The name of the Oracle Solaris Cluster device group that you created (Solaris 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.**

```
# symdg -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.**

```
# symdg -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.

EXAMPLE 30 Creating Replica Pairs

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`.

Run the following commands on all nodes

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

# symdev list pd

Symmetrix ID: 000187990182

        Device Name          Directors                  Device
--------------------------- ------------ -------------------------------------
                                                                         Cap
Sym  Physical               SA :P DA :IT  Config        Attribute    Sts   (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, run the following commands

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

On all nodes on the RDF2 side, run the following commands

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

EXAMPLE 31 Verifying Data Replication Setup

The following commands are performed on one node of the cluster.

```
# 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-assoc'd RBCV's (RBCV RDF) :    0

    Standard (STD) Devices (1):
        {
        ----------------------------------------------------------------
                                                Sym               Cap
        LdevName            PdevName            Dev  Att. Sts     (MB)
        ----------------------------------------------------------------
        DEV001              /dev/rdsk/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
--------------------------------    ------------------------ ----- ------------
             ST                  LI      ST
Standard      A                   N       A
Logical       T  R1 Inv   R2 Inv  K       T  R1 Inv   R2 Inv       RDF Pair
Device  Dev   E  Tracks   Tracks  S Dev   E  Tracks   Tracks MDA   STATE
-------------------------------- -- ------------------------ ----- ------------

DEV001  0067 WD       0        0 RW 0067 RW       0        0 S..   Synchronized

Total          -------- --------           -------- --------
  MB(s)             0.0      0.0                0.0      0.0

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 32 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.

Run these commands on the RDF1 side

```
# symdg show dg1

Group Name:  dg1

    Group Type                                   : RDF1     (RDFA)
...
    Standard (STD) Devices (1):
        {
        ----------------------------------------------------------------
                                                 Sym               Cap
        LdevName           PdevName              Dev  Att. Sts     (MB)
        ----------------------------------------------------------------
        DEV001             /dev/rdsk/c5t6006048000018790002353594D303637d0s2 0067    RW
  4315
        }

    Device Group RDF Information
```

```
…
```

Obtain the corresponding DID

```
# cldevice list | grep c5t6006048000018790002353594D303637d0
217      pmoney1:/dev/rdsk/c5t6006048000018790002353594D303637d0 /dev/did/rdsk/d217
217      pmoney2:/dev/rdsk/c5t6006048000018790002353594D303637d0 /dev/did/rdsk/d217
```

List the corresponding DID

```
# cldevice show d217

=== DID Device Instances ===

DID Device Name:                      /dev/did/rdsk/d217
  Full Device Path:                      pmoney2:/dev/rdsk/
c5t6006048000018790002353594D303637d0
  Full Device Path:                      pmoney1:/dev/rdsk/
c5t6006048000018790002353594D303637d0
  Replication:                           none
  default_fencing:                       global

#
```

Run these commands on the RDF2 side

```
# symdg show dg1

Group Name:  dg1

    Group Type                                  : RDF2     (RDFA)
...
    Standard (STD) Devices (1):
        {
        ----------------------------------------------------------------
                                                  Sym               Cap
        LdevName           PdevName               Dev  Att. Sts     (MB)
        ----------------------------------------------------------------
        DEV001             /dev/rdsk/c5t6006048000018799018253594D303637d0s2 0067    WD
  4315
        }

    Device Group RDF Information
…
```

Obtain the corresponding DID

```
# cldevice list | grep c5t6006048000018799018253594D303637d0
108      pmoney4:/dev/rdsk/c5t6006048000018799018253594D303637d0 /dev/did/rdsk/d108
108      pmoney3:/dev/rdsk/c5t6006048000018799018253594D303637d0 /dev/did/rdsk/d108
```

List the corresponding DID

```
# cldevice show d108
```

```
=== DID Device Instances ===

DID Device Name:            /dev/did/rdsk/d108
  Full Device Path:               pmoney3:/dev/rdsk/
c5t6006048000018799018253594D303637d0
  Full Device Path:               pmoney4:/dev/rdsk/
c5t6006048000018799018253594D303637d0
  Replication:                    none
  default_fencing:                global

#
```

EXAMPLE 33 Combining DID instances

From the RDF2 side, type:

```
# cldevice combine -t srdf -g dg1 -d d217 d108
```

EXAMPLE 34 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/rdsk/d217
  Full Device Path:                      pmoney1:/dev/rdsk/
c5t6006048000018790002353594D303637d0
  Full Device Path:                      pmoney2:/dev/rdsk/
c5t6006048000018790002353594D303637d0
  Full Device Path:                      pmoney4:/dev/rdsk/
c5t6006048000018799018253594D303637d0
  Full Device Path:                      pmoney3:/dev/rdsk/
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 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 assume a role that provides `solaris.cluster.modify` 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 35 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   O 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  O 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 use UFS and can also be automatically mounted at boot. Cluster file systems are only visible from a node in a global cluster.

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` command is not supported on directories that are not empty. For more information, see the `unlink`(1M) man page.
- 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.

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. The `clsetup` utility 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.

Note - You can also bring a device group online and take it offline by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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.

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.

TABLE 7 Task Map: Administering Device Groups

Task	Instructions
Update the global-devices namespace without a reconfiguration reboot by using the `cldevice populate` command	"How to Update the Global-Devices Namespace" on page 129
Change the size of a `lofi` device that is used for the global-devices namespace	"How to Change the Size of a `lofi` Device That Is Used for the Global-Devices Namespace" on page 130
Move an existing global-devices namespace	"How to Migrate the Global-Devices Namespace From a Dedicated Partition to a `lofi` Device" on page 132 "How to Migrate the Global-Devices Namespace From a `lofi` Device to a Dedicated Partition" on page 133
Add Solaris Volume Manager disksets and register them as device groups by using the `metaset` command	"How to Add and Register a Device Group (Solaris Volume Manager)" on page 135
Add and register a raw-disk device group by using the `cldevicegroup` command	"How to Add and Register a Device Group (Raw-Disk)" on page 137
Add a named device group for ZFS by using the `cldevicegroup` command	"How to Add and Register a Replicated Device Group (ZFS)" on page 138
Remove Solaris Volume Manager device groups from the configuration by using the `metaset` and `metaclear` commands	"How to Remove and Unregister a Device Group (Solaris Volume Manager)" on page 141
Remove a node from all device groups by using the `cldevicegroup`, `metaset`, and `clsetup` commands	"How to Remove a Node From All Device Groups" on page 141
Remove a node from a Solaris Volume Manager device group by using the `metaset` command	"How to Remove a Node From a Device Group (Solaris Volume Manager)" on page 143
Remove a node from a raw-disk device group by using the `cldevicegroup` command	"How to Remove a Node From a Raw-Disk Device Group" on page 145

Task	Instructions
Change device group properties by using `clsetup` to generate `cldevicegroup`	"How to Change Device Group Properties" on page 147
Display device groups and properties by using the `cldevicegroup show` command	"How to List a Device Group Configuration" on page 151
Change the desired number of secondaries for a device group by using `clsetup` to generate `cldevicegroup`	"How to Set the Desired Number of Secondaries for a Device Group" on page 148
Switch the primary for a device group by using the `cldevicegroup switch` command	"How to Switch the Primary for a Device Group" on page 152
Put a device group in maintenance state by using the `metaset` command	"How to Put a Device Group in Maintenance State" on page 153

▼ 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. **Assume a role that provides `solaris.cluster.modify` authorization on any node of the cluster.**

2. **On each node in the cluster, run the `devfsadm` command.**

   You can run this command on all nodes in the cluster at the same time. For more information, see the `devfsadm`(1M) man page.

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 36 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. **Assume a role that provides `solaris.cluster.modify` 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 99.

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.

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 132
- "How to Migrate the Global-Devices Namespace From a lofi Device to a Dedicated Partition" on page 133

How to Migrate the Global-Devices Namespace From a Dedicated Partition to a `lofi` Device

1. **Assume the `root`role on the global-cluster 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 99.

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.**

```
# 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. **Assume the `root` role on the global-cluster 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 99.

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 -n` *node- 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/rdsk/d5s3`, the new entry to add to the`/etc/vfstab` file would then be as follows:

   ```
   /dev/did/dsk/d5s3 /dev/did/rdsk/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. **Boot into cluster mode and populate the global-devices namespace.**

    a. **From one node in the cluster, populate the global-devices namespace.**

    ```
    # cldevice populate
    ```

    b. **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 cldevice 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, 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 Manage or raw-disk) must be the same as the name of the replicated device group.

---

1. **Assume a role that provides `solaris.cluster.modify` 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 148 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/rdsk/d`*N*, when you add a drive to a disk set.

   In the following example, the entries for DID device `/dev/did/rdsk/d3` indicate that the drive is shared by `phys-schost-1` and `phys-schost-2`.

   ```
   === DID Device Instances ===
   DID Device Name:                                /dev/did/rdsk/d1
     Full Device Path:                               phys-schost-1:/dev/rdsk/c0t0d0
   DID Device Name:                                /dev/did/rdsk/d2
     Full Device Path:                               phys-schost-1:/dev/rdsk/c0t6d0
   DID Device Name:                                /dev/did/rdsk/d3
     Full Device Path:                               phys-schost-1:/dev/rdsk/c1t1d0
     Full Device Path:                               phys-schost-2:/dev/rdsk/c1t1d0
   …
   ```

6. **Add the drives to the disk set.**

   Use the full DID path name.

   ```
   # metaset -s setname -a /dev/did/rdsk/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 (`c`*N*`t`*X*`d`*Y*) 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 37 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/rdsk/d1` and `/dev/did/rdsk/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/rdsk/d1 /dev/did/rdsk/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 `cldevicegroup` command

Caution - If you are creating a device group on replicated devices, the name of the device group that you create (Solaris 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 devices d*N* and d*X*.

   ```
   phys-schost-1# cldevicegroup disable dsk/dN dsk/dX
   phys-schost-1# cldevicegroup offline dsk/dN dsk/dX
   phys-schost-1# cldevicegroup delete dsk/dN dsk/dX
   ```

2. **Create the new raw-disk device group, including the desired devices.**

The following command creates a global device group, *raw-disk-dg*, which contains devices `d`*N* and `d`*X*.

```
phys-schost-1# cldevicegroup create -n phys-schost-1,phys-schost-2 \
-t rawdisk -d dN,dX raw-disk-dg
```

3. **Verify the raw-disk device group you created.**

```
phys-schost-1# cldevicegroup show raw-disk-dg
```

How to Add and Register a Replicated Device Group (ZFS)

Use this procedure to create a replicated ZFS device group that is managed by HAStoragePlus.

To create a ZFS storage pool (zpool) that does not use HAStoragePlus, instead go to "How to Configure a Local ZFS Storage Pool Without HAStoragePlus" on page 139.

Before You Begin 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 or raw-disk) must be the same as the name of the replicated device group.

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 delete 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 -n pnode1,pnode2 -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.**

```
# clresourcegroup create -n pnode1,pnode2 migrate_srdfdg-rg
```

5. **Create an `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 in Step 2.

```
# clresource create -t HAStoragePlus -x globaldevicepaths=mypool \
-g migrate_srdfdg-rg hasp2migrate_mypool
```

6. **Set the `+++` value in the `rg_affinities` property from this resource group to the resource group you created in Step 4.**

```
# clresourcegroup create -n pnode1,pnode2 \
-p RG_affinities=+++migrate_srdfdg-rg oracle-rg
```

7. **Create an HAStoragePlus resource (`hasp-rs`) for the zpool you created in Step 3 in the resource group that you created in either Step 4 or Step 6.**

   Set the `resource_dependencies` property to the `hasp-rs` resource that you created in Step 5.

```
# clresource create -g oracle-rg -t HAStoragePlus -p zpools=mypool \
-p resource_dependencies=hasp2migrate_mypool \
-p ZpoolsSearchDir=/dev/did/dsk hasp2import_mypool
```

8. **Use the new resource group name where a device group name is required.**

How to Configure a Local ZFS Storage Pool Without HAStoragePlus

This procedure describes how to configure a ZFS storage pool (zpool) on a local device without configuring an HAStoragePlus resource.

Note - To configure a local zpool that uses an HAStoragePlus resource, instead go to "How to Add and Register a Replicated Device Group (ZFS)" on page 138.

---

1. **List the DID mappings and identify the local device to use.**

Choose a device that lists only the cluster node that will use the new zpool. Note both the c*N*t*X*d*Y* device name and the /dev/did/rdsk/d*N* DID device name.

```
phys-schost-1# cldevice show | grep Device
```

In the following example, the entries for DID devices /dev/did/rdsk/d1 and /dev/did/rdsk/d2 show that those drives are used only by phys-schost-1. For the examples in this procedure, DID device /dev/did/rdsk/d2 with device name c0t6d0 will be used and configured for cluster node phys-schost-1.

```
=== DID Device Instances ===
DID Device Name:                                /dev/did/rdsk/d1
  Full Device Path:                               phys-schost-1:/dev/rdsk/c0t0d0
DID Device Name:                                /dev/did/rdsk/d2
  Full Device Path:                               phys-schost-1:/dev/rdsk/c0t6d0
DID Device Name:                                /dev/did/rdsk/d3
  Full Device Path:                               phys-schost-1:/dev/rdsk/c1t1d0
  Full Device Path:                               phys-schost-2:/dev/rdsk/c1t1d0
…
```

2. **Determine the device group name of the DID device that you choose for the zpool.**

   The following example output shows that dsk/d2 is the device group name for DID device /dev/did/rdsk/d2s2. A relationship where the device group name is part of the DID device name is often the case, but not always. This device group has only one node, phys-schost-1, in its node list.

```
phys-schost-1# cldevicegroup show -v
…
Device Group Name:                              dsk/d2
  Type:                                            Disk
  failback:                                        false
  Node List:                                       phys-schost-1
  preferenced:                                     false
  localonly:                                       false
  autogen:                                         true
  numsecondaries:                                  1
  device names:                                    /dev/did/rdsk/d2s2
…
```

3. **Set the `localonly` property for the DID device.**

   Specify the device group name that you identified in Step 2. If you want to disable fencing for the device, also include default_fencing=nofencing in the command.

```
phys-schost-1# cldevicegroup set -p localonly=true \
-p autogen=true [-p default_fencing=nofencing] dsk/d2
```

For more information about `cldevicegroup` properties, see the `cldevicegroup`(1CL) man page.

4. **Verify the device settings.**

```
phys-schost-1# cldevicegroup show dsk/d2
```

5. **Create the zpool.**

```
phys-schost-1# zpool create localpool c0t6d0
```

6. **(Optional) Create a ZFS dataset.**

```
phys-schost-1# zfs create localpool/data
```

7. **Verify the new zpool.**

```
phys-schost-1# zpool list
```

Maintaining Device Groups

You can perform a variety of administrative tasks for your device groups. Some of these tasks can also be performed by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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 *Solaris Volume Manager Administration Guide* 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. **Assume a role that provides `solaris.cluster.modify` 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 143 for each device group of that type.**

4. **Determine the raw-device disk groups of which the node to be removed is a member.**

   ```
   # cldevicegroup list -v
   ```

5. **If any of the device groups listed in Step 4 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 145 for each of these device groups.**

6. **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. **Assume the `root` role 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 38 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

    Assume the root role 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 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. **Assume a role that provides `solaris.cluster.read` and `solaris.cluster.modify` 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 39 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/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 the `clsetup` utility to set or unset the `preferenced` attribute and the `failback` attribute for Solaris Volume Manager 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. **Assume a role that provides `solaris.cluster.read` and `solaris.cluster.modify` 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 for 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 for the option for changing key properties of a Solaris Volume Manager device group.**

   The Change Key Properties Menu is displayed.

5. **To change a device group property, type the number for option for changing the preferences 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 40 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 4.3 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 the `cldevicegroup`(1CL) man page 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.

1. **Assume a role that provides `solaris.cluster.read` and `solaris.cluster.modify` 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 for 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:                                   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 - Device group 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 129.

---

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 41 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 42 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 the `root` role 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 browser interface

  See Chapter 13, "Using the Oracle Solaris Cluster Manager Browser Interface" for more information.

`cldevicegroup show`

  Use `cldevicegroup show` to list the configuration for all device groups in the cluster.

`cldevicegroup show` *devicegroup*

  Use `cldevicegroup show` *devicegroup* to list the configuration of a single device group.

`cldevicegroup status` *devicegroup*

  Use `cldevicegroup status` *devicegroup* to determine the status of a single device group.

`cldevicegroup status +`

  Use `cldevicegroup status +` 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 43 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 44 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.

Note - You can also bring an inactive device group online by using the Oracle Solaris Cluster Manager browser interface. See the Oracle Solaris Cluster Manager online help for more information. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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. **Assume a role that provides `solaris.cluster.modify` 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 45 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 quiesced 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 is being repaired on another node.

For instructions on how to restore a corrupted diskset, see "Restoring a Corrupted Disk Set" on page 287.

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.

Note - You can also take an active device group offline by using the Oracle Solaris Cluster Manager browser interface. See the Oracle Solaris Cluster Manager online help for more information. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

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, manually import that disk set.**

   ```
   # 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.

---

3. **Complete the repair procedure that you need to perform.**

4. **Release ownership of the disk set.**

Caution - Before taking the device group out of maintenance state, you must release ownership of the disk set. Failure to release ownership can result in data loss.

```
# metaset -C release -s diskset
```

5. **Bring the device group online.**

```
# cldevicegroup online devicegroup
# cldevicegroup enable devicegroup
```

Example 46 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.]
For Solaris Volume Manager:
# metaset -C take -f -s dg-schost-1
    [Complete all necessary repair procedures.]
    [Release ownership.]
    For Solaris Volume Manager:
# metaset -C release -s dg-schost-1
    [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 156
- "How to Display the SCSI Protocol of a Single Storage Device" on page 157
- "How to Change the Default Global Fencing Protocol Settings for All Storage Devices" on page 157

- "How to Change the Fencing Protocol for a Single Storage Device" on page 159

▼ 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. **Assume a role that provides `solaris.cluster.read` 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 47 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
  clusterid:                                      0x4FES2C888
  installmode:                                    disabled
  heartbeat_timeout:                              10000
  heartbeat_quantum:                              1000
  private_netaddr:                                172.16.0.0
  private_netmask:                                255.255.111.0
  max_nodes:                                      64
  max_privatenets:                                10
  udp_session_timeout:                            480
  concentrate_load:                               False
  global_fencing:                                 prefer3
  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. **Assume a role that provides `solaris.cluster.read` 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 48 Displaying the SCSI Protocol of a Single Device

The following example displays the SCSI protocol for the device `/dev/rdsk/c4t8d0`.

```
# cldevice show /dev/rdsk/c4t8d0


=== DID Device Instances ===

DID Device Name:                                /dev/did/rdsk/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 159 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-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. **Assume a role that provides `solaris.cluster.modify` 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}
```

`-p global_fencing`

Sets the current global default fencing algorithm for all shared devices.

`prefer3`

Uses the SCSI-3 protocol for devices with two or more paths.

`pathcount`

Determines the fencing protocol by the number of DID paths that are attached to the shared device. The pathcount setting is used for quorum devices.

`nofencing`

Turns fencing off by setting the fencing status for all storage devices.

`nofencing-noscrub`

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 `nofencing-noscrub` option only for storage devices that have severe problems with SCSI reservations.

Example 49 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. **Assume a role that provides `solaris.cluster.modify` authorization.**
2. **Set the fencing protocol of the storage device.**

**cldevice set -p default_fencing ={pathcount | **
scsi3 | global | nofencing | nofencing-noscrub} *device*

`-p default_fencing`

Modifies the `default_fencing` property of the device.

`pathcount`

Determines the fencing protocol by the number of DID paths that are attached to the shared device.

`scsi3`

Uses the SCSI-3 protocol.

`global`

Uses the global default fencing setting. The global setting is used for non-quorum devices.

`nofencing`

Turns fencing off by setting the fencing status for the specified DID instance.

`nofencing-noscrub`

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 `nofencing-noscrub` option only for storage devices that have severe problems with SCSI reservations.

device

Specifies the name of the device path or device name.

For more information, see the `cluster`(1CL) man page.

Example 50 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 8 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 161
Remove a cluster file system	"How to Remove a Cluster File System" on page 164
Check global mount points in a cluster for consistency across nodes	"How to Check Global Mounts in a Cluster" on page 166

▼ 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:

- The `root` role privilege is established on a node in the cluster.
- Volume manager software be installed and configured on the cluster.
- A device group (such as a Solaris Volume Manager device group) or block disk slice exists on which to create the cluster file system.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to add a cluster file system to a zone cluster. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

If you used Oracle Solaris Cluster Manager to install data services, one or more cluster file systems already exist if there were sufficient shared disks on which to create the cluster file 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.

1. **Assume the `root` role on any node in the cluster.**

   Tip - For faster file system creation, become the `root` role on the current primary of the global device for which you create a file system.

2. **Create a UFS file system by using the `newfs` command.**

   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.

   ```
   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/rdsk/d1`	Raw disk device `d1` within the `nfs` disk set
None	`/dev/global/rdsk/d1s3`	Raw disk device `d1s3`

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

   A mount point is required *on each node*, even if the cluster file system 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/mount-point/
   ```

device-group

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

mount-point

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.

   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. **For each cluster file system, ensure that 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, no output is returned.

   For more information, see the `cluster`(1CL) man page.

6. **Mount the cluster file system from any node in the cluster.**

   ```
   phys-schost# mount /global/device-group/mountpoint/
   ```

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.

▼ 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 metadevice 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:

- The `root` role 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.

Note - You can also remove a zone-cluster file system by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume the `root` role 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 51 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/rdsk/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

```
# pfedit /etc/vfstab
#device           device          mount   FS      fsck    mount   mount
#to mount         to fsck         point   type    pass    at boot options
#
/dev/md/oracle/dsk/d1 /dev/md/oracle/rdsk/d1 /global/oracle/d1 ufs 2 yes global,logging
```

Save and exit

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 there are no errors, 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. **Assume the `root` role 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.

The following additional information is available:

Topic	Information
Conceptual information about the disk-path monitoring daemon	Chapter 3, "Key Concepts for System Administrators and Application Developers" in *Oracle Solaris Cluster 4.3 Concepts Guide*
Description of the `cldevice` command options and related commands	`cldevice`(1CL) man page
Tuning the `scdpmd` daemon	`scdpmd.conf`(4) man page
Logged errors that the `syslogd` daemon reports	`syslogd`(1M) man page

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 9 Task Map: Administering Disk-Path Monitoring

Task	Instructions
Monitor a disk path.	"How to Monitor a Disk Path" on page 167
Unmonitor a disk path.	"How to Unmonitor a Disk Path" on page 169
Print the status of faulted disk paths for a node.	"How to Print Failed Disk Paths" on page 170
Monitor disk paths from a file.	"How to Monitor Disk Paths From a File" on page 171
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 173 "How to Disable the Automatic Rebooting of a Node When All Monitored Shared-Disk Paths Fail" on page 174
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 171

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.

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 - You can also enable monitoring of a disk path by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume a role that provides `solaris.cluster.modify` 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 52 Monitoring a Disk Path on a Single Node

The following example monitors the `schost-1:/dev/did/rdsk/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/rdsk/d1   phys-schost-1 Ok
```

Example 53 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 54 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.

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 - You can also disable monitoring of a disk path by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume a role that provides `solaris.cluster.modify` 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 55 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.

1. **Assume the `root` role on any node in the cluster.**
2. **Print the faulted disk paths throughout the cluster.**

   `# cldevice status -s fail`

Example 56 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.

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. **Assume a role that provides `solaris.cluster.modify` 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 57 Monitoring 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 `deviceconfig` XML file shows that the path between `phys-schost–2` and `d3` is not currently monitored.

Export the current cluster configuration

```
# cldevice export -o deviceconfig
<?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>
```

Monitor the path by setting the monitored attribute to true

```
<?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>
```

Read the file and turn on monitoring

```
# cldevice monitor -i deviceconfig
```

Verify that the device is now monitored

```
# cldevice status
```

See Also For more detail about exporting a cluster configuration and using the resulting XML file to set a 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.

Note - You can also edit the `reboot_on_path_failure` node property by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **On any node in the cluster, assume a role that provides `solaris.cluster.modify` 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, assume a role that provides `solaris.cluster.modify` 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 +
```

♦ ♦ ♦ CHAPTER 6

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 4.3 Concepts Guide*.

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` command to perform all quorum device administrative procedures. In addition, you can accomplish some procedures by using the `clsetup` interactive utility or the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see"How to Access Oracle Solaris Cluster Manager" in *Oracle Solaris Cluster 4.3 System Administration Guide*.

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 Oracle Solaris Cluster Manager. For more information, see the `clquorum`(1CL) and `clsetup`(1CL) man pages.

When you work with quorum devices, keep in mind the following guidelines:

- All quorum commands must be run from a global-cluster 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 an Oracle Solaris ZFS storage pool. If a configured quorum device is added to a 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 10 Task List: Administering Quorum

Task	For Instructions
Add a quorum device to a cluster by using the `clsetup` utility	"Adding a Quorum Device" on page 177
Remove a quorum device from a cluster by using the `clsetup` utility (to generate `clquorum`)	"How to Remove a Quorum Device" on page 184
Remove the last quorum device from a cluster by using the `clsetup` utility (to generate `clquorum`)	"How to Remove the Last Quorum Device From a Cluster" on page 185
Replace a quorum device in a cluster by using the add and remove procedures	"How to Replace a Quorum Device" on page 187
Modify a quorum device list by using the add and remove procedures	"How to Modify a Quorum Device Node List" on page 188
Put a quorum device into maintenance state by using the `clsetup` utility (to generate `clquorum`) (While in maintenance state, the quorum device does not participate in voting to establish the quorum.)	"How to Put a Quorum Device Into Maintenance State" on page 189
Reset the quorum configuration to its default state by using the `clsetup` utility (to generate `clquorum`)	"How to Bring a Quorum Device Out of Maintenance State" on page 191

Task	For Instructions
List the quorum devices and vote counts by using the `clquorum` command	"How to List the Quorum Configuration" on page 193

Dynamic Reconfiguration With Quorum Devices

You must consider a few issues when completing dynamic reconfiguration operations on quorum devices in a cluster.

- All of the requirements, procedures, and restrictions that are documented for the Oracle Solaris dynamic reconfiguration feature also apply to Oracle Solaris Cluster dynamic reconfiguration support, except for the operating system quiescence operation. Therefore, review the documentation for the Oracle Solaris dynamic reconfiguration feature *before* using the dynamic reconfiguration feature with Oracle Solaris Cluster software. You should review in particular the issues that affect non-network IO devices during a dynamic reconfiguration detach operation.
- Oracle Solaris Cluster rejects dynamic reconfiguration remove-board operations that are performed when an interface is present that is configured for a quorum device.
- If the dynamic reconfiguration 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 11 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 177
2. Disable the quorum device to be removed.	"How to Remove a Quorum Device" on page 184
3. Perform the dynamic reconfiguration remove-operation on the device being removed.	

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 4.3 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:

- Shared LUNs from the following:
  - Shared SCSI disk
  - Serial Attached Technology Attachment (SATA) storage
  - Oracle ZFS Storage Appliance
  - Supported NAS devices
- 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 179
- "How to Add a Quorum Server Quorum Device" on page 181

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 Cluster software 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.

Note - You can also create a quorum server device or a shared disk quorum device by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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-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. **Assume a role that provides `solaris.cluster.modify` authorization on any node of the cluster.**

2. **Start the `clsetup` utility.**

```
# clsetup
```

The `clsetup` Main Menu is displayed.

3. **Type the number for the option for Quorum.**

   The Quorum Menu is displayed.

4. **Type the number for 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 for 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
```

▼ How to Add an Oracle ZFS Storage Appliance 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.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to add an Oracle ZFS Storage Appliance NAS device. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Consult the installation documentation that shipped with the Oracle ZFS Storage Appliance or the appliance's online Help for instructions on setting up an iSCSI device.**

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.LUN-name,IP-address-of-NASdevice
# devfsadm -i iscsi
# cldevice refresh
```

3. **From one cluster node, configure the DIDs for the iSCSI LUN.**

```
# 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. **Assume a role that provides `solaris.cluster.modify` 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) man page for more information.

▼ 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 Oracle Solaris Cluster Quorum Server Software” in *Oracle Solaris Cluster 4.3 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.

Note - You can also create a quorum server device by using the Oracle Solaris Cluster Manager browser interface. For Oracle Solaris Cluster Manager log-in instructions, see “How to Access Oracle Solaris Cluster Manager” on page 313.

---

1. **Assume a role that provides `solaris.cluster.modify` authorization on any node of the cluster.**

2. **Ensure 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 for the option for Quorum.**

The Quorum Menu is displayed.

5. **Type the number for 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 for the option for a quorum-server quorum device and 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 host of the quorum server.

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
```

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 184
- "How to Remove the Last Quorum Device From a Cluster" on page 185
- "How to Replace a Quorum Device" on page 187

▼ How to Remove a Quorum Device

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 185.

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 - You can also use the Oracle Solaris Cluster Manager browser interface to remove a quorum device. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume a role that provides `solaris.cluster.modify` authorization on any node in the cluster.**

2. **Determine the quorum device to be removed.**

```
# clquorum list -v
```

3. **Execute the `clsetup` utility.**

```
# clsetup
```

The Main Menu is displayed.

4. **Type the number for the option for Quorum.**

5. **Type the number for 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
```

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 199.

▼ 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 184.

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` command enables you to remove the device from the configuration. For more information, see the `clquorum`(1CL) man page. 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 177.

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. **Assume a role that provides `solaris.cluster.modify` 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` 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 256 and the `clsetup`(1CL) man page for more information.

   ---

3. **Verify that the quorum device has been removed.**

   ```
   # clquorum list -v
   ```

4. **Depending on why you are removing the last quorum device, proceed with one of the following steps:**

   - **If you are replacing the quorum device that has been removed, complete the following substeps:**

     a. **Add the new quorum device.**

        See "Adding a Quorum Device" on page 177 for instructions on adding the new quorum device.

b. **Remove the cluster from install mode.**

```
# cluster set -p installmode=disabled
```

- **If you are reducing your cluster to a single-node cluster, remove the cluster from install mode.**

```
# cluster set -p installmode=disabled
```

Example 58 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.

```
     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 177 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 184 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 in your hardware manual for your disk enclosure. See also the *Oracle Solaris Cluster Hardware Administration Manual*.

Maintaining Quorum Devices

This section provides the following procedures for maintaining quorum devices:

▼ How to Modify a Quorum Device Node List

You can use the `clsetup` 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, the `clquorum` command automatically configures the node-to-disk paths for all nodes attached to the disk. For more information, see the `clquorum`(1CL) man page.

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. **Assume a role that provides `solaris.cluster.modify` 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 for the Quorum option.**
   The Quorum Menu is displayed.

5. **Type the number for 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 for 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
```

▼ How to Put a Quorum Device Into Maintenance State

Use the `clquorum` command to put a quorum device into a maintenance state. For more information, see the `clquorum`(1CL) man page. The `clsetup` utility does not currently have this capability.

Put a quorum device into a 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 256.

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 - You can also use the Oracle Solaris Cluster Manager browser interface to disable a quorum device to put it into a maintenance state. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

If your cluster is in installation mode, click Reset Quorum Devices to exit installation mode.

---

1. **Assume a role that provides `solaris.cluster.modify` 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 59 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 191.

To put a node into maintenance state, see "How to Put a Node Into Maintenance State" on page 256.

▼ How to Bring a Quorum Device Out of Maintenance State

Run this procedure each time a quorum device is in a 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 258.
- To learn more about quorum vote counts, see "About Quorum Vote Counts" in *Oracle Solaris Cluster 4.3 Concepts Guide*.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to enable a quorum device to bring it out of a maintenance state. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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. **Assume a role that provides `solaris.cluster.modify` 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 60 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/rdsk/d20s2
Type:                                           shared_disk
Access Mode:                                    scsi3
Hosts (enabled):                                phys-schost-2, phys-schost-3
```

▼ How to List the Quorum Configuration

You do not need to be in the `root` role to list the quorum configuration. You can assume any role that provides `solaris.cluster.read` 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.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to view the quorum configuration. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

● **Use the `clquorum` command to list the quorum configuration.**

```
% clquorum show +
```

Example 61 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/rdsk/d20s2
Type:                                           shared_disk
Access Mode:                                    scsi3
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 177 to add a new quorum device.

   See "How to Remove a Quorum Device" on page 184 to remove a disk device as a quorum device.

2. **Replace the disk device.**

   To replace the disk device, see the procedures for the disk enclosure in the hardware guide. See also the *Oracle Solaris Cluster Hardware Administration Manual*.

3. **Add the replaced disk as a new quorum device.**

   See "Adding a Quorum Device" on page 177 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 184 to remove the quorum device.

Changing the Quorum Default Time-out

A default 25-second time-out exists for the completion of quorum operations during a cluster reconfiguration. You can increase the quorum time-out to a higher value by following the instructions in "How to Configure Quorum Devices" in *Oracle Solaris Cluster 4.3 Software Installation Guide*. Rather than increasing the time-out value, you can also switch to a different quorum device.

Additional troubleshooting information is available in "How to Configure Quorum Devices" in *Oracle Solaris Cluster 4.3 Software Installation Guide*.

Note - Do **not** change the default quorum time-out of 25 seconds for Oracle Real Application Clusters (Oracle RAC). In certain split-brain scenarios, a longer time-out period might lead to the failure of Oracle RAC VIP failover, due to the VIP resource timing out.

If the quorum device being used is not conforming with the default 25–second time-out, use a different 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:

- "Starting and Stopping the Quorum Server Software" on page 195
- "How to Start a Quorum Server" on page 196
- "How to Stop a Quorum Server" on page 197
- "Displaying Information About the Quorum Server" on page 197
- "Cleaning Up Stale Quorum Server Cluster Information" on page 199

For information about installing and configuring Oracle Solaris Cluster quorum servers, see "How to Install and Configure Oracle Solaris Cluster Quorum Server Software" in *Oracle Solaris Cluster 4.3 Software Installation Guide*.

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 installing the Quorum Server software, see "How to Install and Configure Oracle Solaris Cluster Quorum Server Software" in *Oracle Solaris Cluster 4.3 Software Installation Guide*. For information on changing the value of the quorum time-out, see "Changing the Quorum Default Time-out" on page 195.

▼ How to Start a Quorum Server

1. **Assume the `root` role on the host where you want to start the Oracle Solaris Cluster software.**

2. **Use the `clquorumserver start` command to start the software.**

   `# 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 62 Starting All Configured Quorum Servers

The following example starts all the configured quorum servers.

```
# clquorumserver start +
```

Example 63 Starting a Specific Quorum Server

The following example starts the quorum server that listens on port number 2000.

```
# clquorumserver start 2000
```

▼ How to Stop a Quorum Server

1. **Assume the `root` role on the host where you want to start the Oracle Solaris Cluster software.**

2. **Use the `clquorumserver stop` command to stop the software.**

```
# 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 64 Stopping All Configured Quorum Servers

The following example stops all the configured quorum servers.

```
# clquorumserver stop +
```

Example 65 Stopping a Specific Quorum Server

The following example stops the quorum server that listens on port number 2000.

```
# 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. **Assume the `root` role on the host where you want to display the quorum server information.**

   Users other than the `root` role require `solaris.cluster.read` authorization. For more information about rights profiles, see the `rbac(5)` man page.

2. **Display the configuration information of the quorum server by using the `clquorumserver` command.**

   ```
   # 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 66 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.

```
# 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 67 Displaying the Configuration of Several Quorum Servers

The following example displays the configuration information for three quorum servers, `qs1`, `qs2`, and `qs3`.

```
# clquorumserver show qs1 qs2 qs3
```

Example 68 Displaying the Configuration of All Running Quorum Servers

The following example displays the configuration information for all running quorum servers:

```
# 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 184. 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 184.

Caution - If the cluster is still using this quorum server, performing this procedure will compromise cluster quorum.

1. **Assume the `root` role 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 69 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
```

CHAPTER 7

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 all the public network management (PNM) objects, when you initially install and configure the cluster.

PNM objects include the Internet Protocol network multipathing (IPMP) groups, trunk and datalink multipathing (DLMP) link aggregations, and VNICs that are directly backed by link aggregations. Multipathing is installed automatically with the Oracle Solaris 11 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 219. For information about IPMP, see Chapter 3, "Administering IPMP" in *Administering TCP/IP Networks, IPMP, and IP Tunnels in Oracle Solaris 11.3*. For information about link aggregations, see Chapter 2, "Configuring High Availability by Using Link Aggregations" in *Managing Network Datalinks in Oracle Solaris 11.3*.

This chapter provides information and procedures for the following topics.

For a high-level description of the related procedures in this chapter, see Table 12, "Task List: Administering the Cluster Interconnect," on page 204.

Refer to the *Oracle Solaris Cluster 4.3 Concepts Guide* 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 from a global-cluster node.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to perform some of these tasks. For log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

For cluster software installation procedures, see the *Oracle Solaris Cluster 4.3 Software Installation Guide*. For procedures about servicing cluster hardware components, see the *Oracle Solaris Cluster 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 12 Task List: Administering the Cluster Interconnect

Task	Instructions
Administer the cluster transport by using `clsetup`(1CL)	"How to Access the Cluster Configuration Utilities" on page 33
Check the status of the cluster interconnect by using `clinterconnect status`	"How to Check the Status of the Cluster Interconnect" on page 205
Add a cluster transport cable, transport adapter, or switch by using `clsetup`	"How to Add Cluster Transport Cables, Transport Adapters, or Transport Switches" on page 207
Remove a cluster transport cable, transport adapter, or transport switch by using `clsetup`	"How to Remove Cluster Transport Cables, Transport Adapters, and Transport Switches" on page 209
Enable a cluster transport cable by using `clsetup`	"How to Enable a Cluster Transport Cable" on page 212
Disable a cluster transport cable by using `clsetup`	"How to Disable a Cluster Transport Cable" on page 213
Determining an transport adapter's instance number	"How to Determine a Transport Adapter's Instance Number" on page 214
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 215

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 dynamic reconfiguration feature also apply to Oracle Solaris Cluster dynamic reconfiguration support (except for the operating system quiescence operation). Therefore, review the documentation for the Oracle Solaris dynamic reconfiguration feature *before* using the dynamic reconfiguration feature with Oracle Solaris Cluster software. You should review in particular the issues that affect non-network IO devices during a dynamic reconfiguration detach operation.
- The Oracle Solaris Cluster software rejects dynamic reconfiguration remove-board operations performed on active private interconnect interfaces.
- You must completely remove an active adapter from the cluster in order to perform dynamic reconfiguration 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 dynamic reconfiguration operations on public network interfaces.

TABLE 13 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 221
2. Perform the dynamic reconfiguration operation on the public network interface.	

How to Check the Status of the Cluster Interconnect

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 the `root` role to perform this procedure.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to check the status of the cluster interconnect. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

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.
`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 `clinterconnect status` is rerun, take corrective action.

Example 70 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:net0   phys-schost-2:net0   Path online
  Transport path:   phys-schost-1:net4   phys-schost-2:net4   Path online
  Transport path:   phys-schost-1:net0   phys-schost-3:net0   Path online
  Transport path:   phys-schost-1:net4   phys-schost-3:net4   Path online
  Transport path:   phys-schost-2:net0   phys-schost-3:net0   Path online
  Transport path:   phys-schost-2:net4   phys-schost-3:net4   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 Hardware Administration Manual*.

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 - You can also use the Oracle Solaris Cluster Manager browser interface to add cables, transport adapters, and private adapters to your cluster. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

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

   For the procedure on installing a cluster transport cable, see the *Oracle Solaris Cluster Hardware Administration Manual*.

2. **Assume the `root` role on any node in the cluster.**

3. **Start the `clsetup` utility.**

   ```
   # clsetup
   ```

   The Main Menu is displayed.

4. **Type the number for the option for displaying the Cluster Interconnect Menu.**

5. **Type the number for the option for adding a transport cable.**

   Follow the instructions and type the requested information.

6. **Type the number for 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
nge	`set nge:nge_taskq_disable=1`
e1000g	`set e1000g:e1000g_taskq_disable=1`

7. **Type the number for 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 71 Verifying Adding a Cluster Transport Cable, Transport Adapter, or Transport Switch

The following example shows how to verify adding a transport cable, transport adapter, or transport switch to a node. The example contains settings for the Data Link Provider Interface (DLPI) transport type.

```
# clinterconnect show phys-schost-1:net5,hub2
  ===Transport Cables ===
Transport Cable:              phys-schost-1:net5@0,hub2
Endpoint1:                  phys-schost-2:net4@0
Endpoint2:                  hub2@2
State:                      Enabled

# clinterconnect show phys-schost-1:net5
=== Transport Adepters for net5
Transport Adapter:                              net5
Adapter State:                                  Enabled
Adapter Transport Type:                         dlpi
Adapter Property (device_name):                 net6
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 State (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 205.

How to Remove Cluster Transport Cables, Transport Adapters, and Transport Switches

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.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to remove cables, transport adapters, and private adapters from your cluster. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume the `root` role 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 for the option for accessing the Cluster Interconnect menu.**

5. **Type the number for 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 for 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 for 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.

   If you are removing a physical adapter from a node, see the *Oracle Solaris Cluster Hardware Administration Manual* for hardware service procedures.

8. **Type the number for 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 72 Verifying Removing a Transport Cable, Transport Adapter, or Transport Switch

The following example shows how to verify removal of a transport cable, transport adapter, or transport switch.

```
# clinterconnect show phys-schost-1:net5,hub2@0
 ===Transport Cables ===
Transport Cable:                      phys-schost-1:net5,hub2@0
Endpoint1:                            phys-schost-1:net5
Endpoint2:                            hub2@0
State:                                Enabled

# clinterconnect show phys-schost-1:net5
=== Transport Adepters for net5
Transport Adapter:                             net5
Adapter State:                                 Enabled
Adapter Transport Type:                        dlpi
Adapter Property (device_name):                net6
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 State (0):                        Enabled

# clinterconnect show hub2
=== Transport Switches ===
Transport Switch:                                hub2
State:                                            Enabled
Type:                                             switch
Port Names:                                       1 2
Port State(1):                                    Enabled
Port State(2):                                    Enabled
```

▼ How to Enable a Cluster Transport Cable

This option is used to enable an 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.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to enable a cable. Click Private Interconnects, click Cables, click the number of the cable to highlight it, then click Enable. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume the `root` role on any node in the cluster.**

2. **Start the `clsetup` utility.**

   ```
   # clsetup
   ```

   The Main Menu is displayed.

3. **Type the number for the option for accessing the Cluster Interconnect menu.**

4. **Type the number for the option for enabling the transport cable.**

   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
   ```

   Output is similar to the following:

   ```
   # clinterconnect show  phys-schost-1:net5,hub2
   Transport cable:   phys-schost-2:net0@0 ethernet-1@2    Enabled
   Transport cable:   phys-schost-3:net5@1 ethernet-1@3    Enabled
   Transport cable:   phys-schost-1:net5@0 ethernet-1@1    Enabled
   ```

▼ How to Disable a Cluster Transport Cable

You might need to disable a cluster transport cable to temporarily shut down a cluster interconnect path. 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-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 - You can also use the Oracle Solaris Cluster Manager browser interface to disable a cable. Click Private Interconnects, click Cables, click the number of the cable to highlight it, then click Disable. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume the `root` role 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.**

```
# clsetup
```

The Main Menu is displayed.

4. **Type the number for the option for accessing the Cluster Interconnect Menu.**

5. **Type the number for the option for disabling the transport cable.**

   Follow the prompts 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
   ```

   Output is similar to the following:

   ```
   # clinterconnect show  -p phys-schost-1:net5,hub2
   Transport cable:   phys-schost-2:net0@0 ethernet-1@2    Disabled
   Transport cable:   phys-schost-3:net5@1 ethernet-1@3    Enabled
   Transport cable:   phys-schost-1:net5@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
   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,o@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. To perform this task by using the command line, see the `cluster`(1CL) man page.

Before You Begin Ensure that remote shell (`rsh`(1M)) or secure shell (`ssh`(1)) access for the `root` role is enabled to all cluster nodes.

1. **Reboot all cluster nodes into noncluster mode by performing the following substeps on each cluster node:**

   a. **Assume a role that provides `solaris.cluster.admin` 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 the specified node to the next-preferred 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. **Choose the Change Network Addressing and Ranges for the Cluster Transport menu item.**

   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.**

   - **To change the default private network IP address:**

     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.**

   - **To change the IP address range:**

     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. **Provide 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.

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

    ```
    # cluster status -t node
    ```

Troubleshooting Cluster Interconnects

This section provides a troubleshooting procedure to disable and then enable a cluster interconnect, such as cluster transport adapters and transport cables.

Do not use the `ipadm` commands to administer cluster transport adapters. If a transport adapter was disabled by using the `ipadm disable-if` command, you must use the `clinterconnect` commands to disable the transport path and then enable it.

This procedure requires that you have Oracle Solaris Cluster software installed. These commands must be run from a global-cluster node.

▼ How to Enable a Cluster Interconnect

Note - You can also use the Oracle Solaris Cluster Manager browser interface to enable a cluster interconnect. Click Private Interconnects, click Cables, click the number of the cable to highlight it, then click Enable. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Check the status of the cluster interconnect.**

```
% clinterconnect status

=== Cluster Transport Paths===
Endpoint1               Endpoint2               Status
---------               ---------               ------
pnode1:net1             pnode2:net1             waiting
pnode1:net5             pnode2:net5             Path online
```

2. **Disable the cluster interconnect path.**

   a. **Check the cluster interconnect path.**

```
% clinterconnect show | egrep -ie "cable.*pnode1"
Transport Cable: pnode1:net5,switch2@1
Transport Cable: pnode1:net1,switch1@1
```

   b. **Disable the cluster interconnect path.**

```
% clinterconnect disable pnode1:net1,switch1@1
```

3. **Enable the cluster interconnect path.**

```
% clinterconnect enable pnode1:net1,switch1@1
```

4. **Verify that the cluster interconnect is enabled.**

```
% clinterconnect status

=== Cluster Transport Paths===
```

```
Endpoint1            Endpoint2            Status
---------            ---------            ------
pnode1:net1          pnode2:net1          Path online
pnode1:net5          pnode2:net5          Path online
```

Administering the Public Network

Oracle Solaris Cluster software supports the Oracle Solaris software implementation of IPMP, link aggregations, and VNICs for public networks. Basic public network administration is the same for both cluster and noncluster environments.

Multipathing is automatically installed when you install the Oracle Solaris 11 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, link aggregations, and VNICs in an Oracle Solaris Cluster environment.

For information about IPMP, see Chapter 3, "Administering IPMP" in *Administering TCP/IP Networks, IPMP, and IP Tunnels in Oracle Solaris 11.3*. For information about link aggregations, see Chapter 2, "Configuring High Availability by Using Link Aggregations" in *Managing Network Datalinks in Oracle Solaris 11.3*.

How to Administer IP Network Multipathing Groups in a Cluster

Before performing IPMP procedures on a cluster, consider the following guidelines.

- When configuring a scalable service resource (`SCALABLE=TRUE` in the resource type registration file for the resource type) that uses the `SUNW.SharedAddress` network resource, PNM can be configured to monitor IPMP group status on all IPMP groups on the cluster nodes in addition to the one the `SUNW.SharedAddress` is configured to use. This configuration allows the service to be restarted and failed over if any of the IPMP groups on the cluster nodes has failed, in order to maximize service availability for network clients that are co-located on the same subnets as the cluster nodes. For example:

  ```
  # echo ssm_monitor_all > /etc/cluster/pnm/pnm.conf
  ```

  Reboot the node.
- 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.

If you are using iSCSI storage as a quorum device, ensure that the probe-based IPMP device is configured correctly. If the iSCSI network is a private network containing only the cluster nodes and iSCSI storage device and there are no other hosts present on the iSCSI network, then the probe-based IPMP mechanism can break when all but one of the cluster nodes goes down. The problem occurs because there are no other hosts on the iSCSI network for IPMP to probe, so IPMP treats this as a network failure when only one node remains in the cluster. IPMP takes offline the iSCSI network adapter, and then the remaining node loses access to the iSCSI storage and thus the quorum device, To resolve this problem, you could add a router to the iSCSI network so that other hosts outside the cluster respond to the probes and prevent IPMP from taking offline the network adapter. Alternatively, you could configure IPMP with link-based failover instead of probe-based failover.

- Unless there are one or more non-link local IPv6 public network interfaces in the public network configuration, the `scinstall` utility automatically configures a multiple-adapter IPMP group for each set of public-network adapters in the cluster that uses the same subnet. These groups are link-based with transitive probes. Test addresses can be added if probe-based failure detection is required.
- 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`.
- Do not unconfigure (unplumb) or bring 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.
- Do not unplumb or remove a network interface from the IPMP group where the Oracle Solaris Cluster HA IP address is plumbed. This IP address can belong to the logical hostname resource or the shared address resource. However, if you unplumb the active interface by using the `ifconfig` command, Oracle Solaris Cluster now recognizes this event. It fails over the resource group to some other healthy node if the IPMP group has become unusable in the process. Oracle Solaris Cluster could also restart the resource group on the same node if the IPMP group is valid but an HA IP address is missing. The IPMP group becomes unusable for several reasons: loss of IPv4 connectivity, loss of IPv6 connectivity, or both. For more information, see the `if_mpadm`(1M) man page.
- 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 `ipmpstat -g` command. For more information, see Chapter 3, "Administering IPMP" in *Administering TCP/IP Networks, IPMP, and IP Tunnels in Oracle Solaris 11.3*.

For cluster software installation procedures, see the *Oracle Solaris Cluster 4.3 Software Installation Guide*. For procedures about servicing public networking hardware components, see the *Oracle Solaris Cluster 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 dynamic reconfiguration feature also apply to Oracle Solaris Cluster dynamic reconfiguration support (except for the operating system quiescence operation). Therefore, review the documentation for the Oracle Solaris dynamic reconfiguration feature *before* using the dynamic reconfiguration feature with Oracle Solaris Cluster software. You should review in particular the issues that affect non-network IO devices during a dynamic reconfiguration detach operation.
- Dynamic reconfiguration 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` command. For more information, see the `if_mpadm`(1M) man page.
- 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 dynamic reconfiguration remove operation on the disabled network adapter, availability is impacted. The remaining adapter has no place to fail over for the duration of the dynamic reconfiguration operation.

Complete the following procedures in the order indicated when performing dynamic reconfiguration operations on public network interfaces.

TABLE 14 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 `if_mpadm` command	`if_mpadm`(1M) man page. "How to Move an Interface From One IPMP Group to Another IPMP Group" in *Administering TCP/IP Networks, IPMP, and IP Tunnels in Oracle Solaris 11.3*
2. Remove the adapter from the multipathing group by using the `ipadm` command	`ipadm`(1M) man page "How to Remove an Interface From an IPMP Group" in *Administering TCP/IP Networks, IPMP, and IP Tunnels in Oracle Solaris 11.3*
3. Perform the dynamic reconfiguration operation on the public network interface	

♦ ♦ ♦ CHAPTER 8

Administering Cluster Nodes

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 or Zone Cluster" on page 223
- "Restoring Cluster Nodes" on page 226
- "Removing a Node From a Cluster" on page 231

For information about cluster maintenance tasks, see Chapter 9, "Administering the Cluster".

Adding a Node to a Cluster or Zone 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.

Note - The node that you add must run the same version of the Oracle Solaris Cluster software as the cluster it is joining.

Specifying an IP address and NIC for each zone cluster node is optional.

Note - If you do not configure an IP address for each zone cluster node, two things will occur:

1. That specific zone cluster will not be able to configure NAS devices for use in the zone cluster. The cluster uses the IP address of the zone cluster node when communicating with the NAS device, so not having an IP address prevents cluster support for fencing NAS devices.
2. The cluster software will activate any logical host IP address on any NIC.

If the original zone cluster node did not have a IP address or NIC specified, then you do not need to specify that information for the new zone cluster node.

In this chapter, `phys-schost#` 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 15 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	*Oracle Solaris Cluster Hardware Administration Manual*
Add shared storage	Add shared storage manually by following the instructions in the *Oracle Solaris Cluster Hardware Administration Manual*. You can also use Oracle Solaris Cluster Manager to add a shared storage device to a zone cluster. Navigate in Oracle Solaris Cluster Manager to the page for the zone cluster and click the Solaris Resources tab. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.
Add the node to the authorized node list	`claccess allow -h` *node-being-added*
Install and configure the software on the new cluster node	Chapter 2, "Installing Software on Global-Cluster Nodes" in *Oracle Solaris Cluster 4.3 Software Installation Guide*
Add the new node to an existing cluster	"How to Add a Node to an Existing Cluster or Zone Cluster" on page 224
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 *Oracle Solaris Cluster 4.3 Geographic Edition System Administration Guide*

▼ How to Add a Node to an Existing Cluster or Zone Cluster

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 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, assume the `root` role 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 15, "Task Map: Adding a Node to an Existing Global or Zone Cluster," on page 224.**

3. **Install and configure the software on the new cluster node.**

   Use the `scinstall` utility to complete the installation and configuration of the new node, as described in the *Oracle Solaris Cluster 4.3 Software Installation Guide*.

4. **Use the `scinstall` utility on the new node to configure that node in the cluster.**

5. **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 `sc_ipmp0` is the IPMP group name.

```
clzc:sczone>add node
clzc:sczone:node>set physical-host=phys-cluster-3
clzc:sczone:node>set hostname=hostname3
clzc:sczone:node>add net
clzc:sczone:node:net>set address=hostname3
clzc:sczone:node:net>set physical=sc_ipmp0
clzc:sczone:node:net>end
clzc:sczone:node>end
clzc:sczone>exit
```

   For detailed instructions on configuring the node, see "Creating and Configuring a Zone Cluster" in *Oracle Solaris Cluster 4.3 Software Installation Guide*.

6. **If the new zone cluster node will be a `solaris10` brand type and does not have Oracle Solaris Cluster software installed on the zone cluster, provide the path to the DVD image and install the software.**

```
# clzonecluster install-cluster -d dvd-image zone-cluster-name
```

7. **After you configure the node, reboot the node into cluster mode and install the zone cluster on the node.**

```
# clzonecluster install zone-cluster-name
```

8. **To prevent any new machines from being added to the cluster, from the `clsetup` utility type the number for the 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.

9. **Quit the `clsetup` utility.**

See Also `clsetup(1CL)` man page.

For a complete list of tasks for adding a cluster node, see Table 15, "Task Map: Adding a Node to an Existing Global or Zone Cluster," on page 224, "Task Map: Adding a Cluster Node."

To add a node to an existing resource group, see the *Oracle Solaris Cluster 4.3 Data Services Planning and Administration Guide*.

Restoring Cluster Nodes

You can use the Unified Archives to restore a cluster node so that it is exactly the same as the archive. Before you restore the node, you must first create a *recovery* archive on the cluster nodes. Only a *recovery* archive can be used; a *clone* archive cannot be used to restore a cluster node. See Step 1 below for instructions on creating the recovery archive.

This procedure prompts you for the cluster name, node names and their MAC addresses, and the path to the Unified Archives. For each archive that you specify, the `scinstall` utility verifies that the archive's source node name is the same as the node you are restoring. For instructions on restoring the nodes in a cluster from a Unified Archive, see "How to Restore a Node from the Unified Archive" on page 226.

▼ How to Restore a Node from the Unified Archive

This procedure uses the interactive form of the `scinstall` utility on the Automated Installer server. You must have already set up the AI server and installed the `ha-cluster/system/`

`install` packages from the Oracle Solaris Cluster repositories. The node name of the archive must be the same as the node that you are restoring.

Follow these guidelines to use the interactive `scinstall` utility in this procedure:

- Interactive `scinstall` enables you to type ahead. Therefore, do not press the Return key more than once if the next menu screen does not appear immediately.
- Unless otherwise noted, you can press Control-D to return to either the start of a series of related questions or to the Main Menu.
- Default answers or answers to previous sessions are displayed in brackets ([ ]) at the end of a question. Press Return to enter the response that is in brackets without typing it.

1. **Assume the `root` role on a node of the global cluster and create a recovery archive.**

```
phys-schost# archiveadm create -r archive-location
```

When you create an archive, exclude the ZFS datasets that are on the shared storage. If you plan to restore the data on the shared storage, use the traditional method.

For more information on using the `archiveadm` command, see the `archiveadm`(1M) man page.

2. **Log into the Automated Installer server and assume the `root` role.**

3. **Start the `scinstall` utility.**

```
phys-schost# scinstall
```

4. **Type the option number to restore a cluster.**

```
*** Main Menu ***

Please select from one of the following (*) options:

* 1) Install, restore, or replicate a cluster from this Automated Installer server
* 2) Securely install, restore, or replicate a cluster from this Automated Installer
 server
* 3) Print release information for this Automated Installer install server

* ?) Help with menu options
* q) Quit

Option:  2
```

Choose Option 1 to restore a cluster node using a non-secure AI server installation. Choose Option 2 to restore a cluster node by using the secure AI server installation.

The Custom Automated Installer Menu or Custom Secure Automated Installer Menu is displayed.

5. **Type the option number to Restore Cluster Nodes from Unified Archives.**

   The Cluster Name screen is displayed.

6. **Type the cluster name that contains the nodes you want to restore.**

   The Cluster Nodes screen is displayed.

7. **Type the names of the cluster nodes that you want to restore from the Unified Archives.**

   Type one node name per line. When you are done, press Control-D and confirm the list by typing yes and pressing Return. If you want to restore all the nodes in the cluster, specify all the nodes.

   If the `scinstall` utility is unable to find the MAC address of the nodes, type in each address when prompted.

8. **Type the full path to the recovery archive.**

   The archive used to restore a node *must* be a recovery archive. The archive file you use to restore a particular node must be created on the same node. Repeat this for each cluster node you want to restore.

9. **For each node, confirm the options you chose so that the `scinstall` utility performs the necessary configuration to install the cluster nodes from this AI server.**

   The utility also prints instructions to add the DHCP macros on the DHCP server, and adds or clears the security keys for SPARC nodes (if you chose secure installation). Follow those instructions.

10. **(Optional) To customize the target device, update the AI manifest for each node.**

    The AI manifest is located in the following directory:

    ```
    /var/cluster/logs/install/autoscinstall.d/ \
    cluster-name/node-name/node-name_aimanifest.xml
    ```

    a. **To customize the target device, update the `target` element in the manifest file.**

       Update the `target` element in the manifest file based on how you want to use the supported criteria to locate the target device for the installation. For example, you can specify the `disk_name` sub-element.

Note - `scinstall` assumes the existing boot disk in the manifest file to be the target device. To customize the target device, update the `target` element in the manifest file. For more information, see Part 3, "Installing Using an Install Server," in *Installing Oracle Solaris 11.3 Systems* and the `ai_manifest(4)` man page.

---

b. **Run the `installadm` command for each node.**

```
# installadm update-manifest -n cluster-name-{sparc|i386} \
-f /var/cluster/logs/install/autoscinstall.d/cluster-name/node-name/node-
name_aimanifest.xml \
-m node-name_manifest
```

Note that `SPARC` and `i386` is the architecture of the cluster node.

11. **If you are using a cluster administrative console, display a console screen for each node in the cluster.**

- **If `pconsole` software is installed and configured on your administrative console, use the `pconsole` utility to display the individual console screens.**

  As the `root` role, use the following command to start the `pconsole` utility:

  ```
  adminconsole# pconsole host[:port] […]  &
  ```

  The `pconsole` utility also opens a master window from which you can send your input to all individual console windows at the same time.

- **If you do not use the `pconsole` utility, connect to the consoles of each node individually.**

12. **Shut down and boot each node to start the AI installation.**

The Oracle Solaris software is installed with the default configuration.

Note - You cannot use this method if you want to customize the Oracle Solaris installation. If you choose the Oracle Solaris interactive installation, the Automated Installer is bypassed and Oracle Solaris Cluster software is not installed and configured.

To customize Oracle Solaris during installation, instead follow instructions in "How to Install Oracle Solaris Software" in *Oracle Solaris Cluster 4.3 Software Installation Guide*, then install and configure the cluster by following instructions in "How to Install Oracle Solaris Cluster Software Packages" in *Oracle Solaris Cluster 4.3 Software Installation Guide*.

- **SPARC:**

  a. **Shut down each node.**

  ```
  phys-schost# cluster shutdown -g 0 -y
  ```

  b. **Boot the node with the following command**

  ```
  ok boot net:dhcp - install
  ```

Note - Surround the dash (-) in the command with a space on each side.

- **x86**

  a. **Reboot the node.**

  ```
  # reboot -p
  ```

  b. **During PXE boot, press Control-N.**

  The GRUB menu is displayed.

  c. **Immediately select the Automated Install entry.**

Note - If you do not select the Automated Install entry within 20 seconds, installation proceeds, using the default interactive text installer method which will not install and configure the Oracle Solaris Cluster software.

Each node will be automatically rebooted to join the cluster after the installation is finished. The node is restored to the same state as when the archive was created. Oracle Solaris Cluster installation output is logged in the `/var/cluster/logs/install/sc_ai_config.log` file on each node.

13. **From one node, verify that all nodes have joined the cluster.**

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

Output resembles the following.

```
=== Cluster Nodes ===

--- Node Status ---
```

```
Node Name                                       Status
---------                                       ------
phys-schost-1                                   Online
phys-schost-2                                   Online
phys-schost-3                                   Online
```

For more information, see the `clnode(1CL)` man page.

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, remove a node for a RAC configuration, follow the appropriate steps below.

TABLE 16 Task Map: Removing a Node

Task	Instructions
Move all resource groups and device groups off the node to be removed. If you have a zone cluster, log into the zone cluster and evacuate the zone cluster node that is on the physical node getting uninstalled. Then remove the node from the zone cluster before you bring the physical node down. If the affected physical node has already failed, simply remove the node from the cluster.	`clnode evacuate` *node* "How to Remove a Node From a Zone Cluster" on page 232
Verify that the node can be removed by checking the allowed hosts. If the node is not listed by the `claccess show` command, it cannot be removed. Give the node access to the cluster configuration.	`claccess show` `claccess allow -h` *node-to-remove*
Remove the node from all device groups.	"How to Remove a Node From a Device Group (Solaris Volume Manager)" on page 143
Remove all quorum devices connected to the node being removed. **This step is optional if you are removing a node from a two-node cluster.**	"How to Remove a Quorum Device" on page 184 "How to Remove the Last Quorum Device From a Cluster" on page 185

Task	Instructions
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.	
Put the node being removed into noncluster mode.	"How to Put a Node Into Maintenance State" on page 256
Remove a node from the cluster software configuration.	"How to Remove a Node From the Cluster Software Configuration" on page 233
(Optional) Uninstall Oracle Solaris Cluster software from a cluster node.	"How to Uninstall Oracle Solaris Cluster Software From a Cluster Node" on page 260

▼ 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 15, "Task Map: Adding a Node to an Existing Global or Zone Cluster," on page 224. Most of these steps are performed from the global-cluster node.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to shut down a zone-cluster node, but not to remove a zone-cluster node. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume the `root` role 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 zone-cluster-name
   ```

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

3. **Remove the node from all resource groups in the zone cluster.**

   ```
   phys-schost# clresourcegroup remove-node -n zone-hostname -Z zone-cluster-name rg-name
   ```

   If you use the procedure described in the Note in Step 2, the resource groups should be automatically removed so you might be able to skip this step.

4. **Uninstall the zone-cluster node.**

   ```
   phys-schost# clzonecluster uninstall -n node zone-cluster-name
   ```

5. **Remove the zone-cluster node from the configuration.**

Use the following commands:

```
phys-schost# clzonecluster configure zone-cluster-name
clzc:sczone> remove node physical-host=node
clzc:sczone> exit
```

Note - If the zone cluster node you want to remove resides on a system that is inaccessible or unable to join the cluster, remove the node with the `clzonecluster` interactive shell:

```
clzc:sczone> remove -F node physical-host=node
```

If you use this method to remove the last zone cluster node, you will be prompted to delete the zone cluster entirely. If you choose not to, the last node will not be removed. This deletion has the same effect as **`clzonecluster delete -F`** *zone-cluster-name*.

---

6. **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. **Assume a role that provides `solaris.cluster.modify` 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 232 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.

For more information about GRUB based booting, see "Booting a System" in *Booting and Shutting Down Oracle Solaris 11.3 Systems*.

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

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/kernel/#ISADIR/unix -B $ZFS-BOOTFS -x
```

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

The screen displays the edited command.

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.

4. Delete the node from the cluster.

a. Run the following command from an active node:

```
phys-schost# clnode clear -F nodename
```

If you have resource groups that are have `rg_system=true`, you must change them to `rg_system=false` so that the `clnode clear -F` command will succeed. After you run `clnode clear -F`, reset the resource groups back to `rg_system=true`.

b. **Run the following command from the node you want to remove:**

```
phys-schost# clnode remove -F
```

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 the `clnode status` *nodename* command.

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 260.**

  You could also choose to remove the node from the cluster and uninstall the Oracle Solaris Cluster software at the same time. Change to a directory that does not contain any Oracle Solaris Cluster files and type **`scinstall -r`**.

- **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.**

  See *Oracle Solaris Cluster Hardware Administration Manual* for instructions.

Example 73 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 260.

For hardware procedures, see the *Oracle Solaris Cluster Hardware Administration Manual*.

For a comprehensive list of tasks for removing a cluster node, see Table 16, "Task Map: Removing a Node," on page 231.

To add a node to an existing cluster, see "How to Add a Node to an Existing Cluster or Zone Cluster" on page 224.

▼ 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 the specified node to the next-preferred node.

4. **Put the device groups into maintenance state.**
   For the procedure on putting a device group in maintenance state, see "How to Put a Node Into Maintenance State" on page 256.

5. **Remove the node from the device groups.**
   If you use a raw disk, use the `cldevicegroup(1CL)` command to remove the device groups.

6. **For each resource group that contains an `HAStoragePlus` resource, remove the node from the resource group's node list.**

   ```
   phys-schost# clresourcegroup remove-node -n node + | resourcegroup
   ```

   See the *Oracle Solaris Cluster 4.3 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# pkg uninstall /ha-cluster/library/ucmm
```

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`.

13. **On the node, update the device namespace by updating the `/devices` and `/dev` entries.**

```
phys-schost# devfsadm -C
phys-schost# cldevice refresh
```

14. **Bring the device groups back online.**

    For information about bringing a device group online, see "How to Bring a Node Out of Maintenance State" on page 258.

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 2b.
   - 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 141.

   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 260 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 260 and repeat the procedure.**

CHAPTER 9

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 241
- "Performing Zone Cluster Administrative Tasks" on page 273
- "Troubleshooting Procedures for Testing Purposes" on page 285

For information about adding or removing a node from the cluster, see Chapter 8, "Administering Cluster Nodes".

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 17 Task List: Administering the Cluster

Task	Instructions
Add or remove a node from a cluster	Chapter 8, "Administering Cluster Nodes"

Task	Instructions
Change the name of the cluster	"How to Change the Cluster Name" on page 242
List node IDs and their corresponding node names	"How to Map Node ID to Node Name" on page 244
Permit or deny new nodes to add themselves to the cluster	"How to Work With Authentication for New Cluster Nodes" on page 244
Change the time for a cluster by using the NTP	"How to Reset the Time of Day in a Cluster" on page 246
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	"How to Display the OpenBoot PROM (OBP) on a Node" on page 248
Add or change the private hostname	"How to Change the Node Private Hostname" on page 249
	You can also use the Oracle Solaris Cluster Manager browser interface to add a logical hostname to a global cluster or a zone cluster. Click Tasks, then click Logical Hostname to start the wizard. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.
Put a cluster node in maintenance state	"How to Put a Node Into Maintenance State" on page 256
Rename a Node	"How to Rename a Global Cluster Node" on page 252
Bring a cluster node out of maintenance state	"How to Bring a Node Out of Maintenance State" on page 258
Uninstall cluster software from a cluster node	"How to Uninstall Oracle Solaris Cluster Software From a Cluster Node" on page 260
Add and manage an SNMP Event MIB	"How to Enable an SNMP Event MIB" on page 265
	"How to Add an SNMP User on a Node" on page 269
Configure load limits for each node	"How to Configure Load Limits on a Node" on page 272
Move a zone cluster; prepare a zone cluster for applications, remove a zone cluster	"Performing Zone Cluster Administrative Tasks" on page 273

▼ How to Change the Cluster Name

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

Caution - Do not perform this procedure if the cluster is in an Oracle Solaris Cluster Geographic Edition partnership. Instead, follow procedures in "Renaming a Cluster That Is in a Partnership" in *Oracle Solaris Cluster 4.3 Geographic Edition System Administration 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. **Assume the `root` role 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 for 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 74 Changing the Cluster Name

The following example shows the `cluster` command generated from the `clsetup` utility to change to the new cluster name, `dromedary`.

```
phys-schost# cluster rename -c dromedary
```

For more information, see the `cluster`(1CL) and `clsetup`(1CL) man pages.

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 the `root` role 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` command to list the cluster configuration information for the global cluster.**

   ```
   phys-schost# clnode show | grep Node
   ```

   For more information, see the clnode(1CL) man page.

2. **(Optional) 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 sczone clnode -v | grep Node
   ```

Example 75 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 Authentication for New Cluster Nodes

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. **Assume the `root` role on any node in the global cluster.**

2. **Start the `clsetup` utility.**

   ```
   phys-schost# clsetup
   ```

   The Main Menu is displayed.

3. **To work with cluster authentication, type the number for the option for new nodes.**

   The New Nodes menu is displayed.

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

Example 76 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 77 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 78 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 79 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 80 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 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 4.3 Concepts Guide* and the *Network Time Protocol's User's Guide* at `http://download.oracle.com/docs/cd/E19065-01/servers.10k/`.

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`, `rdate`, or `svcadm` commands interactively or within the `cron` scripts. For more information, see the `date`(1), `rdate`(1M), `svcadm`(1M), or `cron`(1M) man pages. The `ntpd(1M)` man page is delivered in the `service/network/ntp` Oracle Solaris 11 package.

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. **Assume the `root` role on any node in the global cluster.**

2. **Shut down the global cluster.**

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

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.

      For more information about GRUB based booting, see "Booting a System" in *Booting and Shutting Down Oracle Solaris 11.3 Systems*.

   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.

   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/kernel/$ISADIR/unix _B $ZFS-BOOTFS -x
      ```

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

The screen displays the edited command.

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 node to the next-preferred 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 4.3 Data Services Planning and Administration Guide*.

2. **If your NTP configuration file refers to the private hostname that you are changing, bring down the NTP daemon on each node of the cluster.**

Use the `svcadm` command to shut down the NTP daemon. See the `svcadm`(1M) man page for more information about the NTP daemon.

```
phys-schost# svcadm disable ntp
```

3. **Run the `clsetup` utility to change the private hostname of the appropriate node.**

Run the utility from only one of the nodes in the cluster. for more information, see the `clsetup`(1CL) man page.

Note - When selecting a new private hostname, ensure that the name is unique to the cluster node.

You can also run the `clnode` command instead of the `clsetup` utility to change the private hostname. In the example below, the cluster node name is `phys-schost-1`. After you run the `clnode` command below, go to Step 6.

```
phys-schost# clnode set -p privatehostname=New-private-nodename phys-schost-1
```

4. **In the `clsetup` utility, type the number for the option for the private hostname.**

5. **In the `clsetup` utility, type the number for 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 or include file, update the NTP file on each node.**

   If you changed a private hostname in your NTP configuration file (`/etc/inet/ntp.conf`) and you have peer host entries or a pointer to the include file for the peer hosts in your NTP configuration file (`/etc/inet/ntp.conf.include`), update the file on each node. If you changed a private hostname in your NTP include file, update the `/etc/inet/ntp.conf.sc` file 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. Typically, the `ntp.conf.sc` 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 svc:network/ntp:default
      ```

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 4.3 Data Services Planning and Administration Guide*.

Example 81 Changing the Private Hostname

The following example changes the private hostname from `clusternode2-priv` to `clusternode4-priv`, on node `phys-schost-2`. Perform this action on each node.

Disable all applications and data services as necessary

```
phys-schost-1# svcadm disable ntp
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# pfedit /etc/inet/ntp.conf.sc
...
peer clusternode1-priv
peer clusternode4-priv
peer clusternode3-priv
phys-schost-1# ping clusternode4-priv
phys-schost-1# svcadm enable ntp
Enable all applications and data services disabled at the beginning of the procedure
```

▼ How to Rename a Global Cluster 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, assume a role that provides `solaris.cluster.modify` authorization.**

2. **(Optional) If you are renaming a node in an Oracle Solaris Cluster Geographic Edition cluster that is in a partnership, determine whether to switch over the protection group.**

   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 4, "Administering Cluster Partnerships" in *Oracle Solaris Cluster 4.3 Geographic Edition System Administration Guide*.

3. **Rename the Oracle Solaris host names.**

   Complete the steps in "How to Change a System's Identity" in *Managing System Information, Processes, and Performance in Oracle Solaris 11.3*, 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 new-node old-node
```

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;reboot
```

9. **Verify the node displays the new name.**

```
# clnode status -v
```

10. **Update your Geographic Edition configuration to use the new cluster node name.**

    The configuration information used by the protection groups and your data replication product might specify the node name.

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 255 for instructions on this optional step.

▼ How to Rename a Zone Cluster Node

You can use this procedure to rename a zone cluster node.

1. **From a global cluster node where the zone cluster node that you want to rename is running, login to the zone cluster.**

```
# zlogin <zone-cluster-name>
```

2. **Rename the zone cluster node.**

```
# hostname <new-zone-cluster-hostname>
```

3. **Confirm that the zone cluster node was renamed.**

```
# hostname
```

This command must return the new hostname given to the zone cluster node.

4. **Exit from the zone cluster.**

```
# exit
```

5. **From one node of the global cluster, use `clzc` command to apply the new zone cluster hostname in the cluster configuration.**

```
# clzc configure <zone-cluster-name>
clzc:<zone-cluster-name>> select node physical-host=<globa-cluster-nodename>
clzc:<zone-cluster-name>:node> set hostname=<new-zone-cluster-hostname>
clzc:<zone-cluster-name>:node> end
clzc:<zone-cluster-name>> exit
#
```

6. **Reboot the zone cluster.**

```
# clzc reboot <zone-cluster-name>
```

7. **Verify the new zone cluster hostname.**

The following commands print the new zone cluster hostname.

```
# clzc status
# zlogin <zone-cluster-name>
# hostname
```

8. **If Geographic Edition is running on the zone cluster where you updated the zone cluster hostname, then update your Geographic Edition configuration to use the new zone cluster hostname.**

The configuration information used by the protection groups and your data replication product might specify the node name. Therefore, they must be updated.

9. **If necessary, change the logical hostname resources' `hostnamelist` property.**

For more information, see "How to Change the Logical Hostnames Used by Existing Oracle Solaris Cluster Logical Hostname Resources" on page 255.

▼ 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 Global Cluster Node" on page 252. This step is optional.

1. **On the global cluster, assume a role that provides `solaris.cluster.modify` 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.**

   ```
   # clresourcegroup offline apache-rg
   ```

   b. **Disable the Apache logical hostname resources.**

   ```
   # clresource disable apache-lh-res
   ```

   c. **Provide the new hostname list.**

   ```
   # clresource 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.**

   ```
   # clresource enable apache-lh-res
   ```

   f. **Bring the Apache resource groups online.**

   ```
   # clresourcegroup online -eM apache-rg
   ```

   g. **Confirm that the application started correctly by running the following command checking a client.**

```
# clresource 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 the `clnode evacuate` and `shutdown` commands. For more information, see the `clnode`(1CL) and `cluster`(1CL) man pages.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to evacuate a node and switch all resource groups and device groups to the next-preferred node. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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.

Note - The Oracle Solaris `shutdown` command shuts down a single node, while the `cluster shutdown` command shuts down the entire cluster.

Use the `clquorum disable` command from another node that is still a cluster member to put a cluster node into maintenance state. For more information, see the `clquorum`(1CL) man page.

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. **Assume a role that provides `solaris.cluster.modify` 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 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 -i0
```

4. **Assume a role that provides `solaris.cluster.modify` 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
```

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 82 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 258.

▼ 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. **Assume a role that provides `solaris.cluster.modify` 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
   ```

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 83 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.

```
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/rdsk/d3s2    1            1             Online
```

```
/dev/did/rdsk/d17s2   0           1            Online
/dev/did/rdsk/d31s2   1           1            Online
`
```

How to Uninstall Oracle Solaris Cluster Software From a Cluster Node

Perform this procedure to unconfigure 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 Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems" in *Oracle Solaris Cluster 4.3 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 16, "Task Map: Removing a Node," on page 231.

   Ensure that you have removed the node from the cluster configuration by using `clnode remove` before you continue with this procedure. Other steps might include adding the node you plan to uninstall to the cluster's node–authentication list, uninstalling a zone cluster, and so on.

   Note - To unconfigure the node but leave Oracle Solaris Cluster software installed on the node, do not proceed further after you run the `clnode remove` command.

   ---

2. **Assume the `root` role on the node to uninstall.**

3. **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 -i0 ok 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
```

4. **In the `/etc/vfstab` file, remove all globally mounted file-system entries *except* the `/global/.devices` global mounts.**

5. **Reboot the node into noncluster mode.**

   - **On SPARC based systems, perform the following command:**

   ```
   ok boot -x
   ```

   - **On x86 based systems, perform 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.**

        For more information about GRUB based booting, see "Booting a System" in *Booting and Shutting Down Oracle Solaris 11.3 Systems*.

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

     c. **Add `-x` to the command to specify that the system boot into noncluster mode.**

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

The screen displays the edited command.

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 add the -x option to the kernel boot parameter command.

---

6. **Change to a directory, such as the root (`/`) directory, that does not contain any files that are delivered by the Oracle Solaris Cluster packages.**

```
phys-schost# cd /
```

7. **To unconfigure the node and remove Oracle Solaris Cluster software, run the following command.**

```
phys-schost# scinstall -r [-b bename]
```

`-r`

Removes cluster configuration information and uninstalls Oracle Solaris Cluster framework and data-service software from the cluster node. You can then reinstall the node or remove the node from the cluster.

`-b` *bootenvironmentname*

Specifies the name of a new boot environment, which is where you boot into after the uninstall process completes. Specifying a name is optional. If you do not specify a name for the boot environment, one is automatically generated.

See the `scinstall(1M)` man page for more information.

8. **If you intend to reinstall the Oracle Solaris Cluster software on this node after the uninstall completes, reboot the node to boot into the new boot environment.**

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 - For more information about migrating a global-devices namespace to a `lofi` device, see "Migrating the Global-Devices Namespace" on page 131.

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 260 and repeat the procedure. Ensure that you successfully complete Step 4 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.
```

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. Because clusters generate numerous notifications, only events with a severity of `min_severity` or greater are sent as trap notifications. By default, the `min_severity` value is set to NOTICE. The `log_number` value specifies the number of events to be logged in the MIB table before retiring older entries. The MIB maintains a read-only table of the most current events for which a trap has been sent. The number of events is limited by the `log_number` value. 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 Cluster Event SNMP Interface uses the common agent container (cacao) SNMP adaptor as its SNMP agent infrastructure. By default, the port number for the SNMP is 11161, and the default port number for the SNMP traps is 11162. These port numbers can be changed by using the `cacaoadm` command. See the `cacaoadm`(1M) man page for more information.

Creating, setting up, and managing an Oracle Solaris Cluster SNMP event MIB can involve the following tasks.

TABLE 18 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 265
Disable an SNMP event MIB	"How to Disable an SNMP Event MIB" on page 265
Change an SNMP event MIB	"How to Change an SNMP Event MIB" on page 266
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 267

Task	Instructions
Remove an SNMP host	"How to Disable an SNMP Host From Receiving SNMP Traps on a Node" on page 268
Add an SNMP user	"How to Add an SNMP User on a Node" on page 269
Remove an SNMP user	"How to Remove an SNMP User From a Node" on page 270

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. **Assume a role that provides `solaris.cluster.modify` 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. **Assume a role that provides `solaris.cluster.modify` authorization.**

2. **Disable the SNMP event MIB.**

```
phys-schost-1# clsnmpmib 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, minimum severity value, and event logging 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. **Assume a role that provides `solaris.cluster.modify` authorization.**

2. **Change the protocol, minimum severity value, and event logging of the SNMP event MIB.**

```
phys-schost-1# clsnmpmib set -n node
-p version=SNMPv3 \
-p min_severity=WARNING \
-p log_number=100 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 the *value* as follows:

- `version=SNMPv2`
- `version=snmpv2`

- `version=2`
- `version=SNMPv3`
- `version=snmpv3`
- `version=3`

`-p min_severity=`*value*

Specifies the minimum severity value to use with MIBs. You specify the *value* as follows:

- `min_severity=NOTICE`
- `min_severity=WARNING`
- `min_severity=ERROR`
- `min_severity=CRITICAL`
- `min_severity=FATAL`

`-p log_number=`*number*

Specifies the number of events to be logged in the MIB table before retiring the older entries. The default value is 100. Values must range from 100-500. You specify the *value* as follows: `log_number=`*100*.

MIB

Specifies the name of the MIB or MIBs to which to apply the subcommand. In this case, you must specify `event`. If you do not specify this operand, the subcommand uses the default plus sign (+), which means all MIBs. If you use the *MIB* operand, specify the MIB in a space-delimited list after all other command-line options.

For more information, see the `clsnmpmib`(1CL) man page.

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. **Assume a role that provides `solaris.cluster.modify` authorization.**

2. **Add the host to the SNMP host list of a community on another node.**

   phys-schost-1# **clsnmphost add -c** *SNMPcommunity* **[-n** *node***]** *host*

   -c *SNMPcommunity*

   Specifies the SNMP community name that is used in conjunction with the hostname. The host is a system in the network which can be configured to receive the traps

   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 cluster *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 default is the node where the command is run.

   host

   Specifies the name, IP address, or IPv6 address of a host that is provided access to the SNMP MIBs in the cluster. This can be a host outside the cluster or a cluster node itself trying to get SNMP traps.

▼ 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. **Assume a role that provides `solaris.cluster.modify` 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 cluster *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 default is the node where the command is run.

host

Specifies the name, IP address, or IPv6 address of the host that is removed from the configuration. This can be a host outside the cluster or a cluster node itself trying to get SNMP traps.

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. **Assume a role that provides `solaris.cluster.modify` 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. **Assume a role that provides `solaris.cluster.modify` 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.

Configuring Load Limits

You can enable the automatic distribution of resource group load across nodes 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 browser interface to configure load limits. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313. The following procedure illustrates how to configure load limits by using the command line.

▼ How to Configure Load Limits on a Node

Note - You can also use the Oracle Solaris Cluster Manager browser interface to create and configure a load limit on a global-cluster node or a zone-cluster node, or to edit or delete an existing node load limit. Click Nodes or Zone Clusters then click the name of the node to access its page. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume a role that provides `solaris.cluster.modify` 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.**

   In the following example command, 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.

   ```
   # clnode create-loadlimit -p limitname=mem_load -Z zc1 \
   -p softlimit=11 -p hardlimit=20 node1 node2 node3
   ```

3. **Assign load factor values to each resource group.**

   In the following example command, 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.

   ```
   # clresourcegroup set -p load_factors=mem_load@50,factor2@1 rg1 rg2
   ```

   You can also perform this step during the creation of the resource group with the `clresourceroup create` command. See the `clresourcegroup`(1CL) man page for more information.

4. **If desired, perform one or more additional optional configuration tasks.**

   - **Redistribute the existing load.**

     ```
     # clresourcegroup remaster rg1 rg2
     ```

     This command can move resource groups off their current master to other nodes to achieve uniform load distribution.

   - **Assign 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.

- **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.

- **Set the `Concentrate_load` flag.**

  ```
  # cluster set -p Concentrate_load=TRUE
  ```

- **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
  ```

5. **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.

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 a node of the global cluster.

You can create a new zone cluster or add a file system or storage device to an existing zone cluster by using the `clsetup` utility to launch the zone cluster configuration wizard. The zones in a zone cluster are configured when you run `clzonecluster install -c` to configure the profiles. See “Creating and Configuring a Zone Cluster” in *Oracle Solaris Cluster 4.3 Software*

Installation Guide for instructions about using the `clsetup` utility or the `-c config_profile` option.

You can also use the Oracle Solaris Cluster Manager browser interface to create a zone cluster, or add a file system or storage device to it. You can also use the Oracle Solaris Cluster Manager browser interface to edit the zone cluster's `Resource Security` property. Click Zone Clusters, click the name of the zone cluster to go to its page, then click the Solaris Resources tab to administer zone-cluster components, or click Properties to administer zone-cluster properties. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

Note - The Oracle Solaris Cluster commands that you run only from a 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 zone clusters.

TABLE 19 Other Zone-Cluster Tasks

Task	Instructions
Move the zone path to a new zone path	`clzonecluster move -f` *zonepath zone-cluster-name*
Prepare the zone cluster to run applications	`clzonecluster ready -n` *nodename zone-cluster-name*
Restore the nodes from the United Archives	"How to Restore a Node from the Unified Archive" on page 226
Configure or install a zone cluster from the United Archive	"How to Configure a Zone Cluster from the Unified Archive" on page 275
	"How to Install a Zone Cluster from the Unified Archive" on page 276
	Use a command:
	`clzonecluster clone -Z` *target- zone-cluster-name* `[-m` *copymethod*`]` *source-zone-cluster-name*
	Halt the source zone cluster before you use the `clone` subcommand. The target zone cluster must already be configured.
Add a network address to a zone cluster	"How to Add a Network Address to a Zone Cluster" on page 277
Add a node to a zone cluster	"How to Add a Node to an Existing Cluster or Zone Cluster" on page 224
Remove a node from a zone cluster	"How to Remove a Node From a Zone Cluster" on page 232
Remove a zone cluster	"How to Remove a Zone Cluster" on page 279
Remove a file system from a zone cluster	"How to Remove a File System From a Zone Cluster" on page 280
Remove a storage device from a zone cluster	"How to Remove a Storage Device From a Zone Cluster" on page 283
Restore zone cluster nodes from the Unified Archive	"How to Restore a Node from the Unified Archive" on page 226

Task	Instructions
Troubleshoot a node uninstallation	"Troubleshooting a Node Uninstallation" on page 263
Create, set up, and manage the Oracle Solaris Cluster SNMP Event MIB	"Creating, Setting Up, and Managing the Oracle Solaris Cluster SNMP Event MIB" on page 264

▼ How to Configure a Zone Cluster from the Unified Archive

Use the `clzonecluster` command to launch an interactive utility to configure a `solaris10` or `labeled` branded zone cluster from the Unified Archive. The `clzonecluster configure` utility lets you specify a *recovery* archive or a *clone* archive.

If you prefer to use the command line rather than the interactive utility to configure a zone cluster from an archive, use the `clzonecluster configure -f` *command-file* command. See the `clzonecluster`(1CL) man page for more information.

Note - If the zone cluster that you want to install was already configured using other supported methods, you do not have to configure the zone cluster from a Unified Archive.

---

1. **Create a recovery or clone archive.**

   ```
   phys-schost# archiveadm create -r archive-location
   ```

   Use the `create` command to create a clone archive or the `-r` option to create a recovery archive. For more information about using the `archiveadm` command, see the `archiveadm`(1M) man page.

2. **Assume the `root` role on a node of the global cluster that will host the zone cluster.**

3. **Configure the zone cluster from the recovered or cloned archive in the Unified Archive.**

   ```
   phys-schost-1# clzonecluster configure zone-cluster-name
   ```

   The `clzonecluster configure` *zone-cluster-name* command launches the interactive utility, where you can specify `create -a` *archive* `[`*other-options-such-as-"-x"*`]`. The archive can be either a clone archive or a recovery archive.

Note - The zone cluster members must be added to the configuration before a zone cluster can be created.

The `configure` subcommand uses the `zonecfg` command to configure a zone on each specified machine. The `configure` subcommand lets you specify properties that apply to each node of the zone cluster. These properties have the same meaning as established by the `zonecfg` command for individual zones. The `configure` subcommand supports the configuration of properties that are unknown to the `zonecfg` command. The `configure` subcommand launches an interactive shell if you do not specify the `-f`option. The `-f` option takes a command file as its argument. The `configure` subcommand uses this file to create or modify zone clusters non-interactively.

▼ How to Install a Zone Cluster from the Unified Archive

You can install a zone cluster from the Unified Archive. The `clzonecluster install` utility lets you specify the absolute path of the archive or the Oracle Solaris 10 image archive to use for the installation. See the `solaris10`(5) man page for details regarding supported archive types. The absolute path of the archive should be accessible on all the physical nodes of the cluster where the zone cluster will be installed. The Unified Archive installation can use either a *recovery* archive or a *clone* archive.

If you prefer to use the command line rather than the interactive utility to install a zone cluster from an archive, use the `clzonecluster create -a` *archive* `-z` *archived-zone* command. See the `clzonecluster`(1CL) man page for more information.

1. **Create a recovery or clone archive.**

   ```
   phys-schost# archiveadm create -r archive-location
   ```

   Use the `create` command to create a clone archive or the `-r` option to create a recovery archive. For more information about using the `archiveadm` command, see the `archiveadm`(1M) man page.

2. **Assume the `root` role on a node of the global cluster that will host the zone cluster.**

3. **Install the zone cluster from the recovered or cloned archive from the Unified Archive.**

   ```
   phys-schost-1# clzonecluster install -a absolute_path_to_archive zone-cluster-name
   ```

The absolute path of the archive should be accessible on all the physical nodes of the cluster where the zone cluster will be installed. If you have an HTTPS Unified Archive location, specify the SSL certificate, Certificate Authority (CA) certificate, and key files using **`-x cert|ca-cert|key=`***file*.

The Unified Archives do not contain zone cluster node resources. Node resources are specified when the cluster is configured. When you configure a zone cluster from a global zone by using the Unified Archives, you must set the zonepath.

If the Unified Archive contains multiple zones, use the *zone-cluster-name* to specify the zone name of the source of the installation. See the `clzonecluster`(1CL) man page for more information.

Note - If the source you used to create the Unified Archive does not contain the Oracle Solaris Cluster packages, you must run **`pkg install`** *ha-cluster-packages* (substituting the specific package name, such as *ha-cluster-minimal* or *ha-cluster-framework-full*). You will need to boot the zone, run the `zlogin` command, and then run the `pkg install` command. This action installs the same packages on the target zone cluster as the global cluster.

---

4. **Boot the new zone cluster.**

   ```
   phys-schost-1# clzonecluster boot zone-cluster-name
   ```

How to Add a Network Address to a Zone Cluster

This procedure adds a network address for use by an existing zone cluster. A network address is used to configure logical host or shared IP address resources in the zone cluster. You can run the `clsetup` utility multiple times to add as many network addresses as you need.

Note - You can also add a network address to a zone cluster by using the Oracle Solaris Cluster Manager browser interface. Click Zone Clusters, click the name of the zone cluster to go to its page, then click the Solaris Resources tab to administer zone-cluster components. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume the `root` role on a node of the global cluster that hosts the zone cluster.**

2. **On the global cluster, configure the cluster file system that you want to use in the zone cluster.**

   Start the `clsetup` utility.

```
phys-schost# clsetup
```

The Main Menu is displayed.

3. **Choose the Zone Cluster menu item.**

4. **Choose the Add Network Address to a Zone Cluster menu item.**

5. **Choose the zone cluster where you want to add the network address.**

6. **Choose the property to specify the network address you want to add.**

   `address=`*value*

   Specifies the network address used to configure logical host or shared IP address resources in the zone cluster. For example, `192.168.100.101`.

   The following types of network addresses are supported:

   - A valid IPv4 address, optionally followed by `/` and a prefix length.
   - A valid IPv6 address, which must be followed by `/` and a prefix length.
   - A hostname which resolves to an IPv4 address. Hostnames that resolve to IPv6 addresses are not supported.

   See the `zonecfg`(1M) man page for more information about network addresses.

7. **To add an additional network address, type `a`.**

8. **Type `c` to save the configuration change.**

   The results of your configuration change are displayed. For example:

```
 >>> Result of Configuration Change to the Zone Cluster(sczone) <<<

Adding network address to the zone cluster...

The zone cluster is being created with the following configuration

/usr/cluster/bin/clzonecluster configure sczone
add net
set address=phys-schost-1
end

All network address added successfully to sczone.
```

9. **When finished, exit the `clsetup` utility.**

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.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to delete a zone cluster. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume a role that provides `solaris.cluster.modify` 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 zone-cluster-name +
   ```

   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 zone-cluster-name
   ```

4. **Uninstall the zone cluster.**

   ```
   phys-schost# clzonecluster uninstall zone-cluster-name
   ```

5. **Unconfigure the zone cluster.**

   ```
   phys-schost# clzonecluster delete zone-cluster-name
   ```

Example 84 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 local file system
- StorageTek QFS standalone file system
- StorageTek QFS shared file system, when used to support Oracle RAC
- Oracle Solaris ZFS (exported as a data set)
- NFS from supported NAS devices

Zone clusters can manage loopback mounts for the following:

- UFS local file system
- StorageTek QFS standalone file system
- StorageTek QFS shared file system, only when used to support Oracle RAC
- UFS cluster file system

You configure an `HAStoragePlus` or `ScalMountPoint` resource to manage the mounting of the file system. For instructions on adding a file system to a zone cluster, see "Adding File Systems to a Zone Cluster" in *Oracle Solaris Cluster 4.3 Software Installation Guide*.

An HAStoragePlus resource does not monitor a ZFS file system if the file system has its `mountpoint` property set to `none` or `legacy`, or its `canmount` property set to `off`. For all other ZFS file systems, the HAStoragePlus resource fault monitor checks if the file system is mounted. If the file system is mounted, the HAStoragePlus resource then probes the file system's accessibility by reading and writing to it, depending on the value of the `IOOption` property called `ReadOnly/ReadWrite`.

If the ZFS file system is not mounted or the probe of the file system fails, the resource fault monitor fails and the resource is set to `Faulted`. The RGM will attempt to restart it, determined by the `retry_count` and `retry_interval` properties of the resource. This action results in remounting the file system if the specific `mountpoint` and `canmount` property settings described above are not in play. If the fault monitor continues to fail and exceeds the `retry_count` within the `retry_interval`, the RGM fails the resource over to another node.

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.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to remove a file system from a zone cluster. Click Zone Clusters, click the name of the zone cluster to go to its page, then click the Solaris Resources tab to administer zone-cluster components. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume the `root` role 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 `HAStoragePlus` and `SUNW.ScalMountPoint`, that are configured for the zone cluster's file system that you are removing.**

   ```
   phys-schost# clresource delete -F -Z zone-cluster-name 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 zone-cluster-name
   ```

4. **Remove the file system from the zone-cluster configuration.**

   ```
   phys-schost# clzonecluster configure zone-cluster-name
   ```

```
clzc:zone-cluster-name> remove fs dir=filesystemdirectory

clzc:zone-cluster-name> commit
```

The file system mount point is specified by `dir=`.

5. **Verify the removal of the file system.**

```
phys-schost# clzonecluster show -v zone-cluster-name
```

Example 85 Removing a Highly Available Local 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 `HAStoragePlus`.

```
phys-schost# clzonecluster show -v sczone
...
Resource Name:                          fs
dir:                                    /local/ufs-1
special:                                /dev/md/ds1/dsk/d0
raw:                                    /dev/md/ds1/rdsk/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 86 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.HAStoragePlus`.

```
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 Solaris Volume Manager disk sets and DID devices, from a zone cluster. Perform this procedure to remove a storage device from a zone cluster.

Note - You can also use the Oracle Solaris Cluster Manager browser interface to remove a storage device from a zone cluster. Click Zone Clusters, click the name of the zone cluster to go to its page, then click the Solaris Resources tab to administer zone-cluster components. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

---

1. **Assume the `root` role 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.HAStoragePlus` and `SUNW.ScalDeviceGroup`, that are configured for the zone cluster's devices that you are removing.

   ```
   phys-schost# clresource delete -F -Z zone-cluster dev_zone_resources
   ```

3. **Determine the match entry for the devices to be removed.**

   ```
   phys-schost# clzonecluster show -v zone-cluster
   …
   Resource Name:       device
   match:               <device_match>
   …
   ```

4. **Remove the devices from the zone-cluster configuration.**

   ```
   phys-schost# clzonecluster configure zone-cluster
   clzc:zone-cluster-name> remove device match=devices-match
   clzc:zone-cluster-name> commit
   clzc:zone-cluster-name> end
   ```

5. **Reboot the zone cluster.**

   ```
   phys-schost# clzonecluster reboot zone-cluster
   ```

6. **Verify the removal of the devices.**

```
phys-schost# clzonecluster show -v zone-cluster
```

Example 87 Removing a Solaris Volume Manager Disk Set From a Zone Cluster

This example shows how to remove a Solaris Volume Manager 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 88 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#
phys-schost# clzonecluster show -v sczone
```

Troubleshooting Procedures for Testing Purposes

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.**

      The following command finds the Persistent Group Reservation Emulation (PGRE) keys. If there are no keys on the disk, an `errno=22` message is displayed.

      ```
      # /usr/cluster/lib/sc/pgre -c pgre_inkeys -d /dev/did/rdsk/dids2
      ```

      After you locate the keys, scrub the PGRE keys.

```
# /usr/cluster/lib/sc/pgre -c pgre_scrub -d /dev/did/rdsk/dids2
```

Caution - If you scrub the active quorum device keys from the disk, the cluster will panic on the next reconfiguration with a `Lost operational quorum` message.

2. **Evacuate the global-cluster node that you want to boot in noncluster mode.**

```
phys-schost# clresourcegroup evacuate -n target-node
```

3. **Take offline any resource group or resource groups that contain 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 resource-group
```

4. **Disable all the resources in the resource groups that you took offline.**

```
phys-schost# clresource disable resource
```

5. **Unmanage the resource groups.**

```
phys-schost# clresourcegroup unmanage resource-group
```

6. **Take offline the corresponding device group or device groups.**

```
phys-schost# cldevicegroup offline device-group
```

7. **Disable the device group or device groups.**

```
phys-schost# cldevicegroup disable device-group
```

8. **Boot the passive node into noncluster mode.**

```
phys-schost# shutdown -g0 -i0 -y
ok> boot -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/rdsk/dids2
```

11. **If any SCSI3 reservations exist on the disks, scrub them.**

```
phys-schost# /usr/cluster/lib/sc/scsi -c scrub -d /dev/did/rdsk/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 device-group
```

17. **Start the resource group or resource groups.**

```
phys-schost# clresourcegroup online -eM resource-group
```

Restoring a Corrupted Disk Set

Use this procedure if a disk set is corrupted or in a state that the nodes in the cluster are unable to take ownership of the disk set. If your attempts to clear the state have failed, use this procedure as a last attempt to fix the disk set.

These procedures work for Solaris Volume Manager metasets and multi-owner Solaris Volume Manager metasets.

▼ 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 Oracle Explorer (`SUNWexplo`) to create a backup. You can then use the saved configuration to recreate the disk set. You should save the current configuration into files (using the prtvtoc and metastat commands), and then recreate the disk set and its components. See "How to Recreate the Solaris Volume Manager Software Configuration" on page 289.

1. **Save the partition table for each disk in the disk set.**

```
# /usr/sbin/prtvtoc /dev/global/rdsk/disk-name > /etc/lvm/disk-name.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 set-name >> /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 Oracle Explorer (SUNWexplo) is run on a node that owns the set, it retrieves the prtvtoc and metaset -p information.

How to Purge the Corrupted Disk Set

Purging a set from a node or all nodes removes the configuration. To purge a disk set from a node, the node must not have ownership of the disk set.

1. **Run the purge command on all nodes.**

```
# /usr/sbin/metaset -s set-name -P
```

Running this command removes the disk set information from the database replicas, as well as the Oracle Solaris Cluster repository. The -P and -C options allow a disk set to be purged without the need to completely rebuild the Solaris Volume Manager environment.

Note - If a multi-owner disk set is purged while the nodes were booted out of cluster mode, you might need to remove the information from the DCS configuration files.

```
# /usr/cluster/lib/sc/dcs_config -c remove -s set-name
```

For more information, see the dcs_config(1M) man page.

---

2. **If you want to remove only the disk set information from the database replicas, use the following command.**

```
# /usr/sbin/metaset -s set-name -C purge
```

You should generally use the `-P` option, rather than the `-C` option. Using the `-C` option can cause a problem recreating the disk set because the Oracle Solaris Cluster software still recognizes the disk set.

a. **If you used the `-C` option with the `metaset` command, first create the disk set to see if a problem occurs.**

b. **If a problem exists, remove the information from the DCS configuration files.**

```
# /usr/cluster/lib/sc/dcs_config -c remove -s setname
```

If the purge options fail, verify that you installed the latest kernel and metadevice updates and contact My Oracle Support.

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 disk set.

Note - Use mediators only on two-node clusters.

---

1. **Create a new disk set.**

   ```
   # /usr/sbin/metaset -s set-name -a -h node1 node2
   ```

   If this is a multi-owner disk set, use the following command to create a new disk set.

   ```
   /usr/sbin/metaset -s set-name -aM -h node1 node2
   ```

2. **On the same host where the set was created, add mediator hosts if required (two nodes only).**

   ```
   /usr/sbin/metaset -s set-name -a -m node1 node2
   ```

3. **Add the same disks back into the disk set from this same host.**

   ```
   /usr/sbin/metaset -s set-name -a /dev/did/rdsk/disk-name /dev/did/rdsk/disk-name
   ```

4. **If you purged the disk set 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/`*disk-name*`.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 metadevice.**

   ```
   # /usr/sbin/metainit -s set-name -n -a metadevice
   ```

6. **Create each metadevice from a saved configuration.**

   ```
   # /usr/sbin/metainit -s set-name -a metadevice
   ```

7. **If a file system exists on the metadevice, run the `fsck` command.**

   ```
   # /usr/sbin/fsck -n /dev/md/set-name/rdsk/metadevice
   ```

   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 metadevice correctly. If you have, review the `fsck` errors to determine if the file system can be recovered. If it cannot, you should restore the data from a backup.

8. **Concatenate all other metasets on all cluster nodes to the `/etc/lvm/md.tab` file and then concatenate the local disk set.**

   ```
   # /usr/sbin/metastat -p >> /etc/lvm/md.tab
   ```

♦ ♦ ♦ CHAPTER 10

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

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 *Introduction to Oracle Solaris Zones*.

On the Oracle Solaris OS, you can do the following:

- Assign CPU shares to resource groups
- Assign processors to resource groups

Note - You can also use the Oracle Solaris Cluster Manager browser interface to view the configuration of a zone-cluster. For Oracle Solaris Cluster Manager log-in instructions, see "How to Access Oracle Solaris Cluster Manager" on page 313.

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`.

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 describes how to control CPU usage in a global-cluster node.

▼ How to Control CPU Usage in a Global-Cluster Node

Perform this procedure to assign CPU shares to a resource group that will be executed in a global–cluster 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 node:

- Augments the number of CPU shares assigned to the node (`zone.`*cpu-shares*) with the specified number of CPU shares, if this has not already been done.
- Creates a project named `SCSLM_`*resource-group* in the 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_`*resource-group* 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 scheduler 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 nodes and the minimum number of CPUs available in the default processor set.**

   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=`*integer*

   Sets the minimum number of CPUs available in the default processor set. The default value is `1`.

   `-p globalzoneshares=`*integer*

   Sets the number of shares assigned to the 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 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
   ```

   `-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 node zone.cpu-shares.

resource-group

Specifies the name of the resource group.

In this procedure, you do not set the RG_SLM_PSET_TYPE property. In the 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 -eM resource-group
```

resource-group Specifies the name of the resource group.

Note - Do not remove or modify the SCSLM_*resource-group* 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.

♦ ♦ ♦ CHAPTER 11

Updating Your Software

This chapter provides information and instructions for updating Oracle Solaris Cluster software in the following sections.

Overview of Updating Oracle Solaris Cluster Software

All cluster member nodes must have the same updates applied for proper cluster operation. When updating a node, you might occasionally need to temporarily remove a node from cluster membership or stop the entire cluster before performing the update.

There are two ways to update Oracle Solaris Cluster software.

- **Upgrade** - Upgrade the cluster to the latest major or minor Oracle Solaris Cluster release and update the Oracle Solaris OS by updating all packages. An example of a major release would be to upgrade from Oracle Solaris Cluster 4.0 to 5.0. An example of a minor release would be to upgrade from Oracle Solaris Cluster 4.1 to 4.2.

  Run the `scinstall` utility or the `scinstall -u update` command to create a new boot environment (a bootable instance of an image), mount the boot environment on a mount point that is not being used, update the bits, and activate the new boot environment. Creating the clone environment initially consumes no additional space and occurs instantaneously. After you perform this update, you must reboot the cluster. The upgrade also upgrades the Oracle Solaris OS to the latest compatible version. For detailed instructions, see the *Oracle Solaris Cluster 4.3 Upgrade Guide*.

  If you have failover zones of brand type `solaris`, follow the instructions in "How to Upgrade a solaris Branded Failover Zone" in *Oracle Solaris Cluster 4.3 Upgrade Guide*.

  If you have a `solaris10` branded zone in a zone cluster, follow the upgrade instructions in "Upgrading a solaris10 Branded Zone in a Zone Cluster" in *Oracle Solaris Cluster 4.3 Upgrade Guide*.

To update a zone cluster that is not branded, follow procedures in "Updating a Specific Package" on page 297 to update the underlying global cluster. When the global cluster is updated, its zone clusters are automatically updated as well.

Note - Applying an Oracle Solaris Cluster Core SRU does not provide the same result as upgrading the software to another Oracle Solaris Cluster release.

- **Update** - Update specific Oracle Solaris Cluster packages to different SRU levels. You can use one of the `pkg` commands to update Image Packaging System (IPS) packages in a Support Repository Update (SRU).

  SRUs are generally released regularly and contain updated packages and defect fixes. The repository contains all IPS packages and the updated packages. Running the `pkg update` command updates both the Oracle Solaris operating system and the Oracle Solaris Cluster software to compatible versions. After you perform this update, you might need to reboot the cluster. For instructions, see "How to Update a Specific Package" on page 298.

  You must be a registered My Oracle Support user to view and download the required software updates for the Oracle Solaris Cluster product. If you do not have a My Oracle Support account, contact your Oracle service representative or sales engineer, or register online at http://support.oracle.com. For information about firmware updates, see your hardware documentation.

Note - Read the software update README before applying or removing any update.

Information for using the Oracle Solaris package management utility, `pkg`, is provided in Chapter 3, "Installing and Updating Software Packages" in *Adding and Updating Software in Oracle Solaris 11.3*.

Updating Oracle Solaris Cluster Software

Consult the following table to determine how to upgrade or update an Oracle Solaris Cluster release or package in the Oracle Solaris Cluster software.

TABLE 20 Updating Oracle Solaris Cluster Software

Task	Instructions
Upgrade the entire cluster to a new major or minor release	"How to Upgrade the Software (Standard Upgrade)" in *Oracle Solaris Cluster 4.3 Upgrade Guide*
Review tips for successful updates	"Updating Tips" on page 297

Task	Instructions
Update a specific package	"How to Update a Specific Package" on page 298
Update a quorum server or AI installation server	"How to Update a Quorum Server or AI Installation Server" on page 301
Update a `solaris` branded zone cluster	"How to Update a `solaris` Branded Zone Cluster" on page 299
Patch a `solaris10` branded zone cluster	"How to Patch a `solaris10` Branded Zone Cluster (clzonecluster)" on page 300
	"How to Patch a `solaris10` Branded Zone in a Zone Cluster (`patchadd`)" on page 300
	"How to Patch an Alternate Boot Environment for a Zone in a `solaris10` Branded Zone Cluster" on page 301
Remove Oracle Solaris Cluster packages	"How to Uninstall a Package" on page 302
	"How to Uninstall Quorum Server or AI Installation Server Packages" on page 302

Updating Tips

Use the following tips to administer Oracle Solaris Cluster updates more efficiently:

- Read the SRU's `README` file before performing the update.
- Check the update requirements of your storage devices.
- Apply all updates before running the cluster in a production environment.
- Check the hardware firmware levels and install any required firmware updates that might be needed. Consult your hardware documentation for information on firmware updates.
- All nodes acting as cluster members must have the same updates.
- Keep cluster subsystem updates up to date. These updates include, for example, volume management, storage device firmware, and cluster transport.
- Test failover after major updates. Be prepared to back out the update if cluster operation is degraded or impaired.

Updating a Specific Package

IPS packages were introduced with the Oracle Solaris 11 operating system. Each IPS package is described by a Fault Managed Resource Indicator (FMRI), and you use the `pkg` command to perform the SRU update. Alternatively, you can also use the `scinstall -u` command to perform an SRU update.

You might want to update a specific package to use an updated Oracle Solaris Cluster data service agent.

▼ How to Update a Specific Package

Perform this procedure on each global cluster node you want to update. Any unbranded zone cluster on the cluster node will automatically also receive this update.

1. **Assume a role that provides `solaris.cluster.admin` authorization.**

2. **Update the package.**

   For example, to update a package from a specific publisher, specify the publisher name in the *pkg-fmri*.

   ```
   # pkg update pkg-fmri
   ```

   Caution - If you use the `pkg update` command with no *pkg-fmri* specified, all installed packages that have updates available are updated.

   If a newer version of an installed package is available and is compatible with the rest of the image, the package is updated to that version. If the package contains binaries that have the `reboot-needed` flag set to true, then performing a `pkg update` *pkg-fmri* automatically creates a new boot environment and after the update you boot into the new boot environment. If the package you are updating does not contain any binaries that force a reboot, then the `pkg update` command updates the live image and a reboot is not necessary.

3. **If you are updating a data service agent (`ha-cluster/data-service/*` or the generic data service agent of `ha-cluster/ha-service/gds`), run the following commands.**

   ```
   # pkg change-facet facet.version-lock.pkg-name=false
   # pkg update pkg-name
   ```

   For example:

   ```
   # pkg change-facet facet.version-lock.ha-cluster/data-service/weblogic=false

   # pkg update ha-cluster/data-service/weblogic
   ```

   Note - If you want to freeze an agent to prevent it from being updated, run the following commands.

   ```
   # pkg change-facet facet.version-lock.pkg-name=false
   # pkg freeze pkg-name
   ```

   For more information on freezing a specific agent, see "Controlling Installation of Optional Components" in *Adding and Updating Software in Oracle Solaris 11.3*.

4. **Verify that the package was updated.**

```
# pkg verify -v pkg-fmri
```

Updating or Patching a Zone Cluster

To update a `solaris` branded zone cluster, you apply an SRU by using the `scinstall-u update` command.

To patch a `solaris10` branded zone cluster, apply a patch by using the `clzonecluster install-cluster -p` command or the `patchadd` command, or by patching an alternate boot environment.

This section provides the following procedures:

- “How to Update a `solaris` Branded Zone Cluster” on page 299
- “How to Patch a `solaris10` Branded Zone Cluster (clzonecluster)” on page 300
- “How to Patch a `solaris10` Branded Zone in a Zone Cluster (`patchadd`)” on page 300
- “How to Patch an Alternate Boot Environment for a Zone in a `solaris10` Branded Zone Cluster” on page 301

▼ How to Update a `solaris` Branded Zone Cluster

Perform this procedure to update a `solaris` branded zone cluster by using the `scinstall-u update` command to apply an SRU.

1. **Assume a role that provides `solaris.cluster.admin` authorization on a node of the global cluster.**

2. **From a node of the global cluster, update the entire node.**

```
phys-schost# scinstall -u update [-b be-name]
```

Repeat this step on each cluster node.

3. **Reboot the cluster.**

```
phys-schost# clzonecluster reboot
```

How to Patch a `solaris10` Branded Zone Cluster (clzonecluster)

Perform this procedure from one node of the global cluster.

1. **Assume a role that provides `solaris.cluster.admin` authorization on a node of the global cluster.**
2. **Boot the zone cluster into Offline Running mode.**
   - **If the zone cluster is not booted, run the following command.**

     `phys-schost#` **`clzonecluster boot -o`** *zone-cluster*

   - **If the zone cluster is booted, run the following command.**

     `phys-schost#` **`clzonecluster reboot -o`** *zone-cluster*

3. **From one node of the global cluster, patch the entire `solaris10` branded zone cluster.**

   `phys-schost#` **`clzonecluster install-cluster -p`** *patch-spec* **[***options***]** *zone-cluster*

   For more information about the `install-cluster` subcommand, see the `clzonecluster`(1CL) man page.

   Note - If the `scinstall -u update` command fails with a message that the `pkg` command is out of date, follow the instructions in the message.

   ---

4. **Reboot the zone cluster.**

   `phys-schost#` **`clzonecluster reboot`** *zone-cluster*

How to Patch a `solaris10` Branded Zone in a Zone Cluster (`patchadd`)

Perform the following steps on each configured zone-cluster node.

1. **Bring the zone cluster into an offline-running state.**

   `phys-schost#` **`clzonecluster reboot -o`** *zone-cluster*

2. **Log in to the zone-cluster node.**

```
phys-schost# zlogin zone-cluster
```

3. **From the command line, type `patchadd`.**

```
zchost# patchadd
```

4. **Install the patches that correspond to the new release.**

How to Patch an Alternate Boot Environment for a Zone in a `solaris10` Branded Zone Cluster

Use this procedure to patch an alternate boot environment (BE) for the zone of zone-cluster node. This method provides the option of backing out of the patched BE if necessary.

1. **Become administrator for the zone of the zone-cluster node you want to patch.**

2. **Create, patch, and activate a new boot environment, then use the `shutdown` command to reboot the zone.**

   Follow instructions for these tasks that are contained in "How to Create and Activate Multiple Boot Environments on a solaris10 Branded Zone" in *Creating and Using Oracle Solaris 10 Zones*.

3. **Repeat the preceding steps for the zones of any other zone-cluster nodes that you need to patch.**

Updating a Quorum Server or AI Installation Server

Use the procedure below to update the packages for your quorum server or Oracle Solaris 11 Automated Installer (AI) installation server. For more information about quorum servers, see "How to Install and Configure Oracle Solaris Cluster Quorum Server Software" in *Oracle Solaris Cluster 4.3 Software Installation Guide*. For more information about using the AI, see "How to Install and Configure Oracle Solaris and Oracle Solaris Cluster Software (IPS Repositories)" in *Oracle Solaris Cluster 4.3 Software Installation Guide*.

How to Update a Quorum Server or AI Installation Server

1. **Assume a role that provides `solaris.cluster.admin` authorization.**

2. **Update the quorum server or AI installation server packages.**

```
# pkg update ha-cluster/*
```

If a newer version of the installed `ha-cluster` packages is available and is compatible with the rest of the image, the packages are updated to that version.

> **Caution -** Running the `pkg update` command updates all `ha-cluster` packages installed on the system.

Uninstalling a Package

You can remove a single package or multiple packages.

How to Uninstall a Package

1. **Assume a role that provides `solaris.cluster.admin` authorization.**
2. **Uninstall one or more existing packages.**

```
# pkg uninstall pkg-fmri [pkg-fmri …]
```

The `pkg uninstall` command will fail if there are other installed packages that depend on the *pkg-fmri* you are uninstalling. To uninstall the *pkg-fmri*, you must supply the `pkg uninstall` command with all the *pkg-fmri* dependents. For additional information on uninstalling packages, see *Adding and Updating Software in Oracle Solaris 11.3* and the `pkg(1)` man page.

How to Uninstall Quorum Server or AI Installation Server Packages

1. **Assume a role that provides `solaris.cluster.admin` authorization.**
2. **Uninstall the quorum server or AI installation server packages.**

```
# pkg uninstall ha-cluster/*
```

> **Caution -** This command uninstalls all `ha-cluster` packages installed on the system.

◆ ◆ ◆ CHAPTER 12

Backing Up and Restoring a Cluster

This chapter provides the following sections:

Backing Up a Cluster

Before you back up your cluster, find the names of the file systems you want to back up, calculate how many tapes you need to contain a full backup, and back up the ZFS root file system.

TABLE 21 Task Map: Backing Up Cluster Files

Task	Instructions
Perform online backup for mirrored or plexed file systems	"How to Perform Online Backups for Mirrors (Solaris Volume Manager)" on page 303
Back up the cluster configuration	"How to Back Up the Cluster Configuration" on page 305
Back up disk partitioning configuration for storage disk	See the documentation for your storage disk

▼ 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. **Assume an equivalent role on the cluster node that you are backing up.**

2. **Use the `metaset` command to determine which node has the ownership on the backed-up volume.**

   ```
   # metaset -s setname
   ```

   `-s` *setname* — Specifies the disk set name.

   For more information, see the `metaset(1M)` man page.

3. **Use the `lockfs` command with the `-w` option to lock the file system from writes.**

   ```
   # lockfs -w mountpoint
   ```

   See the `lockfs(1M)` man page for more information.

4. **Use the `metastat` 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.

   See the `metastat(1M)` man page for more information.

5. **Use the `metadetach` command to take one submirror offline from the mirror.**

   ```
   # metadetach -s setname mirror submirror
   ```

   See the `metadetach(1M)` man page for more information.

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/rdsk/submirror
   ```

8. **Back up the offline submirror to tape or another medium.**

   Note - Use the raw device (`/rdsk`) name for the submirror, rather than the block device (`/dsk`) name.

   ---

9. **Use the `metattach` command to place the metadevice or volume back online.**

   ```
   # metattach -s setname mirror submirror
   ```

   When the metadevice or volume is placed online, it is automatically resynchronized with the mirror. See the `metattach`(1M) man page for more information.

10. **Use the `metastat` command to verify that the submirror is resynchronizing.**

    ```
    # metastat -s setname mirror
    ```

    See *Managing ZFS File Systems in Oracle Solaris 11.3* for more information.

▼ 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 assume a role that provides `solaris.cluster.read` authorization.**

2. **Export the cluster configuration information to a file.**

   `# 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 the `clconfiguration`(5CL) man page.

3. **Verify that the cluster configuration information was successfully exported to the XML file.**

   `# pfedit` *configfile*

Restoring Cluster Files

You can restore the ZFS root file system to a new disk.

You can restore a cluster or node from a unified archive, or you can restore specific files or file systems.

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 22 Task Map: Restoring Cluster Files

Task	Instructions
For Solaris Volume Manager, restore the ZFS root (`/`) file system	"How to Restore the ZFS Root (/) File System (Solaris Volume Manager)" on page 306

▼ How to Restore the ZFS Root (/) File System (Solaris Volume Manager)

Use this procedure to restore the ZFS 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. UFS is supported, except as a root file system. UFS can be used on metadevices in Solaris Volume Manager metasets on shared disks.

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. **Assume a role that provides `solaris.cluster.modify` authorization on a cluster node with access to the disk sets to which the node to be restored is also attached.**

   Use a node *other than* the node that you are restoring.

2. **Remove from all metasets the hostname of the node being restored.**

   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: Rpcbind 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 ZFS root file system (/).**

   For more information, see "Replacing Disks in a ZFS Root Pool" in *Managing ZFS File Systems in Oracle Solaris 11.3*.

   To recover the ZFS root pool or root pool snapshots, follow the procedure in "Replacing Disks in a ZFS Root Pool" in *Managing ZFS File Systems in Oracle Solaris 11.3*.

Note - Ensure that you create the `/global/.devices/node@`*nodeid* file system.

If the /.globaldevices backup file exists in the backup directory, it is restored along with ZFS root restoration. The file is not created automatically by the globaldevices SMF service.

4. **Reboot the node in multiuser mode.**

```
# reboot
```

5. **Replace the device ID.**

```
# cldevice repair root-disk
```

6. **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.
-f *raw-disk-device*	Raw disk device on which to create replicas.
-a	Adds replicas.

See the metadb(1M) man page for more information.

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 89 Restoring the ZFS Root (/) File System (Solaris Volume Manager)

The following example shows the root (/) file system restored to the node phys-schost-1. 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/rdsk/c0t0d0s0, and three state database replicas are recreated on /dev/rdsk/c0t0d0s4. For more information on restoring data, see "Resolving Data Problems in a ZFS Storage Pool" in *Managing ZFS File Systems in Oracle Solaris 11.3*.

```
    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 procedures in Oracle Solaris documentation

Reboot the node
```
# reboot
```

Replace the disk ID
```
# cldevice repair /dev/dsk/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
```

♦ ♦ ♦ CHAPTER 13

Using the Oracle Solaris Cluster Manager Browser Interface

This chapter provides a description of the Oracle Solaris Cluster Manager browser interface that you can use to administer many aspects of a cluster. The chapter also contains procedures to access and use Oracle Solaris Cluster Manager.

Note - Oracle Solaris Cluster Manager uses a private version of Oracle GlassFish Server software, which is shipped with the Oracle Solaris Cluster product. Do not attempt to install, or update to any patchsets, the public version of the Oracle GlassFish Server software. Doing so might cause package problems when updating Oracle Solaris Cluster software or installing Oracle Solaris Cluster SRUs. Any bug fixes to the private version of Oracle GlassFish Server that are needed by Oracle Solaris Cluster Manager are delivered in Oracle Solaris Cluster SRUs.

This chapter includes the following:

Overview of Oracle Solaris Cluster Manager

The Oracle Solaris Cluster Manager browser interface 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, including administering the following Oracle Solaris Cluster components:

- Data services

- Zone clusters
- Nodes
- Private adapters, cables, and switches
- Quorum devices
- Device groups
- Disks
- NAS devices
- Node load limits
- Resource groups and resources
- Geographic Edition partnerships

Oracle Solaris Cluster Manager currently cannot perform all Oracle Solaris Cluster administrative tasks. You must use the command-line interface for some operations.

User Rights for Oracle Solaris Cluster Manager

Oracle Solaris Cluster software provides rights profiles to identify who has rights to administer the cluster, including operations performed by using Oracle Solaris Cluster Manager. You can assign these rights profiles to users or to roles to give users different levels of access to Oracle Solaris Cluster software.

Note the following behaviors for users or roles when logged in to Oracle Solaris Cluster Manager:

- If a user does not have sufficient rights to perform certain actions in Oracle Solaris Cluster Manager, such as Delete, the buttons for those actions are not displayed.
- If you change the assigned rights for a user while that user is logged in to Oracle Solaris Cluster Manager, the user must log out and log back in for the changes to take effect. Otherwise, the user might experience unexpected behavior, such as greyed-out or missing action buttons that should be active for the newly-added rights.

For more information about how to set up and manage user rights for Oracle Solaris Cluster software, see Chapter 2, “Oracle Solaris Cluster and User Rights”.

Accessing the Oracle Solaris Cluster Manager Software

The Oracle Solaris Cluster Manager browser interface provides an easy way to administer many tasks in the Oracle Solaris Cluster software. See the Oracle Solaris Cluster Manager online help for more information.

The common agent container is started automatically when you boot the cluster. If you need to verify that the common agent container is running, see "Troubleshooting Oracle Solaris Cluster Manager" on page 314.

Tip - Do not click Back in the browser to exit from Oracle Solaris Cluster Manager.

How to Access Oracle Solaris Cluster Manager

This procedure shows how to access Oracle Solaris Cluster Manager on your cluster.

1. **From a machine that can make a network connection to the cluster, start a browser.**

   - Ensure that the browser's disk and memory cache sizes are set to a value that is greater than 0.
   - Verify that Java and Javascript are enabled in the browser.

2. **From the browser, connect to the Oracle Solaris Cluster Manager port on one node of the cluster.**

   The default port number is 8998.

   `https://`*node***`:8998/scm`**

3. **Accept any certificates that are presented by the web browser.**

   The Oracle Solaris Cluster Manager login page is displayed.

4. **Enter the name of a node in the cluster that you want to manage or accept the default of `localhost` to manage the cluster node that you are already connected to.**

5. **Enter the user name and password that is authorized to connect to the node.**

   To use Oracle Solaris Cluster Manager to update or manage the cluster, log in as the `root` role or an authorized role that provides `solaris.cluster.modify` and `solaris.cluster.admin` authorizations.

6. **Click Sign In.**

   The Oracle Solaris Cluster Manager application launch page is displayed.

Note - If you have more than one cluster configured, you can select Other from the drop-down list and log into another cluster to display information for that cluster. If a cluster is a member of one or more partnerships, after you visit the Partnerships folder all partner names are automatically added to the drop-down list. After you authenticate, you can select Switch Cluster.

If you cannot connect to Oracle Solaris Cluster Manager, see "Troubleshooting Oracle Solaris Cluster Manager" on page 314. If you choose a restricted network profile during Oracle Solaris installation, external access for Oracle Solaris Cluster Manager is restricted. This network is required to use the Oracle Solaris Cluster Manager browser interface.

Troubleshooting Oracle Solaris Cluster Manager

- Verify that the two manager services are running.

  ```
  # svcs system/cluster/manager\*

  STATE      STIME    FMRI
  online     Oct_30   svc:/system/cluster/manager-glassfish3:default
  online     Oct_30   svc:/system/cluster/manager:default
  ```

  Use the `svcadm` command to disable or enable `system/cluster/manager-glassfish3`. This action stops and restarts the application server, respectively. You should keep `system/cluster/manager` online. You do not need to disable or enable it.
- If you cannot connect to Oracle Solaris Cluster Manager, determine whether the common agent container is running by entering `usr/sbin/cacaoadm status`. If the common agent container is not running, you will get the login page but you cannot authenticate.

  You can manually start the common agent container by entering `/usr/sbin/cacaoadm start`.

▼ How to Configure 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.

Cacao connection errors can occur when you are using the data service configuration wizards in Oracle Solaris Cluster Manager or performing other Oracle Solaris Cluster Manager tasks. This procedure copies the security files for the common agent container to all cluster nodes. This ensures that security files for the common agent container are identical on all cluster nodes and that the copied files retain the correct file permissions. Performing this procedure synchronizes the security keys.

1. **On each node, stop the security file agent.**

```
phys-schost# /usr/sbin/cacaoadm stop
```

2. **On one node, change to the `/etc/cacao/instances/default/` directory.**

```
phys-schost-1# cd /etc/cacao/instances/default/
```

3. **Create a tar file of the `/etc/cacao/instances/default/` directory.**

```
phys-schost-1# tar cf /tmp/SECURITY.tar security
```

4. **Copy the `/tmp/Security.tar` file to each of the cluster nodes.**

5. **On each node where you copied the`/tmp/SECURITY.tar` file, extract the security files.**

   Any security files that already exist in the `/etc/cacao/instances/default/` directory are overwritten.

```
phys-schost-2# cd /etc/cacao/instances/default/
phys-schost-2# tar xf /tmp/SECURITY.tar
```

6. **Delete each copy of the tar file to avoid security risks.**

   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
```

7. **On each node, start the security file agent.**

```
phys-schost# /usr/sbin/cacaoadm start
```

▼ How to Check the Network Bind Address

If you receive a System Error message when you try to view information about a node other than the node running Oracle Solaris Cluster Manager, check whether the common agent container network-bind-address parameter is set to the correct value of `0.0.0.0`.

Perform the following steps on each node of the cluster.

1. **Determine the network bind address.**

```
phys-schost# cacaoadm list-params | grep network
network-bind-address=0.0.0.0
```

If the network bind address is set to anything other than 0.0.0.0, you will need to change it to the desired address.

2. **Stop and start cacao before and after the change.**

```
phys-schost# cacaoadm stop
phys-schost# cacaoadm set-param network-bind-address=0.0.0.0
phys-schost# cacaoadm start
```

Configuring Accessibility Support for Oracle Solaris Cluster Manager

Oracle Solaris Cluster Manager provides the following accessibility settings:

- Screen reader
- High contrast
- Large fonts

To set one or more of these controls, click Accessibility in the Oracle Solaris Cluster Manager top menu bar. Then click the checkbox of the accessibility control you want to enable or disable.

Changes to accessibility control settings do not persist beyond the login session. The settings do persist if you authenticate to another cluster node during the same login session.

Using Topology to Monitor the Cluster

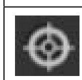
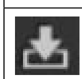
The Topology view helps you monitor your cluster and identify problems. You can quickly view relationships between objects and see which resource groups and resource belong to each node.

▼ How to Use Topology to Monitor and Update Your Cluster

To get to the Topology page, log into Oracle Solaris Cluster Manager, click Resource Groups, and click the Topology tab. Lines represent dependency and colocation relationships.

The online help provides detailed instructions on the elements of the view, as well as how to select an object for filtering the view, and right-clicking to see a context menu of actions for that object. You can collapse or restore the online help by clicking the arrow next to it. You can also collapse or restore the Filter.

The following table provides a list of the controls on the Resource Topology page.

Control	Function	Description
	Zoom	Magnify or shrink a portion of a page.
	Overview	Drag the viewport over the diagram to pan the view.
	Isolate	Single-click a resource group or resource to remove all other objects from the display.
	Drill	Single-click a resource group to drill into its resources.
	Reset	Return to full view after isolating or drilling.
	Filter	Narrow what appears by selecting objects by type, instance, or status.

The following procedure shows how to monitor your cluster nodes for critical errors:

1. **In the Topology tab, locate the Potential Masters area.**

2. **Zoom in to see the status of each node in the cluster.**

3. **Locate a node that has a red Critical status icon and right-click the node and select Show Details.**

4. **From the node's Status page, click** *System Log* **to view and filter the log messages.**

♦ ♦ ♦ A P P E N D I X A

Deployment Example: Configuring Host-Based Data Replication Between Clusters With Availability Suite Software

This appendix provides a method of using host-based replication without the use of Oracle Solaris Cluster Geographic Edition. Use Oracle Solaris Cluster Geographic Edition for host-based replication to simplify the configuration and operation of host-based replication between clusters. See "Understanding Data Replication" on page 103.

The example in this appendix shows how to configure host-based data replication between clusters that use the Availability Suite feature of Oracle Solaris 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. 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) to access the cluster nodes, ensure that you can assume a role that provides authorization for l Oracle Solaris Cluster commands. This series of data replication procedures requires the following Oracle Solaris Cluster authorizations:

- `solaris.cluster.modify`
- `solaris.cluster.admin`
- `solaris.cluster.read`

See the *Securing Users and Processes in Oracle Solaris 11.3* for more information about using roles. See the Oracle Solaris Cluster man pages for the authorization that each Oracle Solaris Cluster subcommand requires.

Understanding Availability Suite Software in a Cluster

This section introduces disaster tolerance and describes the data replication methods that Availability Suite software uses.

Disaster tolerance is the ability to restore an application on an alternate cluster when the primary cluster fails. Disaster tolerance is based on *data replication* and *takeover*. A takeover relocates an application service to a secondary cluster by bringing online one or more resource groups and device groups.

If data is replicated synchronously between the primary and secondary cluster, then no committed data is lost when the primary site fails. However, if data is replicated asynchronously, then some data may not have been replicated to the secondary cluster before the primary site failed, and thus is lost.

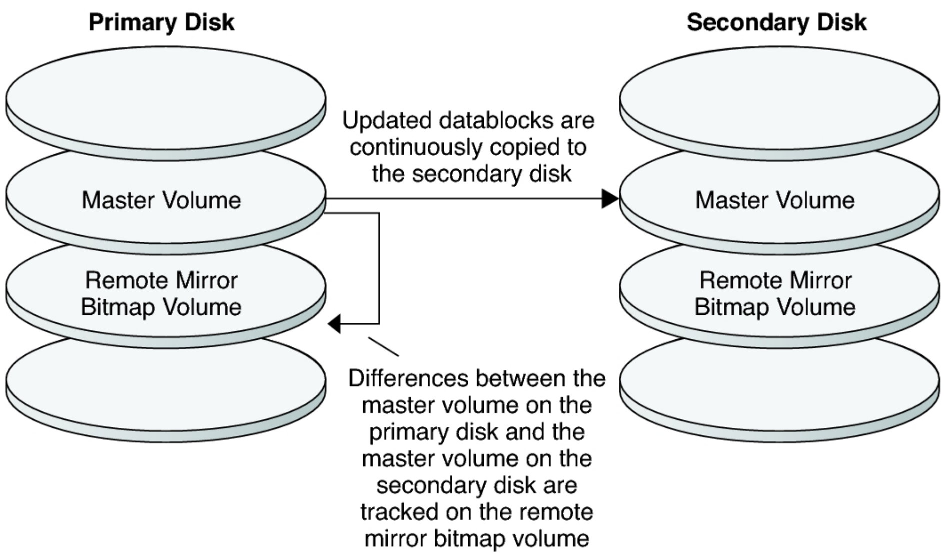
Data Replication Methods Used by Availability Suite Software

This section describes the remote mirror replication method and the point-in-time snapshot method used by Availability Suite software. This software uses the `sndradm` and `iiadm` commands to replicate data. For more information, see the `sndradm`(1M) and `iiadm`(1M) man pages.

Remote Mirror Replication

Figure 2, “Remote Mirror Replication,” on page 321 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 2 Remote Mirror Replication



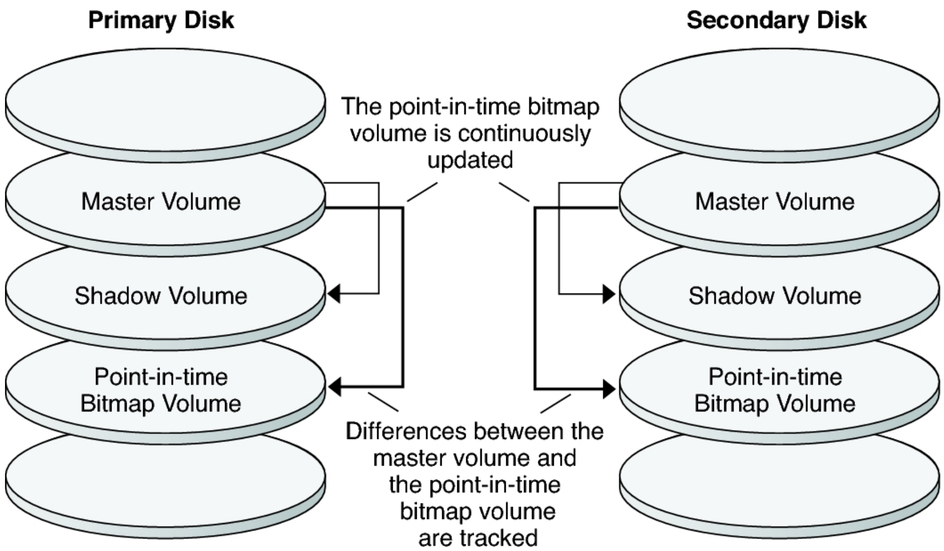
Remote mirror replication can be performed synchronously in real time, or asynchronously. Each volume set in each cluster can be configured individually, for synchronous replication or asynchronous replication.

- 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 3, "Point-in-Time Snapshot," on page 322 shows a 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.

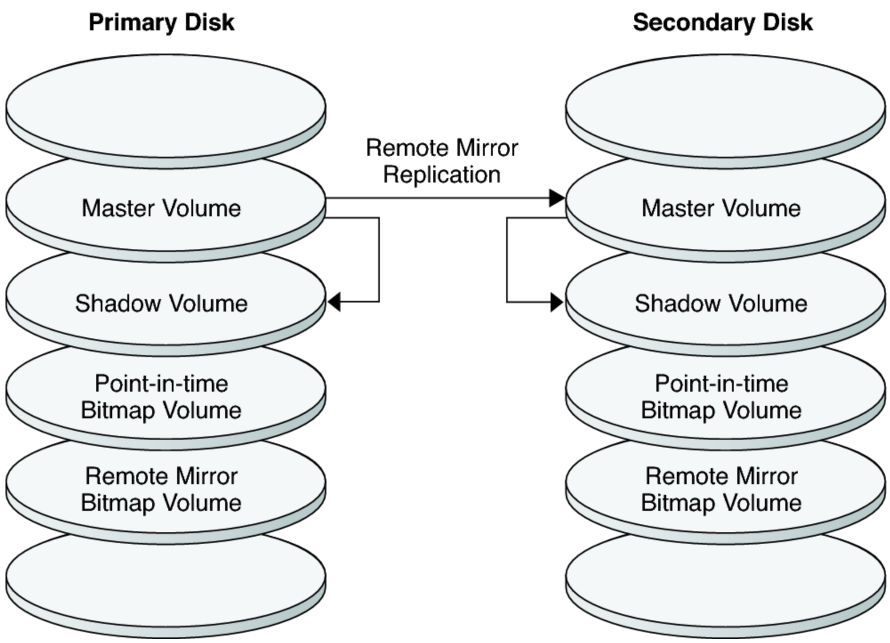
FIGURE 3 Point-in-Time Snapshot



Replication in the Example Configuration

Figure 4, "Replication in the Example Configuration," on page 323 illustrates how remote mirror replication and point-in-time snapshot are used in this example configuration.

FIGURE 4 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 324
- "Configuring Application Resource Groups" on page 324
  - "Configuring Resource Groups for a Failover Application" on page 325
  - "Configuring Resource Groups for a Scalable Application" on page 326
- "Guidelines for Managing a Takeover" on page 328

Configuring Replication Resource Groups

Replication resource groups collocate the device group under Availability Suite software control with a logical hostname resource. A logical hostname must exist on each end of the data replication stream, and must be on the same cluster node that acts as the primary I/O path to the device. 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

  A logical hostname is hosted on one node of each cluster (primary and secondary) and is used to provide source and target addresses for the Availability Suite software data replication stream.
- 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*.

  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.

The `ZPoolsSearchDir` extension property must be defined in the HAStoragePlus resource. This extension property is required to use the ZFS file system.

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 325
- "Configuring Resource Groups for a Scalable Application" on page 326

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 file system or zpool 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.

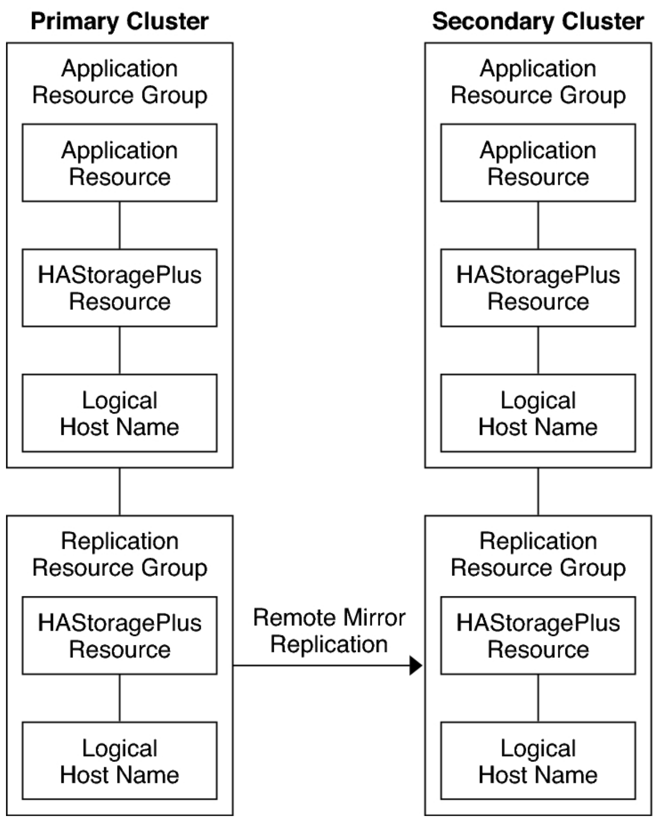
  - If the application data is globally mounted, the presence of an HAStoragePlus resource in the application resource group is not required but is advised.
  - If the application data is mounted locally, the presence of an HAStoragePlus resource in the application resource group is required.

  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 5, "Configuration of Resource Groups in a Failover Application," on page 326 illustrates the configuration of an application resource group and a replication resource group in a failover application.

FIGURE 5 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.

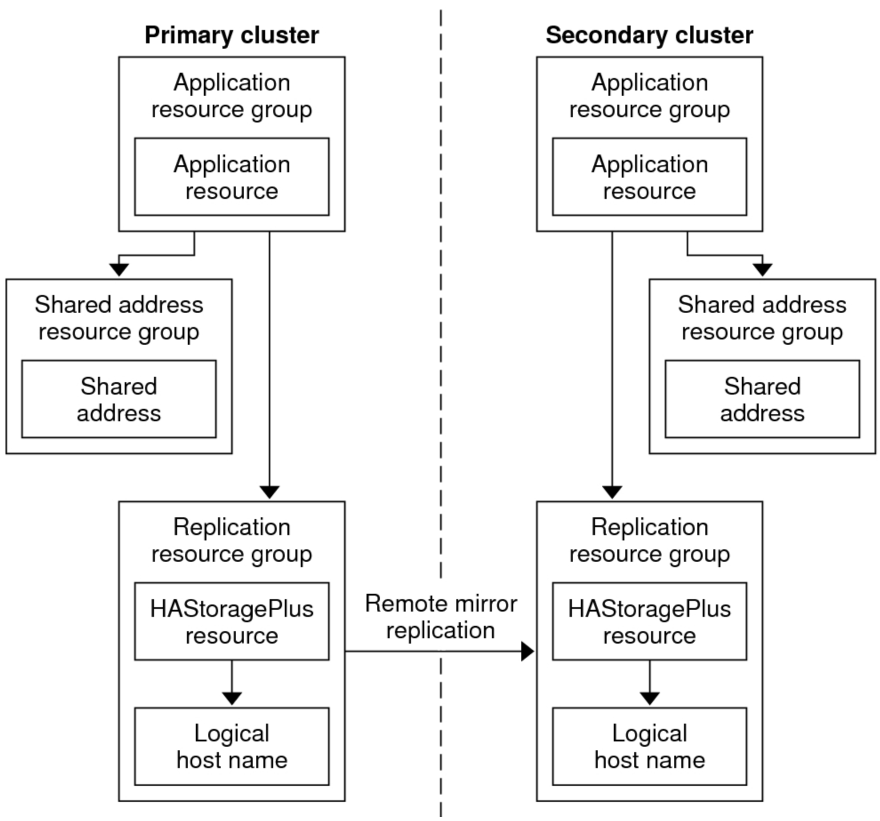
A resource group for a scalable application must have the following characteristics:

- 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 6, “Configuration of Resource Groups in a Scalable Application,” on page 327 illustrates the configuration of resource groups in a scalable application.

FIGURE 6 Configuration of Resource Groups in a Scalable Application

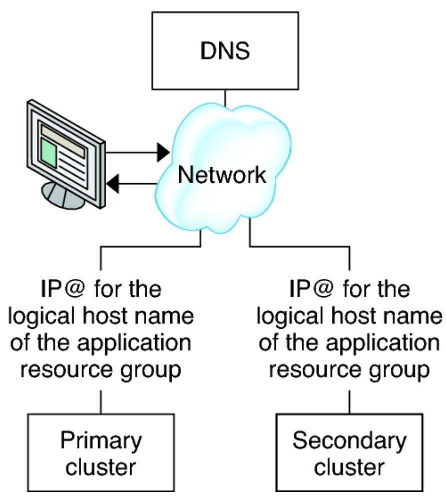


Guidelines for Managing a Takeover

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.

Clients use DNS to map an application's logical hostname to an IP address. After a takeover, where the application is moved to a secondary cluster, the DNS information must be updated to reflect the mapping between the application's logical hostname and the new IP address.

FIGURE 7 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 takeover, see "Example of How to Manage a Takeover" on page 357.

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. You can achieve this by stopping the resource group on the secondary node, so that the replication data stream can drain.
2. Reverse the direction of data replication so that the original primary is now, once again, replicating data to the original secondary.

3. Start the resource group on the primary cluster.
4. Update the DNS so that clients can access the application on the primary cluster.

Task Map: Example of a Data Replication Configuration

Table 23, "Task Map: Example of a Data Replication Configuration," on page 329 lists the tasks in this example of how data replication was configured for an NFS application by using Availability Suite software.

TABLE 23 Task Map: Example of a Data Replication Configuration

Task	Instructions
1. Connect and install the clusters	"Connecting and Installing the Clusters" on page 329
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 331
3. Enable data replication on the primary cluster and on the secondary cluster	"How to Enable Replication on the Primary Cluster" on page 348 "How to Enable Replication on the Secondary Cluster" on page 350
4. Perform data replication	"How to Perform a Remote Mirror Replication" on page 351 "How to Perform a Point-in-Time Snapshot" on page 353
5. Verify the data replication configuration	"How to Verify That Replication Is Configured Correctly" on page 354

Connecting and Installing the Clusters

Figure 8, "Example Cluster Configuration," on page 330 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 8 Example Cluster Configuration

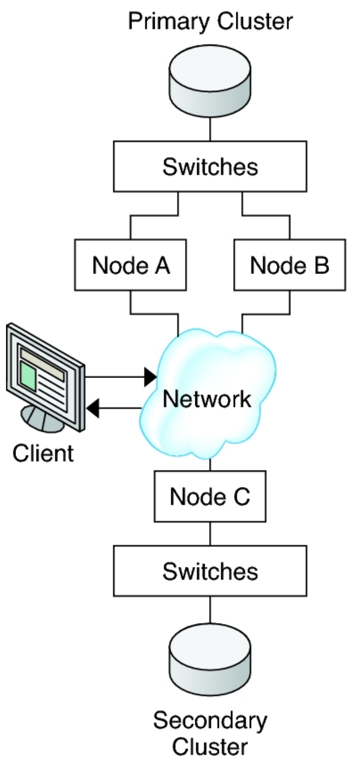


Table 24, "Required Hardware and Software," on page 330 summarizes the hardware and software that the example configuration requires. The Oracle Solaris OS, Oracle Solaris Cluster software, and volume manager software must be installed on the cluster nodes *before* Availability Suite software and software updates are installed.

TABLE 24 Required Hardware and Software

Hardware or Software	Requirement
Node hardware	Availability Suite software is supported on all servers that use Oracle Solaris OS. For information about which hardware to use, see the *Oracle Solaris Cluster Hardware Administration Manual*.
Disk space	Approximately 15 Mbytes.

Hardware or Software	Requirement
Oracle Solaris OS	Oracle Solaris OS releases that are supported by Oracle Solaris Cluster software. All nodes must use the same version of the Oracle Solaris OS. For information about installation, see the *Oracle Solaris Cluster 4.3 Software Installation Guide*
Oracle Solaris Cluster software	At least Oracle Solaris Cluster 4.1 software. For information about installation, see the *Oracle Solaris Cluster 4.3 Software Installation Guide*.
Volume manager software	Solaris Volume Manager software. All nodes must use the same version of volume manager software. For information about installation, see Chapter 4, "Configuring Solaris Volume Manager Software" in *Oracle Solaris Cluster 4.3 Software Installation Guide*.
Availability Suite software	Different clusters can use different versions of Oracle Solaris OS and Oracle Solaris Cluster software, but you must use the same version of Availability Suite software between clusters. For information about how to install the software, see the installation manuals for your release of Availability Suite software.
Availability Suite software updates	For information about the latest software updates, log into My Oracle Support.

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 324 and "Configuring Application Resource Groups" on page 324.

This section contains the following procedures:

- "How to Verify That Replication Is Configured Correctly" on page 354

The following table lists the names of the groups and resources that are created for the example configuration.

TABLE 25 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`.

For information about Solaris Volume Manager software, see the Chapter 4, "Configuring Solaris Volume Manager Software" in *Oracle Solaris Cluster 4.3 Software Installation Guide*.

▼ 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 Availability Suite Software in a Cluster" on page 319
  - "Guidelines for Configuring Host-Based Data Replication Between Clusters" on page 323

- Set up the primary and secondary clusters as described in "Connecting and Installing the Clusters" on page 329.

1. **Access `nodeA` by assuming the role that provides `solaris.cluster.modify` authorization.**

   The node `nodeA` is the first node of the primary cluster. For a reminder of which node is `nodeA`, see Figure 8, "Example Cluster Configuration," on page 330.

2. **Create a metaset to contain the NFS data and associated replication.**

   ```
   nodeA# metaset -s nfsset a -h nodeA nodeB
   ```

3. **Add disks to the metaset.**

   ```
   nodeA# metaset -s nfsset -a /dev/did/dsk/d6 /dev/did/dsk/d7
   ```

4. **Add mediators to the metaset.**

   ```
   nodeA# metaset -s nfsset -a -m nodeA nodeB
   ```

5. **Create the required volumes (or metadevices).**

   Create two components of a mirror:

   ```
   nodeA# metainit -s nfsset d101 1 1 /dev/did/dsk/d6s2
   nodeA# metainit -s nfsset d102 1 1 /dev/did/dsk/d7s2
   ```

   Create the mirror with one of the components:

   ```
   nodeA# metainit -s nfsset d100 -m d101
   ```

   Attach the other component to the mirror and allow it to synchronize:

   ```
   nodeA# metattach -s nfsset d100 d102
   ```

   Create soft partitions from the mirror, following these examples:

   - *d200* - The NFS data (master volume):

     ```
     nodeA# metainit -s nfsset d200 -p d100 50G
     ```

   - *d201* - The point-in-time copy volume for the NFS data:

     ```
     nodeA# metainit -s nfsset d201 -p d100 50G
     ```

   - *d202* - The point-in-time bitmap volume:

     ```
     nodeA# metainit -s nfsset d202 -p d100 10M
     ```

   - *d203* - The remote shadow bitmap volume:

```
nodeA# metainit -s nfsset d203 -p d100 10M
```

- *d204* - The volume for the Oracle Solaris Cluster `SUNW.NFS` configuration information:

  ```
  nodeA# metainit -s nfsset d204 -p d100 100M
  ```

6. **Create file systems for the NFS data and the configuration volume.**

```
nodeA# yes | newfs /dev/md/nfsset/rdsk/d200
nodeA# yes | newfs /dev/md/nfsset/rdsk/d204
```

Next Steps Go to "How to Configure a Device Group on the Secondary Cluster" on page 334.

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 332.

1. **Access `nodeC` by assuming the role that provides `solaris.cluster.modify` authorization.**

2. **Create a metaset to contain the NFS data and associated replication.**

```
nodeC# metaset -s nfsset a -h nodeC
```

3. **Add disks to the metaset.**

In the example below, assume that the disk DID numbers are different.

```
nodeC# metaset -s nfsset -a /dev/did/dsk/d3 /dev/did/dsk/d4
```

Note - Mediators are not required on a single node cluster.

---

4. **Create the required volumes (or metadevices).**

Create two components of a mirror:

```
nodeC# metainit -s nfsset d101 1 1 /dev/did/dsk/d3s2
nodeC# metainit -s nfsset d102 1 1 /dev/did/dsk/d4s2
```

Create the mirror with one of the components:

```
nodeC# metainit -s nfsset d100 -m d101
```

Attach the other component to the mirror and allow it to synchronize:

```
metattach -s nfsset d100 d102
```

Create soft partitions from the mirror, following these examples:

- *d200* - The NFS data master volume:

  ```
  nodeC# metainit -s nfsset d200 -p d100 50G
  ```
- *d201* - The point-in-time copy volume for the NFS data:

  ```
  nodeC# metainit -s nfsset d201 -p d100 50G
  ```
- *d202* - The point-in-time bitmap volume:

  ```
  nodeC# metainit -s nfsset d202 -p d100 10M
  ```
- *d203* - The remote shadow bitmap volume:

  ```
  nodeC# metainit -s nfsset d203 -p d100 10M
  ```
- *d204* - The volume for the Oracle Solaris Cluster `SUNW.NFS` configuration information:

  ```
  nodeC# metainit -s nfsset d204 -p d100 100M
  ```

5. **Create file systems for the NFS data and the configuration volume.**

   ```
   nodeC# yes | newfs /dev/md/nfsset/rdsk/d200
   nodeC# yes | newfs /dev/md/nfsset/rdsk/d204
   ```

Next Steps Go to "How to Configure the File System on the Primary Cluster for the NFS Application" on page 335.

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 334.

1. **On `nodeA` and `nodeB`, assume the role that provides `solaris.cluster.admin` 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 *not* 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/md/nfsset/dsk/d200 /dev/md/nfsset/rdsk/d200 \
   /global/mountpoint ufs 3 no global,logging
   ```

4. **On `nodeA` and `nodeB`, create a mount point for metadevice d204.**

   The following example creates the mount point `/global/etc`.

   ```
   nodeA# mkdir /global/etc
   ```

5. **On `nodeA` and `nodeB`, configure metadevice d204 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/md/nfsset/dsk/d204 /dev/md/nfsset/rdsk/d204 \
   /global/etc ufs 3 yes global,logging
   ```

6. **Mount metadevice d204 on `nodeA`.**

   ```
   nodeA# mount /global/etc
   ```

7. **Create the configuration files and information for the Oracle Solaris Cluster HA for NFS data service.**

   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 337.

▼ 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 335.

1. **On `nodeC`, assume the role that provides `solaris.cluster.admin` 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/md/nfsset/dsk/d200 /dev/md/nfsset/rdsk/d200 \
   /global/mountpoint ufs 3 yes global,logging
   ```

4. **Mount metadevice `d204` on `nodeA`.**

   ```
   nodeC# mount /global/etc
   ```

5. **Create the configuration files and information for the Oracle Solaris Cluster HA for NFS data service.**

   a. **Create a directory called `/global/etc/SUNW.nfs` on `nodeA`.**

   ```
   nodeC# mkdir -p /global/etc/SUNW.nfs
   ```

   b. **Create the file `/global/etc/SUNW.nfs/dfstab.nfs-rs` on `nodeA`.**

   ```
   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 `nodeA`.**

   ```
   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 338.

▼ 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 337.
- Ensure that the `/etc/netmasks` file has IP-address subnet and netmask entries for all logical hostnames. If necessary, edit the `/etc/netmasks` file to add any missing entries.

1. **Access `nodeA` as the role that provides `solaris.cluster.modify`, `solaris.cluster.admin`, and `solaris.cluster.read` 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.HAStoragePlus` resource to the replication resource group.**

```
nodeA# clresource create -g devgrp-stor-rg -t SUNW.HAStoragePlus \
-p GlobalDevicePaths=nfsset \
-p AffinityOn=True \
devgrp-stor
```

`-g`

Specifies the resource group to which resource is added.

`-p GlobalDevicePaths=`

Specifies the device group that Availability Suite software relies on.

`-p AffinityOn=True`

Specifies that the `SUNW.HAStoragePlus` resource must perform an affinity switchover for the global devices and cluster file systems defined by `-p 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.HAStoragePlus(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 -emM -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 339.

▼ 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 338.

- Ensure that the `/etc/netmasks` file has IP-address subnet and netmask entries for all logical hostnames. If necessary, edit the `/etc/netmasks` file to add any missing entries.

1. **Access `nodeC` as the role that provides `solaris.cluster.modify`, `solaris.cluster.admin`, and `solaris.cluster.read` authorization.**

2. **Register `SUNW.HAStoragePlus` as a resource type.**

```
nodeC# clresourcetype register SUNW.HAStoragePlus
```

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.HAStoragePlus` resource to the replication resource group.**

```
nodeC# clresource create \
-t SUNW.HAStoragePlus \
-p GlobalDevicePaths=nfsset \
-p AffinityOn=True \
devgrp-stor
```

`create`

Creates the resource.

`-t`

Specifies the resource type.

`-p GlobalDevicePaths=`

Specifies the device group that Availability Suite software relies on.

-p AffinityOn=True

Specifies that the `SUNW.HAStoragePlus` resource must perform an affinity switchover for the global devices and cluster file systems defined by `-p GlobalDevicePaths=`. Therefore, when the replication resource group fails over or is switched over, the associated device group is switched over.

`devgrp-stor`

The HAStoragePlus resource for the replication resource group.

For more information about these extension properties, see the `SUNW.HAStoragePlus(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 secondary cluster is named `lhost-reprg-sec`.

6. **Enable the resources, manage the resource group, and bring the resource group online.**

```
nodeC# clresourcegroup online -eM -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 342.

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 339.
- Ensure that the `/etc/netmasks` file has IP-address subnet and netmask entries for all logical hostnames. If necessary, edit the `/etc/netmasks` file to add any missing entries.

1. **Access `nodeA` as the role that provides `solaris.cluster.modify`, `solaris.cluster.admin`, and `solaris.cluster.read` authorization.**

2. **Register `SUNW.nfs` as a resource type.**

```
nodeA# clresourcetype register SUNW.nfs
```

3. **If `SUNW.HAStoragePlus` has not been registered as a resource type, register it.**

```
nodeA# clresourcetype register SUNW.HAStoragePlus
```

4. **Create an application resource group for the NFS service.**

```
nodeA# clresourcegroup create \
-p Pathprefix=/global/etc \
-p Auto_start_on_new_cluster=False \
-p RG_affinities=+++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_affinities=+++devgrp-stor-rg`

Specifies the resource group with which the application resource group must be collocated. In this example, the application resource group must be collocated with the replication resource group `devgrp-stor-rg`.

If the replication resource group is switched over to a new primary node, the application resource group is automatically switched over. However, attempts to switch over the application resource group to a new primary node are blocked because that action breaks the collocation requirement.

`nfs-rg`

The name of the application resource group.

5. **Add a `SUNW.HAStoragePlus` resource to the application resource group.**

```
nodeA# clresource create -g nfs-rg \
-t SUNW.HAStoragePlus \
-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.HAStoragePlus`

Specifies that the resource is of the type `SUNW.HAStoragePlus`.

`-p FileSystemMountPoints=/global/`*mountpoint*

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 FileSystemMountPoints`. 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. **Create the `dfstab.`*resource-name* configuration file and place it in the `SUNW.nfs` subdirectory under the `Pathprefix` directory of the containing resource group.**

   a. **Create a directory called `SUNW.nfs` on `nodeA`.**

   ```
   nodeA# mkdir -p /global/etc/SUNW.nfs
   ```

   b. **Create a `dfstab.`*resource-name* file 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
   ```

8. **Bring the application resource group online.**

```
nodeA# clresourcegroup online -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.

9. **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 345.

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 342.
- Ensure that the `/etc/netmasks` file has IP-address subnet and netmask entries for all logical hostnames. If necessary, edit the `/etc/netmasks` file to add any missing entries.

1. **Access `nodeC` as the role that provides `solaris.cluster.modify`, `solaris.cluster.admin`, and `solaris.cluster.read` authorization.**

2. **Register `SUNW.nfs` as a resource type.**

```
nodeC# clresourcetype register SUNW.nfs
```

3. **If `SUNW.HAStoragePlus` has not been registered as a resource type, register it.**

```
nodeC# clresourcetype register SUNW.HAStoragePlus
```

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_affinities=+++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_affinities=+++devgrp-stor-rg`

Specifies the resource group where the application resource group must be collocated. In this example, the application resource group must be collocated with the replication resource group `devgrp-stor-rg`.

If the replication resource group is switched over to a new primary node, the application resource group is automatically switched over. However, attempts to switch over the application resource group to a new primary node are blocked because that breaks the collocation requirement.

`nfs-rg`

The name of the application resource group.

5. **Add a `SUNW.HAStoragePlus` resource to the application resource group.**

```
nodeC# clresource create -g nfs-rg \
-t SUNW.HAStoragePlus \
-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.HAStoragePlus`

Specifies that the resource is of the type `SUNW.HAStoragePlus`.

`-p`

Specifies a property of the resource.

`FileSystemMountPoints=/global/`*mountpoint*

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 `-p FileSystemMountPoints=`. 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.

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. **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 347.

Example of How to Enable Data Replication

This section describes how data replication is enabled for the example configuration. This section uses the Availability Suite software commands `sndradm` and `iiadm`. For more information about these commands, see the Availability Suite documentation.

This section contains the following procedures:

- "How to Enable Replication on the Primary Cluster" on page 348
- "How to Enable Replication on the Secondary Cluster" on page 350

▼ How to Enable Replication on the Primary Cluster

1. **Access `nodeA` as the role that provides `solaris.cluster.read` 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 primary cluster to the secondary cluster. This step enables replication from the master volume (`d200`) on the primary cluster to the master volume (`d200`) on the secondary cluster. In addition, this step enables replication to the remote mirror bitmap on `d203`.

   - If the primary cluster and secondary cluster are unsynchronized, run this command for Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -e lhost-reprg-prim \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 lhost-reprg-sec \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 ip sync
```

   - If the primary cluster and secondary cluster are synchronized, run this command for Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -E lhost-reprg-prim \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 lhost-reprg-sec \
/dev/md/nfsset/rdsk/d200 \
```

```
/dev/md/nfsset/rdsk/d203 ip sync
```

5. **Enable autosynchronization.**

   Run this command for Availability Suite software:

   ```
   nodeA# /usr/sbin/sndradm -n -a on lhost-reprg-prim \
   /dev/md/nfsset/rdsk/d200 \
   /dev/md/nfsset/rdsk/d203 lhost-reprg-sec \
   /dev/md/nfsset/rdsk/d200 \
   /dev/md/nfsset/rdsk/d203 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 Availability Suite software:

   ```
   nodeA# /usr/sbin/sndradm -P
   ```

   The output should resemble the following:

   ```
   /dev/md/nfsset/rdsk/d200 ->
   lhost-reprg-sec:/dev/md/nfsset/rdsk/d200
   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 Availability Suite software:

   ```
   nodeA# /usr/sbin/iiadm -e ind \
   /dev/md/nfsset/rdsk/d200 \
   /dev/md/nfsset/rdsk/d201 \
   /dev/md/nfsset/rdsk/d202
   nodeA# /usr/sbin/iiadm -w \
   /dev/md/nfsset/rdsk/d201
   ```

   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 `d200`, the shadow volume is `d201`, and the point-in-time bitmap volume is `d203`.

8. **Attach the point-in-time snapshot to the remote mirror set.**

   Use the following command for Availability Suite software:

```
nodeA# /usr/sbin/sndradm -I a \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d201 \
/dev/md/nfsset/rdsk/d202
```

This step associates the point-in-time snapshot with the remote mirror volume set. 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 350.

How to Enable Replication on the Secondary Cluster

Before You Begin Complete the procedure "How to Enable Replication on the Primary Cluster" on page 348.

1. **Access `nodeC` as the `root` role.**

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 Availability Suite software:

```
nodeC# /usr/sbin/sndradm -n -e lhost-reprg-prim \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 lhost-reprg-sec \
/dev/md/nfsset/rdsk/d200  \
/dev/md/nfsset/rdsk/d203 ip sync
```

   The primary cluster detects the presence of the secondary cluster and starts synchronization. Refer to the system log file `/var/adm` for Availability Suite for information about the status of the clusters.

4. **Enable independent point-in-time snapshot.**

   Use the following command for Availability Suite software:

```
nodeC# /usr/sbin/iiadm -e ind \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d201 \
/dev/md/nfsset/rdsk/d202
```

```
nodeC# /usr/sbin/iiadm -w \
/dev/md/nfsset/rdsk/d201
```

5. **Attach the point-in-time snapshot to the remote mirror set.**

   Use the following command for Availability Suite software:

```
nodeC# /usr/sbin/sndradm -I a \
/dev/md/nfsset/rdsk/d200  \
/dev/md/nfsset/rdsk/d201 \
/dev/md/nfsset/rdsk/d202
```

Next Steps Go to "Example of How to Perform Data Replication" on page 351.

Example of How to Perform Data Replication

This section describes how data replication is performed for the example configuration. This section uses the Availability Suite software commands `sndradm` and `iiadm`. For more information about these commands, see the Availability Suite documentation.

This section contains the following procedures:

▼ 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 `d200` and the remote mirror bitmap volume is `d203`.

1. **Access `nodeA` as the `root` role.**

2. **Verify that the cluster is in logging mode.**

   Run the following command for Availability Suite software:

```
nodeA# /usr/sbin/sndradm -P
```

   The output should resemble the following:

```
/dev/md/nfsset/rdsk/d200  ->
```

```
lhost-reprg-sec:/dev/md/nfsset/rdsk/d200
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 Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -m lhost-reprg-prim \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 lhost-reprg-sec \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 ip sync
```

6. **Wait until the replication is complete and the volumes are synchronized.**

   Run the following command for Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -w lhost-reprg-prim \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 lhost-reprg-sec \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 ip sync
```

7. **Confirm that the cluster is in replicating mode.**

   Run the following command for Availability Suite software:

```
nodeA# /usr/sbin/sndradm -P
```

   The output should resemble the following:

```
/dev/md/nfsset/rdsk/d200 ->
lhost-reprg-sec:/dev/md/nfsset/rdsk/d200
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 Availability Suite software.

Next Steps Go to "How to Perform a Point-in-Time Snapshot" on page 353.

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 `d200`, the bitmap volume is `d203`, and the shadow volume is `d201`.

Before You Begin Complete the procedure "How to Perform a Remote Mirror Replication" on page 351.

1. **Access `nodeA` as the role that provides `solaris.cluster.modify` and `solaris.cluster.admin` authorization.**

2. **Disable the resource that is running on `nodeA`.**

```
nodeA# clresource disable nfs-rs
```

3. **Change the primary cluster to logging mode.**

   Run the following command for Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -l lhost-reprg-prim \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 lhost-reprg-sec \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 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 Availability Suite software:

```
nodeA# /usr/sbin/iiadm -u s /dev/md/nfsset/rdsk/d201
nodeA# /usr/sbin/iiadm -w /dev/md/nfsset/rdsk/d201
```

5. **Synchronize the shadow volume of the secondary cluster to the master volume of the secondary cluster.**

   Run the following command for Availability Suite software:

```
nodeC# /usr/sbin/iiadm -u s /dev/md/nfsset/rdsk/d201
nodeC# /usr/sbin/iiadm -w /dev/md/nfsset/rdsk/d201
```

6. **Restart the application on `nodeA`.**

```
nodeA# clresource enable nfs-rs
```

7. **Resynchronize the secondary volume with the primary volume.**

   Run the following command for Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -u lhost-reprg-prim \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 lhost-reprg-sec \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 ip sync
```

Next Steps Go to "How to Verify That Replication Is Configured Correctly" on page 354.

▼ How to Verify That Replication Is Configured Correctly

Before You Begin Complete the procedure "How to Perform a Point-in-Time Snapshot" on page 353.

1. **Access `nodeA` and `nodeC` as the role that provides `solaris.cluster.admin` authorization.**

2. **Verify that the primary cluster is in replicating mode, with autosynchronization on.**

   Use the following command for Availability Suite software:

```
nodeA# /usr/sbin/sndradm -P
```

   The output should resemble the following:

```
/dev/md/nfsset/rdsk/d200 ->
lhost-reprg-sec:/dev/md/nfsset/rdsk/d200
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 Availability Suite software.

3. **If the primary cluster is not in replicating mode, put it into replicating mode.**

   Use the following command for Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -u lhost-reprg-prim \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 lhost-reprg-sec \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 ip sync
```

4. **Create a directory on a client machine.**

   a. **Log in to a client machine as the `root` role.**

      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 primary volume on the application directory and display the mounted directory.**

   a. **Mount the primary volume on the application directory.**

      ```
      client-machine# mount -o rw lhost-nfsrg-prim:/global/mountpoint /dir
      ```

   b. **Display the mounted directory.**

      ```
      client-machine# ls /dir
      ```

6. **Unmount the primary volume from the application directory.**

   a. **Unmount the primary volume from the application directory.**

      ```
      client-machine# umount /dir
      ```

   b. **Take the application resource group offline on the primary cluster.**

      ```
      nodeA# clresource disable -g nfs-rg +
      nodeA# clresourcegroup offline nfs-rg
      ```

   c. **Change the primary cluster to logging mode.**

      Run the following command for Availability Suite software:

      ```
      nodeA# /usr/sbin/sndradm -n -l lhost-reprg-prim \
      /dev/md/nfsset/rdsk/d200  \
      ```

```
/dev/md/nfsset/rdsk/d203 lhost-reprg-sec \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 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. **Confirm that the file system is fit to be mounted on the secondary cluster.**

```
nodeC# fsck -y /dev/md/nfsset/rdsk/d200
```

f. **Bring the application into a managed state, and bring it online on the secondary cluster.**

```
nodeC# clresourcegroup online -eM nodeC nfs-rg
```

g. **Access the client machine as the `root` role.**

You see a prompt that resembles the following:

```
client-machine#
```

h. **Mount the application directory that was created in Step 4 to the application directory on the secondary volume.**

```
client-machine# mount -o rw lhost-nfsrg-sec:/global/mountpoint /dir
```

i. **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 volume to the mounted application directory.**

   a. **Take the application resource group offline on the secondary volume.**

```
nodeC# clresource disable -g nfs-rg +
nodeC# clresourcegroup offline nfs-rg
```

b. Ensure that the global volume is unmounted from the secondary volume.

```
nodeC# umount /global/mountpoint
```

c. Bring the application resource group into a managed state, and bring it online on the primary cluster.

```
nodeA# clresourcegroup online -eM nodeA nfs-rg
```

d. Change the primary volume to replicating mode.

Run the following command for Availability Suite software:

```
nodeA# /usr/sbin/sndradm -n -u lhost-reprg-prim \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 lhost-reprg-sec \
/dev/md/nfsset/rdsk/d200 \
/dev/md/nfsset/rdsk/d203 ip sync
```

When the primary volume is written to, the secondary volume is updated by Availability Suite software.

See Also "Example of How to Manage a Takeover" on page 357

Example of How to Manage a Takeover

This section describes how to update the DNS entries. For additional information, see "Guidelines for Managing a Takeover" on page 328.

▼ How to Update the DNS Entry

For an illustration of how DNS maps a client to a cluster, see Figure 7, "DNS Mapping of a Client to a Cluster," on page 328.

1. **Start the `nsupdate` command.**

   For more 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

C

E

F

M

N

O

S

T