Oracle® Solaris Cluster Software Installation Guide

ORACLE®

Contents

Preface

The *Oracle Solaris Cluster Software Installation Guide* contains guidelines and procedures for installing the Oracle Solaris Cluster 4.0 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. In this document, "x86" refers to the larger family of 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.

Bash is the default shell for Oracle Solaris 11. Machine names shown with the Bash shell prompt are displayed for clarity.

Using UNIX Commands

This document contains information about commands that are specific to installing and configuring Oracle Solaris Cluster data services. The document does *not* contain comprehensive information about basic UNIX commands and procedures, such as shutting down the system, booting the system, and configuring devices. Information about basic UNIX commands and procedures is available from the following sources:

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

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

TABLE P–2   Shell Prompts

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

Related Documentation

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

Topic	Documentation
Hardware installation and administration	*Oracle Solaris Cluster 4.0 Hardware Administration Manual*
	Individual hardware administration guides
Concepts	*Oracle Solaris Cluster Concepts Guide*
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
	Oracle Solaris Cluster Geographic Edition Reference Manual
	Oracle Solaris Cluster Quorum Server 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, contact your service provider and provide the following information.

- Your name and email address (if available)
- Your company name, address, and phone number
- The model number and serial number of your systems
- The release number of the operating environment (for example, Oracle Solaris 11)
- The release number of Oracle Solaris Cluster (for example, Oracle Solaris Cluster 4.0)

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
pkg list	Reports which packages are installed
prtdiag -v	Displays system diagnostic information
/usr/cluster/bin/clnode show-rev	Displays Oracle Solaris Cluster release and package version information for each node

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

1

Planning the Oracle Solaris Cluster Configuration

This chapter provides planning information and guidelines specific to an Oracle Solaris Cluster 4.0 configuration.

This chapter contains the following overview information:

- "Finding Oracle Solaris Cluster Installation Tasks" on page 11
- "Planning the Oracle Solaris OS" on page 12
- "Planning the Oracle Solaris Cluster Environment" on page 16
- "Planning Global Devices, Device Groups, and Cluster File Systems" on page 33
- "Planning Volume Management" on page 37

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 4.0 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"
Establish a new global cluster or a new global-cluster node.	"Establishing a New Global Cluster or New Global-Cluster Node" on page 62
Configure Solaris Volume Manager software.	"Configuring Solaris Volume Manager Software" on page 129 *Solaris Volume Manager Administration Guide*
Configure cluster file systems, if used.	"How to Create Cluster File Systems" on page 143

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

Task	Instructions
(Optional) Create zone clusters.	"Configuring a Zone Cluster" on page 147
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
- "System Disk Partitions" on page 14
- "SPARC: Guidelines for Oracle VM Server for SPARC in a Cluster" on page 15

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 Automated Installer (AI) 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 AI installation method. During AI installation of Oracle Solaris software, you choose to either install the OS with defaults accepted or run an interactive installation of the OS where you can customize the installation for components such as the boot disk and the ZFS root pool. If you are installing several cluster nodes, consider a network installation.

See "How to Install and Configure Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 84 for details about the `scinstall` AI installation method. See your Oracle Solaris installation documentation for details about standard Oracle Solaris installation methods and what configuration choices you must make during installation of the OS.

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.

- **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:

  - 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 poweradm(1M) man page for more information.

- **Network Auto-Magic (NWAM)** – The Oracle Solaris Network Auto-Magic (NWAM) feature activates a single network interface and disables all others. For this reason, NWAM cannot coexist with the Oracle Solaris Cluster software and you must disable it before you configure or run your cluster.

- **IP Filter** – Oracle Solaris Cluster relies on IP network multipathing (IPMP) for public network monitoring. Any IP Filter configuration must be made in accordance with IPMP configuration guidelines and restrictions concerning IP Filter.

- **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.

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.

- **root (/)** – The primary space requirements for the root (/) file system are as follows:
  - The Oracle Solaris Cluster software itself occupies less than 40 Mbytes of space in the root (/) file system.
  - Each Oracle Solaris Cluster data service might use between 1 Mbyte and 5 Mbytes.
  - 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.
  - 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 log files.
  - The lofi device for the global-devices namespace requires 100 MBytes of free space. Starting in Oracle Solaris Cluster 4.0, a dedicated partition is no longer used for the global-devices namespace.

    Note – The lofi device that is created for the global-devices namespace is restricted to that use only. Do not use this device for any other purpose, and never unmount the device.

- **/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.

- **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.

- **Volume manager** – Create a 20-Mbyte partition on slice 6 for volume manager use.

To support Solaris Volume Manager, you can create this partition on one of the following locations:

- A local disk other than the ZFS root pool
- The ZFS root pool, if the ZFS root pool is on a partition rather than a disk

Set aside a slice for this purpose on each local disk. However, 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 *Solaris Volume Manager Administration Guide* for more information.

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

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 halts. 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 when you perform initial configuration 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.

For more information about Oracle VM Server for SPARC, see the *Oracle VM Server for SPARC 2.1 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 Updates

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

For general guidelines and procedures for applying software updates, see Chapter 11, "Updating Your Software," 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. The number of addresses that you need depends on which components you include in 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 1–2   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

TABLE 1–2   Oracle Solaris Cluster Components That Use Public-Network IP Addresses   *(Continued)*

Component	Number of IP Addresses Needed
Domain console network interface	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 1, "Planning the Network Deployment," in *Oracle Solaris Administration: IP Services*.

Console-Access Devices

You must have console access to all cluster nodes. A service processor (SP) is used to communicate between the administrative console and the global-cluster node consoles.

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

You can use the Oracle Solaris pconsole utility to connect to the cluster nodes. The utility also provides a master console window from which you can propagate your input to all connections that you opened. For more information, see the pconsole(1) man page that is available when you install the Oracle Solaris 11 terminal/pconsole package.

Public Network Configuration

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.

  Alternatively, create virtual NICs on the same physical interface and assign different virtual NICs to the private and public networks.

- **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.

- **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 additional information about data services and resources, also see the *Oracle Solaris Cluster Concepts Guide*.

- **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.

  Unless there are one or more non-link-local IPv6 public network interfaces in the public network configuration, the scinstall utility automatically configures a multiple-adapter IPMP group for each set of public-network adapters in the cluster that uses the same subnet. These groups are link-based with transitive probes.

  If the configuration contains any non-link-local IPv6 public network interfaces, you must manually configure in IPMP groups all interfaces that will be used for data-service traffic. You can configure the IPMP groups either before or after the cluster is established.

  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. 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 on IPMP groups, see Chapter 14, "Introducing IPMP," in *Oracle Solaris Administration: Network Interfaces and Network Virtualization*. 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 15, "Administering IPMP," in *Oracle Solaris Administration: Network Interfaces and Network Virtualization*.

- **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.
- **Restriction for non-global zones** – In the Oracle Solaris Cluster 4.0 release, a quorum server cannot be installed and configured in a non-global zone.
- **Service to multiple clusters** – You can configure a quorum server 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 for which it provides quorum. 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 11 OS.
- **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 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 situation 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 feature 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.

- **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.

- **Install servers** – Do not use an Oracle Solaris Cluster configuration to provide a highly available installation service on client systems.

- **RARP** – Do not use an Oracle Solaris Cluster configuration to provide an rarpd service.

- **Remote procedure call (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.

See the *Oracle Solaris Cluster Concepts Guide* for further information about cluster time. For more information about NTP, see the ntpd(1M) man page that is delivered in the Oracle Solaris 11 service/network/ntp package.

Oracle Solaris Cluster Configurable Components

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

- "Global-Cluster Name" on page 23
- "Global-Cluster Voting-Node Names and Node IDs" on page 23
- "Private Network Configuration" on page 24

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 30.

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. Node names must be unique throughout the 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 30.

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. However, 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.

- **Oracle VM Server for SPARC** - When guest domains are created on the same physical machine and 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. You do not have to configure a separate VLAN for each cluster. However, for the highest level of fault isolation and interconnect resilience, limit the use of a VLAN to a single cluster.

- **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. However, internode communication on the private network does not use these IPv6 addresses.

See Chapter 1, "Planning the Network Deployment," 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*node-id*-priv, where *node-id* 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 27.
- 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 utility.

For guidelines about cluster interconnect hardware, see "Interconnect Requirements and Restrictions" in *Oracle Solaris Cluster 4.0 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 for scalable data services, 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. You must use the dladm create-vlan command to configure the adapter as a tagged VLAN adapter before you configure it with the cluster.

  To configure a tagged VLAN adapter for the cluster interconnect, 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 net2, the VLAN instance number would be calculated as (1000*73)+2. You would therefore specify the adapter name as net73002 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 4.0 Hardware Administration Manual*. For information about creating and administering VLANs, see the dladm(1M) man page and Chapter 13, "Administering VLANs," in *Oracle Solaris Administration: Network Interfaces and Network Virtualization*.

- **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.

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.

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.

If any device is unable to use the SCSI-3 protocol, the pathcount setting should be used instead, where the fencing protocol for the shared disk is chosen based on the number of DID paths that are attached to the disk. Non-SCSI-3 capable devices are limited to two DID device paths within the cluster. Fencing can be turned off for devices which do not support either SCSI-3 or SCSI-2 fencing. However, data integrity for such devices cannot be guaranteed during split-brain situations.

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 ones described, 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 have the scinstall utility automatically configure an available shared disk in the configuration as a quorum device. The scinstall utility assumes that all available shared disks are supported as quorum devices.

If you want to use a quorum server or a Sun ZFS Storage Appliance NAS device from Oracle as the quorum device, you configure it after scinstall processing is completed.

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 the 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 non-global Oracle Solaris zones. All nodes of a zone cluster are configured as non-global zones of the solaris brand that are set with the cluster attribute. No

other brand type is permitted in a zone cluster. The isolation that is provided by the Oracle Solaris Zones feature enables you to run supported services on the zone cluster in a similar way as running the services in a global cluster.

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

-
-

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 ($cNtXdYsZ$)

  - 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. 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 26.

- **solaris brand** – All nodes of a zone cluster are configured as non-global zones of the solaris brand that is set with the cluster attribute. No other brand type is permitted in a zone cluster.

- **IP type** - A zone cluster is created with the shared IP type. The exclusive IP type is not supported for zone clusters.

- 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 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.

  - By direct mount:

- UFS local file system
- Oracle Solaris ZFS (exported as a data set)
- NFS from supported NAS devices

- By loopback mount:

  - UFS local file system
  - 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.

Planning Global Devices, Device Groups, and Cluster File Systems

This section provides the following information:

- "Planning Global Devices" on page 33
- "Planning Device Groups" on page 34
- "Planning Cluster File Systems" on page 34
- "Choosing Mount Options for UFS Cluster File Systems" on page 35
- "Mount Information for Cluster File Systems" on page 37

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 data from a cluster file system 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 39 for additional guidelines. See "Configuring Solaris Volume Manager Software" on page 129 and your volume-manager documentation for instructions about mirroring.

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.
- **Zone clusters** – You cannot configure cluster file systems that use UFS for use in a zone cluster. Use highly available local file systems instead.
- **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:

- 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. Therefore, do not attempt to use the `fattach` command from any node other than the local node.

  - 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.

- **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 the guidelines in the following list of mount options to determine what mount options to use when you create your UFS cluster file systems.

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 RAC 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. Ignoring 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 is for cluster file systems on UFS. You can nest the mount points if the devices for the two file systems have the same physical host connectivity, for example, different slices on the same disk.

- 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 38
- "Guidelines for Solaris Volume Manager Software" on page 38
- "UFS Cluster File System Logging" on page 39
- "Mirroring Guidelines" on page 39

Oracle Solaris Cluster software uses volume manager software to group disks into device groups that can then be administered as one unit. 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 129 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 39 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 ZFS root pool ensures high availability, but such mirroring is not required. See "Mirroring Guidelines" on page 39 for guidelines to help determine whether to mirror the ZFS root pool.

- **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 software 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:

- **Unique naming** – Disk set names must be unique throughout the cluster.

- **Disk set reserved names** – Do not name a disk set admin or shared.

- **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 or three Oracle Solaris hosts is called a dual-string disk set. This type of 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.

UFS Cluster File System Logging

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

Mirroring Guidelines

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

- "Guidelines for Mirroring Multihost Disks" on page 39
- "Guidelines for Mirroring the ZFS Root Pool" on page 40

Guidelines for Mirroring Multihost Disks

Mirroring 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 or plex 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 or plex.

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

Guidelines for Mirroring the ZFS Root Pool

Oracle Solaris ZFS is the default root file system in the Oracle Solaris 11 release. See "How to Configure a Mirrored Root Pool" in *Oracle Solaris Administration: ZFS File Systems* for instructions about how to mirror the ZFS root pool. Also see Chapter 5, "Managing ZFS Root Pool Components," in *Oracle Solaris Administration: ZFS File Systems* for information about how to manage the different root pool components.

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 ZFS root pool.

Consider the following points when you decide whether to mirror the ZFS root pool:

- **Boot disk** – You can set up the mirror to be a bootable root pool. You can then boot from the mirror if the primary boot disk fails.
- **Backups** – Regardless of whether you mirror the root pool, 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 pool.
- **Separate controllers** – Highest availability includes mirroring the root pool on a separate controller.

2

Installing Software on Global-Cluster Nodes

This chapter provides the following procedures to install Oracle Solaris Cluster 4.0 software on global-cluster voting nodes.

- "How to Prepare for Cluster Software Installation" on page 42
- "How to Install Oracle Solaris Software" on page 43
- "How to Install `pconsole` Software on an Administrative Console" on page 47
- "How to Install and Configure Oracle Solaris Cluster Quorum Server Software" on page 49
- "How to Configure Internal Disk Mirroring" on page 51
- "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 52
- "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53
- "How to Install the Availability Suite Feature of Oracle Solaris 11" on page 57
- "How to Set Up the Root Environment" on page 58
- "How to Configure Solaris IP Filter" on page 59

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 42

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

Task	Instructions
Install the Oracle Solaris OS on all nodes and optionally on an administrative console and a quorum server. Optionally, enable Oracle Solaris I/O multipathing on the nodes.	"How to Install Oracle Solaris Software" on page 43
(Optional) Install pconsole software on an administrative console.	"How to Install pconsole Software on an Administrative Console" on page 47
(*Optional*) Install and configure a quorum server.	"How to Install and Configure Oracle Solaris Cluster Quorum Server Software" on page 49
(*Optional*) Configure internal disk mirroring.	"How to Configure Internal Disk Mirroring" on page 51
(*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 52
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 53
(Optional) Install and configure the Availability Suite feature of Oracle Solaris software	"How to Install the Availability Suite Feature of Oracle Solaris 11" on page 57
Set up directory paths.	"How to Set Up the Root Environment" on page 58
(*Optional*) Configure Oracle Solaris IP Filter.	"How to Configure Solaris IP Filter" on page 59

▼ How to Prepare for Cluster Software Installation

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

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 4.0 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
- 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 you to completely reinstall the Oracle Solaris and Oracle Solaris Cluster software.

---

5   Obtain all necessary updates for your cluster configuration.

See Chapter 11, "Updating Your Software," in *Oracle Solaris Cluster System Administration Guide* for 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 Oracle Solaris Cluster Quorum Server Software" on page 49.
- Otherwise, if you want to use an administrative console to communicate with the cluster nodes, go to "How to Install pconsole Software on an Administrative Console" on page 47.
- 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 43 to first install Oracle Solaris software.
    - To install and configure both Oracle Solaris and Oracle Solaris Cluster software in the same operation (Automated Installer method), go to "How to Install and Configure Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 84.

▼ How to Install Oracle Solaris Software

Use this procedure to install the Oracle Solaris OS on the following systems, as applicable to your cluster configuration:

1. (Optional) An administrative console that you will install with pconsole software. For more information, see "How to Install pconsole Software on an Administrative Console" on page 47.

2. (Optional) A quorum server. For more information, see "How to Install and Configure Oracle Solaris Cluster Quorum Server Software" on page 49.

3. Each node in the global cluster, if you will not use the scinstall custom Automated Installer method to install software. For more information about Automated Installer installation of a cluster, see "How to Install and Configure Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 84.

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 the 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 4.0 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 42 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 17 for planning guidelines. See your Oracle Solaris system administrator documentation for information about using Oracle Solaris naming services.

1   **Connect to the consoles of each node.**

2   **Install the Oracle Solaris OS.**

Follow installation instructions in *Installing Oracle Solaris 11 Systems*.

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 the Oracle Solaris software. During Oracle Solaris software installation, perform the following steps:

a.   (Cluster nodes) Choose Manual Layout to set up the file systems.

- **Specify a slice that is at least 20 Mbytes in size.**

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

b. **(Cluster nodes) For ease of administration, set the same root password on each node.**

3 **Ensure that the `solaris` publisher is valid.**

```
# pkg publisher
PUBLISHER                           TYPE    STATUS  URI
solaris                             origin  online  solaris-repository
```

For information about setting the `solaris` publisher, see *Adding and Updating Oracle Solaris 11 Software Packages*.

4 **(Cluster nodes) If you will use role-based access control (RBAC) instead of the superuser role 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 *Oracle Solaris Administration: 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.

5 **(Cluster nodes) 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.

6 **Install any required Oracle Solaris OS software updates and hardware-related firmware and updates.**

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

See Chapter 11, "Updating Your Software," in *Oracle Solaris Cluster System Administration Guide* for installation instructions.

7   x86: (Cluster nodes) 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/kernel/amd64/unix -B $ZFS-BOOTFS -k
```

For more information, see "How to Boot a System With the Kernel Debugger Enabled (kmdb)" in *Booting and Shutting Down Oracle Solaris on x86 Platforms*.

8   (Cluster nodes) 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) (Cluster nodes) 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 15, "Administering IPMP," in *Oracle Solaris Administration: Network Interfaces and Network Virtualization* 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   (Optional) (Cluster nodes) If the Oracle Solaris Cluster software is not already installed and you want to use Oracle Solaris I/O multipathing, enable multipathing on each node.

⚠ **Caution –** If the 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 you want to use the pconsole utility, go to "How to Install pconsole Software on an Administrative Console" on page 47.

If you want to use a quorum server, go to "How to Install and Configure Oracle Solaris Cluster Quorum Server Software" on page 49.

If your cluster nodes support the mirroring of internal hard drives and you want to configure internal disk mirroring, go to "How to Configure Internal Disk Mirroring" on page 51.

SPARC: If you want to install Oracle VM Server for SPARC, go to "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 52.

Otherwise, install the Oracle Solaris Cluster software on the cluster nodes.

- If you already installed the Oracle Solaris OS on the cluster nodes, go to "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53.
- If you want to use the `scinstall` custom Automated Installer (AI) method to install both Oracle Solaris OS and Oracle Solaris Cluster software on the cluster nodes, go to "How to Install and Configure Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 84.

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 Install pconsole 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 Parallel Console Access (pconsole) software on an administrative console. The pconsole utility is part of the Oracle Solaris 11 `terminal/pconsole` package.

The pconsole utility creates a host terminal window for each remote host that you specify on the command line. The utility also opens a central, or master, console window that you can use to send input to all nodes at one time. For additional information, see the pconsole(1) man page that is installed with the `terminal/pconsole` package.

You can use any desktop machine that runs a version of the Oracle Solaris OS that is supported by Oracle Solaris Cluster 4.0 software as an administrative console.

Before You Begin   Ensure that a supported version of the Oracle Solaris OS and any Oracle Solaris software updates are installed on the administrative console.

1   Become superuser on the administrative console.

2   Ensure that the `solaris` and `ha-cluster` publishers are valid.

```
# pkg publisher
PUBLISHER                          TYPE     STATUS   URI
solaris                            origin   online   solaris-repository
ha-cluster                         origin   online   ha-cluster-repository
```

For information about setting the `solaris` publisher, see "Set the Publisher Origin To the File Repository URI" in *Copying and Creating Oracle Solaris 11 Package Repositories*.

3   Install the `terminal/pconsole` package.

adminconsole# **pkg install terminal/pconsole**

4   (Optional) Install the Oracle Solaris Cluster man page packages.

adminconsole# **pkg install** *pkgname* **...**

Package Name	Description
ha-cluster/system/manual	Oracle Solaris Cluster framework man pages
ha-cluster/system/manual/data-services	Oracle Solaris Cluster data service man pages
ha-cluster/service/quorum-server/manual	Oracle Solaris Cluster Quorum Server man pages
ha-cluster/geo/manual	Oracle Solaris Cluster Geographic Edition 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 on a quorum server.

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

a.   If you installed the `ha-cluster/system/manual/data-services` package, ensure that the `/opt/SUNWcluster/bin/` directory is in the `PATH`.

b.   If you installed any other man page package, ensure that the `/usr/cluster/bin/` directory is in the `PATH`.

6   Start the `pconsole` utility.

Specify in the command each node that you want to connect to.

adminconsole# **pconsole** *host***[:**port**] [...] &**

See the procedures "Logging Into the Cluster Remotely" in *Oracle Solaris Cluster System Administration Guide* and "How to Connect Securely to Cluster Consoles" in *Oracle Solaris Cluster System Administration Guide* for additional information about how to use the pconsole utility. Also see the pconsole(1) man page that is installed as part of the Oracle Solaris 11 terminal/pconsole package.

Next Steps   If you want to use a quorum server, go to "How to Install and Configure Oracle Solaris Cluster Quorum Server Software" on page 49.

If your cluster nodes support the mirroring of internal hard drives and you want to configure internal disk mirroring, go to "How to Configure Internal Disk Mirroring" on page 51.

SPARC: If you want to install Oracle VM Server for SPARC, go to "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 52.

Otherwise, install the Oracle Solaris Cluster software on the cluster nodes.

- If you already installed the Oracle Solaris OS on the cluster nodes, go to "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53.
- If you want to use the scinstall custom Automated Installer (AI) method to install both Oracle Solaris OS and Oracle Solaris Cluster software on the cluster nodes, go to "How to Install and Configure Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 84

▼ How to Install and Configure Oracle Solaris Cluster 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 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 on which you want to install the Oracle Solaris Cluster Quorum Server software.**

2   **Ensure that the solaris and ha-cluster publishers are valid.**

```
# pkg publisher
PUBLISHER                           TYPE     STATUS   URI
solaris                             origin   online   solaris-repository
ha-cluster                          origin   online   ha-cluster-repository
```

For information about setting the solaris publisher, see "Set the Publisher Origin To the File Repository URI" in *Copying and Creating Oracle Solaris 11 Package Repositories*.

3   Install the Quorum Server group package.

```
quorumserver# pkg install ha-cluster-quorum-server-full
```

4   (Optional) Add the Oracle Solaris Cluster Quorum Server binary location to your PATH environment variable.

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

5   Configure the quorum server by adding the following entry to the /etc/scqsd/scqsd.conf file to specify configuration information about the quorum server.

Identify the quorum server by specifying the port number and optionally the instance name.

- 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.

The format for the entry is as follows:

```
/usr/cluster/lib/sc/scqsd [-d quorum-directory] [-i instance-name] -p port
```

-d *quorum-directory*
  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 *instance-name*
  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.

6   (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.

7   Save and close the /etc/scqsd/scqsd.conf file.

8   Start the newly configured quorum server.

```
quorumserver# /usr/cluster/bin/clquorumserver start quorum-server
```

quorum-server
> 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   Oracle Solaris Cluster Quorum Server software consists of the following packages:

- `ha-cluster/service/quorum-server`
- `ha-cluster/service/quorum-server/locale`
- `ha-cluster/service/quorum-server/manual`
- `ha-cluster/service/quorum-server/manual/locale`

These packages are contained in the `ha-cluster/group-package/ha-cluster-quorum-server-full` and `ha-cluster/group-package/ha-cluster-quorum-server-l10n` group packages.

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

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

Next Steps   If your cluster nodes support the mirroring of internal hard drives and you want to configure internal disk mirroring, go to "How to Configure Internal Disk Mirroring" on page 51.

SPARC: If you want to install Oracle VM Server for SPARC, go to "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 52.

Otherwise, install the Oracle Solaris Cluster software on the cluster nodes.

- If you already installed the Oracle Solaris OS on the cluster nodes, go to "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53.
- If you want to use the `scinstall` custom Automated Installer (AI) method to install both Oracle Solaris OS and Oracle Solaris Cluster software on the cluster nodes, go to "How to Install and Configure Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 84.

▼ 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 4.0 Hardware Administration Manual*.

Before You Begin   Ensure that the Oracle Solaris operating system and any necessary updates 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. Provide the name of your primary disk as the first argument and 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: If you want to install Oracle VM Server for SPARC, go to "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 52.

Otherwise, install the Oracle Solaris Cluster software on the cluster nodes.

- If you already installed the Oracle Solaris OS on the cluster nodes, go to "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53.
- If you want to use the scinstall custom Automated Installer (AI) method to install both Oracle Solaris OS and Oracle Solaris Cluster software on the cluster nodes, go to "How to Install and Configure Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 84

▼ 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 *Oracle VM Server for SPARC 2.1 Administration Guide* and *Oracle VM Server for SPARC 2.1 Release Notes*.

- Read the requirements and guidelines in "SPARC: Guidelines for Oracle VM Server for SPARC in a Cluster" on page 15.

1 Become superuser on the machine.

2 Install Oracle VM Server for SPARC software and configure domains by following the procedures in Chapter 2, "Installing and Enabling Software," in *Oracle VM Server for SPARC 2.1 Administration Guide***.**

Observe the following special instructions:

- 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 51.

Otherwise, install the Oracle Solaris Cluster software packages. Go to "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53.

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

Follow this procedure to perform one or more of the following installation tasks:

- To install the Oracle Solaris Cluster framework and data service 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 data services.

Note – You cannot add or remove individual packages that are part of the `ha-cluster-minimal` group package except by complete reinstallation or uninstallation. See "How to Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems" on page 163 and "How to Uninstall Oracle Solaris Cluster Software From a Cluster Node" in *Oracle Solaris Cluster System Administration Guide* for procedures to remove the cluster framework packages.

However, you can add or remove other, optional packages without removing the `ha-cluster-minimal` group package.

Before You Begin   Perform the following tasks:

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

  If the Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for the 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 43 for more information about installing the Oracle Solaris software to meet Oracle Solaris Cluster software requirements.

- Choose which Oracle Solaris Cluster software packages to install.

  The following table lists the primary group packages for the Oracle Solaris Cluster 4.0 software and the principal features that each group package contains. You must install at least the `ha-cluster-framework-minimal` group package.

Feature	ha-cluster-full	ha-cluster-framework-full	ha-cluster-data-services-full	ha-cluster-minimal	ha-cluster-framework-minimal
Framework	X	X	X	X	X
Agents	X		X		
Localization	X	X			
Framework man pages	X	X			
Data Service man pages	X		X		
Agent Builder	X	X			
Generic Data Service	X	X	X		

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

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

     As superuser, use the following command to start the pconsole utility:

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

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

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

2. **Restore external access to remote procedure call (RPC) communication.**

   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 RPC communication service, which is required for cluster communication.

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

   ```
   # svccfg
   svc:> select network/rpc/bind
   svc:/network/rpc/bind> setprop config/local_only=false
   svc:/network/rpc/bind> quit
   # svcadm refresh network/rpc/bind:default
   # 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.

3. **Become superuser on the cluster node to install.**

   Alternatively, if your user account is assigned the System Administrator profile, issue commands as nonroot through a profile shell, or prefix the command with the pfexec command.

4. **Disable Network Auto-Magic (NWAM).**

   NWAM activates a single network interface and disables all others. For this reason, NWAM cannot coexist with the Oracle Solaris Cluster software and you must disable it before you configure or run your cluster. To disable NWAM, you enable the defaultfixed profile.

   ```
   # netadm enable -p ncp defaultfixed
   # netadm list -p ncp defaultfixed
   ```

5. **Set up the repository for the Oracle Solaris Cluster software packages.**

   - **If the cluster nodes have direct access or web proxy access to the Internet, perform the following steps.**

     a. **Go to http://pkg-register.oracle.com.**

b. **Choose `Oracle Solaris Cluster software`.**

c. **Accept the license.**

d. **Request a new certificate by choosing `Oracle Solaris Cluster software` and submitting a request.**

The certification page is displayed with download buttons for the key and the certificate.

e. **Download the key and certificate files and install them as described in the returned certification page.**

f. **Configure the ha-cluster publisher with the downloaded SSL keys and set the location of the Oracle Solaris Cluster 4.0 repository.**

In the following example the repository name is https://pkg.oracle.com/*repository-location*/.

```
# pkg set-publisher \
-k /var/pkg/ssl/Oracle_Solaris_Cluster_4.0.key.pem \
-c /var/pkg/ssl/Oracle_Solaris_Cluster_4.0.certificate.pem \
-O https://pkg.oracle.com/repository-location/ ha-cluster
```

-k /var/pkg/ssl/Oracle_Solaris_Cluster_4.0.key.pem
    Specifies the full path to the downloaded SSL key file.

-c /var/pkg/ssl/Oracle_Solaris_Cluster_4.0.certificate.pem
    Specifies the full path to the downloaded certificate file.

-O https://pkg.oracle.com/*repository-location*/
    Specifies the URL to the Oracle Solaris Cluster 4.0 package repository.

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

- **If you are using an ISO image of the software, perform the following steps.**

  a. **Download the Oracle Solaris Cluster 4.0 ISO image from Oracle Software Delivery Cloud at `http://edelivery.oracle.com/`.**

  Note – A valid Oracle license is required to access Oracle Software Delivery Cloud.

  Oracle Solaris Cluster software is part of the Oracle Solaris Product Pack. Follow online instructions to complete selection of the media pack and download the software.

  b. **Make the Oracle Solaris Cluster 4.0 ISO image available.**

  ```
  # lofiadm -a path-to-iso-image
  /dev/lofi/N
  # mount -F hsfs /dev/lofi/N /mnt
  ```

  -a *path-to-iso-image*
      Specifies the full path and file name of the ISO image.

    c. Set the location of the Oracle Solaris Cluster 4.0 package repository.

```
# pkg set-publisher -g file:///mnt/repo ha-cluster
```

6 Ensure that the `solaris` and `ha-cluster` publishers are valid.

```
# pkg publisher
PUBLISHER                           TYPE    STATUS   URI
solaris                             origin  online   solaris-repository
ha-cluster                          origin  online   ha-cluster-repository
```

For information about setting the `solaris` publisher, see "Set the Publisher Origin To the File Repository URI" in *Copying and Creating Oracle Solaris 11 Package Repositories*.

7 Install the Oracle Solaris Cluster 4.0 software.

```
# /usr/bin/pkg install package
```

8 Verify that the package installed successfully.

```
$ pkg info -r package
```

Package installation succeeded if the state is `Installed`.

9 Perform any necessary updates to the Oracle Solaris Cluster software.

See Chapter 11, "Updating Your Software," in *Oracle Solaris Cluster System Administration Guide* for installation instructions.

Next Steps If you want to use the Availability Suite feature of Oracle Solaris 11, install the Availability Suite software. Go to "How to Install the Availability Suite Feature of Oracle Solaris 11" on page 57.

Otherwise, to set up the root user environment, go to "How to Set Up the Root Environment" on page 58.

▼ How to Install the Availability Suite Feature of Oracle Solaris 11

Before You Begin Ensure that a minimum of Oracle Solaris 11 SRU 1 is installed.

1 Become superuser.

2 Ensure that the `solaris` publishers is valid.

```
# pkg publisher
PUBLISHER                           TYPE    STATUS   URI
solaris                             origin  online   solaris-repository
```

For information about setting the `solaris` publisher, see "Set the Publisher Origin To the File Repository URI" in *Copying and Creating Oracle Solaris 11 Package Repositories*.

3 Install the IPS package for the Availability Suite feature of the Oracle Solaris 11 software.

```
# /usr/bin/pkg install storage/avs
```

4 Configure the Availability Suite feature.

For details, see "Initial Configuration Settings" (http://docs.oracle.com/cd/E19359-01/819-6147-10/config_proc.html#pgfId-998170) in *Sun StorageTek Availability Suite 4.0 Software Installation and Configuration Guide*.

Next Steps To set up the root user environment, go to "How to Set Up the Root Environment" on page 58.

▼ 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: Common Tasks* for more information.

Perform this procedure on each node in the global cluster.

1 Become superuser on a cluster node.

2 Add `/usr/cluster/bin/` and `/usr/sbin/` to the PATH.

> **Note –** Always make /usr/cluster/bin the *first* entry in the PATH. This placement ensures that Oracle Solaris Cluster commands take precedence over any other binaries that have the same name, thus avoiding unexpected behavior.

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 Solaris IP Filter, go to "How to Configure Solaris IP Filter" on page 59.

Otherwise, configure Oracle Solaris Cluster software on the cluster nodes. Go to "Establishing a New Global Cluster or New Global-Cluster Node" on page 62.

▼ How to Configure Solaris IP Filter

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

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

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

Before You Begin   Read the guidelines and restrictions to follow when you configure Solaris 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 show-addr | grep *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 net1 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 Solaris 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 62.

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 147. You must establish a global cluster before you can create a zone cluster.

This chapter contains the following information:

- "Configuring Oracle Solaris Cluster Software on All Nodes (`scinstall`)" on page 64
- "How to Configure Oracle Solaris Cluster Software on All Nodes (XML)" on page 72
- "Installing and Configuring Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 80
- "How to Prepare the Cluster for Additional Global-Cluster Nodes" on page 92
- "How to Change the Private Network Configuration When Adding Nodes or Private Networks" on page 94
- "Configuring Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (`scinstall`)" on page 99
- "How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (XML File)" on page 106
- "How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 111
- "How to Configure Quorum Devices" on page 113
- "How to Verify the Quorum Configuration and Installation Mode" on page 118
- "How to Change Private Hostnames" on page 120
- "Configuring Network Time Protocol (NTP)" on page 120
- "How to Validate the Cluster" on page 123
- "How to Record Diagnostic Data of the Cluster Configuration" on page 127

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, Oracle VM Server for SPARC I/O domains, or 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 41.

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.	"Configuring Oracle Solaris Cluster Software on All Nodes (`scinstall`)" on page 64
■ Use an XML configuration file to establish the cluster.	"How to Configure Oracle Solaris Cluster Software on All Nodes (XML)" on page 72
■ Set up an Automated Installer (AI) install server. Then use the `scinstall` AI option to install the software on each node and establish the cluster.	"Installing and Configuring Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 80
Assign quorum votes and remove the cluster from installation mode, if this operation was not already performed.	"How to Configure Quorum Devices" on page 113
Validate the quorum configuration.	"How to Verify the Quorum Configuration and Installation Mode" on page 118
(Optional) Change a node's private hostname.	"How to Change Private Hostnames" on page 120
Create or modify the NTP configuration file, if not already configured.	"Configuring Network Time Protocol (NTP)" on page 120
If using a volume manager, install 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*

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

Method	Instructions
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 123
Take a baseline recording of the finished cluster configuration.	"How to Record Diagnostic Data of the Cluster Configuration" on page 127

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 92
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 94
Use one of the following methods to add a node to an existing global cluster:	
■ Configure Oracle Solaris Cluster software on the new node by using the scinstall utility.	"Configuring Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (scinstall)" on page 99
■ 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 File)" on page 106
Update the quorum configuration information.	"How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 111
Validate the quorum configuration.	"How to Verify the Quorum Configuration and Installation Mode" on page 118
(Optional) Change a node's private hostname.	"How to Change Private Hostnames" on page 120
Modify the NTP configuration.	"Configuring Network Time Protocol (NTP)" on page 120
If using a volume manager, install 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

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

Method	Instructions
Validate the cluster.	"How to Validate the Cluster" on page 123
Take a baseline recording of the finished cluster configuration.	"How to Record Diagnostic Data of the Cluster Configuration" on page 127

Configuring Oracle Solaris Cluster Software on All Nodes (scinstall)

The `scinstall` utility runs in two modes of installation, Typical or Custom. For the Typical installation of Oracle Solaris Cluster software, `scinstall` automatically specifies the following configuration defaults.

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
Installation security (DES)	Limited

Complete one of the following cluster configuration worksheets to plan your Typical mode or Custom mode installation:

- **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.)*	
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:*
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:*
Cluster Transport Adapters and Cables *(multiple-node cluster only)*	Node name (*the node from which you run scinstall*):	
	Transport adapter name:	*First:* *Second:*
	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?	*First:* Yes | No *Second:* Yes | No
	If no, what is the name of the port that you want to use?	*First:* *Second:*

Component	Description/Example	Answer
	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:*
	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?	*First:* Yes | No *Second:* Yes | No
	If no, what is the name of the port that you want to use?	*First:* *Second:*
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 (`255.255.240.0`)?	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 disable 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.*)	*First:* Yes | No *Second:* 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.*)	*First:* Yes | No *Second:* Yes | No
Check *(multiple-node cluster only)*	Do you want to interrupt cluster creation for `cluster check` errors?	Yes | No

Component	Description/Example	Answer
(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

▼ **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. For information about how to use the noninteractive forms of the `scinstall` command, such as when developing installation scripts, see the `scinstall(1M)` man page.

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.

Before You Begin   Perform the following tasks:

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

  If the Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for the 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 43 for more information about installing the Oracle Solaris software to meet Oracle Solaris Cluster software requirements.

- Ensure that NWAM is disabled. See "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53 for instructions.

- SPARC: If you are configuring Oracle VM Server for SPARC I/O domains or guest domains as cluster nodes, ensure that the 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 52.

- Ensure that Oracle Solaris Cluster software packages and updates are installed on each node. See "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53.

- Ensure that any adapters that you want to use as tagged VLAN adapters are configured and that you have their VLAN IDs.

- Have available your completed Typical Mode or Custom Mode installation worksheet. See "Configuring Oracle Solaris Cluster Software on All Nodes (scinstall)" on page 64.

1 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.

2 On each node to configure in a cluster, become superuser.

Alternatively, if your user account is assigned the System Administrator profile, issue commands as nonroot through a profile shell, or prefix the command with the pfexec command.

3 Ensure that TCP wrappers for RPC are disabled on all nodes of the cluster.

The Oracle Solaris TCP wrappers for RPC feature prevents internode communication that is necessary for cluster configuration.

a. On each node, display the status of TCP wrappers for RPC.

TCP wrappers are enabled if config/enable_tcpwrappers is set to true, as shown in the following example command output.

```
# svccfg -s rpc/bind listprop config/enable_tcpwrappers
config/enable_tcpwrappers  boolean true
```

b. If TCP wrappers for RPC are enabled on a node, disable TCP wrappers and refresh the RPC bind service.

```
# svccfg -s rpc/bind setprop config/enable_tcpwrappers = false
# svcadm refresh rpc/bind
# svcadm restart rpc/bindEntry 2
```

4 Prepare public-network interfaces.

a. Create static IP addresses for each public-network interface.

```
# ipadm create-ip interface
# ipadm create-addr -T static -a local=address/prefix-length addrobj
```

For more information, see "How to Configure an IP Interface" in *Oracle Solaris Administration: Network Interfaces and Network Virtualization*.

b. (Optional) Create IPMP groups for public-network interfaces.

During initial cluster configuration, unless non-link-local IPv6 public network interfaces exist in the cluster, IPMP groups are automatically created based on matching subnets. These groups use transitive probes for interface monitoring and no test addresses are required.

If these automatically created IPMP groups would not meet your needs, or if IPMP groups would not be created because your configuration includes one or more non-link-local IPv6 public network interfaces, do one of the following:

■ **Create the IPMP groups you need before you establish the cluster.**

■ **After the cluster is established, use the `ipadm` command to edit the IPMP groups.**

For more information, see "Configuring IPMP Groups" in *Oracle Solaris Administration: Network Interfaces and Network Virtualization*.

5 From one cluster node, start the `scinstall` utility.

```
phys-schost# scinstall
```

6 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) Print release information for this cluster node

     * ?) Help with menu options
     * q) Quit

   Option:  1
```

The New Cluster and Cluster Node Menu is displayed.

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

The Typical or Custom Mode menu is displayed.

8 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.

9 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.

10 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
```

11 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.

12 Take the cluster out of installmode.

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

13 (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.

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
...
```

14 If you plan to enable RPC use of TCP wrappers, add all `clprivnet0` IP addresses to the `/etc/hosts.allow` file on each cluster node.

Without this addition to the /etc/hosts.allow file, TCP wrappers prevent internode communication over RPC for cluster administration utilities.

a. On each node, display the IP addresses for all `clprivnet0` devices on the node.

```
# /usr/sbin/ipadm show-addr
ADDROBJ          TYPE    STATE      ADDR
clprivnet0/N     static  ok         ip-address/netmask-length
...
```

b. On each cluster node, add to the `/etc/hosts.allow` file the IP addresses of all `clprivnet0` devices in the cluster.

15 If you intend to use the HA for NFS data service (HA for NFS) on a highly available local file system, exclude from the automounter map all shares that are part of the highly available local file system that is exported by HA for NFS.

See "Administrative Tasks Involving Maps" in *Oracle Solaris Administration: Network Services* for more information about modifying the automounter map.

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 net2 and net3. The automatic selection of a quorum device is enabled.

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

 Configuring global device using lofi on pred1: done
  Starting discovery of the cluster transport configuration.

  The following connections were discovered:

      phys-schost-1:net2  switch1  phys-schost-2:net2
```

```
            phys-schost-1:net3  switch2  phys-schost-2:net3

     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".

     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
```

Troubleshooting   **Unsuccessful configuration** – If one or more nodes cannot join the cluster, or if the wrong configuration information was specified, first attempt to perform this procedure again. If that does not correct the problem, perform the procedure "How to Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems" on page 163 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Oracle Solaris Cluster software packages. Then perform this procedure again.

Next Steps   ■   If you installed a single-node cluster, cluster establishment is complete. Go to "Creating Cluster File Systems" on page 143 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 118.

   ■   If you installed a multiple-node cluster and declined automatic quorum configuration, perform postinstallation setup. Go to "How to Configure Quorum Devices" on page 113.

If you intend to configure any quorum devices in your cluster, go to "How to Configure Quorum Devices" on page 113.

Otherwise, go to "How to Verify the Quorum Configuration and Installation Mode" on page 118.

▼ 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 4.0 software.

This procedure configures the following cluster components:

- Cluster name
- Cluster node membership
- Cluster interconnect

Before You Begin Perform the following tasks:

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

  If the Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for the 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 43 for more information about installing the Oracle Solaris software to meet Oracle Solaris Cluster software requirements.

- Ensure that NWAM is disabled. See "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53 for instructions.

- SPARC: If you are configuring Oracle VM Server for SPARC I/O domains or guest domains as cluster nodes, ensure that the 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 52.

- Ensure that any adapters that you want to use as tagged VLAN adapters are configured and that you have their VLAN IDs.

- Ensure that Oracle Solaris Cluster 4.0 software and updates are installed on each node that you will configure. See "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53.

1 **Ensure that the Oracle Solaris Cluster 4.0 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 the 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 the 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 the Oracle Solaris Cluster software is already configured on the node.

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

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

2 **Ensure that TCP wrappers for RPC are disabled on all nodes of the cluster.**

The Oracle Solaris TCP wrappers for RPC feature prevents internode communication that is necessary for cluster configuration.

a. **On each node, display the status of TCP wrappers for RPC.**

TCP wrappers are enabled if config/enable_tcpwrappers is set to true, as shown in the following example command output.

```
# svccfg -s rpc/bind listprop config/enable_tcpwrappers
config/enable_tcpwrappers  boolean true
```

b. **If TCP wrappers for RPC are enabled on a node, disable TCP wrappers and refresh the RPC bind service.**

```
# svccfg -s rpc/bind setprop config/enable_tcpwrappers = false
# svcadm refresh rpc/bind
# svcadm restart rpc/bindEntry 2
```

3 **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.

4 **If you are duplicating an existing cluster than runs the Oracle Solaris Cluster 4.0 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.

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

6   Modify or create the cluster configuration XML file as needed.

Include or modify the values of the XML elements to reflect the cluster configuration that you want to create.

- 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.

- To establish a cluster, the following components must have valid values in the cluster configuration XML file:
  - Cluster name
  - Cluster nodes
  - Cluster transport

- 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.

7   Validate the cluster configuration XML file.

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

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

8 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.

9 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
```

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 Perform any necessary updates to the Oracle Solaris Cluster software.

See Chapter 11, "Updating Your Software," in *Oracle Solaris Cluster System Administration Guide* for installation instructions.

12 If you plan to enable RPC use of TCP wrappers, add all `clprivnet0` IP addresses to the `/etc/hosts.allow` file on each cluster node.

Without this addition to the /etc/hosts.allow file, TCP wrappers prevent internode communication over RPC for cluster administration utilities.

a. On each node, display the IP addresses for all `clprivnet0` devices on the node.

```
# /usr/sbin/ipadm show-addr
ADDROBJ          TYPE    STATE     ADDR
clprivnet0/N     static  ok        ip-address/netmask-length
...
```

b. On each cluster node, add to the `/etc/hosts.allow` file the IP addresses of all `clprivnet0` devices in the cluster.

13   **If you intend to use the HA for NFS data service (HA for NFS) on a highly available local file system, exclude from the automounter map all shares that are part of the highly available local file system that is exported by HA for NFS.**

See "Administrative Tasks Involving Maps" in *Oracle Solaris Administration: Network Services* for more information about modifying the automounter map.

14   **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 113.

    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 Oracle Solaris Cluster Quorum Server Software" on page 49.

    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 4.0 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 device-name
```

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

15   **Remove the cluster from installation mode.**

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

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

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

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

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 Solaris 11 OS. 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 a cluster: 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
```

More Information   Configuring Additional Components

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 describes a list of the cluster components that you can create from a cluster configuration XML file after the cluster is established. The list includes the man page for the command that you use to duplicate the component:

- 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.

- Resource Group Manager components

  - Resources: `clresource(1CL)`
  - Shared address resources: `clressharedaddress(1CL)`
  - Logical hostname resources: `clreslogicalhostname(1CL)`
  - Resource types: `clresourcetype(1CL)`
  - Resource groups: `clresourcegroup(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.

- 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)`

Troubleshooting   **Unsuccessful configuration** – If one or more nodes cannot join the cluster, or if the wrong configuration information was specified, first attempt to perform this procedure again. If that

does not correct the problem, perform the procedure "How to Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems" on page 163 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Oracle Solaris Cluster software packages. Then perform this procedure again.

Next Steps   Go to "How to Verify the Quorum Configuration and Installation Mode" on page 118.

Installing and Configuring Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)

During the scinstall Automated Installer (AI) installation of a cluster, you choose to run installation of the Oracle Solaris software in one of the following ways:

- Run a noninteractive Oracle Solaris installation which applies all default settings.

- Run an interactive Oracle Solaris installation and specify any nondefault settings that you prefer.

See "Installing With the Text Installer" in *Installing Oracle Solaris 11 Systems* for more information about interactive installation of Oracle Solaris software.

The scinstall utility runs in two modes of installation, Typical or Custom. For the Typical installation of Oracle Solaris Cluster software, scinstall automatically specifies the following configuration defaults.

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
Installation security (DES)	Limited

Complete one of the following cluster configuration worksheets to plan your Typical mode or Custom mode installation:

- **Typical Mode Worksheet** – If you will use Typical mode and accept all defaults, complete the following worksheet.

Component	Description/Example	Answer
Custom Automated Installer Boot Image ISO File	What is the full path name of the Automated Installer boot image ISO file?	

Component	Description/Example	Answer
Custom Automated Installer User root Password	What is the password for the root account of the cluster nodes?	
Custom Automated Installer Repositories	What is the repository of publisher `solaris`?	
	What is the repository of publisher `ha-cluster`?	
	Select the Oracle Solaris Cluster components that you want to install. *(Select one or more group packages to install.)*	
	Do you want to select any individual components that are contained in these group packages?	Yes | No
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.)*	
	Confirm that the auto-discovered MAC address for each node is correct.	
Cluster Transport Adapters and Cables	First node name:	
	Transport adapter names:	*First:* *Second:*
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.)*	*First:* Yes | No *Second:* 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
Custom Automated Installer Boot Image ISO File	What is the full path name of the Automated Installer boot image ISO file?	
Custom Automated Installer User root Password	What is the password for the root account of the cluster nodes?	
Custom Automated Installer Repositories	What is the repository of publisher `solaris`?	
	What is the repository of publisher `ha-cluster`?	
	Select the Oracle Solaris Cluster components that you want to install. *(Select one or more group packages to install.)*	
	Do you want to select any individual components that are contained in these group packages?	Yes | No
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.)*	
	Confirm that the auto-discovered MAC address for each node is correct.	
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

Component	Description/Example	Answer
Cluster Switches *(multiple-node cluster only)*	Transport switch name, if used: Defaults: `switch1` and `switch2`	*First:* *Second:*
Cluster Transport Adapters and Cables *(multiple-node cluster only)*	First node name:	
	Transport adapter name:	*First:* *Second:*
	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?	*First:* Yes | No *Second:* Yes | No
	If no, what is the name of the port that you want to use?	*First:* *Second:*
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`	*First:* *Second:*
	If a transport switch, do you want to use the default port name?	*First:* Yes | No *Second:* Yes | No
	If no, what is the name of the port that you want to use?	*First:* *Second:*
Global Fencing	Do you want to disable 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.*	*First:* Yes | No *Second:* 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.*)	*First:* Yes | No *Second:* Yes | No

▼ How to Install and Configure Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)

This procedure describes how to set up and use the scinstall(1M) custom Automated Installer installation method. This method installs both Oracle Solaris OS and Oracle Solaris Cluster framework and data services software on all global-cluster nodes in the same operation and establishes the 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 the Oracle Solaris Cluster software only in I/O domains or guest domains.

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.

Before You Begin Perform the following tasks:

- Ensure that the hardware setup is complete and connections are verified before you install Solaris software. See the *Oracle Solaris Cluster 4.0 Hardware Administration Manual* and your server and storage device documentation for details on how to set up the hardware.

- Ensure that an Automated Installer install server and a DHCP server are configured. See Part III, "Installing Using an Install Server," in *Installing Oracle Solaris 11 Systems*.

- Determine the Ethernet address of the cluster node and the length of the subnet mask of the subnet that the address belongs to.

- Determine the MAC address of each cluster node.

- Ensure that your cluster configuration planning is complete. See "How to Prepare for Cluster Software Installation" on page 42 for requirements and guidelines.

- Have available the root user password for the cluster nodes.

- SPARC: If you are configuring Oracle VM Server for SPARC I/O domains or guest domains as cluster nodes, ensure that the 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 52.

- Determine which Oracle Solaris Cluster software packages you want to install.

The following table lists the group packages for the Oracle Solaris Cluster 4.0 software that you can choose during an AI installation and the principal features that each group package contains. You must install at least the `ha-cluster-framework-minimal` group package.

Feature	ha-cluster-framework-full	ha-cluster-data-services-full	ha-cluster-framework-minimal
Framework	X	X	X
Agents		X	
Localization	X		
Framework man pages	X		
Data Service man pages		X	
Agent Builder	X		
Generic Data Service	X	X	

- Have available your completed Typical Mode or Custom Mode installation worksheet. See "Installing and Configuring Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 80.

1. **Set up your Automated Installer (AI) install server and DHCP server.**

   Ensure that the AI install server meets the following requirements.

   - The install server is on the same subnet as the cluster nodes.
   - The install server is not itself a cluster node.
   - The install server runs a release of the Oracle Solaris OS that is supported by the Oracle Solaris Cluster software.
   - Each new cluster node is configured as a custom AI installation client that uses the custom AI 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 AI install server and DHCP server. See Chapter 8, "Setting Up an Install Server," in *Installing Oracle Solaris 11 Systems* and Part II, "DHCP," in *Oracle Solaris Administration: IP Services*.

2. **On the AI install server, become superuser.**

3. **On the AI install server, install the Oracle Solaris Cluster AI support package.**

   a. **Ensure that the `solaris` and `ha-cluster` publishers are valid.**

   ```
   installserver# pkg publisher
   PUBLISHER        TYPE     STATUS   URI
   ```

```
solaris          origin   online   solaris-repository
ha-cluster       origin   online   ha-cluster-repository
```

b. Install the cluster AI support package.

```
installserver# pkg install ha-cluster/system/install
```

4   On the AI install server, start the `scinstall` utility.

```
installserver# /usr/cluster/bin/scinstall
```

The scinstall Main Menu is displayed.

5   Choose the Install and Configure a Cluster From This Automated Installer Install Server menu item.

```
 *** Main Menu ***

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

     * 1) Install and configure a cluster from this Automated Installer install server
     * 2) Print release information for this Automated Installer install server

     * ?) Help with menu options
     * q) Quit

   Option:  1
```

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

7   To perform any other postinstallation tasks, set up your own AI manifest.

See Chapter 13, "Running a Custom Script During First Boot," in *Installing Oracle Solaris 11 Systems*.

8   Exit from the AI install server.

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

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

  As superuser, use the following command to start the pconsole utility:

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

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

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

10 Ensure that TCP wrappers for RPC are disabled on all nodes of the cluster.

The Oracle Solaris TCP wrappers for RPC feature prevents internode communication that is necessary for cluster configuration.

a. On each node, display the status of TCP wrappers for RPC.

TCP wrappers are enabled if config/enable_tcpwrappers is set to true, as shown in the following example command output.

```
# svccfg -s rpc/bind listprop config/enable_tcpwrappers
config/enable_tcpwrappers  boolean true
```

b. If TCP wrappers for RPC are enabled on a node, disable TCP wrappers and refresh the RPC bind service.

```
# svccfg -s rpc/bind setprop config/enable_tcpwrappers = false
# svcadm refresh rpc/bind
# svcadm restart rpc/bindEntry 2
```

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

The Oracle Solaris software is installed with the default configuration.

Note – You cannot use this method if you want to customize the Oracle Solaris installation. If you choose the Oracle Solaris interactive installation, the Automated Installer is bypassed and Oracle Solaris Cluster software is not installed and configured. To customize Oracle Solaris during installation, instead follow instructions in "How to Install Oracle Solaris Software" on page 43, then install and configure the cluster by following instructions in "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53.

- **SPARC:**

  a. Shut down each node.

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

  b. Boot the node with the following command

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

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

- **x86:**

  a. Reboot the node.

  ```
  # reboot -p
  ```

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

The GRUB menu is displayed with two menu entries similar to the following:

```
Oracle Solaris 11 11/11 Text Installer and command line
Oracle Solaris 11 11/11 Automated Install
```

c. **Immediately select the Automated Install entry and press Return.**

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

On each node, a new boot environment (BE) is created and Automated Installer installs the Oracle Solaris OS and Oracle Solaris Cluster software. 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 on each node.

12 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
```

13 On each node, activate the installed BE and boot into cluster mode.

a. **Activate the installed BE.**

```
# beadm activate BE-name
```

b. **Shut down the node.**

```
# shutdown -y -g0 -i0
```

Note – Do *not* use the reboot or halt command. These commands do not activate a new BE.

---

c. **Boot the node into cluster mode.**

- **SPARC:**

```
ok boot
```

- **x86:**

When the GRUB menu is displayed, select the appropriate Oracle Solaris entry and press Enter.

For more information about GRUB based booting, see *Booting and Shutting Down Oracle Solaris on x86 Platforms*.

14   If you intend to use the HA for NFS data service (HA for NFS) on a highly available local file system, exclude from the automounter map all shares that are part of the highly available local file system that is exported by HA for NFS.

See "Administrative Tasks Involving Maps" in *Oracle Solaris Administration: Network Services* for more information about modifying the automounter map.

15   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/kernel/amd64/unix -B $ZFS-BOOTFS -k
```

For more information, see "How to Boot a System With the Kernel Debugger Enabled (kmdb)" in *Booting and Shutting Down Oracle Solaris on x86 Platforms*.

16   If you performed a task that requires a cluster reboot, reboot the cluster.

The following tasks require a reboot:

- Installing software updates 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 cluster-name
```

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 113.

---

c.   Reboot each node in the cluster.

- **SPARC:**

  ```
  ok boot
  ```

- **x86:**

     When the GRUB menu is displayed, select the appropriate Oracle Solaris entry and press Enter.

     For more information about GRUB based booting, see *Booting and Shutting Down Oracle Solaris on x86 Platforms*.

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.

17 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.

18 If you plan to enable RPC use of TCP wrappers, add all `clprivnet0` IP addresses to the `/etc/hosts.allow` file on each cluster node.

Without this addition to the /etc/hosts.allow file, TCP wrappers prevent internode communication over RPC for cluster administration utilities.

a. On each node, display the IP addresses for all `clprivnet0` devices on the node.

```
# /usr/sbin/ipadm show-addr
ADDROBJ           TYPE     STATE       ADDR
clprivnet0/N      static   ok          ip-address/netmask-length
...
```

b. On each cluster node, add to the `/etc/hosts.allow` file the IP addresses of all `clprivnet0` devices in the cluster.

19 (Optional) On each node, enable automatic node reboot if all monitored shared-disk paths fail.

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 1. **Perform all of the following procedures that are appropriate for your cluster configuration.**

- "How to Configure Internal Disk Mirroring" on page 51
- "SPARC: How to Install Oracle VM Server for SPARC Software and Create Domains" on page 52
- "How to Set Up the Root Environment" on page 58
- "How to Configure Solaris IP Filter" on page 59

2. **Configure quorum, if not already configured, and perform postinstallation tasks.**

- If you installed a multiple-node cluster and accepted automatic quorum configuration, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 118.

- If you installed a multiple-node cluster and declined automatic quorum configuration, perform postinstallation setup. Go to "How to Configure Quorum Devices" on page 113.

- If you added a node to an existing two-node cluster, go to "How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 111.

- If you added a new node to an existing cluster with at least three nodes that uses a quorum device, go to "How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 111.

- If you added a new node to an existing cluster with at least three nodes 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 118.

- If you installed a single-node cluster, cluster establishment is complete. Go to "Creating Cluster File Systems" on page 143 to install volume management software and configure the cluster.

Troubleshooting **Disabled scinstall option** – If the AI option of the scinstall command is not preceded by an asterisk, the option is disabled. This condition indicates that AI 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 7 to correct the AI 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 4.0 Hardware Administration Manual*.
  - Verify that any existing cluster interconnects can support the new node. See the *Oracle Solaris Cluster 4.0 Hardware Administration Manual*.
  - Ensure that any additional storage is installed.

1   **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 New Nodes menu item.**

   d.   **Choose the Specify the Name of a Machine Which May Add Itself menu item.**

   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.**

2   **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 3.**

- 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 Cluster Interconnect menu item.**

    c. **Choose the Add a Transport Cable menu item.**

    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.

3 **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 101.

- **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 94. 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 101 or "How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (XML File)" on page 106.

▼ 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.

---

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.

   a.   Choose the Resource Groups menu item.

   The Resource Group Menu is displayed.

   b.   Choose the Online/Offline or Switchover a Resource Group menu item.

   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.   Choose the Enable/Disable a Resource menu item.

   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.

   ■   **SPARC:**

   ok **boot -x**

- **x86:**

  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 and Shutting Down Oracle Solaris on x86 Platforms*.

  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 **Choose the Change Network Addressing and Ranges for the Cluster Transport menu item.**

   The clsetup utility displays the current private network configuration, then asks if you would like to change this configuration.

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.**

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

  a. **Type `no` in response to the `clsetup` utility question about whether it is okay to accept the default address, then press the Return key.**

  The `clsetup` utility will prompt for the new private-network IP address.

  b. **Type the new IP address and press the Return key.**

  The `clsetup` utility displays the default netmask and then asks if it is okay to accept the default netmask.

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.**

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

  a. **Type `no` in response to the `clsetup` utility's question about whether it is okay to accept the default address range, then press the Return key.**

  When you decline the default netmask, the `clsetup` utility prompts you for the number of nodes and private networks, and zone clusters that you expect to configure in the cluster.

  b. **Provide the number of nodes, private networks, and zone clusters that you expect to configure in the cluster.**

  From these numbers, the `clsetup` utility calculates two proposed netmasks:

  - The first netmask is the minimum netmask to support the number of nodes, private networks, and zone clusters that you specified.
  - The second netmask supports twice the number of nodes, private networks, and zone clusters that you specified, to accommodate possible future growth.

  c. **Specify either of the calculated netmasks, or specify a different netmask that supports the expected number of nodes, private networks, and zone clusters.**

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.

  ■ **SPARC:**

  ok **boot**

  ■ **x86:**

  When the GRUB menu is displayed, select the appropriate Oracle Solaris entry and press
  Enter.

  For more information about GRUB based booting, see *Booting and Shutting Down
  Oracle Solaris on x86 Platforms*.

17   From one node, start the clsetup utility.

```
# clsetup
```

The clsetup Main Menu is displayed.

18   Re-enable all disabled resources.

a. Choose the Resource Groups menu item.

The Resource Group Menu is displayed.

b. Choose the Enable/Disable a Resource menu item.

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. Choose the Online/Offline or Switchover a Resource Group menu item.

**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 Control-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 101
- "How to Install and Configure Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 84
- "How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (XML File)" on page 106

Configuring Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (scinstall)

The `scinstall` utility runs in two modes of installation, Typical or Custom. For the Typical installation of Oracle Solaris Cluster software, `scinstall` automatically specifies the cluster transport switches as `switch1` and `switch2`.

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 16 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:*

Component	Description/Example	Answer
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`	*First:* *Second:*
	For transport switches, do you want to use the default port name?	*First:* Yes | No *Second:* Yes | No
	If no, what is the name of the port that you want to use?	*First:* *Second:*
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`	*First:* *Second:*

Component	Description/Example	Answer
	If a transport switch, do you want to use the default port name?	*First:* Yes | No
		Second: Yes | No
	If no, what is the name of the port that you want to use?	*First:*
		Second:
Automatic Reboot	Do you want scinstall to automatically reboot the node after installation?	Yes | No

▼ 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 Automated Installer to add a new node, follow the instructions in "How to Install and Configure Oracle Solaris and Oracle Solaris Cluster Software (Automated Installer)" on page 84.

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

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.

Before You Begin Perform the following tasks:

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

  If the Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for the 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 43 for more information about installing the Oracle Solaris software to meet Oracle Solaris Cluster software requirements.

- Ensure that NWAM is disabled. See "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53 for instructions.

- SPARC: If you are configuring Oracle VM Server for SPARC I/O domains or guest domains as cluster nodes, ensure that the 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 52.

- Ensure that Oracle Solaris Cluster software packages and updates are installed on the node. See "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53.

- 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 92.

- Have available your completed Typical Mode or Custom Mode installation worksheet. See "Configuring Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (scinstall)" on page 99.

1   On the cluster node to configure, become superuser.

2   Ensure that TCP wrappers for RPC are disabled on all nodes of the cluster.

The Oracle Solaris TCP wrappers for RPC feature prevents internode communication that is necessary for cluster configuration.

a.   On each node, display the status of TCP wrappers for RPC.

TCP wrappers are enabled if config/enable_tcpwrappers is set to true, as shown in the following example command output.

```
# svccfg -s rpc/bind listprop config/enable_tcpwrappers
config/enable_tcpwrappers  boolean true
```

b.   If TCP wrappers for RPC are enabled on a node, disable TCP wrappers and refresh the RPC bind service.

```
# svccfg -s rpc/bind setprop config/enable_tcpwrappers = false
# svcadm refresh rpc/bind
# svcadm restart rpc/bindEntry 2
```

3   Prepare public-network interfaces.

a.   Create static IP addresses for each public-network interface.

```
# ipadm create-ip interface
# ipadm create-addr -T static -a local=address/prefix-length addrobj
```

For more information, see "How to Configure an IP Interface" in *Oracle Solaris Administration: Network Interfaces and Network Virtualization*.

b. (Optional) Create IPMP groups for public-network interfaces.

During initial cluster configuration, unless non-link-local IPv6 public network interfaces exist in the cluster, IPMP groups are automatically created based on matching subnets. These groups use transitive probes for interface monitoring and no test addresses are required.

If these automatically created IPMP groups would not meet your needs, or if IPMP groups would not be created because your configuration includes one or more non-link-local IPv6 public network interfaces, do one of the following:

- **Create the IPMP groups you need before you establish the cluster.**

- **After the cluster is established, use the `ipadm` command to edit the IPMP groups.**

For more information, see "Configuring IPMP Groups" in *Oracle Solaris Administration: Network Interfaces and Network Virtualization*.

4   **Start the `scinstall` utility.**

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

    The scinstall Main Menu is displayed.

5   **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) Print release information for this cluster node

     * ?) Help with menu options
     * q) Quit

    Option:  1
```

    The New Cluster and Cluster Node Menu is displayed.

6   **Type the option number for Add This Machine as a Node in an Existing Cluster and press the Return key.**

7   **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.

8   **Repeat this procedure on any other node to add to the cluster until all additional nodes are fully configured.**

9 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
```

10 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.

11 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.

12 If TCP wrappers are used in the cluster, ensure that the `clprivnet0` IP addresses for all added nodes are added to the `/etc/hosts.allow` file on each cluster node.

Without this addition to the /etc/hosts.allow file, TCP wrappers prevent internode communication over RPC for cluster administration utilities.

a. On each node, display the IP addresses for all `clprivnet0` devices.

```
# /usr/sbin/ipadm show-addr
ADDROBJ         TYPE    STATE     ADDR
clprivnet0/N    static  ok                ip-address/netmask-length
...
```

b. On each node, edit the `/etc/hosts.allow` file with the IP addresses of all `clprivnet0` devices in the cluster.

13 Verify that all necessary software updates are installed.

```
phys-schost# pkg list
```

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

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
...
```

15   If you intend to use the HA for NFS data service (HA for NFS) on a highly available local file system, exclude from the automounter map all shares that are part of the highly available local file system that is exported by HA for NFS.

See "Administrative Tasks Involving Maps" in *Oracle Solaris Administration: Network Services* for more information about modifying the automounter map.

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 node "phys-schost-3" to the cluster configuration ... done
Adding adapter "net2" to the cluster configuration ... done
Adding adapter "net3" 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
Setting the node ID for "phys-schost-3" ... done (id=1)

Verifying the major number for the "did" driver from "phys-schost-1" ... done
Initializing NTP configuration ... done

Updating nsswitch.conf ... done

Adding cluster node 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

Ensure network routing is disabled ... done
Network routing has been disabled on this node by creating /etc/notrouter.
Having a cluster node act as a router is not supported by Oracle Solaris Cluster.
Please do not re-enable network routing.
Updating file ("ntp.conf.cluster") on node phys-schost-1 ... done
Updating file ("hosts") on node phys-schost-1 ... done

Log file - /var/cluster/logs/install/scinstall.log.6952

Rebooting ...
```

Troubleshooting   **Unsuccessful configuration** – If one or more nodes cannot join the cluster, or if the wrong configuration information was specified, first attempt to perform this procedure again. If that does not correct the problem, perform the procedure "How to Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems" on page 163 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Oracle Solaris Cluster software packages. Then perform this procedure again.

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 111.

Otherwise, go to "How to Verify the Quorum Configuration and Installation Mode" on page 118.

▼ How to Configure Oracle Solaris Cluster Software on Additional Global-Cluster Nodes (XML File)

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 the Oracle Solaris Cluster 4.0 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 the Oracle Solaris Cluster software.

If the Oracle Solaris software is already installed on the node, you must ensure that the Oracle Solaris installation meets the requirements for the 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 43 for more information about installing the Oracle Solaris software to meet Oracle Solaris Cluster software requirements.

- Ensure that NWAM is disabled. See "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53 for instructions.

- SPARC: If you are configuring Oracle VM Server for SPARC I/O domains or guest domains as cluster nodes, ensure that the 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 52.

- Ensure that Oracle Solaris Cluster software packages and any necessary updates are installed on the node. See "How to Install Oracle Solaris Cluster Framework and Data Service Software Packages" on page 53.

- 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 92.

1   **Ensure that the 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 the 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.**

         The 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, the 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.**

       - **SPARC:**

         `ok boot -x`

- **x86:**

    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 and Shutting Down Oracle Solaris on x86 Platforms*.

    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 the 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 the Oracle Solaris Cluster 4.0 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   Ensure that TCP wrappers for RPC are disabled on all nodes of the cluster.

The Oracle Solaris TCP wrappers for RPC feature prevents internode communication that is necessary for cluster configuration.

a.   On each node, display the status of TCP wrappers for RPC.

TCP wrappers are enabled if config/enable_tcpwrappers is set to true, as shown in the following example command output.

```
# svccfg -s rpc/bind listprop config/enable_tcpwrappers
config/enable_tcpwrappers  boolean true
```

b.   If TCP wrappers for RPC are enabled on a node, disable TCP wrappers and refresh the RPC bind service.

```
# svccfg -s rpc/bind setprop config/enable_tcpwrappers = false
# svcadm refresh rpc/bind
# svcadm restart rpc/bindEntry 2
```

5   Modify or create the cluster configuration XML file as needed.

- 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.

- 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.

6   Validate the cluster configuration XML file.

phys-schost-new# **xmllint --valid --noout** *clconfigfile*

7   Configure the new cluster node.

phys-schost-new# **clnode add -n** *sponsor-node* **-i** *clconfigfile*

-n *sponsor-node*
  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.

8   If TCP wrappers are used in the cluster, ensure that the `clprivnet0` IP addresses for all added nodes are added to the `/etc/hosts.allow` file on each cluster node.

Without this addition to the /etc/hosts.allow file, TCP wrappers prevent internode communication over RPC for cluster administration utilities.

a.   On each node, display the IP addresses for all `clprivnet0` devices.

```
# /usr/sbin/ipadm show-addr
ADDROBJ          TYPE     STATE     ADDR
clprivnet0/N     static   ok        ip-address/netmask-length
...
```

b.   On each node, edit the `/etc/hosts.allow` file with the IP addresses of all `clprivnet0` devices in the cluster.

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

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 perform this procedure again. If that does not correct the problem, perform the procedure "How to Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems" on page 163 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Oracle Solaris Cluster software packages. Then perform this procedure again.

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 111.

Otherwise, go to "How to Verify the Quorum Configuration and Installation Mode" on page 118.

▼ 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 update 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 the 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 device-name
```

device-name
    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 and choose the shared device from the output.

```
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.   Configure the shared device as a quorum device.

```
phys-schost# clquorum add -t type device-name
```

-t *type*
   Specifies the type of quorum device. If this option is not specified, the default type `shared_disk` is used.

c.   Repeat for each quorum device that you want to configure.

d.   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 118.

▼ 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.

If you chose automatic quorum configuration when you established the cluster, do not perform this procedure. Instead, proceed to "How to Verify the Quorum Configuration and Installation Mode" on page 118.

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   ■   **Quorum servers** – To configure a quorum server as a quorum device, do the following:

- Install the Oracle Solaris Cluster 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 Oracle Solaris Cluster Quorum Server Software" on page 49.

- 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, do the following:

  - Install the NAS device hardware and software. See *Oracle Solaris Cluster 4.0 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, ensure that the correct prefix length is set for the public-network addresses.

- You intend to use a quorum server.

- The public network uses variable-length subnet masking, also called classless inter domain routing (CIDR).

```
# ipadm show-addr
   ADDROBJ         TYPE     STATE       ADDR
   lo0/v4          static   ok          127.0.0.1/8
   ipmp0/v4        static   ok          10.134.94.58/24
```

Note – 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.

2 On one node, become superuser.

Alternatively, if your user account is assigned the System Administrator profile, issue commands as nonroot through a profile shell, or prefix the command with the pfexec command.

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 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 29 for further information about choosing quorum devices.

    Use the cldevice output from Step a to identify the device ID 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, the initial cluster setup was already successfully performed. Skip to Step 11.

---

7   Indicate whether you want to add any quorum disks.

- **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.**

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 ZFS Storage Appliance
quorum_server	Quorum server

9   Specify the name of the device to configure as a quorum device and provide any required additional information.

- 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   Type Yes to verify that it is okay to reset installmode.

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 118.

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 adding 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 the quorum configuration was completed successfully 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 120.

- If you want to install or modify the NTP configuration file, go to "Configuring Network Time Protocol (NTP)" on page 120.

- If you want to install a volume manager, go to Chapter 4, "Configuring Solaris Volume Manager Software," to install volume management software.

- If you want to create cluster file systems, go to "How to Create Cluster File Systems" on page 143.

- To find out how to 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*.

- When your cluster is fully configured, validate the configuration. Go to "How to Validate the Cluster" on page 123.

- 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 127.

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 Node 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 Update the NTP configuration with the changed private hostnames. Go to .

Configuring Network Time Protocol (NTP)

This section contains the following procedures:

-
-

■ "How to Update NTP After Changing a Private Hostname" on page 123

▼ How to Use Your Own /etc/inet/ntp.conf File

Note – If you installed your own /etc/inet/ntp.conf file **before** you installed the Oracle Solaris Cluster software, you do not need to perform this procedure. Proceed to "How to Validate the Cluster" on page 123.

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

2 **Add your /etc/inet/ntp.conf file to each node of the cluster.**

3 **On each node, determine the state of the NTP service.**

    phys-schost# **svcs svc:/network/ntp:default**

4 **Start the NTP daemon on each node.**

    ■ **If the NTP service is disabled, enable the service.**

        phys-schost# **svcadm enable svc:/network/ntp:default**

    ■ **If the NTP service is online, restart the service.**

        phys-schost# **svcadm restart svc:/network/ntp:default**

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 install a volume manager, go to Chapter 4, "Configuring Solaris Volume Manager Software."

■ If you want to create cluster file systems, go to "How to Create Cluster File Systems" on page 143.

■ To find out how to 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*.

■ When your cluster is fully configured, validate the configuration. Go to "How to Validate the Cluster" on page 123.

■ 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 127.

▼ How to Install NTP After Adding a Node to a Single-Node Cluster

When you add 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 **Copy the `/etc/inet/ntp.conf` and `/etc/inet/ntp.conf.sc` files from the added node to the original cluster node.**

These files were created on the added node when it was configured with the cluster.

3 **On the original cluster node, create a symbolic link named `/etc/inet/ntp.conf.include` that points to the `/etc/inet/ntp.conf.sc` file.**

```
phys-schost# ln -s /etc/inet/ntp.conf.sc /etc/inet/ntp.conf.include
```

4 **On each node, determine the state of the NTP service.**

```
phys-schost# svcs svc:/network/ntp:default
```

5 **Start the NTP daemon on each node.**

- **If the NTP service is `disabled`, enable the service.**

```
phys-schost# svcadm enable svc:/network/ntp:default
```

- **If the NTP service is `online`, restart the service.**

```
phys-schost# svcadm restart svc:/network/ntp:default
```

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 install a volume manager, go to Chapter 4, "Configuring Solaris Volume Manager Software."

- If you want to create cluster file systems, go to "How to Create Cluster File Systems" on page 143.

- To find out how to 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*.

- When your cluster is fully configured, validate the configuration. Go to "How to Validate the Cluster" on page 123.

- 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 127.

▼ How to Update NTP After Changing a Private Hostname

1   Become superuser on a cluster node.

2   On each node of the cluster, update the `/etc/inet/ntp.conf.sc` file with the changed private hostname.

3   On each node, determine the state of the NTP service.

```
phys-schost# svcs svc:/network/ntp:default
```

4   Start the NTP daemon on each node.

- **If the NTP service is `disabled`, enable the service.**

  ```
  phys-schost# svcadm enable svc:/network/ntp:default
  ```

- **If the NTP service is `online`, restart the service.**

  ```
  phys-schost# svcadm restart svc:/network/ntp:default
  ```

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 install a volume manager, go to Chapter 4, "Configuring Solaris Volume Manager Software."
- If you want to create cluster file systems, go to "How to Create Cluster File Systems" on page 143.
- To find out how to 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*.
- When your cluster is fully configured, validate the configuration. Go to "How to Validate the Cluster" on page 123.
- 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 127.

▼ 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 software updates.

1   Become superuser on a node of the cluster.

2   Ensure that you have the most current checks.

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

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

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

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

3   Run basic validation checks.

```
# 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 *check-ID*

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

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

d. **Start the functional check.**

cluster check -v -k functional -C *check-ID* **-o** *outputdir*

-k functional   Specifies running functional validation checks

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

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

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

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 127.

▼ 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 you need to troubleshoot the cluster in the future.

1   **Become superuser.**

2   **Install the 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 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 to the Oracle Explorer database 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.

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 37. See your Solaris Volume Manager documentation for additional details.

This chapter contains the following sections:

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 37
Create state database replicas on the local disks.	"How to Create State Database Replicas" on page 129

▼ 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 your Solaris Volume Manager documentation 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   Go to "Creating Disk Sets in a Cluster" on page 130 to create Solaris Volume Manager disk sets.

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: Configuring Solaris Volume Manager Disk Sets

Task	Instructions
Create disk sets by using the metaset command.	"How to Create a Disk Set" on page 131

TABLE 4–2   Task Map: Configuring Solaris Volume Manager Disk Sets   *(Continued)*

Task	Instructions
Add drives to the disk sets.	"How to Add Drives to a Disk Set" on page 133
(*Optional*) Repartition drives in a disk set to allocate space to different slices.	"How to Repartition Drives in a Disk Set" on page 135
List DID pseudo-driver mappings and define volumes in the /etc/lvm/md.tab files.	"How to Create an md.tab File" on page 135
Initialize the md.tab files.	"How to Activate Volumes" on page 137

▼ How to Create a Disk Set

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 138 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 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 129.

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 device-group
```

-p
   Specifies a device-group property.

name
   Specifies the name of a property.

value
   Specifies the value or setting of the property.

device-group
> 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 4–2   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 133.

Adding Drives to a Disk Set

When you add a drive to a disk set, the volume management software repartitions the drive 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 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 131.

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/d***N*

-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 4–3 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 If you want to repartition drives for use in volumes, go to "How to Repartition Drives in a Disk Set" on page 135.

Otherwise, go to "How to Create an md.tab File" on page 135 to find out how 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 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 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, take steps to prevent the metaset command from repartitioning the drive.

a. Create slice 6 for EFI starting at cylinder 0, large enough to hold a state database replica.

Do not allow the target slice to overlap any other slice on the drive.

See your Solaris Volume Manager administration guide to determine the size of a state database replica for your version of the volume-manager software.

b. Set the `Flag` field in the target slice to `wu` (read-write, unmountable).

Do not set this field to read-only.

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 135.

▼ 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 IDs that are used to form disk sets. For example, if the device ID /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 that contains the volume definitions for the disk sets you created.

See Example 4–4 for a sample md.tab file.

> **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. 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 4–4   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 137.

▼ 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 device-group*

-n *node*
  Specifies the node that takes ownership.

device-group
  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 4–5   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 138.

Otherwise, go to "How to Create Cluster File Systems" on page 143 to find out how 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: Configuring Solaris Volume Manager Dual-String Mediators

Task	Instructions
Configure dual-string mediator hosts.	"Requirements for Dual-String Mediators" on page 139 "How to Add Mediator Hosts" on page 139
Check the status of mediator data and, if necessary, fix bad mediator data.	"How to Check For and Fix Bad Mediator Data" on page 141

Requirements for Dual-String Mediators

The following rules apply to dual-string configurations that use mediators.

- Disk sets must be configured with either 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 can 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.

1 If you will use a third mediator host for a dual-string disk set and that host does not already have disk sets configured, modify the `/etc/group` file and create a dummy disk set.

   a. Add the entry `root` to the `sysadmin` group in the `/etc/group` file.

   b. Create a dummy disk set by using the `metaset` command.

      phys-schost-3# **metaset -s** *dummy-diskset-name* **-a -h** *hostname*

     -s *dummy-diskset-net*
       Specifies the dummy disk set name.

     -a
       Adds to the disk set.

     -h *hostname*
       Specifies the name of the node.

2 Become superuser on the node that currently masters the disk set to which you intend to add mediator hosts.

3 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.

  -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 4–6 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 For and Fix Bad Mediator Data" on page 141.

▼ How to Check For and Fix Bad Mediator Data

Perform this procedure to repair bad mediator data.

Before You Begin Ensure that you have added mediator hosts as described in "How to Add Mediator Hosts" on page 139.

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 **Check the Status field of the medstat output for each mediator host.**

- If the status is **Ok**, no error condition exists.

- If the status is **Bad**, repair the affected mediator host.

3 **Become superuser on the node that owns the affected disk set.**

4 **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.

5 **Restore each mediator host that you removed in Step 4.**

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 want to create cluster file systems, go to "How to Create Cluster File Systems" on page 143.

- To find out how to 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*.

5

Creating a Cluster File System

This chapter describes how to create a cluster file system to support data services.

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*.

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 41.

- 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 62.

- If you are using a volume manager, ensure that volume management software is installed and configured. For volume manager installation procedures, see "Configuring Solaris Volume Manager Software" on page 129.

- 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 35.

1 Become superuser on any node in the cluster.

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

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

⚠ **Caution** – Any data on the disks is destroyed when you create a file system. Be sure that you specify the correct disk device name. If you specify the wrong device name, you might erase data that you did not intend to delete.

```
phys-schost# newfs raw-disk-device
```

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

Volume Manager	Sample Disk Device Name	Description
Solaris Volume Manager	`/dev/md/nfs/rdsk/d1`	Raw disk device `d1` within the `nfs` disk set
None	`/dev/global/rdsk/d1s3`	Raw disk device `d1s3`

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

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

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

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

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

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

4 On each node in the cluster, add an entry to the `/etc/vfstab` file for the mount point.

See the `vfstab`(4) man page for details.

a. In each entry, specify the required mount options for the type of file system that you use.

b. **To automatically mount the cluster file system, set the `mount at boot` field to yes.**

c. **For each cluster file system, ensure that the information in its `/etc/vfstab` entry is identical on each node.**

d. **Ensure that the entries in each node's `/etc/vfstab` file list devices in the same order.**

e. **Check the boot order dependencies of the file systems.**

   For example, consider the scenario where phys-schost-1 mounts disk device d0 on /global/oracle/ and phys-schost-2 mounts disk device d1 on /global/oracle/logs/. With this configuration, phys-schost-2 can boot and mount /global/oracle/logs/ only after phys-schost-1 boots and mounts /global/oracle/.

5 **On any node in the cluster, run the configuration check utility.**

   phys-schost# **cluster check -k vfstab**

   The configuration check utility verifies that the mount points exist. The utility also verifies that /etc/vfstab file entries are correct on all nodes of the cluster. If no errors occur, no output is returned.

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

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

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

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

   You can use either the df command or mount command to list mounted file systems. For more information, see the df(1M) man page or mount(1M) man page.

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 completed 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 ■ To find out how to 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*.

◆ ◆ ◆ **C H A P T E R   6**

6

Creating Zone Clusters

This chapter describes the following topic:

- "Configuring a Zone Cluster" on page 147

Configuring a Zone Cluster

This section provides procedures to configure a cluster of Oracle Solaris non-global zones, called a zone cluster. It describes the following topics:

- "Overview of the clzonecluster Utility" on page 147
- "Establishing the Zone Cluster" on page 148
- "Adding File Systems to a Zone Cluster" on page 153
- "Adding Storage Devices to a Zone Cluster" on page 157

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:zone-cluster-name:resource>          cluster-wide setting
clzc:zone-cluster-name:node:resource>     node-specific setting
```

You can use the clzonecluster utility to specify any Oracle Solaris zones resource parameter as well as the parameters that are specific to zone clusters. 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 Create a Zone Cluster

Perform this procedure to create a cluster of non-global zones.

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 30.
- Have available the following information:
  - The unique name to assign to the zone cluster.
  - The zone path that the nodes of the zone cluster will use. For more information, see the description of the zonepath property in "Resource Types and Properties" in *Oracle Solaris Administration: Oracle Solaris Zones, Oracle Solaris 10 Zones, and Resource Management*.
  - 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 IP address 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.

You 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   Create the zone cluster.

Observe the following special instructions:

- By default, whole root zones are created. To create whole root zones, add the -b option to the create command.

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

```
phys-schost-1# clzonecluster configure zone-cluster-name
clzc:zone-cluster-name> create

    Set the zone path for the entire zone cluster
clzc:zone-cluster-name> set zonepath=/zones/zone-cluster-name

    Add the first node and specify node-specific settings
clzc:zone-cluster-name> add node
clzc:zone-cluster-name:node> set physical-host=base-cluster-node1
clzc:zone-cluster-name:node> set hostname=hostname1
clzc:zone-cluster-name:node> add net
clzc:zone-cluster-name:node:net> set address=public-netaddr
clzc:zone-cluster-name:node:net> set physical=adapter
clzc:zone-cluster-name:node:net> end
clzc:zone-cluster-name:node> end

    Add authorization for the public-network addresses that the zone cluster is allowed to use
clzc: zone-cluster-name> add net
clzc: zone-cluster-name:net> set address=IP-address1
clzc: zone-cluster-name:net> end

    Save the configuration and exit the utility
clzc:zone-cluster-name> commit
clzc:zone-cluster-name> exit
```

4 (Optional) Add one or more additional nodes to the zone cluster.

```
phys-schost-1# clzonecluster configure zone-cluster-name
clzc:zone-cluster-name> add node
clzc:zone-cluster-name:node> set physical-host=base-cluster-node2
clzc:zone-cluster-name:node> set hostname=hostname2
clzc:zone-cluster-name:node> add net
clzc:zone-cluster-name:node:net> set address=public-netaddr
clzc:zone-cluster-name:node:net> set physical=adapter
clzc:zone-cluster-name:node:net> end
clzc:zone-cluster-name:node> end
clzc:zone-cluster-name> commit
clzc:zone-cluster-name> exit
```

5 Verify the zone cluster configuration.

The verify subcommand checks for the availability of the specified resources. If the clzonecluster verify command succeeds, no output is displayed.

```
phys-schost-1# clzonecluster verify zone-cluster-name
phys-schost-1# clzonecluster status zone-cluster-name
=== Zone Clusters ===

--- Zone Cluster Status ---

Name     Node Name    Zone HostName   Status    Zone Status
----     ---------    -------------   ------    -----------
zone     basenode1    zone-1          Offline   Configured
         basenode2    zone-2          Offline   Configured
```

6 Install the zone cluster.

```
phys-schost-1# clzonecluster install options zone-cluster-name
Waiting for zone install commands to complete on all the nodes
of the zone cluster "zone-cluster-name"...
```

If needed, include the following options in the clzonecluster install command.

- To include system configuration information, add the following option:

  -c config-profile.xml

  The -c config-profile.xml option provides 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.

- If the base global-cluster nodes for the zone-cluster are not all installed with the same Oracle Solaris Cluster packages but you do not want to change which packages are on the base nodes, add the following option:

  -M manifest.xml

  The -M manifest.xml option specifies a custom Automated Installer manifest that you configure to install the necessary packages on all zone-cluster nodes. If the clzonecluster install command is run without the -M option, zone-cluster installation fails on a base node if it is missing a package that is installed on the issuing base node.

7 Boot the zone cluster.

Installation of the zone cluster might take several minutes

```
phys-schost-1# clzonecluster boot zone-cluster-name
Waiting for zone boot commands to complete on all the nodes of
the zone cluster "zone-cluster-name"...
```

8 If you did not set the profile during zone cluster installation, configure the profile manually.

On each zone-cluster node, issue the following command and progress through the interactive screens.

```
phys-schost-1# zlogin -C zone-cluster-name
```

9 After all zone-cluster nodes are modified, reboot the global-cluster nodes to initialize the changes to the zone-cluster /etc/inet/hosts files.

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

10 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 6–1 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 the 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 net0 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 net1 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=net0
```

```
end
end
add node
set physical-host=phys-schost-2
set hostname=zc-host-2
add net
set address=172.16.0.2
set physical=net1
end
end
commit
exit
```

Example 6–2 Creating a Zone Cluster by Using a Configuration File

The following example shows the commands to create the new zone cluster sczone on the global-cluster node phys-schost-1 by using the configuration file sczone-config. The hostnames of the zone-cluster nodes are zc-host-1 and zc-host-2.

```
phys-schost-1# clzonecluster configure -f sczone-config sczone
phys-schost-1# clzonecluster verify sczone
phys-schost-1# clzonecluster install sczone
Waiting for zone install commands to complete on all the nodes of the
zone cluster "sczone"...
phys-schost-1# clzonecluster boot sczone
Waiting for zone boot commands to complete on all the nodes of the
zone cluster "sczone"...
phys-schost-1# clzonecluster status sczone
=== Zone Clusters ===

--- Zone Cluster Status ---

Name     Node Name      Zone HostName   Status    Zone Status
----     ---------      -------------   ------    -----------
sczone   phys-schost-1  zc-host-1       Offline   Running
         phys-schost-2  zc-host-2       Offline   Running
```

Next Steps If you want to add the use of a file system to the zone cluster, go to "Adding File Systems to a Zone Cluster" on page 153.

If you want to add the use of global storage devices to the zone cluster, go to "Adding Storage Devices to a Zone Cluster" on page 157.

See Also If you want to update a zone cluster, follow procedures in Chapter 11, "Updating Your Software," in *Oracle Solaris Cluster System Administration Guide*. These procedures include special instructions for zone clusters, where needed.

Adding File Systems to a 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`.

This section provides the following procedures to add file systems for use by the zone cluster:

- "How to Add a Local File System to a Zone Cluster" on page 153
- "How to Add a ZFS Storage Pool to a Zone Cluster" on page 155
- "How to Add a Cluster File System to a Zone Cluster" on page 156

In addition, if you want 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 File System Highly Available" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

▼ How to Add a Local File System to a Zone Cluster

Perform this procedure to add a local file system on the global cluster for use by the zone cluster.

Note – To add a ZFS pool to a zone cluster, instead perform procedures in "How to Add a ZFS Storage Pool to a Zone Cluster" on page 155.

Alternatively, 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 File System Highly Available" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

---

1 Become superuser on a node of the global cluster that hosts the zone cluster.

You perform all steps of the procedure from a node of the global cluster.

2 On the global cluster, create a file system that you want to use in the zone cluster.

Ensure that the file system is created on shared disks.

3 Add the file system to the zone-cluster configuration.

```
phys-schost# clzonecluster configure zone-cluster-name
clzc:zone-cluster-name> add fs
clzc:zone-cluster-name:fs> set dir=mount-point
clzc:zone-cluster-name:fs> set special=disk-device-name
clzc:zone-cluster-name:fs> set raw=raw-disk-device-name
clzc:zone-cluster-name:fs> set type=FS-type
clzc:zone-cluster-name:fs> end
clzc:zone-cluster-name> verify
clzc:zone-cluster-name> commit
clzc:zone-cluster-name> exit
```

dir=*mount-point*
  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 zone-cluster-name
```

Example 6–3 Adding a 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

Note – 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 File System Highly Available" in *Oracle Solaris Cluster Data Services Planning and Administration Guide*.

1 Become superuser on a node of the global cluster that hosts the zone cluster.

You perform all steps of this procedure from a node of the global zone.

2 Create the ZFS storage pool on the global cluster.

Ensure that the pool is connected on shared disks that are connected to all nodes of the zone cluster.

See *Oracle Solaris Administration: ZFS File Systems* for procedures to create a ZFS pool.

3 Add the pool to the zone-cluster configuration.

```
phys-schost# clzonecluster configure zone-cluster-name
clzc:zone-cluster-name> add dataset
clzc:zone-cluster-name:dataset> set name=ZFSpoolname
clzc:zone-cluster-name:dataset> end
clzc:zone-cluster-name> verify
clzc:zone-cluster-name> commit
clzc:zone-cluster-name> exit
```

4 Verify the addition of the file system.

```
phys-schost# clzonecluster show -v zone-cluster-name
```

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 file systems in the pool on the zone-cluster

node that currently hosts 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

1 **Become superuser on a voting node of the global cluster that hosts the zone cluster.**

You perform all steps of this procedure from a voting node of the global cluster.

2 **On the global cluster, configure the cluster file system that you want to use in the zone cluster.**

3 **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
...
/dev/global/dsk/d12s0 /dev/global/rdsk/d12s0/ /global/fs ufs 2 no global, logging
```

4 **Configure the cluster file system as a loopback file system for the zone cluster.**

```
phys-schost# clzonecluster configure zone-cluster-name
clzc:zone-cluster-name> add fs
clzc:zone-cluster-name:fs> set dir=zone-cluster-lofs-mountpoint
clzc:zone-cluster-name:fs> set special=global-cluster-mount-point
clzc:zone-cluster-name:fs> set type=lofs
clzc:zone-cluster-name:fs> end
clzc:zone-cluster-name> verify
clzc:zone-cluster-name> commit
clzc:zone-cluster-name> exit
```

dir=*zone-cluster-lofs-mount-point*
   Specifies the file system mount point for LOFS to make the cluster file system available to the zone cluster.

special=*global-cluster-mount-point*
   Specifies the file system mount point of the original cluster file system in the global cluster.

For more information about creating loopback file systems, see "How to Create and Mount an LOFS File System" in *Oracle Solaris Administration: Devices and File Systems*.

5 **Verify the addition of the LOFS file system.**

```
phys-schost# clzonecluster show -v zone-cluster-name
```

Example 6–5   Adding a Cluster File System to a Zone Cluster

The following example shows how to add a cluster file system with mount point `/global/apache` to a zone cluster. The file system is available to a zone cluster using the loopback mount mechanism at the mount point `/zone/apache`.

```
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/oracle/dsk/d1 /dev/md/oracle/rdsk/d1 /global/apache ufs 2 yes global, logging

phys-schost-1# clzonecluster configure zone-cluster-name
clzc:zone-cluster-name> add fs
clzc:zone-cluster-name:fs> set dir=/zone/apache
clzc:zone-cluster-name:fs> set special=/global/apache
clzc:zone-cluster-name:fs> set type=lofs
clzc:zone-cluster-name:fs> end
clzc:zone-cluster-name> verify
clzc:zone-cluster-name> commit
clzc:zone-cluster-name> exit

phys-schost-1# clzonecluster show -v sczone
...
  Resource Name:                          fs
    dir:                                    /zone/apache
    special:                                /global/apache
    raw:
    type:                                   lofs
    options:                                []
    cluster-control:                        true
...
```

Next Steps Configure the cluster file system to be available in the zone cluster by using an HAStoragePlus resource. The HAStoragePlus resource manages the mounting of the file systems in the global cluster, and later performs 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 Storage Devices to a Zone Cluster

This section describes how to add the direct use of global storage devices by a zone cluster. 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.

Note – To import raw-disk devices ($cNtXdYsZ$) into a zone cluster node, use the zonecfg command as you normally would for other brands of non-global zones.

Such devices would not be under the control of the clzonecluster command, but would be treated as local devices of the node. See "Mounting File Systems in Running Non-Global Zones" in *Oracle Solaris Administration: Oracle Solaris Zones, Oracle Solaris 10 Zones, and Resource Management* for more information about importing raw-disk devices into a non-global zone.

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 an Individual Metadevice to a Zone Cluster (Solaris Volume Manager)

Perform this procedure to add an individual metadevice of a Solaris Volume Manager disk set to a zone cluster.

1 Become superuser on a node of the global cluster that hosts the zone cluster.

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

2 Identify the disk set that contains the metadevice to add to the zone cluster and determine whether it is online.

```
phys-schost# cldevicegroup status
```

3 If the disk set that you are adding is not online, bring it online.

```
phys-schost# cldevicegroup online diskset
```

4 Determine the set number that corresponds to the disk set to add.

```
phys-schost# ls -l /dev/md/diskset
lrwxrwxrwx  1 root root  8 Jul 22 23:11 /dev/md/diskset -> shared/set-number
```

5 Add the metadevice for use by the zone cluster.

You must use a separate add device session for each set match= entry.

Note – An asterisk (*) is used as a wildcard character in the path name.

```
phys-schost# clzonecluster configure zone-cluster-name
clzc:zone-cluster-name> add device
clzc:zone-cluster-name:device> set match=/dev/md/diskset/*dsk/metadevice
clzc:zone-cluster-name:device> end
clzc:zone-cluster-name> add device
clzc:zone-cluster-name:device> set match=/dev/md/shared/setnumber/*dsk/metadevice
clzc:zone-cluster-name:device> end
clzc:zone-cluster-name> verify
clzc:zone-cluster-name> commit
clzc:zone-cluster-name> exit
```

match=/dev/md/diskset/*dsk/metadevice
    Specifies the full logical device path of the metadevice

match=/dev/md/shared/*N*/*dsk/*metadevice*
    Specifies the full physical device path of the disk set number

6 **Reboot the zone cluster.**

The change becomes effective after the zone cluster reboots.

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

Example 6–6   Adding a Metadevice to a Zone Cluster

The following example adds the metadevice d1 in the disk set oraset to the sczone zone cluster.
The set number of the disk set is 3.

```
phys-schost-1# clzonecluster configure sczone
clzc:sczone> add device
clzc:sczone:device> set match=/dev/md/oraset/*dsk/d1
clzc:sczone:device> end
clzc:sczone> add device
clzc:sczone:device> set match=/dev/md/shared/3/*dsk/d1
clzc:sczone:device> end
clzc:sczone> verify
clzc:sczone> commit
clzc:sczone> exit

phys-schost-1# clzonecluster reboot sczone
```

▼ **How to Add a Disk Set to a Zone Cluster (Solaris Volume Manager)**

Perform this procedure to add an entire Solaris Volume Manager disk set to a zone cluster.

1 **Become superuser on a node of the global cluster that hosts the zone cluster.**

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

2 **Identify the disk set to add to the zone cluster and determine whether it is online.**

```
phys-schost# cldevicegroup status
```

3 **If the disk set that you are adding is not online, bring it online.**

```
phys-schost# cldevicegroup online diskset
```

4 **Determine the set number that corresponds to the disk set to add.**

```
phys-schost# ls -l /dev/md/diskset
lrwxrwxrwx  1 root root  8 Jul 22 23:11 /dev/md/diskset -> shared/set-number
```

5 **Add the disk set for use by the zone cluster.**

You must use a separate add device session for each set match= entry.

> **Note** – An asterisk (*) is used as a wildcard character in the path name.

```
phys-schost# clzonecluster configure zone-cluster-name
clzc:zone-cluster-name> add device
clzc:zone-cluster-name:device> set match=/dev/md/diskset/*dsk/*
clzc:zone-cluster-name:device> end
clzc:zone-cluster-name> add device
clzc:zone-cluster-name:device> set match=/dev/md/shared/set-number/*dsk/*
clzc:zone-cluster-name:device> end
clzc:zone-cluster-name> verify
clzc:zone-cluster-name> commit
clzc:zone-cluster-name> exit
```

match=/dev/md/*diskset*/*dsk/*
    Specifies the full logical device path of the disk set

match=/dev/md/shared/*N*/*dsk/*
    Specifies the full physical device path of the disk set number

6   Reboot the zone cluster.

The change becomes effective after the zone cluster reboots.

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

Example 6–7   Adding a Disk Set to a Zone Cluster

The following example adds the disk set oraset to the sczone zone cluster. The set number of the disk set is 3.

```
phys-schost-1# clzonecluster configure sczone
clzc:sczone> add device
clzc:sczone:device> set match=/dev/md/oraset/*dsk/*
clzc:sczone:device> end
clzc:sczone> add device
clzc:sczone:device> set match=/dev/md/shared/3/*dsk/*
clzc:sczone:device> end
clzc:sczone> verify
clzc:sczone> commit
clzc:sczone> exit

phys-schost-1# clzonecluster reboot sczone
```

▼   How to Add a DID Device to a Zone Cluster

Perform this procedure to add a DID device to a zone cluster.

1   Become superuser on a node of the global cluster that hosts the zone cluster.

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

2   Identify the DID device to add to the zone cluster.

The device you add must be connected to all nodes of the zone cluster.

```
phys-schost# cldevice list -v
```

3   Add the DID device for use by the zone cluster.

Note – An asterisk (*) is used as a wildcard character in the path name.

```
phys-schost# clzonecluster configure zone-cluster-name
clzc:zone-cluster-name> add device
clzc:zone-cluster-name:device> set match=/dev/did/*dsk/dNs*
clzc:zone-cluster-name:device> end
clzc:zone-cluster-name> verify
clzc:zone-cluster-name> commit
clzc:zone-cluster-name> exit
```

match=/dev/did/*dsk/dNs*
    Specifies the full device path of the DID device

4   Reboot the zone cluster.

The change becomes effective after the zone cluster reboots.

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

Example 6–8   Adding a DID Device to a Zone Cluster

The following example adds the DID device d10 to the sczone zone cluster.

```
phys-schost-1# clzonecluster configure sczone
clzc:sczone> add device
clzc:sczone:device> set match=/dev/did/*dsk/d10s*
clzc:sczone:device> end
clzc:sczone> verify
clzc:sczone> commit
clzc:sczone> exit

phys-schost-1# clzonecluster reboot sczone
```

7

Uninstalling Software From the Cluster

This chapter provides procedures for uninstalling or removing certain software from an Oracle Solaris Cluster configuration. This chapter contains the following procedures:

- "How to Unconfigure Oracle Solaris Cluster Software to Correct Installation Problems" on page 163
- "How to Uninstall Oracle Solaris Cluster Quorum Server Software" on page 166
- "How to Unconfigure a Zone Cluster" on page 167

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 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 118, 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.

- **SPARC:**

ok **boot -x**

- **x86:**

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 and Shutting Down Oracle Solaris on x86 Platforms*.

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.

■ **To unconfigure the node but leave Oracle Solaris Cluster software installed, run the following command:**

```
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.

■ **To unconfigure the node and also remove Oracle Solaris Cluster software, run the following command:**

```
phys-schost# /usr/cluster/bin/scinstall -r [-b BE-name]
```

-r   Removes cluster configuration information and uninstalls Oracle Solaris Cluster framework and data-service software from the cluster node. You can then reinstall the node or remove the node from the cluster.

-b *BE-name*   Specifies the name of a new boot environment, which is where you boot into after the uninstall process completes. Specifying a name is optional. If you do not specify a name for the boot environment, one is automatically generated.

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

7 Repeat Step 2 through Step 6 for each additional node to unconfigure.

Troubleshooting   If the cluster node that you are removing is at least partially configured with the cluster, running the clnode remove command might exit with errors such as Node is still enabled. If such errors occur, add the -F option to the clnode remove command.

Next Steps   Before you 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 4.0 Hardware Administration Manual* and the removal procedure in the Oracle Solaris Cluster manual for your storage array.

▼ How to Uninstall Oracle Solaris Cluster Quorum Server Software

Before You Begin   Before you uninstall Oracle Solaris Cluster 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.**

Alternatively, if your user account is assigned the System Administrator profile, issue commands as nonroot through a profile shell, or prefix the command with the pfexec command.

2   **Uninstall the quorum server software.**

```
quorumserver# pkg uninstall ha-cluster/group-package/ha-cluster-quorum-server* \
ha-cluster/service/quorum-server*
```

3   **(Optional) Clean up or remove the quorum server directories.**

By default, this directory is /var/scqsd.

▼ How to Unconfigure 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 *zone-cluster*" from each command.

---

a.   Take each resource offline.

```
phys-schost# clresource offline -Z zone-cluster resource-group
```

b.   List all enabled resources in the zone cluster.

```
phys-schost# clresource show -Z zone-cluster -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 zone-cluster -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 zone-cluster resource
```

See the clresource(1CL) man page for more information.

e.   Verify that all resources are disabled.

```
phys-schost# clresource show -Z zone-cluster -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** *zone-cluster* *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** *zone-cluster* **-t resource,resourcegroup**

    h. Delete all resource groups and their resources from the zone cluster.

       phys-schost# **clresourcegroup delete -F -Z** *zone-cluster* **+**

3 Halt the zone cluster.

   phys-schost# **clzonecluster halt** *zone-cluster-name*

4 Uninstall the zone cluster.

   phys-schost# **clzonecluster uninstall** *zone-cluster-name*

5 Unconfigure the zone cluster.

   phys-schost# **clzonecluster delete** *zone-cluster-name*

Index