Oracle® Solaris Cluster Software Installation Guide

Part No: E37727
September 2013, E37727–04

ORACLE

130926@25097

Contents

Preface

The *Oracle Solaris Cluster Software Installation Guide* contains guidelines and procedures for installing the Oracle Solaris Cluster software on both SPARC based systems and x86 based systems.

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

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

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

Using UNIX Commands

This document contains information about commands that are used to install, configure, or upgrade an Oracle Solaris Cluster configuration. This document might not contain complete information about basic UNIX commands and procedures such as shutting down the system, booting the system, and configuring devices.

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

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

Typographic Conventions

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

TABLE P–1 Typographic Conventions

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

Shell Prompts in Command Examples

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

TABLE P–2 Shell Prompts

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

Related Documentation

Information about related Oracle Solaris Cluster software topics is available in the documentation that is listed in the following table. All Oracle Solaris Cluster documentation is available at http://www.oracle.com/technetwork/indexes/documentation/index.html.

Topic	Documentation
Concepts	*Oracle Solaris Cluster Concepts Guide*
Hardware installation and administration	*Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual* and individual hardware administration guides
Software installation	*Oracle Solaris Cluster Software Installation Guide*
Data service installation and administration	*Oracle Solaris Cluster Data Services Planning and Administration Guide* and individual data service guides
Data service development	*Oracle Solaris Cluster Data Services Developer's Guide*
System administration	*Oracle Solaris Cluster System Administration Guide*
	Oracle Solaris Cluster Quick Reference
Software upgrade	*Oracle Solaris Cluster Upgrade Guide*
Error messages	*Oracle Solaris Cluster Error Messages Guide*
Command and function references	*Oracle Solaris Cluster Reference Manual*
	Oracle Solaris Cluster Data Services Reference Manual

Access to Oracle Support

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

Getting Help

If you have problems installing or using Oracle Solaris Cluster software, contact your service provider and supply the following information.

- Your name and email address (if available)
- Your company name, address, and phone number
- The model number and serial number of your systems
- The release number of the Oracle Solaris OS (for example, Oracle Solaris 10)

- The release number of Oracle Solaris Cluster (for example, Oracle Solaris Cluster 3.3 3/13)

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

Command	Function
`prtconf -v`	Displays the size of the system memory and reports information about peripheral devices
`psrinfo -v`	Displays information about processors
`showrev -p`	Reports which patches are installed
SPARC: `prtdiag -v`	Displays system diagnostic information
`/usr/cluster/bin/clnode show-rev -v`	Displays Oracle Solaris Cluster release and package version information

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

◆ ◆ ◆ CHAPTER 1

Planning the Oracle Solaris Cluster Configuration

This chapter provides planning information and guidelines specific to an Oracle Solaris Cluster 3.3 3/13 configuration.

The following overview information is in this chapter:

Finding Oracle Solaris Cluster Installation Tasks

The following table shows where to find instructions for various installation tasks for Oracle Solaris Cluster software installation and the order in which you should perform the tasks.

TABLE 1–1 Oracle Solaris Cluster Software Installation Task Information

Task	Instructions
Set up cluster hardware.	*Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual* Documentation that shipped with your server and storage devices
Plan global-cluster software installation.	Chapter 1, "Planning the Oracle Solaris Cluster Configuration"
Install software packages. Optionally, install and configure Sun QFS software.	"Installing the Software" on page 49 Using Sun QFS and Sun Storage Archive Manager With Oracle Solaris Cluster
Establish a new global cluster or a new global-cluster node.	"Establishing a New Global Cluster or New Global-Cluster Node" on page 72

TABLE 1–1 Oracle Solaris Cluster Software Installation Task Information *(Continued)*

Task	Instructions
Configure Solaris Volume Manager software.	"Configuring Solaris Volume Manager Software" on page 149 *Solaris Volume Manager Administration Guide*
Configure cluster file systems, if used.	"How to Create Cluster File Systems" on page 173
(Optional) Create non-global zones.	"Configuring a Non-Global Zone on a Global-Cluster Node" on page 191
(Optional) Create zone clusters.	"Configuring a Zone Cluster" on page 197
Plan, install, and configure resource groups and data services. Create highly available local file systems, if used.	*Oracle Solaris Cluster Data Services Planning and Administration Guide*
Develop custom data services.	*Oracle Solaris Cluster Data Services Developer's Guide*

Planning the Oracle Solaris OS

This section provides the following guidelines for planning Oracle Solaris software installation in a cluster configuration.

- "Guidelines for Selecting Your Oracle Solaris Installation Method" on page 12
- "Oracle Solaris OS Feature Restrictions" on page 13
- "Oracle Solaris Software Group Considerations" on page 14
- "System Disk Partitions" on page 14
- "Guidelines for Non-Global Zones in a Global Cluster" on page 18
- "SPARC: Guidelines for Oracle VM Server for SPARC in a Cluster" on page 19

For more information about Oracle Solaris software, see your Oracle Solaris installation documentation.

Guidelines for Selecting Your Oracle Solaris Installation Method

You can install Oracle Solaris software from a local DVD-ROM or from a network installation server by using the Oracle Solaris JumpStart installation method. In addition, Oracle Solaris Cluster software provides a custom method for installing both the Oracle Solaris OS and Oracle Solaris Cluster software by using the JumpStart installation method. If you are installing several cluster nodes, consider a network installation.

See "How to Install Oracle Solaris and Oracle Solaris Cluster Software (JumpStart)" on page 90 for details about the `scinstall` JumpStart installation method. See your Oracle Solaris installation documentation for details about standard Oracle Solaris installation methods.

Oracle Solaris OS Feature Restrictions

Consider the following points when you plan the use of the Oracle Solaris OS in an Oracle Solaris Cluster configuration:

- **Oracle Solaris Zones** – Install Oracle Solaris Cluster framework software *only* in the global zone.

  To determine whether you can install an Oracle Solaris Cluster data service directly in a non-global zone, see the documentation for that data service.

  If you configure non-global zones on a global-cluster node, the loopback file system (LOFS) must be enabled. See the information for LOFS for additional considerations.
- **Loopback file system (LOFS)** – During cluster creation, LOFS capability is enabled by default. If the cluster meets both of the following conditions, you must disable LOFS to avoid switchover problems or other failures:
  - Oracle Solaris Cluster HA for NFS (HA for NFS) is configured on a highly available local file system.
  - The `automountd` daemon is running.

  If the cluster meets only one of these conditions, you can safely enable LOFS.

  If you require both LOFS and the `automountd` daemon to be enabled, exclude from the automounter map all files that are part of the highly available local file system that is exported by HA for NFS.
- **Power-saving shutdown** – Automatic power-saving shutdown is not supported in Oracle Solaris Cluster configurations and should not be enabled. See the `pmconfig(1M)` and `power.conf(4)` man pages for more information.
- **IP Filter feature** – Oracle Solaris Cluster software does not support the Oracle Solaris IP Filter feature for scalable services, but does support Oracle Solaris IP Filter for failover services. Observe the following guidelines and restrictions when you configure Oracle Solaris IP Filter in a cluster:
  - NAT routing is not supported.
  - The use of NAT for translation of local addresses is supported. NAT translation rewrites packets on-the-wire and is therefore transparent to the cluster software.
  - Stateful filtering rules are not supported; only stateless filtering is supported. Oracle Solaris Cluster relies on IP network multipathing (IPMP) for public network monitoring, which does not work with stateful filtering rules.
- **fssnap** – Oracle Solaris Cluster software does not support the `fssnap` command, which is a feature of UFS. However, you can use the `fssnap` command on local systems that are not controlled by Oracle Solaris Cluster software. The following restrictions apply to `fssnap` support:
  - The `fssnap` command is supported on local files systems that are not managed by Oracle Solaris Cluster software.

- The `fssnap` command is not supported on cluster file systems.
- The `fssnap` command is not supported on local file systems under the control of HAStoragePlus.

Oracle Solaris Software Group Considerations

Oracle Solaris Cluster 3.3 3/13 software requires at least the End User Oracle Solaris Software Group (`SUNWCuser`). However, other components of your cluster configuration might have their own Oracle Solaris software requirements as well. Consider the following information when you decide which Oracle Solaris software group you are installing.

- **Servers** – Check your server documentation for any Oracle Solaris software requirements.
- **Additional Oracle Solaris packages** – You might need to install other Oracle Solaris software packages that are not part of the End User Oracle Solaris Software Group. The Apache HTTP server packages and Trusted Extensions software are two examples that require packages that are in a higher software group than End User. Third-party software might also require additional Oracle Solaris software packages. See your third-party documentation for any Oracle Solaris software requirements.

  Tip – To avoid the need to manually install Oracle Solaris software packages, install the Entire Oracle Solaris Software Group Plus OEM Support.

- **Oracle Solaris package minimization** – See Article 1544605.1 "Solaris Cluster and Solaris OS Minimization Support Required Packages Group" at `(http://support.oracle.com)` for information.

System Disk Partitions

When you install the Oracle Solaris OS, ensure that you create the required Oracle Solaris Cluster partitions and that all partitions meet minimum space requirements.

- **swap** – The combined amount of `swap` space that is allocated for Oracle Solaris and Oracle Solaris Cluster software must be no less than 750 Mbytes. For best results, add at least 512 Mbytes for Oracle Solaris Cluster software to the amount that is required by the Oracle Solaris OS. In addition, allocate any additional `swap` amount that is required by applications that are to run on the Oracle Solaris host.

  Note – If you create an additional `swap` file, do not create the `swap` file on a global device. Use only a local disk as a `swap` device for the host.

- (Optional) `/globaldevices` – By default, a `lofi` device is used for the global devices namespace. However, you can alternatively create a file system at least 512 Mbytes large that is to be used by the `scinstall` utility for global devices. You must name this file system `/globaldevices`.

  Functionality and performance are equivalent for both choices. However, a `lofi` device provides greater ease of use and more flexibility in situations where a disk partition is not available for use.
- **Volume manager** – Create a 20-Mbyte partition on slice 7 for volume manager use.

To meet these requirements, you must customize the partitioning if you are performing interactive installation of the Oracle Solaris OS.

See the following guidelines for additional partition planning information:

- "Guidelines for the Root (/) File System" on page 15
- "Guidelines for the /globaldevices File System" on page 16
- "Volume Manager Requirements" on page 17

Guidelines for the Root (/) File System

As with any other system running the Oracle Solaris OS, you can configure the root (/), `/var`, `/usr`, and `/opt` directories as separate file systems. Or, you can include all the directories in the root (/) file system.

The following describes the software contents of the root (/), `/var`, `/usr`, and `/opt` directories in an Oracle Solaris Cluster configuration. Consider this information when you plan your partitioning scheme.

- root (/) – The Oracle Solaris Cluster software itself occupies less than 40 Mbytes of space in the root (/) file system. Solaris Volume Manager software requires less than 5 Mbytes. To configure ample additional space and inode capacity, add at least 100 Mbytes to the amount of space you would normally allocate for your root (/) file system. This space is used for the creation of both block special devices and character special devices used by the volume management software. You especially need to allocate this extra space if a large number of shared disks are in the cluster.

  On the Oracle Solaris 10 OS, the `lofi` device for the global-devices namespace requires 100 MBytes of free space.
- `/var` – The Oracle Solaris Cluster software occupies a negligible amount of space in the `/var` file system at installation time. However, you need to set aside ample space for log files. Also, more messages might be logged on a clustered node than would be found on a typical stand-alone server. Therefore, allow at least 100 Mbytes for the `/var` file system.
- `/usr` – Oracle Solaris Cluster software occupies less than 25 Mbytes of space in the `/usr` file system. Solaris Volume Manager software requires less than 15 Mbytes.

- `/opt` – Oracle Solaris Cluster framework software uses less than 2 Mbytes in the `/opt` file system. However, each Oracle Solaris Cluster data service might use between 1 Mbyte and 5 Mbytes. Solaris Volume Manager software does not use any space in the `/opt` file system.

  In addition, most database and applications software is installed in the `/opt` file system.

Guidelines for the `/globaldevices` File System

Oracle Solaris Cluster software offers two choices of locations to host the global-devices namespace:

- A `lofi` device, which is the default
- A dedicated file system on one of the local disks

When you use a `lofi` device for the global-devices namespace, observe the following requirements:

- **Dedicated use** – The `lofi` device that hosts the global-devices namespace cannot be used for any other purpose. If you need a `lofi` device for some other use, create a new `lofi` device for that purpose.
- **Mount requirement** – The `lofi` device must not be unmounted.
- **Namespace identification** – After the cluster is configured, you can use the `lofiadm` command to identify the `lofi` device that corresponds to the global-devices namespace, `/.globaldevices`.

If you instead configure a dedicated `/globaldevices` for the global-devices namespace, observe the following guidelines and requirements:

- **Location** - The `/globaldevices` file system is usually located on your root disk. However, if you use different storage on which to locate the global-devices file system, such as a Logical Volume Manager volume, it must not be part of a Solaris Volume Manager shared disk set. This file system is later mounted as a UFS cluster file system. Name this file system `/globaldevices`, which is the default name that is recognized by the `scinstall(1M)` command.
- **Required file-system type** - No file-system type other than UFS is valid for the global-devices file system. Do not attempt to change the file-system type after the global-devices file system is created.

  However, a UFS global-devices file system can coexist on a node with other root file systems that use ZFS.
- **Configured namespace name**- The `scinstall` command later renames the file system `/global/.devices/node@`*nodeid*, where *nodeid* represents the number that is assigned to an Oracle Solaris host when it becomes a global-cluster member. The original `/globaldevices` mount point is removed.

- **Space requirements** - The `/globaldevices` file system must have ample space and ample inode capacity for creating both block special devices and character special devices. This guideline is especially important if a large number of disks are in the cluster. Create a file system size of at least 512 Mbytes and a density of 512, as follows:

  `# newfs -i 512` *globaldevices-partition*

  This number of inodes should suffice for most cluster configurations.

Volume Manager Requirements

For Solaris Volume Manager software, you must set aside a slice on the root disk for use in creating the state database replica. Specifically, set aside a slice for this purpose on each local disk. But, if you have only one local disk on an Oracle Solaris host, you might need to create three state database replicas in the same slice for Solaris Volume Manager software to function properly. See your Solaris Volume Manager documentation for more information.

Example – Sample File-System Allocations

Table 1–2 shows a partitioning scheme for an Oracle Solaris host that has less than 750 Mbytes of physical memory. This scheme is to be installed with the End User Oracle Solaris Software Group, Oracle Solaris Cluster software, and the Oracle Solaris Cluster HA for NFS data service. The last slice on the disk, slice 7, is allocated with a small amount of space for volume-manager use.

If you use a `lofi` device for the global-devices namespace, slice 3 can be used for another purpose or left labeled as unused.

If you use Solaris Volume Manager software, you use slice 7 for the state database replica. This layout provides the necessary two free slices, 4 and 7, as well as provides for unused space at the end of the disk.

TABLE 1–2 Example File-System Allocation

Slice	Contents	Size Allocation	Description
0	`/`	6.75GB	Remaining free space on the disk after allocating space to slices 1 through 7. Used for the Oracle Solaris OS, Oracle Solaris Cluster software, data-services software, volume-manager software, root file systems, and database and application software.
1	`swap`	1GB	512 Mbytes for the Oracle Solaris OS. 512 Mbytes for Oracle Solaris Cluster software.
2	overlap	8.43GB	The entire disk.
3	`/globaldevices`	512MB	The Oracle Solaris Cluster software later assigns this slice a different mount point and mounts the slice as a cluster file system. If you choose to use a lofi device instead of a dedicated partition, leave slice 3 as Unused.

TABLE 1–2 Example File-System Allocation *(Continued)*

Slice	Contents	Size Allocation	Description
4	unused	-	-
5	unused	-	-
6	unused	-	-
7	volume manager	20MB	Used by Solaris Volume Manager software for the state database replica.

Guidelines for Non-Global Zones in a Global Cluster

For information about the purpose and function of Oracle Solaris zones in a cluster, see "Support for Oracle Solaris Zones" in *Oracle Solaris Cluster Concepts Guide*.

For guidelines about configuring a cluster of non-global zones, see "Zone Clusters" on page 36.

Consider the following points when you create an Oracle Solaris 10 non-global zone, simply referred to as a zone, on a global-cluster node.

- **Unique zone name** – The zone name must be unique on the Oracle Solaris host.
- **Reusing a zone name on multiple nodes** – To simplify cluster administration, you can use the same name for a zone on each node where resource groups are to be brought online in that zone.
- **Private IP addresses** – Do not attempt to use more private IP addresses than are available in the cluster.
- **Mounts** – Do not include global mounts in zone definitions. Include only loopback mounts.
- **Failover services** – In multiple-host clusters, while Oracle Solaris Cluster software permits you to specify different zones on the same Oracle Solaris host in a failover resource group's node list, doing so is useful only during testing. If a single host contains all zones in the node list, the node becomes a single point of failure for the resource group. For highest availability, zones in a failover resource group's node list should be on different hosts.

  In single-host clusters, no functional risk is incurred if you specify multiple zones in a failover resource group's node list.
- **Scalable services** – Do not create non-global zones for use in the same scalable service on the same Oracle Solaris host. Each instance of the scalable service must run on a different host.
- **Cluster file systems** - For cluster file systems that use UFS, do not directly add a cluster file system to a non-global zone by using the `zonecfs` command. Instead, configure an HAStoragePlus resource, which manages the mounting of the cluster file system in the global zone and performs a loopback mount of the cluster file system in the non-global zone.

- **LOFS** – Oracle Solaris Zones requires that the loopback file system (LOFS) be enabled. However, the Oracle Solaris Cluster HA for NFS data service requires that LOFS be disabled, to avoid switchover problems or other failures. If you configure both non-global zones and Oracle Solaris Cluster HA for NFS in your cluster, do one of the following to prevent possible problems in the data service:
  - Disable the `automountd` daemon.
  - Exclude from the automounter map all files that are part of the highly available local file system that is exported by Oracle Solaris Cluster HA for NFS.
- **Exclusive-IP zones** – The following guidelines apply specifically to exclusive-IP non-global zones:
  - **Logical-hostname resource groups** – In a resource group that contains a `LogicalHostname` resource, if the node list contains any non-global zone with the `ip-type` property set to `exclusive`, all zones in that node list must have this property set to `exclusive`. Note that a global zone always has the `ip-type` property set to `shared`, and therefore cannot coexist in a node list that contains zones of `ip-type=exclusive`. This restriction applies only to versions of the Oracle Solaris OS that use the Oracle Solaris zones `ip-type` property.
  - **IPMP groups** – For all public-network adapters that are used for data-service traffic in the non-global zone, you must manually configure IPMP groups in all `/etc/hostname.`*adapter* files on the zone. This information is not inherited from the global zone. For guidelines and instructions to configure IPMP groups, follow the procedures in Part V, "IPMP," in *Oracle Solaris Administration: IP Services*.
  - **Private-hostname dependency** - Exclusive-IP zones cannot depend on the private hostnames and private addresses of the cluster.
  - **Shared-address resources** – Shared-address resources cannot use exclusive-IP zones.

SPARC: Guidelines for Oracle VM Server for SPARC in a Cluster

Consider the following points when you create an Oracle VM Server for SPARC I/O domain or guest domain on a physically clustered machine that is SPARC hypervisor capable:

- **SCSI LUN requirement** – The virtual shared storage device, or virtual disk back end, of an Oracle VM Server for SPARC guest domain must be a full SCSI LUN in the I/O domain. You cannot use an arbitrary virtual device.
- **Fencing** – Do not export a storage LUN to more than one guest domain on the same physical machine, unless you also disable fencing for that device. Otherwise, if two different guest domains on the same machine both are visible to a device, the device will be fenced whenever one of the guest domains dies. The fencing of the device will panic any other guest domain that subsequently tries to access the device.

- **Network isolation** – Guest domains that are located on the same physical machine but are configured in different clusters must be network isolated from each other. Use one of the following methods:
  - Configure the clusters to use different network interfaces in the I/O domain for the private network.
  - Use different network addresses for each of the clusters.
- **Networking in guest domains** – Network packets to and from guest domains must traverse service domains to reach the network drivers through virtual switches. Virtual switches use kernel threads that run at system priority. The virtual-switch threads must be able to acquire needed CPU resources to perform critical cluster operations, including heartbeats, membership, checkpoints, and so forth. Configuring virtual switches with the `mode=sc` setting enables expedited handling of cluster heartbeat packets. However, the reliability of other critical cluster operations can be enhanced by adding more CPU resources to the service domain under the following workloads:
  - High-interrupt load, for example, due to network or disk I/O. Under extreme load, virtual switches can preclude system threads from running for a long time, including virtual-switch threads.
  - Real-time threads that are overly aggressive in retaining CPU resources. Real-time threads run at a higher priority than virtual-switch threads, which can restrict CPU resources for virtual-switch threads for an extended time.
- **Non-shared storage** - For non-shared storage, such as for Oracle VM Server for SPARC guest-domain OS images, you can use any type of virtual device. You can back such virtual devices by any implement in the I/O domain, such as files or volumes. However, do not copy files or clone volumes in the I/O domain for the purpose of mapping them into different guest domains of the same cluster. Such copying or cloning would lead to problems because the resulting virtual devices would have the same device identity in different guest domains. Always create a new file or device in the I/O domain, which would be assigned a unique device identity, then map the new file or device into a different guest domain.
- **Exporting storage from I/O domains** – If you configure a cluster that is composed of Oracle VM Server for SPARC I/O domains, do not export its storage devices to other guest domains that also run Oracle Solaris Cluster software.
- **Oracle Solaris I/O multipathing** – Do not run Oracle Solaris I/O multipathing software (MPxIO) from guest domains. Instead, run Oracle Solaris I/O multipathing software in the I/O domain and export it to the guest domains.
- **Virtual disk multipathing** - Do not configure the virtual disk multipathing feature of Oracle VM Server for SPARC on a logical domain that is configured as a cluster node.
- **Private-interconnect IP address range** – The private network is shared by all guest domains that are created on the same physical machine and it is visible to all these domains. Before you specify a private-network IP address range to the `scinstall` utility for use by a guest-domain cluster, ensure that the address range is not already in use by another guest domain on the same physical machine.

For more information about Oracle VM Server for SPARC, see the *Logical Domains (LDoms) 1.0.3 Administration Guide*.

Planning the Oracle Solaris Cluster Environment

This section provides guidelines for planning and preparing the following components for Oracle Solaris Cluster software installation and configuration:

For detailed information about Oracle Solaris Cluster components, see the *Oracle Solaris Cluster Concepts Guide*.

Licensing

Ensure that you have available all necessary license certificates before you begin software installation. Oracle Solaris Cluster software does not require a license certificate, but each node installed with Oracle Solaris Cluster software must be covered under your Oracle Solaris Cluster software license agreement.

For licensing requirements for volume-manager software and applications software, see the installation documentation for those products.

Software Patches

After installing each software product, you must also install any required patches. For proper cluster operation, ensure that all cluster nodes maintain the same patch level.

- For information about current required patches, see "Patches and Required Firmware Levels" in *Oracle Solaris Cluster 3.3 3/13 Release Notes* or consult your Oracle service provider.

- For general guidelines and procedures for applying patches, see Chapter 11, "Patching Oracle Solaris Cluster Software and Firmware," in *Oracle Solaris Cluster System Administration Guide*.

Public-Network IP Addresses

For information about the use of public networks by the cluster, see "Public Network Adapters and IP Network Multipathing" in *Oracle Solaris Cluster Concepts Guide*.

You must set up a number of public-network IP addresses for various Oracle Solaris Cluster components, depending on your cluster configuration. Each Oracle Solaris host in the cluster configuration must have at least one public-network connection to the same set of public subnets.

The following table lists the components that need public-network IP addresses assigned. Add these IP addresses to the following locations:

- Any naming services that are used
- The local `/etc/inet/hosts` file on each global-cluster node, after you install Oracle Solaris software
- The local `/etc/inet/hosts` file on any exclusive-IP non-global zone

TABLE 1–3 Oracle Solaris Cluster Components That Use Public-Network IP Addresses

Component	Number of IP Addresses Needed
Administrative console	1 IP address per subnet.
Global-cluster nodes	1 IP address per node, per subnet.
Zone-cluster nodes	1 IP address per node, per subnet.
Domain console network interface (Sun Fire 15000)	1 IP address per domain.
(Optional) Non-global zones	1 IP address per subnet.
Console-access device	1 IP address.
Logical addresses	1 IP address per logical host resource, per subnet.

For more information about planning IP addresses, see Chapter 2, "Planning Your TCP/IP Network (Tasks)," in *Oracle Solaris Administration: IP Services*.

Console-Access Devices

You must have console access to all cluster nodes. If you install Cluster Control Panel software on an administrative console, you must provide the hostname and port number of the console-access device that is used to communicate with the cluster nodes.

- A terminal concentrator is used to communicate between the administrative console and the global-cluster node consoles.
- A Sun Fire server uses a system controller instead of a terminal concentrator.

For more information about console access, see the *Oracle Solaris Cluster Concepts Guide*.

Alternatively, if you connect an administrative console directly to cluster nodes or through a management network, you instead provide the hostname of each global-cluster node and its serial port number that is used to connect to the administrative console or the management network.

Logical Addresses

Each data-service resource group that uses a logical address must have a hostname specified for each public network from which the logical address can be accessed.

For more information, see the *Oracle Solaris Cluster Data Services Planning and Administration Guide*. For additional information about data services and resources, also see the *Oracle Solaris Cluster Concepts Guide*.

Public Networks

Public networks communicate outside the cluster. Consider the following points when you plan your public-network configuration:

- **Separation of public and private network** – Public networks and the private network (cluster interconnect) must use separate adapters, or you must configure tagged VLAN on tagged-VLAN capable adapters and VLAN-capable switches to use the same adapter for both the private interconnect and the public network.
- **Minimum** – All cluster nodes must be connected to at least one public network. Public-network connections can use different subnets for different nodes.
- **Maximum** – You can have as many additional public-network connections as your hardware configuration allows.
- **Scalable services** – All nodes that run a scalable service must either use the same subnet or set of subnets or use different subnets that are routable among themselves.
- **IPv4** – Oracle Solaris Cluster software supports IPv4 addresses on the public network.

- **IPv6** – Oracle Solaris Cluster software supports IPv6 addresses on the public network for both failover and scalable data services.
- **IPMP groups** – Each public-network adapter that is used for data-service traffic must belong to an IP network multipathing (IPMP) group. If a public-network adapter is not used for data-service traffic, you do not have to configure it in an IPMP group.

  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 probe based.

  The `scinstall` utility ignores adapters that are already configured in an IPMP group. You can use probe-based IPMP groups or link-based IPMP groups in a cluster. But probe-based IPMP groups, which test the target IP address, provide the most protection by recognizing more conditions that might compromise availability.

  If any adapter in an IPMP group that the `scinstall` utility configures will not be used for data-service traffic, you can remove that adapter from the group.

  For guidelines and instructions to configure IPMP groups, follow the procedures in Part V, "IPMP," in *Oracle Solaris Administration: IP Services*. To modify IPMP groups after cluster installation, follow the guidelines in "How to Administer IP Network Multipathing Groups in a Cluster" in *Oracle Solaris Cluster System Administration Guide* and procedures in Chapter 28, "Administering IPMP (Tasks)," in *Oracle Solaris Administration: IP Services*.
- **Local MAC address support** – All public-network adapters must use network interface cards (NICs) that support local MAC address assignment. Local MAC address assignment is a requirement of IPMP.
- `local-mac-address` **setting** – The `local-mac-address?` variable must use the default value `true` for Ethernet adapters. Oracle Solaris Cluster software does not support a `local-mac-address?` value of `false` for Ethernet adapters.

For more information about public-network interfaces, see *Oracle Solaris Cluster Concepts Guide*.

Quorum Server Configuration

You can use Oracle Solaris Cluster Quorum Server software to configure a machine as a quorum server and then configure the quorum server as your cluster's quorum device. You can use a quorum server instead of or in addition to shared disks and NAS filers.

Consider the following points when you plan the use of a quorum server in an Oracle Solaris Cluster configuration.

- **Network connection** – The quorum-server computer connects to your cluster through the public network.
- **Supported hardware** – The supported hardware platforms for a quorum server are the same as for a global-cluster node.

- **Operating system** – Oracle Solaris software requirements for Oracle Solaris Cluster software apply as well to Quorum Server software.
- **Service to multiple clusters** – You can configure a quorum server that is installed with Oracle Solaris Cluster 3.3 3/13 quorum-server software as a quorum device to more than one cluster.
- **Mixed hardware and software** – You do not have to configure a quorum server on the same hardware and software platform as the cluster or clusters that it provides quorum to. For example, a SPARC based machine that runs the Oracle Solaris 10 OS can be configured as a quorum server for an x86 based cluster that runs the Oracle Solaris 10 OS.

  In addition, a cluster running Oracle Solaris Cluster 3.3 3/13 software can use a quorum server that runs a different version of the software than the cluster. See the Quorum Server Interoperability table in the Oracle Solaris Cluster 4 Compatibility Guide (`http://www.oracle.com/technetwork/server-storage/solaris-cluster/overview/solariscluster4-compatibilityguide-1429037.pdf`) for more information about mixed software versions.
- **Spanning tree algorithm** – You must disable the spanning tree algorithm on the Ethernet switches for the ports that are connected to the cluster public network where the quorum server will run.
- **Using a cluster node as a quorum server** – You can configure a quorum server on a cluster node to provide quorum for clusters **other than** the cluster that the node belongs to. However, a quorum server that is configured on a cluster node is not highly available.

NFS Guidelines

Consider the following points when you plan the use of Network File System (NFS) in an Oracle Solaris Cluster configuration.

- **NFS client** – No Oracle Solaris Cluster node can be an NFS client of an Oracle Solaris Cluster HA for NFS (HA for NFS) exported file system that is being mastered on a node in the same cluster. Such cross-mounting of HA for NFS is prohibited. Use the cluster file system to share files among global-cluster nodes.
- **NFSv3 protocol** – If you are mounting file systems on the cluster nodes from external NFS servers, such as NAS filers, and you are using the NFSv3 protocol, you cannot run NFS client mounts and the HA for NFS data service on the same cluster node. If you do, certain HA for NFS data-service activities might cause the NFS daemons to stop and restart, interrupting NFS services. However, you can safely run the HA for NFS data service if you use the NFSv4 protocol to mount external NFS file systems on the cluster nodes.
- **Locking** – Applications that run locally on the cluster must not lock files on a file system that is exported through NFS. Otherwise, local blocking (for example, `flock`(3UCB) or `fcntl`(2)) might interfere with the ability to restart the lock manager (`lockd`(1M)). During

restart, a blocked local process might be granted a lock which might be intended to be reclaimed by a remote client. This would cause unpredictable behavior.

- **NFS security features** – Oracle Solaris Cluster software does not support the following options of the `share_nfs(1M)` command:
  - `secure`
  - `sec=dh`

  However, Oracle Solaris Cluster software does support the following security features for NFS:
  - The use of secure ports for NFS. You enable secure ports for NFS by adding the entry set `nfssrv:nfs_portmon=1` to the `/etc/system` file on cluster nodes.
  - The use of Kerberos with NFS. For more information, see "Securing HA for NFS With Kerberos V5" in *Oracle Solaris Cluster Data Service for Network File System (NFS) Guide*.
  - **Fencing** – Zone clusters support fencing for all supported NAS devices, shared disks, and storage arrays.

Service Restrictions

Observe the following service restrictions for Oracle Solaris Cluster configurations:

- **Routers** – Do not configure cluster nodes as routers (gateways) due to the following reasons:
  - Routing protocols might inadvertently broadcast the cluster interconnect as a publicly reachable network to other routers, despite the setting of the `IFF_PRIVATE` flag on the interconnect interfaces.
  - Routing protocols might interfere with the failover of IP addresses across cluster nodes that impact client accessibility.
  - Routing protocols might compromise proper functionality of scalable services by accepting client network packets and dropping them, instead of forwarding the packets to other cluster nodes.
- **NIS+ servers** – Do not configure cluster nodes as NIS or NIS+ servers. There is no data service available for NIS or NIS+. However, cluster nodes can be NIS or NIS+ clients.
- **Boot and install servers** – Do not use an Oracle Solaris Cluster configuration to provide a highly available boot or installation service on client systems.
- **RARP** – Do not use an Oracle Solaris Cluster configuration to provide an `rarpd` service.
- **RPC program numbers** – If you install an RPC service on the cluster, the service must not use any of the following program numbers:
  - 100141
  - 100142
  - 100248

These numbers are reserved for the Oracle Solaris Cluster daemons `rgmd_receptionist`, `fed`, and `pmfd`, respectively.

If the RPC service that you install also uses one of these program numbers, you must change that RPC service to use a different program number.

- **Scheduling classes** – Oracle Solaris Cluster software does not support the running of high-priority process scheduling classes on cluster nodes. Do not run either of the following types of processes on cluster nodes:
  - Processes that run in the time-sharing scheduling class with a high priority
  - Processes that run in the real-time scheduling class

  Oracle Solaris Cluster software relies on kernel threads that do not run in the real-time scheduling class. Other time-sharing processes that run at higher-than-normal priority or real-time processes can prevent the Oracle Solaris Cluster kernel threads from acquiring needed CPU cycles.

Network Time Protocol (NTP)

Observe the following guidelines for NTP:

- **Synchronization** – The primary requirement when you configure NTP, or any time synchronization facility within the cluster, is that all cluster nodes must be synchronized to the same time.
- **Accuracy** – Consider accuracy of time on individual nodes to be of secondary importance to the synchronization of time among nodes. You are free to configure NTP as best meets your individual needs if this basic requirement for synchronization is met.
- **Error messages about nonexistent nodes** – Unless you have installed your own `/etc/inet/ntp.conf` file, the `scinstall` command installs a default `ntp.conf` file for you. The default file is shipped with references to the maximum number of nodes. Therefore, the `xntpd(1M)` daemon might issue error messages regarding some of these references at boot time. You can safely ignore these messages. See "How to Configure Network Time Protocol (NTP)" on page 140 for information about how to suppress these messages under otherwise normal cluster conditions.

See the *Oracle Solaris Cluster Concepts Guide* for further information about cluster time. See the `/etc/inet/ntp.cluster` template file for additional guidelines about how to configure NTP for an Oracle Solaris Cluster configuration.

Oracle Solaris Cluster Configurable Components

This section provides guidelines for the following Oracle Solaris Cluster components that you configure:

- "Global-Cluster Name" on page 28

Global-Cluster Name

Specify a name for the global cluster during Oracle Solaris Cluster configuration. The global cluster name should be unique throughout the enterprise.

For information about naming a zone cluster, see "Zone Clusters" on page 36.

Global-Cluster Voting-Node Names and Node IDs

The name of a voting node in a global cluster is the same name that you assign to the physical or virtual host when you install it with the Oracle Solaris OS. See the `hosts(4)` man page for information about naming requirements.

In single-host cluster installations, the default cluster name is the name of the voting node.

During Oracle Solaris Cluster configuration, you specify the names of all voting nodes that you are installing in the global cluster.

A node ID number is assigned to each cluster node for intracluster use, beginning with the number 1. Node ID numbers are assigned to each cluster node in the order that the node becomes a cluster member. If you configure all cluster nodes in one operation, the node from which you run the `scinstall` utility is the last node assigned a node ID number. You cannot change a node ID number after it is assigned to a cluster node.

A node that becomes a cluster member is assigned the lowest available node ID number. If a node is removed from the cluster, its node ID becomes available for assignment to a new node. For example, if in a four-node cluster the node that is assigned node ID `3` is removed and a new node is added, the new node is assigned node ID `3`, not node ID `5`.

If you want the assigned node ID numbers to correspond to certain cluster nodes, configure the cluster nodes one node at a time in the order that you want the node ID numbers to be assigned. For example, to have the cluster software assign node ID `1` to `phys-schost-1`, configure that node as the sponsoring node of the cluster. If you next add `phys-schost-2` to the cluster established by `phys-schost-1`, `phys-schost-2` is assigned node ID `2`.

For information about node names in a zone cluster, see "Zone Clusters" on page 36.

Zone Names

A non-global zone of brand `native` is a valid potential node of a resource-group node list. Use the naming convention *nodename*:*zonename* to specify a non-global zone to an Oracle Solaris Cluster command.

- The *nodename* is the name of the Oracle Solaris host.
- The *zonename* is the name that you assign to the non-global zone when you create the zone on the voting node. The zone name must be unique on the node. However, you can use the same zone name on different voting nodes. The different node name in *nodename*:*zonename* makes the complete non-global zone name unique in the cluster.

To specify the global zone, you need to specify only the voting-node name.

For information about a cluster of non-global zones, see "Zone Clusters" on page 36.

You can turn off cluster functionality for a selected non-global zone. A root user logged into one of these zones is not able to discover or disrupt operation of the cluster. For instructions, see "Denying Cluster Services for a Non-Global Zone" in *Oracle Solaris Cluster Data Service for Solaris Containers Guide*.

Private Network Configuration

Note – You do not need to configure a private network for a single-host global cluster. The `scinstall` utility automatically assigns the default private-network address and netmask, even though a private network is not used by the cluster.

Oracle Solaris Cluster software uses the private network for internal communication among nodes and among non-global zones that are managed by Oracle Solaris Cluster software. An Oracle Solaris Cluster configuration requires at least two connections to the cluster interconnect on the private network. When you configure Oracle Solaris Cluster software on the first node of the cluster, you specify the private-network address and netmask in one of the following ways:

- Accept the default private-network address (`172.16.0.0`) and default netmask (`255.255.240.0`). This IP address range supports a combined maximum of 64 voting nodes and non-global zones, a maximum of 12 zone clusters, and a maximum of 10 private networks.

  Note – The maximum number of voting nodes that an IP address range can support does not reflect the maximum number of voting nodes that the hardware or software configuration can currently support.

- Specify a different allowable private-network address and accept the default netmask.

- Accept the default private-network address and specify a different netmask.
- Specify both a different private-network address and a different netmask.

If you choose to specify a different netmask, the `scinstall` utility prompts you for the number of nodes and the number of private networks that you want the IP address range to support. The utility also prompts you for the number of zone clusters that you want to support. The number of global-cluster nodes that you specify should also include the expected number of unclustered non-global zones that will use the private network.

The utility calculates the netmask for the minimum IP address range that will support the number of nodes, zone clusters, and private networks that you specified. The calculated netmask might support more than the supplied number of nodes, including non-global zones, zone clusters, and private networks. The `scinstall` utility also calculates a second netmask that would be the minimum to support twice the number of nodes, zone clusters, and private networks. This second netmask would enable the cluster to accommodate future growth without the need to reconfigure the IP address range.

The utility then asks you what netmask to choose. You can specify either of the calculated netmasks or provide a different one. The netmask that you specify must minimally support the number of nodes and private networks that you specified to the utility.

Note – Changing the cluster private IP-address range might be necessary to support the addition of voting nodes, non-global zones, zone clusters, or private networks.

To change the private-network address and netmask after the cluster is established, see "How to Change the Private Network Address or Address Range of an Existing Cluster" in *Oracle Solaris Cluster System Administration Guide*. You must bring down the cluster to make these changes.

However, the cluster can remain in cluster mode if you use the `cluster set-netprops` command to change only the netmask. For any zone cluster that is already configured in the cluster, the private IP subnets and the corresponding private IP addresses that are allocated for that zone cluster will also be updated.

If you specify a private-network address other than the default, the address must meet the following requirements:

- **Address and netmask sizes** – The private network address cannot be smaller than the netmask. For example, you can use a private network address of `172.16.10.0` with a netmask of `255.255.255.0`. But you cannot use a private network address of `172.16.10.0` with a netmask of `255.255.0.0`.
- **Acceptable addresses** – The address must be included in the block of addresses that RFC 1918 reserves for use in private networks. You can contact the InterNIC to obtain copies of RFCs or view RFCs online at http://www.rfcs.org.

- **Use in multiple clusters** – You can use the same private-network address in more than one cluster, provided that the clusters are on different private networks. Private IP network addresses are not accessible from outside the physical cluster.

  For Oracle VM Server for SPARC guest domains that are created on the same physical machine and that are connected to the same virtual switch, the private network is shared by such guest domains and is visible to all these domains. Proceed with caution before you specify a private-network IP address range to the `scinstall` utility for use by a cluster of guest domains. Ensure that the address range is not already in use by another guest domain that exists on the same physical machine and shares its virtual switch.
- **VLANs shared by multiple clusters** – Oracle Solaris Cluster configurations support the sharing of the same private-interconnect VLAN among multiple clusters. It is not required to configure a separate VLAN for each cluster. However, limiting the use of a VLAN to a single cluster provides better fault isolation and interconnect resilience.
- **IPv6** – Oracle Solaris Cluster software does not support IPv6 addresses for the private interconnect. The system does configure IPv6 addresses on the private-network adapters to support scalable services that use IPv6 addresses. But internode communication on the private network does not use these IPv6 addresses.

See Chapter 2, "Planning Your TCP/IP Network (Tasks)," in *Oracle Solaris Administration: IP Services* for more information about private networks.

Private Hostnames

The private hostname is the name that is used for internode communication over the private-network interface. Private hostnames are automatically created during Oracle Solaris Cluster configuration of a global cluster or a zone cluster. These private hostnames follow the naming convention `clusternode`*nodeid* `-priv`, where *nodeid* is the numeral of the internal node ID. During Oracle Solaris Cluster configuration, the node ID number is automatically assigned to each voting node when the node becomes a cluster member. A voting node of the global cluster and a node of a zone cluster can both have the same private hostname, but each hostname resolves to a different private-network IP address.

After a global cluster is configured, you can rename its private hostnames by using the `clsetup(1CL)` utility. Currently, you cannot rename the private hostname of a zone-cluster node.

The creation of a private hostname for a non-global zone is optional. There is no required naming convention for the private hostname of a non-global zone.

Cluster Interconnect

The cluster interconnects provide the hardware pathways for private-network communication between cluster nodes. Each interconnect consists of a cable that is connected in one of the following ways:

- Between two transport adapters

- Between a transport adapter and a transport switch

For more information about the purpose and function of the cluster interconnect, see "Cluster Interconnect" in *Oracle Solaris Cluster Concepts Guide*.

Note – You do not need to configure a cluster interconnect for a single-host cluster. However, if you anticipate eventually adding more voting nodes to a single-host cluster configuration, you might want to configure the cluster interconnect for future use.

During Oracle Solaris Cluster configuration, you specify configuration information for one or two cluster interconnects.

- If the number of available adapter ports is limited, you can use tagged VLANs to share the same adapter with both the private and public network. For more information, see the guidelines for tagged VLAN adapters in "Transport Adapters" on page 32.
- You can set up from one to six cluster interconnects in a cluster. While a single cluster interconnect reduces the number of adapter ports that are used for the private interconnect, it provides no redundancy and less availability. If a single interconnect fails, the cluster is at a higher risk of having to perform automatic recovery. Whenever possible, install two or more cluster interconnects to provide redundancy and scalability, and therefore higher availability, by avoiding a single point of failure.

You can configure additional cluster interconnects, up to six interconnects total, after the cluster is established by using the `clsetup(1CL)` utility.

For guidelines about cluster interconnect hardware, see "Interconnect Requirements and Restrictions" in *Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual*. For general information about the cluster interconnect, see "Cluster Interconnect" in *Oracle Solaris Cluster Concepts Guide*.

Transport Adapters

For the transport adapters, such as ports on network interfaces, specify the transport adapter names and transport type. If your configuration is a two-host cluster, you also specify whether your interconnect is a point-to-point connection (adapter to adapter) or uses a transport switch.

Consider the following guidelines and restrictions:

- **IPv6** – Oracle Solaris Cluster software does not support IPv6 communications over the private interconnects.
- **Local MAC address assignment** – All private network adapters must use network interface cards (NICs) that support local MAC address assignment. Link-local IPv6 addresses, which are required on private-network adapters to support IPv6 public-network addresses, are derived from the local MAC addresses.

- **Tagged VLAN adapters** – Oracle Solaris Cluster software supports tagged Virtual Local Area Networks (VLANs) to share an adapter between the private cluster interconnect and the public network. To configure a tagged VLAN adapter for the cluster interconnect, specify the adapter name and its VLAN ID (VID) in one of the following ways:
  - Specify the usual adapter name, which is the device name plus the instance number or physical point of attachment (PPA). For example, the name of instance 2 of a Cassini Gigabit Ethernet adapter would be `ce2`. If the `scinstall` utility asks whether the adapter is part of a shared virtual LAN, answer **`yes`** and specify the adapter's VID number.
  - Specify the adapter by its VLAN virtual device name. This name is composed of the adapter name plus the VLAN instance number. The VLAN instance number is derived from the formula (1000**V*)+*N*, where *V* is the VID number and *N* is the PPA.

    As an example, for VID 73 on adapter `ce2`, the VLAN instance number would be calculated as (1000*73)+2. You would therefore specify the adapter name as `ce73002` to indicate that it is part of a shared virtual LAN.

  For information about configuring VLAN in a cluster, see "Configuring VLANs as Private Interconnect Networks" in *Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual*. For general information about VLAN, see "Administering Virtual Local Area Networks" in *Oracle Solaris Administration: IP Services*.
- **SPARC: Oracle VM Server for SPARC guest domains** – Specify adapter names by their virtual names, `vnet`*N*, such as `vnet0` and `vnet1`. Virtual adapter names are recorded in the `/etc/path_to_inst` file.
- **Logical network interfaces** – Logical network interfaces are reserved for use by Oracle Solaris Cluster software.

See the `scconf_trans_adap_*`(1M) family of man pages for information about a specific transport adapter.

Transport Switches

If you use transport switches, such as a network switch, specify a transport switch name for each interconnect. You can use the default name `switch`*N*, where *N* is a number that is automatically assigned during configuration, or create another name.

Also specify the switch port name or accept the default name. The default port name is the same as the internal node ID number of the Oracle Solaris host that hosts the adapter end of the cable. However, you cannot use the default port name for certain adapter types.

Note – Clusters with three or more voting nodes *must* use transport switches. Direct connection between voting cluster nodes is supported only for two-host clusters.

If your two-host cluster is direct connected, you can still specify a transport switch for the interconnect.

Tip – If you specify a transport switch, you can more easily add another voting node to the cluster in the future.

Global Fencing

Fencing is a mechanism that is used by the cluster to protect the data integrity of a shared disk during split-brain situations. By default, the `scinstall` utility in Typical Mode leaves global fencing enabled, and each shared disk in the configuration uses the default global fencing setting of `prefer3`. With the `prefer3` setting, the SCSI-3 protocol is used.

In Custom Mode, the `scinstall` utility prompts you whether to disable global fencing. For most situations, respond **No** to keep global fencing enabled. However, you can disable global fencing in certain situations.

Caution – If you disable fencing under other situations than the following, your data might be vulnerable to corruption during application failover. Examine this data corruption possibility carefully when you consider turning off fencing.

The situations in which you can disable global fencing are as follows:

- The shared storage does not support SCSI reservations.

  If you turn off fencing for a shared disk that you then configure as a quorum device, the device uses the software quorum protocol. This is true regardless of whether the disk supports SCSI-2 or SCSI-3 protocols. Software quorum is a protocol in Oracle Solaris Cluster software that emulates a form of SCSI Persistent Group Reservations (PGR).
- You want to enable systems that are outside the cluster to gain access to storage that is attached to the cluster.

If you disable global fencing during cluster configuration, fencing is turned off for all shared disks in the cluster. After the cluster is configured, you can change the global fencing protocol or override the fencing protocol of individual shared disks. However, to change the fencing protocol of a quorum device, you must first unconfigure the quorum device. Then set the new fencing protocol of the disk and reconfigure it as a quorum device.

For more information about fencing behavior, see "Failfast Mechanism" in *Oracle Solaris Cluster Concepts Guide*. For more information about setting the fencing protocol of individual shared disks, see the `cldevice`(1CL) man page. For more information about the global fencing setting, see the `cluster`(1CL) man page.

Quorum Devices

Oracle Solaris Cluster configurations use quorum devices to maintain data and resource integrity. If the cluster temporarily loses connection to a voting node, the quorum device

prevents amnesia or split-brain problems when the voting cluster node attempts to rejoin the cluster. For more information about the purpose and function of quorum devices, see "Quorum and Quorum Devices" in *Oracle Solaris Cluster Concepts Guide*.

During Oracle Solaris Cluster installation of a two-host cluster, you can choose to let the `scinstall` utility automatically configure as a quorum device an available shared disk in the configuration. Shared disks include any Sun NAS device that is configured for use as a shared disk. The `scinstall` utility assumes that all available shared disks are supported as quorum devices.

If you want to use a quorum server as the quorum device, you add it to the cluster configuration after `scinstall` processing is completed. For more information about quorum servers, see "Quorum Server Configuration" on page 24.

After installation, you can also configure additional quorum devices by using the `clsetup` utility.

Note – You do not need to configure quorum devices for a single-host cluster.

If your cluster configuration includes third-party shared storage devices that are not supported for use as quorum devices, you must use the `clsetup` utility to configure quorum manually.

Consider the following points when you plan quorum devices.

- **Minimum** – A two-host cluster must have at least one quorum device, which can be a shared disk, a quorum server, or a NAS device. For other topologies, quorum devices are optional.
- **Odd-number rule** – If more than one quorum device is configured in a two-host cluster, or in a pair of hosts directly connected to the quorum device, configure an odd number of quorum devices. This configuration ensures that the quorum devices have completely independent failure pathways.
- **Distribution of quorum votes** – 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 voting nodes. Otherwise, the nodes cannot form a cluster if all quorum devices are unavailable, even if all nodes are functioning.
- **Connection** – You must connect a quorum device to at least two voting nodes.
- **SCSI fencing protocol** – When a SCSI shared-disk quorum device is configured, its fencing protocol is automatically set to SCSI-2 in a two-host cluster or SCSI-3 in a cluster with three or more voting nodes.
- **Changing the fencing protocol of quorum devices** – For SCSI disks that are configured as a quorum device, you must unconfigure the quorum device before you can enable or disable its SCSI fencing protocol.

- **Software quorum protocol** – You can configure supported shared disks that do not support SCSI protocol, such as SATA disks, as quorum devices. You must disable fencing for such disks. The disks would then use software quorum protocol, which emulates SCSI PGR.

  The software quorum protocol would also be used by SCSI shared disks if fencing is disabled for such disks.
- **Replicated devices** – Oracle Solaris Cluster software does not support replicated devices as quorum devices.
- **ZFS storage pools** – Do not add a configured quorum device to a ZFS storage pool. When 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. The disk can then no longer provide a quorum vote to the cluster.

  After a disk is in a storage pool, you can configure that disk as a quorum device. Or, you can unconfigure the quorum device, add it to the storage pool, then reconfigure the disk as a quorum device.

For more information about quorum devices, see "Quorum and Quorum Devices" in *Oracle Solaris Cluster Concepts Guide*.

Zone Clusters

A zone cluster is a cluster of Oracle Solaris non-global zones. All nodes of a zone cluster are configured as non-global zones of the `cluster` brand. No other brand type is permitted in a zone cluster. You can run supported services on the zone cluster similar to a global cluster, with the isolation that is provided by Oracle Solaris zones.

You can use the `clsetup` utility to create a zone cluster and add a network address, file system, ZFS storage pool, or storage device. You can also use a command line interface (the `clzonecluster` utility) to create a zone cluster, make configuration changes, and remove a zone cluster. For more information about using the `clzonecluster` utility, see the `clzonecluster(1CL)` man page.

Consider the following points when you plan the creation of a zone cluster.

- "Global-Cluster Requirements and Guidelines" on page 36
- "Zone-Cluster Requirements and Guidelines" on page 37
- "Guidelines for Trusted Extensions in a Zone Cluster" on page 39

Global-Cluster Requirements and Guidelines

- **Global cluster** – The zone cluster must be configured on a global Oracle Solaris Cluster configuration. A zone cluster cannot be configured without an underlying global cluster.

- **Cluster mode** – The global-cluster voting node from which you create or modify a zone cluster must be in cluster mode. If any other voting nodes are in noncluster mode when you administer a zone cluster, the changes that you make are propagated to those nodes when they return to cluster mode.
- **Adequate private-IP addresses** – The private IP-address range of the global cluster must have enough free IP-address subnets for use by the new zone cluster. If the number of available subnets is insufficient, the creation of the zone cluster fails.
- **Changes to the private IP-address range** – The private IP subnets and the corresponding private IP-addresses that are available for zone clusters are automatically updated if the global cluster's private IP-address range is changed. If a zone cluster is deleted, the cluster infrastructure frees the private IP-addresses that were used by that zone cluster, making the addresses available for other use within the global cluster and by any other zone clusters that depend on the global cluster.
- **Supported devices** – Devices that are supported with Oracle Solaris zones can be exported to a zone cluster. Such devices include the following:
  - Oracle Solaris disk devices (c*N*t*X*d*Y*s*Z*)
  - DID devices (`/dev/did/*dsk/d`*N*)
  - Solaris Volume Manager and Solaris Volume Manager for Sun Cluster multi-owner disk sets (`/dev/md/`*setname*`/*dsk/d`*N*)

Zone-Cluster Requirements and Guidelines

- **Distribution of nodes** – You cannot host multiple nodes of the same zone cluster on the same host machine. A host can support multiple zone-cluster nodes as long as each zone-cluster node on that host is a member of a different zone cluster.
- **Node creation** – You must create at least one zone-cluster node at the time that you create the zone cluster. You can use the `clsetup` utility or the `clzonecluster` command to create the zone cluster. The name of the zone-cluster node must be unique within the zone cluster. The infrastructure automatically creates an underlying non-global zone on each host that supports the zone cluster. Each non-global zone is given the same zone name, which is derived from, and identical to, the name that you assign to the zone cluster when you create the cluster. For example, if you create a zone cluster that is named `zc1`, the corresponding non-global zone name on each host that supports the zone cluster is also `zc1`.
- **Cluster name** – Each zone-cluster name must be unique throughout the cluster of machines that host the global cluster. The zone-cluster name cannot also be used by a non-global zone elsewhere in the cluster of machines, nor can the zone-cluster name be the same as that of a global-cluster node. You cannot use “all” or “global” as a zone-cluster name, because these are reserved names.
- **Public-network IP addresses** – You can optionally assign a specific public-network IP address to each zone-cluster node.

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

- 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.
- The cluster software will activate any Logical Host IP address on any NIC.

- **Private hostnames** – During creation of the zone cluster, a private hostname is automatically created for each node of the zone cluster, in the same way that hostnames are created in global clusters. Currently, you cannot rename the private hostname of a zone-cluster node. For more information about private hostnames, see "Private Hostnames" on page 31.
- **Oracle Solaris zones brands** – All nodes of a zone cluster are configured as non-global zones of the `cluster` brand. No other brand type is permitted in a zone cluster.
- `Global_zone=TRUE` **resource-type property** – To register a resource type that uses the `Global_zone=TRUE` resource-type property, the resource-type file must reside in the `/usr/cluster/global/rgm/rtreg/` directory of the zone cluster. If that resource-type file resides in any other location, the command to register the resource type is rejected.
- **Conversion to a zone-cluster node** – You cannot add to a zone cluster a non-global zone that resides outside that zone cluster. You must use only the `clzonecluster` command to add new nodes to a zone cluster.
- **File systems** – You can use the `clsetup` utility or the `clzonecluster` command to add the following types of file systems for use by a zone cluster. A file system is exported to a zone cluster by using either a direct mount or a loopback mount. Adding a file system with the `clsetup` utility is done in cluster scope, which affects the entire zone cluster.
  - By direct mount:
    - UFS local file system
    - QFS standalone file system
    - QFS shared file system, only when used to support Oracle Real Application Clusters
    - ZFS (exported as a data set)
    - NFS from supported NAS devices
  - By loopback mount:
    - UFS local file system
    - QFS standalone file system
    - QFS shared file system, only when used to support Oracle Real Application Clusters
    - UFS cluster file system

You configure an HAStoragePlus or `ScalMountPoint` resource to manage the mounting of the file system.

- **Fencing** – Zone clusters support fencing for all supported NAS devices, shared disks, and storage arrays.

Guidelines for Trusted Extensions in a Zone Cluster

Consider the following points when you use the Trusted Extensions feature of Oracle Solaris in a zone cluster:

- **Only zone-cluster support** – In an Oracle Solaris Cluster configuration with Trusted Extensions enabled, applications must run only in a zone cluster. No other non-global zones can be used on the cluster. You must use only the `clzonecluster` command to create a zone cluster. Do not use the `txzonemgr` command to create a non-global zone on a cluster that has Trusted Extensions enabled.
- **Trusted Extensions scope** – You can either enable or disable Trusted Extensions for the entire cluster configuration. When Trusted Extensions is enabled, all non-global zones in the cluster configuration must belong to one of the zone clusters in the cluster. You cannot configure any other kind of non-global zone without compromising security.
- **IP addresses** – Each zone cluster that uses Trusted Extensions must use its own IP addresses. The special networking feature in Trusted Extensions that enables an IP address to be shared between multiple non-global zones is not supported with Oracle Solaris Cluster software.
- **Loopback mounts** – You cannot use loopback mounts that have write permissions in a zone cluster that uses Trusted Extensions. Use only direct mounts of file systems that permit write access, or use loopback mounts that have only read permissions.
- **File systems** – Do not configure in the zone cluster the global device that underlies a file system. Configure only the file system itself in the zone cluster.
- **Storage device name** – Do not add an individual slice of a storage device to a zone cluster. You must add the entire device to a single zone cluster. The use of slices of the same storage device in different zone clusters compromises the security of those zone clusters.
- **Application installation** – Install applications only in the zone cluster or in the global cluster and then exported to the zone cluster by using read-only loopback mounts.
- **Zone cluster isolation** – When Trusted Extensions is used, the name of a zone cluster is a security label. In some cases, the security label itself might be information that cannot be disclosed, and the name of a resource or resource group might be a sensitive piece of information that cannot be disclosed. When an inter-cluster resource dependency or inter-cluster resource-group affinity is configured, the name of the other cluster becomes visible as well as the name of any affected resource or resource group. Therefore, before you establish any inter-cluster relationships, evaluate whether this information can be made visible according to the your requirements.

Planning the Global Devices, Device Groups, and Cluster File Systems

This section provides the following guidelines for planning global devices and for planning cluster file systems:

Planning Global Devices

For information about the purpose and function of global devices, see "Global Devices" in *Oracle Solaris Cluster Concepts Guide*.

Oracle Solaris Cluster software does not require any specific disk layout or file system size. Consider the following points when you plan your layout for global devices.

- **Mirroring** – You must mirror all global devices for the global device to be considered highly available. You do not need to use software mirroring if the storage device provides hardware RAID as well as redundant paths to disks.
- **Disks** – When you mirror, lay out file systems so that the file systems are mirrored across disk arrays.
- **Availability** – You must physically connect a global device to more than one voting node in the cluster for the global device to be considered highly available. A global device with multiple physical connections can tolerate a single-node failure. A global device with only one physical connection is supported, but the global device becomes inaccessible from other voting nodes if the node with the connection is down.
- **Swap devices** – Do not create a swap file on a global device.
- **Non-global zones** – Global devices are not directly accessible from a non-global zone. Only cluster-file-system data is accessible from a non-global zone.

Planning Device Groups

For information about the purpose and function of device groups, see "Device Groups" in *Oracle Solaris Cluster Concepts Guide*.

Consider the following points when you plan device groups.

- **Failover** – You can configure multihost disks and properly configured volume-manager devices as failover devices. Proper configuration of a volume-manager device includes multihost disks and correct setup of the volume manager itself. This configuration ensures that multiple voting nodes can host the exported device. You cannot configure tape drives, CD-ROMs or DVD-ROMs, or single-ported devices as failover devices.
- **Mirroring** – You must mirror the disks to protect the data from disk failure. See "Mirroring Guidelines" on page 46 for additional guidelines. See "Configuring Solaris Volume Manager Software" on page 149 and your volume-manager documentation for instructions about mirroring.
- **Storage-based replication** – Disks in a device group must be either all replicated or none replicated. A device group cannot use a mix of replicated and nonreplicated disks.

Planning Cluster File Systems

For information about the purpose and function of cluster file systems, see "Cluster File Systems" in *Oracle Solaris Cluster Concepts Guide*.

Note – You can alternatively configure highly available local file systems. This can provide better performance to support a data service with high I/O, or to permit use of certain file-system features that are not supported in a cluster file system. For more information, see "Enabling Highly Available Local File Systems" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

Consider the following points when you plan cluster file systems.

- **Quotas** – Quotas are not supported on cluster file systems. However, quotas are supported on highly available local file systems.
- **Non-global zones** – If a cluster file system is to be accessed from a non-global zone, it must first be mounted in the global zone. The cluster file system is then mounted in the non-global zone by using a loopback mount. Therefore, the loopback file system (LOFS) must be enabled in a cluster that contains non-global zones.
- **Zone clusters** – You cannot configure cluster file systems that use UFS for use in a zone cluster. Use highly available local file systems instead. You can use a QFS shared file system in a zone cluster, but only to support Oracle RAC.
- **Loopback file system (LOFS)** – During cluster creation, LOFS is enabled by default. You must manually disable LOFS on each voting cluster node if the cluster meets both of the following conditions:
  - Oracle Solaris Cluster HA for NFS (HA for NFS) is configured on a highly available local file system.
  - The `automountd` daemon is running.

If the cluster meets both of these conditions, you must disable LOFS to avoid switchover problems or other failures. If the cluster meets only one of these conditions, you can safely enable LOFS.

If you require both LOFS and the `automountd` daemon to be enabled, exclude from the automounter map all files that are part of the highly available local file system that is exported by HA for NFS.

- **Process accounting log files** – Do not locate process accounting log files on a cluster file system or on a highly available local file system. A switchover would be blocked by writes to the log file, which would cause the node to hang. Use only a local file system to contain process accounting log files.
- **Communication endpoints** – The cluster file system does not support any of the file-system features of Oracle Solaris software by which one would put a communication endpoint in the file-system namespace.
  - Although you can create a UNIX domain socket whose name is a path name into the cluster file system, the socket would not survive a node failover.
  - Any FIFOs or named pipes that you create on a cluster file system would not be globally accessible.

  Therefore, do not attempt to use the `fattach` command from any node other than the local node.
- **Device special files** – Neither block special files nor character special files are supported in a cluster file system. To specify a path name to a device node in a cluster file system, create a symbolic link to the device name in the `/dev` directory. Do not use the `mknod` command for this purpose.
- **atime** – Cluster file systems do not maintain `atime`.
- **ctime** – When a file on a cluster file system is accessed, the update of the file's `ctime` might be delayed.
- **Installing applications** - If you want the binaries of a highly available application to reside on a cluster file system, wait to install the application until after the cluster file system is configured.

Choosing Mount Options for UFS Cluster File Systems

This section describes requirements and restrictions for mount options of UFS cluster file systems.

Note – You can alternatively configure this and other types of file systems as highly available local file systems. For more information, see "Enabling Highly Available Local File Systems" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

Follow these guidelines to determine what mount options to use when you create your cluster file systems.

Mount Option	Usage	Description
`global`	Required	This option makes the file system globally visible to all nodes in the cluster.
`logging`	Required	This option enables logging.
`forcedirectio`	Conditional	This option is required only for cluster file systems that will host Oracle Real Application Clusters RDBMS data files, log files, and control files.
`onerror=panic`	Required	You do not have to explicitly specify the `onerror=panic` mount option in the `/etc/vfstab` file. This mount option is already the default value if no other `onerror` mount option is specified. **Note –** Only the `onerror=panic` mount option is supported by Oracle Solaris Cluster software. Do not use the `onerror=umount` or `onerror=lock` mount options. These mount options are not supported on cluster file systems for the following reasons: ■ Use of the `onerror=umount` or `onerror=lock` mount option might cause the cluster file system to lock or become inaccessible. This condition might occur if the cluster file system experiences file corruption. ■ The `onerror=umount` or `onerror=lock` mount option might cause the cluster file system to become unmountable. This condition might thereby cause applications that use the cluster file system to hang or prevent the applications from being killed. A node might require rebooting to recover from these states.
`syncdir`	Optional	If you specify `syncdir`, you are guaranteed POSIX-compliant file system behavior for the `write()` system call. If a `write()` succeeds, then this mount option ensures that sufficient space is on the disk. If you do not specify `syncdir`, the same behavior occurs that is seen with UFS file systems. When you do not specify `syncdir`, performance of writes that allocate disk blocks, such as when appending data to a file, can significantly improve. However, in some cases, without `syncdir` you would not discover an out-of-space condition (`ENOSPC`) until you close a file. You see `ENOSPC` on close only during a very short time after a failover. With `syncdir`, as with POSIX behavior, the out-of-space condition would be discovered before the close.

See the `mount_ufs(1M)` man page for more information about UFS mount options.

Mount Information for Cluster File Systems

Consider the following points when you plan mount points for cluster file systems.

- **Mount-point location** – Create mount points for cluster file systems in the `/global` directory, unless you are prohibited by other software products. By using the `/global` directory, you can more easily distinguish cluster file systems, which are globally available, from local file systems.
- **Nesting mount points** – Normally, you should not nest the mount points for cluster file systems. For example, do not set up one file system that is mounted on `/global/a` and another file system that is mounted on `/global/a/b`. To ignore this rule can cause availability and node boot-order problems. These problems would occur if the parent mount point is not present when the system attempts to mount a child of that file system.

  The only exception to this rule, for cluster file systems on UFS, is if the devices for the two file systems have the same physical host connectivity. An example is different slices on the same disk.

  Note – This restriction still applies to QFS shared file systems, even if the two file-system devices have the same physical host connectivity.

- `forcedirectio` – Oracle Solaris Cluster software does not support the execution of binaries off cluster file systems that are mounted by using the `forcedirectio` mount option.

Planning Volume Management

This section provides the following guidelines for planning volume management of your cluster configuration:

- "Guidelines for Volume-Manager Software" on page 45
- "Guidelines for Solaris Volume Manager Software" on page 45
- "File-System Logging" on page 46
- "Mirroring Guidelines" on page 46

Oracle Solaris Cluster software uses volume-manager software to group disks into device groups which can then be administered as one unit. Oracle Solaris Cluster software supports Solaris Volume Manager software You must install Solaris Volume Manager software on all voting nodes of the cluster.

See your volume-manager documentation and "Configuring Solaris Volume Manager Software" on page 149 for instructions about how to install and configure the volume-manager software. For more information about the use of volume management in a cluster configuration, see "Multihost Devices" in *Oracle Solaris Cluster Concepts Guide* and "Device Groups" in *Oracle Solaris Cluster Concepts Guide*.

Guidelines for Volume-Manager Software

Consider the following general guidelines when you configure your disks with volume-manager software:

- **Software RAID** – Oracle Solaris Cluster software does not support software RAID 5.
- **Mirrored multihost disks** – You must mirror all multihost disks across disk expansion units. See "Guidelines for Mirroring Multihost Disks" on page 46 for guidelines on mirroring multihost disks. You do not need to use software mirroring if the storage device provides hardware RAID as well as redundant paths to devices.
- **Mirrored root** – Mirroring the root disk ensures high availability, but such mirroring is not required. See "Mirroring Guidelines" on page 46 for guidelines about deciding whether to mirror the root disk.
- **Unique naming** – You might have local Solaris Volume Manager volumes that are used as devices on which the `/global/.devices/node@`*nodeid* file systems are mounted. If so, the name of each local volume on which a `/global/.devices/node@`*nodeid* file system is to be mounted must be unique throughout the cluster.
- **Node lists** – To ensure high availability of a device group, make its node lists of potential masters and its failback policy identical to any associated resource group. Or, if a scalable resource group uses more nodes than its associated device group, make the scalable resource group's node list a superset of the device group's node list. See the resource group planning information in the *Oracle Solaris Cluster Data Services Planning and Administration Guide* for information about node lists.
- **Multihost disks** – You must connect, or port, all devices that are used to construct a device group to all of the nodes that are configured in the node list for that device group. Solaris Volume Manager software can automatically check for this connection at the time that devices are added to a disk set.
- **Hot-spare disks** – You can use hot-spare disks to increase availability, but hot spare disks are not required.

See your volume-manager documentation for disk layout recommendations and any additional restrictions.

Guidelines for Solaris Volume Manager Software

Consider the following points when you plan Solaris Volume Manager configurations:

- **Local volume names** – The name of each local Solaris Volume Manager volume on which a global-devices file system, `/global/.devices/node@`*nodeid*, is mounted must be unique throughout the cluster. Also, the name cannot be the same as any device-ID name.

- **Dual-string mediators** – A disk string consists of a disk enclosure, its physical disks, cables from the enclosure to the host or hosts, and the interface adapter cards. Each disk set configured with exactly two disk strings and mastered by exactly two Oracle Solaris hosts is called a dual-string disk set. Such a disk set must have Solaris Volume Manager dual-string mediators configured. Observe the following rules when you configure dual-string mediators:
  - You must configure each disk set with two or three hosts that act as mediator hosts.
  - You must use the hosts that can master a disk set as mediators for that disk set. If you have a campus cluster, you can also configure a third node or a non-clustered host on the cluster network as a third mediator host to improve availability.
  - Mediators cannot be configured for disk sets that do not meet the two-string and two-host requirements.

  See the `mediator(7D)` man page for details.

File-System Logging

Logging is required for UFS cluster file systems. Oracle Solaris Cluster software supports Oracle Solaris UFS logging. See the `mount_ufs(1M)` man page for more information.

Solaris Volume Manager supports both types of file-system logging.

Mirroring Guidelines

This section provides the following guidelines for planning the mirroring of your cluster configuration:

- "Guidelines for Mirroring Multihost Disks" on page 46
- "Guidelines for Mirroring the Root Disk" on page 47

Guidelines for Mirroring Multihost Disks

To mirror all multihost disks in an Oracle Solaris Cluster configuration enables the configuration to tolerate single-device failures. Oracle Solaris Cluster software requires that you mirror all multihost disks across expansion units. You do not need to use software mirroring if the storage device provides hardware RAID as well as redundant paths to devices.

Consider the following points when you mirror multihost disks:

- **Separate disk expansion units** – Each submirror of a given mirror should reside in a different multihost expansion unit.
- **Disk space** – Mirroring doubles the amount of necessary disk space.
- **Three-way mirroring** – Solaris Volume Manager software supports three-way mirroring. However, Oracle Solaris Cluster software requires only two-way mirroring.

- **Differing device sizes** – If you mirror to a device of a different size, your mirror capacity is limited to the size of the smallest submirror.

For more information about multihost disks, see "Multihost Devices" in *Oracle Solaris Cluster Concepts Guide*.

Guidelines for Mirroring the Root Disk

For maximum availability, mirror root (`/`), `/usr`, `/var`, `/opt`, and `swap` on the local disks. However, Oracle Solaris Cluster software does not require that you mirror the root disk.

Before you decide whether to mirror the root disk, consider the risks, complexity, cost, and service time for the various alternatives that concern the root disk. No single mirroring strategy works for all configurations. You might want to consider your local Oracle service representative's preferred solution when you decide whether to mirror root.

See your volume-manager documentation and "Configuring Solaris Volume Manager Software" on page 149 for instructions about how to mirror the root disk.

Consider the following points when you decide whether to mirror the root disk.

- **Boot disk** – You can set up the mirror to be a bootable root disk. You can then boot from the mirror if the primary boot disk fails.
- **Complexity** – To mirror the root disk adds complexity to system administration. To mirror the root disk also complicates booting in single-user mode.
- **Backups** – Regardless of whether you mirror the root disk, you also should perform regular backups of root. Mirroring alone does not protect against administrative errors. Only a backup plan enables you to restore files that have been accidentally altered or deleted.
- **Quorum devices** – Do not use a disk that was configured as a quorum device to mirror a root disk.
- **Quorum** – Under Solaris Volume Manager software, in failure scenarios in which state database quorum is lost, you cannot reboot the system until maintenance is performed. See your Solaris Volume Manager documentation for information about the state database and state database replicas.
- **Separate controllers** – Highest availability includes mirroring the root disk on a separate controller.
- **Secondary root disk** – With a mirrored root disk, the primary root disk can fail but work can continue on the secondary (mirror) root disk. Later, the primary root disk might return to service, for example, after a power cycle or transient I/O errors. Subsequent boots are then performed by using the primary root disk that is specified for the `eeprom(1M)` `boot-device` parameter. In this situation, no manual repair task occurs, but the drive starts working well enough to boot. With Solaris Volume Manager software, a resync does occur. A resync requires a manual step when the drive is returned to service.

If changes were made to any files on the secondary (mirror) root disk, they would not be reflected on the primary root disk during boot time. This condition would cause a stale submirror. For example, changes to the `/etc/system` file would be lost. With Solaris Volume Manager software, some administrative commands might have changed the `/etc/system` file while the primary root disk was out of service.

The boot program does not check whether the system is booting from a mirror or from an underlying physical device. The mirroring becomes active partway through the boot process, after the volumes are loaded. Before this point, the system is therefore vulnerable to stale submirror problems.

◆ ◆ ◆ CHAPTER 2

Installing Software on Global-Cluster Nodes

This chapter provides procedures to install Oracle Solaris Cluster 3.3 3/13 software on global-cluster voting nodes and optionally on the administrative console.

The following procedures are in this chapter:

- "Installing the Software" on page 49

Installing the Software

This section provides information and procedures to install software on the cluster nodes.

The following task map lists the tasks that you perform to install software on multiple-host or single-host global clusters. Complete the procedures in the order that is indicated.

TABLE 2–1 Task Map: Installing the Software

Task	Instructions
Plan the layout of your cluster configuration and prepare to install software.	"How to Prepare for Cluster Software Installation" on page 50
(*Optional*) Install and configure a quorum server.	"How to Install and Configure Quorum Server Software" on page 51
(*Optional*) Install Cluster Control Panel (CCP) software on the administrative console.	"How to Install Cluster Control Panel Software on an Administrative Console" on page 54
Install the Oracle Solaris OS on all nodes. Optionally, enable Oracle Solaris I/O multipathing.	"How to Install Oracle Solaris Software" on page 57
(*Optional*) Configure internal disk mirroring.	"How to Configure Internal Disk Mirroring" on page 61
(*Optional*) Install Oracle VM Server for SPARC software and create domains.	"SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 62

TABLE 2–1 Task Map: Installing the Software *(Continued)*

Task	Instructions
Install Oracle Solaris Cluster software and any data services that you will use.	"How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63
(*Optional*) Install Sun QFS software.	"How to Install Sun QFS Software" on page 66
Set up directory paths.	"How to Set Up the Root Environment" on page 66
(*Optional*) Configure the IP Filter feature of Oracle Solaris.	"How to Configure IP Filter" on page 67

▼ How to Prepare for Cluster Software Installation

Before you begin to install software, make the following preparations.

1 **Ensure that the combination of hardware and software that you choose for your cluster is currently a supported Oracle Solaris Cluster configuration.**

- See "Cluster Nodes" in *Oracle Solaris Cluster Concepts Guide* for information about physical and virtual machines that are supported as cluster nodes.
- Contact your Oracle sales representative for the most current information about supported cluster configurations.

2 **Read the following manuals for information that can help you plan your cluster configuration and prepare your installation strategy.**

- *Oracle Solaris Cluster 3.3 3/13 Release Notes* - Restrictions, bug workarounds, and other late-breaking information.
- *Oracle Solaris Cluster Concepts Guide* - Overview of the Oracle Solaris Cluster product.
- *Oracle Solaris Cluster Software Installation Guide* (this manual) - Planning guidelines and procedures for installing and configuring Oracle Solaris, Oracle Solaris Cluster, and volume-manager software.
- *Oracle Solaris Cluster Data Services Planning and Administration Guide* - Planning guidelines and procedures to install and configure data services.

3 **Have available all related documentation, including third-party documents.**

The following is a partial list of products whose documentation you might need to reference during cluster installation:

- Oracle Solaris OS
- Solaris Volume Manager software
- Sun QFS software
- Third-party applications

4 **Plan your cluster configuration.**

Use the planning guidelines in Chapter 1, "Planning the Oracle Solaris Cluster Configuration," and in the *Oracle Solaris Cluster Data Services Planning and Administration Guide* to determine how to install and configure your cluster.

Caution – Plan your cluster installation completely. Identify requirements for all data services and third-party products **before** you begin Oracle Solaris and Oracle Solaris Cluster software installation. Failure to do so might result in installation errors that require that you completely reinstall the Oracle Solaris and Oracle Solaris Cluster software. You must accommodate these requirements before you install Oracle Solaris Cluster software because you cannot change hostnames after you install Oracle Solaris Cluster software.

5 **Obtain all necessary patches for your cluster configuration.**

See "Patches and Required Firmware Levels" in *Oracle Solaris Cluster 3.3 3/13 Release Notes* for the location of patches and installation instructions.

Next Steps If you want to install a machine as a quorum server to use as the quorum device in your cluster, go next to "How to Install and Configure Quorum Server Software" on page 51.

Otherwise, if you want to use Cluster Control Panel software to connect from an administrative console to your cluster nodes, go to "How to Install Cluster Control Panel Software on an Administrative Console" on page 54.

Otherwise, choose the Oracle Solaris installation procedure to use.

- To configure Oracle Solaris Cluster software by using the `scinstall(1M)` utility, go to "How to Install Oracle Solaris Software" on page 57 to first install Oracle Solaris software.
- To install and configure both Oracle Solaris and Oracle Solaris Cluster software in the same operation (JumpStart method), go to "How to Install Oracle Solaris and Oracle Solaris Cluster Software (JumpStart)" on page 90.

▼ How to Install and Configure Quorum Server Software

Perform this procedure to configure a host server as a quorum server.

Before You Begin Perform the following tasks:

- Ensure that the machine that you choose for the quorum server has at least 1 Mbyte of disk space available for Oracle Java Web Console software installation.
- Ensure that the quorum-server machine is connected to a public network that is accessible to the cluster nodes.
- Disable the spanning tree algorithm on the Ethernet switches for the ports that are connected to the cluster public network where the quorum server will run.

1. **Become superuser on the machine to install.**

   Use the following command if you want to ensure that the `installer` program can display the GUI.

   ```
   # ssh -X [-l root] quorumserver
   ```

2. **Load the installation media into the drive.**

   If the volume management daemon (`vold`(1M)) is running and is configured to manage CD-ROM or DVD devices, the daemon automatically mounts the media on the `/cdrom/cdrom0` directory.

3. **Change to the installation wizard directory of the media.**

   - If you are installing the software packages on the SPARC platform, type the following command:

     ```
     phys-schost# cd /cdrom/cdrom0/Solaris_sparc
     ```

   - If you are installing the software packages on the x86 platform, type the following command:

     ```
     phys-schost# cd /cdrom/cdrom0/Solaris_x86
     ```

4. **Start the installation wizard.**

   ```
   phys-schost# ./installer
   ```

5. **Follow instructions on the screen to install Quorum Server software on the host server.**

   Choose the Configure Later option.

   Note – If the installer does not allow you to choose the Configure Later option, choose Configure Now.

   After installation is finished, you can view any available installation log. See the *Sun Java Enterprise System 7 Installation and Upgrade Guide* for additional information about using the `installer` program.

6. **Apply any required Quorum Server patches.**

7. **Unload the installation media from the drive.**

   a. **To ensure that the installation media is not being used, change to a directory that does *not* reside on the media.**

   b. **Eject the media.**

      ```
      phys-schost# eject cdrom
      ```

8 **Apply any necessary patches to support the Quorum Server software.**

See "Patches and Required Firmware Levels" in *Oracle Solaris Cluster 3.3 3/13 Release Notes* for the location of patches and installation instructions.

9 **(Optional) Add the Quorum Server binary location to your `PATH` environment variable.**

```
quorumserver# PATH=$PATH:/usr/cluster/bin
```

10 **(Optional) Add the Quorum Server man-page location to your `MANPATH` environment variable.**

```
quorumserver# MANPATH=$MANPATH:/usr/cluster/man
```

11 **Configure the quorum server.**

Add the following entry to the `/etc/scqsd/scqsd.conf` file to specify configuration information about the quorum server.

Identify the quorum server by using at least one of either an instance name or a port number. You must provide the port number, but the instance name is optional.

- If you provide an instance name, that name must be unique among your quorum servers.
- If you do not provide an instance name, always refer to this quorum server by the port on which it listens.

`/usr/cluster/lib/sc/scqsd [-d` *quorumdirectory*`] [-i` *instancename*`] -p` *port*

`-d` *quorumdirectory*
: The path to the directory where the quorum server can store quorum data.

  The quorum-server process creates one file per cluster in this directory to store cluster-specific quorum information.

  By default, the value of this option is `/var/scqsd`. This directory must be unique for each quorum server that you configure.

`-i` *instancename*
: A unique name that you choose for the quorum-server instance.

`-p` *port*
: The port number on which the quorum server listens for requests from the cluster.

12 **(Optional) To serve more than one cluster but use a different port number or instance, configure an additional entry for each additional instance of the quorum server that you need.**

13 **Save and close the `/etc/scqsd/scqsd.conf` file.**

14 **Start the newly configured quorum server.**

```
quorumserver# /usr/cluster/bin/clquorumserver start quorumserver
```

quorumserver
Identifies the quorum server. You can use the port number on which the quorum server listens. If you provided an instance name in the configuration file, you can use that name instead.

- To start a single quorum server, provide either the instance name or the port number.
- To start all quorum servers when you have multiple quorum servers configured, use the + operand.

Troubleshooting The installer performs a simple `pkgadd` installation of the Quorum Server packages and sets up the necessary directories. The software consists of the following packages:

- `SUNWscqsr`
- `SUNWscqsu`
- `SUNWscqsman`

The installation of these packages adds software to the `/usr/cluster` and `/etc/scqsd` directories. You cannot modify the location of the Quorum Server software.

If you receive an installation error message regarding the Quorum Server software, verify that the packages were properly installed.

Next Steps If you want to use an administrative console to communicate with the cluster nodes, go to "How to Install Cluster Control Panel Software on an Administrative Console" on page 54.

Otherwise, go to "How to Install Oracle Solaris Software" on page 57.

▼ How to Install Cluster Control Panel Software on an Administrative Console

Note – You are not required to use an administrative console. If you do not use an administrative console, perform administrative tasks from one designated node in the cluster.

You cannot use this software to connect to Oracle VM Server for SPARC guest domains.

This procedure describes how to install the Cluster Control Panel (CCP) software on an administrative console. The CCP provides a single interface from which to start the `cconsole`, `cssh`, `ctelnet`, and `crlogin` tools. Each of these tools provides a multiple-window connection to a set of nodes, as well as a common window. You can use the common window to send input to all nodes at one time. For additional information, see the `ccp(1M)` man page.

You can use any desktop machine that runs a version of the Oracle Solaris OS that is supported by Oracle Solaris Cluster 3.3 3/13 software as an administrative console.

Before You Begin Ensure that a supported version of the Oracle Solaris OS and any Oracle Solaris patches are installed on the administrative console. All platforms require at least the End User Oracle Solaris Software Group.

1 **Become superuser on the administrative console.**

2 **Load the DVD-ROM into the DVD-ROM drive.**

If the volume management daemon `vold(1M)` is running and is configured to manage CD-ROM or DVD devices, the daemon automatically mounts the media on the `/cdrom/cdrom0` directory.

3 **Change to the `Solaris_`*arch*`/Product/sun_cluster/Solaris_`*ver*`/Packages/` directory, where *arch* is `sparc` or `x86`, and where *ver* is `10` for Oracle Solaris 10 .**

```
adminconsole# cd /cdrom/cdrom0/Solaris_arch/Product/sun_cluster/Solaris_ver/Packages/
```

4 **Install the `SUNWccon` package.**

```
adminconsole# pkgadd -d . SUNWccon
```

5 **(Optional) Install Oracle Solaris Cluster man-page packages.**

```
adminconsole# pkgadd -d . pkgname ...
```

Package Name	Description
`SUNWscman`	Oracle Solaris Cluster framework man pages
`SUNWscdsman`	Oracle Solaris Cluster data-service man pages
`SUNWscqsman`	Quorum Server man pages

When you install the Oracle Solaris Cluster man-page packages on the administrative console, you can view them from the administrative console before you install Oracle Solaris Cluster software on the cluster nodes or quorum server.

6 **Unload the DVD-ROM from the DVD-ROM drive.**

a. **To ensure that the DVD-ROM is not being used, change to a directory that does *not* reside on the DVD-ROM.**

b. **Eject the DVD-ROM.**

```
adminconsole# eject cdrom
```

7 **Create an `/etc/clusters` file on the administrative console.**

Add your cluster name and the physical node name of each cluster node to the file.

```
adminconsole# vi /etc/clusters
clustername  node1  node2
```

See the /opt/SUNWcluster/bin/clusters(4) man page for details.

8 Create an /etc/serialports file.

Add an entry for each node in the cluster to the file. Specify the physical node name, the hostname of the console-access device, and the port number. Examples of a console-access device are a terminal concentrator (TC) and a Sun Fire system controller.

```
adminconsole# vi /etc/serialports
node1  ca-dev-hostname  port
node2  ca-dev-hostname  port
```

node1, node2
Physical names of the cluster nodes.

ca-dev-hostname
Hostname of the console-access device.

port
Serial port number, or the Secure Shell port number for Secure Shell connections.

Note these special instructions to create an /etc/serialports file:

- For a Sun Fire 15000 system controller, use telnet(1) port number 23 for the serial port number of each entry.
- For all other console-access devices, to connect to the console through a telnet connection, use the telnet serial port number, not the physical port number. To determine the telnet serial port number, add 5000 to the physical port number. For example, if a physical port number is 6, the telnet serial port number is 5006.
- For Secure Shell connections to node consoles, specify for each node the name of the console-access device and the port number to use for secure connection. The default port number for Secure Shell is 22.
- To connect the administrative console directly to the cluster nodes or through a management network, specify for each node its hostname and the port number that the node uses to connect to the administrative console or the management network.

9 (Optional) For convenience, set the directory paths on the administrative console.

a. Add the /opt/SUNWcluster/bin/ directory to the PATH.

b. Add the /opt/SUNWcluster/man/ directory to the MANPATH.

c. If you installed the SUNWscman package, also add the /usr/cluster/man/ directory to the MANPATH.

10 Start the CCP utility.

```
adminconsole# /opt/SUNWcluster/bin/ccp &
```

Click the `cconsole`, `cssh`, `crlogin`, or `ctelnet` button in the CCP window to launch that tool. Alternately, you can start any of these tools directly. For example, to start `ctelnet`, type the following command:

```
adminconsole# /opt/SUNWcluster/bin/ctelnet &
```

The CCP software supports the following Secure Shell connections:

- For secure connection to the node consoles, start the `cconsole` tool. Then from the Options menu of the Cluster Console window, enable the Use SSH check box.
- For secure connection to the cluster nodes, use the `cssh` tool.

See the procedure "How to Log Into the Cluster Remotely" in *Oracle Solaris Cluster System Administration Guide* for additional information about how to use the CCP utility. Also see the `ccp`(1M) man page.

Next Steps Determine whether the Oracle Solaris OS is already installed to meet Oracle Solaris Cluster software requirements. See "Planning the Oracle Solaris OS" on page 12 for information about Oracle Solaris Cluster installation requirements for the Oracle Solaris OS.

- If the Oracle Solaris OS meets Oracle Solaris Cluster requirements, go to "How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63.
- If the Oracle Solaris OS does not meet Oracle Solaris Cluster requirements, install, reconfigure, or reinstall the Oracle Solaris OS as needed.
  - To install the Oracle Solaris OS alone, go to "How to Install Oracle Solaris Software" on page 57.
  - To use the `scinstall` custom JumpStart method to install both the Oracle Solaris OS and Oracle Solaris Cluster software, go to "How to Install Oracle Solaris and Oracle Solaris Cluster Software (JumpStart)" on page 90

▼ How to Install Oracle Solaris Software

If you do not use the `scinstall` custom JumpStart installation method to install software, perform this procedure to install the Oracle Solaris OS on each node in the global cluster. See "How to Install Oracle Solaris and Oracle Solaris Cluster Software (JumpStart)" on page 90 for more information about JumpStart installation of a cluster.

Tip – To speed installation, you can install the Oracle Solaris OS on each node at the same time.

If your nodes are already installed with the Oracle Solaris OS but do not meet Oracle Solaris Cluster installation requirements, you might need to reinstall the Oracle Solaris software. Follow the steps in this procedure to ensure subsequent successful installation of Oracle Solaris

Cluster software. See "Planning the Oracle Solaris OS" on page 12 for information about required root-disk partitioning and other Oracle Solaris Cluster installation requirements.

Before You Begin Perform the following tasks:

- Ensure that the hardware setup is complete and that connections are verified before you install Oracle Solaris software. See the *Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual* and your server and storage device documentation for details.
- Ensure that your cluster configuration planning is complete. See "How to Prepare for Cluster Software Installation" on page 50 for requirements and guidelines.
- If you use a naming service, add address-to-name mappings for all public hostnames and logical addresses to any naming services that clients use for access to cluster services. See "Public-Network IP Addresses" on page 22 for planning guidelines. See your Oracle Solaris system-administrator documentation for information about using Oracle Solaris naming services.

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

- **If Cluster Control Panel (CCP) software is installed and configured on your administrative console, use the `cconsole(1M)` utility to display the individual console screens.**

  As superuser, use the following command to start the `cconsole` utility:

  ```
  adminconsole# /opt/SUNWcluster/bin/cconsole clustername &
  ```

  The `cconsole` 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 `cconsole` utility, connect to the consoles of each node individually.**

2 Install the Oracle Solaris OS as instructed in your Oracle Solaris installation documentation.

Note – You must install all nodes in a cluster with the same version of the Oracle Solaris OS.

You can use any method that is normally used to install Oracle Solaris software. During Oracle Solaris software installation, perform the following steps:

a. Install at least the End User Oracle Solaris Software Group.

Tip – To avoid the need to manually install Oracle Solaris software packages, install the Entire Oracle Solaris Software Group Plus OEM Support.

See "Oracle Solaris Software Group Considerations" on page 14 for information about additional Oracle Solaris software requirements.

b. Choose Manual Layout to set up the file systems.

- **Specify that slice 7 is at least 20 Mbytes in size.**
- **(Optional) Create a file system of at least 512 Mbytes for use by the global-device subsystem.**

  Note – Do not create this file system if you plan to use a `lofi` device, which is the default. You specify the use of a `lofi` device to the `scinstall` command when you establish the cluster.

- **Create any other file-system partitions that you need, as described in "System Disk Partitions" on page 14.**

c. For ease of administration, set the same root password on each node.

3 If you will use role-based access control (RBAC) instead of superuser to access the cluster nodes, set up an RBAC role that provides authorization for all Oracle Solaris Cluster commands.

This series of installation procedures requires the following Oracle Solaris Cluster RBAC authorizations if the user is not superuser:

- `solaris.cluster.modify`
- `solaris.cluster.admin`
- `solaris.cluster.read`

See "Role-Based Access Control (Overview)" in *System Administration Guide: Security Services* for more information about using RBAC roles. See the Oracle Solaris Cluster man pages for the RBAC authorization that each Oracle Solaris Cluster subcommand requires.

4 If you are adding a node to an existing cluster, add mount points for cluster file systems to the new node.

a. From the active cluster node, display the names of all cluster file systems.

```
phys-schost-1# mount | grep global | egrep -v node@ | awk '{print $1}'
```

b. On the new node, create a mount point for each cluster file system in the cluster.

```
phys-schost-new# mkdir -p mountpoint
```

For example, if the mount command returned the file-system name `/global/dg-schost-1`, run `mkdir -p /global/dg-schost-1` on the new node you are adding to the cluster.

5 If you installed the End User Oracle Solaris Software Group and you want to use any of the following Oracle Solaris Cluster features, install additional Oracle Solaris software packages to support these features.

Feature	Mandatory Oracle Solaris Software Packages
`scsnapshot`	`SUNWp15u SUNWp15v SUNWp15p`
Oracle Solaris Cluster Manager	`SUNWapchr SUNWapchu`

```
phys-schost# pkgadd -G -d . package ...
```

You must add these packages only to the global zone. The `-G` option adds packages to the current zone only. This option also specifies that the packages are *not* propagated to any existing non-global zone or to any non-global zone that is created later.

6 Install any required Oracle Solaris OS patches and hardware-related firmware and patches.

Include those patches for storage-array support. Also download any needed firmware that is contained in the hardware patches.

See "Patches and Required Firmware Levels" in *Oracle Solaris Cluster 3.3 3/13 Release Notes* for the location of patches and installation instructions.

7 x86: Set the default boot file.

The setting of this value enables you to reboot the node if you are unable to access a login prompt.

```
grub edit> kernel /platform/i86pc/multiboot kmdb
```

8 Update the `/etc/inet/hosts` file on each node with all public IP addresses that are used in the cluster.

Perform this step regardless of whether you are using a naming service.

Note – During establishment of a new cluster or new cluster node, the `scinstall` utility automatically adds the public IP address of each node that is being configured to the `/etc/inet/hosts` file.

---

9 (Optional) Configure public-network adapters in IPMP groups.

If you do not want to use the multiple-adapter IPMP groups that the `scinstall` utility configures during cluster creation, configure custom IPMP groups as you would in a stand-alone system. See Chapter 28, "Administering IPMP (Tasks)," in *Oracle Solaris Administration: IP Services* for details.

During cluster creation, the `scinstall` utility configures each set of public-network adapters that use the same subnet and are not already configured in an IPMP group into a single multiple-adapter IPMP group. The `scinstall` utility ignores any existing IPMP groups.

10 If you want to use Oracle Solaris I/O multipathing, enable multipathing on each node.

Caution – If Oracle Solaris Cluster software is already installed, do not issue this command. Running the `stmsboot` command on an active cluster node might cause Oracle Solaris services to go into the maintenance state. Instead, follow instructions in the `stmsboot(1M)` man page for using the `stmsboot` command in an Oracle Solaris Cluster environment.

```
phys-schost# /usr/sbin/stmsboot -e
```

`-e`
Enables Oracle Solaris I/O multipathing.

See the `stmsboot(1M)` man page for more information.

Next Steps If your server supports the mirroring of internal hard drives and you want to configure internal disk mirroring, go to "How to Configure Internal Disk Mirroring" on page 61.

Otherwise, install the Oracle Solaris Cluster software packages. Go to "How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63.

See Also See the *Oracle Solaris Cluster System Administration Guide* for procedures to perform dynamic reconfiguration tasks in an Oracle Solaris Cluster configuration.

▼ How to Configure Internal Disk Mirroring

Perform this procedure on each node of the global cluster to configure internal hardware RAID disk mirroring to mirror the system disk. This procedure is optional.

Note – Do not perform this procedure under either of the following circumstances:

- Your servers do not support the mirroring of internal hard drives.
- You have already established the cluster. Instead, perform "Mirroring Internal Disks on Servers that Use Internal Hardware Disk Mirroring or Integrated Mirroring" in *Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual*.

Before You Begin Ensure that the Oracle Solaris operating system and any necessary patches are installed.

1 Become superuser.

2 Configure an internal mirror.

```
phys-schost# raidctl -c clt0d0 clt1d0
```

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

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

Next Steps SPARC: To create Oracle VM Server for SPARC, go to "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 62.

Otherwise, install the Oracle Solaris Cluster software packages. Go to "How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63.

▼ SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains

Perform this procedure to install Oracle VM Server for SPARC software on a physically clustered machine and to create I/O and guest domains.

Before You Begin Perform the following tasks:

- Ensure that the machine is SPARC hypervisor capable.
- Have available *Logical Domains (LDoms) 1.0.3 Administration Guide* and *Logical Domains (LDoms) 1.0.3 Release Notes*.
- Read the requirements and guidelines in "SPARC: Guidelines for Oracle VM Server for SPARC in a Cluster" on page 19.

1 **Become superuser on the machine.**

2 **Install Oracle VM Server for SPARC software and configure domains.**

- **Follow the procedures in "Installing and Enabling Software" in *Logical Domains (LDoms) 1.0.3 Administration Guide*.**

  If you create guest domains, adhere to the Oracle Solaris Cluster guidelines for creating guest domains in a cluster.

- **Use the `mode=sc` option for all virtual switch devices that connect the virtual network devices that are used as the cluster interconnect.**

- **For shared storage, map only the full SCSI disks into the guest domains.**

Next Steps If your server supports the mirroring of internal hard drives and you want to configure internal disk mirroring, go to "How to Configure Internal Disk Mirroring" on page 61.

Otherwise, install the Oracle Solaris Cluster software packages. Go to "How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63.

How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages

Follow this procedure to use the `installer` program to perform one or more of the following installation tasks:

- To install the Oracle Solaris Cluster framework software packages on each node in the global cluster. These nodes can be physical machines or (SPARC only) Oracle VM Server for SPARC I/O domains or guest domains, or a combination of any of these types of nodes.

  Note – If your physically clustered machines are configured with Oracle VM Server for SPARC, install Oracle Solaris Cluster software only in I/O domains or guest domains.

- To install Oracle Solaris Cluster framework software on the master node where you will create a flash archive for a JumpStart installation. See "How to Install Oracle Solaris and Oracle Solaris Cluster Software (JumpStart)" on page 90 for more information about a JumpStart installation of a global cluster.
- To install data services.

  Note – This procedure installs data services only to the global zone. To install data services to be visible only from within a certain non-global zone, see "How to Create a Non-Global Zone on a Global-Cluster Node" on page 191.

Note – This procedure uses the interactive form of the `installer` program. To use the noninteractive form of the `installer` program, such as when developing installation scripts, see Chapter 5, "Installing in Silent Mode," in *Sun Java Enterprise System 5 Update 1 Installation Guide for UNIX*.

Before You Begin Perform the following tasks:

- Ensure that the Oracle Solaris OS is installed to support Oracle Solaris Cluster software.

  If Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for Oracle Solaris Cluster software and any other software that you intend to install on the cluster. See "How to Install Oracle Solaris Software" on page 57 for more information about installing Oracle Solaris software to meet Oracle Solaris Cluster software requirements.
- Have available the DVD-ROM.

1 Restore external access to RPC communication and optionally to Oracle Java Web Console.

During the installation of the Oracle Solaris OS, a restricted network profile is used that disables external access for certain network services. The restricted services include the following services that affect cluster functionality:

- The RPC communication service, which is required for cluster communication
- The Oracle Java Web Console service, which is required to use the Oracle Solaris Cluster Manager GUI

The following steps restore Oracle Solaris functionality that is used by the Oracle Solaris Cluster framework but which is prevented if a restricted network profile is used.

a. Perform the following commands to restore external access to RPC communication.

```
phys-schost# svccfg
svc:> select network/rpc/bind
svc:/network/rpc/bind> setprop config/local_only=false
svc:/network/rpc/bind> quit
phys-schost# svcadm refresh network/rpc/bind:default
phys-schost# svcprop network/rpc/bind:default | grep local_only
```

The output of the last command should show that the `local_only` property is now set to false.

b. (Optional) Perform the following commands to restore external access to Oracle Java Web Console.

```
phys-schost# svccfg
svc:> select system/webconsole
svc:/system/webconsole> setprop options/tcp_listen=true
svc:/system/webconsole> quit
phys-schost# /usr/sbin/smcwebserver restart
phys-schost# netstat -a | grep 6789
```

The output of the last command should return an entry for 6789, which is the port number that is used to connect to Oracle Java Web Console.

For more information about what services the restricted network profile restricts to local connections, see "Planning Network Security" in *Oracle Solaris 10 1/13 Installation Guide: Planning for Installation and Upgrade*.

2 Become superuser on the machine to install.

Use the following command if you want to ensure that the `installer` program can display the GUI.

```
# ssh -X [-l root] nodename
```

3 Load the DVD-ROM into the DVD-ROM drive.

If the volume management daemon `vold(1M)` is running and is configured to manage CD-ROM or DVD devices, the daemon automatically mounts the media on the `/cdrom/cdrom0` directory.

4 **Change to the installation wizard directory of the DVD-ROM.**

- If you are installing the software packages on the SPARC platform, type the following command:

```
phys-schost# cd /cdrom/cdrom0/Solaris_sparc
```

- If you are installing the software packages on the x86 platform, type the following command:

```
phys-schost# cd /cdrom/cdrom0/Solaris_x86
```

5 **Start the installation wizard program.**

```
phys-schost# ./installer
```

See the *Sun Java Enterprise System 7 Installation and Upgrade Guide* for additional information about using the different forms and features of the `installer` program.

6 **Follow instructions on the screen to install Oracle Solaris Cluster framework software and data services on the node.**

- If you do not want to install Oracle Solaris Cluster Manager, formerly SunPlex Manager, deselect it.

Note – You must install Oracle Solaris Cluster Manager either on all nodes of the cluster or on none.

- If you want to install Oracle Solaris Cluster Geographic Edition software, select it.

  After the cluster is established, see *Oracle Solaris Cluster Geographic Edition Installation Guide* for further installation procedures.
- Choose Configure Later when prompted whether to configure Oracle Solaris Cluster framework software.

After installation is finished, you can view any available installation log.

7 **Unload the DVD-ROM from the DVD-ROM drive.**

a. **To ensure that the DVD-ROM is not being used, change to a directory that does *not* reside on the DVD-ROM.**

b. **Eject the DVD-ROM.**

```
phys-schost# eject cdrom
```

8 **Apply any necessary patches to support Oracle Solaris Cluster software.**

See "Patches and Required Firmware Levels" in *Oracle Solaris Cluster 3.3 3/13 Release Notes* for the location of patches and installation instructions.

9 **If you will use any of the following adapters for the cluster interconnect, uncomment the relevant entry in the `/etc/system` file on each node.**

Adapter	Entry
ipge	set ipge:ipge_taskq_disable=1
ixge	set ixge:ixge_taskq_disable=1

This entry becomes effective after the next system reboot.

Next Steps If you want to install Sun QFS file system software, follow the procedures for initial installation. See "How to Install Sun QFS Software" on page 66.

Otherwise, to set up the root user environment, go to "How to Set Up the Root Environment" on page 66.

▼ How to Install Sun QFS Software

Perform this procedure on each node in the global cluster.

1 **Ensure that Oracle Solaris Cluster software is installed.**

See "How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63.

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

3 **Install Sun QFS file system software.**

Follow procedures for initial installation in your Sun QFS documentation.

Next Steps Set up the root user environment. Go to "How to Set Up the Root Environment" on page 66.

▼ How to Set Up the Root Environment

Note – In an Oracle Solaris Cluster configuration, user initialization files for the various shells must verify that they are run from an interactive shell. The files must verify this before they attempt to output to the terminal. Otherwise, unexpected behavior or interference with data services might occur. See "Customizing a User's Work Environment" in *Oracle Solaris Administration: Basic Administration* for more information.

Perform this procedure on each node in the global cluster.

1. **Become superuser on a cluster node.**
2. **Modify `PATH` and `MANPATH` entries in the `.cshrc` or `.profile` file.**
   a. **Add `/usr/sbin/` and `/usr/cluster/bin/` to the `PATH`.**
   b. **Add `/usr/cluster/man/` to the `MANPATH`.**

   See your Oracle Solaris OS documentation, volume manager documentation, and other application documentation for additional file paths to set.
3. **(Optional) For ease of administration, set the same root password on each node, if you have not already done so.**

Next Steps If you want to use the IP Filter feature of Oracle Solaris, go to "How to Configure IP Filter" on page 67.

Otherwise, configure Oracle Solaris Cluster software on the cluster nodes. Go to "Establishing a New Global Cluster or New Global-Cluster Node" on page 72.

▼ How to Configure IP Filter

Perform this procedure to configure the IP Filter feature of Oracle Solaris on the global cluster.

Note – Only use IP Filter with failover data services. The use of IP Filter with scalable data services is not supported.

For more information about the IP Filter feature, see Part IV, "IP Security," in *Oracle Solaris Administration: IP Services*.

Before You Begin Read the guidelines and restrictions to follow when you configure IP Filter in a cluster. See the "IP Filter" bullet item in "Oracle Solaris OS Feature Restrictions" on page 13.

1. **Become superuser.**
2. **Add filter rules to the `/etc/ipf/ipf.conf` file on all affected nodes.**

   Observe the following guidelines and requirements when you add filter rules to Oracle Solaris Cluster nodes.

- In the `ipf.conf` file on each node, add rules to explicitly allow cluster interconnect traffic to pass unfiltered. Rules that are not interface specific are applied to all interfaces, including cluster interconnects. Ensure that traffic on these interfaces is not blocked mistakenly. If interconnect traffic is blocked, the IP Filter configuration interferes with cluster handshakes and infrastructure operations.

  For example, suppose the following rules are currently used:

  ```
  # Default block TCP/UDP unless some later rule overrides
  block return-rst in proto tcp/udp from any to any

  # Default block ping unless some later rule overrides
  block return-rst in proto icmp all
  ```

  To unblock cluster interconnect traffic, add the following rules. The subnets used are for example only. Derive the subnets to use by using the `ifconfig` *interface* command.

  ```
  # Unblock cluster traffic on 172.16.0.128/25 subnet (physical interconnect)
  pass in quick proto tcp/udp from 172.16.0.128/25 to any
  pass out quick proto tcp/udp from 172.16.0.128/25 to any

  # Unblock cluster traffic on 172.16.1.0/25 subnet (physical interconnect)
  pass in quick proto tcp/udp from 172.16.1.0/25 to any
  pass out quick proto tcp/udp from 172.16.1.0/25 to any

  # Unblock cluster traffic on 172.16.4.0/23 (clprivnet0 subnet)
  pass in quick proto tcp/udp from 172.16.4.0/23 to any
  pass out quick proto tcp/udp from 172.16.4.0/23 to any
  ```

- You can specify either the adapter name or the IP address for a cluster private network. For example, the following rule specifies a cluster private network by its adapter's name:

  ```
  # Allow all traffic on cluster private networks.
  pass in quick on e1000g1 all
  ...
  ```

- Oracle Solaris Cluster software fails over network addresses from node to node. No special procedure or code is needed at the time of failover.
- All filtering rules that reference IP addresses of logical hostname and shared address resources must be identical on all cluster nodes.
- Rules on a standby node will reference a nonexistent IP address. This rule is still part of the IP filter's active rule set and will become effective when the node receives the address after a failover.
- All filtering rules must be the same for all NICs in the same IPMP group. In other words, if a rule is interface-specific, the same rule must also exist for all other interfaces in the same IPMP group.

For more information about IP Filter rules, see the `ipf(4)` man page.

3 Enable the `ipfilter` SMF service.

```
phys-schost# svcadm enable /network/ipfilter:default
```

Next Steps Configure Oracle Solaris Cluster software on the cluster nodes. Go to "Establishing a New Global Cluster or New Global-Cluster Node" on page 72.

◆ ◆ ◆ CHAPTER 3

Establishing the Global Cluster

This chapter provides procedures for how to establish a global cluster or a new global-cluster node.

Note – To create a zone cluster, see "Configuring a Zone Cluster" on page 197. You must establish a global cluster before you can create a zone cluster.

The following procedures are in this chapter:

Establishing a New Global Cluster or New Global-Cluster Node

This section provides information and procedures to establish a new global cluster or to add a node to an existing cluster. Global-cluster nodes can be physical machines, (SPARC only) Oracle VM Server for SPARC I/O domains, or (SPARC only) Oracle VM Server for SPARC guest domains. A cluster can consist of a combination of any of these node types. Before you start to perform these tasks, ensure that you installed software packages for the Oracle Solaris OS, Oracle Solaris Cluster framework, and other products as described in "Installing the Software" on page 49.

The following task maps list the tasks to perform for either a new global cluster or a node added to an existing global cluster. Complete the procedures in the order that is indicated.

- Task Map: Establish a New Global Cluster
- Task Map: Add a Node to an Existing Global Cluster

TABLE 3–1 Task Map: Establish a New Global Cluster

Method	Instructions
Use one of the following methods to establish a new global cluster:	
■ Use the `scinstall` utility to establish the cluster.	"How to Configure Oracle Solaris Cluster Software on All Nodes (`scinstall`)" on page 74
■ Use an XML configuration file to establish the cluster.	"How to Configure Oracle Solaris Cluster Software on All Nodes (XML)" on page 83
■ Set up a JumpStart install server. Then create a flash archive of the installed system. Finally, use the `scinstall` JumpStart option to install the flash archive on each node and establish the cluster.	"How to Install Oracle Solaris and Oracle Solaris Cluster Software (JumpStart)" on page 90
Assign quorum votes and remove the cluster from installation mode, if this operation was not already performed.	"How to Configure Quorum Devices" on page 127
Validate the quorum configuration.	"How to Verify the Quorum Configuration and Installation Mode" on page 132
(Optional) Change a node's private hostname.	"How to Change Private Hostnames" on page 134
Create or modify the NTP configuration file, if not already configured.	"How to Configure Network Time Protocol (NTP)" on page 140
(Optional) Configure IPsec to secure the private interconnect.	"How to Configure IP Security Architecture (IPsec) on the Cluster Private Interconnect" on page 142
If using Solaris Volume Manager, configure the volume management software.	Chapter 4, "Configuring Solaris Volume Manager Software"

TABLE 3–1 Task Map: Establish a New Global Cluster *(Continued)*

Method	Instructions
Create cluster file systems or highly available local file systems as needed.	Chapter 5, "Creating a Cluster File System," or "Enabling Highly Available Local File Systems" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*
Install third-party applications, register resource types, set up resource groups, and configure data services.	*Oracle Solaris Cluster Data Services Planning and Administration Guide* Documentation that is supplied with the application software
Validate the cluster.	"How to Validate the Cluster" on page 144
Take a baseline recording of the finished cluster configuration.	"How to Record Diagnostic Data of the Cluster Configuration" on page 147

TABLE 3–2 Task Map: Add a Node to an Existing Global Cluster

Method	Instructions
Use the `clsetup` command to add the new node to the cluster authorized-nodes list. If necessary, also configure the cluster interconnect and reconfigure the private network address range.	"How to Prepare the Cluster for Additional Global-Cluster Nodes" on page 106
Reconfigure the cluster interconnect and the private network address range as needed to accommodate the added node.	"How to Change the Private Network Configuration When Adding Nodes or Private Networks" on page 108
Use one of the following methods to add a node to an existing global cluster:	
■ Set up a JumpStart install server. Then create a flash archive of the installed system. Finally, use the `scinstall` JumpStart option to install the flash archive on the node you are adding to the cluster.	"How to Install Oracle Solaris and Oracle Solaris Cluster Software (JumpStart)" on page 90
■ Configure Oracle Solaris Cluster software on the new node by using the `scinstall` utility.	"How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (`scinstall`)" on page 114
■ Configure Oracle Solaris Cluster software on the new node by using an XML configuration file.	"How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (XML)" on page 121
Update the quorum configuration information.	"How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 124
Validate the quorum configuration.	"How to Verify the Quorum Configuration and Installation Mode" on page 132
(Optional) Change a node's private hostname.	"How to Change Private Hostnames" on page 134
Modify the NTP configuration.	"How to Configure Network Time Protocol (NTP)" on page 140
If IPsec is configured in the cluster, configure IPsec on the added node.	"How to Configure IP Security Architecture (IPsec) on the Cluster Private Interconnect" on page 142

TABLE 3–2 Task Map: Add a Node to an Existing Global Cluster *(Continued)*

Method	Instructions
If using Solaris Volume Manager, configure the volume management software.	Chapter 4, "Configuring Solaris Volume Manager Software"
Create cluster file systems or highly available local file systems as needed.	Chapter 5, "Creating a Cluster File System," or "Enabling Highly Available Local File Systems" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*
Install third-party applications, register resource types, set up resource groups, and configure data services.	*Oracle Solaris Cluster Data Services Planning and Administration Guide* Documentation that is supplied with the application software
Validate the cluster.	"How to Validate the Cluster" on page 144
Take a baseline recording of the finished cluster configuration.	"How to Record Diagnostic Data of the Cluster Configuration" on page 147

▼ How to Configure Oracle Solaris Cluster Software on All Nodes (`scinstall`)

Perform this procedure from one node of the global cluster to configure Oracle Solaris Cluster software on all nodes of the cluster.

Note – This procedure uses the interactive form of the `scinstall` command. To use the noninteractive forms of the `scinstall` command, such as when developing installation scripts, see the `scinstall(1M)` man page.

Ensure that Oracle Solaris Cluster software packages are installed on the node, either manually or by using the silent-mode form of the `installer` program, before you run the `scinstall` command. For information about running the `installer` program from an installation script, see Chapter 5, "Installing in Silent Mode," in *Sun Java Enterprise System 5 Update 1 Installation Guide for UNIX*.

Before You Begin Perform the following tasks:

- Ensure that the Oracle Solaris OS is installed to support Oracle Solaris Cluster software.

  If Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for Oracle Solaris Cluster software and any other software that you intend to install on the cluster. See "How to Install Oracle Solaris Software" on page 57 for more information about installing Oracle Solaris software to meet Oracle Solaris Cluster software requirements.

- SPARC: If you are configuring Oracle VM Server for SPARC I/O domains or guest domains as cluster nodes, ensure that Oracle VM Server for SPARC software is installed on each physical machine and that the domains meet Oracle Solaris Cluster requirements. See "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 62.
- Ensure that Oracle Solaris Cluster software packages and patches are installed on each node. See "How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63.
- Determine which mode of the `scinstall` utility you will use, Typical or Custom.

  For the Typical installation of Oracle Solaris Cluster software, `scinstall` automatically specifies the following configuration defaults.

Component	Default Value
Private-network address	`172.16.0.0`
Private-network netmask	`255.255.240.0`
Cluster-transport adapters	Exactly two adapters
Cluster-transport switches	`switch1` and `switch2`
Global fencing	Enabled
Global-devices namespace	A `lofi` device
Installation security (DES)	Limited

- Complete one of the following cluster configuration worksheets, depending on whether you run the `scinstall` utility in Typical mode or Custom mode.
  - **Typical Mode Worksheet** – If you will use Typical mode and accept all defaults, complete the following worksheet.

Component	Description/Example	Answer
Cluster Name	What is the name of the cluster that you want to establish?	
Cluster Nodes	List the names of the other cluster nodes planned for the initial cluster configuration.*(For a single-node cluster, press Control-D alone.)*	

Component	Description/Example	Answer	
Cluster Transport Adapters and Cables	What are the names of the two cluster-transport adapters that attach the node to the private interconnect?	*First*	*Second*
	Will this be a dedicated cluster transport adapter? *(Answer No if using tagged VLAN adapters.)*	Yes | No	Yes | No
	If no, what is the VLAN ID for this adapter?		
Quorum Configuration *(two-node cluster only)*	Do you want to disable automatic quorum device selection? *(Answer Yes if any shared storage is not qualified to be a quorum device or if you want to configure a quorum server as a quorum device.)*	Yes | No	
Check	Do you want to interrupt cluster creation for `cluster check` errors?	Yes | No	

- **Custom Mode Worksheet** – If you will use Custom mode and customize the configuration data, complete the following worksheet.

Note – If you are installing a single-node cluster, the `scinstall` utility automatically assigns the default private network address and netmask, even though the cluster does not use a private network.

Component	Description/Example	Answer	
Cluster Name	What is the name of the cluster that you want to establish?		
Cluster Nodes	List the names of the other cluster nodes planned for the initial cluster configuration.*(For a single-node cluster, press Control-D alone.)*		
Authenticating Requests to Add Nodes *(multiple-node cluster only)*	Do you need to use DES authentication?	No | Yes	
Minimum Number of Private Networks *(multiple-node cluster only)*	Should this cluster use at least two private networks?	Yes | No	
Point-to-Point Cables *(multiple-node cluster only)*	If this is a two-node cluster, does this cluster use switches?	Yes | No	
Cluster Switches *(multiple-node cluster only)*	Transport switch name: Defaults: `switch1` and `switch2`	*First*	*Second*

Component	Description/Example	Answer	
Cluster Transport Adapters and Cables *(multiple-node cluster only)*	Node name (*the node from which you run scinstall*):		
	Transport adapter name:	*First*	*Second*
	Will this be a dedicated cluster transport adapter? *(Answer No if using tagged VLAN adapters.)*	Yes | No	Yes | No
	If no, what is the VLAN ID for this adapter?		
	Where does each transport adapter connect to (*a switch or another adapter*)? Switch defaults: `switch1` and `switch2`	*First*	*Second*
	If a transport switch, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		
	Do you want to use autodiscovery to list the available adapters for the other nodes? If no, supply the following information for each additional node:	Yes | No	
Specify for each additional node *(multiple-node cluster only)*	Node name:		
	Transport adapter name:	*First*	*Second*
	Will this be a dedicated cluster transport adapter? *(Answer No if using tagged VLAN adapters.)*	Yes | No	Yes | No
	If no, what is the VLAN ID for this adapter?		
	Where does each transport adapter connect to (*a switch or another adapter*)? Defaults: `switch1` and `switch2`	*First*	*Second*
	If a transport switch, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		

Component	Description/Example	Answer	
Network Address for the Cluster Transport *(multiple-node cluster only)*	Do you want to accept the default network address (`172.16.0.0`)?	Yes | No	
	If no, which private network address do you want to use?	___.___.___.___	
	Do you want to accept the default netmask?	Yes | No	
	If no, what are the maximum numbers of nodes, private networks, and zone clusters that you expect to configure in the cluster?	_____ nodes _____ networks _____ zone clusters	
	Which netmask do you want to use? *(Choose from the values calculated by scinstall or supply your own.)*	___.___.___.___	
Global Fencing	Do you want to turn off global fencing? *(Answer No unless the shared storage does not support SCSI reservations or unless you want systems that are outside the cluster to access the shared storage.)*	Yes | No	Yes | No
Quorum Configuration *(two-node cluster only)*	Do you want to disable automatic quorum device selection? *(Answer Yes if any shared storage is not qualified to be a quorum device or if you want to configure a quorum server as a quorum device.)*	Yes | No	Yes | No
Global Devices File System (*specify for each node*)	Do you want to use the default `lofi` method?	Yes | No	
	If no, do you want to use the default global-devices file system, `/globaldevices`?	Yes | No	
	If no, do you want to select another file system?	Yes | No	
	What is the name of the file system that you want to use?		
Check *(multiple-node cluster only)*	Do you want to interrupt cluster creation for `cluster check` errors?	Yes | No	
(single-node cluster only)	Do you want to run the `cluster check` utility to validate the cluster?	Yes | No	
Automatic Reboot *(single-node cluster only)*	Do you want `scinstall` to automatically reboot the node after installation?	Yes | No	

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. **If you disabled remote configuration during Oracle Solaris Cluster software installation, re-enable remote configuration.**

   Enable remote shell (rsh(1M)) or secure shell (ssh(1)) access for superuser to all cluster nodes.

2. **If you are using switches in the private interconnect of your new cluster, ensure that Neighbor Discovery Protocol (NDP) is disabled.**

   Follow the procedures in the documentation for your switches to determine whether NDP is enabled and to disable NDP.

   During cluster configuration, the software checks that there is no traffic on the private interconnect. If NDP sends any packages to a private adapter when the private interconnect is being checked for traffic, the software will assume that the interconnect is not private and cluster configuration will be interrupted. NDP must therefore be disabled during cluster creation.

   After the cluster is established, you can re-enable NDP on the private-interconnect switches if you want to use that feature.

3. **From one cluster node, start the `scinstall` utility.**

   ```
   phys-schost# /usr/cluster/bin/scinstall
   ```

4. **Type the option number for Create a New Cluster or Add a Cluster Node and press the Return key.**

   ```
    *** Main Menu ***

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

         * 1) Create a new cluster or add a cluster node
         * 2) Configure a cluster to be JumpStarted from this install server
           3) Manage a dual-partition upgrade
           4) Upgrade this cluster node
         * 5) Print release information for this cluster node

         * ?) Help with menu options
         * q) Quit

       Option:  1
   ```

   The New Cluster and Cluster Node Menu is displayed.

5. **Type the option number for Create a New Cluster and press the Return key.**

   The Typical or Custom Mode menu is displayed.

6. **Type the option number for either Typical or Custom and press the Return key.**

   The Create a New Cluster screen is displayed. Read the requirements, then press Control-D to continue.

7 Follow the menu prompts to supply your answers from the configuration planning worksheet.

The `scinstall` utility installs and configures all cluster nodes and reboots the cluster. The cluster is established when all nodes have successfully booted into the cluster. Oracle Solaris Cluster installation output is logged in a `/var/cluster/logs/install/scinstall.log.`*N* file.

8 Verify on each node that multiuser services for the Service Management Facility (SMF) are online.

If services are not yet online for a node, wait until the state changes to online before you proceed to the next step.

```
phys-schost# svcs multi-user-server node
STATE          STIME    FMRI
online         17:52:55 svc:/milestone/multi-user-server:default
```

9 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.

10 (Optional) Enable the automatic node reboot feature.

This feature automatically reboots a node if all monitored shared-disk paths fail, provided that at least one of the disks is accessible from a different node in the cluster.

Note – At initial configuration time, disk-path monitoring is enabled by default for all discovered devices.

---

a. Enable automatic reboot.

```
phys-schost# clnode set -p reboot_on_path_failure=enabled
```

`-p`
Specifies the property to set

`reboot_on_path_failure=enable`
Enables automatic node reboot if failure of all monitored shared-disk paths occurs.

b. Verify that automatic reboot on disk-path failure is enabled.

```
phys-schost# clnode show
=== Cluster Nodes ===

Node Name:                                      node
...
  reboot_on_path_failure:                          enabled
...
```

11 If you intend to use Oracle Solaris Cluster HA for NFS (HA for NFS) on a highly available local file system, ensure that the loopback file system (LOFS) is disabled.

To disable LOFS, add the following entry to the `/etc/system` file on each node of the cluster.

```
exclude:lofs
```

The change to the `/etc/system` file becomes effective after the next system reboot.

Note – You cannot have LOFS enabled if you use HA for NFS on a highly available local file system *and* have `automountd` running. LOFS can cause switchover problems for HA for NFS. If you choose to add HA for NFS on a highly available local file system, you must make one of the following configuration changes.

However, if you configure non-global zones in your cluster, you must enable LOFS on all cluster nodes. If HA for NFS on a highly available local file system must coexist with LOFS, use one of the other solutions instead of disabling LOFS.

- Disable LOFS.
- Disable the `automountd` daemon.
- Exclude from the automounter map all files that are part of the highly available local file system that is exported by HA for NFS. This choice enables you to keep both LOFS and the `automountd` daemon enabled.

See "The Loopback File System" in *System Administration Guide: Devices and File Systems* for more information about loopback file systems.

Example 3–1 Configuring Oracle Solaris Cluster Software on All Nodes

The following example shows the `scinstall` progress messages that are logged as `scinstall` completes configuration tasks on the two-node cluster, `schost`. The cluster is installed from `phys-schost-1` by using the `scinstall` utility in Typical Mode. The other cluster node is `phys-schost-2`. The adapter names are `bge2` and `bge3`. The automatic selection of a quorum device is enabled.

```
  Installation and Configuration

    Log file - /var/cluster/logs/install/scinstall.log.24747

    Configuring global device using lofi on phys-schost-1: done
```

```
    Starting discovery of the cluster transport configuration.
    The Oracle Solaris Cluster software is already installed on "phys-schost-1".
    The Oracle Solaris Cluster software is already installed on "phys-schost-2".
    Starting discovery of the cluster transport configuration.

    The following connections were discovered:

        phys-schost-1:bge2  switch1  phys-schost-2:bge2
        phys-schost-1:bge3  switch2  phys-schost-2:bge3

    Completed discovery of the cluster transport configuration.

    Started cluster check on "phys-schost-1".
    Started cluster check on "phys-schost-2".

    cluster check completed with no errors or warnings for "phys-schost-1".
    cluster check completed with no errors or warnings for "phys-schost-2".

    Removing the downloaded files ... done

    Configuring "phys-schost-2" ... done
    Rebooting "phys-schost-2" ... done

    Configuring "phys-schost-1" ... done
    Rebooting "phys-schost-1" ...

Log file - /var/cluster/logs/install/scinstall.log.24747

Rebooting ...
```

Troubleshooting

Unsuccessful configuration – If one or more nodes cannot join the cluster, or if the wrong configuration information was specified, first attempt to rerun this procedure. If that does not correct the problem, perform the procedure "How to Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems" on page 225 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Oracle Solaris Cluster software packages. Then rerun this procedure.

Next Steps

- If you installed a single-node cluster, cluster establishment is complete. Go to "Creating Cluster File Systems" on page 173 to install volume management software and configure the cluster.
- If you installed a multiple-node cluster and chose automatic quorum configuration, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 132.
- If you installed a multiple-node cluster and declined automatic quorum configuration, perform postinstallation setup. Go to "How to Configure Quorum Devices" on page 127.

If you intend to configure any quorum devices in your cluster, go to "How to Configure Quorum Devices" on page 127.

Otherwise, go to "How to Verify the Quorum Configuration and Installation Mode" on page 132.

▼ How to Configure Oracle Solaris Cluster Software on All Nodes (XML)

Perform this procedure to configure a new global cluster by using an XML cluster configuration file. The new cluster can be a duplication of an existing cluster that runs Oracle Solaris Cluster 3.3 3/13 software.

This procedure configures the following cluster components:

- Cluster name
- Cluster node membership
- Cluster interconnect
- Global devices

Before You Begin Perform the following tasks:

- Ensure that the Oracle Solaris OS is installed to support Oracle Solaris Cluster software.

  If Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for Oracle Solaris Cluster software and any other software that you intend to install on the cluster. See "How to Install Oracle Solaris Software" on page 57 for more information about installing Oracle Solaris software to meet Oracle Solaris Cluster software requirements.

  Ensure that the Oracle Solaris OS is installed to support Oracle Solaris Cluster software.

  If Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for Oracle Solaris Cluster software and any other software that you intend to install on the cluster. See "How to Install Oracle Solaris Software" on page 57 for more information about installing Oracle Solaris software to meet Oracle Solaris Cluster software requirements.
- SPARC: If you are configuring Oracle VM Server for SPARC I/O domains or guest domains as cluster nodes, ensure that Oracle VM Server for SPARC software is installed on each physical machine and that the domains meet Oracle Solaris Cluster requirements. See "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 62.
- Ensure that Oracle Solaris Cluster 3.3 3/13 software and patches are installed on each node that you will configure. See "How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63.

1 Ensure that Oracle Solaris Cluster 3.3 3/13 software is not yet configured on each potential cluster node.

a. Become superuser on a potential node that you want to configure in the new cluster.

b. Determine whether Oracle Solaris Cluster software is already configured on the potential node.

```
phys-schost# /usr/sbin/clinfo -n
```

- **If the command returns the following message, proceed to Step c.**

  ```
  clinfo: node is not configured as part of acluster: Operation not applicable
  ```

  This message indicates that Oracle Solaris Cluster software is not yet configured on the potential node.

- **If the command returns the node ID number, do not perform this procedure.**

  The return of a node ID indicates that Oracle Solaris Cluster software is already configured on the node.

  If the cluster is running an older version of Oracle Solaris Cluster software and you want to install Oracle Solaris Cluster 3.3 3/13 software, instead perform upgrade procedures in *Oracle Solaris Cluster Upgrade Guide*.

c. Repeat Step a and Step b on each remaining potential node that you want to configure in the new cluster.

If Oracle Solaris Cluster software is not yet configured on any of the potential cluster nodes, proceed to Step 2.

2 If you are using switches in the private interconnect of your new cluster, ensure that Neighbor Discovery Protocol (NDP) is disabled.

Follow the procedures in the documentation for your switches to determine whether NDP is enabled and to disable NDP.

During cluster configuration, the software checks that there is no traffic on the private interconnect. If NDP sends any packages to a private adapter when the private interconnect is being checked for traffic, the software will assume that the interconnect is not private and cluster configuration will be interrupted. NDP must therefore be disabled during cluster creation.

After the cluster is established, you can re-enable NDP on the private-interconnect switches if you want to use that feature.

3 If you are duplicating an existing cluster than runs Oracle Solaris Cluster 3.3 3/13 software, use a node in that cluster to create a cluster configuration XML file.

a. Become superuser on an active member of the cluster that you want to duplicate.

b. Export the existing cluster's configuration information to a file.

```
phys-schost# cluster export -o clconfigfile
```

`-o`
Specifies the output destination.

clconfigfile
The name of the cluster configuration XML file. The specified file name can be an existing file or a new file that the command will create.

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

c. Copy the configuration file to the potential node from which you will configure the new cluster.

You can store the file in any directory that is accessible to the other hosts that you will configure as cluster nodes.

4 Become superuser on the potential node from which you will configure the new cluster.

5 Modify the cluster configuration XML file as needed.

a. Open your cluster configuration XML file for editing.

- **If you are duplicating an existing cluster, open the file that you created with the `cluster export` command.**
- **If you are not duplicating an existing cluster, create a new file.**

  Base the file on the element hierarchy that is shown in the `clconfiguration`(5CL) man page. You can store the file in any directory that is accessible to the other hosts that you will configure as cluster nodes.

b. Modify the values of the XML elements to reflect the cluster configuration that you want to create.

- To establish a cluster, the following components must have valid values in the cluster configuration XML file:
  - Cluster name
  - Cluster nodes
  - Cluster transport
- By default, the cluster is created with the global-devices namespace configured on a `lofi` device. If you instead need to use a dedicated file system on which to create the global devices, add the following property to the `<propertyList>` element for each node that will use a partition instead of a `lofi` device.

```
...
  <nodeList>
    <node name="node" id="N">
      <propertyList>
...
```

```
            <property name="globaldevfs" value="/filesystem-name">
...
      </propertyList>
    </node>
...
```

- If you are modifying configuration information that was exported from an existing cluster, some values that you must change to reflect the new cluster, such as node names, are used in the definitions of more than one cluster object.

See the `clconfiguration(5CL)` man page for details about the structure and content of the cluster configuration XML file.

6 Validate the cluster configuration XML file.

```
phys-schost# /usr/share/src/xmllint --valid --noout clconfigfile
```

See the `xmllint`(1) man page for more information.

7 From the potential node that contains the cluster configuration XML file, create the cluster.

```
phys-schost# cluster create -i clconfigfile
```

`-i` *clconfigfile*
Specifies the name of the cluster configuration XML file to use as the input source.

8 Verify on each node that multiuser services for the Service Management Facility (SMF) are online.

If services are not yet online for a node, wait until the state changes to online before you proceed to the next step.

```
phys-schost# svcs multi-user-server node
STATE          STIME    FMRI
online         17:52:55 svc:/milestone/multi-user-server:default
```

9 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.

10 Install any necessary patches to support Oracle Solaris Cluster software, if you have not already done so.

See "Patches and Required Firmware Levels" in *Oracle Solaris Cluster 3.3 3/13 Release Notes* for the location of patches and installation instructions.

11 If you intend to use Oracle Solaris Cluster HA for NFS (HA for NFS) on a highly available local file system, ensure that the loopback file system (LOFS) is disabled.

To disable LOFS, add the following entry to the `/etc/system` file on each node of the cluster.

```
exclude:lofs
```

The change to the `/etc/system` file becomes effective after the next system reboot.

Note – You cannot have LOFS enabled if you use HA for NFS on a highly available local file system *and* have `automountd` running. LOFS can cause switchover problems for HA for NFS. If you choose to add HA for NFS on a highly available local file system, you must make one of the following configuration changes.

However, if you configure non-global zones in your cluster, you must enable LOFS on all cluster nodes. If HA for NFS on a highly available local file system must coexist with LOFS, use one of the other solutions instead of disabling LOFS.

- Disable LOFS.
- Disable the `automountd` daemon.
- Exclude from the automounter map all files that are part of the highly available local file system that is exported by HA for NFS. This choice enables you to keep both LOFS and the `automountd` daemon enabled.

See "The Loopback File System" in *System Administration Guide: Devices and File Systems* for more information about loopback file systems.

12 To duplicate quorum information from an existing cluster, configure the quorum device by using the cluster configuration XML file.

You must configure a quorum device if you created a two-node cluster. If you choose not to use the cluster configuration XML file to create a required quorum device, go instead to "How to Configure Quorum Devices" on page 127.

a. If you are using a quorum server for the quorum device, ensure that the quorum server is set up and running.

Follow instructions in "How to Install and Configure Quorum Server Software" on page 51.

b. If you are using a NAS device for the quorum device, ensure that the NAS device is set up and operational.

i. Observe the requirements for using a NAS device as a quorum device.

See *Oracle Solaris Cluster 3.3 3/13 With Network-Attached Storage Device Manual*.

ii. Follow instructions in your device's documentation to set up the NAS device.

c. Ensure that the quorum configuration information in the cluster configuration XML file reflects valid values for the cluster that you created.

d. If you made changes to the cluster configuration XML file, validate the file.

```
phys-schost# xmllint --valid --noout clconfigfile
```

e. Configure the quorum device.

```
phys-schost# clquorum add -i clconfigfile devicename
```

devicename
Specifies the name of the device to configure as a quorum device.

13 Remove the cluster from installation mode.

```
phys-schost# clquorum reset
```

14 Close access to the cluster configuration by machines that are not configured cluster members.

```
phys-schost# claccess deny-all
```

15 (Optional) Enable automatic node reboot if all monitored shared-disk paths fail.

Note – At initial configuration time, disk-path monitoring is enabled by default for all discovered devices.

---

a. Enable automatic reboot.

```
phys-schost# clnode set -p reboot_on_path_failure=enabled
```

`-p`
Specifies the property to set

`reboot_on_path_failure=enable`
Enables automatic node reboot if failure of all monitored shared-disk paths occurs.

b. Verify that automatic reboot on disk-path failure is enabled.

```
phys-schost# clnode show
=== Cluster Nodes ===

Node Name:                                      node
...
```

```
     reboot_on_path_failure:                          enabled
...
```

Example 3–2 Configuring Oracle Solaris Cluster Software on All Nodes By Using an XML File

The following example duplicates the cluster configuration and quorum configuration of an existing two-node cluster to a new two-node cluster. The new cluster is installed with the Oracle Solaris 10 OS and is not configured with non-global zones. The cluster configuration is exported from the existing cluster node, `phys-oldhost-1`, to the cluster configuration XML file `clusterconf.xml`. The node names of the new cluster are `phys-newhost-1` and `phys-newhost-2`. The device that is configured as a quorum device in the new cluster is `d3`.

The prompt name `phys-newhost-`*N* in this example indicates that the command is performed on both cluster nodes.

```
phys-newhost-N# /usr/sbin/clinfo -n
clinfo: node is not configured as part of acluster: Operation not applicable

phys-oldhost-1# cluster export -o clusterconf.xml
     Copy clusterconf.xml to phys-newhost-1 and modify the file with valid values

phys-newhost-1# xmllint --valid --noout clusterconf.xml
     No errors are reported

phys-newhost-1# cluster create -i clusterconf.xml
phys-newhost-N# svcs multi-user-server
STATE          STIME    FMRI
online         17:52:55 svc:/milestone/multi-user-server:default
phys-newhost-1# clnode status
     Output shows that both nodes are online

phys-newhost-1# clquorum add -i clusterconf.xml d3
phys-newhost-1# clquorum reset
```

Troubleshooting **Unsuccessful configuration** – If one or more nodes cannot join the cluster, or if the wrong configuration information was specified, first attempt to rerun this procedure. If that does not correct the problem, perform the procedure "How to Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems" on page 225 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Oracle Solaris Cluster software packages. Then rerun this procedure.

Next Steps Go to "How to Verify the Quorum Configuration and Installation Mode" on page 132.

See Also After the cluster is fully established, you can duplicate the configuration of the other cluster components from the existing cluster. If you did not already do so, modify the values of the XML elements that you want to duplicate to reflect the cluster configuration you are adding the

component to. For example, if you are duplicating resource groups, ensure that the `<resourcegroupNodeList>` entry contains the valid node names for the new cluster, and not the node names from the cluster that you duplicated unless the node names are the same.

To duplicate a cluster component, run the `export` subcommand of the object-oriented command for the cluster component that you want to duplicate. For more information about the command syntax and options, see the man page for the cluster object that you want to duplicate. The following table lists the cluster components that you can create from a cluster configuration XML file after the cluster is established and the man page for the command that you use to duplicate the component.

Cluster Component	Man Page	Special Instructions
Device groups: Solaris Volume Manager	`cldevicegroup`(1CL)	For Solaris Volume Manager, first create the disk sets that you specify in the cluster configuration XML file.
Resources	`clresource`(1CL)	You can use the `-a` option of the `clresource`, `clressharedaddress`, or `clreslogicalhostname` command to also duplicate the resource type and resource group that are associated with the resource that you duplicate. Otherwise, you must first add the resource type and resource group to the cluster before you add the resource.
Shared address resources	`clressharedaddress`(1CL)	
Logical hostname resources	`clreslogicalhostname`(1CL)	
Resource types	`clresourcetype`(1CL)	
Resource groups	`clresourcegroup`(1CL)	
NAS devices	`clnasdevice`(1CL)	You must first set up the NAS device as described in the device's documentation.
SNMP hosts	`clsnmphost`(1CL)	The `clsnmphost create -i` command requires that you specify a user password file with the `-f` option.
SNMP users	`clsnmpuser`(1CL)	
Thresholds for monitoring system resources on cluster objects	`cltelemetryattribute`(1CL)	

▼ How to Install Oracle Solaris and Oracle Solaris Cluster Software (JumpStart)

This procedure describes how to set up and use the `scinstall`(1M) custom JumpStart installation method. This method installs both Oracle Solaris OS and Oracle Solaris Cluster software on all global-cluster nodes and establishes the cluster. You can also use this procedure to add new nodes to an existing cluster.

Before You Begin Perform the following tasks:

- Ensure that the hardware setup is complete and connections are verified before you install Oracle Solaris software. See the Oracle Solaris Cluster hardware documentation and your server and storage device documentation for details on how to set up the hardware.
- Determine the Ethernet address of each cluster node.
- If you use a naming service, ensure that the following information is added to any naming services that clients use to access cluster services. See "Public-Network IP Addresses" on page 22 for planning guidelines. See your Oracle Solaris system-administrator documentation for information about using Oracle Solaris naming services.
  - Address-to-name mappings for all public hostnames and logical addresses
  - The IP address and hostname of the JumpStart install server
- Ensure that your cluster configuration planning is complete. See "How to Prepare for Cluster Software Installation" on page 50 for requirements and guidelines.
- On the server from which you will create the flash archive, ensure that all Oracle Solaris OS software, patches, and firmware that is necessary to support Oracle Solaris Cluster software is installed.

  If Oracle Solaris software is already installed on the server, you must ensure that the Oracle Solaris installation meets the requirements for Oracle Solaris Cluster software and any other software that you intend to install on the cluster. See "How to Install Oracle Solaris Software" on page 57 for more information about installing Oracle Solaris software to meet Oracle Solaris Cluster software requirements.
- SPARC: If you are configuring Oracle VM Server for SPARC I/O domains or guest domains as cluster nodes, ensure that Oracle VM Server for SPARC software is installed on each physical machine and that the domains meet Oracle Solaris Cluster requirements. See "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 62.
- Ensure that Oracle Solaris Cluster software packages and patches are installed on the server from which you will create the flash archive. See "How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63.
- Determine which mode of the `scinstall` utility you will use, Typical or Custom. For the Typical installation of Oracle Solaris Cluster software, `scinstall` automatically specifies the following configuration defaults.

Component	Default Value
Private-network address	`172.16.0.0`
Private-network netmask	`255.255.240.0`
Cluster-transport adapters	Exactly two adapters
Cluster-transport switches	`switch1` and `switch2`

Component	Default Value
Global fencing	Enabled
Global-devices namespace	A `lofi` device
Installation security (DES)	Limited

- Complete one of the following cluster configuration worksheets, depending on whether you run the `scinstall` utility in Typical mode or Custom mode. See "Planning the Oracle Solaris Cluster Environment" on page 21 for planning guidelines.
  - **Typical Mode Worksheet** – If you will use Typical mode and accept all defaults, complete the following worksheet.

Component	Description/Example	Answer	
JumpStart Directory	What is the name of the JumpStart directory to use?		
Cluster Name	What is the name of the cluster that you want to establish?		
Cluster Nodes	List the names of the cluster nodes that are planned for the initial cluster configuration. *(For a single-node cluster, press Control-D alone.)*		
Cluster Transport Adapters and Cables	First node name:		
	Transport adapter names:	*First*	*Second*
VLAN adapters only	Will this be a dedicated cluster transport adapter? *(Answer No if using tagged VLAN adapters.)*	Yes | No	Yes | No
	If no, what is the VLAN ID for this adapter?		
Specify for each additional node	Node name:		
	Transport adapter names:	*First*	*Second*
Quorum Configuration *(two-node cluster only)*	Do you want to disable automatic quorum device selection? *(Answer Yes if any shared storage is not qualified to be a quorum device or if you want to configure a quorum server as a quorum device.)*	Yes | No	Yes | No

  - **Custom Mode Worksheet** – If you will use Custom mode and customize the configuration data, complete the following worksheet.

Note – If you are installing a single-node cluster, the `scinstall` utility automatically uses the default private network address and netmask, even though the cluster does not use a private network.

Component	Description/Example	Answer	
JumpStart Directory	What is the name of the JumpStart directory to use?		
Cluster Name	What is the name of the cluster that you want to establish?		
Cluster Nodes	List the names of the cluster nodes that are planned for the initial cluster configuration. *(For a single-node cluster, press Control-D alone.)*		
Authenticating Requests to Add Nodes *(multiple-node cluster only)*	Do you need to use DES authentication?	No | Yes	
Network Address for the Cluster Transport *(multiple-node cluster only)*	Do you want to accept the default network address (`172.16.0.0`)?	Yes | No	
	If no, which private network address do you want to use?	___.___.___.___	
	Do you want to accept the default netmask?	Yes | No	
	If no, what are the maximum numbers of nodes, private networks, and zone clusters that you expect to configure in the cluster?	_____ nodes _____ networks _____ zone clusters	
	Which netmask do you want to use? *Choose from the values that are calculated by scinstall or supply your own.*	___.___.___.___	
Minimum Number of Private Networks *(multiple-node cluster only)*	Should this cluster use at least two private networks?	Yes | No	
Point-to-Point Cables *(two-node cluster only)*	Does this cluster use switches?	Yes | No	
Cluster Switches *(multiple-node cluster only)*	Transport switch name, if used: Defaults: `switch1` and `switch2`	*First*	*Second*

Component	Description/Example	Answer	
Cluster Transport Adapters and Cables *(multiple-node cluster only)*	First node name:		
	Transport adapter name:	*First*	*Second*
(VLAN adapters only)	Will this be a dedicated cluster transport adapter? *(Answer No if using tagged VLAN adapters.)*	Yes | No	Yes | No
	If no, what is the VLAN ID for this adapter?		
	Where does each transport adapter connect to (*a switch or another adapter*)? Switch defaults: `switch1` and `switch2`		
	If a transport switch, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		
Specify for each additional node *(multiple-node cluster only)*	Node name:		
	Transport adapter name:	*First*	*Second*
	Where does each transport adapter connect to (*a switch or another adapter*)? Switch defaults: `switch1` and `switch2`		
	If a transport switch, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		
Global Devices File System *Specify for each node*	Do you want to use the default `lofi` method?	Yes | No	
	If no, do you want to use the default global-devices file system, `/globaldevices`?	Yes | No	
	If no, do you want to select another file system?	Yes | No	
	What is the name of the file system that you want to use?		
Global Fencing	Do you want to turn off global fencing? *Answer No unless the shared storage does not support SCSI reservations or unless you want systems that are outside the cluster to access the shared storage.*	Yes | No	Yes | No
Quorum Configuration *(two-node cluster only)*	Do you want to disable automatic quorum device selection? *(Answer Yes if any shared storage is not qualified to be a quorum device or if you want to configure a quorum server as a quorum device.)*	Yes | No	Yes | No

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 Set up your JumpStart install server.

Ensure that the JumpStart install server meets the following requirements.

- The install server is on the same subnet as the cluster nodes, or on the Oracle Solaris boot server for the subnet that the cluster nodes use.
- The install server is not itself a cluster node.
- The install server installs a release of the Oracle Solaris OS that is supported by the Oracle Solaris Cluster software.
- A custom JumpStart directory exists for JumpStart installation of Oracle Solaris Cluster software. This *jumpstart-dir* directory must meet the following requirements:
  - Contain a copy of the `check` utility.
  - Be NFS exported for reading by the JumpStart install server.
- Each new cluster node is configured as a custom JumpStart installation client that uses the custom JumpStart directory that you set up for Oracle Solaris Cluster installation.

Follow the appropriate instructions for your software platform and OS version to set up the JumpStart install server. See "Creating a Profile Server for Networked Systems" in *Oracle Solaris 10 1/13 Installation Guide: JumpStart Installations*.

See also the `setup_install_server(1M)` and `add_install_client(1M)` man pages.

2 If you are installing a new node to an existing cluster, add the node to the list of authorized cluster nodes.

a. Switch to another cluster node that is active and start the `clsetup` utility.

b. Use the `clsetup` utility to add the new node's name to the list of authorized cluster nodes.

For more information, see "How to Add a Node to an Existing Cluster" in *Oracle Solaris Cluster System Administration Guide*.

3 On a cluster node or another machine of the same server platform, install the Oracle Solaris OS and any necessary patches, if you have not already done so.

If Oracle Solaris software is already installed on the server, you must ensure that the Oracle Solaris installation meets the requirements for Oracle Solaris Cluster software and any other software that you intend to install on the cluster. See "How to Install Oracle Solaris Software" on page 57 for more information about installing Oracle Solaris software to meet Oracle Solaris Cluster software requirements.

Follow procedures in "How to Install Oracle Solaris Software" on page 57.

4 **(Optional) SPARC: On the installed system, install Oracle VM Server for SPARC software and create domains, if you have not already done so.**

Follow the procedures in "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 62.

5 **On the installed system, install Oracle Solaris Cluster software and any necessary patches, if you have not already done so.**

Follow procedures in "How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63.

See "Patches and Required Firmware Levels" in *Oracle Solaris Cluster 3.3 3/13 Release Notes* for the location of patches and installation instructions.

6 **Enable the common agent container daemon to start automatically during system boots.**

```
machine# cacaoadm enable
```

7 **On the installed system, update the `/etc/inet/hosts` file with all public IP addresses that are used in the cluster.**

Perform this step regardless of whether you are using a naming service. See "Public-Network IP Addresses" on page 22 for a listing of Oracle Solaris Cluster components whose IP addresses you must add.

8 **On the installed system, reset Oracle Java Web Console to its initial unconfigured state.**

The following command removes configuration information from the web console. Some of this configuration information is specific to the installed system. You must remove this information before you create the flash archive. Otherwise, the configuration information that is transferred to the cluster node might prevent the web console from starting or from interacting correctly with the cluster node.

```
# /usr/share/webconsole/private/bin/wcremove -i console
```

After you install the unconfigured web console on the cluster node and start the web console for the first time, the web console automatically runs its initial configuration and uses information from the cluster node.

For more information about the `wcremove` command, see "Oracle Java Web Console User Identity" in *Oracle Solaris Administration: Basic Administration*.

9 **Create the flash archive of the installed system.**

Follow procedures in Chapter 3, "Creating Flash Archives (Tasks)," in *Oracle Solaris 10 1/13 Installation Guide: Flash Archives (Creation and Installation)*.

```
machine# flarcreate -n name archive
```

`-n name`
Name to give the flash archive.

archive
File name to give the flash archive, with the full path. By convention, the file name ends in `.flar`.

10 Ensure that the flash archive is NFS exported for reading by the JumpStart install server.

See Chapter 4, "Managing Network File Systems (Overview)," in *System Administration Guide: Network Services* for more information about automatic file sharing.

See also the `share(1M)` and `dfstab(4)` man pages.

11 On the JumpStart install server, become superuser.

12 From the JumpStart install server, start the `scinstall(1M)` utility.

In the media path, replace *arch* with `sparc` or `x86` and replace *ver* with `10` for Oracle Solaris 10.

```
installserver# cd /cdrom/cdrom0/Solaris_arch/Product/sun_cluster/ \
Solaris_ver/Tools/

installserver# ./scinstall
```

The `scinstall` Main Menu is displayed.

13 Choose the menu item, Configure a Cluster to be JumpStarted From This Install Server.

This option is used to configure custom JumpStart finish scripts. JumpStart uses these finish scripts to install the Oracle Solaris Cluster software.

```
 *** Main Menu ***

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

      * 1) Create a new cluster or add a cluster node
      * 2) Configure a cluster to be JumpStarted from this install server
        3) Manage a dual-partition upgrade
        4) Upgrade this cluster node
      * 5) Print release information for this cluster node

      * ?) Help with menu options
      * q) Quit

    Option:  2
```

14 Follow the menu prompts to supply your answers from the configuration planning worksheet.

The `scinstall` command stores your configuration information and copies the `autoscinstall.class` default `class` file in the */jumpstart-dir*`/autoscinstall.d/3.2/` directory. This file is similar to the following example.

```
install_type    initial_install
system_type     standalone
partitioning    explicit
```

```
filesys         rootdisk.s0 free /
filesys         rootdisk.s1 750  swap
filesys         rootdisk.s3 512  /globaldevices
filesys         rootdisk.s7 20
cluster         SUNWCuser        add
package         SUNWman          add
```

15 If necessary, make adjustments to the `autoscinstall.class` file to configure JumpStart to install the flash archive.

Modify entries as necessary to match configuration choices that you made when you installed the Oracle Solaris OS on the flash archive machine or when you ran the `scinstall` utility.

a. To use a lofi device for the global-devices namespace, delete the `filesys` entry for the `/globaldevices` partition.

b. Change the following entries in the `autoscinstall.class` file.

Existing Entry to Replace		New Entry to Add	
`install_type`	`initial_install`	`install_type`	`flash_install`
`system_type`	`standalone`	`archive_location`	*retrieval_type location*

See "archive_location Keyword" in *Oracle Solaris 10 1/13 Installation Guide: JumpStart Installations* for information about valid values for *retrieval_type* and *location* when used with the `archive_location` keyword.

c. Remove all entries that would install a specific package, such as the following entries.

```
cluster         SUNWCuser        add
package         SUNWman          add
```

d. If your configuration has additional Oracle Solaris software requirements, change the `autoscinstall.class` file accordingly.

The `autoscinstall.class` file installs the End User Oracle Solaris Software Group (`SUNWCuser`).

e. If you install the End User Oracle Solaris Software Group (`SUNWCuser`), add to the `autoscinstall.class` file any additional Oracle Solaris software packages that you might need.

The following table lists Oracle Solaris packages that are required to support some Oracle Solaris Cluster functionality. These packages are not included in the End User Oracle Solaris Software Group. See "Oracle Solaris Software Group Considerations" on page 14 for more information.

Feature	Mandatory Oracle Solaris Software Packages
`scsnapshot`	`SUNWp15u SUNWp15v SUNWp15p`
Oracle Solaris Cluster Manager	`SUNWapchr SUNWapchu`

You can change the default `class` file in one of the following ways:

- Edit the `autoscinstall.class` file directly. These changes are applied to all nodes in all clusters that use this custom JumpStart directory.
- Update the `rules` file to point to other profiles, then run the `check` utility to validate the `rules` file.

As long as the Oracle Solaris OS installation profile meets minimum Oracle Solaris Cluster file-system allocation requirements, Oracle Solaris Cluster software places no restrictions on other changes to the installation profile. See "System Disk Partitions" on page 14 for partitioning guidelines and requirements to support Oracle Solaris Cluster software.

For more information about JumpStart profiles, see Chapter 3, "Preparing JumpStart Installations (Tasks)," in *Oracle Solaris 10 1/13 Installation Guide: JumpStart Installations*.

16 To perform any other postinstallation tasks, set up your own finish script.

Your own finish script runs after the standard finish script that is installed by the `scinstall` command. See Chapter 3, "Preparing JumpStart Installations (Tasks)," in *Oracle Solaris 10 1/13 Installation Guide: JumpStart Installations* for information about creating a JumpStart finish script.

a. Ensure that any dependency Oracle Solaris packages will be installed by the default `class` file.

See Step 15.

b. Name your finish script `finish`.

c. Make any modifications for postinstallation tasks that you want the `finish` script to perform.

d. Copy your `finish` script to each *jumpstart-dir*`/autoscinstall.d/nodes/`*node* directory.

Create one *node* directory for each node in the cluster. Or, use this naming convention to create symbolic links to a shared `finish` script.

17 Exit from the JumpStart install server.

18 If you are using switches in the private interconnect of your new cluster, ensure that Neighbor Discovery Protocol (NDP) is disabled.

Follow the procedures in the documentation for your switches to determine whether NDP is enabled and to disable NDP.

During cluster configuration, the software checks that there is no traffic on the private interconnect. If NDP sends any packages to a private adapter when the private interconnect is being checked for traffic, the software will assume that the interconnect is not private and cluster configuration will be interrupted. NDP must therefore be disabled during cluster creation.

After the cluster is established, you can re-enable NDP on the private-interconnect switches if you want to use that feature.

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

- **If Cluster Control Panel (CCP) software is installed and configured on your administrative console, use the `cconsole(1M)` utility to display the individual console screens.**

  As superuser, use the following command to start the `cconsole` utility:

  ```
  adminconsole# /opt/SUNWcluster/bin/cconsole clustername &
  ```

  The `cconsole` 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 `cconsole` utility, connect to the consoles of each node individually.**

20 Shut down each node.

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

21 Boot each node to start the JumpStart installation.

- **On SPARC based systems, do the following:**

  ```
  ok boot net - install
  ```

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

- **On x86 based systems, do the following:**

  a. **Press any key to begin the booting sequence.**

  ```
  Press any key to reboot.
      keystroke
  ```

  b. **As soon as the BIOS information screen appears, immediately press Esc+2 or press the F2 key.**

  After the initialization sequence completes, the BIOS Setup Utility screen appears.

  c. **In the BIOS Setup Utility menu bar, navigate to the Boot menu item.**

  The list of boot devices is displayed.

d. Navigate to the listed IBA that is connected to the same network as the JumpStart PXE install server and move it to the top of the boot order.

The lowest number to the right of the IBA boot choices corresponds to the lower Ethernet port number. The higher number to the right of the IBA boot choices corresponds to the higher Ethernet port number.

e. Save your change and exit the BIOS.

The boot sequence begins again. After further processing, the GRUB menu is displayed.

f. Immediately select the Oracle Solaris JumpStart entry and press Enter.

Note – If the Oracle Solaris JumpStart entry is the only entry listed, you can alternatively wait for the selection screen to time out. If you do not respond in 30 seconds, the system automatically continues the boot sequence.

After further processing, the installation type menu is displayed.

g. From the installation type menu, immediately type the menu number for Custom JumpStart.

Note – If you do not type the number for Custom JumpStart before the 30–second timeout period ends, the system automatically begins the Oracle Solaris interactive installation.

JumpStart installs the Oracle Solaris OS and Oracle Solaris Cluster software on each node. When the installation is successfully completed, each node is fully installed as a new cluster node. Oracle Solaris Cluster installation output is logged in a `/var/cluster/logs/install/scinstall.log.` *N* file.

h. When the BIOS screen again appears, immediately press Esc+2 or press the F2 key.

Note – If you do not interrupt the BIOS at this point, it automatically returns to the installation type menu. There, if no choice is typed within 30 seconds, the system automatically begins an interaction installation.

After further processing, the BIOS Setup Utility is displayed.

i. In the menu bar, navigate to the Boot menu.

The list of boot devices is displayed.

j. Navigate to the Hard Drive entry and move it back to the top of the boot order.

k. Save your change and exit the BIOS.

The boot sequence begins again. No further interaction with the GRUB menu is needed to complete booting into cluster mode.

22 Verify on each node that multiuser services for the Service Management Facility (SMF) are online.

If services are not yet online for a node, wait until the state changes to online before you proceed to the next step.

```
phys-schost# svcs multi-user-server node
STATE          STIME    FMRI
online         17:52:55 svc:/milestone/multi-user-server:default
```

23 If you are installing a new node to an existing cluster, create mount points on the new node for all existing cluster file systems.

a. From another cluster node that is active, display the names of all cluster file systems.

```
phys-schost# mount | grep global | egrep -v node@ | awk '{print $1}'
```

b. On the node that you added to the cluster, create a mount point for each cluster file system in the cluster.

```
phys-schost-new# mkdir -p mountpoint
```

For example, if a file-system name that is returned by the mount command is `/global/dg-schost-1`, run `mkdir -p /global/dg-schost-1` on the node that is being added to the cluster.

Note – The mount points become active after you reboot the cluster in Step 27.

---

24 If you intend to use Oracle Solaris Cluster HA for NFS (HA for NFS) on a highly available local file system, ensure that the loopback file system (LOFS) is disabled.

To disable LOFS, add the following entry to the `/etc/system` file on each node of the cluster.

```
exclude:lofs
```

The change to the `/etc/system` file becomes effective after the next system reboot.

Note – You cannot have LOFS enabled if you use HA for NFS on a highly available local file system *and* have `automountd` running. LOFS can cause switchover problems for HA for NFS. If you choose to add HA for NFS on a highly available local file system, you must make one of the following configuration changes.

However, if you configure non-global zones in your cluster, you must enable LOFS on all cluster nodes. If HA for NFS on a highly available local file system must coexist with LOFS, use one of the other solutions instead of disabling LOFS.

- Disable LOFS.
- Disable the `automountd` daemon.
- Exclude from the automounter map all files that are part of the highly available local file system that is exported by HA for NFS. This choice enables you to keep both LOFS and the `automountd` daemon enabled.

See "The Loopback File System" in *System Administration Guide: Devices and File Systems* for more information about loopback file systems.

25 If you will use any of the following adapters for the cluster interconnect, uncomment the relevant entry in the `/etc/system` file on each node.

Adapter	Entry
`ipge`	`set ipge:ipge_taskq_disable=1`
`ixge`	`set ixge:ixge_taskq_disable=1`

This entry becomes effective after the next system reboot.

26 x86: Set the default boot file.

The setting of this value enables you to reboot the node if you are unable to access a login prompt.

```
grub edit> kernel /platform/i86pc/multiboot kmdb
```

27 If you performed a task that requires a cluster reboot, follow these steps to reboot the cluster.

The following are some of the tasks that require a reboot:

- Adding a new node to an existing cluster
- Installing patches that require a node or cluster reboot
- Making configuration changes that require a reboot to become active

a. On one node, become superuser.

b. Shut down the cluster.

```
phys-schost-1# cluster shutdown -y -g0 clustername
```

Note – Do not reboot the first-installed node of the cluster until *after* the cluster is shut down. Until cluster installation mode is disabled, only the first-installed node, which established the cluster, has a quorum vote. In an established cluster that is still in installation mode, if the cluster is not shut down before the first-installed node is rebooted, the remaining cluster nodes cannot obtain quorum. The entire cluster then shuts down.

Cluster nodes remain in installation mode until the first time that you run the `clsetup` command. You run this command during the procedure "How to Configure Quorum Devices" on page 127.

---

c. Reboot each node in the cluster.

- **On SPARC based systems, do the following:**

  ```
  ok boot
  ```

- **On x86 based systems, do the following:**

  When the GRUB menu is displayed, select the appropriate Oracle Solaris entry and press Enter.

  For more information about GRUB based booting, see "Booting an x86 Based System by Using GRUB (Task Map)" in *Oracle Solaris Administration: Basic Administration*.

The `scinstall` utility installs and configures all cluster nodes and reboots the cluster. The cluster is established when all nodes have successfully booted into the cluster. Oracle Solaris Cluster installation output is logged in a `/var/cluster/logs/install/scinstall.log.`*N* file.

28 (Optional) If you did not perform Step 27 to reboot the nodes, start the Oracle Java Web Console web server manually on each node.

```
phys-schost# smcwebserver start
```

For more information, see the `smcwebserver(1M)` man page.

29 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.

30 (Optional) On each node, enable automatic node reboot if all monitored shared-disk paths fail.

Note – At initial configuration time, disk-path monitoring is enabled by default for all discovered devices.

---

a. Enable automatic reboot.

```
phys-schost# clnode set -p reboot_on_path_failure=enabled
```

`-p`
Specifies the property to set

`reboot_on_path_failure=enable`
Enables automatic node reboot if failure of all monitored shared-disk paths occurs.

b. Verify that automatic reboot on disk-path failure is enabled.

```
phys-schost# clnode show
=== Cluster Nodes ===

Node Name:                                      node
...
  reboot_on_path_failure:                          enabled
...
```

Next Steps If you added a node to a two-node cluster, go to "How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 124.

Otherwise, go to the next appropriate procedure:

- If you installed a multiple-node cluster and chose automatic quorum configuration, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 132.
- If you installed a multiple-node cluster and declined automatic quorum configuration, perform postinstallation setup. Go to "How to Configure Quorum Devices" on page 127.
- If you added a new node to an existing cluster that uses a quorum device, go to "How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 124.
- If you added a new node to an existing cluster that does not use a quorum device, verify the state of the cluster. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 132.
- If you installed a single-node cluster, cluster establishment is complete. Go to "Creating Cluster File Systems" on page 173 to install volume management software and configure the cluster.

Troubleshooting **Disabled scinstall option** – If the JumpStart option of the `scinstall` command does not have an asterisk in front, the option is disabled. This condition indicates that JumpStart setup is not complete or that the setup has an error. To correct this condition, first quit the `scinstall` utility. Repeat Step 1 through Step 16 to correct JumpStart setup, then restart the `scinstall` utility.

▼ How to Prepare the Cluster for Additional Global-Cluster Nodes

Perform this procedure on existing global-cluster nodes to prepare the cluster for the addition of new cluster nodes.

Before You Begin Perform the following tasks:

- Ensure that all necessary hardware is installed.
  - Ensure that the host adapter is installed on the new node. See the *Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual*.
  - Verify that any existing cluster interconnects can support the new node. See the *Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual*.
  - Ensure that any additional storage is installed. See the appropriate Oracle Solaris Cluster storage manual.

1 If you use the Cluster Control Panel (CCP), update the configuration files on the administrative console.

a. Add to the cluster's entry in the `/etc/clusters` file the name of the node that you are adding.

b. Add to the `/etc/serialports` files an entry with the new node name, the host name of the node's console-access device, and the port number.

2 Add the name of the new node to the cluster's authorized-nodes list.

a. On any node, become superuser.

b. Start the `clsetup` utility.

```
phys-schost# clsetup
```

The Main Menu is displayed.

c. Choose the menu item, New Nodes.

d. Choose the menu item, Specify the Name of a Machine Which May Add Itself.

e. Follow the prompts to add the node's name to the list of recognized machines.

The `clsetup` utility displays the message `Command completed successfully` if the task is completed without error.

f. Quit the `clsetup` utility.

3 If you are adding a node to a single-node cluster, ensure that two cluster interconnects already exist by displaying the interconnect configuration.

```
phys-schost# clinterconnect show
```

You must have at least two cables or two adapters configured before you can add a node.

- **If the output shows configuration information for two cables or for two adapters, proceed to Step 4.**
- **If the output shows no configuration information for either cables or adapters, or shows configuration information for only one cable or adapter, configure new cluster interconnects.**

  a. On one node, start the `clsetup` utility.

  ```
  phys-schost# clsetup
  ```

  b. Choose the menu item, Cluster Interconnect.

  c. Choose the menu item, Add a Transport Cable.

  Follow the instructions to specify the name of the node to add to the cluster, the name of a transport adapter, and whether to use a transport switch.

  d. If necessary, repeat Step c to configure a second cluster interconnect.

  e. When finished, quit the `clsetup` utility.

  f. Verify that the cluster now has two cluster interconnects configured.

  ```
  phys-schost# clinterconnect show
  ```

  The command output should show configuration information for at least two cluster interconnects.

4 Ensure that the private-network configuration can support the nodes and private networks that you are adding.

a. Display the maximum numbers of nodes, private networks, and zone clusters that the current private-network configuration supports.

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

The output looks similar to the following:

```
=== Private Network ===

private_netaddr:                                172.16.0.0
  private_netmask:                                255.255.240.0
  max_nodes:                                      64
  max_privatenets:                                10
  max_zoneclusters:                               12
```

b. Determine whether the current private-network configuration can support the increased number of nodes, including non-global zones, and private networks.

- **If the current IP address range is sufficient, you are ready to install the new node.**

  Go to "How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (scinstall)" on page 114.

- **If the current IP address range is not sufficient, reconfigure the private IP-address range.**

  Go to "How to Change the Private Network Configuration When Adding Nodes or Private Networks" on page 108. You must shut down the cluster to change the private IP-address range. This involves switching each resource group offline, disabling all resources in the cluster, then rebooting into noncluster mode before you reconfigure the IP address range.

Next Steps Configure Oracle Solaris Cluster software on the new cluster nodes. Go to "How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (scinstall)" on page 114 or "How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (XML)" on page 121.

▼ How to Change the Private Network Configuration When Adding Nodes or Private Networks

Perform this task to change the global-cluster's private IP-address range to accommodate an increase in one or more of the following cluster components:

- The number of nodes or non-global zones
- The number of private networks
- The number of zone clusters

You can also use this procedure to decrease the private IP-address range.

Note – This procedure requires you to shut down the entire cluster. If you need to change only the netmask, for example, to add support for zone clusters, do not perform this procedure. Instead, run the following command from a global-cluster node that is running in cluster mode to specify the expected number of zone clusters:

```
phys-schost# cluster set-netprops num_zoneclusters=N
```

This command does not require you to shut down the cluster.

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

1. **Become superuser on a node of the cluster.**
2. **From one node, start the `clsetup` utility.**

   ```
   # clsetup
   ```

   The `clsetup` Main Menu is displayed.
3. **Switch each resource group offline.**

   If the node contains non-global zones, any resource groups in the zones are also switched offline.

   a. **Type the number that corresponds to the option for Resource groups and press the Return key.**

      The Resource Group Menu is displayed.

   b. **Type the number that corresponds to the option for Online/Offline or Switchover a resource group and press the Return key.**

   c. **Follow the prompts to take offline all resource groups and to put them in the unmanaged state.**

   d. **When all resource groups are offline, type `q` to return to the Resource Group Menu.**
4. **Disable all resources in the cluster.**

   a. **Type the number that corresponds to the option for Enable/Disable a resource and press the Return key.**

   b. **Choose a resource to disable and follow the prompts.**

   c. **Repeat the previous step for each resource to disable.**

d. When all resources are disabled, type `q` to return to the Resource Group Menu.

5 Quit the `clsetup` utility.

6 Verify that all resources on all nodes are `Offline` and that all resource groups are in the `Unmanaged` state.

```
# cluster status -t resource,resourcegroup
```

`-t`	Limits output to the specified cluster object
`resource`	Specifies resources
`resourcegroup`	Specifies resource groups

7 From one node, shut down the cluster.

```
# cluster shutdown -g0 -y
```

`-g`	Specifies the wait time in seconds
`-y`	Prevents the prompt that asks you to confirm a shutdown from being issued

8 Boot each 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 an x86 Based System by Using GRUB (Task Map)" in *Oracle Solaris Administration: Basic Administration*.

  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.

9 From one node, start the `clsetup` utility.

When run in noncluster mode, the `clsetup` utility displays the Main Menu for noncluster-mode operations.

10 Type the number that corresponds to the option for Change Network Addressing and Ranges for the Cluster Transport and press the Return key.

The `clsetup` utility displays the current private-network configuration, then asks if you would like to change this configuration.

11 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.

12 Change or accept the private-network IP address.

- **To accept the default private-network IP address and proceed to changing the IP address range, type `yes` and press the Return key.**

  The `clsetup` utility will ask if it is okay to accept the default netmask. Skip to the next step to enter your response.

- **To change the default private-network IP address, perform the following substeps.**

  a. **Type `no` in response to the `clsetup` utility question about whether it is okay to accept the default address, then press the Return key.**

     The `clsetup` utility will prompt for the new private-network IP address.

  b. **Type the new IP address and press the Return key.**

     The `clsetup` utility displays the default netmask and then asks if it is okay to accept the default netmask.

13 Change or accept the default private-network IP address range.

The default netmask is `255.255.240.0`. This default IP address range supports up to 64 nodes, 12 zone clusters, and 10 private networks in the cluster.

- **To accept the default IP address range, type `yes` and press the Return key.**

  Then skip to the next step.

- **To change the IP address range, perform the following substeps.**

  a. Type `no` in response to the `clsetup` utility's question about whether it is okay to accept the default address range, then press the Return key.

  When you decline the default netmask, the `clsetup` utility prompts you for the number of nodes and private networks, and zone clusters that you expect to configure in the cluster.

  b. Enter the number of nodes, private networks, and zone clusters that you expect to configure in the cluster.

  From these numbers, the `clsetup` utility calculates two proposed netmasks:

  - The first netmask is the minimum netmask to support the number of nodes, private networks, and zone clusters that you specified.
  - The second netmask supports twice the number of nodes, private networks, and zone clusters that you specified, to accommodate possible future growth.

  c. Specify either of the calculated netmasks, or specify a different netmask that supports the expected number of nodes, private networks, and zone clusters.

14 Type `yes` in response to the `clsetup` utility's question about proceeding with the update.

15 When finished, exit the `clsetup` utility.

16 Reboot each node back into the cluster.

a. Shut down each node.

```
# shutdown -g0 -y
```

b. Boot each node into cluster mode.

- **On SPARC based systems, do the following:**

```
ok boot
```

- **On x86 based systems, do the following:**

  When the GRUB menu is displayed, select the appropriate Oracle Solaris entry and press Enter.

  For more information about GRUB based booting, see "Booting an x86 Based System by Using GRUB (Task Map)" in *Oracle Solaris Administration: Basic Administration*.

17 From one node, start the `clsetup` utility.

```
# clsetup
```

The `clsetup` Main Menu is displayed.

18 Re-enable all disabled resources.

a. Type the number that corresponds to the option for Resource groups and press the Return key.

The Resource Group Menu is displayed.

b. Type the number that corresponds to the option for Enable/Disable a resource and press the Return key.

c. Choose a resource to enable and follow the prompts.

d. Repeat for each disabled resource.

e. When all resources are re-enabled, type `q` to return to the Resource Group Menu.

19 Bring each resource group back online.

If the node contains non-global zones, also bring online any resource groups that are in those zones.

a. Type the number that corresponds to the option for Online/Offline or Switchover a resource group and press the Return key.

b. Follow the prompts to put each resource group into the managed state and then bring the resource group online.

20 When all resource groups are back online, exit the `clsetup` utility.

Type **`q`** to back out of each submenu, or press Ctrl-C.

Next Steps To add a node to an existing cluster, go to one of the following procedures:

- "How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (`scinstall`)" on page 114

- "How to Install Oracle Solaris and Oracle Solaris Cluster Software (JumpStart)" on page 90
- "How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (XML)" on page 121

To create a non-global zone on a cluster node, go to "Configuring a Non-Global Zone on a Global-Cluster Node" on page 191.

▼ How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (`scinstall`)

Perform this procedure to add a new node to an existing global cluster. To use JumpStart to add a new node, instead follow procedures in "How to Install Oracle Solaris and Oracle Solaris Cluster Software (JumpStart)" on page 90.

Note – This procedure uses the interactive form of the `scinstall` command. To use the noninteractive forms of the `scinstall` command, such as when developing installation scripts, see the `scinstall(1M)` man page.

Ensure that Oracle Solaris Cluster software packages are installed on the node, either manually or by using the silent-mode form of the `installer` program, before you run the `scinstall` command. For information about running the `installer` program from an installation script, see Chapter 5, "Installing in Silent Mode," in *Sun Java Enterprise System 5 Update 1 Installation Guide for UNIX*.

Before You Begin Perform the following tasks:

- Ensure that the Oracle Solaris OS is installed to support Oracle Solaris Cluster software.

  If Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for Oracle Solaris Cluster software and any other software that you intend to install on the cluster. See "How to Install Oracle Solaris Software" on page 57 for more information about installing Oracle Solaris software to meet Oracle Solaris Cluster software requirements.
- SPARC: If you are configuring Oracle VM Server for SPARC I/O domains or guest domains as cluster nodes, ensure that Oracle VM Server for SPARC software is installed on each physical machine and that the domains meet Oracle Solaris Cluster requirements. See "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 62.
- Ensure that Oracle Solaris Cluster software packages and patches are installed on the node. See "How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63.

- Ensure that the cluster is prepared for the addition of the new node. See "How to Prepare the Cluster for Additional Global-Cluster Nodes" on page 106.
- Determine which mode of the `scinstall` utility you will use, Typical or Custom. For the Typical installation of Oracle Solaris Cluster software, `scinstall` automatically specifies the following configuration defaults.

Component	Default Value
Cluster-transport switches	`switch1` and `switch2`
Global-devices namespace	A `lofi` device

- Complete one of the following configuration planning worksheets. See "Planning the Oracle Solaris OS" on page 12 and "Planning the Oracle Solaris Cluster Environment" on page 21 for planning guidelines.
  - **Typical Mode Worksheet** – If you will use Typical mode and accept all defaults, complete the following worksheet.

Component	Description/Example	Answer	
Sponsoring Node	What is the name of the sponsoring node? *Choose any node that is active in the cluster.*		
Cluster Name	What is the name of the cluster that you want the node to join?		
Check	Do you want to run the `cluster check` validation utility?	Yes | No	
Autodiscovery of Cluster Transport	Do you want to use autodiscovery to configure the cluster transport? If no, supply the following additional information:	Yes | No	
Point-to-Point Cables	Does the node that you are adding to the cluster make this a two-node cluster?	Yes | No	
	Does the cluster use switches?	Yes | No	
Cluster Switches	If used, what are the names of the two switches? Defaults: `switch1` and `switch2`	*First*	*Second*
Cluster Transport Adapters and Cables	Transport adapter names:	*First*	*Second*
	Where does each transport adapter connect to (*a switch or another adapter*)? Switch defaults: `switch1` and `switch2`		
	For transport switches, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		

Component	Description/Example	Answer
Automatic Reboot	Do you want `scinstall` to automatically reboot the node after installation?	Yes | No

- **Custom Mode Worksheet** – If you will use Custom mode and customize the configuration data, complete the following worksheet.

Component	Description/Example	Answer	
Sponsoring Node	What is the name of the sponsoring node? *Choose any node that is active in the cluster.*		
Cluster Name	What is the name of the cluster that you want the node to join?		
Check	Do you want to run the `cluster check` validation utility?	Yes | No	
Autodiscovery of Cluster Transport	Do you want to use autodiscovery to configure the cluster transport? If no, supply the following additional information:	Yes | No	
Point-to-Point Cables	Does the node that you are adding to the cluster make this a two-node cluster?	Yes | No	
	Does the cluster use switches?	Yes | No	
Cluster Switches	Transport switch name, if used: Defaults: `switch1` and `switch2`	*First*	*Second*
Cluster Transport Adapters and Cables	Transport adapter name:	*First*	*Second*
	Where does each transport adapter connect to (*a switch or another adapter*)? Switch defaults: `switch1` and `switch2`		
	If a transport switch, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		
Global Devices File System	Do you want to use the default `lofi` method?	Yes | No	
	If no, do you want to select another file system?	Yes | No	
Automatic Reboot	Do you want `scinstall` to automatically reboot the node after installation?	Yes | No	

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 **On the cluster node to configure, become superuser.**

2 **Start the `scinstall` utility.**

```
phys-schost-new# /usr/cluster/bin/scinstall
```

The `scinstall` Main Menu is displayed.

3 **Type the option number for Create a New Cluster or Add a Cluster Node and press the Return key.**

```
*** Main Menu ***

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

     * 1) Create a new cluster or add a cluster node
       2) Configure a cluster to be JumpStarted from this install server
       3) Manage a dual-partition upgrade
       4) Upgrade this cluster node
     * 5) Print release information for this cluster node

     * ?) Help with menu options
     * q) Quit

   Option:  1
```

The New Cluster and Cluster Node Menu is displayed.

4 **Type the option number for Add This Machine as a Node in an Existing Cluster and press the Return key.**

5 **Follow the menu prompts to supply your answers from the configuration planning worksheet.**

The `scinstall` utility configures the node and boots the node into the cluster.

6 **Unload the DVD-ROM from the DVD-ROM drive.**

a. **To ensure that the DVD-ROM is not being used, change to a directory that does *not* reside on the DVD-ROM.**

b. **Eject the DVD-ROM.**

```
phys-schost# eject cdrom
```

7 **Repeat this procedure on any other node to add to the cluster until all additional nodes are fully configured.**

8 Verify on each node that multiuser services for the Service Management Facility (SMF) are online.

If services are not yet online for a node, wait until the state changes to online before you proceed to the next step.

```
phys-schost# svcs multi-user-server node
STATE          STIME    FMRI
online         17:52:55 svc:/milestone/multi-user-server:default
```

9 From an active cluster member, prevent any other nodes from joining the cluster.

```
phys-schost# claccess deny-all
```

Alternately, you can use the `clsetup` utility. See "How to Add a Node to an Existing Cluster" in *Oracle Solaris Cluster System Administration Guide* for procedures.

10 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.

11 Verify that all necessary patches are installed.

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

12 (Optional) Enable automatic node reboot if all monitored shared-disk paths fail.

Note – At initial configuration time, disk-path monitoring is enabled by default for all discovered devices.

---

a. Enable automatic reboot.

```
phys-schost# clnode set -p reboot_on_path_failure=enabled
```

`-p`
Specifies the property to set

`reboot_on_path_failure=enable`
Enables automatic node reboot if failure of all monitored shared-disk paths occurs.

b. Verify that automatic reboot on disk-path failure is enabled.

```
phys-schost# clnode show
=== Cluster Nodes ===

Node Name:                                      node
...
  reboot_on_path_failure:                          enabled
...
```

13 If you intend to use Oracle Solaris Cluster HA for NFS (HA for NFS) on a highly available local file system, ensure that the loopback file system (LOFS) is disabled.

To disable LOFS, add the following entry to the `/etc/system` file on each node of the cluster.

```
exclude:lofs
```

The change to the `/etc/system` file becomes effective after the next system reboot.

Note – You cannot have LOFS enabled if you use HA for NFS on a highly available local file system *and* have `automountd` running. LOFS can cause switchover problems for HA for NFS. If you choose to add HA for NFS on a highly available local file system, you must make one of the following configuration changes.

However, if you configure non-global zones in your cluster, you must enable LOFS on all cluster nodes. If HA for NFS on a highly available local file system must coexist with LOFS, use one of the other solutions instead of disabling LOFS.

- Disable LOFS.
- Disable the `automountd` daemon.
- Exclude from the automounter map all files that are part of the highly available local file system that is exported by HA for NFS. This choice enables you to keep both LOFS and the `automountd` daemon enabled.

See "The Loopback File System" in *System Administration Guide: Devices and File Systems* for more information about loopback file systems.

Example 3–3 Configuring Oracle Solaris Cluster Software on an Additional Node

The following example shows the node `phys-schost-3` added to the cluster `schost`. The sponsoring node is `phys-schost-1`.

```
*** Adding a Node to an Existing Cluster ***
Fri Feb  4 10:17:53 PST 2005


scinstall -ik -C schost -N phys-schost-1 -A trtype=dlpi,name=bge2 -A trtype=dlpi,name=bge3
-m endpoint=:bge2,endpoint=switch1 -m endpoint=:bge3,endpoint=switch2


Checking device to use for global devices file system ... done
```

```
Adding node "phys-schost-3" to the cluster configuration ... done
Adding adapter "bge2" to the cluster configuration ... done
Adding adapter "bge3" to the cluster configuration ... done
Adding cable to the cluster configuration ... done
Adding cable to the cluster configuration ... done

Copying the config from "phys-schost-1" ... done

Copying the postconfig file from "phys-schost-1" if it exists ... done
Copying the Common Agent Container keys from "phys-schost-1" ... done


Setting the node ID for "phys-schost-3" ... done (id=1)

Setting the major number for the "did" driver ...
Obtaining the major number for the "did" driver from "phys-schost-1" ... done
"did" driver major number set to 300

Checking for global devices global file system ... done
Updating vfstab ... done

Verifying that NTP is configured ... done
Initializing NTP configuration ... done

Updating nsswitch.conf ...
done

Adding clusternode entries to /etc/inet/hosts ... done


Configuring IP Multipathing groups in "/etc/hostname.<adapter>" files

Updating "/etc/hostname.hme0".

Verifying that power management is NOT configured ... done

Ensure that the EEPROM parameter "local-mac-address?" is set to "true" ... done
The "local-mac-address?" parameter setting has been changed to "true".

Ensure network routing is disabled ... done

Updating file ("ntp.conf.cluster") on node phys-schost-1 ... done
Updating file ("hosts") on node phys-schost-1 ... done

Rebooting ...
```

Troubleshooting **Unsuccessful configuration** – If one or more nodes cannot join the cluster, or if the wrong configuration information was specified, first attempt to rerun this procedure. If that does not correct the problem, perform the procedure "How to Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems" on page 225 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Oracle Solaris Cluster software packages. Then rerun this procedure.

Next Steps If you added a node to an existing cluster that uses a quorum device, go to "How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 124.

Otherwise, go to "How to Verify the Quorum Configuration and Installation Mode" on page 132.

▼ How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (XML)

Perform this procedure to configure a new global-cluster node by using an XML cluster configuration file. The new node can be a duplication of an existing cluster node that runs Oracle Solaris Cluster 3.3 3/13 software.

This procedure configures the following cluster components on the new node:

- Cluster node membership
- Cluster interconnect
- Global devices

Before You Begin Perform the following tasks:

- Ensure that the Oracle Solaris OS is installed to support Oracle Solaris Cluster software.

  If Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for Oracle Solaris Cluster software and any other software that you intend to install on the cluster. See "How to Install Oracle Solaris Software" on page 57 for more information about installing Oracle Solaris software to meet Oracle Solaris Cluster software requirements.
- SPARC: If you are configuring Oracle VM Server for SPARC I/O domains or guest domains as cluster nodes, ensure that Oracle VM Server for SPARC software is installed on each physical machine and that the domains meet Oracle Solaris Cluster requirements. See "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 62.
- Ensure that Oracle Solaris Cluster software packages and any necessary patches are installed on the node. See "How to Install Oracle Solaris Cluster Framework and Data-Service Software Packages" on page 63.
- Ensure that the cluster is prepared for the addition of the new node. See "How to Prepare the Cluster for Additional Global-Cluster Nodes" on page 106.

1 Ensure that Oracle Solaris Cluster software is not yet configured on the potential node that you want to add to a cluster.

a. Become superuser on the potential node.

b. Determine whether Oracle Solaris Cluster software is configured on the potential node.

```
phys-schost-new# /usr/sbin/clinfo -n
```

- **If the command fails, go to Step 2.**

  Oracle Solaris Cluster software is not yet configured on the node. You can add the potential node to the cluster.

- **If the command returns a node ID number, proceed to Step c.**

  Oracle Solaris Cluster software is already a configured on the node. Before you can add the node to a different cluster, you must remove the existing cluster configuration information.

c. Boot the potential node into noncluster mode.

- **On SPARC based systems, perform the following command:**

  ```
  ok boot -x
  ```

- **On x86 based systems, perform the following commands:**

  i. **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 an x86 Based System by Using GRUB (Task Map)" in *Oracle Solaris Administration: Basic Administration*.

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

  iii. **Add `-x` to the command to specify that the system boot into noncluster mode.**

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

  The screen displays the edited command.

  v. **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.

d. Unconfigure Oracle Solaris Cluster software from the potential node.

```
phys-schost-new# /usr/cluster/bin/clnode remove
```

2 **If you are duplicating a node that runs Oracle Solaris Cluster 3.3 3/13 software, create a cluster configuration XML file.**

a. **Become superuser on the cluster node that you want to duplicate.**

b. **Export the existing node's configuration information to a file.**

```
phys-schost# clnode export -o clconfigfile
```

`-o`
Specifies the output destination.

clconfigfile
The name of the cluster configuration XML file. The specified file name can be an existing file or a new file that the command will create.

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

c. **Copy the cluster configuration XML file to the potential node that you will configure as a new cluster node.**

3 **Become superuser on the potential node.**

4 **Modify the cluster configuration XML file as needed.**

a. **Open your cluster configuration XML file for editing.**

- **If you are duplicating an existing cluster node, open the file that you created with the `clnode export` command.**
- **If you are not duplicating an existing cluster node, create a new file.**

  Base the file on the element hierarchy that is shown in the `clconfiguration(5CL)` man page. You can store the file in any directory.

b. **Modify the values of the XML elements to reflect the node configuration that you want to create.**

See the `clconfiguration(5CL)` man page for details about the structure and content of the cluster configuration XML file.

5 **Validate the cluster configuration XML file.**

```
phys-schost-new# xmllint --valid --noout clconfigfile
```

6 **Configure the new cluster node.**

```
phys-schost-new# clnode add -n sponsornode -i clconfigfile
```

`-n` *sponsornode*
Specifies the name of an existing cluster member to act as the sponsor for the new node.

`-i` *clconfigfile*
Specifies the name of the cluster configuration XML file to use as the input source.

7 (Optional) Enable automatic node reboot if all monitored shared-disk paths fail.

Note – At initial configuration time, disk-path monitoring is enabled by default for all discovered devices.

---

a. Enable automatic reboot.

```
phys-schost# clnode set -p reboot_on_path_failure=enabled
```

`-p`
Specifies the property to set

`reboot_on_path_failure=enable`
Enables automatic node reboot if failure of all monitored shared-disk paths occurs.

b. Verify that automatic reboot on disk-path failure is enabled.

```
phys-schost# clnode show
=== Cluster Nodes ===

Node Name:                                      node
...
  reboot_on_path_failure:                          enabled
...
```

Troubleshooting **Unsuccessful configuration** – If one or more nodes cannot join the cluster, or if the wrong configuration information was specified, first attempt to rerun this procedure. If that does not correct the problem, perform the procedure "How to Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems" on page 225 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Oracle Solaris Cluster software packages. Then rerun this procedure.

Next Steps If you added a node to a cluster that uses a quorum device, go to "How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 124.

Otherwise, go to "How to Verify the Quorum Configuration and Installation Mode" on page 132.

▼ How to Update Quorum Devices After Adding a Node to a Global Cluster

If you added a node to a global cluster, you must update the configuration information of the quorum devices, regardless of whether you use shared disks, NAS devices, a quorum server, or a combination. To do this, you remove all quorum devices and update the global-devices

namespace. You can optionally reconfigure any quorum devices that you still want to use. This registers the new node with each quorum device, which can then recalculate its vote count based on the new number of nodes in the cluster.

Any newly configured SCSI quorum devices will be set to SCSI-3 reservations.

Before You Begin Ensure that you have completed installation of Oracle Solaris Cluster software on the added node.

1 On any node of the cluster, become superuser.

2 Ensure that all cluster nodes are online.

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

3 View the current quorum configuration.

Command output lists each quorum device and each node. The following example output shows the current SCSI quorum device, `d3`.

```
phys-schost# clquorum list
d3
...
```

4 Note the name of each quorum device that is listed.

5 Remove the original quorum device.

Perform this step for each quorum device that is configured.

```
phys-schost# clquorum remove devicename
```

devicename
Specifies the name of the quorum device.

6 Verify that all original quorum devices are removed.

If the removal of the quorum devices was successful, no quorum devices are listed.

```
phys-schost# clquorum status
```

7 Update the global-devices namespace.

```
phys-schost# cldevice populate
```

Note – This step is necessary to prevent possible node panic.

8 On each node, verify that the `cldevice populate` command has completed processing before you attempt to add a quorum device.

The `cldevice populate` command executes remotely on all nodes, even through the command is issued from just one node. To determine whether the `cldevice populate` command has completed processing, run the following command on each node of the cluster.

```
phys-schost# ps -ef | grep scgdevs
```

9 (Optional) Add a quorum device.

You can configure the same device that was originally configured as the quorum device or choose a new shared device to configure.

a. (Optional) If you want to choose a new shared device to configure as a quorum device, display all devices that the system checks.

Otherwise, skip to Step c.

```
phys-schost# cldevice list -v
```

Output resembles the following:

```
DID Device          Full Device Path
----------          ----------------
d1                  phys-schost-1:/dev/rdsk/c0t0d0
d2                  phys-schost-1:/dev/rdsk/c0t6d0
d3                  phys-schost-2:/dev/rdsk/c1t1d0
d3                  phys-schost-1:/dev/rdsk/c1t1d0
...
```

b. From the output, choose a shared device to configure as a quorum device.

c. Configure the shared device as a quorum device.

```
phys-schost# clquorum add -t type devicename
```

`-t` *type*
Specifies the type of quorum device. If this option is not specified, the default type `shared_disk` is used.

d. Repeat for each quorum device that you want to configure.

e. Verify the new quorum configuration.

```
phys-schost# clquorum list
```

Output should list each quorum device and each node.

Example 3–4 Updating SCSI Quorum Devices After Adding a Node to a Two-Node Cluster

The following example identifies the original SCSI quorum device `d2`, removes that quorum device, lists the available shared devices, updates the global-device namespace, configures `d3` as a new SCSI quorum device, and verifies the new device.

```
phys-schost# clquorum list
d2
phys-schost-1
phys-schost-2

phys-schost# clquorum remove d2
phys-schost# clquorum status
...
--- Quorum Votes by Device ---

Device Name       Present      Possible      Status
-----------       -------      --------      ------

phys-schost# cldevice list -v
DID Device          Full Device Path
----------          ----------------
...
d3                  phys-schost-2:/dev/rdsk/c1t1d0
d3                  phys-schost-1:/dev/rdsk/c1t1d0
...
phys-schost# cldevice populate
phys-schost# ps -ef - grep scgdevs
phys-schost# clquorum add d3
phys-schost# clquorum list
d3
phys-schost-1
phys-schost-2
```

Next Steps Go to "How to Verify the Quorum Configuration and Installation Mode" on page 132.

▼ How to Configure Quorum Devices

Note – You do not need to configure quorum devices in the following circumstances:

- You chose automatic quorum configuration during Oracle Solaris Cluster software configuration.
- You installed a single-node global cluster.
- You added a node to an existing global cluster and already have sufficient quorum votes assigned.

Instead, proceed to "How to Verify the Quorum Configuration and Installation Mode" on page 132.

Perform this procedure one time only, after the new cluster is fully formed. Use this procedure to assign quorum votes and then to remove the cluster from installation mode.

Before You Begin

- Perform the following preparations to configure a quorum server or a NAS device as a quorum device.

- **Quorum servers** – To configure a quorum server as a quorum device, do the following:
  - Install the Quorum Server software on the quorum server host machine and start the quorum server. For information about installing and starting the quorum server, see "How to Install and Configure Quorum Server Software" on page 51.
  - 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.
  - Have available the following information:
    - A name to assign to the configured quorum device
    - The IP address of the quorum server host machine
    - The port number of the quorum server
- **NAS devices** – To configure a network-attached storage (NAS) device as a quorum device, install the NAS device hardware and software. See *Oracle Solaris Cluster 3.3 3/13 With Network-Attached Storage Device Manual* and your device documentation for requirements and installation procedures for NAS hardware and software.

1 If both of the following conditions apply, modify the netmask file entries for the public network on each cluster node.

- You intend to use a quorum server.
- The public network uses variable-length subnet masking, also called classless inter domain routing (CIDR).

If you use a quorum server but the public network uses classful subnets, as defined in RFC 791, you do not need to perform this step.

a. Add to the `/etc/inet/netmasks` file an entry for each public subnet that the cluster uses.

The following is an example entry that contains a public-network IP address and netmask:

```
10.11.30.0    255.255.255.0
```

b. Append `netmask + broadcast +` to the hostname entry in each `/etc/hostname.`*adapter* file.

nodename **`netmask + broadcast +`**

2 On one node, become superuser.

3 **Ensure that all cluster nodes are online.**

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

4 **To use a shared disk as a quorum device, verify device connectivity to the cluster nodes and choose the device to configure.**

a. **From one node of the cluster, display a list of all the devices that the system checks.**

You do not need to be logged in as superuser to run this command.

```
phys-schost-1# cldevice list -v
```

Output resembles the following:

```
DID Device          Full Device Path
----------          ----------------
d1                  phys-schost-1:/dev/rdsk/c0t0d0
d2                  phys-schost-1:/dev/rdsk/c0t6d0
d3                  phys-schost-2:/dev/rdsk/c1t1d0
d3                  phys-schost-1:/dev/rdsk/c1t1d0
...
```

b. **Ensure that the output shows all connections between cluster nodes and storage devices.**

c. **Determine the global device-ID name of each shared disk that you are configuring as a quorum device.**

Note – Any shared disk that you choose must be qualified for use as a quorum device. See "Quorum Devices" on page 34 for further information about choosing quorum devices.

Use the `scdidadm` output from Step a to identify the device–ID name of each shared disk that you are configuring as a quorum device. For example, the output in Step a shows that global device `d3` is shared by `phys-schost-1` and `phys-schost-2`.

5 **To use a shared disk that does not support the SCSI protocol, ensure that fencing is disabled for that shared disk.**

a. **Display the fencing setting for the individual disk.**

```
phys-schost# cldevice show device

=== DID Device Instances ===
DID Device Name:                                    /dev/did/rdsk/dN
...
  default_fencing:                                   nofencing
...
```

- **If fencing for the disk is set to `nofencing` or `nofencing-noscrub`, fencing is disabled for that disk. Go to Step 6.**

- **If fencing for the disk is set to `pathcount` or `scsi`, disable fencing for the disk. Skip to Step c.**
- **If fencing for the disk is set to `global`, determine whether fencing is also disabled globally. Proceed to Step b.**

  Alternatively, you can simply disable fencing for the individual disk, which overrides for that disk whatever value the `global_fencing` property is set to. Skip to Step c to disable fencing for the individual disk.

b. Determine whether fencing is disabled globally.

```
phys-schost# cluster show -t global

=== Cluster ===
Cluster name:                                   cluster
...
   global_fencing:                                nofencing
...
```

- **If global fencing is set to `nofencing` or `nofencing-noscrub`, fencing is disabled for the shared disk whose `default_fencing` property is set to `global`. Go to Step 6.**
- **If global fencing is set to `pathcount` or `prefer3`, disable fencing for the shared disk. Proceed to Step c.**

Note – If an individual disk has its `default_fencing` property set to `global`, the fencing for that individual disk is disabled only while the cluster-wide `global_fencing` property is set to `nofencing` or `nofencing-noscrub`. If the `global_fencing` property is changed to a value that enables fencing, then fencing becomes enabled for all disks whose `default_fencing` property is set to `global`.

---

c. Disable fencing for the shared disk.

```
phys-schost# cldevice set \
-p default_fencing=nofencing-noscrub device
```

d. Verify that fencing for the shared disk is now disabled.

```
phys-schost# cldevice show device
```

6 Start the `clsetup` utility.

```
phys-schost# clsetup
```

The Initial Cluster Setup screen is displayed.

Note – If the Main Menu is displayed instead, initial cluster setup was already successfully performed. Skip to Step 11.

---

7 **Answer the prompt `Do you want to add any quorum devices?`.**

- **If your cluster is a two-node cluster, you must configure at least one shared quorum device. Type `Yes` to configure one or more quorum devices.**
- **If your cluster has three or more nodes, quorum device configuration is optional.**
  - **Type `No` if you do not want to configure additional quorum devices. Then skip to Step 10.**
  - **Type `Yes` to configure additional quorum devices. Then proceed to Step 8.**

8 **Specify what type of device you want to configure as a quorum device.**

Quorum Device Type	Description
`shared_disk`	Shared LUNs from the following: ■ Shared SCSI disk ■ Serial Attached Technology Attachment (SATA) storage ■ Sun NAS ■ Sun ZFS Storage Appliance
`quorum_server`	Quorum server

9 **Specify the name of the device to configure as a quorum device.**

For a quorum server, also specify the following information:

- The IP address of the quorum server host
- The port number that is used by the quorum server to communicate with the cluster nodes

10 **At the prompt `Is it okay to reset "installmode"?`, type `Yes`.**

After the `clsetup` utility sets the quorum configurations and vote counts for the cluster, the message `Cluster initialization is complete` is displayed. The utility returns you to the Main Menu.

11 **Quit the `clsetup` utility.**

Next Steps Verify the quorum configuration and that installation mode is disabled. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 132.

Troubleshooting **Interrupted clsetup processing** - If the quorum setup process is interrupted or fails to be completed successfully, rerun `clsetup`.

Changes to quorum vote count – If you later 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 by removing each quorum device and then add it back into the configuration, one quorum device at a time. 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. See the procedure "How to Modify a Quorum Device Node List" in Chapter 6, "Administering Quorum," in *Oracle Solaris Cluster System Administration Guide*.

Unreachable quorum device – If you see messages on the cluster nodes that a quorum device is unreachable, or if you see failures of cluster nodes with the message `CMM: Unable to acquire the quorum device`, there might be a problem with the quorum device or the path to it. Check that both the quorum device and the path to it are functional.

If the problem persists, use a different quorum device. Or, if you want to use the same quorum device, increase the quorum timeout to a high value, as follows:

Note – For Oracle Real Application Clusters (Oracle RAC), do not change the default quorum timeout of 25 seconds. In certain split-brain scenarios, a longer timeout 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 timeout, use a different quorum device.

---

1. **Become superuser.**
2. **On each cluster node, edit the `/etc/system` file as superuser to set the timeout to a high value.**

   The following example sets the timeout to 700 seconds.

```
phys-schost# vi /etc/system
...
set cl_haci:qd_acquisition_timer=700
```

3. **From one node, shut down the cluster.**

```
phys-schost-1# cluster shutdown -g0 -y
```

4. **Boot each node back into the cluster.**

   Changes to the `/etc/system` file are initialized after the reboot.

▼ How to Verify the Quorum Configuration and Installation Mode

Perform this procedure to verify that quorum configuration was completed successfully, if quorum was configured, and that cluster installation mode is disabled.

You do not need to be superuser to run these commands.

1 **From any global-cluster node, verify the device and node quorum configurations.**

```
phys-schost% clquorum list
```

Output lists each quorum device and each node.

2 **From any node, verify that cluster installation mode is disabled.**

```
phys-schost% cluster show -t global | grep installmode
  installmode:                                    disabled
```

Cluster installation and creation is complete.

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- If you want to change any private hostnames, go to "How to Change Private Hostnames" on page 134.
- If you did not install your own `/etc/inet/ntp.conf` file before you installed Oracle Solaris Cluster software, install or create the NTP configuration file. Go to "How to Configure Network Time Protocol (NTP)" on page 140.
- If you want to configure IPsec on the private interconnect, go to "How to Configure IP Security Architecture (IPsec) on the Cluster Private Interconnect" on page 142.
- To configure Solaris Volume Manager software, go to Chapter 4, "Configuring Solaris Volume Manager Software."
- To create cluster file systems, go to "How to Create Cluster File Systems" on page 173.
- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 191.
- Install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation that is supplied with the application software and the *Oracle Solaris Cluster Data Services Planning and Administration Guide*.
- Before you put the cluster into production, make a baseline recording of the cluster configuration for future diagnostic purposes. Go to "How to Record Diagnostic Data of the Cluster Configuration" on page 147.

See Also Make a backup of your cluster configuration.

An archived backup of your cluster configuration facilitates easier recovery of the your cluster configuration. For more information, see "How to Back Up the Cluster Configuration" in *Oracle Solaris Cluster System Administration Guide*.

How to Change Private Hostnames

Perform this task if you do not want to use the default private hostnames, `clusternode`*nodeid*`-priv`, that are assigned during Oracle Solaris Cluster software installation.

Note – Do *not* perform this procedure after applications and data services have been configured and have been started. Otherwise, an application or data service might continue to use the old private hostname after the hostname is renamed, which would cause hostname conflicts. If any applications or data services are running, stop them before you perform this procedure.

Perform this procedure on one active node of the cluster.

1. **Become superuser on a global-cluster node.**

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

   ```
   phys-schost# clsetup
   ```

   The `clsetup` Main Menu is displayed.

3. **Type the option number for Private Hostnames and press the Return key.**

   The Private Hostname Menu is displayed.

4. **Type the option number for Change a Private Hostname and press the Return key.**

5. **Follow the prompts to change the private hostname.**

   Repeat for each private hostname to change.

6. **Verify the new private hostnames.**

   ```
   phys-schost# clnode show -t node | grep privatehostname
     privatehostname:                                clusternode1-priv
     privatehostname:                                clusternode2-priv
     privatehostname:                                clusternode3-priv
   ```

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- If you did not install your own `/etc/inet/ntp.conf` file before you installed Oracle Solaris Cluster software, install or create the NTP configuration file. Go to "How to Configure Network Time Protocol (NTP)" on page 140.
- If you want to configure IPsec on the private interconnect, go to "How to Configure IP Security Architecture (IPsec) on the Cluster Private Interconnect" on page 142.
- To configure Solaris Volume Manager software, go to Chapter 4, "Configuring Solaris Volume Manager Software."

- To create cluster file systems, go to "How to Create Cluster File Systems" on page 173.
- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 191.
- Install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation that is supplied with the application software and the *Oracle Solaris Cluster Data Services Planning and Administration Guide*.
- Before you put the cluster into production, make a baseline recording of the cluster configuration for future diagnostic purposes. Go to "How to Record Diagnostic Data of the Cluster Configuration" on page 147.

Configuring the Distribution of Resource Group Load Across Nodes

You can enable the automatic distribution of resource group load across nodes or zones by setting load limits. 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. Each resource group is started on a node from its node list. The Resource Group Manager (RGM) chooses a node that best satisfies the configured load distribution policy. 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.

You can configure load limits in a global cluster or a zone cluster.

The factors you set to control load distribution on each node include load limits, resource group priority, and preemption mode. In the global cluster, you can set the `Concentrate_load` property to choose the preferred load distribution policy: to concentrate resource group load onto as few nodes as possible without exceeding load limits or to spread the load out as evenly as possible across all available nodes. The default behavior is to spread out the resource group load. Each resource group is still limited to running only on nodes in its node list, regardless of load factor and load limit settings.

Note – You can use the command line, the Oracle Solaris Cluster Manager interface, or the `clsetup` utility to configure load distribution for resource groups. The following procedure illustrates how to configure load distribution for resource groups using the `clsetup` utility. For instructions on using the command line to perform these procedures, see "Configuring Load Limits" in *Oracle Solaris Cluster System Administration Guide*.

This section contains the following procedures:

How to Configure Load Limits for a Node

Each cluster node or zone can have its own set of load limits. You assign load factors to resource groups, and the load factors correspond to the defined load limits of the nodes. You can set soft load limits (which can be exceeded) or hard load limits (which cannot be exceeded).

1. **Become superuser on one active node of the cluster.**

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

   ```
   phys-schost# clsetup
   ```

   The `clsetup` menu is displayed.

3. **Choose the menu item, Other Cluster Tasks.**

   The Other Cluster Tasks Menu is displayed.

4. **Choose the menu item, Manage Resource Group Load Distribution.**

   The Manage Resource Group Load Distribution Menu is displayed.

5. **Choose the menu item, Manage Load Limits.**

   The Manage load limits Menu is displayed.

6. **Type `yes` and press the Return key to continue.**

7. **Type the option number that corresponds to the operation you want to perform and press the Return key.**

   You can create a load limit, modify a load limit, or delete a load limit.

8. **If you chose to create a load limit, select the option number that corresponds to the node where you want to set the load limit.**

   If you want to set a load limit on a second node, select the option number that corresponds to the second node and press the Return key. After you have selected all the nodes where you want to configure load limits, type `q` and press the Return key.

9. **Type `yes` and press the Return key to confirm the nodes you selected in Step 8.**

10. **Type the name of the load limit and press the Return key.**

    For example, type `mem_load` as the name of a load limit.

11 Type `yes` or `no` to specify a soft limit value and press the Return key.

If you typed **`yes`**, type the soft limit value and press Enter.

12 Type `yes` or `no` to specify a hard limit value and press the Return key.

If you typed **`yes`**, type the hard limit value and press Enter.

13 Type `yes` and press the Return key to proceed with creating the load limit.

14 Type `yes` to proceed with the update and press the Return key.

The message `Command completed successfully` is displayed, along with the soft and hard load limits for the nodes you selected. Press the Return key to continue.

15 You can modify or delete a load limit by following the prompts in the `clsetup` utility.

Return to the previous menu by typing `q` and pressing the Return key.

▼ How to Set Priority for a Resource Group

You can configure a resource group to have a higher priority so that it is less likely to be displaced from a specific node. If load limits are exceeded, lower-priority resource groups might be forced offline.

1 Become superuser on one active node of the cluster.

2 Start the `clsetup` utility.

```
phys-schost# clsetup
```

The `clsetup` menu is displayed.

3 Choose the menu item, Other Cluster Tasks.

The Other Cluster Tasks Menu is displayed.

4 Choose the menu item, Manage Resource Group Load Distribution.

The Manage Resource Group Load Distribution Menu is displayed.

5 Choose the menu item, Set Priority Per Resource Group.

The Set the Priority of a Resource Group Menu is displayed.

6 Type `yes` and Press the Return key.

7 Type the option that corresponds to the resource group and press the Return key.

The existing Priority value is displayed. The default Priority value is 500.

8 Type the new Priority value and press the Return key.

9. **Type `yes` to confirm your entry and press the Return key.**

10. **Press the Return key to return to the previous menu.**

    The Manage Resource Group Load Distribution Menu is displayed.

How to Set Load Factors for a Resource Group

A load factor is a value that you assign to the load on a load limit. Load factors are assigned to a resource group, and those load factors correspond to the defined load limits of the nodes.

1. **Become superuser on one active node of the cluster.**

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

   ```
   phys-schost# clsetup
   ```

   The `clsetup` menu is displayed.

3. **Choose the menu item, Other Cluster Tasks.**

   The Other Cluster Tasks Menu is displayed.

4. **Choose the menu item, Manage Resource Group Load Distribution.**

   The Manage Resource Group Load Distribution Menu is displayed.

5. **Choose the menu item, Set Load Factors Per Resource Group.**

   The Set the load factors of a Resource Group Menu is displayed.

6. **Type `yes` and press the Return key.**

7. **Type the option number for the resource group and press the Return key.**

8. **Type the desired load factor.**

   For example, you can set a load factor called `mem_load` on the resource group you selected by typing `mem_load@50`. Press `Ctrl-D` when you are done.

9. **Press the Return key to proceed with the update.**

10. **Press the Return key to return to the previous menu.**

    The Manage Resource Group Load Distribution Menu is displayed.

How to Set Preemption Mode for a Resource Group

The `preemption_mode` property determines if a resource group will be preempted from a node by a higher-priority resource group because of node overload. The property indicates the cost of moving a resource group from one node to another.

1. **Become superuser on one active node of the cluster.**

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

   ```
   phys-schost# clsetup
   ```

   The `clsetup` menu is displayed.

3. **Choose the menu item, Other Cluster Tasks.**

   The Other Cluster Tasks Menu is displayed.

4. **Choose the menu item, Manage Resource Group Load Distribution.**

   The Manage Resource Group Load Distribution Menu is displayed.

5. **Choose the menu item, Set Preemption Mode per Resource Group.**

   The Set the Preemption Mode of a Resource Group Menu is displayed.

6. **Type `yes` and press the Return key to continue.**

7. **Type the option number for the resource group and press the Return key.**

   If the resource group has a preemption mode set, it is displayed, similar to the following:

```
The preemption mode property of "rg11" is currently set to the following: preemption mode:  Has_Cost
```

8. **Type the option number for the preemption mode you want and press the Return key.**

   The three choices are `Has_cost`, `No_cost`, or `Never`.

9. **Type `yes` to proceed with the update and press the Return key.**

10. **Press the Return key to return to the previous menu.**

    The Manage Resource Group Load Distribution Menu is displayed.

▼ How to Concentrate Load Onto Fewer Nodes in the Cluster

Setting the `Concentrate_load` property to false causes the cluster to spread resource group loads evenly across all available nodes. If you set this property to `True`, the cluster attempts to concentrate resource group load on the fewest possible nodes without exceeding load limits. By default, the `Concentrate_load` property is set to *False*. You can only set the `Concentrate_load` property in a global cluster; you cannot set this property in a zone cluster. In a zone cluster, the default setting is always *False*.

1. **Become superuser on one active node of the cluster.**

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

   ```
   phys-schost# clsetup
   ```

The `clsetup` menu is displayed.

3 **Choose the menu item, Other cluster tasks.**
The Other Cluster Tasks Menu is displayed.

4 **Choose the menu item, Set the concentrate_load Property of the Cluster.**
The Set the Concentrate Load Property of the Cluster Menu is displayed.

5 **Type yes and press the Return key.**
The current value of `TRUE` or `FALSE` is displayed.

6 **Type yes to change the value and press the Return key.**

7 **Type yes to proceed with the update and press the Return key.**

8 **Press the Return key to return to the previous menu.**
The Other Cluster Tasks Menu is displayed.

▼ How to Configure Network Time Protocol (NTP)

Note – If you installed your own `/etc/inet/ntp.conf` file **before** you installed Oracle Solaris Cluster software, you do not need to perform this procedure. Determine your next step:

Perform this task to create or modify the NTP configuration file after you perform any of the following tasks:

- Install Oracle Solaris Cluster software
- Add a node to an existing global cluster
- Change the private hostname of a node in the global cluster

If you added a node to a single-node cluster, you must ensure that the NTP configuration file that you use is copied to the original cluster node as well as to the new node.

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

2 **If you have your own `/etc/inet/ntp.conf` file, copy your file to each node of the cluster.**

3 **If you do not have your own `/etc/inet/ntp.conf` file to install, use the `/etc/inet/ntp.conf.cluster` file as your NTP configuration file.**

Note – Do not rename the `ntp.conf.cluster` file as `ntp.conf`.

If the `/etc/inet/ntp.conf.cluster` file does not exist on the node, you might have an `/etc/inet/ntp.conf` file from an earlier installation of Oracle Solaris Cluster software. Oracle Solaris Cluster software creates the `/etc/inet/ntp.conf.cluster` file as the NTP configuration file if an `/etc/inet/ntp.conf` file is not already present on the node. If so, perform the following edits instead on that `ntp.conf` file.

a. **Use your preferred text editor to open the NTP configuration file on one node of the cluster for editing.**

b. **Ensure that an entry exists for the private hostname of each cluster node.**

   If you changed any node's private hostname, ensure that the NTP configuration file contains the new private hostname.

c. **If necessary, make other modifications to meet your NTP requirements.**

d. **Copy the NTP configuration file to all nodes in the cluster.**

   The contents of the NTP configuration file must be identical on all cluster nodes.

4 **Stop the NTP daemon on each node.**

Wait for the command to complete successfully on each node before you proceed to Step 5.

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

5 **Restart the NTP daemon on each node.**

- **If you use the `ntp.conf.cluster` file, run the following command:**

  ```
  phys-schost# /etc/init.d/xntpd.cluster start
  ```

  The `xntpd.cluster` startup script first looks for the `/etc/inet/ntp.conf` file.

  - If the `ntp.conf` file exists, the script exits immediately without starting the NTP daemon.
  - If the `ntp.conf` file does not exist but the `ntp.conf.cluster` file does exist, the script starts the NTP daemon. In this case, the script uses the `ntp.conf.cluster` file as the NTP configuration file.

- **If you use the `ntp.conf` file, run the following command:**

  ```
  phys-schost# svcadm enable ntp
  ```

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- If you want to configure IPsec on the private interconnect, go to "How to Configure IP Security Architecture (IPsec) on the Cluster Private Interconnect" on page 142.
- To configure Solaris Volume Manager software, go to Chapter 4, "Configuring Solaris Volume Manager Software."
- To create cluster file systems, go to "How to Create Cluster File Systems" on page 173.
- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 191.
- Install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation that is supplied with the application software and the *Oracle Solaris Cluster Data Services Planning and Administration Guide*.
- Before you put the cluster into production, make a baseline recording of the cluster configuration for future diagnostic purposes. Go to "How to Record Diagnostic Data of the Cluster Configuration" on page 147.

▼ How to Configure IP Security Architecture (IPsec) on the Cluster Private Interconnect

You can configure IP Security Architecture (IPsec) for the `clprivnet` interface to provide secure TCP/IP communication on the cluster interconnect.

For information about IPsec, see Part IV, "IP Security," in *Oracle Solaris Administration: IP Services* and the `ipsecconf(1M)` man page. For information about the `clprivnet` interface, see the `clprivnet(7)` man page.

Perform this procedure on each global-cluster voting node that you want to configure to use IPsec.

1 Become superuser.

2 On each node, determine the IP address of the `clprivnet` interface of the node.

```
phys-schost# ifconfig clprivnet0
```

3 On each node, configure the `/etc/inet/ipsecinit.conf` policy file and add Security Associations (SAs) between each pair of private-interconnect IP addresses that you want to use IPsec.

Follow the instructions in "How to Secure Traffic Between Two Systems With IPsec" in *Oracle Solaris Administration: IP Services*. In addition, observe the following guidelines:

- Ensure that the values of the configuration parameters for these addresses are consistent on all the partner nodes.
- Configure each policy as a separate line in the configuration file.

- To implement IPsec without rebooting, follow the instructions in the procedure's example, "Securing Traffic With IPsec Without Rebooting."

For more information about the `sa unique` policy, see the `ipsecconf(1M)` man page.

a. In each file, add one entry for each `clprivnet` IP address in the cluster to use IPsec.

Include the `clprivnet` IP address of the local node.

b. If you use VNICs, also add one entry for the IP address of each physical interface that is used by the VNICs.

c. (Optional) To enable striping of data over all links, include the `sa unique` policy in the entry.

This feature helps the driver to optimally utilize the bandwidth of the cluster private network, which provides a high granularity of distribution and better throughput. The `clprivnet`interface uses the Security Parameter Index (SPI) of the packet to stripe the traffic.

4 On each node, edit the `/etc/inet/ike/config` file to set the `p2_idletime_secs` parameter.

Add this entry to the policy rules that are configured for cluster transports. This setting provides the time for security associations to be regenerated when a cluster node reboots, and limits how quickly a rebooted node can rejoin the cluster. A value of 30 seconds should be adequate.

```
phys-schost# vi /etc/inet/ike/config
...
{
    label "clust-priv-interconnect1-clust-priv-interconnect2"
...
p2_idletime_secs 30
}
...
```

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- To configure Solaris Volume Manager software, go to Chapter 4, "Configuring Solaris Volume Manager Software."
- To create cluster file systems, go to "How to Create Cluster File Systems" on page 173.
- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 191.
- Install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation that is supplied with the application software and the *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

Otherwise, if you have completed all hardware and software installation and configuration tasks, validate the cluster. Go to "How to Validate the Cluster" on page 144.

How to Validate the Cluster

After you complete all configuration of the cluster, use the `cluster check` command to validate the cluster configuration and functionality. For more information, see the `cluster(1CL)` man page.

Tip – For ease of future reference or troubleshooting, for each validation that you run, use the `-o` *outputdir* option to specify a subdirectory for log files. Reuse of an existing subdirectory name will remove all existing files in the subdirectory. Therefore, to ensure that log files are available for future reference, specify a unique subdirectory name for each cluster check that you run.

Before You Begin Ensure that you have completed the installation and configuration of all hardware and software components in the cluster, including firmware and patches.

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

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

   Go to the Patches & Updates tab of My Oracle Support. Using the Advanced Search, select "Solaris Cluster" as the Product and specify "check" in the Description field to locate Oracle Solaris Cluster patches that contain checks. Apply any patches that are not already installed on your cluster.

3. **Run basic validation checks.**

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

   `-v` Verbose mode

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

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

4. **Run interactive validation checks.**

   ```
   # 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.**

      ```
      # 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.

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

`-C` *checkID* 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.

```
# cluster check -v -k functional -C checkid -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.

Example 3–5 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 3–6 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
'/usr/cluster/bin/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 check -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

Next Steps Before you put the cluster into production, make a baseline recording of the cluster configuration for future diagnostic purposes. Go to "How to Record Diagnostic Data of the Cluster Configuration" on page 147.

▼ How to Record Diagnostic Data of the Cluster Configuration

After you finish configuring the global cluster but before you put it into production, use the Oracle Explorer utility to record baseline information about the cluster. This data can be used if there is a future need to troubleshoot the cluster.

1 **Become superuser.**

2 **Install Oracle Explorer software, if it is not already installed.**

The Services Tools Bundle contains the Oracle Explorer packages `SUNWexplo` and `SUNWexplu`. See `http://www.oracle.com/us/support/systems/premier/services-tools-bundle-sun-systems-163717.html` for software download and installation information.

3 **Run the `explorer` utility on each node in the cluster.**

Use the appropriate command for your platform. For example, to collect information on a Sun Fire T1000 server from Oracle, run the following command:

```
# explorer -i -w default,Tx000
```

For more information, see the `explorer`(1M) man page in the `/opt/SUNWexplo/man/man1m/` directory and *Oracle Explorer Data Collector User's Guide* which is available through Note 1153444.1 on My Oracle Support:

`https://support.oracle.com`

The `explorer` output file is saved in the `/opt/SUNWexplo/output/` directory as `explorer.`*hostid*`.`*hostname*`-`*date*`.tar.gz`.

4 **Save the files to a location that you can access if the entire cluster is down.**

5 **Send all `explorer` files by email to the Oracle Explorer database alias for your geographic location.**

Follow the procedures in *Oracle Explorer Data Collector User's Guide* to use FTP or HTTPS to submit Oracle Explorer files.

The Oracle Explorer database makes your `explorer` output available to Oracle technical support if the data is needed to help diagnose a technical problem with your cluster.

◆ ◆ ◆ CHAPTER 4

Configuring Solaris Volume Manager Software

Configure your local and multihost disks for Solaris Volume Manager software by using the procedures in this chapter, along with the planning information in "Planning Volume Management" on page 44. See your Solaris Volume Manager documentation for additional details.

Note – The Enhanced Storage module of Solaris Management Console is not compatible with Oracle Solaris Cluster software. Use the command-line interface or Oracle Solaris Cluster utilities to configure Solaris Volume Manager software.

The following sections are in this chapter:

Configuring Solaris Volume Manager Software

The following table lists the tasks that you perform to configure Solaris Volume Manager software for Oracle Solaris Cluster configurations. Complete the procedures in the order that is indicated.

TABLE 4–1 Task Map: Configuring Solaris Volume Manager Software

Task	Instructions
Plan the layout of your Solaris Volume Manager configuration.	"Planning Volume Management" on page 44
Create state database replicas on the local disks.	"How to Create State Database Replicas" on page 150
(*Optional*) Mirror file systems on the root disk.	"Mirroring the Root Disk" on page 150

▼ How to Create State Database Replicas

Perform this procedure on each node in the global cluster.

1 **Become superuser.**

2 **Create state database replicas on one or more local devices for each cluster node.**

Use the physical name (c*N*t*X*d*Y* s*Z*), not the device-ID name (d*N*), to specify the slices to use.

```
phys-schost# metadb -af slice-1 slice-2 slice-3
```

Tip – To provide protection of state data, which is necessary to run Solaris Volume Manager software, create at least three replicas for each node. Also, you can place replicas on more than one device to provide protection if one of the devices fails.

See the `metadb`(1M) man page and "Creating State Database Replicas" in *Solaris Volume Manager Administration Guide* for details.

3 **Verify the replicas.**

```
phys-schost# metadb
```

The `metadb` command displays the list of replicas.

Example 4–1 Creating State Database Replicas

The following example shows three state database replicas. Each replica is created on a different device.

```
phys-schost# metadb -af c0t0d0s7 c0t1d0s7 c1t0d0s7
phys-schost# metadb
flags           first blk       block count
     a       u       16          8192         /dev/dsk/c0t0d0s7
     a       u       16          8192         /dev/dsk/c0t1d0s7
     a       u       16          8192         /dev/dsk/c1t0d0s7
```

Next Steps To mirror file systems on the root disk, go to "Mirroring the Root Disk" on page 150.

Otherwise, go to "Creating Disk Sets in a Cluster" on page 161 to create Solaris Volume Manager disk sets.

Mirroring the Root Disk

Mirroring the root disk prevents the cluster node itself from shutting down because of a system disk failure. Four types of file systems can reside on the root disk. Each file-system type is mirrored by using a different method.

Use the following procedures to mirror each type of file system.

- "How to Mirror the Root (/) File System" on page 151
- "How to Mirror the Global Devices Namespace on a Dedicated Partition" on page 153
- "How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted" on page 156
- "How to Mirror File Systems That Can Be Unmounted" on page 158

Caution – For local disk mirroring, do not use `/dev/global` as the path when you specify the disk name. If you specify this path for anything other than cluster file systems, the system cannot boot.

How to Mirror the Root (/) File System

Use this procedure to mirror the root (`/`) file system.

Note – If the global-devices namespace is on a `lofi`-created file, this procedure includes the mirroring of the global-devices namespace.

1. **Become superuser.**

2. **Place the root slice in a single-slice (one-way) concatenation.**

   Specify the physical disk name of the root-disk slice (c*N*t*X*d*Y*s*Z*).

   ```
   phys-schost# metainit -f submirror1 1 1 root-disk-slice
   ```

3. **Create a second concatenation.**

   ```
   phys-schost# metainit submirror2 1 1 submirror-disk-slice
   ```

4. **Create a one-way mirror with one submirror.**

   ```
   phys-schost# metainit mirror -m submirror1
   ```

   Note – If the device is a local device to be used to mount a global-devices file system, `/global/.devices/node@`*nodeid*, the volume name for the mirror *must* be unique throughout the cluster.

5. **Set up the system files for the root (`/`) directory.**

   ```
   phys-schost# metaroot mirror
   ```

   This command edits the `/etc/vfstab` and `/etc/system` files so the system can be booted with the root (`/`) file system on a metadevice or volume. For more information, see the `metaroot(1M)` man page.

6 **Flush all file systems.**

```
phys-schost# lockfs -fa
```

This command flushes all transactions out of the log and writes the transactions to the master file system on all mounted UFS file systems. For more information, see the `lockfs(1M)` man page.

7 **Move any resource groups or device groups from the node.**

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

from-node
Specifies the name of the node from which to evacuate resource or device groups.

8 **Reboot the node.**

This command remounts the newly mirrored root (`/`) file system.

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

9 **Attach the second submirror to the mirror.**

```
phys-schost# metattach mirror submirror2
```

See the `metattach(1M)` man page for more information.

10 **If the disk that is used to mirror the root disk is physically connected to more than one node (multihosted), disable fencing for that disk.**

Disabling fencing for the device prevents unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

```
phys-schost# cldevice set -p default_fencing=nofencing submirror-disk
```

`-p`
Specifies a device property.

`default_fencing=nofencing`
Disables fencing for the specified device.

For more information about the `default_fencing` property, see the `cldevice(1CL)` man page.

11 **Record the alternate boot path for possible future use.**

If the primary boot device fails, you can then boot from this alternate boot device. See "Creating a RAID-1 Volume" in *Solaris Volume Manager Administration Guide* for more information about alternate boot devices.

```
phys-schost# ls -l /dev/rdsk/root-disk-slice
```

12 **Repeat Step 1 through Step 11 on each remaining node of the cluster.**

Ensure that each volume name for a mirror on which a global-devices file system, `/global/.devices/node@`*nodeid*, is to be mounted is unique throughout the cluster.

Example 4–2 Mirroring the Root (/) File System

The following example shows the creation of mirror `d0` on the node `phys-schost-1`, which consists of submirror `d10` on partition `c0t0d0s0` and submirror `d20` on partition `c2t2d0s0`. Device `c2t2d0` is a multihost disk, so fencing is disabled. The example also displays the alternate boot path for recording.

```
phys-schost# metainit -f d10 1 1 c0t0d0s0
d11: Concat/Stripe is setup
phys-schost# metainit d20 1 1 c2t2d0s0
d12: Concat/Stripe is setup
phys-schost# metainit d0 -m d10
d10: Mirror is setup
phys-schost# metaroot d0
phys-schost# lockfs -fa
phys-schost# clnode evacuate phys-schost-1
phys-schost# shutdown -g0 -y -i6
phys-schost# metattach d0 d20
d0: Submirror d20 is attachedphys-schost# cldevice set -p default_fencing=nofencing c2t2d0
phys-schost# ls -l /dev/rdsk/c2t2d0s0
lrwxrwxrwx  1 root     root          57 Apr 25 20:11 /dev/rdsk/c2t2d0s0
–> ../../devices/node@1/pci@1f,0/pci@1/scsi@3,1/disk@2,0:a,raw
```

Next Steps To mirror a global devices namespace that is configured on a dedicated partition, `/global/.devices/node@`*nodeid*, go to "How to Mirror the Global Devices Namespace on a Dedicated Partition" on page 153.

To mirror file systems than cannot be unmounted, go to "How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted" on page 156.

To mirror user-defined file systems, go to "How to Mirror File Systems That Can Be Unmounted" on page 158.

Otherwise, go to "Creating Disk Sets in a Cluster" on page 161 to create a disk set.

Troubleshooting Some of the steps in this mirroring procedure might cause an error message similar to `metainit: dg-schost-1: d1s0: not a metadevice`. Such an error message is harmless and can be ignored.

▼ How to Mirror the Global Devices Namespace on a Dedicated Partition

If you configured a dedicated partition for the global devices namespace, use this procedure to mirror the namespace, `/global/.devices/node@`*nodeid*`/`.

Note – Do not use this procedure if the global-devices namespace is on a `lofi`-based file. Instead, go to "How to Mirror the Root (/) File System" on page 151.

1 **Become superuser.**

2 **Place the global devices namespace slice in a single-slice (one-way) concatenation.**

Use the physical disk name of the disk slice (c*N*t*X*d*Y* s*Z*).

```
phys-schost# metainit -f submirror1 1 1 diskslice
```

3 **Create a second concatenation.**

```
phys-schost# metainit submirror2 1 1 submirror-diskslice
```

4 **Create a one-way mirror with one submirror.**

```
phys-schost# metainit mirror -m submirror1
```

Note – The volume name for a mirror on which a global-devices file system, `/global/.devices/node@`*nodeid*, is to be mounted *must* be unique throughout the cluster.

5 **Attach the second submirror to the mirror.**

This attachment starts a synchronization of the submirrors.

```
phys-schost# metattach mirror submirror2
```

6 **Edit the `/etc/vfstab` file entry for the `/global/.devices/node@`*nodeid* file system.**

Replace the names in the `device to mount` and `device to fsck` columns with the mirror name.

```
phys-schost# vi /etc/vfstab
#device        device        mount    FS     fsck    mount    mount
#to mount      to fsck       point    type   pass    at boot  options
#
/dev/md/dsk/mirror /dev/md/rdsk/mirror /global/.devices/node@nodeid ufs 2 no global
```

7 **Repeat Step 1 through Step 6 on each remaining node of the cluster.**

8 **Wait for the synchronization of the mirrors, started in Step 5, to be completed.**

Use the `metastat(1M)` command to view mirror status and to verify that mirror synchronization is complete.

```
phys-schost# metastat mirror
```

9 If the disk that is used to mirror the global devices namespace is physically connected to more than one node (multihosted), disable fencing for that disk.

Disabling fencing for the device prevents unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

```
phys-schost# cldevice set -p default_fencing=nofencing submirror-disk
```

`-p`
Specifies a device property.

`default_fencing=nofencing`
Disables fencing for the specified device.

For more information about the `default_fencing` property, see the `cldevice`(1CL) man page.

Example 4–3 Mirroring a Global Devices Namespace That Is Configured on a Dedicated Partition

The following example shows creation of mirror `d101`, which consists of submirror `d111` on partition `c0t0d0s3` and submirror `d121` on partition `c2t2d0s3`. The `/etc/vfstab` file entry for `/global/.devices/node@1` is updated to use the mirror name `d101`. Device `c2t2d0` is a multihost disk, so fencing is disabled.

```
phys-schost# metainit -f d111 1 1 c0t0d0s3
d111: Concat/Stripe is setup
phys-schost# metainit d121 1 1 c2t2d0s3
d121: Concat/Stripe is setup
phys-schost# metainit d101 -m d111
d101: Mirror is setup
phys-schost# metattach d101 d121
d101: Submirror d121 is attached
phys-schost# vi /etc/vfstab
#device        device        mount    FS     fsck    mount    mount
#to mount      to fsck       point    type   pass    at boot  options
#
/dev/md/dsk/d101 /dev/md/rdsk/d101 /global/.devices/node@1 ufs 2 no global
phys-schost# metastat d101
d101: Mirror
      Submirror 0: d111
         State: Okay
      Submirror 1: d121
         State: Resyncing
      Resync in progress: 15 % done
...
phys-schost# cldevice show phys-schost-3:/dev/rdsk/c2t2d0
=== DID Device Instances ===

DID Device Name:                                /dev/did/rdsk/d2
  Full Device Path:                               phys-schost-1:/dev/rdsk/c2t2d0
  Full Device Path:                               phys-schost-3:/dev/rdsk/c2t2d0
...

phys-schost# cldevicegroup show | grep dsk/d2
Device Group Name:                              dsk/d2
...
```

```
  Node List:                                      phys-schost-1, phys-schost-3
...
  localonly:                                      false
phys-schost# cldevicegroup remove-node -n phys-schost-3 dsk/d2
phys-schost# cldevice set -p default_fencing=nofencing c2t2d0
```

Next Steps To mirror file systems other than root (/) that cannot be unmounted, go to "How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted" on page 156.

To mirror user-defined file systems, go to "How to Mirror File Systems That Can Be Unmounted" on page 158

Otherwise, go to "Creating Disk Sets in a Cluster" on page 161 to create a disk set.

Troubleshooting Some of the steps in this mirroring procedure might cause an error message similar to `metainit: dg-schost-1: d1s0: not a metadevice`. Such an error message is harmless and can be ignored.

▼ How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted

Use this procedure to mirror file systems other than root (/) that cannot be unmounted during normal system usage, such as /usr, /opt, or swap.

1 Become superuser.

2 Place the slice on which an unmountable file system resides in a single-slice (one-way) concatenation.

Specify the physical disk name of the disk slice (c*N*t*X* d*Y*s*Z*).

```
phys-schost# metainit -f submirror1 1 1 diskslice
```

3 Create a second concatenation.

```
phys-schost# metainit submirror2 1 1 submirror-diskslice
```

4 Create a one-way mirror with one submirror.

```
phys-schost# metainit mirror -m submirror1
```

Note – The volume name for this mirror does *not* need to be unique throughout the cluster.

5 Repeat Step 1 through Step 4 for each remaining unmountable file system that you want to mirror.

6 **On each node, edit the `/etc/vfstab` file entry for each unmountable file system you mirrored.**

Replace the names in the `device to mount` and `device to fsck` columns with the mirror name.

```
phys-schost# vi /etc/vfstab
#device        device        mount    FS     fsck    mount    mount
#to mount      to fsck       point    type   pass    at boot  options
#
/dev/md/dsk/mirror /dev/md/rdsk/mirror /filesystem ufs 2 no global
```

7 **Move any resource groups or device groups from the node.**

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

from-node
Specifies the name of the node from which to move resource or device groups.

8 **Reboot the node.**

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

9 **Attach the second submirror to each mirror.**

This attachment starts a synchronization of the submirrors.

```
phys-schost# metattach mirror submirror2
```

10 **Wait for the synchronization of the mirrors, started in Step 9, to complete.**

Use the `metastat(1M)` command to view mirror status and to verify that mirror synchronization is complete.

```
phys-schost# metastat mirror
```

11 **If the disk that is used to mirror the unmountable file system is physically connected to more than one node (multihosted), disable fencing for that disk.**

Disabling fencing for the device prevents unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

```
phys-schost# cldevice set -p default_fencing=nofencing submirror-disk
```

`-p`
Specifies a device property.

`default_fencing=nofencing`
Disables fencing for the specified device.

For more information about the `default_fencing` property, see the `cldevice(1CL)` man page.

Example 4–4 Mirroring File Systems That Cannot Be Unmounted

The following example shows the creation of mirror `d1` on the node `phys-schost-1` to mirror `/usr`, which resides on `c0t0d0s1`. Mirror `d1` consists of submirror `d11` on partition `c0t0d0s1`

and submirror `d21` on partition `c2t2d0s1`. The `/etc/vfstab` file entry for `/usr` is updated to use the mirror name `d1`. Device `c2t2d0` is a multihost disk, so fencing is disabled.

```
phys-schost# metainit -f d11 1 1 c0t0d0s1
d11: Concat/Stripe is setup
phys-schost# metainit d21 1 1 c2t2d0s1
d21: Concat/Stripe is setup
phys-schost# metainit d1 -m d11
d1: Mirror is setup
phys-schost# vi /etc/vfstab
#device        device        mount    FS     fsck    mount    mount
#to mount      to fsck       point    type   pass    at boot  options
#
/dev/md/dsk/d1 /dev/md/rdsk/d1 /usr ufs  2       no global
...
phys-schost# clnode evacuate phys-schost-1
phys-schost# shutdown -g0 -y -i6
phys-schost# metattach d1 d21
d1: Submirror d21 is attached
phys-schost# metastat d1
d1: Mirror
      Submirror 0: d11
         State: Okay
      Submirror 1: d21
         State: Resyncing
      Resync in progress: 15 % done
...
phys-schost# cldevice show phys-schost-3:/dev/rdsk/c2t2d0
...
DID Device Name:                                /dev/did/rdsk/d2
phys-schost# cldevicegroup show dsk/d2
Device Group Name:                              dsk/d2
...
  Node List:                                      phys-schost-1, phys-schost-3
...
  localonly:                                      false
phys-schost# cldevicegroup remove-node -n phys-schost-3 dsk/d2
phys-schost# cldevice set -p default_fencing=nofencing c2t2d0
```

Next Steps To mirror user-defined file systems, go to "How to Mirror File Systems That Can Be Unmounted" on page 158.

Otherwise, go to "Creating Disk Sets in a Cluster" on page 161 to create a disk set.

Troubleshooting Some of the steps in this mirroring procedure might cause an error message similar to `metainit: dg-schost-1: d1s0: not a metadevice`. Such an error message is harmless and can be ignored.

▼ How to Mirror File Systems That Can Be Unmounted

Use this procedure to mirror user-defined file systems that can be unmounted. In this procedure, the nodes do not need to be rebooted.

1. **Become superuser.**

2. **Unmount the file system to mirror.**

   Ensure that no processes are running on the file system.

   ```
   phys-schost# umount /mount-point
   ```

   See the umount(1M) man page and "Mounting and Unmounting Oracle Solaris File Systems" in *System Administration Guide: Devices and File Systems* for more information.

3. **Place in a single-slice (one-way) concatenation the slice that contains a user-defined file system that can be unmounted.**

   Specify the physical disk name of the disk slice (c*N*t*X* d*Y*s*Z*).

   ```
   phys-schost# metainit -f submirror1 1 1 diskslice
   ```

4. **Create a second concatenation.**

   ```
   phys-schost# metainit submirror2 1 1 submirror-diskslice
   ```

5. **Create a one-way mirror with one submirror.**

   ```
   phys-schost# metainit mirror -m submirror1
   ```

   Note – The volume name for this mirror does *not* need to be unique throughout the cluster.

6. **Repeat Step 1 through Step 5 for each mountable file system to be mirrored.**

7. **On each node, edit the `/etc/vfstab` file entry for each file system you mirrored.**

   Replace the names in the `device to mount` and `device to fsck` columns with the mirror name.

   ```
   phys-schost# vi /etc/vfstab
   #device        device        mount    FS     fsck    mount    mount
   #to mount      to fsck       point    type   pass    at boot  options
   #
   /dev/md/dsk/mirror /dev/md/rdsk/mirror /filesystem ufs 2 no global
   ```

8. **Attach the second submirror to the mirror.**

   This attachment starts a synchronization of the submirrors.

   ```
   phys-schost# metattach mirror submirror2
   ```

9. **Wait for the synchronization of the mirrors, started in Step 8, to be completed.**

   Use the metastat(1M) command to view mirror status.

   ```
   phys-schost# metastat mirror
   ```

10 If the disk that is used to mirror the user-defined file system is physically connected to more than one node (multihosted), disable fencing for that disk.

Disabling fencing for the device prevents unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

```
phys-schost# cldevice set -p default_fencing=nofencing submirror-disk
```

`-p`
Specifies a device property.

`default_fencing=nofencing`
Disables fencing for the specified device.

For more information about the `default_fencing` property, see the `cldevice`(1CL) man page.

11 Mount the mirrored file system.

```
phys-schost# mount /mount-point
```

See the `mount`(1M) man page and "Mounting and Unmounting Oracle Solaris File Systems" in *System Administration Guide: Devices and File Systems* for more information.

Example 4–5 Mirroring File Systems That Can Be Unmounted

The following example shows creation of mirror `d4` to mirror `/export`, which resides on `c0t0d0s4`. Mirror `d4` consists of submirror `d14` on partition `c0t0d0s4` and submirror `d24` on partition `c2t2d0s4`. The `/etc/vfstab` file entry for `/export` is updated to use the mirror name `d4`. Device `c2t2d0` is a multihost disk, so fencing is disabled.

```
phys-schost# umount /export
phys-schost# metainit -f d14 1 1 c0t0d0s4
d14: Concat/Stripe is setup
phys-schost# metainit d24 1 1 c2t2d0s4
d24: Concat/Stripe is setup
phys-schost# metainit d4 -m d14
d4: Mirror is setup
phys-schost# vi /etc/vfstab
#device        device        mount    FS     fsck    mount    mount
#to mount      to fsck       point    type   pass    at boot  options
#
# /dev/md/dsk/d4 /dev/md/rdsk/d4 /export ufs 2 no    global
phys-schost# metattach d4 d24
d4: Submirror d24 is attached
phys-schost# metastat d4
d4: Mirror
       Submirror 0: d14
          State: Okay
       Submirror 1: d24
          State: Resyncing
       Resync in progress: 15 % done
...
phys-schost# cldevice show phys-schost-3:/dev/rdsk/c2t2d0
...
DID Device Name:                                /dev/did/rdsk/d2
```

```
phys-schost# cldevicegroup show dsk/d2
Device Group Name:                                dsk/d2
...
  Node List:                                        phys-schost-1, phys-schost-2
...
  localonly:                                        false
phys-schost# cldevicegroup remove-node -n phys-schost-3 dsk/d2
phys-schost# cldevice set -p default_fencing=nofencing c2t2d0
phys-schost# mount /export
```

Next Steps To create a disk set, go to "Creating Disk Sets in a Cluster" on page 161. Alternatively, if you will create a multi-owner disk set for use by Oracle Real Application Clusters, go to "How to Create a Multi-Owner Disk Set in Solaris Volume Manager for Sun Cluster for the Oracle RAC Database" in *Oracle Solaris Cluster Data Service for Oracle Real Application Clusters Guide*.

If you have sufficient disk sets for your needs, go to one of the following:

- If your cluster contains disk sets that are configured with exactly two disk enclosures and two nodes, you must add dual-string mediators. Go to "Configuring Dual-String Mediators" on page 169.
- If your cluster configuration does not require dual-string mediators, go to "How to Create Cluster File Systems" on page 173.

Troubleshooting Some of the steps in this mirroring procedure might cause an error message that is similar to `metainit: dg-schost-1: d1s0: not a metadevice`. Such an error message is harmless and can be ignored.

Creating Disk Sets in a Cluster

This section describes how to create disk sets for a cluster configuration. When you create a Solaris Volume Manager disk set in an Oracle Solaris Cluster environment, the disk set is automatically registered with the Oracle Solaris Cluster software as a device group of type `svm`. To create or delete an `svm` device group, you must use Solaris Volume Manager commands and utilities to create or delete the underlying disk set of the device group.

The following table lists the tasks that you perform to create disk sets. Complete the procedures in the order that is indicated.

TABLE 4–2 Task Map: Installing and Configuring Solaris Volume Manager Software

Task	Instructions
Create disk sets by using the `metaset` command.	"How to Create a Disk Set" on page 162
Add drives to the disk sets.	"How to Add Drives to a Disk Set" on page 164
(*Optional*) Repartition drives in a disk set to allocate space to different slices.	"How to Repartition Drives in a Disk Set" on page 166

TABLE 4–2 Task Map: Installing and Configuring Solaris Volume Manager Software *(Continued)*

Task	Instructions
List DID pseudo-driver mappings and define volumes in the `/etc/lvm/md.tab` files.	"How to Create an `md.tab` File" on page 166
Initialize the `md.tab` files.	"How to Activate Volumes" on page 168

▼ How to Create a Disk Set

Perform this procedure to create disk sets.

Before You Begin The disk set that you intend to create must meet one of the following requirements:

- If the disk set is configured with exactly two disk strings, the disk set must connect to exactly two nodes and use two or three mediator hosts. These mediator hosts must include the two hosts attached to the enclosures containing the disk set. See "Configuring Dual-String Mediators" on page 169 for details on how to configure dual-string mediators.
- If the disk set is configured with more than two disk strings, ensure that for any two disk strings S1 and S2, the sum of the number of drives on those strings exceeds the number of drives on the third string S3. Stated as a formula, the requirement is that count(S1) + count(S2) > count(S3).

1. **On each node in the cluster, run the `devfsadm(1M)` command.**

   You can run this command on all nodes in the cluster at the same time.

2. **From one node of the cluster, update the global-devices namespace.**

   ```
   phys-schost# cldevice populate
   ```

   See the `cldevice(1CL)` man page for more information.

3. **On each node, verify that the command has completed processing before you attempt to create any disk sets.**

   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.

   ```
   phys-schost# ps -ef | grep scgdevs
   ```

4. **Ensure that the local state database replicas exist.**

   For instructions, see "How to Create State Database Replicas" on page 150.

5. **Become superuser on the cluster node that will master the disk set.**

6 **Create the disk set.**

The following command creates the disk set and registers the disk set as an Oracle Solaris Cluster device group.

```
phys-schost# metaset -s setname -a -h node1 node2
```

`-s setname`
Specifies the disk set name.

`-a`
Adds (creates) the disk set.

`-h` *node1*
Specifies the name of the primary node to master the disk set.

node2
Specifies the name of the secondary node to master the disk set

---

Note – When you run the `metaset` command to configure a Solaris Volume Manager device group on a cluster, the command designates one secondary node by default. You can change the desired number of secondary nodes in the device group by using the `clsetup` utility after the device group is created. Refer to "Administering Device Groups" in *Oracle Solaris Cluster System Administration Guide* for more information about how to change the `numsecondaries` property.

---

7 **If you are configuring a replicated Solaris Volume Manager device group, set the replication property for the device group.**

```
phys-schost# cldevicegroup sync device-group-name
```

For more information about data replication, see Chapter 4, "Data Replication Approaches," in *Oracle Solaris Cluster System Administration Guide*.

8 **Verify the status of the new disk set.**

```
phys-schost# metaset -s setname
```

9 **As needed, set device group properties.**

```
phys-schost# cldevicegroup set -p name=value devicegroup
```

`-p`
Specifies a device-group property.

name
Specifies the name of a property.

value
Specifies the value or setting of the property.

devicegroup
Specifies the name of the device group. The device-group name is the same as the disk-set name.

See the `cldevicegroup(1CL)` for information about device-group properties.

Example 4–6 Creating a Disk Set

The following command creates two disk sets, `dg-schost-1` and `dg-schost-2`, with the nodes `phys-schost-1` and `phys-schost-2` specified as the potential primaries.

```
phys-schost# metaset -s dg-schost-1 -a -h phys-schost-1 phys-schost-2
phys-schost# metaset -s dg-schost-2 -a -h phys-schost-1 phys-schost-2
```

Next Steps Add drives to the disk set. Go to "Adding Drives to a Disk Set" on page 164.

Adding Drives to a Disk Set

When you add a drive to a disk set, the volume management software repartitions the drive as follows so that the state database for the disk set can be placed on the drive.

- A small portion of each drive is reserved for use by Solaris Volume Manager software. In volume table of contents (VTOC) labeled devices, slice 7 is used. In Extensible Firmware Interface (EFI) labeled devices, slice 6 is used. The remainder of the space on each drive is placed into slice 0.
- Drives are repartitioned when they are added to the disk set only if the target slice is not configured correctly.
- Any existing data on the drives is lost by the repartitioning.
- If the target slice starts at cylinder 0, and the drive partition is large enough to contain a state database replica, the drive is not repartitioned.

▼ How to Add Drives to a Disk Set

Before You Begin Ensure that the disk set has been created. For instructions, see "How to Create a Disk Set" on page 162.

1 Become superuser.

2 List the DID mappings.

```
phys-schost# 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
...
```

3 Become owner of the disk set.

```
phys-schost# cldevicegroup switch -n node devicegroup
```

`-n` *node*
Specifies the node to take ownership of the device group.

devicegroup
Specifies the device group name, which is the same as the disk set name.

4 Add the drives to the disk set.

Use the full DID path name.

```
phys-schost# 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.

---

5 Verify the status of the disk set and drives.

```
phys-schost# metaset -s setname
```

Example 4–7 Adding Drives to a Disk Set

The `metaset` command adds the drives `/dev/did/rdsk/d1` and `/dev/did/rdsk/d2` to the disk set `dg-schost-1`.

```
phys-schost# metaset -s dg-schost-1 -a /dev/did/rdsk/d1 /dev/did/rdsk/d2
```

Next Steps To repartition drives for use in volumes, go to "How to Repartition Drives in a Disk Set" on page 166.

Otherwise, go to "How to Create an `md.tab` File" on page 166 to define metadevices or volumes by using an `md.tab` file.

How to Repartition Drives in a Disk Set

The `metaset(1M)` command repartitions drives in a disk set so that a small portion of each drive is reserved for use by Solaris Volume Manager software. In volume table of contents (VTOC) labeled devices, slice 7 is used. In Extensible Firmware Interface (EFI) labeled devices, slice 6 is used. The remainder of the space on each drive is placed into slice 0. To make more effective use of the drive, use this procedure to modify the disk layout. If you allocate space to VTOC slices 1 through 6 or EFI slices 1 through 5, you can use these slices when you set up Solaris Volume Manager volumes.

1 **Become superuser.**

2 **Use the `format` command to change the disk partitioning for each drive in the disk set.**

When you repartition a drive, you must meet the following conditions to prevent the `metaset(1M)` command from repartitioning the drive.

- Create slice 7 for VTOC or slice 6 for EFI starting at cylinder 0, large enough to hold a state database replica. See your Solaris Volume Manager administration guide to determine the size of a state database replica for your version of the volume-manager software.
- Set the `Flag` field in the target slice to `wu` (read-write, unmountable). Do not set it to read-only.
- Do not allow the target slice to overlap any other slice on the drive.

See the `format(1M)` man page for details.

Next Steps Define volumes by using an `md.tab` file. Go to "How to Create an `md.tab` File" on page 166.

How to Create an `md.tab` File

Create an `/etc/lvm/md.tab` file on each node in the cluster. Use the `md.tab` file to define Solaris Volume Manager volumes for the disk sets that you created.

Note – If you are using local volumes, ensure that local volume names are distinct from the device-ID names that are used to form disk sets. For example, if the device-ID name `/dev/did/dsk/d3` is used in a disk set, do not use the name `/dev/md/dsk/d3` for a local volume. This requirement does not apply to shared volumes, which use the naming convention `/dev/md/`*setname*`/{r}dsk/d#`.

---

1 **Become superuser.**

2 List the DID mappings for reference when you create your `md.tab` file.

Use the full DID device names in the `md.tab` file in place of the lower-level device names (c*N* t*X*d*Y*). The DID device name takes the form `/dev/did/rdsk/d`*N*.

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

=== 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
...
```

3 Create an `/etc/lvm/md.tab` file and edit it with your preferred text editor.

Note – If you have existing data on the drives that will be used for the submirrors, you must back up the data before volume setup. Then restore the data onto the mirror.

To avoid possible confusion between local volumes on different nodes in a cluster environment, use a naming scheme that makes each local volume name unique throughout the cluster. For example, for node 1 choose names from `d100` to `d199`. And for node 2 use `d200` to `d299`.

See your Solaris Volume Manager documentation and the `md.tab`(4) man page for details about how to create an `md.tab` file.

Example 4–8 Sample `md.tab` File

The following sample `md.tab` file defines the disk set that is named `dg-schost-1`. The ordering of lines in the `md.tab` file is not important.

```
dg-schost-1/d0 -m dg-schost-1/d10 dg-schost-1/d20
    dg-schost-1/d10 1 1 /dev/did/rdsk/d1s0
    dg-schost-1/d20 1 1 /dev/did/rdsk/d2s0
```

The sample `md.tab` file is constructed as follows.

1. The first line defines the device `d0` as a mirror of volumes `d10` and `d20`. The `-m` signifies that this device is a mirror device.

   ```
   dg-schost-1/d0 -m dg-schost-1/d0 dg-schost-1/d20
   ```

2. The second line defines volume `d10`, the first submirror of `d0`, as a one-way stripe.

   ```
   dg-schost-1/d10 1 1 /dev/did/rdsk/d1s0
   ```

3. The third line defines volume `d20`, the second submirror of `d0`, as a one-way stripe.

   ```
   dg-schost-1/d20 1 1 /dev/did/rdsk/d2s0
   ```

Next Steps Activate the volumes that are defined in the `md.tab` files. Go to "How to Activate Volumes" on page 168.

How to Activate Volumes

Perform this procedure to activate Solaris Volume Manager volumes that are defined in `md.tab` files.

1. **Become superuser.**

2. **Ensure that `md.tab` files are located in the `/etc/lvm` directory.**

3. **Ensure that you have ownership of the disk set on the node where the command will be executed.**

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

   ```
   phys-schost# cldevicegroup switch -n node devicegroup
   ```

   `-n` *node*
   Specifies the node that takes ownership.

   devicegroup
   Specifies the disk set name.

5. **Activate the disk set's volumes, which are defined in the `md.tab` file.**

   ```
   phys-schost# metainit -s setname -a
   ```

   `-s` *setname*
   Specifies the disk set name.

   `-a`
   Activates all volumes in the `md.tab` file.

6. **Repeat Step 3 through Step 5 for each disk set in the cluster.**

   If necessary, run the `metainit(1M)` command from another node that has connectivity to the drives. This step is required for cluster-pair topologies, where the drives are not accessible by all nodes.

7. **Check the status of the volumes.**

   ```
   phys-schost# metastat -s setname
   ```

   See the `metastat(1M)` man page for more information.

8. **(Optional) Capture the disk partitioning information for future reference.**

   ```
   phys-schost# prtvtoc /dev/rdsk/cNtXdYsZ > filename
   ```

Store the file in a location outside the cluster. If you make any disk configuration changes, run this command again to capture the changed configuration. If a disk fails and needs replacement, you can use this information to restore the disk partition configuration. For more information, see the `prtvtoc(1M)` man page.

9 (Optional) Make a backup of your cluster configuration.

An archived backup of your cluster configuration facilitates easier recovery of the your cluster configuration. For more information, see "How to Back Up the Cluster Configuration" in *Oracle Solaris Cluster System Administration Guide*.

Example 4–9 Activating Volumes in the `md.tab` File

In the following example, all volumes that are defined in the `md.tab` file for disk set `dg-schost-1` are activated.

```
phys-schost# metainit -s dg-schost-1 -a
```

Next Steps If your cluster contains disk sets that are configured with exactly two disk enclosures and two nodes, add dual-string mediators. Go to "Configuring Dual-String Mediators" on page 169.

Otherwise, go to "How to Create Cluster File Systems" on page 173 to create a cluster file system.

Configuring Dual-String Mediators

This section provides information and procedures to configure dual-string mediator hosts.

A single *disk string* consists of a disk enclosure, its physical drives, cables from the enclosure to the node or nodes, and the interface adapter cards. A dual-string disk set includes disks in two disk strings, and is attached to exactly two nodes. If a single disk string in a dual-string disk set fails, such that exactly half the Solaris Volume Manager replicas remain available, the disk set will stop functioning. Dual-string mediators are therefore required for all Solaris Volume Manager dual-string disk sets. The use of mediators enables the Oracle Solaris Cluster software to ensure that the most current data is presented in the instance of a single-string failure in a dual-string configuration.

A *dual-string mediator*, or mediator host, is a cluster node that stores mediator data. Mediator data provides information about the location of other mediators and contains a commit count that is identical to the commit count that is stored in the database replicas. This commit count is used to confirm that the mediator data is in sync with the data in the database replicas.

The following table lists the tasks that you perform to configure dual-string mediator hosts. Complete the procedures in the order that is indicated.

TABLE 4–3 Task Map: Installing and Configuring Solaris Volume Manager Software

Task	Instructions
Configure dual-string mediator hosts.	"Requirements for Dual-String Mediators" on page 170 "How to Add Mediator Hosts" on page 170
Check the status of mediator data.	"How to Check the Status of Mediator Data" on page 171
If necessary, fix bad mediator data.	"How to Fix Bad Mediator Data" on page 172

Requirements for Dual-String Mediators

The following rules apply to dual-string configurations that use mediators.

- Disk sets must be configured with two or three mediator hosts. Two of these mediator hosts must be the same two cluster nodes that are used for the disk set. The third may be another node in the cluster or a non-clustered host on the cluster's public network, such as a quorum server.
- Mediators cannot be configured for disk sets that do not meet the two-string and two-host criteria.

These rules do not require that the entire cluster consist of only two nodes. An N+1 cluster and many other topologies are permitted under these rules.

▼ How to Add Mediator Hosts

Perform this procedure if your configuration requires dual-string mediators.

Before You Begin

- If you will use a third mediator host for a dual-string disk set, and that host does not already have disk sets configured, perform the following steps:
- Add the entry `root` to the `sysadmin`group in the `/etc/group` file
- Create a dummy disk set by using the command:

  ```
  phys-schost-3# metaset -s dummy-diskset-name -a -h hostname
  ```

1 Become superuser on the node that currently masters the disk set to which you intend to add mediator hosts.

2 Add each node with connectivity to the disk set as a mediator host for that disk set.

```
phys-schost# metaset -s setname -a -m mediator-host-list
```

`-s` *setname*
Specifies the disk set name.

`-a`
Adds to the disk set.

`-m` *mediator-host-list*
Specifies the name of the node to add as a mediator host for the disk set.

See the `mediator(7D)` man page for details about mediator-specific options to the `metaset` command.

Example 4–10 Adding Mediator Hosts

The following example adds the nodes `phys-schost-1` and `phys-schost-2` as mediator hosts for the disk set `dg-schost-1`. If needed, repeat the command a third time for a third mediator host. All commands are run from the node that masters the disk set for which you are adding mediator hosts, in this case `phys-schost-1`.

```
phys-schost# metaset -s dg-schost-1 -a -m phys-schost-1
phys-schost# metaset -s dg-schost-1 -a -m phys-schost-2
phys-schost# metaset -s dg-schost-1 -a -m phys-schost-3
```

Next Steps Check the status of mediator data. Go to "How to Check the Status of Mediator Data" on page 171.

▼ How to Check the Status of Mediator Data

Before You Begin Ensure that you have added mediator hosts as described in "How to Add Mediator Hosts" on page 170.

1 Display the status of the mediator data.

```
phys-schost# medstat -s setname
```

`-s` *setname*
Specifies the disk set name.

See the `medstat(1M)` man page for more information.

2 If `Bad` is the value in the Status field of the `medstat` output, repair the affected mediator host.

Go to "How to Fix Bad Mediator Data" on page 172.

Next Steps Go to "How to Create Cluster File Systems" on page 173 to create a cluster file system.

How to Fix Bad Mediator Data

Perform this procedure to repair bad mediator data.

1. **Identify all mediator hosts with bad mediator data.**

   See "How to Check the Status of Mediator Data" on page 171.

2. **Become superuser on the node that owns the affected disk set.**

3. **Remove all mediator hosts with bad mediator data from all affected disk sets.**

   ```
   phys-schost# metaset -s setname -d -m mediator-host-list
   ```

   `-s` *setname*
   Specifies the disk set name.

   `-d`
   Deletes from the disk set.

   `-m` *mediator-host-list*
   Specifies the name of the node to remove as a mediator host for the disk set.

4. **Restore each mediator host that you removed in Step 3.**

   ```
   phys-schost# metaset -s setname -a -m mediator-host-list
   ```

   `-a`
   Adds to the disk set.

   `-m` *mediator-host-list*
   Specifies the name of the node to add as a mediator host for the disk set.

   See the `mediator(7D)` man page for details about mediator-specific options to the `metaset` command.

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- To create cluster file systems, go to "How to Create Cluster File Systems" on page 173.
- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 191.
- Install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation that is supplied with the application software and the *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

◆ ◆ ◆ CHAPTER 5

Creating a Cluster File System

This chapter describes how to create a cluster file system to support data services.

Note – Alternatively, you can use a highly available local file system to support a data service. For information about choosing between creating a cluster file system or a highly available local file system to support a particular data service, see the manual for that data service. For general information about creating a highly available local file system, see "Enabling Highly Available Local File Systems" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

The following sections are in this chapter:

Creating Cluster File Systems

This section provides procedures to create cluster file systems to support data services.

▼ How to Create Cluster File Systems

Perform this procedure for each cluster file system that you want to create. Unlike a local file system, a cluster file system is accessible from any node in the global cluster.

Before You Begin Perform the following tasks:

- Ensure that you installed software packages for the Oracle Solaris OS, Oracle Solaris Cluster framework, and other products as described in "Installing the Software" on page 49.
- Ensure that you established the new cluster or cluster node as described in "Establishing a New Global Cluster or New Global-Cluster Node" on page 72.

- If you are using Solaris Volume Manager software, ensure that volume-management software is configured. For procedures, see "Configuring Solaris Volume Manager Software" on page 149.
- Determine the mount options to use for each cluster file system that you want to create. See "Choosing Mount Options for UFS Cluster File Systems" on page 42.

1 Become superuser on any node in the cluster.

Perform this procedure from the global zone if non-global zones are configured in the cluster.

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

---

2 Create a file system.

Caution – Any data on the disks is destroyed when you create a file system. Be sure that you specify the correct disk device name. If you specify the wrong device name, you might erase data that you did not intend to delete.

- **For a UFS file system, use the `newfs(1M)` command.**

  ```
  phys-schost# newfs raw-disk-device
  ```

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

Volume Manager	Sample Disk Device Name	Description
Solaris Volume Manager	`/dev/md/nfs/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/mountpoint/
```

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

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

4 On each node in the cluster, add an entry to the `/etc/vfstab` file for the mount point.

See the `vfstab(4)` man page for details.

Note – If non-global zones are configured in the cluster, ensure that you mount cluster file systems in the global zone on a path in the global zone's root directory.

---

a. In each entry, specify the required mount options for the type of file system that you use.

b. To automatically mount the cluster file system, set the `mount at boot` field to `yes`.

c. Ensure that, for each cluster file system, the information in its `/etc/vfstab` entry is identical on each node.

d. Ensure that the entries in each node's `/etc/vfstab` file list devices in the same order.

e. Check the boot order dependencies of the file systems.

For example, consider the scenario where `phys-schost-1` mounts disk device `d0` on `/global/oracle/`, and `phys-schost-2` mounts disk device `d1` on `/global/oracle/logs/`. With this configuration, `phys-schost-2` can boot and mount `/global/oracle/logs/` only after `phys-schost-1` boots and mounts `/global/oracle/`.

5 On any node in the cluster, run the configuration check utility.

```
phys-schost# cluster check -k vfstab
```

The configuration check utility verifies that the mount points exist. The utility also verifies that `/etc/vfstab` file entries are correct on all nodes of the cluster. If no errors occur, nothing is returned.

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

6 Mount the cluster file system.

For UFS and QFS, 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.

Cluster file systems are accessible from both the global zone and the non-global zone.

Example 5–1 Creating a UFS Cluster File System

The following example creates a UFS cluster file system on the Solaris Volume Manager volume `/dev/md/oracle/rdsk/d1`. An entry for the cluster file system is added to the `vfstab` file on each node. Then from one node the `cluster check` command is run. After configuration check processing is completes successfully, the cluster file system is mounted from one node and verified on all nodes.

```
phys-schost# newfs /dev/md/oracle/rdsk/d1
...
phys-schost# mkdir -p /global/oracle/d1
phys-schost# vi /etc/vfstab
#device           device        mount   FS      fsck    mount   mount
#to mount         to fsck       point   type    pass    at boot options
#
/dev/md/oracle/dsk/d1 /dev/md/oracle/rdsk/d1 /global/oracle/d1 ufs 2 yes global,logging
...
phys-schost# cluster check -k vfstab
phys-schost# mount /global/oracle/d1
phys-schost# mount
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/largefiles
on Sun Oct 3 08:56:16 2005
```

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 191.
- Install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation that is supplied with the application software and the *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

Creating Oracle ACFS File Systems

This section provides procedures to create an Oracle Automatic Storage Management Cluster File System (Oracle ACFS) file system to support data services. This file system can be used as a general-purpose file system or as the database home for an Oracle database. An Oracle ACFS file system is supported for use in the global cluster and in zone clusters.

Note – A minimum of Oracle ASM version 11*g* release 2 is required.

The following table lists the tasks to perform to create an Oracle ACFS file system. Complete the procedures in the order that is indicated.

TABLE 5–1 Task Map: Creating Oracle ACFS File Systems

Task	Instructions
Register and configure the framework resource groups.	"How to Register and Configure the Framework Resource Groups" on page 177
Create an Oracle ACFS file system.	"How to Create an Oracle ACFS File System" on page 180
Register and configure the scalable device-group resource group.	"How to Register and Configure the Scalable Device-Group Resource Group" on page 182
Register and configure the Oracle ASM resource groups.	"How to Register and Configure the Oracle ASM Resource Group" on page 183
Configure interoperability between Oracle Grid Infrastructure and Oracle Solaris Cluster	"How to Create an Oracle Grid Infrastructure Resource for Interoperation With Oracle Solaris Cluster" on page 187
Register and configure the Oracle ACFS proxy resource group.	"How to Register and Configure the Oracle ACFS Proxy Resource Group" on page 185

▼ How to Register and Configure the Framework Resource Groups

Use this procedure to configure a scalable Oracle RAC framework resource group and, if Solaris Volume Manager for Sun Cluster is used, a multiple-owner volume-manager framework resource group.

Perform all steps from one node of the global cluster.

1 Become superuser or assume a role that provides `solaris.cluster.admin` and `solaris.cluster.modify` RBAC authorizations.

2 Create a scalable Oracle RAC framework resource group.

This framework resource group is used regardless of whether the file system is for general purpose or for a database home.

a. Create the Oracle RAC framework resource group.

```
# clresourcegroup create -n nodelist \
  -p maximum_primaries=num-in-list \
  -p desired_primaries=num-in-list \
  [-p rg_description="description"] \
  -p rg_mode=Scalable rac-fmwk-rg
```

`-n nodelist=`*nodelist*
Specifies a comma-separated list of cluster nodes on which the data service is to be enabled. The data service software packages must be installed on each node in this list.

`-p maximum_primaries=`*num-in-list*
Specifies the number of nodes on which the data service is to be enabled. This number must equal the number of nodes in *nodelist*.

`-p desired_primaries=`*num-in-list*
Specifies the number of nodes on which the data service is to be enabled. This number must equal the number of nodes in *nodelist*.

`-p rg_description="`*description*"
Specifies an optional brief description of the resource group. This description is displayed when you use Oracle Solaris Cluster maintenance commands to obtain information about the resource group.

`-p rg_mode=Scalable`
Specifies that the resource group is scalable.

rac-fmwk-rg
Specifies the name that you are assigning to the Oracle RAC framework resource group.

b. Register the `SUNW.rac_framework` resource type.

```
# clresourcetype register SUNW.rac_framework
```

c. Add an instance of the `SUNW.rac_framework` resource type to the Oracle RAC framework resource group.

```
# clresource create -g rac-fmwk-rg -t SUNW.rac_framework rac-fmwk-rs
```

`-g` *rac-fmwk-rg* — Specifies the Oracle RAC framework resource group to which you are adding the resource.

rac-fmwk-rs — Specifies the name that you are assigning to the `SUNW.rac_framework` resource.

d. Register the Oracle Clusterware framework resource type

```
# clresourcetype register SUNW.crs_framework
```

e. Add an instance of the `SUNW.crs_framework` resource type to the Oracle RAC framework resource group.

```
# clresource create -g rac-fmwk-rg \
-p resource_dependencies=rac-fmwk-rs \
-t SUNW.crs_framework \
crs-fmwk-rs
```

crs-fmwk-rs — Specifies the name that you are assigning to the `SUNW.crs_framework` resource.

3 If you are using Solaris Volume Manager for Sun Cluster, create a scalable multiple-owner volume-manager framework resource group.

This resource group manages Solaris Volume Manager for Sun Cluster.

If you are using hardware RAID, skip to Step 4.

a. Create the resource group

```
# clresourcegroup create -n nodelist -S vucmm-fmwk-rg
```

`-n nodelist=`*nodelist*
Specifies the same node list that you configured for the scalable Oracle RAC framework resource group.

vucmm-fmwk-rg
Specifies the name that you are assigning to the multiple-owner volume-manager framework resource group.

b. Register the `SUNW.vucmm_framework` resource type.

```
# clresourcetype register SUNW.vucmm_framework
```

c. Add an instance of the `SUNW.vucmm_framework` resource type to the multiple-owner volume-manager resource group.

```
# clresource create -g vucmm-fmwk-rg -t SUNW.vucmm_framework vucmm-fmwk-rs
```

`-g` *vucmm-fmwk-rg*	Specifies the multiple-owner volume-manager resource group to which you are adding the resource.
vucmm-fmwk-rs	Specifies the name that you are assigning to the `SUNW.vucmm_framework` resource.

d. Register the `SUNW.vucmm_svm` resource type.

```
# clresourcetype register SUNW.vucmm_svm
```

e. Add an instance of the `SUNW.vucmm_svm` resource type to the `SUNW.vucmm_framework` resource group.

Ensure that this instance depends on the `SUNW.vucmm_framework` resource that you created.

```
# clresource create -g vucmm-fmwk-rg \
  -t svm-rt \
  -p resource_dependencies=vucmm-fmwk-rs svm-rs
```

`-g` *vucmm-fmwk-rg*
Specifies the multiple-owner volume-manager resource group.

`-p resource_dependencies=`*vucmm-fmwk-rs*
Specifies that this instance depends on the `SUNW.vucmm_framework` resource.

svm-rs
Specifies the name that you are assigning to the `SUNW.vucmm_svm` resource.

4 Bring online and in a managed state the Oracle RAC framework resource group, the multiple-owner volume-manager framework resource group if used, and their resources.

```
# clresourcegroup online -eM rac-fmwk-rg [vucmm-fmwk-rg]
```

rac-fmwk-rg	Specifies that the Oracle RAC framework resource group is to be moved to the `MANAGED` state and brought online.
vucmm-fmwk-rg	Specifies that the multiple-owner volume-manager framework resource group is to be moved to the `MANAGED` state and brought online.

Next Steps Go to "How to Create an Oracle ACFS File System" on page 180.

▼ How to Create an Oracle ACFS File System

Use this procedure to create an Oracle ACFS file system. Perform all steps from the one node of the global cluster.

Before You Begin

- Ensure that the framework resource groups are configured. See "How to Register and Configure the Framework Resource Groups" on page 177.
- Observe the following guidelines and restrictions for configuring an Oracle ACFS file system in an Oracle Solaris Cluster configuration.
  - Ensure that a minimum of Oracle ASM version 11*g* version 2 is installed.
  - An Oracle ACFS file system is supported in a global cluster and a zone cluster, but not in individual non-global zones.
  - An Oracle ACFS file system must be managed by an Oracle Clusterware resource.

1 Create an Oracle ACFS file system.

Follow procedures in "Creating an Oracle ACFS File System" in *Oracle Automatic Storage Management Administrator's Guide*.

Observe the following special instructions:

- Do not register the Oracle ACFS file-system mount point with the Oracle ACFS registry, if you will use the file system as a general purpose file system. Only register the mount point with the Oracle ACFS registry if the file system will be used as the database home.
- Configure the Oracle ACFS file system only in the global zone. To use the file system in a zone cluster, you direct mount the file system in the zone cluster.
- Configure the Oracle ACFS resource on the same nodes on which you configure the clustered Oracle ASM disk-group proxy resource for the Oracle ASM disk group that contains the Oracle ACFS volume.

2 Create a mount point for the Oracle ACFS file system.

Note – For a zone cluster, create the mount point under the zone root path:

```
# mkdir -p /zonepath/root/path-to-filesystem
```

3 For a zone cluster, ensure that the zone cluster is online.

```
# clzonecluster status zonecluster
```

4 Start and mount the Oracle ACFS file system and verify the status.

```
# /Grid_home/bin/srvctl add filesystem -d /dev/asm/volume-dev-path -v volume-name \
-g device-group-name -m mount-point
# /Grid_home/bin/srvctl start filesystem -d /dev/asm/volume-dev-path
# /Grid_home/bin/srvctl status filesystem -d /dev/asm/volume-dev-path
```

5 *(Oracle ACFS 11g release 2 only)* **For a zone cluster, add the file system to the zone cluster.**

Perform these steps from the global zone of one node.

a. Add the Oracle ACFS file system to the zone cluster.

```
# clzonecluster configure zonecluster
clzc:zonecluster> add fs
clzc:zonecluster:fs> set dir=mountpoint
clzc:zonecluster:fs> set special=/dev/asm/volume-dev-path
clzc:zonecluster:fs> set type=acfs
clzc:zonecluster:fs> end
clzc:zonecluster> exit
```

b. Verify that the Oracle ACFS file system is added to the zone cluster.

```
# clzonecluster show zonecluster
...
   Resource Name:                fs
   dir:                          mountpoint
   special                       volume
   raw:
   type:                         acfs
   options:                      []
   cluster-control:              true
...
```

Next Steps If you are using Solaris Volume Manager for Sun Cluster, go to "How to Register and Configure the Scalable Device-Group Resource Group" on page 182.

Otherwise, go to "How to Register and Configure the Oracle ASM Resource Group" on page 183

▼ How to Register and Configure the Scalable Device-Group Resource Group

If you use Solaris Volume Manager for Sun Cluster, register and configure a scalable device-group resource group. Perform all steps from one node of the global cluster.

If your configuration does not use Solaris Volume Manager for Sun Cluster, do not perform this procedure. Proceed to "How to Register and Configure the Oracle ASM Resource Group" on page 183.

Before You Begin Ensure that the Oracle ACFS file system is created. See "How to Create an Oracle ACFS File System" on page 180.

1. **Become superuser or assume a role that provides `solaris.cluster.admin` and `solaris.cluster.modify` RBAC authorizations.**

2. **Create a scalable resource group to contain the scalable device-group resource.**

   Set a strong positive affinity by the resource group for the multiple-owner volume-manager framework resource group.

   ```
   # clresourcegroup create -p nodelist=nodelist \
   -p desired_primaries=num-in-list \
   -p maximum_primaries=num-in-list \
   -p rg_affinities=++vucmm-fmwk-rg \
   [-p rg_description="description"] \
   -p rg_mode=Scalable \
   scal-dg-rg
   ```

3. **Register the `SUNW.ScalDeviceGroup` resource type.**

   ```
   # clresourcetype register SUNW.ScalDeviceGroup
   ```

4. **Add an instance of the `SUNW.ScalDeviceGroup` resource type to the `SUNW.ScalDeviceGroup` resource group.**

   Set a strong dependency for the instance of `SUNW.ScalDeviceGroup` on the *svm-rs* resource in the `SUNW.vucmm_svm` framework resource group. Limit the scope of this dependency to only the node where the `SUNW.ScalDeviceGroup` resource is running.

   ```
   # clresource create -t SUNW.ScalDeviceGroup -g scal-dg-rg \
   -p resource_dependencies=svm-rs{local_node} \
   -p diskgroupname=disk-group  scal-dg-rs
   ```

5. **Bring online and in a managed state the scalable device-group resource group.**

   ```
   # clresourcegroup online -eM scal-dg-rg
   ```

6. **Set an offline-restart dependency on *crs-fmwk-rs* by *scal-dg-rs*.**

   ```
   # clresource set -p resource_dependencies_offline_restart=scal-dg-rs  crs-fmwk-rs
   ```

Next Steps Go to "How to Register and Configure the Oracle ASM Resource Group" on page 183.

How to Register and Configure the Oracle ASM Resource Group

Use this procedure to register and configure the Oracle Automatic Storage Management (Oracle ASM) resource group. Perform all steps from one node of the global cluster.

Before You Begin

- Ensure that the framework resource groups are created. See "How to Register and Configure the Framework Resource Groups" on page 177.
- If you use Solaris Volume Manager for Sun Cluster, ensure that the scalable device–group resource group is created. See "How to Register and Configure the Scalable Device-Group Resource Group" on page 182.

1. **Become superuser or assume a role that provides `solaris.cluster.admin` and `solaris.cluster.modify` RBAC authorizations.**

2. **Register the Oracle ASM resource types for the data service.**

   a. **Register the scalable Oracle ASM instance proxy resource type.**

      ```
      # clresourcetype register SUNW.scalable_asm_instance_proxy
      ```

   b. **Register the Oracle ASM disk-group resource type.**

      ```
      # clresourcetype register SUNW.scalable_asm_diskgroup_proxy
      ```

3. **Create resource groups** *asm-inst-rg* **and** *asm-dg-rg*.

   ```
   # clresourcegroup create -S asm-inst-rg asm-dg-rg
   ```

   asm-inst-rg
   Specifies the name of the Oracle ASM instance resource group.

   asm-dg-rg
   Specifies the name of the Oracle ASM disk-group resource group.

4. **Set a strong positive affinity on** *rac-fmwk-rg* **by** *asm-inst-rg*.

   ```
   # clresourcegroup set -p rg_affinities=++rac-fmwk-rg asm-inst-rg
   ```

5. **Set a strong positive affinity by** *asm-dg-rg*.

   - **If you are using hardware RAID, set the affinity on** *asm-inst-rg*

     ```
     # clresourcegroup set -p rg_affinities=++asm-inst-rg asm-dg-rg
     ```

- **If you are using Solaris Volume Manager for Sun Cluster, set the affinity on** *scal-dg-rg* **and** *asm-inst-rg***.**

  ```
  # clresourcegroup set -p rg_affinities=++asm-inst-rg,++scal-dg-rg  asm-dg-rg
  ```

6 Create a `SUNW.scalable_asm_instance_proxy` resource and set the resource dependencies.

```
# clresource create -g asm-inst-rg \
-t SUNW.scalable_asm_instance_proxy \
-p ORACLE_HOME=Grid_home \
-p CRS_HOME=Grid_home \
-p "ORACLE_SID{node1}"=+ASM1 \
-p "ORACLE_SID{node2}"=+ASM2 \
-p resource_dependencies_offline_restart=crs-fmwk-rs \
-d asm-inst-rs
```

`-t SUNW.scalable_asm_instance_proxy`
: Specifies the type of the resource to add.

`-p ORACLE_HOME=`*Grid_home*
: Sets the path to the Oracle Grid Infrastructure home directory, where Oracle ASM is installed.

`-p CRS_HOME=`*Grid_home*
: Sets the path to the Oracle Grid Infrastructure home directory, where Oracle Clusterware is installed.

`-p ORACLE_SID=+ASM`*n*
: Sets the Oracle ASM system identifier.

`-d` *asm-inst-rs*
: Specifies the name of the Oracle ASM instance resource that you create.

7 Bring online the *asm-inst-rg* **resource group in a managed state on a cluster node.**

```
# clresourcegroup online -eM asm-inst-rg
```

8 Add an Oracle ASM disk-group resource to the *asm-dg-rg* **resource group.**

- **For hardware RAID, use the following command:**

  ```
  # clresource create -g asm-dg-rg \
  -t SUNW.scalable_asm_diskgroup_proxy \
  -p asm_diskgroups=dg[,dg...] \
  -p resource_dependencies_offline_restart=asm-inst-rs \
  -d asm-dg-rs
  ```

- **For Solaris Volume Manager for Sun Cluster, use the following command:**

  ```
  # clresource create -g asm-dg-rg \
  -t SUNW.scalable_asm_diskgroup_proxy \
  -p asm_diskgroups=dg[,dg...] \
  -p resource_dependencies_offline_restart=asm-inst-rs,scal-dg-rs \
  -d asm-dg-rs
  ```

9 **Bring online the *asm-dg-rg* resource group in a managed state on a cluster node.**

```
# clresourcegroup online -eM asm-dg-rg
```

10 **For a zone cluster, from the global zone of one node, create a `SUNW.wait_zc_boot` resource group.**

If the Oracle ACFS file system will not be used by a zone cluster, omit this step.

```
# clresourcetype register SUNW.wait_zc_boot
# clresourcegroup create -S scal-wait-zc-rg
# clresource create -g scal-wait-zc-rg \
-t SUNW.wait_zc_boot \
-p zcname=zonecluster \
wait-zc-rs
# clresourcegroup online -eM scal-wait-zc-rg
```

Next Steps Go to "How to Register and Configure the Oracle ACFS Proxy Resource Group" on page 185.

▼ How to Register and Configure the Oracle ACFS Proxy Resource Group

Use this procedure to register and configure the Oracle ACFS proxy resource group. Perform all steps from one node of the global cluster.

Before You Begin Ensure that you have registered and configured the Oracle Grid Infrastructure resource. See "How to Create an Oracle Grid Infrastructure Resource for Interoperation With Oracle Solaris Cluster" on page 187.

1 **Become superuser or assume a role that provides `solaris.cluster.admin` and `solaris.cluster.modify` RBAC authorizations.**

2 **If the file system will be used by a zone cluster, from the global zone of one node, create an Oracle Clusterware proxy resource.**

a. **Create the resource.**

```
# /Grid_home/bin/crsctl add type sun.zcboot_proxy.type -basetype local_resource
# /Grid_home/bin/crsctl add res sun.wait-zc-rs \
-type sun.zcboot_proxy.type \
-attr "ACTION_SCRIPT='/opt/SUNWscor/dsconfig/bin/scproxy_crs_action' \
ACL='owner:root:rwx,pgrp:oinstall:rwx,other::r--' \
SCRIPT_TIMEOUT='20' \
RESTART_ATTEMPTS='60' "
```

b. **Verify the resource.**

```
# /Grid_home/bin/crsctl stat res sun.wait-zc-rs -p
NAME=sun.wait-zc-rs
TYPE=sun.zcboot_proxy.type
```

```
ACL=owner:root:rwx,pgrp:oinstall:rwx,other::r-
...
```

c. Bring the resource online.

```
# /Grid_home/bin/crsctl start res sun.wait-zc-rs
```

3 Register the `SUNW.scalable_acfs_proxy` resource type.

- **If the file system is for use in the global cluster, use the following command:**

```
# clresourcetype register SUNW.scalable_acfs_proxy
```

- **If the file system is for use by a zone cluster, use the following command:**

```
# clresourcetype register -Z zonecluster SUNW.scalable_acfs_proxy
```

4 Create the Oracle ACFS resource group with offline-restart resource dependencies.

- **If the file system is for use in the global cluster, use the following command:**

```
# clresourcegroup create -S -p rg_affinities=++asm-dg-rg acfs-rg
```

asm-dg-rg Specifies the name of the Oracle ASM disk-group resource group.

acfs-rg Specifies the name of the Oracle ACFS resource group.

- **If the file system is for use by a zone cluster, use the following command:**

```
# clresourcegroup create -Z zonecluster -S \
-p rg_affinities=++global:asm-dg-rg,++global:scal-wait-zc-rg \
acfs-rg
```

scal-wait-zc-rg For a zone cluster, specifies the `SUNW.wait_zc_boot` resource group.

5 Add an instance of the `SUNW.scalable_acfs_proxy` resource type to the Oracle ACFS resource group.

- **If the file system is for use in the global cluster, use the following command:**

```
# clresource create -g acfs-rg \
-t SUNW.scalable_acfs_proxy \
-p acfs_mountpoint=/acfs-mountpoint \
-p resource_dependencies_offline_restart=asm-dg-rs \
-d acfs-rs
```

- **If the file system is for use by a zone cluster, use the following command:**

```
# clresource create -Z zonecluster -g acfs-rg \
-t SUNW.scalable_acfs_proxy \
-p acfs_mountpoint=/acfs-mountpoint \
-p resource_dependencies_offline_restart=global:asm-dg-rs \
-p resource_dependencies=global:wait-zc-rs \
-d acfs-rs
```

6 **Bring online the *acfs-rg* resource group in a managed state on a cluster node.**

```
# clresourcegroup online -eM acfs-rg
```

7 **Verify the Oracle ACFS configuration.**

```
# clresource status +
```

Next Steps Go to "How to Create an Oracle Grid Infrastructure Resource for Interoperation With Oracle Solaris Cluster" on page 187.

▼ How to Create an Oracle Grid Infrastructure Resource for Interoperation With Oracle Solaris Cluster

Use this procedure to create an Oracle Grid Infrastructure resource. This resource coordinates operations that are managed by Oracle Clusterware with operations that are managed by Oracle Solaris Cluster. Perform all steps from one node of the global cluster.

Before You Begin Ensure that you have registered and configured the Oracle ASM resource groups. See "How to Register and Configure the Oracle ASM Resource Group" on page 183.

1 **Become superuser or assume a role that provides `solaris.cluster.admin` and `solaris.cluster.modify` RBAC authorizations.**

2 **If using Solaris Volume Manager for Sun Cluster, configure the Oracle Grid Infrastructure storage proxy resource.**

a. **Create the Oracle Grid Infrastructure `sun.storage_proxy.type` resource type.**

```
# /Grid_home/bin/crsctl \
add type sun.storage_proxy.type \
-basetype cluster_resource \
-attr \
"ATTRIBUTE=ACTION_SCRIPT,TYPE=string", \
"ATTRIBUTE=HOSTING_MEMBERS,TYPE=string", \
"ATTRIBUTE=CARDINALITY,TYPE=string", \
"ATTRIBUTE=PLACEMENT,TYPE=string", \
"ATTRIBUTE=SCRIPT_TIMEOUT,TYPE=int", \
"ATTRIBUTE=RESTART_ATTEMPTS,TYPE=int", \
"ATTRIBUTE=ACL,TYPE=string", \
"ATTRIBUTE=VERSION,TYPE=string"
```

b. Create the Oracle Grid Infrastructure `sun.`*storage-proxy-resource* resource of type `sun.storage_proxy.type`.

The Oracle Grid Infrastructure resource name uses the form `sun.`*storage-proxy-resource*, where *storage-proxy-resource* is the name of the `SUNW.ScalDeviceGroup` resource.

```
# /Grid_home/bin/crsctl add resource sun.storage-proxy-resource \
-type sun.storage_proxy.type \
-attr "ACTION_SCRIPT='/opt/SUNWscor/dsconfig/bin/scproxy_crs_action' \
CARDINALITY='number-nodes' \
SCRIPT_TIMEOUT='timeout' \
PLACEMENT='restricted' \
RESTART_ATTEMPTS='restarts' \
HOSTING_MEMBERS='nodelist' \
VERSION='1' "
```

`CARDINALITY`	The number of nodes in the cluster membership
`HOSTING_MEMBERS`	The list of nodes in the cluster membership

c. Bring online the Oracle Grid Infrastructure storage proxy resource.

```
# /Grid_home/bin/crsctl start resource sun.storage-proxy-resource
```

3 Create the Oracle Grid Infrastructure stop trigger for the Oracle Solaris Cluster ACFS proxy resource.

Perform this step from one node of the global cluster.

a. Create the stop trigger resource.

```
# /Grid_home/bin/crsctl add type sun.stoptrigger.type -basetype cluster_resource
# /Grid_home/bin/crsctl add res sun.acfs-rs -type sun.stoptrigger.type \
-attr "ACTION_SCRIPT='/opt/SUNWscor/dsconfig/bin/crs_stoptrigger_action' \
HOSTING_MEMBERS='node1 node2 [...]' \
CARDINALITY='number-nodes' \
PLACEMENT='restricted' \
ACL='owner:root:rwx,pgrp:oinstall:rwx,other::r--' \
SCRIPT_TIMEOUT='20' \
RESTART_ATTEMPTS='60' \
START_DEPENDENCIES='hard(ora.ASMdg.ASMvolume.acfs) pullup:always(ora.ASMdg.ASMvolume.acfs)' \
STOP_DEPENDENCIES='hard(ora.ASMdg.ASMvolume.acfs)' "
```

b. Verify the stop trigger resource.

```
# /Grid_home/bin/crsctl stat res sun.acfs-rs -p
NAME=sun.acfs-rs
TYPE=sun.stoptrigger.type
...
```

c. Start the stop trigger resource.

```
# /Grid_home/bin/crsctl start res sun.acfs-rs
```

d. Verify that the resource is online on all nodes.

```
# /Grid_home/bin/crsctl stat res sun.acfs-rs
```

4 **If you want to use the Oracle ACFS file system for Oracle RAC, configure an Oracle Grid Infrastructure resource.**

Follow procedures in "How to Create an Oracle Grid Infrastructure Resource for Interoperation With Oracle Solaris Cluster" in *Oracle Solaris Cluster Data Service for Oracle Real Application Clusters Guide*.

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- To create a zone cluster, go to "Configuring a Zone Cluster" on page 197.
- To create non-global zones on a node, go to "Configuring a Non-Global Zone on a Global-Cluster Node" on page 191.
- Install third-party applications, register resource types, set up resource groups, and configure data services. See the documentation that is supplied with the application software and the *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

◆ ◆ ◆ CHAPTER 6

Creating Non-Global Zones and Zone Clusters

This chapter describes the following topics:

Configuring a Non-Global Zone on a Global-Cluster Node

This section provides the following procedures to create a non-global zone on a global-cluster node.

▼ How to Create a Non-Global Zone on a Global-Cluster Node

Perform this procedure for each non-global zone that you create in the global cluster.

Note – For complete information about installing a zone, refer to *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

You can configure an Oracle Solaris Containers non-global zone, simply referred to as a zone, on a cluster node while the node is booted in either cluster mode or in noncluster mode.

- If you create a zone while the node is booted in noncluster mode, the cluster software discovers the zone when the node joins the cluster.
- If you create or remove a zone while the node is in cluster mode, the cluster software dynamically changes its list of zones that can master resource groups.

Before You Begin Perform the following tasks:

- Plan your non-global zone configuration. Observe the requirements and restrictions in "Guidelines for Non-Global Zones in a Global Cluster" on page 18.
- Have available the following information:
  - The total number of non-global zones that you will create.
  - The public adapter and public IP address that each zone will use.
  - The zone path for each zone. This path must be a local file system, not a cluster file system or a highly available local file system.
  - One or more devices that should appear in each zone.
  - (Optional) The name that you will assign each zone.
- If you will assign the zone a private IP address, ensure that the cluster IP address range can support the additional private IP addresses that you will configure. Use the `cluster show-netprops` command to display the current private-network configuration.

  If the current IP address range is not sufficient to support the additional private IP addresses that you will configure, follow the procedures in "How to Change the Private Network Configuration When Adding Nodes or Private Networks" on page 108 to reconfigure the private IP-address range.

Note – You can turn off cluster functionality for a selected non-global zone, so that a root user logged into one of these zones will not able to discover or disrupt operation of the cluster. For instructions, see "How to Deny Cluster Services For a Non-Global Zone" in *Oracle Solaris Cluster Data Services Planning and Administration Guide* and "How to Allow Cluster Services For a Non-Global Zone" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

For additional information, see "Zone Components" in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

1 Become superuser on the global-cluster node where you are creating the non-voting node.

You must be working in the global zone.

2 Verify on each node that multiuser services for the Service Management Facility (SMF) are online.

If services are not yet online for a node, wait until the state changes to online before you proceed to the next step.

```
phys-schost# svcs multi-user-server node
STATE          STIME    FMRI
online         17:52:55 svc:/milestone/multi-user-server:default
```

3 Configure, install, and boot the new zone.

Note – You must set the `autoboot` property to `true` to support resource-group functionality in the non-voting node on the global cluster.

Follow procedures in the Oracle Solaris documentation:

a. **Perform procedures in Chapter 18, "Planning and Configuring Non-Global Zones (Tasks)," in** *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

b. **Perform procedures in "Installing and Booting Zones" in** *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

c. **Perform procedures in "How to Boot a Zone" in** *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

4 Verify that the zone is in the `ready` state.

```
phys-schost# zoneadm list -v
ID  NAME     STATUS       PATH
 0  global   running      /
 1  my-zone  ready        /zone-path
```

5 (Optional) For a shared-IP zone, assign a private IP address and a private hostname to the zone.

The following command chooses and assigns an available IP address from the cluster's private IP-address range. The command also assigns the specified private hostname, or host alias, to the zone and maps it to the assigned private IP address.

```
phys-schost# clnode set -p zprivatehostname=hostalias  node:zone
```

`-p`	Specifies a property.
`zprivatehostname=`*hostalias*	Specifies the zone private hostname, or host alias.
node	The name of the node.
zone	The name of the global-cluster non-voting node.

6 Perform the initial internal zone configuration.

Follow the procedures in "Performing the Initial Internal Zone Configuration" in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*. Choose either of the following methods:

- Log in to the zone.
- Use an `/etc/sysidcfg` file.

7 In the non-voting node, modify the `nsswitch.conf` file.

These changes enable the zone to resolve searches for cluster-specific hostnames and IP addresses.

a. Log in to the zone.

```
phys-schost# zlogin -c zonename
```

b. Open the `/etc/nsswitch.conf` file for editing.

```
sczone# vi /etc/nsswitch.conf
```

c. Add the `cluster` switch to the beginning of the lookups for the `hosts` and `netmasks` entries, followed by the `files` switch.

The modified entries should appear similar to the following:

```
...
hosts:      cluster files nis [NOTFOUND=return]
...
netmasks:   cluster files nis [NOTFOUND=return]
...
```

d. For all other entries, ensure that the `files` switch is the first switch that is listed in the entry.

e. Exit the zone.

8 If you created an exclusive-IP zone, configure IPMP groups in each `/etc/hostname.`*interface* file that is on the zone.

You must configure an IPMP group for each public-network adapter that is used for data-service traffic in the zone. This information is not inherited from the global zone. See "Public Networks" on page 23 for more information about configuring IPMP groups in a cluster.

9 Set up name-to-address mappings for all logical hostname resources that are used by the zone.

a. Add name-to-address mappings to the `/etc/inet/hosts` file on the zone.

This information is not inherited from the global zone.

b. If you use a name server, add the name-to-address mappings.

Next Steps To install an application in a non-global zone, use the same procedure as for a stand-alone system. See your application's installation documentation for procedures to install the software in a non-global zone. Also see "Adding and Removing Packages and Patches on an Oracle Solaris System With Zones Installed (Task Map)" in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

To install and configure a data service in a non-global zone, see the Oracle Solaris Cluster manual for the individual data service.

How to Configure an HAStoragePlus Resource for a Cluster File System That is Used by Non-Global Zones

Use this procedure to make a cluster file system available for use by a `native` brand non-global zone that is configured on a cluster node.

Note – Use this procedure with only the `native` brand of non-global zones. You cannot perform this task with any other brand of non-global zone, such as the `solaris8` brand or the `cluster` brand which is used for zone clusters.

---

1. **On one node of the global cluster, become superuser or assume a role that provides `solaris.cluster.modify` RBAC authorization.**

2. **Create a resource group with a node list of `native` brand non-global zones.**

   - **Use the following command to create a failover resource group:**

     ```
     phys-schost# clresourcegroup create -n node:zone[,...] resource-group
     ```

     `-n` *`node`*`:`*`zone`*
     Specifies the names of the non-global zones in the resource-group node list.

     resource-group
     The name of the resource group that you create.

   - **Use the following command to create a scalable resource group:**

     ```
     phys-schost# clresourcegroup create -S -n node:zone[,...] resource-group
     ```

     `-S`
     Specifies that the resource group is scalable.

3. **Register the HAStoragePlus resource type.**

   ```
   phys-schost# clresourcetype register SUNW.HAStoragePlus
   ```

4. **On each global-cluster node where a non-global zone in the node list resides, add the cluster file system entry to the `/etc/vfstab` file.**

   Entries in the `/etc/vfstab` file for a cluster file system must contain the `global` keyword in the mount options.

   See "Sample Entries in /etc/vfstab for Cluster File Systems" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

5 **Create the HAStoragePlus resource and define the file-system mount points.**

```
phys-schost# clresource create -g resource-group -t SUNW.HAStoragePlus \
-p FileSystemMountPoints="mount-point-list" hasp-resource
```

-g *resource-group*
Specifies the name of the resource group that the new resource is added to.

-p FileSystemMountPoints="*mount-point-list***"**
Specifies one or more file-system mount points for the resource.

hasp-resource
The name of the HAStoragePlus resource that you create.

The resource is created in the enabled state.

6 **Add a resource to** *resource-group* **and set a dependency for the resource on** *hasp-resource***.**

If you have more than one resource to add to the resource group, use a separate command for each resource.

```
phys-schost# clresource create -g resource-group -t resource-type \
-p Network_resources_used=hasp-resource resource
```

`-t` *resource-type*
Specifies the resource type that you create the resource for.

`-p Network_resources_used=`*hasp-resource*
Specifies that the resource has a dependency on the HAStoragePlus resource, *hasp-resource*.

resource
The name of the resource that you create.

7 **Bring online and in a managed state the resource group that contains the HAStoragePlus resource.**

```
phys-schost# clresourcegroup online -eM resource-group
```

-M
Specifies that the resource group is managed.

Example 6–1 Configuring an HAStoragePlus Resource for a Cluster File System That is Used by Non-Global Zones

The following example creates a failover resource group, `cfs-rg`, to manage an HA-Apache data service. The resource-group node list contains two non-global zones, `sczone1` on `phys-schost-1` and `sczone1` on `phys-schost-2`. The resource group contains an HAStoragePlus resource, `hasp-rs`, and a data-service resource, `apache-rs`. The file-system mount point is `/global/local-fs/apache`.

```
phys-schost-1# clresourcegroup create -n phys-schost-1:sczone1,phys-schost-2:sczone1 cfs-rg
phys-schost-1# clresourcetype register SUNW.HAStoragePlus
```

Add the cluster file system entry to the /etc/vfstab file on phys-schost-1

```
phys-schost-1# vi /etc/vfstab
#device            device         mount          FS    fsck    mount    mount
#to mount          to fsck        point          type  pass    at boot  options
#
/dev/md/kappa-1/dsk/d0 /dev/md/kappa-1/rdsk/d0 /global/local-fs/apache ufs 5 yes logging,global
```

Add the cluster file system entry to the /etc/vfstab file on phys-schost-2

```
phys-schost-2# vi /etc/vfstab

...
phys-schost-1# clresource create -g cfs-rg -t SUNW.HAStoragePlus \
-p FileSystemMountPoints="/global/local-fs/apache" hasp-rs
phys-schost-1# clresource create -g cfs-rg -t SUNW.apache \
-p Network_resources_used=hasp-rs apache-rs
phys-schost-1# clresourcegroup online -eM cfs-rg
```

Configuring a Zone Cluster

This section provide procedures to configure a cluster of Oracle Solaris Containers non-global zones, called a zone cluster.

Overview of the `clzonecluster` Utility

The `clzonecluster` utility creates, modifies, and removes a zone cluster. The `clzonecluster` utility actively manages a zone cluster. For example, the `clzonecluster` utility both boots and halts a zone cluster. Progress messages for the `clzonecluster` utility are output to the console, but are not saved in a log file.

The utility operates in the following levels of scope, similar to the `zonecfg` utility:

- The cluster scope affects the entire zone cluster.
- The node scope affects only the one zone-cluster node that is specified.
- The resource scope affects either a specific node or the entire zone cluster, depending on which scope you enter the resource scope from. Most resources can only be entered from the node scope. The scope is identified by the following prompts:

  `clzc:`*zoneclustername*`:`*resource*`>` *cluster-wide setting*
  `clzc:`*zoneclustername*`:node:`*resource*`>` *node-specific setting*

You can specify any Oracle Solaris zones resource parameter, as well as parameters that are specific to zone clusters, by using the `clzonecluster` utility. For information about parameters that you can set in a zone cluster, see the `clzonecluster(1CL)`man page. Additional information about Oracle Solaris zones resource parameters is in the `zonecfg(1M)` man page.

Establishing the Zone Cluster

This section describes how to configure a cluster of non-global zones.

How to Prepare for Trusted Extensions Use With Zone Clusters

This procedure prepares the global cluster to use the Trusted Extensions feature of Oracle Solaris software with zone clusters and enables the Trusted Extensions feature.

If you do not plan to enable Trusted Extensions, proceed to "How to Create a Zone Cluster" on page 201.

Perform this procedure on each node in the global cluster.

Before You Begin Perform the following tasks:

- Ensure that the Oracle Solaris OS is installed to support Oracle Solaris Cluster and Trusted Extensions software.

  If Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for Oracle Solaris Cluster software and any other software that you intend to install on the cluster. Trusted Extensions software is not included in the Oracle Solaris End User software group.

  See "How to Install Oracle Solaris Software" on page 57 for more information about installing Oracle Solaris software to meet Oracle Solaris Cluster software requirements.
- Ensure that an LDAP naming service is configured for use by Trusted Extensions. See Chapter 5, "Configuring LDAP for Trusted Extensions (Tasks)," in *Trusted Extensions Configuration Guide*.
- Review guidelines for Trusted Extensions in a zone cluster. See "Guidelines for Trusted Extensions in a Zone Cluster" on page 39.

1 Become superuser on a node of the global cluster.

2 Disable the Trusted Extensions `zoneshare` and `zoneunshare` scripts.

The Trusted Extensions `zoneshare` and `zoneunshare` scripts support the ability to export home directories on the system. An Oracle Solaris Cluster configuration does not support this feature.

Disable this feature by replacing each script with a symbolic link to the `/bin/true` utility. Do this on each global-cluster node.

```
phys-schost# ln -s /usr/lib/zones/zoneshare /bin/true
phys-schost# ln -s /usr/lib/zones/zoneunshare /bin/true
```

3 Configure all logical-hostname shared-IP addresses that are in the global cluster.

See "Run the txzonemgr Script" in *Trusted Extensions Configuration Guide*.

4 Ensure that the administrative console is defined in the `/etc/security/tsol/tnrhdb` file as `admin_low`.

ipaddress`:admin_low`

5 Ensure that no `/etc/hostname.`*interface* file contains the `-failover` option in an entry.

Delete the `-failover` option from any entry that contains that option.

6 Modify the `/etc/security/tsol/tnrhdb` file to authorize communication with global-cluster components.

Use the Security Templates wizard in Solaris Management Console as described in "How to Construct a Remote Host Template" in *Trusted Extensions Administrator's Procedures* to perform the following tasks.

- Create a new entry for IP addresses used by cluster components and assign each entry a CIPSO template.

  Add entries for each of the following IP addresses that exist in the global-cluster node's `/etc/inet/hosts` file:

  - Each global-cluster node private IP address
  - All `cl_privnet` IP addresses in the global cluster
  - Each logical-hostname public IP address for the global cluster
  - Each shared-address public IP address for the global cluster

  Entries would look similar to the following.

  ```
  127.0.0.1:cipso
  172.16.4.1:cipso
  172.16.4.2:cipso
  ...
  ```

- Add an entry to make the default template internal.

  ```
  0.0.0.0:internal
  ```

For more information about CIPSO templates, see "Configure the Domain of Interpretation" in *Trusted Extensions Configuration Guide*.

7 Enable the Trusted Extensions SMF service and reboot the global-cluster node.

```
phys-schost# svcadm enable -s svc:/system/labeld:default
phys-schost# shutdown -g0 -y -i6
```

For more information, see "Enable Trusted Extensions" in *Trusted Extensions Configuration Guide*.

8 Verify that the Trusted Extensions SMF service is enabled.

```
phys-schost# svcs labeld
STATE          STIME    FMRI
online         17:52:55 svc:/system/labeld:default
```

9 Repeat Step 1 through Step 8 on each remaining node of the global cluster.

When all steps are completed on all global-cluster nodes, perform the remaining steps of this procedure on each node of the global cluster.

10 Add the IP address of the Trusted Extensions-enabled LDAP server to the `/etc/inet/hosts` file on each global-cluster node.

The LDAP server is used by the global zone and by the nodes of the zone cluster.

11 Enable remote login by the LDAP server to the global-cluster node.

a. In the `/etc/default/login` file, comment out the `CONSOLE` entry.

b. Enable remote login.

```
phys-schost# svcadm enable rlogin
```

c. Modify the `/etc/pam.conf` file.

Modify the account management entries by appending a Tab and typing `allow_remote` or `allow_unlabeled` respectively, as shown below.

```
other   account requisite       pam_roles.so.1          Tab  allow_remote
other   account required        pam_unix_account.so.1   Tab  allow_unlabeled
```

12 Modify the `/etc/nsswitch.ldap` file.

- Ensure that the `passwd` and `group` lookup entries have `files` first in the lookup order.

```
...
passwd:      files ldap
group:       files ldap
...
```

- Ensure that the `hosts` and `netmasks` lookup entries have `cluster` listed first in the lookup order.

```
...
hosts:       cluster files ldap
...
netmasks:    cluster files ldap
...
```

13 **Make the global-cluster node an LDAP client.**

See "Make the Global Zone an LDAP Client in Trusted Extensions" in *Trusted Extensions Configuration Guide*.

14 **Add Trusted Extensions users to the `/etc/security/tsol/tnzonecfg` file.**

Use the Add User wizard in Solaris Management Console as described in "Creating Roles and Users in Trusted Extensions" in *Trusted Extensions Configuration Guide*.

Next Steps Create the zone cluster. Go to "How to Create a Zone Cluster" on page 201.

▼ How to Create a Zone Cluster

Perform this procedure to create a cluster of non-global zones.

To modify the zone cluster after it is installed, see "Performing Zone Cluster Administrative Tasks" in *Oracle Solaris Cluster System Administration Guide* and the `clzonecluster`(1CL) man page.

Before You Begin

- Create a global cluster. See Chapter 3, "Establishing the Global Cluster."
- Read the guidelines and requirements for creating a zone cluster. See "Zone Clusters" on page 36.
- If the zone cluster will use Trusted Extensions, ensure that you have configured and enabled Trusted Extensions as described in "How to Prepare for Trusted Extensions Use With Zone Clusters" on page 198.
- Have available the following information:
  - The unique name to assign to the zone cluster.

    Note – To configure a zone cluster when Trusted Extensions is enabled, you must use the name of the Trusted Extensions security label that the zone cluster will use as the name of the zone cluster itself. Create a separate zone cluster for each Trusted Extensions security label that you want to use.

  - The zone path that the nodes of the zone cluster will use. For more information, see the description of the `zonepath` property in "Resource and Property Types" in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.
  - The name of each node in the global cluster on which to create a zone-cluster node.
  - The zone public hostname, or host alias, that you assign to each zone-cluster node.
  - If applicable, the public-network IPMP group that each zone-cluster node uses.
  - If applicable, the name of the public-network adapter that each zone-cluster node uses to connect to the public network.

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

- 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.
- The cluster software will activate any Logical Host IP address on any NIC.

1 Become superuser on an active member node of a global cluster.

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

2 Ensure that the node of the global cluster is in cluster mode.

If any node is in noncluster mode, changes that you make are propagated when the node returns to cluster mode. Therefore, you can create a zone cluster even if some global-cluster nodes are in noncluster mode. When those nodes return to cluster mode, the system performs zone-cluster creation tasks on those nodes.

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

--- Node Status ---

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

3 Start the `clsetup` utility.

```
phys-schost# clsetup
```

The Main Menu is displayed.

4 Choose the Zone Cluster menu item.

5 Choose the Create a Zone Cluster menu item.

6 Type the name of the zone cluster you want to add.

A zone cluster name can contain ASCII letters (a-z and A-Z), numbers, a dash, or an underscore. The maximum length of the name is 20 characters.

7 Choose the property you want to change.

Note – The `brand` and `ip-type` properties are set by default and cannot be changed.

You can set the following properties:

Property	Description
`zonepath=`*zone-cluster-node-path*	Specifies the path to the zone cluster node. For example, `/zones/sczone`.
`enable_priv_net=`*value*	When set to `true`, Oracle Solaris Cluster private network communication is enabled between the nodes of the zone cluster. The Oracle Solaris Cluster private hostnames and IP addresses for the zone cluster nodes are automatically generated by the system. Private network communication is disabled if the value is set to `false`. The default value is `true`.
`limitpriv=`*privilege*`[,…]`	Specifies the maximum set of privileges any process in this zone can obtain. See the `zonecfg(1M)` man page for more information.

8 (Optional) Choose the Zone System Resource Control properties that you want to change.

You can set the following properties:

Property	Description
`max-lwps=`*value*	Specifies the maximum number of lightweight processes (LWPs) simultaneously available to this zone cluster.
`max-shm-memory=`*value*	Specifies the maximum amount of shared memory in GBytes allowed for this zone cluster.
`max-shm-ids=`*value*	Specifies the maximum number of shared memory IDs allowed for this zone cluster.
`max-msg-ids=`*value*	Specifies the maximum number of message queue IDs allowed for this zone cluster.
`max-sem-ids=`*value*	Specifies the maximum number of semaphore IDs allowed for this zone cluster.
`cpu-shares=`*value*	Specifies the number of Fair Share Scheduler (FSS) shares to allocate to this zone cluster.

9 (Optional) Choose the Zone CPU Resource Control property that you want to change.

You can set the following properties:

Property	Description
`scope=`*scope-type*	Specifies whether the `ncpus` property used in a zone cluster is `dedicated-cpu` or `capped-cpu`.

Property	Description
`ncpus=`*value*	Specifies the limit for the scope type. ■ If the `scope` property is set to `dedicated-cpu`, the `ncpus` property sets a limit on the number of CPUs that should be assigned for this zone's exclusive use. The zone will create a pool and processor set when it boots. See the `pooladm(1M)` and `poolcfg(1M)` man pages for more information on resource pools. ■ If the `scope` property is set to `capped-cpu`, the `ncpus` property sets a limit on the amount of CPU time that can be used by a zone cluster. The unit used translates to the percentage of a single CPU that can be used by all user threads in a zone, expressed as a fraction (for example, .75) or a mixed number (whole number and fraction, for example, 1.25). An `ncpus` value of 1 means 100% of a CPU. See the `pooladm(1M)`, `pooladm(1M)`, and `poolcfg(1M)` man pages for more information on resource pools.

10 (Optional) Choose the capped-memory property that you want to change.

You can set the following properties:

Property	Description
`physical=`*value*	Specifies the GByte limit for physical memory.
`swap=`*value*	Specifies the GByte limit for swap memory.
`locked=`*value*	Specifies the GByte limit for locked memory.

11 Choose a physical host from the list of available physical hosts.

You can select one or all of the available physical nodes (or hosts), and then configure one zone-cluster node at a time.

You can set the following properties:

Property	Description
`hostname=`*hostname*	Specifies the zone-cluster node hostname. For example, `zc-host-1`.
`address=`*public-network-address*	Specifies the public network address for the zone-cluster node on a shared-IP type zone cluster. For example, `172.1.1.1`.
`physical=`*physical-interface*	Specifies a network physical interface for the public network from the available network interfaces that are discovered on the physical nodes, for example, `bge0`.

Property	Description
`defrouter=`*default-router*	Specifies the default router for the network address, if your zone is configured in a different subnet. Each zone or set of zones that uses a different `defrouter` setting must be on a different subnet, for example, `192.168.0.1`. See the `zonecfg(1M)` man page for more information about the `defrouter` property.

12 Specify the network addresses for the zone cluster.

The network addresses can be used to configure a logical hostname or shared-IP cluster resources in the zone cluster. The network address is in the zone cluster global scope.

13 At the Review Configuration screen, press Return to continue and then type c to create the zone cluster.

The results of your configuration change are displayed, similar to the following:

```
>>> Result of the Creation for the Zone Cluster(sczone) <<<

    The zone cluster is being created with the following configuration

        /usr/cluster/bin/clzonecluster configure sczone
        create
        set brand=cluster
        set zonepath=/zones/sczone
        set ip-type=shared
        set enable_priv_net=true
        add capped-memory
        set physical=2G
        end
        add node
        set physical-host=phys-schost-1
        set hostname=zc-host-1
        add net
        set address=172.1.1.1
        set physical=net0
        end
        end
        add net
        set address=172.1.1.2
              end

    Zone cluster, zc2 has been created and configured successfully.

    Continue to install the zone cluster(yes/no) ?
```

14 Type yes to continue.

The `clsetup` utility performs a standard installation of a zone cluster and you cannot specify any options.

15 When finished, exit the `clsetup` utility.

16 Verify the zone cluster configuration.

The `verify` subcommand checks for the availability of the specified resources. If the `clzonecluster verify` command succeeds, there is no output.

```
phys-schost-1# clzonecluster verify zoneclustername
phys-schost-1# clzonecluster status zoneclustername
=== Zone Clusters ===

--- Zone Cluster Status ---

Name      Node Name   Zone HostName   Status    Zone Status
----      ---------   -------------   ------    -----------
zone      basenode1   zone-1        Offline   Configured
          basenode2   zone-2        Offline   Configured
```

17 For Trusted Extensions, make the password files writable on each zone-cluster node.

From the global zone, launch the `txzonemgr` GUI.

```
phys-schost# txzonemgr
```

Select the global zone, then select the item, Configure per-zone name service.

18 Install the zone cluster.

```
phys-schost-1# clzonecluster install [-c config-profile.xml] zoneclustername
Waiting for zone install commands to complete on all the nodes
of the zone cluster "zoneclustername"...
```

The `-c` *config-profile*`.xml` option specifies a configuration profile for all non-global zones of the zone cluster. Using this option changes only the hostname of the zone, which is unique for each zone in the zone cluster. All profiles must have a `.xml` extension.

19 Boot the zone cluster.

Installation of the zone cluster might take several minutes

```
phys-schost-1# clzonecluster boot zoneclustername
Waiting for zone boot commands to complete on all the nodes of
the zone cluster "zoneclustername"...
```

20 If you did *not* use the `-c` *config-profile*`.xml` option when you installed the zone cluster, perform `sysid` configuration.

Perform the following steps on each zone-cluster node.

Note – In the following steps, the non-global zone *zcnode* and *zone-cluster-name* share the same name.

---

a. Unconfigure the Oracle Solaris instance and reboot the zone.

```
phys-schost# zlogin zcnode
zcnode# sysconfig unconfigure
zcnode# reboot
```

The `zlogin` session terminates during the reboot.

b. Issue the `zlogin` command and progress through the interactive screens.

```
phys-schost# zlogin -C zcnode
```

c. When finished, exit the zone console.

For information about methods to exit from a non-global zone, see "How to Exit a Non-Global Zone" in *System Administration Guide: Oracle Solaris Containers-Resource Management and Oracle Solaris Zones*.

d. Repeat for each remaining zone-cluster node.

21 If you use Trusted Extensions, complete IP-address mappings for the zone cluster.

Perform this step on each node of the zone cluster.

a. From a node of the global cluster, display the node's ID.

```
phys-schost# cat /etc/cluster/nodeid
N
```

b. Log in to a zone-cluster node on the same global-cluster node.

Ensure that the SMF service has been imported and all services are up before you log in.

c. Determine the IP addresses used by this zone-cluster node for the private interconnect.

The cluster software automatically assigns these IP addresses when the cluster software configures a zone cluster.

In the `ifconfig -a` output, locate the `clprivnet0` logical interface that belongs to the zone cluster. The value for `inet` is the IP address that was assigned to support the use of the cluster private interconnect by this zone cluster.

```
zc1# ifconfig -a
lo0:3: flags=20010008c9<UP,LOOPBACK,RUNNING,NOARP,MULTICAST,IPv4,VIRTUAL> mtu 8232 index 1
        zone zc1
        inet 127.0.0.1 netmask ff000000
bge0: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 2
        inet 10.11.166.105 netmask ffffff00 broadcast 10.11.166.255
        groupname sc_ipmp0
        ether 0:3:ba:19:fa:b7
ce0: flags=9040843<UP,BROADCAST,RUNNING,MULTICAST,DEPRECATED,IPv4,NOFAILOVER> mtu 1500 index 4
        inet 10.11.166.109 netmask ffffff00 broadcast 10.11.166.255
        groupname sc_ipmp0
        ether 0:14:4f:24:74:d8
ce0:3: flags=1000843<UP,BROADCAST,RUNNING,MULTICAST,IPv4> mtu 1500 index 4
        zone zc1
        inet 10.11.166.160 netmask ffffff00 broadcast 10.11.166.255
clprivnet0: flags=1009843<UP,BROADCAST,RUNNING,MULTICAST,MULTI_BCAST,PRIVATE,IPv4> mtu 1500 index 7
        inet 172.16.0.18 netmask fffffff8 broadcast 172.16.0.23
        ether 0:0:0:0:0:2
clprivnet0:3: flags=1009843<UP,BROADCAST,RUNNING,MULTICAST,MULTI_BCAST,PRIVATE,IPv4> mtu 1500 index 7
        zone zc1
        inet 172.16.0.22 netmask fffffffc broadcast 172.16.0.23
```

d. Add to the zone-cluster node's `/etc/inet/hosts` file the IP addresses of the zone-cluster node.

- The hostname for the private interconnect, which is `clusternode`*N*`-priv`, where *N* is the global-cluster node ID

  ```
  172.16.0.22     clusternodeN-priv
  ```

- Each `net` resource that was specified to the `clzonecluster` command when you created the zone cluster

e. Repeat on the remaining zone-cluster nodes.

22 Modify the `/etc/security/tsol/tnrhdb` file to authorize communication with zone-cluster components.

Use the Security Templates wizard in Solaris Management Console as described in "How to Construct a Remote Host Template" in *Trusted Extensions Administrator's Procedures* to perform the following tasks.

- Create a new entry for IP addresses used by zone-cluster components and assign each entry a CIPSO template.

  Add entries for each of the following IP addresses that exist in the zone-cluster node's `/etc/inet/hosts` file:

  - Each zone-cluster node private IP address
  - All `cl_privnet` IP addresses in the zone cluster
  - Each logical-hostname public IP address for the zone cluster
  - Each shared-address public IP address for the zone cluster

  Entries would look similar to the following.

  ```
  127.0.0.1:cipso
  172.16.4.1:cipso
  172.16.4.2:cipso
  ...
  ```

- Add an entry to make the default template internal.

  ```
  0.0.0.0:internal
  ```

For more information about CIPSO templates, see "Configure the Domain of Interpretation" in *Trusted Extensions Configuration Guide*.

23 Enable DNS and `rlogin` access to the zone-cluster nodes.

Perform the following commands on each node of the zone cluster.

```
phys-schost# zlogin zcnode
zcnode# svcadm enable svc:/network/dns/client:default
zcnode# svcadm enable svc:/network/login:rlogin
zcnode# reboot
```

Example 6–2 Configuration File to Create a Zone Cluster

The following example shows the contents of a command file that can be used with the `clzonecluster` utility to create a zone cluster. The file contains the series of `clzonecluster` commands that you would input manually.

In the following configuration, the zone cluster `sczone` is created on the global-cluster node `phys-schost-1`. The zone cluster uses `/zones/sczone` as the zone path and public IP address 172.16.2.2. The first node of the zone cluster is assigned the hostname `zc-host-1` and uses the network address `172.16.0.1` and the `bge0` adapter. The second node of the zone cluster is created on the global-cluster node `phys-schost-2`. This second zone-cluster node is assigned the hostname `zc-host-2` and uses the network address `172.16.0.2` and the `bge1` adapter.

```
create
set zonepath=/zones/sczone
add net
set address=172.16.2.2
end
add node
set physical-host=phys-schost-1
set hostname=zc-host-1
add net
set address=172.16.0.1
set physical=bge0
end
end
add sysid
set root_password=encrypted_password
end
add node
set physical-host=phys-schost-2
set hostname=zc-host-2
add net
set address=172.16.0.2
set physical=bge1
end
end
commit
exit
```

Next Steps To add the use of a file system to the zone cluster, go to "Adding File Systems to a Zone Cluster" on page 210.

To add the use of global storage devices to the zone cluster, go to "Adding Storage Devices to a Zone Cluster" on page 220.

See Also To patch a zone cluster, follow procedures in Chapter 11, "Patching Oracle Solaris Cluster Software and Firmware," in *Oracle Solaris Cluster System Administration Guide*. These procedures include special instructions for zone clusters, where needed.

Adding File Systems to a Zone Cluster

This section provides procedures to add file systems for use by the zone cluster.

After a file system is added to a zone cluster and brought online, the file system is authorized for use from within that zone cluster. To mount the file system for use, configure the file system by using cluster resources such as `SUNW.HAStoragePlus` or `SUNW.ScalMountPoint`.

The following procedures are in this section:

In addition, to configure a ZFS storage pool to be highly available in a zone cluster, see "How to Set Up the HAStoragePlus Resource Type to Make a Local Solaris ZFS Highly Available" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

How to Add a Highly Available Local File System to a Zone Cluster

Perform this procedure to configure a highly available local file system on the global cluster for use by the zone cluster. The file system is added to the zone cluster and is configured with an HAStoragePlus resource to make the local file system highly available.

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

1 Become superuser on a node of the global cluster that hosts the zone cluster.

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

---

2 Start the `clsetup` utility.

```
phys-schost# clsetup
```

The Main Menu is displayed.

Tip – To return to a previous screen, type the < key and press Return.

---

3 Choose the Zone Cluster menu item.

The Zone Cluster Tasks Menu is displayed.

4 Choose the Add File System/Storage Device to a Zone Cluster menu item.

The Select Zone Cluster menu is displayed.

5 Choose the zone cluster where you want to add the file system.

The Storage Type Selection menu is displayed.

6 **Choose the File System menu item.**

The File System Selection for the Zone Cluster menu is displayed.

7 **Choose the file system you want to add to the zone cluster.**

The file systems in the list are those that are configured on the shared disks and can be accessed by the nodes where the zone cluster is configured. You can also type **e** to manually specify all properties for a file system.

The Mount Type Selection menu is displayed.

8 **Choose the Loopback mount type.**

The File System Properties for the Zone Cluster menu is displayed.

9 **Change the properties that you are allowed to change for the file system you are adding.**

Note – For UFS file systems, enable logging.

When, finished, type **d** and press Return.

10 **Type c to save the configuration change.**

The results of your configuration change are displayed.

11 **When finished, exit the `clsetup` utility.**

12 **Verify the addition of the file system.**

```
phys-schost# clzonecluster show -v zoneclustername
```

Example 6–3 Adding a Highly Available Local File System to a Zone Cluster

This example adds the local file system `/global/oracle/d1` for use by the `sczone` zone cluster.

```
phys-schost-1# clzonecluster configure sczone
clzc:sczone> add fs
clzc:sczone:fs> set dir=/global/oracle/d1
clzc:sczone:fs> set special=/dev/md/oracle/dsk/d1
clzc:sczone:fs> set raw=/dev/md/oracle/rdsk/d1
clzc:sczone:fs> set type=ufs
clzc:sczone:fs> add options [logging]
clzc:sczone:fs> end
clzc:sczone> verify
clzc:sczone> commit
clzc:sczone> exit

phys-schost-1# clzonecluster show -v sczone
...
  Resource Name:                          fs
    dir:                                     /global/oracle/d1
```

```
    special:                                /dev/md/oracle/dsk/d1
    raw:                                    /dev/md/oracle/rdsk/d1
    type:                                   ufs
    options:                                [logging]
    cluster-control:                        [true]
...
```

Next Steps Configure the file system to be highly available by using an HAStoragePlus resource. The HAStoragePlus resource manages the mounting of the file system on the zone-cluster node that currently host the applications that are configured to use the file system. See "Enabling Highly Available Local File Systems" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

▼ How to Add a ZFS Storage Pool to a Zone Cluster

Perform this procedure to add a ZFS storage pool for use by a zone cluster. The pool can be local to a single zone-cluster node or configured with HAStoragePlus to be highly available.

The `clsetup` utility discovers and displays all configured ZFS pools on the shared disks that can be accessed by the nodes where the selected zone cluster is configured. After you use the `clsetup` utility to add a ZFS storage pool in cluster scope to an existing zone cluster, you can use the `clzonecluster` command to modify the configuration or to add a ZFS storage pool in node-scope.

Before You Begin Ensure that the ZFS pool is connected on shared disks that are connected to all nodes of the zone cluster. See *Oracle Solaris ZFS Administration Guide* for procedures to create a ZFS pool.

1 Become superuser on a node of the global cluster that hosts the zone cluster.

Note – Perform all steps of this procedure from a node of the global zone.

---

2 Start the `clsetup` utility.

```
phys-schost# clsetup
```

The Main Menu is displayed.

Tip – To return to a previous screen, type the < key and press Return.

---

3 Choose the Zone Cluster menu item.

The Zone Cluster Tasks Menu is displayed.

4 Choose the Add File System/Storage Device to a Zone Cluster menu item.

The Select Zone Cluster menu is displayed.

5 **Choose the zone cluster where you want to add the ZFS storage pool.**

The Storage Type Selection menu is displayed.

6 **Choose the ZFS menu item.**

The ZFS Pool Selection for the Zone Cluster menu is displayed.

7 **Choose the ZFS pool you want to add to the zone cluster.**

The ZFS pools in the list are those that are configured on the shared disks and can be accessed by the nodes where the zone cluster is configured. You can also type **e** to manually specify properties for a ZFS pool.

The ZFS Pool Dataset Property for the Zone Cluster menu is displayed. The selected ZFS pool is assigned to the `name` property.

8 **Type d and press Return.**

The Review File Systems/Storage Devices for the Zone Cluster menu is displayed.

9 **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 file systems or storage devices to sczone zone cluster...

    The zone cluster is being created with the following configuration

        /usr/cluster/bin/clzonecluster configure sczone
        add dataset
        set name=myzpool5
        end

    Configuration change to sczone zone cluster succeeded.
```

10 **When finished, exit the `clsetup` utility.**

11 **Verify the addition of the file system.**

```
phys-schost# clzonecluster show -v zoneclustername
```

Example 6–4 Adding a ZFS Storage Pool to a Zone Cluster

The following example shows the ZFS storage pool `zpool1` added to the zone cluster `sczone`.

```
phys-schost-1# clzonecluster configure sczone
clzc:sczone> add dataset
clzc:sczone:dataset> set name=zpool1
clzc:sczone:dataset> end
clzc:sczone> verify
clzc:sczone> commit
clzc:sczone> exit
```

```
phys-schost-1# clzonecluster show -v sczone
...
  Resource Name:                                dataset
    name:                                          zpool1
...
```

Next Steps Configure the ZFS storage pool to be highly available by using an HAStoragePlus resource. The HAStoragePlus resource manages the mounting of the file systems that are in the pool on the zone-cluster node that currently host the applications that are configured to use the file system. See "Enabling Highly Available Local File Systems" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

▼ How to Add a Cluster File System to a Zone Cluster

The `clsetup` utility discovers and displays the available file systems that are configured on the cluster nodes where the selected zone cluster is configured. When you use the `clsetup` utility to add a file system, the file system is added in cluster scope.

You can add the following types of cluster file systems to a zone cluster:

- UFS cluster file system - You specify the file system type in the `/etc/vfstab` file, using the `global` mount option. This file system can be located on the shared disk or on a Solaris Volume Manager device.
- Sun QFS shared file system - You specify the file system type in the `/etc/vfstab` file, using the `shared` mount option.

  Note – At this time, QFS shared file systems are only supported for use in clusters that are configured with Oracle Real Application Clusters (RAC). On clusters that are not configured with Oracle RAC, you can use a single-machine QFS file system that is configured as a highly available local file system.

- ACFS - Discovered automatically, based on the `ORACLE_HOME` path you provide.

Before You Begin Ensure that the cluster file system you want to add to the zone cluster is configured. See "Planning Cluster File Systems" on page 41 and Chapter 5, "Creating a Cluster File System."

1 Become superuser on a node of the global cluster that hosts the zone cluster.

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

2 **On each node of the global cluster that hosts a zone-cluster node, add an entry to the `/etc/vfstab` file for the file system that you want to mount on the zone cluster.**

```
phys-schost# vi /etc/vfstab
```

- **For a UFS entry, include the `global` mount option, similar to the following example:**

```
/dev/md/datadg/dsk/d0 /dev/md/datadg/rdsk/d0 /global/fs ufs 2 no global, logging
```

- **For a shared QFS entry, include the `shared` mount option, similar to the following example:**

```
Data-cz1    -    /db_qfs/Data1 samfs - no shared,notrace
```

3 **On the global cluster, start the `clsetup` utility.**

```
phys-schost# clsetup
```

The Main Menu is displayed.

Tip – To return to a previous screen, type the < key and press Return.

---

4 **Choose the Zone Cluster menu item.**

The Zone Cluster Tasks Menu is displayed.

5 **Choose the Add File System/Storage Device to a Zone Cluster menu item.**

The Select Zone Cluster menu is displayed.

6 **Choose the zone cluster where you want to add the file system.**

The Storage Type Selection menu is displayed.

7 **Choose the File System menu item.**

The File System Selection for the Zone Cluster menu is displayed.

8 **Choose a file system from the list.**

The Mount Type Selection menu is displayed.

You can also type **e** to manually specify all properties for a file system.

Note – If you are using an ACFS file system, type **a** to select `Discover ACFS` and then specify the `ORACLE_HOME` directory.

---

9 **Choose the Loopback file system mount type for the zone cluster.**

Note – If you chose an ACFS file system in Step 8, the `clsetup` utility skips this step because ACFS supports only the direct mount type.

For more information about creating loopback file systems, see "How to Create and Mount an LOFS File System" in *System Administration Guide: Devices and File Systems*.

The File System Properties for the Zone Cluster menu is displayed.

10 Specify the mount point directory.

Type the number for the `dir` property and press Return. Then type the LOFS mount point directory name in the New Value field and press Return.

When finished, type `d` and press Return. The Review File Systems/Storage Devices for the Zone Cluster menu is displayed.

11 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 file systems or storage devices to sczone zone cluster...

  The zone cluster is being created with the following configuration

      /usr/cluster/bin/clzonecluster configure sczone
      add fs
      set dir=/dev/md/ddg/dsk/d9
      set special=/dev/md/ddg/dsk/d10
      set raw=/dev/md/ddg/rdsk/d10
      set type=lofs
      end

  Configuration change to sczone zone cluster succeeded.
```

12 When finished, exit the `clsetup` utility.

13 Verify the addition of the LOFS file system.

```
phys-schost# clzonecluster show -v zoneclustername
```

Next Steps (Optional) Configure the cluster file system to be managed by an HAStoragePlus resource. The HAStoragePlus resource manages by mounting the file system in the global cluster, and later performing a loopback mount on the zone-cluster nodes that currently host the applications that are configured to use the file system. For more information, see "Configuring an HAStoragePlus Resource for Cluster File Systems" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

Adding Local File Systems to a Specific Zone-Cluster Node

This section describes how to add file systems that are dedicated to a single zone-cluster node. To instead configure file systems for use by the entire zone cluster, go to "Adding File Systems to a Zone Cluster" on page 210.

This section contains the following procedures:

- "How to Add a Local File System to a Specific Zone-Cluster Node" on page 217
- "How to Add a Local ZFS Storage Pool to a Specific Zone-Cluster Node" on page 218

▼ How to Add a Local File System to a Specific Zone-Cluster Node

Perform this procedure to add a local file system to a single, specific zone-cluster node of a specific zone cluster. The file system is not managed by Oracle Solaris Cluster software but is instead passed to the underlying Oracle Solaris zone.

Note – To add a highly available local file system to a zone cluster, perform procedures in "How to Add a Highly Available Local File System to a Zone Cluster" on page 210.

---

1 Become superuser on a node of the global cluster that hosts the zone cluster.

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

---

2 Create the local file system that you want to configure to a specific zone-cluster node.

Use local disks of the global-cluster node that hosts the intended zone-cluster node.

3 Add the file system to the zone-cluster configuration in the node scope.

```
phys-schost# clzonecluster configure zoneclustername
clzc:zoneclustername> select node physical-host=baseclusternode
clzc:zoneclustername:node> add fs
clzc:zoneclustername:node:fs> set dir=mountpoint
clzc:zoneclustername:node:fs> set special=disk-device-name
clzc:zoneclustername:node:fs> set raw=raw-disk-device-name
clzc:zoneclustername:node:fs> set type=FS-type
clzc:zoneclustername:node:fs> end
clzc:zoneclustername:node> end
clzc:zoneclustername> verify
clzc:zoneclustername> commit
clzc:zoneclustername> exit
```

`dir=`*mountpoint*
: Specifies the file-system mount point

`special=`*disk-device-name*
: Specifies the name of the disk device

`raw=`*raw-disk-device-name*
Specifies the name of the raw-disk device

`type=`*FS-type*
Specifies the type of file system

Note – Enable logging for UFS file systems.

---

4 Verify the addition of the file system.

```
phys-schost# clzonecluster show -v zoneclustername
```

Example 6–5 Adding a Local File System to a Zone-Cluster Node

This example adds a local UFS file system `/local/data` for use by a node of the `sczone` zone cluster. This zone-cluster node is hosted on global—cluster node `phys-schost-1`.

```
phys-schost-1# clzonecluster configure sczone
clzc:sczone> select node physical-host=phys-schost-1
clzc:sczone:node> add fs
clzc:sczone:node:fs> set dir=/local/data
clzc:sczone:node:fs> set special=/dev/md/localdg/dsk/d1
clzc:sczone:node:fs> set raw=/dev/md/localdg/rdsk/d1
clzc:sczone:node:fs> set type=ufs
clzc:sczone:node:fs> add options [logging]
clzc:sczone:node:fs> end
clzc:sczone:node> end
clzc:sczone> verify
clzc:sczone> commit
clzc:sczone> exit

phys-schost-1# clzonecluster show -v sczone
...
     --- Solaris Resources for phys-schost-1 ---
...
   Resource Name:                                fs
     dir:                                           /local/data
     special:                                       /dev/md/localdg/dsk/d1
     raw:                                           /dev/md/localdg/rdsk/d1
     type:                                          ufs
     options:                                       [logging]
     cluster-control:                               false ...
```

▼ How to Add a Local ZFS Storage Pool to a Specific Zone-Cluster Node

Perform this procedure to add a local ZFS storage pool to a specific zone-cluster node. The local ZFS pool is not managed by Oracle Solaris Cluster software but is instead passed to the underlying Oracle Solaris zone.

Note – To add a highly available local ZFS pool to a zone cluster, see "How to Add a Highly Available Local File System to a Zone Cluster" on page 210.

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

1 **Become superuser on a node of the global cluster that hosts the zone cluster.**

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

2 **Create the local ZFS pool that you want to configure to a specific zone-cluster node.**

Use local disks of the global-cluster node that hosts the intended zone-cluster node.

3 **Add the pool to the zone-cluster configuration in the node scope.**

```
phys-schost# clzonecluster configure zoneclustername
clzc:zoneclustername> select node physical-host=baseclusternode
clzc:zoneclustername:node> add dataset
clzc:zoneclustername:node:dataset> set name=localZFSpoolname
clzc:zoneclustername:node:dataset> end
clzc:zoneclustername:node> end
clzc:zoneclustername> verify
clzc:zoneclustername> commit
clzc:zoneclustername> exit
```

`set name=`*localZFSpoolname*
Specifies the name of the local ZFS pool

4 **Verify the addition of the ZFS pool.**

```
phys-schost# clzonecluster show -v zoneclustername
```

Example 6–6 Adding a Local ZFS Pool to a Zone-Cluster Node

This example adds the local ZFS pool `local_pool` for use by a node of the `sczone` zone cluster. This zone-cluster node is hosted on global—cluster node `phys-schost-1`.

```
phys-schost-1# clzonecluster configure sczone
clzc:sczone> select node physical-host=phys-schost-1
clzc:sczone:node> add dataset
clzc:sczone:node:dataset> set name=local_pool
clzc:sczone:node:dataset> end
clzc:sczone:node> end
clzc:sczone> verify
clzc:sczone> commit
clzc:sczone> exit

phys-schost-1# clzonecluster show -v sczone
...
     --- Solaris Resources for phys-schost-1 ---
```

```
...
  Resource Name:                                dataset
    name:                                         local_pool
```

Adding Storage Devices to a Zone Cluster

This section describes how to add the direct use of global storage devices by a zone cluster or add storage devices that are dedicated to a single zone-cluster node. Global devices are devices that can be accessed by more than one node in the cluster, either one node at a time or multiple nodes concurrently.

After a device is added to a zone cluster, the device is visible only from within that zone cluster.

This section contains the following procedures:

- "How to Add a Global Storage Device to a Zone Cluster" on page 220
- "How to Add a Raw-Disk Device to a Specific Zone—Cluster Node" on page 222

How to Add a Global Storage Device to a Zone Cluster

Perform this procedure to add one of the following types of storage devices in cluster scope:

- Raw-disk devices
- Solaris Volume Manager disk sets (not including multi-owner)

Note – To add a raw-disk device to a specific zone-cluster node, go instead to "How to Add a Raw-Disk Device to a Specific Zone—Cluster Node" on page 222.

The `clsetup` utility discovers and displays the available storage devices that are configured on the cluster nodes where the selected zone cluster is configured. After you use the `clsetup` utility to add a storage device to an existing zone cluster , use the `clzonecluster` command to modify the configuration. For instructions on using the `clzonecluster` command to remove a storage device from a zone cluster, see "How to Remove a Storage Device From a Zone Cluster" in *Oracle Solaris Cluster System Administration Guide*.

1 Become superuser on a node of the global cluster that hosts the zone cluster.

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

---

2 Identify the device to add to the zone cluster and determine whether it is online.

```
phys-schost# cldevicegroup status
```

3 If the device that you are adding is not online, bring it online.

```
phys-schost# cldevicegroup online device
```

4 **Start the `clsetup` utility.**

```
phys-schost# clsetup
```

The Main Menu is displayed.

Tip – To return to a previous screen, type the < key and press Return.

---

5 **Choose the Zone Cluster menu item.**

The Zone Cluster Tasks Menu is displayed.

6 **Choose the Add File System/Storage Device to a Zone Cluster menu item.**

The Select Zone Cluster menu is displayed.

7 **Choose the zone cluster where you want to add the storage device.**

The Storage Type Selection menu is displayed.

8 **Choose the Device menu item.**

A list of the available devices is displayed.

9 **Choose a storage device from the list.**

You can also type **e** to manually specify properties for a storage device.

The Storage Device Property for the Zone Cluster menu is displayed.

10 **Add or change any properties for the storage device you are adding.**

Note – An asterisk (*) is used as a wildcard character in the path name.

When, finished, type **d** and press Return. The Review File Systems/Storage Devices for the Zone Cluster menu is displayed.

11 **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 file systems or storage devices to sczone zone cluster...

    The zone cluster is being created with the following configuration

        /usr/cluster/bin/clzonecluster configure sczone
        add device
        set match=/dev/md/ddg/*dsk/*
        end
        add device
        set match=/dev/md/shared/1/*dsk/*
        end
```

```
Configuration change to sczone zone cluster succeeded.
The change will become effective after the zone cluster reboots.
```

12 **When finished, exit the `clsetup` utility.**

13 **Verify the addition of the device.**

```
phys-schost# clzonecluster show -v zoneclustername
```

▼ How to Add a Raw-Disk Device to a Specific Zone—Cluster Node

Perform this procedure to add a raw-disk device to a specific zone-cluster node. This device would not be under Oracle Solaris Cluster control. Perform all steps of the procedure from a node of the global cluster.

Note – To add a raw-disk device for use by the full zone cluster, go instead to "How to Add a Global Storage Device to a Zone Cluster" on page 220.

---

1 **Become superuser on a node of the global cluster that hosts the zone cluster.**

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

---

2 **Identify the device (c*N*t*X*d*Y*s*Z*) to add to the zone cluster and determine whether it is online.**

3 **Add the device to the zone-cluster configuration in the node scope.**

Note – An asterisk (*) is used as a wildcard character in the path name.

```
phys-schost# clzonecluster configure zone-cluster-name
clzc:zone-cluster-name> select node physical-host=baseclusternode
clzc:zone-cluster-name:node> add device
clzc:zone-cluster-name:node:device> set match=/dev/*dsk/cNtXdYs*
clzc:zone-cluster-name:node:device> end
clzc:zone-cluster-name:node> end
clzc:zone-cluster-name> verify
clzc:zone-cluster-name> commit
clzc:zone-cluster-name> exit
```

`match=/dev/*dsk/c`*N*`t`*X*`d`*Y*`s*`

Specifies the full device path of the raw-disk device

4 **Verify the addition of the device.**

```
phys-schost# clzonecluster show -v zoneclustername
```

Example 6–7 Adding a Raw-Disk Device to a Specific Zone-Cluster Node

The following example adds the raw–disk device `c1t1d0s0` for use by a node of the `sczone` zone cluster. This zone-cluster node is hosted on global—cluster node `phys-schost-1`.

```
phys-schost-1# clzonecluster configure sczone
clzc:sczone> select node physical-host=phys-schost-1
clzc:sczone:node> add device
clzc:sczone:node:device> set match=/dev/*dsk/c1t1d0s0
clzc:sczone:node:device> end
clzc:sczone:node> end
clzc:sczone> verify
clzc:sczone> commit
clzc:sczone> exit

phys-schost-1# clzonecluster show -v sczone
...
     --- Solaris Resources for phys-schost-1 ---
...
   Resource Name:                                device
     name:                                          /dev/*dsk/c1t1d0s0
```

◆ ◆ ◆ CHAPTER 7

Uninstalling Software From the Cluster

This chapter provides procedures for uninstalling or removing certain software from an Oracle Solaris Cluster configuration.

Note – If you want to uninstall a node from an established cluster, see "Removing a Node From a Cluster" in *Oracle Solaris Cluster System Administration Guide*.

Uninstalling the Software

This section provides the following procedures to uninstall or remove certain software products from a global cluster.

▼ How to Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems

Perform this procedure if the installed node cannot join the cluster or if you need to correct configuration information. For example, perform this procedure on all nodes to reconfigure the transport adapters or the private-network address.

Note – If the node has already joined the cluster and is no longer in installation mode, as described in Step 2 of "How to Verify the Quorum Configuration and Installation Mode" on page 132, do not perform this procedure. Instead, go to "How to Uninstall Oracle Solaris Cluster Software From a Cluster Node" in *Oracle Solaris Cluster System Administration Guide*.

Before You Begin Attempt to rerun cluster configuration of the node by using the `scinstall` utility. You can correct certain cluster node configuration failures by repeating Oracle Solaris Cluster software configuration on the node.

1 Add to the cluster's node-authentication list each node that you intend to unconfigure.

If you are unconfiguring a single-node cluster, skip to Step 2.

a. On an active cluster member other than the node that you are unconfiguring, become superuser.

b. Specify the name of the node to add to the authentication list.

```
phys-schost# /usr/cluster/bin/claccess allow -h nodename
```

`-h` *nodename*
Specifies the name of the node to add to the authentication list.

You can also use the `clsetup` utility to perform this task. See "How to Add a Node to an Existing Cluster" in *Oracle Solaris Cluster System Administration Guide* for procedures.

2 On a node that you intend to unconfigure, become superuser.

3 Shut down the node.

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

4 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 an x86 Based System by Using GRUB (Task Map)" in *Oracle Solaris Administration: Basic Administration*.

  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.

---

5 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 /
```

6 Remove the node from the cluster configuration.

```
phys-schost# /usr/cluster/bin/clnode remove
```

The node is removed from the cluster configuration but Oracle Solaris Cluster software is not removed from the node.

See the `clnode`(1CL) man page for more information.

7 Repeat Step 2 through Step 6 for each additional node to unconfigure.

8 (Optional) Uninstall Oracle Solaris Cluster framework and data service software packages.

Note – If you do not need to remove or reinstall Oracle Solaris Cluster software, you can skip this step.

This step also removes the Oracle Solaris Cluster entry from the `installer` program product registry. If the product registry contains a record that Oracle Solaris Cluster software is installed, the `installer` program will show the Oracle Solaris Cluster component grayed out and will not permit reinstallation.

a. Start the `uninstall` program.

Run the following command, where *ver* is the version of the Java ES distribution from which you installed Oracle Solaris Cluster software.

```
phys-schost# /var/sadm/prod/SUNWentsysver/uninstall
```

b. Follow the onscreen instructions to select the Oracle Solaris Cluster components to uninstall.

Note – If Oracle Solaris Cluster Geographic Edition software is installed, you must uninstall it as well.

For more information about using the `uninstall` program, see Chapter 8, "Uninstalling," in *Sun Java Enterprise System 5 Update 1 Installation Guide for UNIX*.

Troubleshooting If the `uninstall` program does not exist on the node, package installation might not have fully completed. Remove the `/var/sadm/install/productregistry` file, then use the `pkgrm` command to manually remove all Oracle Solaris Cluster packages.

Next Steps To reinstall or reconfigure Oracle Solaris Cluster software on the node, refer to Table 2–1. This table lists all installation tasks and the order in which to perform the tasks.

To physically remove the node from the cluster, see "How to Remove an Interconnect Component" in *Oracle Solaris Cluster 3.3 3/13 Hardware Administration Manual* and the removal procedure in the Oracle Solaris ClusterHardware Administration Collection manual for your storage array.

▼ How to Remove Oracle Solaris Cluster Information From a JumpStart Install Server

Perform this procedure to remove Oracle Solaris Cluster information from a JumpStart install server that was used to install and configure a cluster. You can remove information for one or more individual nodes or for one or more entire clusters. For more information about the JumpStart feature, see *Oracle Solaris 10 1/13 Installation Guide: JumpStart Installations*.

1 Become superuser on the JumpStart install server.

2 Change to the JumpStart directory that you used to install Oracle Solaris Cluster software.

```
installserver# cd jumpstart-dir
```

3 Remove from the `rules` file any entries that were created by the `scinstall` command that contain the name of a node that you want to remove.

The Oracle Solaris Cluster entries reference `autostinstall.class` or `autoscinstall.finish` or both. Entries appear similar to the following, where *release* is the version of Oracle Solaris Cluster software:

```
hostname phys-schost-1  -  autoscinstall.d/release/autoscinstall.class  \
autoscinstall.d/release/autoscinstall.finish
```

4 **Regenerate the `rules.ok` file.**

Run the `check` command that is located in the *jumpstart-dir/* directory to regenerate the `rules.ok` file.

```
installserver# ./check
```

5 **Remove from the appropriate `clusters/`*clustername*`/` directory the symbolic link for each node that you want to remove.**

- To remove the symbolic link for one or more nodes in a cluster, remove the link that is named for each node to remove.

  ```
  installserver# rm -f autoscinstall.d/clusters/clustername/nodename
  ```
- To remove the symbolic links for an entire cluster, recursively remove the directory that is named for the cluster to remove.

  ```
  installserver# rm -rf autoscinstall.d/clusters/clustername
  ```
- To remove the symbolic links for all clusters, recursively remove the `clusters/` directory.

  ```
  installserver# rm -rf autoscinstall.d/clusters
  ```

6 **Remove from the `autoscinstall.d/` directory the node configuration directory that is named for each node that you want to remove.**

If you are removing information for an entire cluster, remove the directory for each node in the cluster.

- To remove information for one or more nodes in a cluster, recursively remove the directory for each node.

  ```
  installserver# rm -rf autoscinstall.d/nodes/nodename
  ```
- To remove all entries for all clusters, recursively remove the `autoscinstall.d` directory.

  ```
  installserver# rm -rf autoscinstall.d
  ```

7 **Remove the `.autoscinstall.log.3` file.**

```
installserver# rm .autoscinstall.log.3
```

8 **(Optional) If you used a flash archive to JumpStart install the cluster, remove the flash archive if you no longer need the file.**

```
installserver# rm filename.flar
```

Next Steps If you intend to use custom JumpStart to reinstall a cluster from which you removed information for one or more nodes that were removed from that cluster, you must rerun interactive `scinstall` to update the cluster node list. See "How to Install Oracle Solaris and Oracle Solaris Cluster Software (JumpStart)" on page 90.

How to Uninstall SunPlex Manager Software

Use this procedure to uninstall SunPlex Manager software that was installed by the `installer` utility up to and including the Sun Java Enterprise System 2005Q4 distribution or by any other installation method.

To remove Oracle Solaris Cluster Manager software that was installed with the Sun Java Enterprise System 5 or compatible distribution of the `installer` utility, instead use the `uninstall` utility to remove these packages. For more information, see Chapter 8, "Uninstalling," in *Sun Java Enterprise System 5 Update 1 Installation Guide for UNIX*.

Note – Do not uninstall SunPlex Manager or Oracle Solaris Cluster Manager or its shared components if you intend to use the graphical user interface (GUI) to Oracle Solaris Cluster Geographic Edition software or to Oracle Solaris Cluster features that are available through the GUI. These features include the data-service configuration wizards or system resource monitoring.

However, if you use the command-line interface to administer these features, it is safe to uninstall SunPlex Manager or Oracle Solaris Cluster Manager software.

Perform this procedure on each node in the cluster to uninstall SunPlex Manager software and related Sun Java Enterprise System shared components.

Note – SunPlex Manager software must be installed on all cluster nodes or on none.

---

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

2. **Remove the SunPlex Manager software packages.**

   ```
   phys-schost# pkgrm SUNWscspm SUNWscspmu SUNWscspmr
   ```

3. **(Optional) Remove Oracle Java Web Console software packages, if you have no other need for them.**

   ```
   phys-schost# pkgrm SUNWmctag SUNWmconr SUNWmcon SUNWmcos SUNWmcosx
   ```

4. **(Optional) If you removed Oracle Java Web Console packages, remove Apache Tomcat and Java Studio Enterprise Web Application Framework (Java ATO) software packages, if you have no other need for them.**

   Remove the packages listed below for each additional product that you want to uninstall, in the order that the set of packages is listed.

   ```
   phys-schost# pkgrm packages
   ```

Product	Package Names
Apache Tomcat	SUNWtcatu
Java ATO	SUNWjato SUNWjatodmo SUNWjatodoc

▼ How to Uninstall Oracle Solaris Cluster Quorum Server Software

Before You Begin Before you uninstall Quorum Server software, ensure that you have completed the following tasks:

- On each cluster that uses the quorum server, remove the quorum server as a quorum device. Perform the steps in "How to Remove a Quorum Device" in *Oracle Solaris Cluster System Administration Guide*.

  In normal operation, this step also removes the quorum server information on the quorum-server host. If communications between the cluster and the quorum-server host computer are down during this step, you must clean up the invalid quorum server information on the quorum-server host computer. Perform the steps in "Cleaning Up Stale Quorum Server Cluster Information" in *Oracle Solaris Cluster System Administration Guide*.
- On each quorum-server host computer, stop the quorum server by following the steps in "How to Stop a Quorum Server" in *Oracle Solaris Cluster System Administration Guide*.

1 Become superuser on the quorum server host computer to uninstall.

Use the following command if you want to ensure that the `installer` program can display the GUI.

```
# ssh -X [-l root] quorumserver
```

2 Navigate to the directory where the uninstaller is located.

```
quorumserver# cd /var/sadm/prod/SUNWentsysver
```

ver The version that is installed on your system.

3 Start the uninstallation wizard.

```
quorumserver# ./uninstall
```

4 Follow instructions on the screen to uninstall the Quorum Server software from the quorum-server host computer.

After removal is finished, you can view any available log. See Chapter 8, "Uninstalling," in *Sun Java Enterprise System 5 Update 1 Installation Guide for UNIX* for additional information about using the `uninstall` program.

5 **(Optional) Clean up or remove the quorum server directories.**

By default, this directory is `/var/scqsd`.

How to Unconfigure a Zone Cluster

Perform this procedure to remove a zone cluster.

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

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

2 **Take offline each resource group in the zone cluster and disable its resources.**

Note – The following steps are performed from a global-cluster node. To instead perform these steps from a node of the zone cluster, log in to the zone-cluster node and omit "`-Z` *zonecluster*" from each command.

a. **Take each resource offline.**

```
phys-schost# clresource offline -Z zonecluster resource-group
```

b. **List all enabled resources in the zone cluster.**

```
phys-schost# clresource show -Z zonecluster -p Enabled
=== Resources ===

Resource:                                       resource
  Enabled{nodename1}:                               True
  Enabled{nodename2}:                               True
...
```

c. **Identify those resources that depend on other resources.**

```
phys-schost# clresource show -Z zonecluster -p resource_dependencies
=== Resources ===

Resource:                                       node
  Resource_dependencies:                          node
...
```

You must disable dependent resources first before you disable the resources that they depend on.

d. **Disable each enabled resource in the cluster.**

```
phys-schost# clresource disable -Z zonecluster resource
```

See the `clresource`(1CL) man page for more information.

e. Verify that all resources are disabled.

```
phys-schost# clresource show -Z zonecluster -p Enabled
=== Resources ===

Resource:                                       resource
  Enabled{nodename1}:                               False
  Enabled{nodename2}:                               False
...
```

f. Move each resource group to the unmanaged state.

```
phys-schost# clresourcegroup unmanage -Z zonecluster resource-group
```

g. Verify that all resources on all nodes are `Offline` and that all resource groups are in the `Unmanaged` state.

```
phys-schost# cluster status -Z zonecluster -t resource,resourcegroup
```

h. Delete all resource groups and their resources from the zone cluster.

```
phys-schost# clresourcegroup delete -F -Z zonecluster +
```

3 Halt the zone cluster.

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

4 Uninstall the zone cluster.

```
phys-schost# clzonecluster uninstall zoneclustername
```

5 Unconfigure the zone cluster.

```
phys-schost# clzonecluster delete zoneclustername
```

Index

D

E

F

M

N

O

P

Q

R

S

T

U

V

X

Z