Sun Cluster Software Installation Guide for Solaris OS

Contents

Preface

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

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

This document is intended for experienced system administrators with extensive knowledge of Sun 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 Solaris™ Operating System (Solaris OS) and expertise with the volume-manager software that is used with Sun Cluster software.

Using UNIX Commands

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

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

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

Typographic Conventions

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

TABLE P–1 Typographic Conventions

Typeface	Meaning	Example
AaBbCc123	The names of commands, files, and directories, and onscreen computer output	Edit your `.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 the C shell, Bourne shell, and Korn shell.

TABLE P–2 Shell Prompts

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

Related Documentation

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

Topic	Documentation
Overview	*Sun Cluster Overview for Solaris OS*
	Sun Cluster 3.2 1/09 Documentation Center
Concepts	*Sun Cluster Concepts Guide for Solaris OS*
Hardware installation and administration	*Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*
	Individual hardware administration guides
Software installation	*Sun Cluster Software Installation Guide for Solaris OS*
	Sun Cluster Quick Start Guide for Solaris OS
Data service installation and administration	*Sun Cluster Data Services Planning and Administration Guide for Solaris OS*
	Individual data service guides
Data service development	*Sun Cluster Data Services Developer's Guide for Solaris OS*
System administration	*Sun Cluster System Administration Guide for Solaris OS*
	Sun Cluster Quick Reference
Software upgrade	*Sun Cluster Upgrade Guide for Solaris OS*
Error messages	*Sun Cluster Error Messages Guide for Solaris OS*
Command and function references	*Sun Cluster Reference Manual for Solaris OS*
	Sun Cluster Data Services Reference Manual for Solaris OS
	Sun Cluster Quorum Server Reference Manual for Solaris OS

For a complete list of Sun Cluster documentation, see the release notes for your release of Sun Cluster software at http://wikis.sun.com/display/SunCluster/Home/.

Related Third-Party Web Site References

Sun is not responsible for the availability of third-party web sites mentioned in this document. Sun does not endorse and is not responsible or liable for any content, advertising, products, or other materials that are available on or through such sites or resources. Sun will not be responsible or liable for any actual or alleged damage or loss caused or alleged to be caused by or in connection with use of or reliance on any such content, goods, or services that are available on or through such sites or resources.

Documentation, Support, and Training

The Sun web site provides information about the following additional resources:

- Documentation (http://www.sun.com/documentation/)
- Support (http://www.sun.com/support/)
- Training (http://www.sun.com/training/)

Sun Welcomes Your Comments

Sun is interested in improving its documentation and welcomes your comments and suggestions. To share your comments, go to http://docs.sun.com and click Feedback.

Getting Help

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

- Your name and email address (if available)
- Your company name, address, and phone number
- The model number and serial number of your systems
- The release number of the Solaris OS (for example, Solaris 10)
- The release number of Sun Cluster (for example, Sun Cluster 3.2 1/09)

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

Command	Function
psrinfo -v	Displays information about processors
showrev -p	Reports which patches are installed
SPARC: prtdiag -v	Displays system diagnostic information
/usr/cluster/bin/clnode show-rev	Displays Sun Cluster release and package version information

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

Planning the Sun Cluster Configuration

This chapter provides planning information and guidelines for installing a Sun Cluster configuration.

The following overview information is in this chapter:

- "Finding Sun Cluster Installation Tasks" on page 15
- "Planning the Solaris OS" on page 16
- "Planning the Sun Cluster Environment" on page 25
- "Planning the Global Devices, Device Groups, and Cluster File Systems" on page 41
- "Planning Volume Management" on page 46

Finding Sun Cluster Installation Tasks

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

TABLE 1–1   Sun Cluster Software Installation Task Information

Task	Instructions
Set up cluster hardware.	*Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS* Documentation that shipped with your server and storage devices
Plan global-cluster software installation.	Chapter 1, "Planning the Sun Cluster Configuration" "Installation and Configuration Worksheets" on page 246
Install software packages. Optionally, install and configure Sun StorageTek™ QFS software.	"Installing the Software" on page 55 *Sun StorageTek QFS Installation and Upgrade Guide, Version 4, Update 6*

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

Task	Instructions
Establish a new global cluster or a new global-cluster node.	"Establishing a New Global Cluster or New Global-Cluster Node" on page 82
Configure Solaris Volume Manager software.	"Configuring Solaris Volume Manager Software" on page 157 Solaris Volume Manager documentation
Install and configure Veritas Volume Manager (VxVM) software.	"Installing and Configuring VxVM Software" on page 187 VxVM documentation
Configure cluster file systems, if used.	"How to Create Cluster File Systems" on page 205
(Optional) On the Solaris 10 OS, create non-global zones.	"Configuring a Non-Global Zone on a Global-Cluster Node" on page 209
(Optional) On the Solaris 10 OS, create zone clusters.	"Configuring a Zone Cluster" on page 213
(Optional) SPARC: Install and configure the Sun Cluster module to Sun Management Center.	"SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 227 Sun Management Center documentation
Plan, install, and configure resource groups and data services. Create highly available local file systems, if used.	*Sun Cluster Data Services Planning and Administration Guide for Solaris OS*
Develop custom data services.	*Sun Cluster Data Services Developer's Guide for Solaris OS*

Planning the Solaris OS

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

- "Guidelines for Selecting Your Solaris Installation Method" on page 17
- "Solaris OS Feature Restrictions" on page 17
- "Solaris Software Group Considerations" on page 18
- "System Disk Partitions" on page 19
- "Guidelines for Non-Global Zones in a Global Cluster" on page 22
- "SPARC: Guidelines for Sun Logical Domains in a Cluster" on page 23

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

Guidelines for Selecting Your Solaris Installation Method

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

See "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 100 for details about the scinstall JumpStart installation method. See your Solaris installation documentation for details about standard Solaris installation methods.

Solaris OS Feature Restrictions

Consider the following points when you plan the use of the Solaris OS in a Sun Cluster configuration:

- **Solaris 10 Zones** – Install Sun Cluster framework software *only* in the global zone.

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

  If you configure non-global zones on a global-cluster node, the loopback file system (LOFS) must be enabled. See the information for LOFS for additional considerations.

- **Loopback file system (LOFS)** – During cluster creation with the Solaris 9 version of Sun Cluster software , LOFS capability is disabled by default. During cluster creation with the Solaris 10 version of Sun Cluster software, LOFS capability is not disabled by default.

  If the cluster meets both of the following conditions, you must disable LOFS to avoid switchover problems or other failures:

  - Sun Cluster 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 Sun Cluster HA for NFS.

- **Interface groups** – Solaris interface groups are not supported in a Sun Cluster configuration. The Solaris interface groups feature is disabled by default during Solaris software installation. Do not re-enable Solaris interface groups. See the ifconfig(1M) man page for more information about Solaris interface groups.

- **Power-saving shutdown** – Automatic power-saving shutdown is not supported in Sun Cluster configurations and should not be enabled. See the pmconfig(1M) and power.conf(4) man pages for more information.

- **IP Filter** – Sun Cluster software does not support the Solaris IP Filter feature for scalable services, but does support Solaris IP Filter for failover services.

- **fssnap** – Sun 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 Sun Cluster software. The following restrictions apply to fssnap support:

  - The fssnap command is supported on local files systems that are not managed by Sun 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.

Solaris Software Group Considerations

Sun Cluster 3.2 1/09 software requires at least the End User Solaris Software Group (SUNWCuser). However, other components of your cluster configuration might have their own Solaris software requirements as well. Consider the following information when you decide which Solaris software group you are installing.

- **Servers** – Check your server documentation for any Solaris software requirements. For example , Sun Enterprise™ 10000 servers require the Entire Solaris Software Group Plus OEM Support.

- **SCI-PCI adapters** – To use SCI-PCI adapters, which are available for use in SPARC based clusters only, or the Remote Shared Memory Application Programming Interface (RSMAPI), ensure that you install the RSMAPI software packages, which are SUNWrsm and SUNWrsmo, and for the Solaris 9 OS on SPARC based platforms also SUNWrsmx and SUNWrsmox. The RSMAPI software packages are included only in some Solaris software groups. For example, the Developer Solaris Software Group includes the RSMAPI software packages but the End User Solaris Software Group does not.

  If the software group that you install does not include the RSMAPI software packages, install the RSMAPI software packages manually before you install Sun Cluster software. Use the pkgadd(1M) command to manually install the software packages. See the Section (3RSM) man pages for information about using the RSMAPI.

- **Additional Solaris packages** – You might need to install other Solaris software packages that are not part of the End User Solaris Software Group. The Apache HTTP server packages are one example. Third-party software, such as ORACLE®, might also require additional Solaris software packages. See your third-party documentation for any Solaris software requirements.

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

System Disk Partitions

Add this information to the appropriate "Local File System Layout Worksheet" on page 247.

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

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

- /globaldevices – Create a file system at least 512 Mbytes large that is to be used by the scinstall(1M) utility for global devices.
- **Volume manager** – Create a 20-Mbyte partition on slice 7 for volume manager use. If your cluster uses Veritas Volume Manager (VxVM) and you intend to encapsulate the root disk, you need to have two unused slices available for use by VxVM.

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

See the following guidelines for additional partition planning information:

- "Guidelines for the Root (/) File System" on page 19
- "Guidelines for the /globaldevices File System" on page 20
- "Volume Manager Requirements" on page 21

Guidelines for the Root (/) File System

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

Note – No file-system type other than UFS is valid for the root (/) file system. Do not attempt to change the file-system type after the root (/) file system is created.

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

- root (/) – The Sun Cluster software itself occupies less than 40 Mbytes of space in the root (/) file system. Solaris Volume Manager software requires less than 5 Mbytes, and VxVM software requires less than 15 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.

- /var – The Sun Cluster software occupies a negligible amount of space in the /var file system at installation time. However, you need to set aside ample space for log files. Also, more messages might be logged on a clustered node than would be found on a typical stand-alone server. Therefore, allow at least 100 Mbytes for the /var file system.

- /usr – Sun Cluster software occupies less than 25 Mbytes of space in the /usr file system. Solaris Volume Manager and VxVM software each require less than 15 Mbytes.

- /opt – Sun Cluster framework software uses less than 2 Mbytes in the /opt file system. However, each Sun Cluster data service might use between 1 Mbyte and 5 Mbytes. Solaris Volume Manager software does not use any space in the /opt file system. VxVM software can use over 40 Mbytes if all of its packages and tools are installed.

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

    SPARC: If you use Sun Management Center software to monitor the cluster, you need an additional 25 Mbytes of space on each Solaris host to support the Sun Management Center agent and Sun Cluster module packages.

Guidelines for the /globaldevices **File System**

Sun Cluster software requires you to set aside a dedicated file system on one of the local disks for use in managing global devices. This file system is usually located on your root disk. However, if you use different storage on which to locate the global-devices file system, such as a Logical Volume Manager volume, it must not be part of a Solaris Volume Manager shared disk set or part of a VxVM disk group other than a root disk group. This file system is later mounted as a UFS cluster file system. Name this file system /globaldevices, which is the default name that is recognized by the scinstall(1M) command.

Note – No file-system type other than UFS is valid for the global-devices file system. Do not attempt to change the file-system type after the global-devices file system is created.

The `scinstall` command later renames the file system /global/.devices/node@*nodeid*, where *nodeid* represents the number that is assigned to a Solaris host when it becomes a global-cluster member. The original /globaldevices mount point is removed.

The /globaldevices file system must have ample space and ample inode capacity for creating both block special devices and character special devices. This guideline is especially important if a large number of disks are in the cluster. A file system size of 512 Mbytes should suffice for most cluster configurations.

Volume Manager Requirements

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

If you use Veritas Volume Manager (VxVM) and you intend to encapsulate the root disk, you need to have two unused slices that are available for use by VxVM. Additionally, you need to have some additional unassigned free space at either the beginning or the end of the disk. See your VxVM documentation for more information about root disk encapsulation.

Example – Sample File-System Allocations

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

This layout allows for the use of either Solaris Volume Manager software or VxVM software. If you use Solaris Volume Manager software, you use slice 7 for the state database replica. If you use VxVM, you later free slice 7 by assigning the slice a zero length. This layout provides the necessary two free slices, 4 and 7, as well as provides for unused space at the end of the disk.

TABLE 1–2 Example File-System Allocation

Slice	Contents	Size Allocation	Description
0	/	6.75GB	Remaining free space on the disk after allocating space to slices 1 through 7. Used for the Solaris OS, Sun Cluster software, data-services software, volume-manager software, Sun Management Center agent and Sun Cluster module agent packages, root file systems, and database and application software.
1	swap	1GB	512 Mbytes for the Solaris OS.
			512 Mbytes for Sun Cluster software.
2	overlap	8.43GB	The entire disk.
3	/globaldevices	512MB	The Sun Cluster software later assigns this slice a different mount point and mounts the slice as a cluster file system.
4	unused	-	Available as a free slice for encapsulating the root disk under VxVM.
5	unused	-	-
6	unused	-	-
7	volume manager	20MB	Used by Solaris Volume Manager software for the state database replica, or used by VxVM for installation after you free the slice.

Guidelines for Non-Global Zones in a Global Cluster

For information about the purpose and function of Solaris 10 zones in a cluster, see "Support for Solaris Zones" in *Sun Cluster Concepts Guide for Solaris OS*.

For guidelines about configuring a cluster of non-global zones, see "Zone Clusters" on page 39.

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

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

becomes a single point of failure for the resource group. For highest availability, zones in a failover resource group's node list should be on different hosts.

In single-host clusters, no functional risk is incurred if you specify multiple zones in a failover resource group's node list.

- **Scalable services** – Do not create non-global zones for use in the same scalable service on the same Solaris host. Each instance of the scalable service must run on a different host.

- **Cluster file systems** - Do not directly add a cluster file system from the global zone to a non-global zone. Instead, add a loopback mount of the cluster file system from the global zone to the non-global zone. This restriction does not apply to QFS shared file systems.

- **LOFS** – Solaris Zones requires that the loopback file system (LOFS) be enabled. However, the Sun Cluster HA for NFS data service requires that LOFS be disabled, to avoid switchover problems or other failures. If you configure both non-global zones and Sun Cluster HA for NFS in your cluster, do one of the following to prevent possible problems in the data service:

  - Disable the `automountd` daemon.

  - Exclude from the automounter map all files that are part of the highly available local file system that is exported by Sun Cluster HA for NFS.

- **Exclusive-IP zones** – The following guidelines apply specifically to exclusive-IP non-global zones:

  - **Logical-hostname resource groups** – In a resource group that contains a `LogicalHostname` resource, if the node list contains any non-global zone with the `ip-type` property set to `exclusive`, all zones in that node list must have this property set to `exclusive`. Note that a global zone always has the `ip-type` property set to `shared`, and therefore cannot coexist in a node list that contains zones of `ip-type=exclusive`. This restriction applies only to versions of the Solaris OS that use the Solaris zones `ip-type` property.

  - **IPMP groups** – For all public-network adapters that are used for data-service traffic in the non-global zone, you must manually configure IPMP groups in all `/etc/hostname.`*adapter* files on the zone. This information is not inherited from the global zone. For guidelines and instructions to configure IPMP groups, follow the procedures in Part VI, "IPMP," in *System Administration Guide: IP Services*.

  - **Private-hostname dependency** - Exclusive-IP zones cannot depend on the private hostnames and private addresses of the cluster.

  - **Shared-address resources** – Shared-address resources cannot use exclusive-IP zones.

SPARC: Guidelines for Sun Logical Domains in a Cluster

Consider the following points when you create a Sun Logical Domains (LDoms) 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 a Sun LDoms guest domain must be a full SCSI LUN in the I/O domain. You cannot use an arbitrary virtual device.

- **Fencing** – Do not export a storage LUN to more than one guest domain on the same physical machine, unless you also disable fencing for that device. Otherwise, if two different guest domains on the same machine both are visible to a device, the device will be fenced whenever one of the guest domains dies. The fencing of the device will panic any other guest domain that subsequently tries to access the device.

- **Network isolation** – Guest domains that are located on the same physical machine but are configured in different clusters must be network isolated from each other. Use one of the following methods:

  - Configure the clusters to use different network interfaces in the I/O domain for the private network.

  - Use different network addresses for each of the clusters.

- **Networking in guest domains** – Network packets to and from guest domains must traverse service domains to reach the network drivers through virtual switches. Virtual switches use kernel threads that run at system priority. The virtual-switch threads must be able to acquire needed CPU resources to perform critical cluster operations, including heartbeats, membership, checkpoints, and so forth. Configuring virtual switches with the mode=sc setting enables expedited handling of cluster heartbeat packets. However, the reliability of other critical cluster operations can be enhanced by adding more CPU resources to the service domain under the following workloads:

  - High-interrupt load, for example, due to network or disk I/O. Under extreme load, virtual switches can preclude system threads from running for a long time, including virtual-switch threads.

  - Real-time threads that are overly aggressive in retaining CPU resources. Real-time threads run at a higher priority than virtual-switch threads, which can restrict CPU resources for virtual-switch threads for an extended time.

- **Exporting storage from I/O domains** – If you configure a cluster that is composed of Sun Logical Domains I/O domains, do not export its storage devices to other guest domains that also run Sun Cluster software.

- **Solaris I/O multipathing** – Do not run Solaris I/O multipathing software (MPxIO) from guest domains. Instead, run Solaris I/O multipathing software in the I/O domain and export it to the guest domains.

- **Private-interconnect IP address range** – The private network is shared by all guest domains that are created on the same physical machine and it is visible to all these domains. Before you specify a private-network IP address range to the scinstall utility for use by a guest-domain cluster, ensure that the address range is not already in use by another guest domain on the same physical machine.

For more information about Sun Logical Domains, see the *Logical Domains (LDoms) 1.0.3 Administration Guide*.

Planning the Sun Cluster Environment

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

- "Licensing" on page 25
- "Software Patches" on page 25
- "Public-Network IP Addresses" on page 25
- "Console-Access Devices" on page 26
- "Logical Addresses" on page 27
- "Public Networks" on page 27
- "Quorum Servers" on page 28
- "NFS Guidelines" on page 29
- "Service Restrictions" on page 30
- "Sun Cluster Configurable Components" on page 31
- "Zone Clusters" on page 39

For detailed information about Sun Cluster components, see the *Sun Cluster Overview for Solaris OS* and the *Sun Cluster Concepts Guide for Solaris OS*.

Licensing

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

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

Software Patches

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

- For information about current required patches, see "Patches and Required Firmware Levels" in *Sun Cluster Release Notes* or consult your Sun service provider.
- For general guidelines and procedures for applying patches, see Chapter 10, "Patching Sun Cluster Software and Firmware," in *Sun Cluster System Administration Guide for Solaris OS*.

Public-Network IP Addresses

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

You must set up a number of public-network IP addresses for various Sun Cluster components, depending on your cluster configuration. Each 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 Solaris software
- For IPv6 IP addresses on the Solaris 9 OS, the local /etc/inet/ipnodes file on each global-cluster node, after you install Solaris software
- The local /etc/inet/hosts file on any exclusive-IP non-global zone

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

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

For more information about planning IP addresses, see Chapter 3, "Planning Your TCP/IP Network (Task)," in *System Administration Guide: IP Services* (Solaris 9) or Chapter 2, "Planning Your TCP/IP Network (Tasks)," in *System Administration Guide: IP Services* (Solaris 10).

Console-Access Devices

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

- A terminal concentrator is used to communicate between the administrative console and the global-cluster node consoles.

- A Sun Enterprise 10000 server uses a System Service Processor (SSP) instead of a terminal concentrator.
- A Sun Fire server uses a system controller instead of a terminal concentrator.

For more information about console access, see the *Sun Cluster Concepts Guide for Solaris OS*.

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

Logical Addresses

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

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

Public Networks

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

- **Separation of public and private network** – Public networks and the private network (cluster interconnect) must use separate adapters, or you must configure tagged VLAN on tagged-VLAN capable adapters and VLAN-capable switches to use the same adapter for both the private interconnect and the public network.
- **Minimum** – All cluster nodes must be connected to at least one public network. Public-network connections can use different subnets for different nodes.
- **Maximum** – You can have as many additional public-network connections as your hardware configuration allows.
- **Scalable services** – All nodes that run a scalable service must either use the same subnet or set of subnets or use different subnets that are routable among themselves.
- **IPv4** – Sun Cluster software supports IPv4 addresses on the public network.
- **IPv6** – Sun Cluster software supports IPv6 addresses on the public network under the following conditions or restrictions:
  - Sun Cluster software does not support IPv6 addresses on the public network if the private interconnect uses SCI adapters.
  - Sun Cluster software supports IPv6 addresses 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.

  In the Sun Cluster 3.2 1/09 release, the `scinstall` utility no longer automatically configures a single-adapter IPMP group on each unconfigured public-network adapter during Sun Cluster creation. Instead, 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. On the Solaris 10 OS, these groups are probe based.

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

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

  For guidelines and instructions to configure IPMP groups, follow the procedures in Part VI, "IPMP," in *System Administration Guide: IP Services*. To modify IPMP groups after cluster installation, follow the guidelines in "How to Administer IP Network Multipathing Groups in a Cluster" in *Sun Cluster System Administration Guide for Solaris OS* and procedures in Chapter 28, "Administering Network Multipathing (Task)," in *System Administration Guide: IP Services* (Solaris 9) or Chapter 31, "Administering IPMP (Tasks)," in *System Administration Guide: IP Services* (Solaris 10).

- **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. Sun Cluster software does not support a `local-mac-address?` value of `false` for Ethernet adapters. This requirement is a change from Sun Cluster 3.0, which did require a `local-mac-address?` value of `false`.

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

Quorum Servers

You can use Sun 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 a Sun 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** – Solaris software requirements for Sun Cluster software apply as well to Quorum Server software.

- **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 that it provides quorum to. For example, a SPARC based machine that runs the Solaris 9 OS can be configured as a quorum server for an x86 based cluster that runs the Solaris 10 OS.

- **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 a Sun Cluster configuration.

- **NFS client** – No Sun Cluster node can be an NFS client of a Sun Cluster HA for NFS-exported file system that is being mastered on a node in the same cluster. Such cross-mounting of Sun Cluster 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 Sun Cluster HA for NFS data service on the same cluster node. If you do, certain Sun Cluster HA for NFS data-service activities might cause the NFS daemons to stop and restart, interrupting NFS services. However, you can safely run the Sun Cluster HA for NFS data service if you use the NFSv4 protocol to mount external NFS file systems on the cluster nodes.

- **Locking** – Applications that run locally on the cluster must not lock files on a file system that is exported through NFS. Otherwise, local blocking (for example, flock(3UCB) or fcntl(2)) might interfere with the ability to restart the lock manager (lockd(1M)). During restart, a blocked local process might be granted a lock which might be intended to be reclaimed by a remote client. This would cause unpredictable behavior.

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

  However, Sun Cluster software does support the following security features for NFS:

- The use of secure ports for NFS. You enable secure ports for NFS by adding the entry set nfssrv:nfs_portmon=1 to the /etc/system file on cluster nodes.

- The use of Kerberos with NFS. For more information, see "Securing Sun Cluster HA for NFS With Kerberos V5" in *Sun Cluster Data Service for NFS Guide for Solaris OS*.

- **No fencing support for NAS devices in non-global zones** – Sun Cluster software does not provide fencing support for NFS-exported file systems from a NAS device when such file systems are used in a non-global zone, including nodes of a zone cluster. Fencing support is provided only for NFS-exported file systems in the global zone.

Service Restrictions

Observe the following service restrictions for Sun Cluster configurations:

- **Routers** – Do not configure cluster nodes as routers (gateways). If the system goes down, the clients cannot find an alternate router and cannot recover.

- **NIS+ servers** – Do not configure cluster nodes as NIS or NIS+ servers. There is no data service available for NIS or NIS+. However, cluster nodes can be NIS or NIS+ clients.

- **Boot and install servers** – Do not use a Sun Cluster configuration to provide a highly available boot or installation service on client systems.

- **RARP** – Do not use a Sun Cluster configuration to provide an rarpd service.

- **RPC program numbers** – If you install an RPC service on the cluster, the service must not use any of the following program numbers:

    - 100141
    - 100142
    - 100248

    These numbers are reserved for the Sun 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** – Sun 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

    Sun 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 Sun Cluster kernel threads from acquiring needed CPU cycles.

Sun Cluster Configurable Components

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

- "Global-Cluster Name" on page 31
- "Global-Cluster Voting-Node Names" on page 31
- "Zone Names" on page 31
- "Private Network" on page 32
- "Private Hostnames" on page 34
- "Cluster Interconnect" on page 34
- "Global Fencing" on page 36
- "Quorum Devices" on page 37

Add this information to the appropriate configuration planning worksheet.

Global-Cluster Name

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

For information about naming a zone cluster, see "Zone Clusters" on page 39.

Global-Cluster Voting-Node Names

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

For information about node names in a zone cluster, see "Zone Clusters" on page 39.

Zone Names

On the Solaris 10 OS in versions that support Solaris brands, a non-global zone of brand `native` is a valid potential node of a resource-group node list. Use the naming convention *nodename*:*zonename* to specify a non-global zone to a Sun Cluster command.

- The *nodename* is the name of the Solaris host.

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

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

For information about a cluster of non-global zones, see "Zone Clusters" on page 39.

Private Network

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.

Sun Cluster software uses the private network for internal communication among nodes and among non-global zones that are managed by Sun Cluster software. A Sun Cluster configuration requires at least two connections to the cluster interconnect on the private network. When you configure Sun 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.
    - On the Solaris 10 OS, the default netmask is 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.
    - On the Solaris 9 OS, the default netmask is 255.255.248.0. This IP address range supports a combined maximum of 64 nodes 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. On the Solaris 10 OS, 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 *Sun Cluster System Administration Guide for Solaris OS*. You must bring down the cluster to make these changes.

However, on the Solaris 10 OS the cluster can remain in cluster mode if you use the `cluster set-netprops` command to change only the netmask. For any zone cluster that is already configured in the cluster, the private IP subnets and the corresponding private IP addresses that are allocated for that zone cluster will also be updated.

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

- **Address and netmask sizes** – The private network address cannot be smaller than the netmask. For example, you can use a private network address of `172.16.10.0` with a netmask of `255.255.255.0`. But you cannot use a private network address of `172.16.10.0` with a netmask of `255.255.0.0`.

- **Acceptable addresses** – The address must be included in the block of addresses that RFC 1918 reserves for use in private networks. You can contact the InterNIC to obtain copies of RFCs or view RFCs online at http://www.rfcs.org.

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

    For Sun Logical Domains (LDoms) guest domains that are created on the same physical machine and that are connected to the same virtual switch, the private network is shared by such guest domains and is visible to all these domains. Proceed with caution before you specify a private-network IP address range to the `scinstall` utility for use by a cluster of guest domains. Ensure that the address range is not already in use by another guest domain that exists on the same physical machine and shares its virtual switch.

- **IPv6** – Sun Cluster software does not support IPv6 addresses for the private interconnect. The system does configure IPv6 addresses on the private-network adapters to support scalable services that use IPv6 addresses. But internode communication on the private network does not use these IPv6 addresses.

See "Planning Your TCP/IP Network (Tasks)," in *System Administration Guide: IP Services* (Solaris 9 or Solaris 10) 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 Sun Cluster configuration of a global cluster or a zone cluster. These private hostnames follow the naming convention clusternode*nodeid*-priv, where *nodeid* is the numeral of the internal node ID. During Sun 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.

For the Solaris 10 OS, 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 *Sun Cluster Concepts Guide for Solaris OS*.

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 Sun 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 35.
- You can set up from one to six cluster interconnects in a cluster. While a single cluster interconnect reduces the number of adapter ports that are used for the private interconnect, it provides no redundancy and less availability. If a single interconnect fails, the cluster is at a higher risk of having to perform automatic recovery. Whenever possible, install two or more cluster interconnects to provide redundancy and scalability, and therefore higher availability, by avoiding a single point of failure.

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

For guidelines about cluster interconnect hardware, see "Interconnect Requirements and Restrictions" in *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*. For general information about the cluster interconnect, see "Cluster-Interconnect Components" in *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS*.

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** – Sun Cluster software does not support IPv6 communications over the private interconnects.

- **Local MAC address assignment** – All private network adapters must use network interface cards (NICs) that support local MAC address assignment. Link-local IPv6 addresses, which are required on private-network adapters to support IPv6 public-network addresses, are derived from the local MAC addresses.

- **Tagged VLAN adapters** – Sun Cluster software supports tagged Virtual Local Area Networks (VLANs) to share an adapter between the private cluster interconnect and the public network. To configure a tagged VLAN adapter for the cluster interconnect, specify the adapter name and its VLAN ID (VID) in one of the following ways:

  - Specify the usual adapter name, which is the device name plus the instance number or physical point of attachment (PPA). For example, the name of instance 2 of a Cassini Gigabit Ethernet adapter would be ce2. If the scinstall utility asks whether the adapter is part of a shared virtual LAN, answer **yes** and specify the adapter's VID number.

  - Specify the adapter by its VLAN virtual device name. This name is composed of the adapter name plus the VLAN instance number. The VLAN instance number is derived from the formula $(1000*V)+N$, where V is the VID number and N is the PPA.

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

  For information about configuring VLAN in a cluster, see "Configuring VLANs as Private Interconnect Networks" in *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*. For general information about VLAN, see "Administering Virtual Local Area Networks" in *System Administration Guide: IP Services*.

- **SPARC: Sun LDoms 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.

- **SBus SCI adapters** – The SBus Scalable Coherent Interface (SCI) is not supported as a cluster interconnect. However, the SCI-PCI interface is supported.

- **Logical network interfaces** – Logical network interfaces are reserved for use by Sun Cluster software.

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

Transport Switches

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

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

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

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

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

Global Fencing

Fencing is a mechanism that is used by the cluster to protect the data integrity of a shared disk during split-brain situations. By default, the scinstall utility in Typical Mode leaves global fencing enabled, and each shared disk in the configuration uses the default global fencing setting of pathcount. With the pathcount setting, the fencing protocol for each shared disk is chosen based on the number of DID paths that are attached to the disk.

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 to support the following situations:

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

- The 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 Sun 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 *Sun Cluster Concepts Guide for Solaris OS*. 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

Sun 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 *Sun Cluster Concepts Guide for Solaris OS*.

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

If you want to use a quorum server or a Network Appliance NAS device 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(1CL) utility.

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

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

Consider the following points when you plan quorum devices.

- **Minimum** – A two-host cluster must have at least one quorum device, which can be a shared disk, a quorum server, or a NAS device. For other topologies, quorum devices are optional.

- **Odd-number rule** – If more than one quorum device is configured in a two-host cluster, or in a pair of hosts directly connected to the quorum device, configure an odd number of quorum devices. This configuration ensures that the quorum devices have completely independent failure pathways.

- **Distribution of quorum votes** – For highest availability of the cluster, ensure that the total number of votes that are contributed by quorum devices is less than the total number of votes that are contributed by voting nodes. Otherwise, the nodes cannot form a cluster if all quorum devices are unavailable, even if all nodes are functioning.

- **Connection** – You must connect a quorum device to at least two voting nodes.

- **SCSI fencing protocol** – When a SCSI shared-disk quorum device is configured, its fencing protocol is automatically set to SCSI-2 in a two-host cluster or SCSI-3 in a cluster with three or more voting nodes.

- **Changing the fencing protocol of quorum devices** – For SCSI disks that are configured as a quorum device, you must unconfigure the quorum device before you can enable or disable its SCSI fencing protocol.

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

  The software quorum protocol would also be used by SCSI shared disks if fencing is disabled for such disks.

- **Replicated devices** – Sun 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 *Sun Cluster Concepts Guide for Solaris OS* and "Quorum Devices" in *Sun Cluster Overview for Solaris OS*.

Zone Clusters

On the Solaris 10 OS, a zone cluster is a cluster of non-global zones. All nodes of a zone cluster are configured as non-global zones of the cluster brand. No other brand type is permitted in a zone cluster. You can run supported services on the zone cluster similar to a global cluster, with the isolation that is provided by Solaris zones.

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 Sun Cluster configuration. A zone cluster cannot be configured without an underlying global cluster.

- **Minimum Solaris OS** – The global cluster must run at least the Solaris 10 5/08 OS.

- **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 Solaris zones can be exported to a zone cluster. Such devices include the following:

    - Solaris disk devices (cNtXdYsZ)

    - DID devices (/dev/did/*dsk/dN)

    - Solaris Volume Manager and Solaris Volume Manager for Sun Cluster multi-owner disk sets (/dev/md/*setname*/*dsk/dN)

Zone-Cluster Requirements and Guidelines

- **Distribution of nodes** – You cannot host multiple nodes of the same zone cluster on the same node of the global cluster. A global-cluster node can host multiple zone-cluster nodes as long as each node 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 names of the nodes must be unique within the zone cluster. The infrastructure automatically creates an underlying non-global zone on each global-cluster node that hosts 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 global-cluster node that hosts the zone cluster is also zc1.

- **Cluster name** – The name of the zone cluster must be unique throughout the global cluster. The name cannot also be used by a non-global zone elsewhere in the global cluster, nor can the 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 assign a specific public-network IP address to each zone-cluster node.

- **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 34.

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

- **Conversion to a zone-cluster node** – You cannot add an existing non-global zone to a zone cluster.

- **File systems** – You can use the clzonecluster command to add only the following types of file systems for use by a zone cluster:

  - Highly available local file systems

  - QFS shared file systems, which are supported for use by Oracle Real Application Clusters

  Do not directly add a cluster file system from the global zone to a zone cluster node. Instead, add a loopback mount of the cluster file system from the global zone to the non-global zone.

  To add a local file system to a zone cluster, you must instead use the zonecfg command as you normally would in a stand-alone system.

- **No fencing support for NAS devices in non-global zones** – Sun Cluster software does not provide fencing support for NFS-exported file systems from a NAS device when such file systems are used in a non-global zone, including nodes of a zone cluster. Fencing support is provided only for NFS-exported file systems in the global zone.

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

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

- "Global Devices" on page 41
- "Device Groups" on page 41
- "Cluster File Systems" on page 42
- "Choosing Mount Options for Cluster File Systems" on page 44
- "Mount Information for Cluster File Systems" on page 45

Global Devices

For information about the purpose and function of global devices, see "Global Devices, Local Devices, and Device Groups" in *Sun Cluster Overview for Solaris OS* and "Global Devices" in *Sun Cluster Concepts Guide for Solaris OS*.

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

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

Device Groups

For information about the purpose and function of device groups, see "Global Devices, Local Devices, and Device Groups" in *Sun Cluster Overview for Solaris OS* and "Device Groups" in *Sun Cluster Concepts Guide for Solaris OS*.

Add this planning information to the "Device Group Configurations Worksheet" on page 251.

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 51 for additional guidelines. See "Configuring Solaris Volume Manager Software" on page 157 or "Installing and Configuring VxVM Software" on page 187 and your volume-manager documentation for instructions about mirroring.

- **Storage-based replication** – Disks in a device group must be either all replicated or none replicated. A device group cannot use a mix of replicated and nonreplicated disks.

Cluster File Systems

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

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 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

Consider the following points when you plan cluster file systems.

- **Quotas** – Quotas are not supported on cluster file systems. However, quotas are supported on highly available local file systems.

- **Non-global zones** – If a cluster file system is to be accessed from a non-global zone, it must first be mounted in the global zone. The cluster file system is then mounted in the non-global zone by using a loopback mount. Therefore, the loopback file system (LOFS) must be enabled in a cluster that contains non-global zones.

- **Zone clusters** – You cannot configure cluster file systems for use in a zone cluster. Use highly available local file systems instead. If the zone cluster is configured with Oracle Real Application Clusters (RAC), you can use shared QFS in that zone cluster to support Oracle RAC.

- **Loopback file system (LOFS)** – During cluster creation with the Solaris 9 version of Sun Cluster software, LOFS is disabled by default. During cluster creation with the Solaris 10 version of Sun Cluster software, LOFS is enabled by default.

You must manually disable LOFS on each voting cluster node if the cluster meets both of the following conditions:

- Sun Cluster 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 Sun Cluster 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 Solaris software by which one would put a communication endpoint in the file-system namespace.

  - Although you can create a UNIX domain socket whose name is a path name into the cluster file system, the socket would not survive a node failover.

  - Any FIFOs or named pipes that you create on a cluster file system would not be globally accessible.

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

- **Device special files** – Neither block special files nor character special files are supported in a cluster file system. To specify a path name to a device node in a cluster file system, create a symbolic link to the device name in the /dev directory. Do not use the mknod command for this purpose.

- **atime** – Cluster file systems do not maintain atime.

- **ctime** – When a file on a cluster file system is accessed, the update of the file's ctime might be delayed.

- **Installing applications** - If you want the binaries of a highly available application to reside on a cluster file system, wait to install the application until after the cluster file system is configured. Also, if the application is installed by using the Sun Java System installer program and the application depends on any shared components, install those shared components on all nodes in the cluster that are not installed with the application.

Choosing Mount Options for Cluster File Systems

This section describes requirements and restrictions for the following types of cluster file systems:

- "UFS Cluster File Systems" on page 44
- "VxFS Cluster File Systems" on page 45

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

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

UFS Cluster File Systems

Mount Option	Usage	Description
global	Required	This option makes the file system globally visible to all nodes in the cluster.
logging	Required	This option enables logging.
forcedirectio	Conditional	This option is required only for cluster file systems that will host Oracle Real Application Clusters RDBMS data files, log files, and control files.
		Note – Oracle Real Application Clusters is supported for use only in SPARC based clusters.
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 Sun 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.

Mount Option	Usage	Description
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.

VxFS Cluster File Systems

Mount Option	Usage	Description
global	Required	This option makes the file system globally visible to all nodes in the cluster.
log	Required	This option enables logging.

See the VxFS mount_vxfs man page and "Overview of Administering Cluster File Systems" in *Sun Cluster System Administration Guide for Solaris OS* for more information about VxFS 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.

- **SPARC: VxFS mount requirement** – If you use Veritas File System (VxFS), globally mount and unmount a VxFS file system from the primary node. The primary node is the Solaris host that masters the disk on which the VxFS file system resides. This method ensures that the mount or unmount operation succeeds. A VxFS file-system mount or unmount operation that is performed from a secondary node might fail.

- **SPARC: VxFS feature restrictions** –

The following VxFS features are not supported in a Sun Cluster 3.2 cluster file system. They are, however, supported in a local file system.

- Quick I/O
- Snapshots
- Storage checkpoints
- VxFS-specific mount options:
  - `convosync` (Convert O_SYNC)
  - `mincache`
  - `qlog`, `delaylog`, `tmplog`
- Veritas cluster file system (requires VxVM cluster feature & Veritas Cluster Server)

Cache advisories can be used, but the effect is observed on the given node only.

All other VxFS features and options that are supported in a cluster file system are supported by Sun Cluster 3.2 software. See VxFS documentation for details about VxFS options that are supported in a cluster configuration.

- **Nesting mount points** – Normally, you should not nest the mount points for cluster file systems. For example, do not set up one file system that is mounted on `/global/a` and another file system that is mounted on `/global/a/b`. To ignore this rule can cause availability and node boot-order problems. These problems would occur if the parent mount point is not present when the system attempts to mount a child of that file system. The only exception to this rule is if the devices for the two file systems have the same physical host connectivity. An example is different slices on the same disk.
- `forcedirectio` – Sun 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

Add this planning information to the "Device Group Configurations Worksheet" on page 251 and the "Volume-Manager Configurations Worksheet" on page 253. For Solaris Volume Manager, also add this planning information to the "Volumes Worksheet (Solaris Volume Manager)" on page 255.

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

- "Guidelines for Volume-Manager Software" on page 47
- "Guidelines for Solaris Volume Manager Software" on page 48
- "Guidelines for Veritas Volume Manager Software" on page 50
- "File-System Logging" on page 51
- "Mirroring Guidelines" on page 51

Sun Cluster software uses volume-manager software to group disks into device groups which can then be administered as one unit. Sun Cluster software supports Solaris Volume Manager software and Veritas Volume Manager (VxVM) software that you install or use in the following ways.

TABLE 1–4 Supported Use of Volume Managers With Sun Cluster Software

Volume-Manager Software	Requirements
Solaris Volume Manager	You must install Solaris Volume Manager software on all voting nodes of the cluster, regardless of whether you use VxVM on some nodes to manage disks.
SPARC: VxVM with the cluster feature	You must install and license VxVM with the cluster feature on all voting nodes of the cluster.
VxVM without the cluster feature	You are only required to install and license VxVM on those voting nodes that are attached to storage devices that VxVM manages.
Both Solaris Volume Manager and VxVM	If you install both volume managers on the same voting node, you must use Solaris Volume Manager software to manage disks that are local to each node. Local disks include the root disk. Use VxVM to manage all shared disks.

See your volume-manager documentation and "Configuring Solaris Volume Manager Software" on page 157 or "Installing and Configuring VxVM Software" on page 187 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 *Sun Cluster Concepts Guide for Solaris OS* and "Device Groups" in *Sun Cluster Concepts Guide for Solaris OS*.

Guidelines for Volume-Manager Software

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

- **Software RAID** – Sun 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 51 for guidelines on mirroring multihost disks. You do not need to use software mirroring if the storage device provides hardware RAID as well as redundant paths to devices.

- **Mirrored root** – Mirroring the root disk ensures high availability, but such mirroring is not required. See "Mirroring Guidelines" on page 51 for guidelines about deciding whether to mirror the root disk.

- **Unique naming** – You might have local Solaris Volume Manager or VxVM volumes that are used as devices on which the /global/.devices/node@*nodeid* file systems are mounted. If so, the name of each local volume on which a /global/.devices/node@*nodeid* file system is to be mounted must be unique throughout the cluster.

- **Node lists** – To ensure high availability of a device group, make its node lists of potential masters and its failback policy identical to any associated resource group. Or, if a scalable resource group uses more nodes than its associated device group, make the scalable resource group's node list a superset of the device group's node list. See the resource group planning information in the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* 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. However, configured VxVM disk groups do not have an association to any particular set of nodes.

- **Hot-spare disks** – You can use hot-spare disks to increase availability, but hot spare disks are not required.

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

Guidelines for Solaris Volume Manager Software

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

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

- **Dual-string mediators** – Each disk set configured with exactly two disk strings and mastered by exactly two Solaris hosts must have Solaris Volume Manager mediators configured for the disk set. 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. Observe the following rules to configure dual-string mediators:

  - You must configure each disk set with exactly two hosts that act as mediator hosts.

  - You must use the same two hosts for all disk sets that require mediators. Those two hosts must master those disk sets.

  - 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.

■ /kernel/drv/md.conf **settings** – SPARC: On the Solaris 9 OS, Solaris Volume Manager volumes that are used by each disk set are created in advance, at reconfiguration boot time. This reconfiguration is based on the configuration parameters that exist in the /kernel/drv/md.conf file.

Note – With the Solaris 10 release, Solaris Volume Manager has been enhanced to configure volumes dynamically. You no longer need to edit the nmd and the md_nsets parameters in the /kernel/drv/md.conf file. New volumes are dynamically created, as needed.

You must modify the nmd and md_nsets fields as follows to support a Sun Cluster configuration on the Solaris 9 OS:

Caution – All voting cluster nodes must have identical /kernel/drv/md.conf files, regardless of the number of disk sets that are served by each node. Failure to follow this guideline can result in serious Solaris Volume Manager errors and possible loss of data.

■ md_nsets – The md_nsets field defines the total number of disk sets that can be created for a system to meet the needs of the entire cluster. Set the value of md_nsets to the expected number of disk sets in the cluster plus one additional disk set. Solaris Volume Manager software uses the additional disk set to manage the private disks on the local host.

The maximum number of disk sets that are allowed per cluster is 32. This number allows for 31 disk sets for general use plus one disk set for private disk management. The default value of md_nsets is 4.

■ nmd – The nmd field defines the highest predicted value of any volume name that will exist in the cluster. For example, if the highest value of the volume names that are used in the first 15 disk sets of a cluster is 10, but the highest value of the volume in the 16th disk set is 1000, set the value of nmd to at least 1000. Also, the value of nmd must be large enough to ensure that enough numbers exist for each device–ID name. The number must also be large enough to ensure that each local volume name can be unique throughout the cluster.

The highest allowed value of a volume name per disk set is 8192. The default value of nmd is 128.

Set these fields at installation time to allow for all predicted future expansion of the cluster. To increase the value of these fields after the cluster is in production is time consuming. The value change requires a reconfiguration reboot for each voting node. To raise these values later also increases the possibility of inadequate space allocation in the root (/) file system to create all of the requested devices.

At the same time, keep the value of the nmd field and the md_nsets field as low as possible. Memory structures exist for all possible devices as determined by nmd and md_nsets, even if

you have not created those devices. For optimal performance, keep the value of nmd and md_nsets only slightly higher than the number of volumes that you plan to use.

See "System Files and Startup Files" in *Solaris Volume Manager Administration Guide* (Solaris 9 or Solaris 10) for more information about the md.conf file.

Guidelines for Veritas Volume Manager Software

Consider the following points when you plan Veritas Volume Manager (VxVM) configurations.

- **Accessibility to nodes** – You must configure all volume-manager disk groups as either Sun Cluster device groups or as local-only disk groups. If you do not configure the disk group in one of these ways, the devices in the disk group will not be accessible to any node in the cluster.
  - A device group enables a secondary node to host multihost disks if the primary node fails.
  - A local-only disk group functions outside the control of Sun Cluster software and can be accessed from only one node at a time.
- **Enclosure-Based Naming** – If you use Enclosure-Based Naming of devices, ensure that you use consistent device names on all cluster nodes that share the same storage. VxVM does not coordinate these names, so the administrator must ensure that VxVM assigns the same names to the same devices from different nodes. Failure to assign consistent names does not interfere with correct cluster behavior. However, inconsistent names greatly complicate cluster administration and greatly increase the possibility of configuration errors, potentially leading to loss of data.
- **Root disk group** – The creation of a root disk group is optional.

  A root disk group can be created on the following disks:
  - The root disk, which must be encapsulated
  - One or more local nonroot disks, which you can encapsulate or initialize
  - A combination of root and local nonroot disks

  The root disk group must be local to the Solaris host.
- **Simple root disk groups** – Simple root disk groups, which are created on a single slice of the root disk, are not supported as disk types with VxVM on Sun Cluster software. This is a general VxVM software restriction.
- **Encapsulation** – Disks to be encapsulated must have two disk-slice table entries free.
- **Number of volumes** – Estimate the maximum number of volumes any given device group can use at the time the device group is created.
  - If the number of volumes is less than 1000, you can use default minor numbering.

- If the number of volumes is 1000 or greater, you must carefully plan the way in which minor numbers are assigned to device group volumes. No two device groups can have overlapping minor number assignments.

- **Dirty Region Logging** – The use of Dirty Region Logging (DRL) decreases volume recovery time after a node failure. Using DRL might decrease I/O throughput.

- **Dynamic Multipathing (DMP)** – The use of DMP alone to manage multiple I/O paths per Solaris host to the shared storage is not supported. The use of DMP is supported only in the following configurations:

  - A single I/O path per host is configured to the cluster's shared storage.
  - A supported multipathing solution is used, such as Solaris I/O multipathing software (MPxIO), EMC PowerPath, or Hitachi HDLM, that manages multiple I/O paths per host to the shared cluster storage.

See your VxVM installation documentation for additional information.

File-System Logging

Logging is required for UFS and VxFS cluster file systems. Sun Cluster software supports the following choices of file-system logging:

- Solaris UFS logging – See the `mount_ufs(1M)` man page for more information.
- SPARC: Veritas File System (VxFS) logging – See the `mount_vxfs` man page provided with VxFS software for more information.

Both Solaris Volume Manager and Veritas Volume Manager support both types of file-system logging.

Mirroring Guidelines

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

Guidelines for Mirroring Multihost Disks

To mirror all multihost disks in a Sun Cluster configuration enables the configuration to tolerate single-device failures. Sun 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 and Veritas Volume Manager (VxVM) software support three-way mirroring. However, Sun 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 Disk Storage" in *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS*.

Guidelines for Mirroring the Root Disk

Add this planning information to the "Local File System Layout Worksheet" on page 247.

For maximum availability, mirror root (/), /usr, /var, /opt, and swap on the local disks. Under VxVM, you encapsulate the root disk and mirror the generated subdisks. However, Sun Cluster software does not require that you mirror the root disk.

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

See your volume-manager documentation and "Configuring Solaris Volume Manager Software" on page 157 or "Installing and Configuring VxVM Software" on page 187 for instructions about how to mirror the root disk.

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

- **Boot disk** – You can set up the mirror to be a bootable root disk. You can then boot from the mirror if the primary boot disk fails.
- **Complexity** – To mirror the root disk adds complexity to system administration. To mirror the root disk also complicates booting in single-user mode.
- **Backups** – Regardless of whether you mirror the root disk, you also should perform regular backups of root. Mirroring alone does not protect against administrative errors. Only a backup plan enables you to restore files that have been accidentally altered or deleted.
- **Quorum devices** – Do not use a disk that was configured as a quorum device to mirror a root disk.
- **Quorum** – Under Solaris Volume Manager software, in failure scenarios in which state database quorum is lost, you cannot reboot the system until maintenance is performed. See your Solaris Volume Manager documentation for information about the state database and state database replicas.

- **Separate controllers** – Highest availability includes mirroring the root disk on a separate controller.

- **Secondary root disk** – With a mirrored root disk, the primary root disk can fail but work can continue on the secondary (mirror) root disk. Later, the primary root disk might return to service, for example, after a power cycle or transient I/O errors. Subsequent boots are then performed by using the primary root disk that is specified for the `eeprom(1M)` `boot-device` parameter. In this situation, no manual repair task occurs, but the drive starts working well enough to boot. With Solaris Volume Manager software, a resync does occur. A resync requires a manual step when the drive is returned to service.

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

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

◆ ◆ ◆ **C H A P T E R 2**

2

Installing Software on Global-Cluster Nodes

This chapter provides procedures for how to install software on global-cluster voting nodes and the administration console.

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 56
(*Optional*) Install and configure a quorum server.	"How to Install and Configure Quorum Server Software" on page 57
(*Optional*) Install Cluster Control Panel (CCP) software on the administrative console.	"How to Install Cluster Control Panel Software on an Administrative Console" on page 60
Install the Solaris OS on all nodes.	"How to Install Solaris Software" on page 64
(*Optional*) Configure internal disk mirroring.	"How to Configure Internal Disk Mirroring" on page 68
(*Optional*) Install Sun Logical Domains (LDoms) software and create domains.	"SPARC: How to Install Sun Logical Domains Software and Create Domains" on page 69
(*Optional*) SPARC: Install and configure Solaris I/O multipathing software.	"How to Install Solaris I/O Multipathing Software" on page 70

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

Task	Instructions
(*Optional*) SPARC: Install Veritas File System software.	"SPARC: How to Install Veritas File System Software" on page 72
Install Sun Cluster software and any data services that you will use.	"How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73
Set up directory paths.	"How to Set Up the Root Environment" on page 78
(*Optional*) Configure Solaris IP Filter.	"How to Configure Solaris IP Filter" on page 78

▼ How to Prepare for Cluster Software Installation

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

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

   Contact your Sun 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.**

   - *Sun Cluster Release Notes* - Restrictions, bug workarounds, and other late-breaking information.

   - *Sun Cluster Overview for Solaris OS* and *Sun Cluster Concepts Guide for Solaris OS* - Overviews of the Sun Cluster product.

   - *Sun Cluster Software Installation Guide for Solaris OS* (this manual) - Planning guidelines and procedures for installing and configuring Solaris, Sun Cluster, and volume-manager software.

   - *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* - 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:

   - Solaris OS
   - Solaris Volume Manager software
   - Sun StorageTek QFS software
   - Veritas Volume Manager
   - Third-party applications

4 **Plan your cluster configuration.**

⚠️ **Caution** – Plan your cluster installation completely. Identify requirements for all data services and third-party products **before** you begin Solaris and Sun Cluster software installation. Failure to do so might result in installation errors that require that you completely reinstall the Solaris and Sun Cluster software.

For example, the Oracle Real Application Clusters Guard option of Oracle Real Application Clusters has special requirements for the hostnames that you use in the cluster. Another example with special requirements is Sun Cluster HA for SAP. You must accommodate these requirements before you install Sun Cluster software because you cannot change hostnames after you install Sun Cluster software.

- Use the planning guidelines in Chapter 1, "Planning the Sun Cluster Configuration," and in the *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* to determine how to install and configure your cluster.
- Fill out the cluster framework and data-services configuration worksheets that are referenced in the planning guidelines. Use your completed worksheets for reference during the installation and configuration tasks.

5 Obtain all necessary patches for your cluster configuration.

See "Patches and Required Firmware Levels" in *Sun Cluster Release Notes* for the location of patches and installation instructions.

Next Steps   If you want to use Cluster Control Panel software to connect from an administrative console to your cluster nodes, go to "How to Install Cluster Control Panel Software on an Administrative Console" on page 60.

Otherwise, choose the Solaris installation procedure to use.

- To configure Sun Cluster software by using the `scinstall(1M)` utility, go to "How to Install Solaris Software" on page 64 to first install Solaris software.
- To install and configure Solaris and Sun Cluster software in the same operation (JumpStart method), go to "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 100.

▼ How to Install and Configure Quorum Server Software

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

Before You Begin   Perform the following tasks:

- Ensure that the machine that you choose for the quorum server has at least 1 Mbyte of disk space available for Quorum Server software installation.
- Ensure that the quorum-server machine is connected to a public network that is accessible to the cluster nodes.

■ Disable the spanning tree algorithm on the Ethernet switches for the ports that are connected to the cluster public network where the quorum server will run.

1 **Become superuser on the machine to install with Quorum Server software.**

2 **(Optional) To use the** `installer` **program with a GUI, ensure that the display environment of the host server to install is set to display the GUI.**

```
# xhost +
# setenv DISPLAY nodename:0.0
```

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

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

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

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

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

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

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

5 **Start the installation wizard.**

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

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

Choose the Configure Later option.

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

After installation is finished, you can view any available installation log. See the *Sun Java Enterprise System 2006Q4 Installation Guide for UNIX* for additional information about using the Java Enterprise System `installer` program.

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

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

b. Eject the media.

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

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

See "Patches and Required Firmware Levels" in *Sun Cluster Release Notes* for the location of patches and installation instructions.

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

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

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

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

11   Configure the quorum server.

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

Identify the quorum server by using at least one of either an instance name or a port number. You must provide the port number, but the instance name is optional. If you provide an instance name, that name must be unique among your quorum servers. If you do not provide a name, always refer to this quorum server by the port on which it listens.

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

-d */var/scqsd*	The path to the directory where the quorum server can store quorum data.
	The quorum-server process creates one file per cluster in this directory to store cluster-specific quorum information.
	By default, the value of this option is /var/scqsd. This directory must be unique for each quorum server that you configure.
-i *instancename*	A unique name that you choose for the quorum-server instance.
-p *port*	The port number on which the quorum server listens for requests from the cluster.

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

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

14   Start the newly configured quorum server.

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

<blockquote>

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

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

</blockquote>

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

- SUNWscqsr
- SUNWscqsu
- SUNWscqsman

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

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

Next Steps   If you want to use an administrative console to communicate with the cluster nodes, go to "How to Install Cluster Control Panel Software on an Administrative Console" on page 60.

Otherwise, go to "How to Install Solaris Software" on page 64.

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

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

You cannot use this software to connect to Sun Logical Domains (LDoms) guest domains.

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

You can use any desktop machine that runs a version of the Solaris OS that is supported by Sun Cluster 3.2 1/09 software as an administrative console. If you are using Sun Cluster software on a SPARC based system, you can also use the administrative console as a Sun Management Center console or server as well. See Sun Management Center documentation for information about how to install Sun Management Center software.

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

1   Become superuser on the administrative console.

2   Load the Sun Java Availability Suite DVD-ROM into the DVD-ROM drive.

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

3   Change to the /cdrom/suncluster_3_0Packages/ **directory, where** *arch* **is** sparc **or** x86 **(Solaris 10 only), and where** *ver* **is** 9 **for Solaris 9 or** 10 **for Solaris 10 .**

adminconsole# **cd /cdrom/cdrom0//cdrom/suncluster_3_0Packages/**

4   Install the SUNWccon **package.**

adminconsole# **pkgadd -d . SUNWccon**

5   (Optional) Install Sun Cluster man-page packages.

adminconsole# **pkgadd -d .** *pkgname* **...**

Package Name	Description
SUNWscman	Sun Cluster framework man pages
SUNWscdsman	Sun Cluster data-service man pages
SUNWscqsman	Sun Cluster Quorum Server man pages

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

6   Unload the Sun Java Availability Suite DVD-ROM from the DVD-ROM drive.

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

b. Eject the DVD-ROM.

```
adminconsole# eject cdrom
```

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

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

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

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

8 **Create an** /etc/serialports **file.**

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

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

node1, node2	Physical names of the cluster nodes.
ca-dev-hostname	Hostname of the console-access device.
port	Serial port number, or the Secure Shell port number for Secure Shell connections.

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

- For a Sun Fire 15000 system controller, use telnet(1) port number 23 for the serial port number of each entry.

- For all other console-access devices, to connect to the console through a telnet connection, use the telnet serial port number, not the physical port number. To determine the telnet serial port number, add 5000 to the physical port number. For example, if a physical port number is 6, the telnet serial port number is 5006.

- For Sun Enterprise 10000 servers, also see the /opt/SUNWcluster/bin/serialports(4) man page for details and special considerations.

- For Secure Shell connections to node consoles, specify for each node the name of the console-access device and the port number to use for secure connection. The default port number for Secure Shell is 22.

- To connect the administrative console directly to the cluster nodes or through a management network, specify for each node its hostname and the port number that the node uses to connect to the administrative console or the management network.

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

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

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

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

10 **Start the CCP utility.**

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

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

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

The CCP software supports the following Secure Shell connections:

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

- For secure connection to the cluster nodes, use the cssh tool.

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

Next Steps   Determine whether the Solaris OS is already installed to meet Sun Cluster software requirements. See "Planning the Solaris OS" on page 16 for information about Sun Cluster installation requirements for the Solaris OS.

- If the Solaris OS meets Sun Cluster requirements, go to "How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73.

- If the Solaris OS does not meet Sun Cluster requirements, install, reconfigure, or reinstall the Solaris OS as needed.

  - To install the Solaris OS alone, go to "How to Install Solaris Software" on page 64.

  - To use the scinstall custom JumpStart method to install both the Solaris OS and Sun Cluster software, go to "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 100

▼ How to Install Solaris Software

If you do not use the scinstall custom JumpStart installation method to install software, perform this procedure to install the Solaris OS on each node in the global cluster. See "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 100 for more information about JumpStart installation of a cluster.

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

If your nodes are already installed with the Solaris OS but do not meet Sun Cluster installation requirements, you might need to reinstall the Solaris software. Follow the steps in this procedure to ensure subsequent successful installation of Sun Cluster software. See "Planning the Solaris OS" on page 16 for information about required root-disk partitioning and other Sun 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 Solaris software. See the *Sun Cluster Hardware Administration Collection* 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 56 for requirements and guidelines.

- Complete the "Local File System Layout Worksheet" on page 247.

- 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 25 for planning guidelines. See your Solaris system-administrator documentation for information about using Solaris naming services.

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

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

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

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

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

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

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

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

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

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

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

See "Solaris Software Group Considerations" on page 18 for information about additional Solaris software requirements.

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

- **Create a file system of at least 512 Mbytes for use by the global-device subsystem.**

  Note – Sun Cluster software requires a global-devices file system for installation to succeed.

- **Specify that slice 7 is at least 20 Mbytes in size.**

- **Create any other file-system partitions that you need, as described in "System Disk Partitions" on page 19.**

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

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

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

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

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

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

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

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

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

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

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

5   If you are adding a node and VxVM is installed on any node in the cluster, perform the following tasks.

   a. Ensure that the same vxio number is used on the VxVM-installed nodes.

```
phys-schost# grep vxio /etc/name_to_major
vxio NNN
```

   b. Ensure that the vxio number is available for use on each of the nodes that do not have VxVM installed.

   c. If the vxio number is already in use on a node that does not have VxVM installed, change the /etc/name_to_major entry to use a different number.

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

- Remote Shared Memory Application Programming Interface (RSMAPI)
- RSMRDT drivers
- SPARC: SCI-PCI adapters

- **SPARC: For the Solaris 9 OS, use the following command:**

```
phys-schost# pkgadd -d . SUNWrsm SUNWrsmc SUNWrsmo SUNWrsmox
```

- **For the Solaris 10 OS, use the following command:**

```
phys-schost# pkgadd -G -d . SUNWrsm SUNWrsmo
```

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

7   Install any required Solaris OS patches and hardware-related firmware and patches, including those for storage-array support. Also download any needed firmware that is contained in the hardware patches.

See "Patches and Required Firmware Levels" in *Sun Cluster Release Notes* for the location of patches and installation instructions.

8   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.

- **On the Solaris 9 OS, set the default to** kadb**.**

  phys-schost# **eeprom boot-file=kadb**

- **On the Solaris 10 OS, set the default to** kmdb **in the GRUB boot parameters menu.**

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

9   Update the /etc/inet/hosts **or** /etc/inet/ipnodes **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. The ipnodes file can contain both IPv4 and IPv6 addresses. See "Public-Network IP Addresses" on page 25 for a listing of Sun Cluster components whose IP addresses you must add.

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. Adding these IP addresses to the /etc/inet/ipnodes file is optional.

---

10   If you will use ce **adapters for the cluster interconnect, add the following entry to the** /etc/system **file.**

set ce:ce_taskq_disable=1

This entry becomes effective after the next system reboot.

11   (Optional) On Sun Enterprise 10000 servers, configure the /etc/system **file to use dynamic reconfiguration.**

Add the following entry to the /etc/system file on each node of the cluster:

set kernel_cage_enable=1

This entry becomes effective after the next system reboot. See your server documentation for more information about dynamic reconfiguration.

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

If you do not want to use the multiple-adapter IPMP groups that the scinstall utility configures during cluster creation, configure custom IPMP groups as you would in a stand-alone system. See Chapter 31, "Administering IPMP (Tasks)," in *System Administration Guide: IP Services* for details.

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

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 68.

Otherwise, to use Solaris I/O multipathing software, go to "How to Install Solaris I/O Multipathing Software" on page 70.

Otherwise, to install VxFS, go to "SPARC: How to Install Veritas File System Software" on page 72.

Otherwise, install the Sun Cluster software packages. Go to "How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73.

See Also See the *Sun Cluster System Administration Guide for Solaris OS* for procedures to perform dynamic reconfiguration tasks in a Sun Cluster configuration.

▼ How to Configure Internal Disk Mirroring

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

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

- Your servers do not support the mirroring of internal hard drives.
- You have already established the cluster. Instead, perform "Mirroring Internal Disks on Servers that Use Internal Hardware Disk Mirroring or Integrated Mirroring" in *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*.

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

1 Become superuser.

2   Configure an internal mirror.

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

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

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

Next Steps   SPARC: To create Sun Logical Domains (LDoms), go to "SPARC: How to Install Sun Logical
Domains Software and Create Domains" on page 69.

To use Solaris I/O multipathing software, go to "How to Install Solaris I/O Multipathing
Software" on page 70.

Otherwise, to install VxFS, go to "SPARC: How to Install Veritas File System Software" on
page 72.

Otherwise, install the Sun Cluster software packages. Go to "How to Install Sun Cluster
Framework and Data-Service Software Packages" on page 73.

▼ SPARC: How to Install Sun Logical Domains Software and Create Domains

Perform this procedure to install Sun Logical Domains (LDoms) 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.
- Ensure that the Solaris 10 OS is installed.
- Have available *Logical Domains (LDoms) 1.0.3 Administration Guide* and *Logical Domains
  (LDoms) 1.0.3 Release Notes*.
- Read the requirements and guidelines in "SPARC: Guidelines for Sun Logical Domains in a
  Cluster" on page 23.

1   Become superuser on the machine.

2   Install Sun Logical Domains software and configure domains.

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

    If you create guest domains, adhere to the Sun 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 68.

Otherwise, to use Solaris I/O multipathing software, go to "How to Install Solaris I/O Multipathing Software" on page 70.

Otherwise, to install VxFS, go to "SPARC: How to Install Veritas File System Software" on page 72.

Otherwise, install the Sun Cluster software packages. Go to "How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73.

▼ How to Install Solaris I/O Multipathing Software

Perform this procedure on each node of the global cluster to install and configure Solaris I/O multipathing software (MPxIO) for Fibre Channel (FC) storage. Multipathing software manages multiple I/O paths to the shared cluster storage. This procedure is optional.

- SPARC: For the Solaris 9 OS, you install and configure Sun StorEdge Traffic Manager software.

- For the Solaris 10 OS, you enable the Solaris multipathing feature, which is installed by default as part of the Solaris 10 software.

Before You Begin   Perform the following tasks:

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

    If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 64 for more information about installing Solaris software to meet Sun Cluster software requirements.

- SPARC: For the Solaris 9 OS, have available your software packages, patches, and documentation for Sun StorEdge Traffic Manager software and Sun StorEdge SAN Foundation software. See `http://www.sun.com/products-n-solutions/hardware/docs/` for links to documentation.

- For the Solaris 10 OS, have available the *Solaris Fibre Channel Storage Configuration and Multipathing Support Guide*.

1 Become superuser.

Note – SPARC: If you installed Sun Logical Domains (LDoms) software, perform this procedure in the I/O domain and export it to the guest domains. Do not enable Solaris I/O multipathing software directly in guest domains.

2 SPARC: For the Solaris 9 OS, install on each node Sun StorEdge Traffic Manager software and any necessary patches.

- For the procedure about how to install Sun StorEdge Traffic Manager software, see the *Sun StorEdge Traffic Manager Installation and Configuration Guide* at `http://www.sun.com/products-n-solutions/hardware/docs/`.

- For a list of required patches for Sun StorEdge Traffic Manager software, see the *Sun StorEdge Traffic Manager Software Release Notes* at `http://www.sun.com/storage/san/`.

3 Enable multipathing functionality.

- **SPARC: For the Solaris 9 OS, change the value of the** `mpxio-disable` **parameter to** `no`**.**

  Modify this entry in the `/kernel/drv/scsi_vhci.conf` file on each node.

  `set mpxio-disable=no`

- **For the Solaris 10 OS, issue the following command on each node:**

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

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

  `-e`   Enables Solaris I/O multipathing.

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

4 SPARC: For the Solaris 9 OS, determine whether your version of Sun StorEdge SAN Foundation software includes built-in support for your storage array.

If the software does not include built-in support for your storage array, edit the /kernel/drv/scsi_vhci.conf file on each node to include the necessary entries. For more information, see the release notes for your storage device.

5 SPARC: For the Solaris 9 OS, shut down each node and perform a reconfiguration boot.

The reconfiguration boot creates the new Solaris device files and links.

```
phys-schost# shutdown -y -g0 -i0
ok boot -r
```

6 After the reconfiguration reboot is finished on all nodes, perform any additional tasks that are necessary to complete the configuration of your storage array.

See installation instructions for your storage array in the *Sun Cluster Hardware Administration Collection* for details.

Troubleshooting If you installed Solaris I/O multipathing software after Sun Cluster software was installed on the cluster, DID mappings might require updating. Issue the following commands on each node of the cluster to regenerate the DID namespace.

```
phys-schost# cldevice clearphys-schost# cldevice refresh
(Solaris 9 only) phys-schost# cfgadm -c configure
phys-schost# cldevice populate
```

See the cfgadm(1M) and cldevice(1CL) man pages for more information.

Next Steps To install VxFS, go to "SPARC: How to Install Veritas File System Software" on page 72.

Otherwise, install the Sun Cluster software packages. Go to "How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73.

▼ SPARC: How to Install Veritas File System Software

To use Veritas File System (VxFS) software in the cluster, perform this procedure on each node of the global cluster.

1 Follow the procedures in your VxFS installation documentation to install VxFS software on each node of the cluster.

2 Install any Sun Cluster patches that are required to support VxFS.

See "Patches and Required Firmware Levels" in *Sun Cluster Release Notes* for the location of patches and installation instructions.

3 In the /etc/system **file on each node, set the following values.**

```
set rpcmod:svc_default_stksize=0x8000
set lwp_default_stksize=0x6000
```

These changes become effective at the next system reboot.

- Sun Cluster software requires a minimum rpcmod:svc_default_stksize setting of 0x8000. Because VxFS installation sets the value of the rpcmod:svc_default_stksize variable to 0x4000, you must manually set the value to 0x8000 after VxFS installation is complete.

- You must set the lwp_default_stksize variable in the /etc/system file to override the VxFS default value of 0x4000.

Next Steps Install the Sun Cluster software packages. Go to "How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73.

▼ How to Install Sun Cluster Framework and Data-Service Software Packages

Note – You can alternatively deploy the Sun Cluster Plug-in for Sun N1™ Service Provisioning System to install Sun Cluster framework and data-service software. Follow instructions in the documentation that is provided with the plug-in. You can also access this information at http://wikis.sun.com/display/SunCluster/Sun+Cluster+Framework+Plug-in.

Follow this procedure to use the Sun Java™ Enterprise System (Java ES) installer program to perform one or more of the following installation tasks:

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

- To install Sun Cluster framework software on the master node where you will create a flash archive for a JumpStart installation. See "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 100 for more information about a JumpStart installation of a global cluster.

- To install data services.

> **Note** – For the Solaris 10 OS, this procedure installs data services only to the global zone. To install data services to be visible only from within a certain non-global zone, see "How to Create a Non-Global Zone on a Global-Cluster Node" on page 210.

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

Before You Begin Perform the following tasks:

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

  If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 64 for more information about installing Solaris software to meet Sun Cluster software requirements.

- Have available the Sun Java Availability Suite DVD-ROM.

1 (Solaris 10 only) Restore external access to RPC communication and optionally to Sun Java Web Console.

During the installation of the Solaris 10 OS, if you choose not to enable network services for remote clients, a restricted network profile is used that disables external access for certain network services. The restricted services include the following services that affect cluster functionality:

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

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

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

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

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

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

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

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

For more information about what services the restricted network profile restricts to local connections, see "Planning Network Security" in *Solaris 10 5/08 Installation Guide: Planning for Installation and Upgrade*.

2 (Optional) To use the `installer` **program with a GUI, ensure that the display environment of the cluster node to install is set to display the GUI.**

```
% xhost +
% setenv DISPLAY nodename:0.0
```

If you do not make these settings, the `installer` program runs in text-based mode.

3 Become superuser on the cluster node to install.

Note – If your physically clustered machines are configured with Sun LDoms, install Sun Cluster software only in I/O domains or guest domains.

---

4 Load the Sun Java Availability Suite DVD-ROM into the DVD-ROM drive.

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

5 Change to the installation wizard directory of the DVD-ROM.

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

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

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

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

6 Start the installation wizard program.

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

See the *Sun Java Enterprise System 5 Installation Guide for UNIX* for additional information about using the different forms and features of the Java ES `installer` program.

7 Follow instructions on the screen to install Sun Cluster framework software and data services on the node.

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

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

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

  After the cluster is established, see *Sun Cluster Geographic Edition Installation Guide* for further installation procedures.

- Choose Configure Later when prompted whether to configure Sun Cluster framework software.

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

8 Install additional packages to use any of the following features.

- Remote Shared Memory Application Programming Interface (RSMAPI)
- SCI-PCI adapters for the interconnect transport
- RSMRDT drivers

Note – Use of the RSMRDT driver is restricted to clusters that run an Oracle9i Release 2 SCI configuration with RSM enabled. Refer to Oracle9i Release 2 user documentation for detailed installation and configuration instructions.

---

a. Determine which packages you must install.

The following table lists the Sun Cluster 3.2 1/09 packages that each feature requires, in the order in which you must install each group of packages. The Java ES `installer` program does not automatically install these packages.

Note – Install packages in the order in which they are listed in the following table.

Feature	Additional Sun Cluster 3.2 1/09 Packages to Install
RSMAPI	SUNWscrif

Feature	Additional Sun Cluster 3.2 1/09 Packages to Install
SCI-PCI adapters	■ *Solaris 9*: SUNWsci SUNWscid SUNWscidx ■ *Solaris 10*: SUNWscir SUNWsci SUNWscidr SUNWscid
RSMRDT drivers	SUNWscrdt

b. Change to the /cdrom/suncluster_3_0Packages/ **directory, where** *arch* **is** sparc **or** x86 **(Solaris 10 only), and where** *ver* **is** 9 **for Solaris 9 or** 10 **for Solaris 10 .**

phys-schost# **cd /cdrom/cdrom0//cdrom/suncluster_3_0Packages/**

c. Install the additional packages.

■ **SPARC: For the Solaris 9 OS, use the following command:**

phys-schost# **pkgadd -d .** *packages*

■ **For the Solaris 10 OS, use the following command:**

phys-schost# **pkgadd -G -d .** *packages*

9 Unload the Sun Java Availability Suite DVD-ROM from the DVD-ROM drive.

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

b. Eject the DVD-ROM.

phys-schost# **eject cdrom**

10 Apply any necessary patches to support Sun Cluster software.

See "Patches and Required Firmware Levels" in *Sun Cluster Release Notes* for the location of patches and installation instructions.

Next Steps If you want to install Sun StorageTek QFS file system software, follow the procedures for initial installation in the *Sun StorEdge QFS Installation and Upgrade Guide*.

Otherwise, to set up the root user environment, go to "How to Set Up the Root Environment" on page 78.

▼ How to Set Up the Root Environment

Note – In a Sun 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 *System Administration Guide: Basic Administration* (Solaris 9 or Solaris 10) for more information.

Perform this procedure on each node in the global cluster.

1   Become superuser on a cluster node.

2   Modify PATH **and** MANPATH **entries in the** .cshrc **or** .profile **file.**

   a.   Add /usr/sbin/ **and** /usr/cluster/bin/ **to the** PATH.

   b.   Add /usr/cluster/man/ **to the** MANPATH.

   See your 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 78.

Otherwise, configure Sun Cluster software on the cluster nodes. Go to "Establishing a New Global Cluster or New Global-Cluster Node" on page 82.

▼ 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.

Observe the following guidelines:

- NAT routing is not supported.
- The use of NAT for translation of local addresses is supported. NAT translation rewrites packets on-the-wire and is therefore transparent to the cluster software.

- Only stateless filtering is supported.

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

1   *(Solaris 10 11/06 only)* **Modify entries in the** /etc/iu.ap **file to ensure that the IP filter works in noncluster mode**

   a. **Modify the public NIC entries to list** clhbsndr pfil **as the module list.**

      The pfil must be the last module in the list.

      Note – If you have the same type of adapter for private and public network, your edits to the /etc/iu.ap file will push pfil to the private network streams. However, the cluster transport module will automatically remove all unwanted modules at stream creation, so pfil will be removed from the private network streams.

      ---

   b. **Add public-network interfaces to the** /etc/ipf/pfil.ap **file.**

      See Chapter 26, "Solaris IP Filter (Tasks)," in *System Administration Guide: IP Services* for more information.

   c. **Reboot all affected nodes.**

      You can boot the nodes in a rolling fashion.

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 Sun Cluster nodes.

   - *(Solaris 10 8/07 only)* 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. For example, suppose the following rules are currently used:

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

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

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

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

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

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

- Sun 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 non-existent 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 Sun Cluster software on the cluster nodes. Go to "Establishing a New Global Cluster or New Global-Cluster Node" on page 82.

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

The following procedures are in this chapter:

Establishing a New Global Cluster or New Global-Cluster Node

This section provides information and procedures to establish a new global cluster or to add a node to an existing cluster. Global-cluster nodes can be physical machines, (SPARC only) Sun Logical Domains (LDoms) I/O domains, or Sun LDoms 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 Solaris OS, Sun Cluster framework, and other products as described in "Installing the Software" on page 55.

Note – You can alternatively deploy the Sun Cluster Plug-in for Sun N1™ Service Provisioning System to create a multiple-node cluster or add a node to an existing cluster. Follow instructions in the documentation that is provided with the plug-in. You can also access this information at

`http://wikis.sun.com/display/SunCluster/Sun+Cluster+Framework+Plug-in`.

The following task map lists the tasks to perform. Complete the procedures in the order that is indicated.

TABLE 3–1   Task Map: Establish the Global Cluster

Method	Instructions
Use one of the following methods to establish a new global cluster or add a node to an existing global cluster:	
■ (*New clusters only*) Use the scinstall utility to establish the cluster.	"How to Configure Sun Cluster Software on All Nodes (scinstall)" on page 83
■ (*New clusters only*) Use an XML configuration file to establish the cluster.	"How to Configure Sun Cluster Software on All Nodes (XML)" on page 92
■ (*New clusters or added nodes*) Set up a JumpStart install server. Then create a flash archive of the installed system. Finally, use the scinstall JumpStart option to install the flash archive on each node and establish the cluster.	"How to Install Solaris and Sun Cluster Software (JumpStart)" on page 100
■ (*Added nodes only*) 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. Configure Sun Cluster software on a new node by using the scinstall utility or by using an XML configuration file.	"How to Prepare the Cluster for Additional Global-Cluster Nodes" on page 117 "How to Change the Private Network Configuration When Adding Nodes or Private Networks" on page 120 "How to Configure Sun Cluster Software on Additional Global-Cluster Nodes (scinstall)" on page 127 "How to Configure Sun Cluster Software on Additional Global-Cluster Nodes (XML)" on page 134

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

Method	Instructions
If you added a node to a cluster, update the quorum configuration information.	"How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 139
Assign quorum votes and remove the cluster from installation mode, if this operation was not already performed.	"How to Configure Quorum Devices" on page 142
Validate the quorum configuration.	"How to Verify the Quorum Configuration and Installation Mode" on page 147
(Optional) Change a node's private hostname.	"How to Change Private Hostnames" on page 148
Create or modify the NTP configuration file, if not already configured.	"How to Configure Network Time Protocol (NTP)" on page 150
(Optional) Configure IPsec to secure the private interconnect.	"How to Configure IP Security Architecture (IPsec) on the Cluster Interconnect" on page 153
If using a volume manager, install the volume management software.	Chapter 4, "Configuring Solaris Volume Manager Software," or Chapter 5, "Installing and Configuring Veritas Volume Manager"
Create cluster file systems or highly available local file systems as needed.	"How to Create Cluster File Systems" on page 205 or "Enabling Highly Available Local File Systems" in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*
(Optional) SPARC: Configure Sun Management Center to monitor the cluster.	"SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 227
Install third-party applications, register resource types, set up resource groups, and configure data services.	*Sun Cluster Data Services Planning and Administration Guide for Solaris OS* Documentation that is supplied with the application software
Take a baseline recording of the finished cluster configuration.	"How to Record Diagnostic Data of the Cluster Configuration" on page 154

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

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

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

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

Before You Begin Perform the following tasks:

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

  If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 64 for more information about installing Solaris software to meet Sun Cluster software requirements.

- SPARC: If you are configuring Sun Logical Domains (LDoms) I/O domains or guest domains as cluster nodes, ensure that Sun LDoms software is installed on each physical machine and that the domains meet Sun Cluster requirements. See "SPARC: How to Install Sun Logical Domains Software and Create Domains" on page 69.

- Ensure that Sun Cluster software packages and patches are installed on the node. See "How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73.

- Determine which mode of the scinstall utility you will use, Typical or Custom. For the Typical installation of Sun Cluster software, scinstall automatically specifies the following configuration defaults.

Component	Default Value
Private-network address	172.16.0.0
Private-network netmask	Solaris 9: 255.255.248.0 Solaris 10: 255.255.240.0
Cluster-transport adapters	Exactly two adapters
Cluster-transport switches	switch1 and switch2
Global fencing	Enabled
Global-devices file-system name	/globaldevices
Installation security (DES)	Limited

- Complete one of the following cluster configuration worksheets, depending on whether you run the `scinstall` utility in Typical mode or Custom mode.

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

Component	Description/Example	Answer	
Cluster Name	What is the name of the cluster that you want to establish?		
Cluster Nodes	List the names of the other cluster nodes planned for the initial cluster configuration. *(For a single-node cluster, press Control-D alone.)*		
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*
(VLAN adapters only)	Will this be a dedicated cluster transport adapter? *(Answer No if using tagged VLAN adapters.)*	Yes | No	Yes | No
	If no, what is the VLAN ID for this adapter?		
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 or a Network Appliance NAS device as a quorum device.)*	Yes | No	
Check	Do you want to interrupt cluster creation for `sccheck` 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

Component	Description/Example	Answer	
Network Address for the Cluster Transport *(multiple-node cluster only)*	Do you want to accept the default network address (`172.16.0.0`)?	Yes | No	
	If no, which private network address do you want to use?	___.___.___.___	
	Do you want to accept the default netmask?	Yes | No	
	If no, what are the maximum numbers of nodes and private networks that you expect to configure in the cluster? On the Solaris 10 OS, what is the maximum number of 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.*	___.___.___.___	
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*
(VLAN adapters only)	Will this be a dedicated cluster transport adapter? (*Answer No if using tagged VLAN adapters.*)	Yes | No	Yes | No
	If no, what is the VLAN ID for this adapter?		
	Where does each transport adapter connect to (*a switch or another adapter*)? Switch defaults: `switch1` and `switch2`	*First*	*Second*
	If a transport switch, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		
	Do you want to use autodiscovery to list the available adapters for the other nodes? If no, supply the following information for each additional node:	Yes | No	

Component	Description/Example	Answer	
Specify for each additional node	Node name:		
(multiple-node cluster only)		*First*	*Second*
	Transport adapter name:		
(VLAN adapters only)	Will this be a dedicated cluster transport adapter? *(Answer No if using tagged VLAN adapters.)*	Yes | No	Yes | No
	If no, what is the VLAN ID for this adapter?		
	Where does each transport adapter connect to (*a switch or another adapter*)? Defaults: `switch1` and `switch2`	*First*	*Second*
	If a transport switch, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		
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.*	Yes | No	Yes | No
Quorum Configuration *(two-node cluster only)*	Do you want to disable automatic quorum device selection? *(Answer Yes if any shared storage is not qualified to be a quorum device or if you want to configure a quorum server or a Network Appliance NAS device as a quorum device.)*	Yes | No	Yes | No
Global Devices File System *(specify for each node)*	Do you want to use the default name of the global-devices file system (`/globaldevices`)?	Yes | No	
	If no, do you want to use an already-existing file system?	Yes | No	
	What is the name of the file system that you want to use?		
Check *(multiple-node cluster only)*	Do you want to interrupt cluster creation for `sccheck` errors?	Yes | No	
(single-node cluster only)	Do you want to run the `sccheck` utility to validate the cluster?	Yes | No	
Automatic Reboot *(single-node cluster only)*	Do you want `scinstall` to automatically reboot the node after installation?	Yes | No	

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

- Interactive `scinstall` enables you to type ahead. Therefore, do not press the Return key more than once if the next menu screen does not appear immediately.

- Unless otherwise noted, you can press Control-D to return to either the start of a series of related questions or to the Main Menu.

- Default answers or answers to previous sessions are displayed in brackets ([ ]) at the end of a question. Press Return to enter the response that is in brackets without typing it.

1   If you disabled remote configuration during Sun Cluster software installation, re-enable remote configuration.

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

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

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

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

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

3   On the cluster node from which you intend to configure the cluster, become superuser.

4   Start the scinstall utility.

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

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

      * ?) Help with menu options
      * q) Quit

      Option:  1
```

The New Cluster and Cluster Node Menu is displayed.

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

The Typical or Custom Mode menu is displayed.

7   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.

8   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. Sun Cluster installation output is logged in a /var/cluster/logs/install/scinstall.log.*N* file.

9   For the Solaris 10 OS, 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 becomes 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   On one node, become superuser.

11   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   (Optional) Enable automatic node reboot if all monitored disk paths fail.

a. Enable the automatic reboot feature.

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

-p   Specifies the property to set

reboot_on_path_failure=enable   Specifies that the node will reboot if all monitored disk paths fail, provided that at least one of the disks is accessible from a different node in the cluster.

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

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

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

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

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

```
exclude:lofs
```

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

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

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

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

See "The Loopback File System" in *System Administration Guide: Devices and File Systems* (Solaris 9 or Solaris 10) for more information about loopback file systems.

Example 3–1 Configuring Sun 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 Typical mode. The other cluster node is phys-schost-2. The adapter names are qfe2 and qfe3. The automatic selection of a quorum device is enabled.

```
  Installation and Configuration
```

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

        Testing for "/globaldevices" on "phys-schost-1" ... done
        Testing for "/globaldevices" on "phys-schost-2" ... done
        Checking installation status ... done

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

        The following connections were discovered:

            phys-schost-1:qfe2  switch1  phys-schost-2:qfe2
            phys-schost-1:qfe3  switch2  phys-schost-2:qfe3

        Completed discovery of the cluster transport configuration.

        Started sccheck on "phys-schost-1".
        Started sccheck on "phys-schost-2".

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

        Removing the downloaded files ... done

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

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

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

Rebooting ...
```

Troubleshooting   **Unsuccessful configuration** – If one or more nodes cannot join the cluster, or if the wrong configuration information was specified, first attempt to rerun this procedure. If that does not correct the problem, perform the procedure "How to Unconfigure Sun Cluster Software to Correct Installation Problems" on page 233 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Sun Cluster software packages. Then rerun this procedure.

Next Steps   ■ If you installed a single-node cluster, cluster establishment is complete. Go to "Creating Cluster File Systems" on page 205 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 147.

- If you installed a multiple-node cluster and declined automatic quorum configuration, perform postinstallation setup. Go to "How to Configure Quorum Devices" on page 142.

▼ How to Configure Sun 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 Sun Cluster 3.2 1/09 software.

This procedure configures the following cluster components:

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

Before You Begin   Perform the following tasks:

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

  If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 64 for more information about installing Solaris software to meet Sun Cluster software requirements.

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

  If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 64 for more information about installing Solaris software to meet Sun Cluster software requirements.

- SPARC: If you are configuring Sun Logical Domains (LDoms) I/O domains or guest domains as cluster nodes, ensure that Sun LDoms software is installed on each physical machine and that the domains meet Sun Cluster requirements. See "SPARC: How to Install Sun Logical Domains Software and Create Domains" on page 69.

- Ensure that Sun Cluster 3.2 1/09 software and patches are installed on each node that you will configure. See "How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73.

1 Ensure that Sun Cluster 3.2 1/09 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 Sun 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 Sun 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 Sun Cluster software is already configured on the node.

      If the cluster is running an older version of Sun Cluster software and you want to install Sun Cluster 3.2 1/09 software, instead perform upgrade procedures in *Sun Cluster Upgrade Guide for Solaris OS*.

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

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

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

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

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

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

3 If you are duplicating an existing cluster than runs Sun Cluster 3.2 1/09 software, use a node in that cluster to create a cluster configuration XML file.

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

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

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

-o   Specifies the output destination.

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

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

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

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

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

5   **Modify the cluster configuration XML file as needed.**

a. Open your cluster configuration XML file for editing.

- **If you are duplicating an existing cluster, open the file that you created with the** cluster export **command.**

- **If you are not duplicating an existing cluster, create a new file.**

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

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

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

- The cluster is created with the assumption that the partition /globaldevices exists on each node that you configure as a cluster node. The global-devices namespace is created on this partition. If you need to use a different file-system name on which to create the global devices, add the following property to the <propertyList> element for each node that does not have a partition that is named /globaldevices.

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

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

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

6 Validate the cluster configuration XML file.

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

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

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

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

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

8 For the Solaris 10 OS, 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 becomes online before you proceed to the next step.

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

9 On one node, become superuser.

10 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   Install any necessary patches to support Sun Cluster software, if you have not already done so.
See "Patches and Required Firmware Levels" in *Sun Cluster Release Notes* for the location of patches and installation instructions.

12   If you intend to use Sun Cluster HA for NFS on a highly available local file system, ensure that the loopback file system (LOFS) is disabled.
To disable LOFS, add the following entry to the /etc/system file on each node of the cluster.

```
exclude:lofs
```

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

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

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

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

See "The Loopback File System" in *System Administration Guide: Devices and File Systems* (Solaris 9 or Solaris 10) for more information about loopback file systems.

13 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 142.

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

Follow instructions in "How to Install and Configure Quorum Server Software" on page 57.

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 *Sun Cluster 3.1 - 3.2 With Network-Attached Storage Devices Manual for Solaris OS*.

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

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

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

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

e. Configure the quorum device.

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

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

14 Remove the cluster from installation mode.

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

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

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

16 (Optional) Enable automatic node reboot if all monitored disk paths fail.

a. Enable the automatic reboot feature.

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

-p   Specifies the property to set

```
                          reboot_on_path_failure=enable        Specifies that the node will reboot if all monitored
                                                               disk paths fail, provided that at least one of the disks
                                                               is accessible from a different node in the cluster.
```

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

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

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

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

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

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

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

Troubleshooting   **Unsuccessful configuration** – If one or more nodes cannot join the cluster, or if the wrong configuration information was specified, first attempt to rerun this procedure. If that does not correct the problem, perform the procedure "How to Unconfigure Sun Cluster Software to Correct Installation Problems" on page 233 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Sun Cluster software packages. Then rerun this procedure.

Next Steps   Go to "How to Verify the Quorum Configuration and Installation Mode" on page 147.

See Also   After the cluster is fully established, you can duplicate the configuration of the other cluster components from the existing cluster. If you did not already do so, modify the values of the XML elements that you want to duplicate to reflect the cluster configuration you are adding the component to. For example, if you are duplicating resource groups, ensure that the <resourcegroupNodeList> entry contains the valid node names for the new cluster, and not the node names from the cluster that you duplicated unless the node names are the same.

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

Note – This table provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix B, "Sun Cluster Object-Oriented Commands," in *Sun Cluster System Administration Guide for Solaris OS*.

Cluster Component	Man Page	Special Instructions
Device groups: Solaris Volume Manager and Veritas Volume Manager	cldevicegroup(1CL)	For Solaris Volume Manager, first create the disk sets that you specify in the cluster configuration XML file. For VxVM, first install and configure VxVM software and create the disk groups that you specify in the cluster configuration XML file.

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

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

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

Before You Begin Perform the following tasks:

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

- Determine the Ethernet address of each cluster node.

- If you use a naming service, ensure that the following information is added to any naming services that clients use to access cluster services. See "Public-Network IP Addresses" on page 25 for planning guidelines. See your Solaris system-administrator documentation for information about using Solaris naming services.

  - Address-to-name mappings for all public hostnames and logical addresses
  - The IP address and hostname of the JumpStart install server

- Ensure that your cluster configuration planning is complete. See "How to Prepare for Cluster Software Installation" on page 56 for requirements and guidelines.

- On the server from which you will create the flash archive, ensure that all Solaris OS software, patches, and firmware that is necessary to support Sun Cluster software is installed.

  If Solaris software is already installed on the server, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 64 for more information about installing Solaris software to meet Sun Cluster software requirements.

- SPARC: If you are configuring Sun Logical Domains (LDoms) I/O domains or guest domains as cluster nodes, ensure that Sun LDoms software is installed on each physical machine and that the domains meet Sun Cluster requirements. See "SPARC: How to Install Sun Logical Domains Software and Create Domains" on page 69.

- Ensure that Sun Cluster software packages and patches are installed on the server from which you will create the flash archive. See "How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73.

- Determine which mode of the `scinstall` utility you will use, Typical or Custom. For the Typical installation of Sun Cluster software, `scinstall` automatically specifies the following configuration defaults.

Component	Default Value
Private-network address	`172.16.0.0`
Private-network netmask	Solaris 9: `255.255.248.0` Solaris 10: `255.255.240.0`
Cluster-transport adapters	Exactly two adapters
Cluster-transport switches	`switch1` and `switch2`
Global fencing	Enabled
Global-devices file-system name	`/globaldevices`
Installation security (DES)	Limited

- Complete one of the following cluster configuration worksheets, depending on whether you run the `scinstall` utility in Typical mode or Custom mode. See "Planning the Sun Cluster Environment" on page 25 for planning guidelines.

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

Component	Description/Example	Answer
JumpStart Directory	What is the name of the JumpStart directory to use?	

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

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

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

Component	Description/Example	Answer
JumpStart Directory	What is the name of the JumpStart directory to use?	
Cluster Name	What is the name of the cluster that you want to establish?	
Cluster Nodes	List the names of the cluster nodes that are planned for the initial cluster configuration. *(For a single-node cluster, press Control-D alone.)*	
Authenticating Requests to Add Nodes *(multiple-node cluster only)*	Do you need to use DES authentication?	No | Yes

Component	Description/Example	Answer	
Network Address for the Cluster Transport *(multiple-node cluster only)*	Do you want to accept the default network address (`172.16.0.0`)?	Yes | No	
	If no, which private network address do you want to use?	___.___.___.___	
	Do you want to accept the default netmask?	Yes | No	
	If no, what are the maximum numbers of nodes and private networks that you expect to configure in the cluster? On the Solaris 10 OS, what is the maximum number of zone clusters that you expect to configure in the cluster?	_____ nodes _____ networks _____ zone clusters	
	Which netmask do you want to use? *Choose from the values that are calculated by scinstall or supply your own.*	___.___.___.___	
Minimum Number of Private Networks *(multiple-node cluster only)*	Should this cluster use at least two private networks?	Yes | No	
Point-to-Point Cables *(two-node cluster only)*	Does this cluster use switches?	Yes | No	
Cluster Switches *(multiple-node cluster only)*	Transport switch name, if used: Defaults: `switch1` and `switch2`	*First*	*Second*
Cluster Transport Adapters and Cables *(multiple-node cluster only)*	First node name:		
	Transport adapter name:	*First*	*Second*
(VLAN adapters only)	Will this be a dedicated cluster transport adapter? *(Answer No if using tagged VLAN adapters.)*	Yes | No	Yes | No
	If no, what is the VLAN ID for this adapter?		
	Where does each transport adapter connect to (*a switch or another adapter*)? Switch defaults: `switch1` and `switch2`		
	If a transport switch, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		

Component	Description/Example	Answer	
Specify for each additional node *(multiple-node cluster only)*	Node name:		
	Transport adapter name:	*First*	*Second*
	Where does each transport adapter connect to (*a switch or another adapter*)? Switch defaults: `switch1` and `switch2`		
	If a transport switch, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		
Global Devices File System *Specify for each node*	Do you want to use the default name of the global-devices file system (`/globaldevices`)?	Yes | No	
	If no, do you want to use an already-existing file system?	Yes | No	
	If no, do you want to create a new file system on an unused partition?	Yes | No	
	What is the name of the file system?		
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.*	Yes | No	Yes | No
Quorum Configuration *(two-node cluster only)*	Do you want to disable automatic quorum device selection? (*Answer Yes if any shared storage is not qualified to be a quorum device or if you want to configure a quorum server or a Network Appliance NAS device as a quorum device.*)	Yes | No	Yes | No

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

- Interactive `scinstall` enables you to type ahead. Therefore, do not press the Return key more than once if the next menu screen does not appear immediately.

- Unless otherwise noted, you can press Control-D to return to either the start of a series of related questions or to the Main Menu.

- Default answers or answers to previous sessions are displayed in brackets ([ ]) at the end of a question. Press Return to enter the response that is in brackets without typing it.

1   Set up your JumpStart install server.

Ensure that the JumpStart install server meets the following requirements.

- The install server is on the same subnet as the cluster nodes, or on the Solaris boot server for the subnet that the cluster nodes use.

- The install server is not itself a cluster node.

- The install server installs a release of the Solaris OS that is supported by the Sun Cluster software.

- A custom JumpStart directory exists for JumpStart installation of Sun Cluster software. This *jumpstart-dir* directory must meet the following requirements:

  - Contain a copy of the check utility.
  - Be NFS exported for reading by the JumpStart install server.

- Each new cluster node is configured as a custom JumpStart installation client that uses the custom JumpStart directory that you set up for Sun Cluster installation.

Follow the appropriate instructions for your software platform and OS version to set up the JumpStart install server. See "Creating a Profile Server for Networked Systems" in *Solaris 9 9/04 Installation Guide* or "Creating a Profile Server for Networked Systems" in *Solaris 10 5/08 Installation Guide: Custom JumpStart and Advanced Installations*.

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

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

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

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

For more information, see "How to Add a Node to the Authorized Node List" in *Sun Cluster System Administration Guide for Solaris OS*.

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

If Solaris software is already installed on the server, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 64 for more information about installing Solaris software to meet Sun Cluster software requirements.

Follow procedures in "How to Install Solaris Software" on page 64.

4   (Optional) SPARC: On the installed system, install Sun Logical Domains (LDoms) software and create domains, if you have not already done so.

Follow the procedures in "SPARC: How to Install Sun Logical Domains Software and Create Domains" on page 69.

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

Follow procedures in "How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73.

See "Patches and Required Firmware Levels" in *Sun Cluster Release Notes* for the location of patches and installation instructions.

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

```
machine# cacaoadm enable
```

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

Perform this step regardless of whether you are using a naming service. See "Public-Network IP Addresses" on page 25 for a listing of Sun Cluster components whose IP addresses you must add.

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

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

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

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

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

9   Create the flash archive of the installed system.

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

-n name   Name to give the flash archive.

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

Follow procedures in one of the following manuals:

- Chapter 21, "Creating Solaris Flash Archives (Tasks)," in *Solaris 9 9/04 Installation Guide*
- Chapter 3, "Creating Solaris Flash Archives (Tasks)," in *Solaris 10 5/08 Installation Guide: Solaris Flash Archives (Creation and Installation)*

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

See "Managing Network File Systems (Overview)," in *System Administration Guide: Network Services* (Solaris 9 or Solaris 10) for more information about automatic file sharing.

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

11   On the JumpStart install server, become superuser.

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

In the media path, replace *arch* with sparc or x86 (Solaris 10 only) and replace *ver* with 9 for Solaris 9 or 10 for Solaris 10.

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

installserver# ./scinstall
```

The scinstall Main Menu is displayed.

13   Type the option number for Configure a Cluster to be JumpStarted From This Install Server and press the Return key.

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

```
 *** Main Menu ***

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

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

      * ?) Help with menu options
      * q) Quit

    Option:  2
```

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

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

```
install_type     initial_install
system_type      standalone
partitioning     explicit
filesys          rootdisk.s0 free /
filesys          rootdisk.s1 750  swap
filesys          rootdisk.s3 512  /globaldevices
filesys          rootdisk.s7 20
cluster          SUNWCuser        add
package          SUNWman          add
```

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

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

For example, if you assigned slice 4 for the global-devices file system and specified to `scinstall` that the file-system name is /gdevs, you would change the /globaldevices entry of the `autoscinstall.class` file to the following:

```
filesys         rootdisk.s4 512  /gdevs
```

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

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

See "`archive_location` Keyword" in *Solaris 9 9/04 Installation Guide* or *Solaris 10 Installation Guide: Custom JumpStart and Advanced Installations* for information about valid values for *retrieval_type* and *location* when used with the `archive_location` keyword.

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

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

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

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

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

The following table lists Solaris packages that are required to support some Sun Cluster functionality. These packages are not included in the End User Solaris Software Group. See "Solaris Software Group Considerations" on page 18 for more information.

Feature	Mandatory Solaris Software Packages
RSMAPI, RSMRDT drivers, or SCI-PCI adapters (SPARC based clusters only)	SPARC: *Solaris 9*: SUNWrsm SUNWrsmx SUNWrsmo SUNWrsmox *Solaris 10*: SUNWrsm SUNWrsmo
Sun Cluster Manager (formerly SunPlex™ Manager)	SUNWapchr SUNWapchu

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

- Edit the `autoscinstall.class` file directly. These changes are applied to all nodes in all clusters that use this custom JumpStart directory.

- Update the `rules` file to point to other profiles, then run the `check` utility to validate the `rules` file.

As long as the Solaris OS installation profile meets minimum Sun Cluster file-system allocation requirements, Sun Cluster software places no restrictions on other changes to the installation profile. See "System Disk Partitions" on page 19 for partitioning guidelines and requirements to support Sun Cluster software.

For more information about JumpStart profiles, see Chapter 26, "Preparing Custom JumpStart Installations (Tasks)," in *Solaris 9 9/04 Installation Guide* or Chapter 3, "Preparing Custom JumpStart Installations (Tasks)," in *Solaris 10 5/08 Installation Guide: Custom JumpStart and Advanced Installations*.

16 **To install required packages for any of the following features or to perform other postinstallation tasks, set up your own finish script.**

- Remote Shared Memory Application Programming Interface (RSMAPI)
- SCI-PCI adapters for the interconnect transport
- RSMRDT drivers

Note – Use of the RSMRDT driver is restricted to clusters that run an Oracle9i release 2 SCI configuration with RSM enabled. Refer to Oracle9i release 2 user documentation for detailed installation and configuration instructions.

Your own finish script runs after the standard finish script that is installed by the `scinstall` command. See Preparing Custom JumpStart Installations in Chapter 26, "Preparing Custom JumpStart Installations (Tasks)," in *Solaris 9 9/04 Installation Guide* or Chapter 3, "Preparing Custom JumpStart Installations (Tasks)," in *Solaris 10 5/08 Installation Guide: Custom JumpStart and Advanced Installations* for information about creating a JumpStart finish script.

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

See Step 15.

b. Name your finish script `finish`.

c. Modify the `finish` script to install the software packages listed in the following table that support the features that you intend to use.

Feature	Additional Sun Cluster 3.2 1/09 Packages to Install
RSMAPI	`SUNWscrif`
SCI-PCI adapters	■ *Solaris 9*: `SUNWsci SUNWscid SUNWscidx` ■ *Solaris 10*: `SUNWscir SUNWsci SUNWscidr SUNWscid`
RSMRDT drivers	`SUNWscrdt`

- Install the packages in the order that is used in the table.

- Install the packages from the `/cdrom/suncluster_3_0Packages/` directory, where *arch* is `sparc` or `x86` (Solaris 10 only), and where *ver* is `9` for Solaris 9 or `10` for Solaris 10 .

   d. Make any additional modifications for other postinstallation tasks that you want the `finish` **script to perform.**

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

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

17 **Exit from the JumpStart install server.**

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

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

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

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

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

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

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

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

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

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

20   **Shut down each node.**

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

21   **Boot each node to start the JumpStart installation.**

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

```
ok boot net - install
```

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

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

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

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

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

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

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

   The list of boot devices is displayed.

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

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

   e. **Save your change and exit the BIOS.**

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

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

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

```
GNU GRUB version 0.95 (631K lower / 2095488K upper memory)
+-------------------------------------------------------------------+
| Solaris_10 Jumpstart                                              |
|                                                                   |
|                                                                   |
+-------------------------------------------------------------------+
Use the ^ and v keys to select which entry is highlighted.
Press enter to boot the selected OS, 'e' to edit the
commands before booting, or 'c' for a command-line.
```

After further processing, the installation type menu is displayed.

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

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

```
      Select the type of installation you want to perform:

          1 Solaris Interactive
          2 Custom JumpStart
          3 Solaris Interactive Text (Desktop session)
          4 Solaris Interactive Text (Console session)
          5 Apply driver updates
          6 Single user shell

          Enter the number of your choice.
2
```

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

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

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

After further processing, the BIOS Setup Utility is displayed.

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

The list of boot devices is displayed.

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

k. Save your change and exit the BIOS.

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

22 For the Solaris 10 OS, 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 becomes online before you proceed to the next step.

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

23 If you are installing a new node to an existing cluster, create mount points on the new node for all existing cluster file systems.

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

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

b. On the node that you added to the cluster, create a mount point for each cluster file system in the cluster.

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

For example, if a file-system name that is returned by the mount command is /global/dg-schost-1, run mkdir -p /global/dg-schost-1 on the node that is being added to the cluster.

Note – The mount points become active after you reboot the cluster in Step 27.

---

c. If Veritas Volume Manager (VxVM) is installed on any nodes that are already in the cluster, view the vxio **number on each VxVM–installed node.**

```
phys-schost# grep vxio /etc/name_to_major
vxio NNN
```

- Ensure that the same vxio number is used on each of the VxVM-installed nodes.
- Ensure that the vxio number is available for use on each of the nodes that do not have VxVM installed.
- If the vxio number is already in use on a node that does not have VxVM installed, free the number on that node. Change the /etc/name_to_major entry to use a different number.

24 **(Optional) To use dynamic reconfiguration on Sun Enterprise 10000 servers, add the following entry to the** `/etc/system` **file on each node in the cluster.**

```
set kernel_cage_enable=1
```

This entry becomes effective after the next system reboot. See the *Sun Cluster System Administration Guide for Solaris OS* for procedures to perform dynamic reconfiguration tasks in a Sun Cluster configuration. See your server documentation for more information about dynamic reconfiguration.

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

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

```
exclude:lofs
```

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

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

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

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

See "The Loopback File System" in *System Administration Guide: Devices and File Systems* (Solaris 9 or Solaris 10) for more information about loopback file systems.

26 **x86: Set the default boot file.**

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

- **On the Solaris 9 OS, set the default to** kadb**.**

  ```
  phys-schost# eeprom boot-file=kadb
  ```

- **On the Solaris 10 OS, set the default to** kmdb **in the GRUB boot parameters menu.**

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

27 If you performed a task that requires a cluster reboot, follow these steps to reboot the cluster.

The following are some of the tasks that require a reboot:

- Adding a new node to an existing cluster
- Installing patches that require a node or cluster reboot
- Making configuration changes that require a reboot to become active

a. On one node, become superuser.

b. Shut down the cluster.

```
phys-schost-1# cluster shutdown -y -g0 clustername
```

Note – Do not reboot the first-installed node of the cluster until *after* the cluster is shut down. Until cluster installation mode is disabled, only the first-installed node, which established the cluster, has a quorum vote. In an established cluster that is still in installation mode, if the cluster is not shut down before the first-installed node is rebooted, the remaining cluster nodes cannot obtain quorum. The entire cluster then shuts down.

Cluster nodes remain in installation mode until the first time that you run the clsetup command. You run this command during the procedure "How to Configure Quorum Devices" on page 142.

c. Reboot each node in the cluster.

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

  ```
  ok boot
  ```

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

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

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

  For more information about GRUB based booting, see Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration*.

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

28 (Optional) If you did not perform Step 27 to reboot the nodes, start the Sun Java Web Console web server manually on each node.

```
phys-schost# smcwebserver start
```

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

29 On one node, become superuser.

30 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.

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

a. Enable the automatic reboot feature.

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

-p	Specifies the property to set
reboot_on_path_failure=enable	Specifies that the node will reboot if all monitored disk paths fail, provided that at least one of the disks is accessible from a different node in the cluster.

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

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

Node Name:                                        node
...
```

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

Next Steps If you added a node to a two-node cluster, go to "How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 139.

Otherwise, go to the next appropriate procedure:

- If you installed a multiple-node cluster and chose automatic quorum configuration, postinstallation setup is complete. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 147.

- If you installed a multiple-node cluster and declined automatic quorum configuration, perform postinstallation setup. Go to "How to Configure Quorum Devices" on page 142.

- If you added a new node to an existing cluster that uses a quorum device, go to "How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 139.

- If you added a new node to an existing cluster that does not use a quorum device, verify the state of the cluster. Go to "How to Verify the Quorum Configuration and Installation Mode" on page 147.

- If you installed a single-node cluster, cluster establishment is complete. Go to "Creating Cluster File Systems" on page 205 to install volume management software and configure the cluster.

Troubleshooting **Disabled scinstall option** – If the JumpStart option of the scinstall command does not have an asterisk in front, the option is disabled. This condition indicates that JumpStart setup is not complete or that the setup has an error. To correct this condition, first quit the scinstall utility. Repeat Step 1 through Step 16 to correct JumpStart setup, then restart the scinstall utility.

Error messages about nonexistent nodes – Unless you have installed your own /etc/inet/ntp.conf file, the scinstall command installs a default ntp.conf file for you. The default file is shipped with references to the maximum number of nodes. Therefore, the xntpd(1M) daemon might issue error messages regarding some of these references at boot time. You can safely ignore these messages. See "How to Configure Network Time Protocol (NTP)" on page 150 for information about how to suppress these messages under otherwise normal cluster conditions.

▼ 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 *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*.
    - Verify that any existing cluster interconnects can support the new node. See the *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS*.
    - Ensure that any additional storage is installed. See the appropriate manual from the Sun Cluster 3.1 - 3.2 Hardware Administration Collection.

1 If you use the Cluster Control Panel (CCP), update the configuration files on the administrative console.

a. Add to the cluster's entry in the `/etc/clusters` **file the name of the node that you are adding.**

b. Add to the `/etc/serialports` **files an entry with the new node name, the host name of the node's console-access device, and the port number.**

2 Add the name of the new node to the cluster's authorized-nodes list.

a. On any node, become superuser.

b. Start the `clsetup` **utility.**

```
phys-schost# clsetup
```

The Main Menu is displayed.

c. Choose the menu item, New nodes.

d. Choose the menu item, Specify the name of a machine which may add itself.

e. Follow the prompts to add the node's name to the list of recognized machines.

The `clsetup` utility displays the message Command completed successfully if the task is completed without error.

f. Quit the `clsetup` **utility.**

3 If you are adding a node to a single-node cluster, ensure that two cluster interconnects already exist by displaying the interconnect configuration.

```
phys-schost# clinterconnect show
```

You must have at least two cables or two adapters configured before you can add a node.

- **If the output shows configuration information for two cables or for two adapters, proceed to Step 4.**

- **If the output shows no configuration information for either cables or adapters, or shows configuration information for only one cable or adapter, configure new cluster interconnects.**

    a. **On one node, start the** `clsetup` **utility.**

    ```
    phys-schost# clsetup
    ```

    b. **Choose the menu item, Cluster interconnect.**

    c. **Choose the menu item, Add a transport cable.**

    Follow the instructions to specify the name of the node to add to the cluster, the name of a transport adapter, and whether to use a transport switch.

    d. **If necessary, repeat Step c to configure a second cluster interconnect.**

    e. **When finished, quit the** `clsetup` **utility.**

    f. **Verify that the cluster now has two cluster interconnects configured.**

    ```
    phys-schost# clinterconnect show
    ```

    The command output should show configuration information for at least two cluster interconnects.

4 **Ensure that the private-network configuration can support the nodes and private networks that you are adding.**

a. **Display the maximum numbers of nodes and private networks, and zone clusters on the Solaris 10 OS, that the current private-network configuration supports.**

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

The output looks similar to the following, which shows the default values on the Solaris 10 OS:

```
=== 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 Sun Cluster Software on Additional Global-Cluster Nodes (`scinstall`)" on page 127.

- **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 120. 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 Sun Cluster software on the new cluster nodes. Go to "How to Configure Sun Cluster Software on Additional Global-Cluster Nodes (`scinstall`)" on page 127 or "How to Configure Sun Cluster Software on Additional Global-Cluster Nodes (XML)" on page 134.

▼ 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
- On the Solaris 10 OS, 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. On the Solaris 10 OS, 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
```

Before You Begin   Ensure that remote shell (rsh(1M)) or secure shell (ssh(1)) access for superuser is enabled for all cluster nodes.

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

2   **From one node, start the** clsetup **utility.**

     # **clsetup**

     The clsetup Main Menu is displayed.

3   **Switch each resource group offline.**

     If the node contains non-global zones, any resource groups in the zones are also switched offline.

     a.   **Type the number that corresponds to the option for Resource groups and press the Return key.**

        The Resource Group Menu is displayed.

     b.   **Type the number that corresponds to the option for Online/Offline or Switchover a resource group and press the Return key.**

     c.   **Follow the prompts to take offline all resource groups and to put them in the unmanaged state.**

     d.   **When all resource groups are offline, type q to return to the Resource Group Menu.**

4   **Disable all resources in the cluster.**

     a.   **Type the number that corresponds to the option for Enable/Disable a resource and press the Return key.**

     b.   **Choose a resource to disable and follow the prompts.**

     c.   **Repeat the previous step for each resource to disable.**

     d.   **When all resources are disabled, type q to return to the Resource Group Menu.**

5   **Quit the** clsetup **utility.**

6   **Verify that all resources on all nodes are** Offline **and that all resource groups are in the** Unmanaged **state.**

     # **cluster status -t resource,resourcegroup**

     -t   Limits output to the specified cluster object

resource   Specifies resources

resourcegroup   Specifies resource groups

7 From one node, shut down the cluster.

    `# cluster shutdown -g0 -y`

    -g   Specifies the wait time in seconds

    -y   Prevents the prompt that asks you to confirm a shutdown from being issued

8 Boot each node into noncluster mode.

■   **On SPARC based systems, perform the following command:**

    ok `boot -x`

■   **On x86 based systems, perform the following commands:**

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

    The GRUB menu appears similar to the following:

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

    For more information about GRUB based booting, see Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration*.

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

    The GRUB boot parameters screen appears similar to the following:

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

```
after ('O' for before) the selected line, 'd' to remove the
selected line, or escape to go back to the main menu.
```

c. Add -x to the command to specify that the system boot into noncluster mode.

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

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

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

The screen displays the edited command.

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

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

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

---

9   From one node, start the clsetup **utility.**

When run in noncluster mode, the clsetup utility displays the Main Menu for noncluster-mode operations.

10   Type the number that corresponds to the option for Change IP Address Range and press the Return key.

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

11 To change either the private-network IP address or the IP address range, type yes **and press the Return key.**

The clsetup utility displays the default private-network IP address, 172.16.0.0, and asks if it is okay to accept this default.

12 Change or accept the private-network IP address.

- **To accept the default private-network IP address and proceed to changing the IP address range, type** yes **and press the Return key.**

  The clsetup utility will ask if it is okay to accept the default netmask. Skip to the next step to enter your response.

- **To change the default private-network IP address, perform the following substeps.**

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

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

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

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

13 Change or accept the default private-network IP address range.

On the Solaris 9 OS, the default netmask is 255.255.248.0. This default IP address range supports up to 64 nodes and up to 10 private networks in the cluster. On the Solaris 10 OS, the default netmask is 255.255.240.0. This default IP address range supports up to 64 nodes, up to 12 zone clusters, and up to 10 private networks in the cluster.

- **To accept the default IP address range, type** yes **and press the Return key.**

  Then skip to the next step.

- **To change the IP address range, perform the following substeps.**

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

  When you decline the default netmask, the clsetup utility prompts you for the number of nodes and private networks, and zone clusters on the Solaris 10 OS, that you expect to configure in the cluster.

    b. Enter the number of nodes and private networks, and zone clusters on the Solaris 10 OS, 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 and private networks, and zone clusters on the Solaris 10 OS, that you specified.
- The second netmask supports twice the number of nodes and private networks, and zone clusters on the Solaris 10 OS, 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 and private networks, and zone clusters on the Solaris 10 OS.

14 **Type** yes **in response to the** clsetup **utility's question about proceeding with the update.**

15 **When finished, exit the** clsetup **utility.**

16 **Reboot each node back into the cluster.**

    a. Shut down each node.

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

    b. Boot each node into cluster mode.

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

```
ok boot
```

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

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

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

For more information about GRUB based booting, see Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration*.

17   From one node, start the `clsetup` **utility.**

```
# clsetup
```

The `clsetup` Main Menu is displayed.

18   Re-enable all disabled resources.

    a.   Type the number that corresponds to the option for Resource groups and press the Return key.

      The Resource Group Menu is displayed.

    b.   Type the number that corresponds to the option for Enable/Disable a resource and press the Return key.

    c.   Choose a resource to enable and follow the prompts.

    d.   Repeat for each disabled resource.

    e.   When all resources are re-enabled, type q to return to the Resource Group Menu.

19   Bring each resource group back online.

If the node contains non-global zones, also bring online any resource groups that are in those zones.

    a.   Type the number that corresponds to the option for Online/Offline or Switchover a resource group and press the Return key.

    b.   Follow the prompts to put each resource group into the managed state and then bring the resource group online.

20   When all resource groups are back online, exit the `clsetup` **utility.**

Type **q** to back out of each submenu, or press Ctrl-C.

Next Steps   To add a node to an existing cluster, go to one of the following procedures:

- "How to Configure Sun Cluster Software on Additional Global-Cluster Nodes (`scinstall`)" on page 127
- "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 100
- "How to Configure Sun Cluster Software on Additional Global-Cluster Nodes (XML)" on page 134

To create a non-global zone on a cluster node, go to "Configuring a Non-Global Zone on a Global-Cluster Node" on page 209.

▼ How to Configure Sun Cluster Software on Additional Global-Cluster Nodes (`scinstall`)

Perform this procedure to add a new node to an existing global cluster. To use JumpStart to add a new node, instead follow procedures in "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 100.

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

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

Before You Begin   Perform the following tasks:

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

  If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 64 for more information about installing Solaris software to meet Sun Cluster software requirements.

- SPARC: If you are configuring Sun Logical Domains (LDoms) I/O domains or guest domains as cluster nodes, ensure that Sun LDoms software is installed on each physical machine and that the domains meet Sun Cluster requirements. See "SPARC: How to Install Sun Logical Domains Software and Create Domains" on page 69.

- Ensure that Sun Cluster software packages and patches are installed on the node. See "How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73.

- 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 117.

- Determine which mode of the `scinstall` utility you will use, Typical or Custom. For the Typical installation of Sun Cluster software, `scinstall` automatically specifies the following configuration defaults.

Component	Default Value
Cluster-transport switches	`switch1` and `switch2`

Component	Default Value
Global-devices file-system name	`/globaldevices`

- Complete one of the following configuration planning worksheets. See "Planning the Solaris OS" on page 16 and "Planning the Sun Cluster Environment" on page 25 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 `sccheck` validation utility?	Yes | No	
Autodiscovery of Cluster Transport	Do you want to use autodiscovery to configure the cluster transport? If no, supply the following additional information:	Yes | No	
Point-to-Point Cables	Does the node that you are adding to the cluster make this a two-node cluster?	Yes | No	
	Does the cluster use switches?	Yes | No	
Cluster Switches	If used, what are the names of the two switches? Defaults: `switch1` and `switch2`	*First*	*Second*
Cluster Transport Adapters and Cables	 Transport adapter names:	*First*	*Second*
	Where does each transport adapter connect to (*a switch or another adapter*)? Switch defaults: `switch1` and `switch2`		
	For transport switches, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		
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 `sccheck` validation utility?	Yes | No	
Autodiscovery of Cluster Transport	Do you want to use autodiscovery to configure the cluster transport? If no, supply the following additional information:	Yes | No	
Point-to-Point Cables	Does the node that you are adding to the cluster make this a two-node cluster?	Yes | No	
	Does the cluster use switches?	Yes | No	
Cluster Switches	Transport switch name, if used: Defaults: `switch1` and `switch2`	*First*	*Second*
Cluster Transport Adapters and Cables	Transport adapter name:	*First*	*Second*
	Where does each transport adapter connect to (*a switch or another adapter*)? Switch defaults: `switch1` and `switch2`		
	If a transport switch, do you want to use the default port name?	Yes | No	Yes | No
	If no, what is the name of the port that you want to use?		
Global Devices File System	What is the name of the global-devices file system? Default: `/globaldevices`		
Automatic Reboot	Do you want `scinstall` to automatically reboot the node after installation?	Yes | No	

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

- Interactive `scinstall` enables you to type ahead. Therefore, do not press the Return key more than once if the next menu screen does not appear immediately.

- Unless otherwise noted, you can press Control-D to return to either the start of a series of related questions or to the Main Menu.

- Default answers or answers to previous sessions are displayed in brackets ([ ]) at the end of a question. Press Return to enter the response that is in brackets without typing it.

1   On the cluster node to configure, become superuser.

2   Start the `scinstall` utility.

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

The `scinstall` Main Menu is displayed.

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

```
   *** Main Menu ***

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

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

    * ?) Help with menu options
    * q) Quit

    Option:  1
```

The New Cluster and Cluster Node Menu is displayed.

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

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

The scinstall utility configures the node and boots the node into the cluster.

6   Unload the Sun Java Availability Suite DVD-ROM from the DVD-ROM drive.

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

   b.   Eject the DVD-ROM.

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

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

8   For the Solaris 10 OS, 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 becomes online before you proceed to the next step.

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

9   From an active cluster member, prevent any other nodes from joining the cluster.

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

Alternately, you can use the clsetup utility. See "How to Add a Node to the Authorized Node List" in *Sun Cluster System Administration Guide for Solaris OS* for procedures.

10   On one node, become superuser.

11   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   Verify that all necessary patches are installed.

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

13   (Optional) Enable automatic node reboot if all monitored disk paths fail.

a.   Enable the automatic reboot feature.

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

-p	Specifies the property to set
reboot_on_path_failure=enable	Specifies that the node will reboot if all monitored disk paths fail, provided that at least one of the disks is accessible from a different node in the cluster.

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 intend to use Sun Cluster HA for NFS on a highly available local file system, ensure that the loopback file system (LOFS) is disabled.**

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

```
exclude:lofs
```

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

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

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

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

See "The Loopback File System" in *System Administration Guide: Devices and File Systems* (Solaris 9 or Solaris 10) for more information about loopback file systems.

Example 3–3 Configuring Sun Cluster Software on an Additional Node

The following example shows the node `phys-schost-3` added to the cluster `schost`. The sponsoring node is `phys-schost-1`.

```
*** Adding a Node to an Existing Cluster ***
Fri Feb  4 10:17:53 PST 2005


scinstall -ik -C schost -N phys-schost-1 -A trtype=dlpi,name=qfe2 -A trtype=dlpi,name=qfe3
-m endpoint=:qfe2,endpoint=switch1 -m endpoint=:qfe3,endpoint=switch2


Checking device to use for global devices file system ... done

Adding node "phys-schost-3" to the cluster configuration ... done
Adding adapter "qfe2" to the cluster configuration ... done
Adding adapter "qfe3" to the cluster configuration ... done
Adding cable to the cluster configuration ... done
```

```
Adding cable to the cluster configuration ... done

Copying the config from "phys-schost-1" ... done

Copying the postconfig file from "phys-schost-1" if it exists ... done
Copying the Common Agent Container keys from "phys-schost-1" ... done


Setting the node ID for "phys-schost-3" ... done (id=1)

Setting the major number for the "did" driver ...
Obtaining the major number for the "did" driver from "phys-schost-1" ... done
"did" driver major number set to 300

Checking for global devices global file system ... done
Updating vfstab ... done

Verifying that NTP is configured ... done
Initializing NTP configuration ... done

Updating nsswitch.conf ...
done

Adding clusternode entries to /etc/inet/hosts ... done


Configuring IP Multipathing groups in "/etc/hostname.<adapter>" files

Updating "/etc/hostname.hme0".

Verifying that power management is NOT configured ... done

Ensure that the EEPROM parameter "local-mac-address?" is set to "true" ... done
The "local-mac-address?" parameter setting has been changed to "true".

Ensure network routing is disabled ... done

Updating file ("ntp.conf.cluster") on node phys-schost-1 ... done
Updating file ("hosts") on node phys-schost-1 ... done

Rebooting ...
```

Troubleshooting **Unsuccessful configuration** – If one or more nodes cannot join the cluster, or if the wrong configuration information was specified, first attempt to rerun this procedure. If that does not correct the problem, perform the procedure "How to Unconfigure Sun Cluster Software to

Correct Installation Problems" on page 233 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Sun Cluster software packages. Then rerun this procedure.

Next Steps   If you added a node to an existing cluster that uses a quorum device, go to "How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 139.

Otherwise, go to "How to Verify the Quorum Configuration and Installation Mode" on page 147.

▼ How to Configure Sun Cluster Software on Additional Global-Cluster Nodes (XML)

Perform this procedure to configure a new global-cluster node by using an XML cluster configuration file. The new node can be a duplication of an existing cluster node that runs Sun Cluster 3.2 1/09 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 Solaris OS is installed to support Sun Cluster software.

  If Solaris software is already installed on the node, you must ensure that the Solaris installation meets the requirements for Sun Cluster software and any other software that you intend to install on the cluster. See "How to Install Solaris Software" on page 64 for more information about installing Solaris software to meet Sun Cluster software requirements.

- SPARC: If you are configuring Sun Logical Domains (LDoms) I/O domains or guest domains as cluster nodes, ensure that Sun LDoms software is installed on each physical machine and that the domains meet Sun Cluster requirements. See "SPARC: How to Install Sun Logical Domains Software and Create Domains" on page 69.

- Ensure that Sun Cluster software packages and any necessary patches are installed on the node. See "How to Install Sun Cluster Framework and Data-Service Software Packages" on page 73.

- 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 117.

1   Ensure that Sun 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 Sun Cluster software is configured on the potential node.

phys-schost-new# **/usr/sbin/clinfo -n**

■ **If the command fails, go to Step 2.**

Sun Cluster software is not yet configured on the node. You can add the potential node to the cluster.

■ **If the command returns a node ID number, proceed to Step c.**

Sun Cluster software is already a configured on the node. Before you can add the node to a different cluster, you must remove the existing cluster configuration information.

c.   Boot the potential node into noncluster mode.

■ **On SPARC based systems, perform the following command:**

ok **boot -x**

■ **On x86 based systems, perform the following commands:**

i.   In the GRUB menu, use the arrow keys to select the appropriate Solaris entry and type e to edit its commands.

The GRUB menu appears similar to the following:

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

For more information about GRUB based booting, see Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration*.

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

The GRUB boot parameters screen appears similar to the following:

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

iii. Add -x to the command to specify that the system boot into noncluster mode.

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

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

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

The screen displays the edited command.

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

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 to add the -x option to the kernel boot parameter command.

---

    d. Unconfigure Sun Cluster software from the potential node.

        `phys-schost-new#` **`/usr/cluster/bin/clnode remove`**

2 If you are duplicating a node that runs Sun Cluster 3.2 1/09 software, create a cluster configuration XML file.

    a. Become superuser on the cluster node that you want to duplicate.

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

        `phys-schost#` **`clnode export -o`** *clconfigfile*

        `-o`   Specifies the output destination.

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

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

    c. Copy the cluster configuration XML file to the potential node that you will configure as a new cluster node.

3 Become superuser on the potential node.

4 Modify the cluster configuration XML file as needed.

    a. Open your cluster configuration XML file for editing.

      ■ **If you are duplicating an existing cluster node, open the file that you created with the** `clnode export` **command.**

      ■ **If you are not duplicating an existing cluster node, create a new file.**

         Base the file on the element hierarchy that is shown in the `clconfiguration(5CL)` man page. You can store the file in any directory.

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

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

5 Validate the cluster configuration XML file.

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

6 Configure the new cluster node.

    `phys-schost-new#` **`clnode add -n`** *sponsornode* **`-i`** *clconfigfile*

-n *sponsornode*   Specifies the name of an existing cluster member to act as the sponsor for the new node.

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

7   **(Optional) Enable automatic node reboot if all monitored disk paths fail.**

a.   **Enable the automatic reboot feature.**

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

-p   Specifies the property to set

reboot_on_path_failure=enable   Specifies that the node will reboot if all monitored disk paths fail, provided that at least one of the disks is accessible from a different node in the cluster.

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

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

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

Troubleshooting   **Unsuccessful configuration** – If one or more nodes cannot join the cluster, or if the wrong configuration information was specified, first attempt to rerun this procedure. If that does not correct the problem, perform the procedure "How to Unconfigure Sun Cluster Software to Correct Installation Problems" on page 233 on each misconfigured node to remove it from the cluster configuration. You do not need to uninstall the Sun Cluster software packages. Then rerun this procedure.

Next Steps   If you added a node to a cluster that uses a quorum device, go to "How to Update Quorum Devices After Adding a Node to a Global Cluster" on page 139.

Otherwise, go to "How to Verify the Quorum Configuration and Installation Mode" on page 147.

▼ How to Update Quorum Devices After Adding a Node to a Global Cluster

If you added a node to a global cluster, you must update the configuration information of the quorum devices, regardless of whether you use shared disks, NAS devices, a quorum server, or a combination. To do this, you remove all quorum devices and update the global-devices namespace. You can optionally reconfigure any quorum devices that you still want to use. This registers the new node with each quorum device, which can then recalculate its vote count based on the new number of nodes in the cluster.

Any newly configured SCSI quorum devices will be set to SCSI-3 reservations.

Before You Begin   Ensure that you have completed installation of Sun Cluster software on the added node.

1   On any node of the cluster, become superuser.

2   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
...
```

3   Note the name of each quorum device that is listed.

4   Remove the original quorum device.
Perform this step for each quorum device that is configured.

```
phys-schost# clquorum remove devicename
```

devicename   Specifies the name of the quorum device.

5   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
```

6   Update the global-devices namespace.

```
phys-schost# cldevice populate
```

Note – This step is necessary to prevent possible node panic.

---

7   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
```

8   (Optional) Add a quorum device.

You can configure the same device that was originally configured as the quorum device or choose a new shared device to configure.

a.   (Optional) If you want to choose a new shared device to configure as a quorum device, display all devices that the system checks.

Otherwise, skip to Step c.

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

Output resembles the following:

```
DID Device         Full Device Path
----------         ----------------
d1                 phys-schost-1:/dev/rdsk/c0t0d0
d2                 phys-schost-1:/dev/rdsk/c0t6d0
d3                 phys-schost-2:/dev/rdsk/c1t1d0
d3                 phys-schost-1:/dev/rdsk/c1t1d0
...
```

b.   From the output, choose a shared device to configure as a quorum device.

c.   Configure the shared device as a quorum device.

```
phys-schost# clquorum add -t type devicename
```

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

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

e.   Verify the new quorum configuration.

```
phys-schost# clquorum list
```

Output should list each quorum device and each node.

Example 3–4 Updating SCSI Quorum Devices After Adding a Node to a Two-Node Cluster

The following example identifies the original SCSI quorum device d2, removes that quorum device, lists the available shared devices, updates the global-device namespace, configures d3 as a new SCSI quorum device, and verifies the new device.

```
phys-schost# clquorum list
d2
phys-schost-1
phys-schost-2

phys-schost# clquorum remove d2
phys-schost# clquorum status
...
--- Quorum Votes by Device ---

Device Name       Present       Possible      Status
-----------       -------       --------      ------
phys-schost# cldevice list -v
DID Device        Full Device Path
----------        ----------------
...
d3                phys-schost-2:/dev/rdsk/c1t1d0
d3                phys-schost-1:/dev/rdsk/c1t1d0
...
phys-schost# cldevice populate
phys-schost# ps -ef - grep scgdevs
phys-schost# clquorum add d3
phys-schost# clquorum list
d3
phys-schost-1
phys-schost-2
```

Next Steps Go to "How to Verify the Quorum Configuration and Installation Mode" on page 147.

▼ How to Configure Quorum Devices

Note – You do not need to configure quorum devices in the following circumstances:

- You chose automatic quorum configuration during Sun Cluster software configuration.
- You installed a single-node global cluster.
- You added a node to an existing global cluster and already have sufficient quorum votes assigned.

Instead, proceed to "How to Verify the Quorum Configuration and Installation Mode" on page 147.

Perform this procedure one time only, after the cluster is fully formed. Use this procedure to assign quorum votes and then to remove the cluster from installation mode.

Before You Begin
- Perform the following preparations to configure a quorum server or a NAS device as a quorum device.
  - To configure a quorum server as a quorum device, do the following:
    - Install the Sun 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 Quorum Server Software" on page 57.
    - 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
  - To configure a network-attached storage (NAS) device as a quorum device, do the following:

- Install the NAS device hardware and software. See *Sun Cluster 3.1 - 3.2 With Network-Attached Storage Devices Manual for Solaris OS* and your device documentation for requirements and installation procedures for NAS hardware and software.

- For Network Appliance NAS devices, also have available the following information:
  - The name of the NAS device
  - The LUN ID of the NAS device

1   To use a quorum server as a quorum device, prepare the cluster to communicate with the quorum server.

a.   Edit the quorum-server configuration file `/etc/scqsd/scqsd.conf.`

When you install the Sun Cluster software, a default configuration file, `/etc/scqsd/scqsd.conf`, is created that contains information about a single default quorum server. Each line in the `/etc/scqsd/scqsd.conf` file has the following format:

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

`/usr/cluster/lib/sc/scqsd`	The full path to where you installed the Sun Cluster software. This value must be `/usr/cluster/lib/sc/scqsd`.
`-d` *quorumdirectory*	The path to the directory where the quorum server can store quorum data.
	The quorum server process creates one file per cluster in this directory to store cluster-specific quorum information. By default, the value of this option is `/var/scqsd`. This directory must be unique for each quorum server that you configure.
`-i` *instancename*	A unique name that you choose for the quorum server instance.
`-p` *port*	The port number on which the quorum server listens for requests from the cluster. The default port is 9000.

Instance names are optional. If you specify a name for the quorum server, that name must be unique among all quorum servers in your system. If you choose to omit the instance name option, you must refer to the quorum server by the port on which it listens.

b.   **If the public network uses variable-length subnetting, also called Classless Inter-Domain Routing (CIDR), modify the following files on each node.**

If you use classful subnets, as defined in RFC 791, you do not need to perform these steps.

   i.   **Add to the** `/etc/inet/netmasks` **file an entry for each public subnet that the cluster uses.**

   The following is an example entry which contains a public-network IP address and netmask:

   ```
   10.11.30.0    255.255.255.0
   ```

   ii.   **Append** `netmask + broadcast +` **to the hostname entry in each** `/etc/hostname.`*adapter* **file.**

   nodename **netmask + broadcast +**

c.   **Ensure that the IP address of the quorum server is included in the** `/etc/inet/hosts` **or** `/etc/inet/ipnodes` **file on each node in the cluster.**

d.   **If you use a naming service, ensure that the quorum server is included in the name-to-address mappings.**

2   **On one node, become superuser.**

3   **To use a shared disk as a quorum device, verify device connectivity to the cluster nodes and choose the device to configure.**

a.   **From one node of the cluster, display a list of all the devices that the system checks.**

   You do not need to be logged in as superuser to run this command.

   ```
   phys-schost-1# cldevice list -v
   ```

   Output resembles the following:

   ```
   DID Device        Full Device Path
   ----------        ----------------
   d1                phys-schost-1:/dev/rdsk/c0t0d0
   d2                phys-schost-1:/dev/rdsk/c0t6d0
   d3                phys-schost-2:/dev/rdsk/c1t1d0
   d3                phys-schost-1:/dev/rdsk/c1t1d0
   ...
   ```

b.   **Ensure that the output shows all connections between cluster nodes and storage devices.**

c.   **Determine the global device-ID name of each shared disk that you are configuring as a quorum device.**

> **Note –** Any shared disk that you choose must be qualified for use as a quorum device. See "Quorum Devices" on page 37 for further information about choosing quorum devices.

Use the scdidadm output from Step a to identify the device–ID name of each shared disk that you are configuring as a quorum device. For example, the output in Step a shows that global device d2 is shared by phys-schost-1 and phys-schost-2.

4 If you intend to configure a quorum device from 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.

```
# cldevice show device

=== DID Device Instances ===
DID Device Name:                                      /dev/did/rdsk/dN
...
  default_fencing:                                        nofencing
...
```

b. If default_fencing **is not set to** nofencing **or** nofencing-noscrub**, determine whether fencing is disabled globally.**

```
# cluster show -t global

=== Cluster ===
Cluster name:                                          cluster
...
    global_fencing:                                      nofencing
...
```

c. If neither default_fencing **nor** global_fencing **is set to** nofencing **or** nofencing-noscrub**, disable fencing for the shared disk.**

```
# cldevice set -p default_fencing=nofencing-noscrub device
```

d. Verify that fencing for the shared disk is now disabled.

```
# cldevice show device
```

5 Start the clsetup **utility.**

```
phys-schost# clsetup
```

The Initial Cluster Setup screen is displayed.

> **Note** – If the Main Menu is displayed instead, initial cluster setup was already successfully
> performed. Skip to Step 10.

6 Answer the prompt `Do 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 9.**

  - **Type Yes to configure additional quorum devices. Then proceed to Step 7.**

7 Specify what type of device you want to configure as a quorum device.

Quorum Device Type	Description
shared_disk	Sun NAS device or shared disk
quorum_server	Quorum server
netapp_nas	Network Appliance NAS device

8 Specify the name of the device to configure as a quorum device.

- For a quorum server, also specify the following information:
  - The IP address of the quorum server host
  - The port number that is used by the quorum server to communicate with the cluster nodes
- For a Network Appliance NAS device, also specify the following information:
  - The name of the NAS device
  - The LUN ID of the NAS device

9 At the prompt `Is it okay to reset "installmode"?`**, type Yes.**

After the `clsetup` utility sets the quorum configurations and vote counts for the cluster, the
message `Cluster initialization is complete` is displayed. The utility returns you to the
Main Menu.

10 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 147.

Troubleshooting **Interrupted clsetup processing** - If the quorum setup process is interrupted or fails to be completed successfully, rerun `clsetup`.

Changes to quorum vote count – If you later increase or decrease the number of node attachments to a quorum device, the quorum vote count is not automatically recalculated. You can reestablish the correct quorum vote by removing each quorum device and then add it back into the configuration, one quorum device at a time. For a two-node cluster, temporarily add a new quorum device before you remove and add back the original quorum device. Then remove the temporary quorum device. See the procedure "How to Modify a Quorum Device Node List" in Chapter 6, "Administering Quorum," in *Sun Cluster System Administration Guide for Solaris OS*.

▼ How to Verify the Quorum Configuration and Installation Mode

Perform this procedure to verify that 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 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 148.

- If you did not install your own /etc/inet/ntp.conf file before you installed Sun Cluster software, install or create the NTP configuration file. Go to "How to Configure Network Time Protocol (NTP)" on page 150.

- If you want to configure IPsec on the private interconnect, go to "How to Configure IP Security Architecture (IPsec) on the Cluster Interconnect" on page 153.

- To install a volume manager, go to Chapter 4, "Configuring Solaris Volume Manager Software," and Chapter 5, "Installing and Configuring Veritas Volume Manager," to install volume management software.

Note – If you added a new node to a cluster that uses VxVM, you must perform one of the following tasks:

- Install VxVM on that node.
- Modify that node's /etc/name_to_major file to support coexistence with VxVM.

Follow the procedures in "How to Install Veritas Volume Manager Software" on page 189 to perform one of these required tasks.

- To create cluster file systems, go to "How to Create Cluster File Systems" on page 205.
- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 210.
- SPARC: To configure Sun Management Center to monitor the cluster, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 227.
- 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 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.
- 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 154.

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 *Sun Cluster System Administration Guide for Solaris OS*.

▼ 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 Sun Cluster software installation.

Note – Do *not* perform this procedure after applications and data services have been configured and have been started. Otherwise, an application or data service might continue to use the old private hostname after the hostname is renamed, which would cause hostname conflicts. If any applications or data services are running, stop them before you perform this procedure.

Perform this procedure on one active node of the cluster.

1 Become superuser on a global-cluster node.

2 Start the clsetup **utility.**

```
phys-schost# clsetup
```

The clsetup Main Menu is displayed.

3 Type the option number for Private Hostnames and press the Return key.

The Private Hostname Menu is displayed.

4 Type the option number for Change a Private Hostname and press the Return key.

5 Follow the prompts to change the private hostname.

Repeat for each private hostname to change.

6 Verify the new private hostnames.

```
phys-schost# clnode show -t node | grep privatehostname
  privatehostname:                      clusternode1-priv
  privatehostname:                      clusternode2-priv
  privatehostname:                      clusternode3-priv
```

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- If you did not install your own /etc/inet/ntp.conf file before you installed Sun Cluster software, install or create the NTP configuration file. Go to "How to Configure Network Time Protocol (NTP)" on page 150.

- If you want to configure IPsec on the private interconnect, go to "How to Configure IP Security Architecture (IPsec) on the Cluster Interconnect" on page 153.

- To install a volume manager, go to Chapter 4, "Configuring Solaris Volume Manager Software," and Chapter 5, "Installing and Configuring Veritas Volume Manager," to install volume management software.

Note – If you added a new node to a cluster that uses VxVM, you must perform one of the following tasks:

- Install VxVM on that node.
- Modify that node's /etc/name_to_major file to support coexistence with VxVM.

Follow the procedures in "How to Install Veritas Volume Manager Software" on page 189 to perform one of these required tasks.

- To create cluster file systems, go to "How to Create Cluster File Systems" on page 205.
- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 210.
- SPARC: To configure Sun Management Center to monitor the cluster, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 227.
- 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 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.
- 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 154.

▼ How to Configure Network Time Protocol (NTP)

Note – If you installed your own /etc/inet/ntp.conf file **before** you installed Sun Cluster software, you do not need to perform this procedure. Determine your next step:

Perform this task to create or modify the NTP configuration file after you perform any of the following tasks:

- Install Sun Cluster software
- Add a node to an existing global cluster
- Change the private hostname of a node in the global cluster

If you added a node to a single-node cluster, you must ensure that the NTP configuration file that you use is copied to the original cluster node as well as to the new node.

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

1   Become superuser on a cluster node.

2   If you have your own `/etc/inet/ntp.conf` **file, copy your file to each node of the cluster.**

3   If you do not have your own `/etc/inet/ntp.conf` **file to install, use the**
`/etc/inet/ntp.conf.cluster` **file as your NTP configuration file.**

Note – Do not rename the `ntp.conf.cluster` file as `ntp.conf`.

If the `/etc/inet/ntp.conf.cluster` file does not exist on the node, you might have an `/etc/inet/ntp.conf` file from an earlier installation of Sun Cluster software. Sun Cluster software creates the `/etc/inet/ntp.conf.cluster` file as the NTP configuration file if an `/etc/inet/ntp.conf` file is not already present on the node. If so, perform the following edits instead on that `ntp.conf` file.

a.   Use your preferred text editor to open the NTP configuration file on one node of the cluster for editing.

b.   Ensure that an entry exists for the private hostname of each cluster node.

If you changed any node's private hostname, ensure that the NTP configuration file contains the new private hostname.

c.   If necessary, make other modifications to meet your NTP requirements.

d.   Copy the NTP configuration file to all nodes in the cluster.

The contents of the NTP configuration file must be identical on all cluster nodes.

4   Stop the NTP daemon on each node.

Wait for the command to complete successfully on each node before you proceed to Step 5.

- SPARC: For the Solaris 9 OS, use the following command:

  phys-schost# **/etc/init.d/xntpd stop**

- For the Solaris 10 OS, use the following command:

  phys-schost# **svcadm disable ntp**

5 Restart the NTP daemon on each node.

- **If you use the** `ntp.conf.cluster` **file, run the following command:**

  phys-schost# **/etc/init.d/xntpd.cluster start**

  The `xntpd.cluster` startup script first looks for the `/etc/inet/ntp.conf` file.

  - If the `ntp.conf` file exists, the script exits immediately without starting the NTP daemon.
  - If the `ntp.conf` file does not exist but the `ntp.conf.cluster` file does exist, the script starts the NTP daemon. In this case, the script uses the `ntp.conf.cluster` file as the NTP configuration file.

- **If you use the** `ntp.conf` **file, run one of the following commands:**

  - SPARC: For the Solaris 9 OS, use the following command:

    phys-schost# **/etc/init.d/xntpd start**

  - For the Solaris 10 OS, use the following command:

    phys-schost# **svcadm enable ntp**

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- If you want to configure IPsec on the private interconnect, go to "How to Configure IP Security Architecture (IPsec) on the Cluster Interconnect" on page 153.
- To install a volume manager, go to Chapter 4, "Configuring Solaris Volume Manager Software," and Chapter 5, "Installing and Configuring Veritas Volume Manager," to install volume management software.

  Note – If you added a new node to a cluster that uses VxVM, you must perform one of the following tasks:

  - Install VxVM on that node.
  - Modify that node's /etc/name_to_major file to support coexistence with VxVM.

  Follow the procedures in "How to Install Veritas Volume Manager Software" on page 189 to perform one of these required tasks.

- To create cluster file systems, go to "How to Create Cluster File Systems" on page 205.
- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 210.

- SPARC: To configure Sun Management Center to monitor the cluster, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 227.

- 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 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

- 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 154.

▼ How to Configure IP Security Architecture (IPsec) on the Cluster Interconnect

You can configure IP Security Architecture (IPsec) for the clprivnet interface to provide secure communication on the cluster interconnect. For information about IPsec, see Part IV, "IP Security," in *System Administration Guide: IP Services* and the ipsecconf(1M) man page. For information about the clprivnet interface, see the clprivnet(7) man page.

Perform this procedure on each global-cluster voting node that you want to configure to use IPsec.

1 **Become superuser.**

2 **On each node, determine the IP address of the** clprivnet **interface of the node.**

```
phys-schost# ifconfig clprivnet0
```

3 **On each node, configure the** /etc/inet/ipsecinit.conf **policy file and add Security Associations (SAs) between each pair of** clprivnet **IP addresses that you want to use IPsec.**

Follow the instructions in "How to Secure Traffic Between Two Systems With IPsec" in *System Administration Guide: IP Services*.

Note – To implement IPsec without rebooting, follow the instructions in the procedure's example, "Securing Traffic With IPsec Without Rebooting."

Observe the following guidelines when you add entries to the configuration file:

- In each file, add one entry for each clprivnet IP address to use IPsec, including the clprivnet IP address of the local node.

- Configure each policy as a separate line in the configuration file.

- Ensure that the values of the configuration parameters for these addresses are consistent on all the partner nodes.

- To enable striping of data over all links, include the sa unique policy in the entry. This features helps the driver to optimally utilize the bandwidth of the cluster private network, which provides a high granularity of distribution and better throughput. The clprivnet interface uses the Security Parameter Index (SPI) of the packet to stripe the traffic. For more information about the sa unique policy, see the ipsecconf(1M) man page.

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- To install a volume manager, go to Chapter 4, "Configuring Solaris Volume Manager Software," and Chapter 5, "Installing and Configuring Veritas Volume Manager," to install volume management software.

Note – If you added a new node to a cluster that uses VxVM, you must perform one of the following tasks:

- Install VxVM on that node.
- Modify that node's /etc/name_to_major file to support coexistence with VxVM.

Follow the procedures in "How to Install Veritas Volume Manager Software" on page 189 to perform one of these required tasks.

- To create cluster file systems, go to "How to Create Cluster File Systems" on page 205.
- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 210.
- SPARC: To configure Sun Management Center to monitor the cluster, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 227.
- 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 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.
- 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 154.

▼ 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 Sun Explorer utility to record baseline information about the cluster. This data can be used if there is a future need to troubleshoot the cluster.

1 Become superuser.

2 Run the `explorer` **utility on each node in the cluster.**

Use the appropriate command for your platform:

Server	Command
Sun Fire 3800 through 6800	`# explorer -i -w default,scextended`
Sun Fire V1280 and E2900	`# explorer -i -w default,1280extended`
Sun Fire T1000 and T2000	`# explorer -i -w default,Tx000`
Sun Fire X4x00 and X8x00	`# explorer -i -w default,ipmi`
All other platforms	`# explorer -i`

For more information, see the explorer(1M) man page in the `/opt/SUNWexplo/man/man1m/` directory and *Sun Explorer User's Guide*.

The `explorer` output file is saved in the `/opt/SUNWexplo/output/` directory as explorer.*hostid*.*hostname-date*.tar.gz.

3 Save the files to a location that you can access if the entire cluster is down.

4 Send all `explorer` **files by email to the Sun Explorer database alias for your geographic location.**

This database makes your `explorer` output available to Sun technical support if the data is needed to help diagnose a technical problem with your cluster.

Location	Email Address
North, Central, and South America (AMER)	`explorer-database-americas@sun.com`
Europe, Middle East, and Africa (EMEA)	`explorer-database-emea@sun.com`
Asia, Australia, New Zealand, and Pacific (APAC)	`explorer-database-apac@sun.com`

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

Note – The Enhanced Storage module of Solaris Management Console is not compatible with Sun Cluster software. Use the command-line interface or Sun Cluster utilities to configure Solaris Volume Manager software.

The following sections are in this chapter:

- "Configuring Solaris Volume Manager Software" on page 157
- "Creating Disk Sets in a Cluster" on page 172
- "Configuring Dual-String Mediators" on page 182

Configuring Solaris Volume Manager Software

The following table lists the tasks that you perform to configure Solaris Volume Manager software for Sun 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 46

TABLE 4–1   Task Map: Configuring Solaris Volume Manager Software   *(Continued)*

Task	Instructions
(*Solaris 9 only*) Calculate the number of volume names and disk sets that you need for your configuration, and modify the `/kernel/drv/md.conf` file.	"SPARC: How to Set the Number of Volume Names and Disk Sets" on page 158
Create state database replicas on the local disks.	"How to Create State Database Replicas" on page 160
(*Optional*) Mirror file systems on the root disk.	"Mirroring the Root Disk" on page 160

▼ SPARC: How to Set the Number of Volume Names and Disk Sets

Note – This procedure is only required for the Solaris 9 OS. If the cluster runs on the Solaris 10 OS, proceed to "How to Create State Database Replicas" on page 160.

With the Solaris 10 release, Solaris Volume Manager has been enhanced to configure volumes dynamically. You no longer need to edit the nmd and the md_nsets parameters in the `/kernel/drv/md.conf` file. New volumes are dynamically created, as needed.

This procedure describes how to determine the number of Solaris Volume Manager volume names and disk sets that you need for your configuration. This procedure also describes how to modify the `/kernel/drv/md.conf` file to specify these numbers.

Tip – The default number of volume names per disk set is 128, but many configurations need more than the default. Increase this number before you implement a configuration, to save administration time later.

At the same time, keep the value of the nmd field and the md_nsets field as low as possible. Memory structures exist for all possible devices as determined by nmd and md_nsets, even if you have not created those devices. For optimal performance, keep the value of nmd and md_nsets only slightly higher than the number of volumes that you plan to use.

Before You Begin   Have available the completed "Device Group Configurations Worksheet" on page 251.

1   **Calculate the total number of disk sets that you expect to need in the cluster, then add one more disk set for private disk management.**

The cluster can have a maximum of 32 disk sets, 31 disk sets for general use plus one disk set for private disk management. The default number of disk sets is 4. You supply this value for the md_nsets field in Step 3.

2   **Calculate the largest volume name that you expect to need for any disk set in the cluster.**

Each disk set can have a maximum of 8192 volume names. You supply this value for the nmd field in Step 3.

a.   **Determine the quantity of volume names that you expect to need for each disk set.**

If you use local volumes, ensure that each local volume name on which a global-devices file system, /global/.devices/node@ *nodeid*, is mounted is unique throughout the cluster and does not use the same name as any device-ID name in the cluster.

Tip – Choose a range of numbers to use exclusively for device-ID names and a range for each node to use exclusively for its local volume names. For example, device-ID names might use the range from d1 to d100. Local volumes on node 1 might use names in the range from d100 to d199. And local volumes on node 2 might use d200 to d299.

b.   **Calculate the highest of the volume names that you expect to use in any disk set.**

The quantity of volume names to set is based on the volume name *value* rather than on the *actual quantity* . For example, if your volume names range from d950 to d1000, Solaris Volume Manager software requires that you set the value at 1000 names, not 50.

3   **On each node, become superuser and edit the** /kernel/drv/md.conf **file.**

Caution – All cluster nodes (or cluster pairs in the cluster-pair topology) must have identical /kernel/drv/md.conf files, regardless of the number of disk sets served by each node. Failure to follow this guideline can result in serious Solaris Volume Manager errors and possible loss of data.

a.   **Set the** md_nsets **field to the value that you determined in Step 1.**

b.   **Set the** nmd **field to the value that you determined in Step 2.**

4   **On each node, perform a reconfiguration reboot.**

```
phys-schost# touch /reconfigure
phys-schost# shutdown -g0 -y -i6
```

Changes to the /kernel/drv/md.conf file become operative after you perform a reconfiguration reboot.

Next Steps   Create local state database replicas. Go to "How to Create State Database Replicas" on page 160.

▼ 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 To mirror file systems on the root disk, go to "Mirroring the Root Disk" on page 160.

Otherwise, go to "Creating Disk Sets in a Cluster" on page 172 to create Solaris Volume Manager disk sets.

Mirroring the Root Disk

Mirroring the root disk prevents the cluster node itself from shutting down because of a system disk failure. Four types of file systems can reside on the root disk. Each file-system type is mirrored by using a different method.

Use the following procedures to mirror each type of file system.

- "How to Mirror the Root (/) File System" on page 161
- "How to Mirror the Global Devices Namespace" on page 164
- "How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted" on page 166
- "How to Mirror File Systems That Can Be Unmounted" on page 169

⚠ **Caution** – For local disk mirroring, do not use /dev/global as the path when you specify the disk name. If you specify this path for anything other than cluster file systems, the system cannot boot.

▼ How to Mirror the Root (/) File System

Use this procedure to mirror the root (/) file system.

Note – This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix B, "Sun Cluster Object-Oriented Commands," in *Sun Cluster System Administration Guide for Solaris OS*.

1 Become superuser.

2 Place the root slice in a single-slice (one-way) concatenation.

Specify the physical disk name of the root-disk slice (c*N*t*X*d*Y*s*Z*).

phys-schost# **metainit -f** *submirror1* **1 1** *root-disk-slice*

3 Create a second concatenation.

phys-schost# **metainit** *submirror2* **1 1** *submirror-disk-slice*

4 Create a one-way mirror with one submirror.

phys-schost# **metainit** *mirror* **-m** *submirror1*

Note – If the device is a local device to be used to mount a global-devices file system, /global/.devices/node@*nodeid*, the volume name for the mirror *must* be unique throughout the cluster.

5 Set up the system files for the root (/) directory.

phys-schost# **metaroot** *mirror*

This command edits the /etc/vfstab and /etc/system files so the system can be booted with the root (/) file system on a metadevice or volume. For more information, see the metaroot(1M) man page.

6   Flush all file systems.

```
phys-schost# lockfs -fa
```

This command flushes all transactions out of the log and writes the transactions to the master file system on all mounted UFS file systems. For more information, see the lockfs(1M) man page.

7   Move any resource groups or device groups from the node.

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

from-node   Specifies the name of the node from which to evacuate resource or device groups.

8   Reboot the node.

This command remounts the newly mirrored root (/) file system.

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

9   Attach the second submirror to the mirror.

```
phys-schost# metattach mirror submirror2
```

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

10   If the disk that is used to mirror the root disk is physically connected to more than one node (multihosted), disable fencing for that disk.

Disabling fencing for the device prevents unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

```
phys-schost# cldevice set -p default_fencing=nofencing submirror-disk
```

-p   Specifies a device property.

default_fencing=nofencing   Disables fencing for the specified device.

For more information about the default_fencing property, see the cldevice(1CL) man page.

11   Record the alternate boot path for possible future use.

If the primary boot device fails, you can then boot from this alternate boot device. See "Special Considerations for Mirroring root (/)" in *Solaris Volume Manager Administration Guide* or "Creating a RAID-1 Volume" in *Solaris Volume Manager Administration Guide* for more information about alternate boot devices.

```
phys-schost# ls -l /dev/rdsk/root-disk-slice
```

12 Repeat Step 1 through Step 11 on each remaining node of the cluster.

Ensure that each volume name for a mirror on which a global-devices file system, /global/.devices/node@*nodeid*, is to be mounted is unique throughout the cluster.

Example 4–2 Mirroring the Root (/) File System

The following example shows the creation of mirror d0 on the node phys-schost-1, which consists of submirror d10 on partition c0t0d0s0 and submirror d20 on partition c2t2d0s0. Device c2t2d0 is a multihost disk, so fencing is disabled. The example also displays the alternate boot path for recording.

```
phys-schost# metainit -f d10 1 1 c0t0d0s0
d11: Concat/Stripe is setup
phys-schost# metainit d20 1 1 c2t2d0s0
d12: Concat/Stripe is setup
phys-schost# metainit d0 -m d10
d10: Mirror is setup
phys-schost# metaroot d0
phys-schost# lockfs -fa
phys-schost# clnode evacuate phys-schost-1
phys-schost# shutdown -g0 -y -i6
phys-schost# metattach d0 d20
d0: Submirror d20 is attachedphys-schost# cldevice set -p default_fencing=nofencing c2t2d0
phys-schost# ls -l /dev/rdsk/c2t2d0s0
lrwxrwxrwx  1 root     root          57 Apr 25 20:11 /dev/rdsk/c2t2d0s0
–> ../../devices/node@1/pci@1f,0/pci@1/scsi@3,1/disk@2,0:a,raw
```

Next Steps To mirror the global devices namespace, /global/.devices/node@*nodeid*, go to "How to Mirror the Global Devices Namespace" on page 164.

To mirror file systems than cannot be unmounted, go to "How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted" on page 166.

To mirror user-defined file systems, go to "How to Mirror File Systems That Can Be Unmounted" on page 169.

Otherwise, go to "Creating Disk Sets in a Cluster" on page 172 to create a disk set.

Troubleshooting Some of the steps in this mirroring procedure might cause an error message similar to metainit: dg-schost-1: d1s0: not a metadevice. Such an error message is harmless and can be ignored.

▼ How to Mirror the Global Devices Namespace

Use this procedure to mirror the global devices namespace, /global/.devices/node@*nodeid*/.

Note – This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix B, "Sun Cluster Object-Oriented Commands," in *Sun Cluster System Administration Guide for Solaris OS*.

---

1   Become superuser.

2   Place the global devices namespace slice in a single-slice (one-way) concatenation.

Use the physical disk name of the disk slice (c*N*t*X*d*Y*s*Z*).

```
phys-schost# metainit -f submirror1 1 1 diskslice
```

3   Create a second concatenation.

```
phys-schost# metainit submirror2 1 1 submirror-diskslice
```

4   Create a one-way mirror with one submirror.

```
phys-schost# metainit mirror -m submirror1
```

Note – The volume name for a mirror on which a global-devices file system, /global/.devices/node@*nodeid*, is to be mounted *must* be unique throughout the cluster.

---

5   Attach the second submirror to the mirror.

This attachment starts a synchronization of the submirrors.

```
phys-schost# metattach mirror submirror2
```

6   Edit the /etc/vfstab **file entry for the** /global/.devices/node@*nodeid* **file system.**

Replace the names in the device to mount and device to fsck columns with the mirror name.

```
phys-schost# vi /etc/vfstab
#device          device          mount     FS      fsck    mount     mount
#to mount        to fsck         point     type    pass    at boot   options
#
/dev/md/dsk/mirror /dev/md/rdsk/mirror /global/.devices/node@nodeid ufs 2 no global
```

7   Repeat Step 1 through Step 6 on each remaining node of the cluster.

8   Wait for the synchronization of the mirrors, started in Step 5, to be completed.

Use the metastat(1M) command to view mirror status and to verify that mirror synchronization is complete.

phys-schost# **metastat** *mirror*

9   If the disk that is used to mirror the global devices namespace is physically connected to more than one node (multihosted), disable fencing for that disk.

Disabling fencing for the device prevents unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

phys-schost# **cldevice set -p default_fencing=nofencing** *submirror-disk*

-p   Specifies a device property.

default_fencing=nofencing   Disables fencing for the specified device.

For more information about the default_fencing property, see the cldevice(1CL) man page.

Example 4–3   Mirroring the Global Devices Namespace

The following example shows creation of mirror d101, which consists of submirror d111 on partition c0t0d0s3 and submirror d121 on partition c2t2d0s3. The /etc/vfstab file entry for /global/.devices/node@1 is updated to use the mirror name d101. Device c2t2d0 is a multihost disk, so fencing is disabled.

```
phys-schost# metainit -f d111 1 1 c0t0d0s3
d111: Concat/Stripe is setup
phys-schost# metainit d121 1 1 c2t2d0s3
d121: Concat/Stripe is setup
phys-schost# metainit d101 -m d111
d101: Mirror is setup
phys-schost# metattach d101 d121
d101: Submirror d121 is attached
phys-schost# vi /etc/vfstab
#device         device          mount   FS      fsck    mount   mount
#to mount       to fsck         point   type    pass    at boot options
#
/dev/md/dsk/d101 /dev/md/rdsk/d101 /global/.devices/node@1 ufs 2 no global
phys-schost# metastat d101
d101: Mirror
      Submirror 0: d111
         State: Okay
      Submirror 1: d121
         State: Resyncing
      Resync in progress: 15 % done
...
phys-schost# cldevice show phys-schost-3:/dev/rdsk/c2t2d0
```

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

DID Device Name:                               /dev/did/rdsk/d2
  Full Device Path:                              phys-schost-1:/dev/rdsk/c2t2d0
  Full Device Path:                              phys-schost-3:/dev/rdsk/c2t2d0
...

phys-schost# cldevicegroup show | grep dsk/d2
Device Group Name:                             dsk/d2
...
  Node List:                                     phys-schost-1, phys-schost-3
...
  localonly:                                     false
phys-schost# cldevicegroup remove-node -n phys-schost-3 dsk/d2
phys-schost# cldevice set -p default_fencing=nofencing c2t2d0
```

Next Steps   To mirror file systems other than root (/) that cannot be unmounted, go to "How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted" on page 166.

To mirror user-defined file systems, go to "How to Mirror File Systems That Can Be Unmounted" on page 169

Otherwise, go to "Creating Disk Sets in a Cluster" on page 172 to create a disk set.

Troubleshooting   Some of the steps in this mirroring procedure might cause an error message similar to metainit: dg-schost-1: d1s0: not a metadevice. Such an error message is harmless and can be ignored.

▼ How to Mirror File Systems Other Than Root (/) That Cannot Be Unmounted

Use this procedure to mirror file systems other than root (/) that cannot be unmounted during normal system usage, such as /usr, /opt, or swap.

Note – This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix B, "Sun Cluster Object-Oriented Commands," in *Sun Cluster System Administration Guide for Solaris OS*.

---

1   **Become superuser.**

2 Place the slice on which an unmountable file system resides in a single-slice (one-way) concatenation.

Specify the physical disk name of the disk slice (c*N*t*X* d*Y*s*Z*).

```
phys-schost# metainit -f submirror1 1 1 diskslice
```

3 Create a second concatenation.

```
phys-schost# metainit submirror2 1 1 submirror-diskslice
```

4 Create a one-way mirror with one submirror.

```
phys-schost# metainit mirror -m submirror1
```

Note – The volume name for this mirror does *not* need to be unique throughout the cluster.

---

5 Repeat Step 1 through Step 4 for each remaining unmountable file system that you want to mirror.

6 On each node, edit the /etc/vfstab **file entry for each unmountable file system you mirrored.**

Replace the names in the device to mount and device to fsck columns with the mirror name.

```
phys-schost# vi /etc/vfstab
#device        device        mount   FS    fsck    mount   mount
#to mount      to fsck       point   type  pass    at boot options
#
/dev/md/dsk/mirror /dev/md/rdsk/mirror /filesystem ufs 2 no global
```

7 Move any resource groups or device groups from the node.

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

from-node   Specifies the name of the node from which to move resource or device groups.

8 Reboot the node.

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

9 Attach the second submirror to each mirror.

This attachment starts a synchronization of the submirrors.

```
phys-schost# metattach mirror submirror2
```

10 Wait for the synchronization of the mirrors, started in Step 9, to complete.

Use the metastat(1M) command to view mirror status and to verify that mirror synchronization is complete.

```
phys-schost# metastat mirror
```

11   If the disk that is used to mirror the unmountable file system is physically connected to more than one node (multihosted), disable fencing for that disk.

Disabling fencing for the device prevents unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

```
phys-schost# cldevice set -p default_fencing=nofencing submirror-disk
```

-p   Specifies a device property.

default_fencing=nofencing   Disables fencing for the specified device.

For more information about the default_fencing property, see the cldevice(1CL) man page.

Example 4–4   Mirroring File Systems That Cannot Be Unmounted

The following example shows the creation of mirror d1 on the node phys-schost-1 to mirror /usr, which resides on c0t0d0s1. Mirror d1 consists of submirror d11 on partition c0t0d0s1 and submirror d21 on partition c2t2d0s1. The /etc/vfstab file entry for /usr is updated to use the mirror name d1. Device c2t2d0 is a multihost disk, so fencing is disabled.

```
phys-schost# metainit -f d11 1 1 c0t0d0s1
d11: Concat/Stripe is setup
phys-schost# metainit d21 1 1 c2t2d0s1
d21: Concat/Stripe is setup
phys-schost# metainit d1 -m d11
d1: Mirror is setup
phys-schost# vi /etc/vfstab
#device         device         mount   FS    fsck    mount    mount
#to mount       to fsck        point   type  pass    at boot  options
#
/dev/md/dsk/d1 /dev/md/rdsk/d1 /usr ufs  2       no global
...
phys-schost# clnode evacuate phys-schost-1
phys-schost# shutdown -g0 -y -i6
phys-schost# metattach d1 d21
d1: Submirror d21 is attached
phys-schost# metastat d1
d1: Mirror
      Submirror 0: d11
        State: Okay
      Submirror 1: d21
        State: Resyncing
      Resync in progress: 15 % done
...
phys-schost# cldevice show phys-schost-3:/dev/rdsk/c2t2d0
...
DID Device Name:                             /dev/did/rdsk/d2
phys-schost# cldevicegroup show dsk/d2
```

```
Device Group Name:                              dsk/d2
...
  Node List:                                    phys-schost-1, phys-schost-3
...
  localonly:                                    false
phys-schost# cldevicegroup remove-node -n phys-schost-3 dsk/d2
phys-schost# cldevice set -p default_fencing=nofencing c2t2d0
```

Next Steps To mirror user-defined file systems, go to "How to Mirror File Systems That Can Be Unmounted" on page 169.

Otherwise, go to "Creating Disk Sets in a Cluster" on page 172 to create a disk set.

Troubleshooting Some of the steps in this mirroring procedure might cause an error message similar to metainit: dg-schost-1: d1s0: not a metadevice. Such an error message is harmless and can be ignored.

▼ How to Mirror File Systems That Can Be Unmounted

Use this procedure to mirror user-defined file systems that can be unmounted. In this procedure, the nodes do not need to be rebooted.

Note – This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix B, "Sun Cluster Object-Oriented Commands," in *Sun Cluster System Administration Guide for Solaris OS*.

---

1   Become superuser.

2   Unmount the file system to mirror.
Ensure that no processes are running on the file system.

phys-schost# **umount** */mount-point*

See the umount(1M) man page and Chapter 18, "Mounting and Unmounting File Systems (Tasks)," in *System Administration Guide: Devices and File Systems* for more information.

3   Place in a single-slice (one-way) concatenation the slice that contains a user-defined file system that can be unmounted.
Specify the physical disk name of the disk slice (cNtXdYsZ).

phys-schost# **metainit -f** *submirror1* **1 1** *diskslice*

4   Create a second concatenation.

```
phys-schost# metainit submirror2 1 1 submirror-diskslice
```

5   Create a one-way mirror with one submirror.

```
phys-schost# metainit mirror -m submirror1
```

Note – The volume name for this mirror does *not* need to be unique throughout the cluster.

---

6   Repeat Step 1 through Step 5 for each mountable file system to be mirrored.

7   On each node, edit the /etc/vfstab **file entry for each file system you mirrored.**

Replace the names in the device to mount and device to fsck columns with the mirror name.

```
phys-schost# vi /etc/vfstab
#device       device       mount   FS     fsck    mount   mount
#to mount     to fsck      point   type   pass    at boot options
#
/dev/md/dsk/mirror /dev/md/rdsk/mirror /filesystem ufs 2 no global
```

8   Attach the second submirror to the mirror.

This attachment starts a synchronization of the submirrors.

```
phys-schost# metattach mirror submirror2
```

9   Wait for the synchronization of the mirrors, started in Step 8, to be completed.

Use the metastat(1M) command to view mirror status.

```
phys-schost# metastat mirror
```

10   If the disk that is used to mirror the user-defined file system is physically connected to more than one node (multihosted), disable fencing for that disk.

Disabling fencing for the device prevents unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

```
phys-schost# cldevice set -p default_fencing=nofencing submirror-disk
```

-p   Specifies a device property.

default_fencing=nofencing   Disables fencing for the specified device.

For more information about the default_fencing property, see the cldevice(1CL) man page.

11   Mount the mirrored file system.

```
phys-schost# mount /mount-point
```

See the mount(1M) man page and Chapter 18, "Mounting and Unmounting File Systems (Tasks)," in *System Administration Guide: Devices and File Systems* for more information.

Example 4–5 Mirroring File Systems That Can Be Unmounted

The following example shows creation of mirror d4 to mirror /export, which resides on c0t0d0s4. Mirror d4 consists of submirror d14 on partition c0t0d0s4 and submirror d24 on partition c2t2d0s4. The /etc/vfstab file entry for /export is updated to use the mirror name d4. Device c2t2d0 is a multihost disk, so fencing is disabled.

```
phys-schost# umount /export
phys-schost# metainit -f d14 1 1 c0t0d0s4
d14: Concat/Stripe is setup
phys-schost# metainit d24 1 1 c2t2d0s4
d24: Concat/Stripe is setup
phys-schost# metainit d4 -m d14
d4: Mirror is setup
phys-schost# vi /etc/vfstab
#device         device          mount   FS     fsck    mount    mount
#to mount       to fsck         point   type   pass    at boot  options
#
# /dev/md/dsk/d4 /dev/md/rdsk/d4 /export ufs 2 no     global
phys-schost# metattach d4 d24
d4: Submirror d24 is attached
phys-schost# metastat d4
d4: Mirror
      Submirror 0: d14
        State: Okay
      Submirror 1: d24
        State: Resyncing
      Resync in progress: 15 % done
...
phys-schost# cldevice show phys-schost-3:/dev/rdsk/c2t2d0
...
DID Device Name:                              /dev/did/rdsk/d2
phys-schost# cldevicegroup show dsk/d2
Device Group Name:                    dsk/d2
...
  Node List:                                  phys-schost-1, phys-schost-2
...
  localonly:                          false
phys-schost# cldevicegroup remove-node -n phys-schost-3 dsk/d2
phys-schost# cldevice set -p default_fencing=nofencing c2t2d0
phys-schost# mount /export
```

Next Steps   If you need to create disk sets, go to one of the following:

- To create a Solaris Volume Manager for Sun Cluster disk set for use by Oracle Real Application Clusters, go to "How to Create a Multi-Owner Disk Set in Solaris Volume Manager for Sun Cluster for the Oracle RAC Database" in *Sun Cluster Data Service for Oracle RAC Guide for Solaris OS*.

- To create a disk set for any other application, go to "Creating Disk Sets in a Cluster" on page 172.

If you have sufficient disk sets for your needs, go to one of the following:

- If your cluster contains disk sets that are configured with exactly two disk enclosures and two nodes, you must add dual-string mediators. Go to "Configuring Dual-String Mediators" on page 182.

- If your cluster configuration does not require dual-string mediators, go to "How to Create Cluster File Systems" on page 205.

Troubleshooting Some of the steps in this mirroring procedure might cause an error message that is similar to `metainit: dg-schost-1: d1s0: not a metadevice`. Such an error message is harmless and can be ignored.

Creating Disk Sets in a Cluster

This section describes how to create disk sets for a cluster configuration. When you create a Solaris Volume Manager disk set in a Sun Cluster environment, the disk set is automatically registered with the Sun 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.

Note – To create a Solaris Volume Manager for Sun Cluster disk set for use by Oracle Real Application Clusters, do not use these procedures. Instead, perform the procedures in "How to Create a Multi-Owner Disk Set in Solaris Volume Manager for Sun Cluster for the Oracle RAC Database" in *Sun Cluster Data Service for Oracle RAC Guide for Solaris OS*.

The following table lists the tasks that you perform to create disk sets. Complete the procedures in the order that is indicated.

TABLE 4–2 Task Map: Installing and Configuring Solaris Volume Manager Software

Task	Instructions
Create disk sets by using the `metaset` command.	"How to Create a Disk Set" on page 173

TABLE 4–2 Task Map: Installing and Configuring Solaris Volume Manager Software   *(Continued)*

Task	Instructions
Add drives to the disk sets.	"How to Add Drives to a Disk Set" on page 177
(*Optional*) Repartition drives in a disk set to allocate space to different slices.	"How to Repartition Drives in a Disk Set" on page 178
List DID pseudo-driver mappings and define volumes in the /etc/lvm/md.tab files.	"How to Create an md.tab File" on page 179
Initialize the md.tab files.	"How to Activate Volumes" on page 180

▼ How to Create a Disk Set

Perform this procedure to create disk sets.

Note – This procedure provides the long forms of the Sun Cluster commands. Most commands also have short forms. Except for the forms of the command names, the commands are identical. For a list of the commands and their short forms, see Appendix B, "Sun Cluster Object-Oriented Commands," in *Sun Cluster System Administration Guide for Solaris OS*.

---

1 **SPARC:** *(Solaris 9)* **Determine whether, after you create the new disk sets, the global cluster will have more than three disk sets.**

- If the cluster will have no more than three disk sets, skip to Step 9.

- If the cluster will have four or more disk sets, proceed to Step 2 to prepare the cluster. You must perform this task whether you are installing disk sets for the first time or whether you are adding more disk sets to a fully configured cluster.

- If the cluster runs on the Solaris 10 OS, Solaris Volume Manager automatically makes the necessary configuration changes. Skip to Step 9.

2 **On any node of the cluster, check the value of the** md_nsets **variable in the** /kernel/drv/md.conf **file.**

3 **If the total number of disk sets in the cluster will be greater than the existing value of** md_nsets **minus one, increase the value of** md_nsets **to the desired value.**

The maximum permissible number of disk sets is one less than the configured value of md_nsets. The maximum possible value of md_nsets is 32, therefore the maximum permissible number of disk sets that you can create is 31.

4 **Ensure that the** /kernel/drv/md.conf **file is identical on each node of the cluster.**

> ⚠️ **Caution** – Failure to follow this guideline can result in serious Solaris Volume Manager errors and possible loss of data.

5 If you made changes to the md.conf **file on any node, perform the following steps to make those changes active.**

a. On one node, become superuser.

b. From one node, shut down the cluster.

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

c. Reboot each node of the cluster.

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

```
ok boot
```

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

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

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

For more information about GRUB based booting, see Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration*.

6 On each node in the cluster, run the devfsadm**(1M) command.**

You can run this command on all nodes in the cluster at the same time.

7 From one node of the cluster, update the global-devices namespace.

```
phys-schost# cldevice populate
```

See the cldevice(1CL) man page for more information.

8   **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
```

9   **Ensure that the disk set that you intend to create meets 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 exactly two mediator hosts. These mediator hosts must be the same two hosts used for the disk set. See "Configuring Dual-String Mediators" on page 182 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).

10   **Ensure that the local state database replicas exist.**

For instructions, see "How to Create State Database Replicas" on page 160.

11   **Become superuser on the cluster node that will master the disk set.**

12   **Create the disk set.**

The following command creates the disk set and registers the disk set as a Sun 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 *Sun Cluster System Administration Guide for Solaris OS* for more information about how to change the numsecondaries property.

---

13   If you are configuring a replicated Solstice DiskSuite or 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 *Sun Cluster System Administration Guide for Solaris OS*.

14   Verify the status of the new disk set.

phys-schost# **metaset -s** *setname*

15   As needed, set device group properties.

phys-schost# **cldevicegroup set -p** *name*=*value* *devicegroup*

-p   Specifies a device-group property.

name   Specifies the name of a property.

value   Specifies the value or setting of the property.

devicegroup   Specifies the name of the device group. The device-group name is the same as the disk-set name.

See the cldevicegroup(1CL) for information about device-group properties.

Example 4–6   Creating a Disk Set

The following command creates two disk sets, dg-schost-1 and dg-schost-2, with the nodes phys-schost-1 and phys-schost-2 specified as the potential primaries.

```
phys-schost# metaset -s dg-schost-1 -a -h phys-schost-1 phys-schost-2
phys-schost# metaset -s dg-schost-2 -a -h phys-schost-1 phys-schost-2
```

Next Steps   Add drives to the disk set. Go to "Adding Drives to a Disk Set" on page 176.

Adding Drives to a Disk Set

When you add a drive to a disk set, the volume management software repartitions the drive as follows so that the state database for the disk set can be placed on the drive.

- A small portion of each drive is reserved for use by Solaris Volume Manager software. In volume table of contents (VTOC) labeled devices, slice 7 is used. In Extensible Firmware Interface (EFI) labeled devices, slice 6 is used. The remainder of the space on each drive is placed into slice 0.

- Drives are repartitioned when they are added to the disk set only if the target slice is not configured correctly.

- Any existing data on the drives is lost by the repartitioning.

- If the target slice starts at cylinder 0, and the drive partition is large enough to contain a state database replica, the drive is not repartitioned.

▼ How to Add Drives to a Disk Set

Before You Begin Ensure that the disk set has been created. For instructions, see "How to Create a Disk Set" on page 173.

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 4–7   Adding Drives to a Disk Set

The metaset command adds the drives /dev/did/rdsk/d1 and /dev/did/rdsk/d2 to the disk set dg-schost-1.

```
phys-schost# metaset -s dg-schost-1 -a /dev/did/rdsk/d1 /dev/did/rdsk/d2
```

Next Steps   To repartition drives for use in volumes, go to "How to Repartition Drives in a Disk Set" on page 178.

Otherwise, go to "How to Create an md.tab File" on page 179 to define metadevices or volumes by using an md.tab file.

▼ How to Repartition Drives in a Disk Set

The metaset(1M) command repartitions drives in a disk set so that a small portion of each drive is reserved for use by Solaris Volume Manager software. In volume table of contents (VTOC) labeled devices, slice 7 is used. In Extensible Firmware Interface (EFI) labeled devices, slice 6 is used. The remainder of the space on each drive is placed into slice 0. To make more effective use of the drive, use this procedure to modify the disk layout. If you allocate space to VTOC slices 1 through 6 or EFI slices 1 through 5, you can use these slices when you set up Solaris Volume Manager volumes.

1   Become superuser.

2   Use the format **command to change the disk partitioning for each drive in the disk set.**

When you repartition a drive, you must meet the following conditions to prevent the metaset(1M) command from repartitioning the drive.

- Create slice 7 for VTOC or slice 6 for EFI starting at cylinder 0, large enough to hold a state database replica. See your Solaris Volume Manager administration guide to determine the size of a state database replica for your version of the volume-manager software.

- Set the Flag field in the target slice to wu (read-write, unmountable). Do not set it to read-only.
- Do not allow the target slice to overlap any other slice on the drive.

See the format(1M) man page for details.

Next Steps Define volumes by using an md.tab file. Go to

▼ **How to Create an** md.tab **File**

Create an /etc/lvm/md.tab file on each node in the cluster. Use the md.tab file to define Solaris Volume Manager volumes for the disk sets that you created.

Note – If you are using local volumes, ensure that local volume names are distinct from the device-ID names that are used to form disk sets. For example, if the device-ID name /dev/did/dsk/d3 is used in a disk set, do not use the name /dev/md/dsk/d3 for a local volume. This requirement does not apply to shared volumes, which use the naming convention /dev/md/*setname*/{r}dsk/d#.

---

1 **Become superuser.**

2 **List the DID mappings for reference when you create your** md.tab **file.**

Use the full DID device names in the md.tab file in place of the lower-level device names (c*N*t*X*d*Y*). The DID device name takes the form /dev/did/rdsk/d*N*.

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

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

3 **Create an** /etc/lvm/md.tab **file and edit it with your preferred text editor.**

Note – If you have existing data on the drives that will be used for the submirrors, you must back up the data before volume setup. Then restore the data onto the mirror.

To avoid possible confusion between local volumes on different nodes in a cluster environment, use a naming scheme that makes each local volume name unique throughout the cluster. For example, for node 1 choose names from d100 to d199. And for node 2 use d200 to d299.

See your Solaris Volume Manager documentation and the md.tab(4) man page for details about how to create an md.tab file.

Example 4–8   Sample md.tab File

The following sample md.tab file defines the disk set that is named dg-schost-1. The ordering of lines in the md.tab file is not important.

```
dg-schost-1/d0 -m dg-schost-1/d10 dg-schost-1/d20
    dg-schost-1/d10 1 1 /dev/did/rdsk/d1s0
    dg-schost-1/d20 1 1 /dev/did/rdsk/d2s0
```

The sample md.tab file is constructed as follows.

1. The first line defines the device d0 as a mirror of volumes d10 and d20. The -m signifies that this device is a mirror device.

   ```
   dg-schost-1/d0 -m dg-schost-1/d0 dg-schost-1/d20
   ```

2. The second line defines volume d10, the first submirror of d0, as a one-way stripe.

   ```
   dg-schost-1/d10 1 1 /dev/did/rdsk/d1s0
   ```

3. The third line defines volume d20, the second submirror of d0, as a one-way stripe.

   ```
   dg-schost-1/d20 1 1 /dev/did/rdsk/d2s0
   ```

Next Steps   Activate the volumes that are defined in the md.tab files. Go to “How to Activate Volumes” on page 180.

▼ How to Activate Volumes

Perform this procedure to activate Solaris Volume Manager volumes that are defined in md.tab files.

1 **Become superuser.**

2 **Ensure that** md.tab **files are located in the** /etc/lvm **directory.**

3 **Ensure that you have ownership of the disk set on the node where the command will be executed.**

4 Take ownership of the disk set.

phys-schost# **cldevicegroup switch -n** *node* *devicegroup*

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

devicegroup   Specifies the disk set name.

5 Activate the disk set's volumes, which are defined in the md.tab **file.**

phys-schost# **metainit -s** *setname* **-a**

-s *setname*   Specifies the disk set name.

-a   Activates all volumes in the md.tab file.

6 Repeat Step 3 through Step 5 for each disk set in the cluster.

If necessary, run the metainit(1M) command from another node that has connectivity to the drives. This step is required for cluster-pair topologies, where the drives are not accessible by all nodes.

7 Check the status of the volumes.

phys-schost# **metastat -s** *setname*

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

8 (Optional) Capture the disk partitioning information for future reference.

phys-schost# **prtvtoc /dev/rdsk/c***N***t***X***d***Y***s***Z* **>** *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 *Sun Cluster System Administration Guide for Solaris OS*.

Example 4–9   Activating Volumes in the md.tab File

In the following example, all volumes that are defined in the md.tab file for disk set dg-schost-1 are activated.

phys-schost# **metainit -s dg-schost-1 -a**

Next Steps   If your cluster contains disk sets that are configured with exactly two disk enclosures and two nodes, add dual-string mediators. Go to .

Otherwise, go to to create a cluster file system.

Configuring Dual-String Mediators

This section provides information and procedures to configure dual-string mediator hosts. Dual-string mediators are required for all Solaris Volume Manager disk sets that are configured with exactly two disk strings and two cluster nodes. The use of mediators enables the Sun 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.

A *disk string* consists of a disk enclosure, its physical drives, cables from the enclosure to the node or nodes, and the interface adapter cards.

The following table lists the tasks that you perform to configure dual-string mediator hosts. complete the procedures in the order that is indicated.

TABLE 4–3   Task Map: Installing and Configuring Solaris Volume Manager Software

Task	Instructions
Configure dual-string mediator hosts.	"Requirements for Dual-String Mediators" on page 182 "How to Add Mediator Hosts" on page 183
Check the status of mediator data.	"How to Check the Status of Mediator Data" on page 183
If necessary, fix bad mediator data.	"How to Fix Bad Mediator Data" on page 184

Requirements for Dual-String Mediators

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

- Disk sets must be configured with exactly two mediator hosts. Those two mediator hosts must be the same two cluster nodes that are used for the disk set.

- A disk set cannot have more than two mediator hosts.

■ 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 must have exactly two nodes. Rather, only those disk sets that have two disk strings must be connected to exactly 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 Become superuser on the node that currently masters the disk set to which you intend to add mediator hosts.

2 Add each node with connectivity to the disk set as a mediator host for that disk set.

phys-schost# **metaset -s** *setname* **-a -m** *mediator-host-list*

-s setname	Specifies the disk set name.
-a	Adds to the disk set.
-m mediator-host-list	Specifies the name of the node to add as a mediator host for the disk set.

See the mediator(7D) man page for details about mediator-specific options to the metaset command.

Example 4–10 Adding Mediator Hosts

The following example adds the nodes phys-schost-1 and phys-schost-2 as mediator hosts for the disk set dg-schost-1. Both commands are run from the node 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
```

Next Steps Check the status of mediator data. Go to "How to Check the Status of Mediator Data" on page 183.

▼ How to Check the Status of Mediator Data

Before You Begin Ensure that you have added mediator hosts as described in "How to Add Mediator Hosts" on page 183.

1 Display the status of the mediator data.

phys-schost# **medstat -s** *setname*

-s setname   Specifies the disk set name.

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

2 If Bad **is the value in the Status field of the** medstat **output, repair the affected mediator host.**

Go to "How to Fix Bad Mediator Data" on page 184.

Next Steps   Go to "How to Create Cluster File Systems" on page 205 to create a cluster file system.

▼ How to Fix Bad Mediator Data

Perform this procedure to repair bad mediator data.

1 Identify all mediator hosts with bad mediator data as described in the procedure "How to Check the Status of Mediator Data" on page 183.

2 Become superuser on the node that owns the affected disk set.

3 Remove all mediator hosts with bad mediator data from all affected disk sets.

phys-schost# **metaset -s** *setname* **-d -m** *mediator-host-list*

-s setname   Specifies the disk set name.

-d   Deletes from the disk set.

-m mediator-host-list   Specifies the name of the node to remove as a mediator host for the disk set.

4 Restore each mediator host that you removed in Step 3.

phys-schost# **metaset -s** *setname* **-a -m** *mediator-host-list*

-a   Adds to the disk set.

-m mediator-host-list   Specifies the name of the node to add as a mediator host for the disk set.

See the mediator(7D) man page for details about mediator-specific options to the metaset command.

Next Steps   Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- To create cluster file systems, go to "How to Create Cluster File Systems" on page 205.

- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 210.

- SPARC: To configure Sun Management Center to monitor the cluster, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 227.

- 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 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

5

Installing and Configuring Veritas Volume Manager

Install and configure your local and multihost disks for Veritas Volume Manager (VxVM) by using the procedures in this chapter, along with the planning information in "Planning Volume Management" on page 46. See your VxVM documentation for additional details.

The following sections are in this chapter:

- "Installing and Configuring VxVM Software" on page 187
- "Creating Disk Groups in a Cluster" on page 194
- "Unencapsulating the Root Disk" on page 201

Installing and Configuring VxVM Software

This section provides information and procedures to install and configure VxVM software on a Sun Cluster configuration.

The following table lists the tasks to perform to install and configure VxVM software for Sun Cluster configurations. Complete the procedures in the order that is indicated.

TABLE 5–1 Task Map: Installing and Configuring VxVM Software

Task	Instructions
Plan the layout of your VxVM configuration.	"Planning Volume Management" on page 46
(Optional) Determine how you will create the root disk group on each node.	"Setting Up a Root Disk Group Overview" on page 188
Install VxVM software.	"How to Install Veritas Volume Manager Software" on page 189
	VxVM installation documentation

TABLE 5–1   Task Map: Installing and Configuring VxVM Software   *(Continued)*

Task	Instructions
(Optional) Create a root disk group. You can either encapsulate the root disk or create the root disk group on local, nonroot disks.	"SPARC: How to Encapsulate the Root Disk" on page 190 "How to Create a Root Disk Group on a Nonroot Disk" on page 191
(Optional) Mirror the encapsulated root disk.	"How to Mirror the Encapsulated Root Disk" on page 192
Create disk groups.	"Creating Disk Groups in a Cluster" on page 194

Setting Up a Root Disk Group Overview

The creation of a root disk group is optional. If you do not intend to create a root disk group, proceed to "How to Install Veritas Volume Manager Software" on page 189.

- Access to a node's root disk group must be restricted to only that node.
- Remote nodes must never access data stored in another node's root disk group.
- Do not use the cldevicegroup command to register the root disk group as a device group.
- Whenever possible, configure the root disk group for each node on a nonshared disk.

Sun Cluster software supports the following methods to configure the root disk group.

- **Encapsulate the node's root disk** – This method enables the root disk to be mirrored, which provides a boot alternative if the root disk is corrupted or damaged. To encapsulate the root disk you need two free disk slices as well as free cylinders, preferably at the beginning or the end of the disk.
- **Use local nonroot disks** – This method provides an alternative to encapsulating the root disk. If a node's root disk is encapsulated, certain tasks you might later perform, such as upgrade the Solaris OS or perform disaster recovery procedures, could be more complicated than if the root disk is not encapsulated. To avoid this potential added complexity, you can instead initialize or encapsulate local nonroot disks for use as root disk groups.

  A root disk group that is created on local nonroot disks is local to that node, neither globally accessible nor highly available. As with the root disk, to encapsulate a nonroot disk you need two free disk slices as well as free cylinders at the beginning or the end of the disk.

See your VxVM installation documentation for more information.

▼ How to Install Veritas Volume Manager Software

Perform this procedure to install Veritas Volume Manager (VxVM) software on each global-cluster node that you want to install with VxVM. You can install VxVM on all nodes of the cluster, or install VxVM just on the nodes that are physically connected to the storage devices that VxVM will manage.

Before You Begin   Perform the following tasks:

- Ensure that all nodes in the cluster are running in cluster mode.
- Obtain any Veritas Volume Manager (VxVM) license keys that you need to install.
- Have available your VxVM installation documentation.

1   **Become superuser on a cluster node that you intend to install with VxVM.**

2   **Insert the VxVM CD-ROM in the CD-ROM drive on the node.**

3   **Follow procedures in your VxVM installation guide to install and configure VxVM software and licenses.**

4   **Run the** clvxvm **utility in noninteractive mode.**

    phys-schost# **clvxvm initialize**

The clvxvm utility performs necessary postinstallation tasks. The clvxvm utility also selects and configures a cluster-wide vxio driver major number. See the clvxvm(1CL) man page for more information.

5   **SPARC: To enable the VxVM cluster feature, supply the cluster feature license key, if you did not already do so.**

See your VxVM documentation for information about how to add a license.

6   **(Optional) Install the VxVM GUI.**

See your VxVM documentation for information about installing the VxVM GUI.

7   **Eject the CD-ROM.**

8   **Install any VxVM patches to support Sun Cluster software.**

See "Patches and Required Firmware Levels" in *Sun Cluster Release Notes* for the location of patches and installation instructions.

9   **Repeat Step 1 through Step 8 to install VxVM on any additional nodes.**

> **Note –** SPARC: To enable the VxVM cluster feature, you *must* install VxVM on all nodes of the cluster.

10 **If you do *not* install one or more nodes with VxVM, modify the** /etc/name_to_major **file on each non-VxVM node.**

    a. **On a node that is installed with VxVM, determine the** vxio **major number setting.**

       `phys-schost# `**`grep vxio /etc/name_to_major`**

    b. **Become superuser on a node that you do *not* intend to install with VxVM.**

    c. **Edit the** /etc/name_to_major **file and add an entry to set the** vxio **major number to** *NNN*, **the number derived in Step a.**

       `phys-schost# `**`vi /etc/name_to_major`**
       `vxio `*NNN*

    d. **Initialize the** vxio **entry.**

       `phys-schost# `**`drvconfig -b -i vxio -m `*NNN*

    e. **Repeat Step a through Step d on all other nodes that you do *not* intend to install with VxVM.**

       When you finish, each node of the cluster should have the same vxio entry in its /etc/name_to_major file.

11 **To create a root disk group, go to "SPARC: How to Encapsulate the Root Disk" on page 190 or "How to Create a Root Disk Group on a Nonroot Disk" on page 191.**

    Otherwise, proceed to Step 12.

> **Note –** A root disk group is optional.

12 **Reboot each node on which you installed VxVM.**

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

Next Steps   To create a root disk group, go to "SPARC: How to Encapsulate the Root Disk" on page 190 or "How to Create a Root Disk Group on a Nonroot Disk" on page 191.

    Otherwise, create disk groups. Go to "Creating Disk Groups in a Cluster" on page 194.

▼ SPARC: How to Encapsulate the Root Disk

Perform this procedure to create a root disk group by encapsulating the root disk. Root disk groups are optional. See your VxVM documentation for more information.

Note – If you want to create a root disk group on nonroot disks, instead perform procedures in "How to Create a Root Disk Group on a Nonroot Disk" on page 191.

Before You Begin Ensure that you have installed VxVM as described in "How to Install Veritas Volume Manager Software" on page 189.

1 Become superuser on a node that you installed with VxVM.

2 Encapsulate the root disk.

phys-schost# **clvxvm encapsulate**

See the clvxvm(1CL) man page for more information.

3 Repeat for any other node on which you installed VxVM.

Next Steps To mirror the encapsulated root disk, go to "How to Mirror the Encapsulated Root Disk" on page 192.

Otherwise, go to "Creating Disk Groups in a Cluster" on page 194.

▼ How to Create a Root Disk Group on a Nonroot Disk

Use this procedure to create a root disk group by encapsulating or initializing local disks other than the root disk. The creation of a root disk group is optional.

Note – If you want to create a root disk group on the root disk, instead perform procedures in "SPARC: How to Encapsulate the Root Disk" on page 190.

Before You Begin If the disks are to be encapsulated, ensure that each disk has at least two slices with 0 cylinders. If necessary, use the format(1M) command to assign 0 cylinders to each VxVM slice.

1 Become superuser.

2 Start the vxinstall **utility.**

phys-schost# **vxinstall**

3 When prompted by the vxinstall **utility, make the following choices or entries.**

- SPARC: To enable the VxVM cluster feature, supply the cluster feature license key.
- Choose Custom Installation.

- Do not encapsulate the boot disk.
- Choose any disks to add to the root disk group.
- Do not accept automatic reboot.

4 If the root disk group that you created contains one or more disks that connect to more than one node, ensure that fencing is disabled for such disks.

Use the following command to disable fencing for each shared disk in the root disk group.

```
phys-schost# cldevice set -p default_fencing=nofencing device
```

-p   Specifies a device property.

default_fencing=nofencing   Disables fencing for the specified device.

Disabling fencing for the device prevents unintentional fencing of the node from the disk that is used by the root disk group if that disk is connected to multiple nodes.

For more information about the default_fencing property, see the cldevice(1CL) man page.

5 Evacuate any resource groups or device groups from the node.

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

from-node   Specifies the name of the node from which to move resource or device groups.

6 Reboot the node.

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

7 Use the vxdiskadm command to add multiple disks to the root disk group.

The root disk group becomes tolerant of a disk failure when it contains multiple disks. See VxVM documentation for procedures.

Next Steps Create disk groups. Go to "Creating Disk Groups in a Cluster" on page 194.

▼ How to Mirror the Encapsulated Root Disk

After you install VxVM and encapsulate the root disk, perform this procedure on each node on which you mirror the encapsulated root disk.

Before You Begin Ensure that you have encapsulated the root disk as described in "SPARC: How to Encapsulate the Root Disk" on page 190.

1 Become superuser.

2   List the devices.

phys-schost# **cldevice list -v**

Output looks similar to 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
```

3   Mirror the encapsulated root disk.

Follow the procedures in your VxVM documentation.

For maximum availability and simplified administration, use a local disk for the mirror. See for additional guidelines.

Caution – Do not use a quorum device to mirror a root disk. Using a quorum device to mirror a root disk might prevent the node from booting from the root-disk mirror under certain circumstances.

4   View the node list of the raw-disk device group for the device that you used to mirror the root disk.

The name of the device group is the form dsk/dN, where dN is the DID device name.

phys-schost# **cldevicegroup list -v dsk/d**N

-v   Displays verbose output.

Output looks similar to the following.

```
Device group      Type              Node list
------------      ----              ---------
dsk/dN            Local_Disk         phys-schost-1, phys-schost-3
```

5   If the node list contains more than one node name, remove from the node list all nodes except the node whose root disk you mirrored.

Only the node whose root disk you mirrored should remain in the node list for the raw-disk device group.

phys-schost# **cldevicegroup remove-node -n** *node* **dsk/d**N

-n *node*   Specifies the node to remove from the device-group node list.

6   **Disable fencing for all disks in the raw-disk device group that connect to more than one node.**

Disabling fencing for a device prevents unintentional fencing of the node from its boot device if the boot device is connected to multiple nodes.

phys-schost# **cldevice set -p default_fencing=nofencing** *device*

-p   Sets the value of a device property.

default_fencing=nofencing   Disables fencing for the specified device.

For more information about the default_fencing property, see the cldevice(1CL) man page.

7   **Repeat this procedure for each node in the cluster whose encapsulated root disk you want to mirror.**

Example 5–1   Mirroring the Encapsulated Root Disk

The following example shows a mirror created of the root disk for the node phys-schost-1. The mirror is created on the disk c0t0d0, whose raw-disk device-group name is dsk/d2. Disk c0t0d0 is a multihost disk, so the node phys-schost-3 is removed from the disk's node list and fencing is disabled.

```
phys-schost# cldevice list -v
DID Device        Full Device Path
----------        ---------------
d2                pcircinus1:/dev/rdsk/c0t0d0
...
    Create the mirror by using VxVM procedures
phys-schost# cldevicegroup list -v dsk/d2
Device group      Type               Node list
------------      ----               ---------
dsk/d2            Local_Disk         phys-schost-1, phys-schost-3
phys-schost# cldevicegroup remove-node -n phys-schost-3 dsk/d2
phys-schost# cldevice set -p default_fencing=nofencing c0t0d0
```

Next Steps   Create disk groups. Go to "Creating Disk Groups in a Cluster" on page 194.

Creating Disk Groups in a Cluster

This section describes how to create VxVM disk groups in a cluster. The following table describes the types of VxVM disk groups you can configure in a Sun Cluster configuration and their characteristics.

Disk Group Type	Use	Registered with Sun Cluster?	Storage Requirement
VxVM disk group	Device groups for failover or scalable data services, global devices, or cluster file systems	Yes	Shared storage
Local VxVM disk group	Applications that are not highly available and are confined to a single node	No	Shared or unshared storage
VxVM shared disk group	Oracle Real Application Clusters (also requires the VxVM cluster feature)	No	Shared storage

The following table lists the tasks to perform to create VxVM disk groups in a Sun Cluster configuration. complete the procedures in the order that is indicated.

TABLE 5–2   Task Map: Creating VxVM Disk Groups

Task	Instructions
Create disk groups and volumes.	"How to Create a Disk Group" on page 195
Register as Sun Cluster device groups those disk groups that are not local and that do not use the VxVM cluster feature.	"How to Register a Disk Group" on page 197
If necessary, resolve any minor-number conflicts between device groups by assigning a new minor number.	"How to Assign a New Minor Number to a Device Group" on page 198
Verify the disk groups and volumes.	"How to Verify the Disk Group Configuration" on page 199

▼ How to Create a Disk Group

Use this procedure to create your VxVM disk groups and volumes.

Perform this procedure from a node that is physically connected to the disks that make the disk group that you add.

Before You Begin Perform the following tasks:

- Make mappings of your storage disk drives. See the appropriate manual in the *Sun Cluster Hardware Administration Collection* to perform an initial installation of your storage device.

- Complete the following configuration planning worksheets.
    - "Local File System Layout Worksheet" on page 247
    - "Device Group Configurations Worksheet" on page 251
    - "Volume-Manager Configurations Worksheet" on page 253

    See "Planning Volume Management" on page 46 for planning guidelines.

■ If you did not create root disk groups, ensure that you have rebooted each node on which you installed VxVM, as instructed in Step 12 of "How to Install Veritas Volume Manager Software" on page 189.

1   Become superuser on the node that will own the disk group.

2   Create the VxVM disk groups and volumes.

Observe the following special instructions:

■ SPARC: If you are installing Oracle Real Application Clusters, create shared VxVM disk groups by using the cluster feature of VxVM. Observe guidelines and instructions in "How to Create a VxVM Shared-Disk Group for the Oracle RAC Database" in *Sun Cluster Data Service for Oracle RAC Guide for Solaris OS* and in the *Veritas Volume Manager Administrator's Reference Guide*.

■ Otherwise, create VxVM disk groups by using the standard procedures that are documented in the VxVM documentation.

Note – You can use Dirty Region Logging (DRL) to decrease volume recovery time if a node failure occurs. However, DRL might decrease I/O throughput.

---

3   For local disk groups, set the localonly **property and add a single node to the disk group's node list.**

Note – A disk group that is configured to be local only is not highly available or globally accessible.

---

a.   Start the clsetup **utility.**

```
phys-schost# clsetup
```

b.   Choose the menu item, Device groups and volumes.

c.   Choose the menu item, Set localonly on a VxVM disk group.

d.   Follow the instructions to set the localonly **property and to specify the single node that will exclusively master the disk group.**

Only one node at any time is permitted to master the disk group. You can later change which node is the configured master.

e.   When finished, quit the clsetup **utility.**

Next Steps   Determine your next step:

- SPARC: If the VxVM cluster feature is enabled, go to "How to Verify the Disk Group Configuration" on page 199.

- If you created disk groups that are not local and the VxVM cluster feature is *not* enabled, register the disk groups as Sun Cluster device groups. Go to "How to Register a Disk Group" on page 197.

- If you created only local disk groups, go to "How to Verify the Disk Group Configuration" on page 199.

▼ How to Register a Disk Group

If the VxVM cluster feature is *not* enabled, perform this procedure to register disk groups that are not local as Sun Cluster device groups.

Note – SPARC: If the VxVM cluster feature is enabled or you created a local disk group, do not perform this procedure. Instead, proceed to "How to Verify the Disk Group Configuration" on page 199.

---

1 Become superuser on a node of the cluster.

2 Register the global disk group as a Sun Cluster device group.

a. Start the clsetup **utility.**

```
phys-schost# clsetup
```

b. Choose the menu item, Device groups and volumes.

c. Choose the menu item, Register a VxVM disk group.

d. Follow the instructions to specify the VxVM disk group that you want to register as a Sun Cluster device group.

e. When finished, quit the clsetup **utility.**

f. Deport and re-import each local disk group.

```
phys-schost# vxdg deport diskgroup
# vxdg import dg
```

g. Restart each local disk group.

```
phys-schost# vxvol -g diskgroup startall
```

h. Verify the local-only status of each local disk group.

If the value of the flags property of the disk group is `nogdl`, the disk group is correctly configured for local-only access.

```
phys-schost# vxdg list diskgroup | grep flags
flags: nogdl
```

3   **Verify that the device group is registered.**

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

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

Next Steps   Go to .

Troubleshooting   **Stack overflow** – If a stack overflows when the device group is brought online, the default value of the thread stack size might be insufficient. On each node, add the entry `set cl_haci:rm_thread_stacksize=0x`*size* to the `/etc/system` file, where *size* is a number greater than 8000, which is the default setting.

Configuration changes – If you change any configuration information for a VxVM device group or its volumes, you must register the configuration changes by using the `clsetup` utility. Configuration changes that you must register include adding or removing volumes and changing the group, owner, or permissions of existing volumes. See "Administering Device Groups" in *Sun Cluster System Administration Guide for Solaris OS* for procedures to register configuration changes that are made to a VxVM device group.

▼ How to Assign a New Minor Number to a Device Group

If device group registration fails because of a minor-number conflict with another disk group, you must assign the new disk group a new, unused minor number. Perform this procedure to reminor a disk group.

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

2   **Determine the minor numbers in use.**

```
phys-schost# ls -l /global/.devices/node@1/dev/vx/dsk/*
```

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

4   **Assign the new base minor number to the disk group.**

```
phys-schost# vxdg reminor diskgroup base-minor-number
```

Example 5–2 How to Assign a New Minor Number to a Device Group

This example uses the minor numbers 16000-16002 and 4000-4001. The vxdg reminor command reminors the new device group to use the base minor number 5000.

```
phys-schost# ls -l /global/.devices/node@1/dev/vx/dsk/*
/global/.devices/node@1/dev/vx/dsk/dg1
brw-------  1 root     root      56,16000 Oct  7 11:32 dg1v1
brw-------  1 root     root      56,16001 Oct  7 11:32 dg1v2
brw-------  1 root     root      56,16002 Oct  7 11:32 dg1v3

/global/.devices/node@1/dev/vx/dsk/dg2
brw-------  1 root     root      56,4000 Oct  7 11:32 dg2v1
brw-------  1 root     root      56,4001 Oct  7 11:32 dg2v2
phys-schost# vxdg reminor dg3 5000
```

Next Steps Register the disk group as a Sun Cluster device group. Go to "How to Register a Disk Group" on page 197.

▼ How to Verify the Disk Group Configuration

Perform this procedure on each node of the cluster.

1 Become superuser.

2 List the disk groups.
```
phys-schost# vxdisk list
```

3 List the device groups.
```
phys-schost# cldevicegroup list -v
```

4 Verify that all disk groups are correctly configured.
Ensure that the following requirements are met:

- The root disk group includes only local disks.
- All disk groups and any local disk groups are imported on the current primary node only.

5 Verify that all volumes have been started.
```
phys-schost# vxprint
```

6 Verify that all disk groups have been registered as Sun Cluster device groups and are online.
```
phys-schost# cldevicegroup status
```

Output should not display any local disk groups.

7 (Optional) Capture the disk partitioning information for future reference.

phys-schost# **prtvtoc /dev/rdsk/c**N**t**X**d**Y**s**Z **>** *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.

8 (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 *Sun Cluster System Administration Guide for Solaris OS*.

More Information Guidelines for Administering VxVM Disk Groups

Observe the following guidelines for administering VxVM disk groups in a Sun Cluster configuration:

- **VxVM device groups** – VxVM disk groups that have been registered as device groups are managed by Sun Cluster software. After a disk group is registered as a device group, you should never import or deport that VxVM disk group by using VxVM commands. The Sun Cluster software can handle all cases where device groups need to be imported or deported. See "Administering Device Groups" in *Sun Cluster System Administration Guide for Solaris OS* for procedures about how to manage device groups.

- **Local disk groups** – Local VxVM disk groups are not managed by Sun Cluster software. Use VxVM commands to administer local disk groups as you would in a nonclustered system.

Troubleshooting If the output of the cldevicegroup status command includes any local disk groups, the displayed disk groups are not configured correctly for local-only access. Return to "How to Create a Disk Group" on page 195 to reconfigure the local disk group.

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

- To create cluster file systems, go to "How to Create Cluster File Systems" on page 205.

- To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 210.

- SPARC: To configure Sun Management Center to monitor the cluster, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 227.

- 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 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

Unencapsulating the Root Disk

This section describes how to unencapsulate the root disk in a Sun Cluster configuration.

▼ How to Unencapsulate the Root Disk

Perform this procedure to unencapsulate the root disk.

Before You Begin Perform the following tasks:

- Ensure that only Solaris root file systems are present on the root disk. The Solaris root file systems are root (/), swap, the global devices namespace, /usr, /var, /opt, and /home.
- Back up and remove from the root disk any file systems other than Solaris root file systems that reside on the root disk.

1 **Become superuser on the node that you intend to unencapsulate.**

2 **Evacuate all resource groups and device groups from the node.**

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

from-node   Specifies the name of the node from which to move resource or device groups.

3 **Determine the node-ID number of the node.**

```
phys-schost# clinfo -n
```

4 **Unmount the global-devices file system for this node, where N is the node ID number that is returned in Step 3.**

```
phys-schost# umount /global/.devices/node@N
```

5 **View the** /etc/vfstab **file and determine which VxVM volume corresponds to the global-devices file system.**

```
phys-schost# vi /etc/vfstab
#device         device         mount   FS     fsck    mount    mount
#to mount       to fsck        point   type   pass    at boot  options
#
#NOTE: volume rootdiskxNvol (/global/.devices/node@N) encapsulated
#partition cNtXdYsZ
```

6 **Remove from the root disk group the VxVM volume that corresponds to the global-devices file system.**

```
phys-schost# vxedit -g rootdiskgroup -rf rm rootdiskxNvol
```

> ⚠️ **Caution –** Do not store data other than device entries for global devices in the global-devices file system. All data in the global-devices file system is destroyed when you remove the VxVM volume. Only data that is related to global devices entries is restored after the root disk is unencapsulated.

7   Unencapsulate the root disk.

> **Note –** Do **not** accept the shutdown request from the command.

```
phys-schost# /etc/vx/bin/vxunroot
```

See your VxVM documentation for details.

8   Use the format(1M) **command to add a 512-Mbyte partition to the root disk to use for the global-devices file system.**

> **Tip –** Use the same slice that was allocated to the global-devices file system before the root disk was encapsulated, as specified in the /etc/vfstab file.

9   Set up a file system on the partition that you created in Step 8.

```
phys-schost# newfs /dev/rdsk/cNtXdYsZ
```

10   Determine the DID name of the root disk.

```
phys-schost# cldevice list cNtXdY
dN
```

11   In the /etc/vfstab **file, replace the path names in the global-devices file system entry with the DID path that you identified in Step 10.**

The original entry would look similar to the following.

```
phys-schost# vi /etc/vfstab
/dev/vx/dsk/rootdiskxNvol /dev/vx/rdsk/rootdiskxNvol /global/.devices/node@N ufs 2 no global
```

The revised entry that uses the DID path would look similar to the following.

```
/dev/did/dsk/dNsX /dev/did/rdsk/dNsX /global/.devices/node@N ufs 2 no global
```

12   Mount the global-devices file system.

```
phys-schost# mount /global/.devices/node@N
```

13   From one node of the cluster, repopulate the global-devices file system with device nodes for any raw-disk devices and Solaris Volume Manager devices.

phys-schost# **cldevice populate**

VxVM devices are recreated during the next reboot.

14   On each node, verify that the cldevice populate **command has completed processing before you proceed to the next step.**

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**

15   Reboot the node.

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

16   Repeat this procedure on each node of the cluster to unencapsulate the root disk on those nodes.

Creating Cluster File Systems, Non-Global Zones, and Zone Clusters

This chapter describes the following topics:

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.

Note – Alternatively, you can use a highly available local file system to support a data service. For information about choosing between creating a cluster file system or a highly available local file system to support a particular data service, see the manual for that data service. For general information about creating a highly available local file system, see "Enabling Highly Available Local File Systems" in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

You cannot add a cluster file system to a zone cluster.

Before You Begin   Perform the following tasks:

- Ensure that you installed software packages for the Solaris OS, Sun Cluster framework, and other products as described in "Installing the Software" on page 55.

- 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 82.

- 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 157 or "Installing and Configuring VxVM Software" on page 187.

Note – If you added a new node to a cluster that uses VxVM, you must perform one of the following tasks:

- Install VxVM on that node.
- Modify that node's /etc/name_to_major file to support coexistence with VxVM.

Follow the procedures in "How to Install Veritas Volume Manager Software" on page 189 to perform one of these required tasks.

- Determine the mount options to use for each cluster file system that you want to create. See "Choosing Mount Options for Cluster File Systems" on page 44.

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

For Solaris, you must perform this procedure from the global zone if non-global zones are configured in the cluster.

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

---

2 **Create a file system.**

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

- **For a UFS file system, use the** newfs**(1M) command.**

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

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

Volume Manager	Sample Disk Device Name	Description
Solaris Volume Manager	/dev/md/nfs/rdsk/d1	Raw disk device d1 within the nfs disk set

Volume Manager	Sample Disk Device Name	Description
Veritas Volume Manager	/dev/vx/rdsk/oradg/vol01	Raw disk device vol01 within the oradg disk group
None	/dev/global/rdsk/d1s3	Raw disk device d1s3

- **SPARC: For a Veritas File System (VxFS) file system, follow the procedures that are provided in your VxFS documentation.**

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

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

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

```
phys-schost# mkdir -p /global/device-group/mountpoint/
```

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

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

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

See the vfstab(4) man page for details.

Note – If non-global zones are configured in the cluster, ensure that you mount cluster file systems in the global zone on a path in the global zone's root directory.

---

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

Note – Do *not* use the logging mount option for Solaris Volume Manager transactional volumes. Transactional volumes provide their own logging.

In addition, Solaris Volume Manager transactional-volume logging is removed from the Solaris 10 OS. Solaris UFS logging provides the same capabilities but superior performance, as well as lower system administration requirements and overhead.

---

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

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

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

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

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

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

   `phys-schost#` **`sccheck`**

   The configuration check utility verifies that the mount points exist. The utility also verifies that `/etc/vfstab` file entries are correct on all nodes of the cluster. If no errors occur, nothing is returned.

   For more information, see the `sccheck(1M)` man page.

6 Mount the cluster file system.

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

   ■ **For UFS, mount the cluster file system from any node in the cluster.**

   ■ **SPARC: For VxFS, mount the cluster file system from the current master of** *device-group* **to ensure that the file system mounts successfully.**

     In addition, unmount a VxFS file system from the current master of *device-group* to ensure that the file system unmounts successfully.

     Note – To manage a VxFS cluster file system in a Sun Cluster environment, run administrative commands only from the primary node on which the VxFS cluster file system is mounted.

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

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

   For the Solaris 10 OS, cluster file systems are accessible from both the global zone and the non-global zone.

Example 6–1 Creating a 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 sccheck command is run. After configuration check processing is completes successfully, the cluster file system is mounted from one node and verified on all nodes.

```
phys-schost# newfs /dev/md/oracle/rdsk/d1
...
phys-schost# mkdir -p /global/oracle/d1
phys-schost# vi /etc/vfstab
#device         device         mount FS     fsck    mount   mount
#to mount       to fsck        point type   pass    at boot options
#
/dev/md/oracle/dsk/d1 /dev/md/oracle/rdsk/d1 /global/oracle/d1 ufs 2 yes global,logging
...
phys-schost# sccheck
phys-schost# mount /global/oracle/d1
phys-schost# mount
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/largefiles
on Sun Oct 3 08:56:16 2005
```

Next Steps Determine from the following list the next task to perform that applies to your cluster configuration. If you need to perform more than one task from this list, go to the first of those tasks in this list.

■ To create non-global zones on a node, go to "How to Create a Non-Global Zone on a Global-Cluster Node" on page 210.

■ SPARC: To configure Sun Management Center to monitor the cluster, go to "SPARC: Installing the Sun Cluster Module for Sun Management Center" on page 227.

■ 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 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

Configuring a Non-Global Zone on a Global-Cluster Node

This section provides procedures to create a non-global zone on a global-cluster node.

▼ How to Create a Non-Global Zone on a Global-Cluster Node

Perform this procedure for each non-global zone that you create in the global cluster.

Note – For complete information about installing a zone, refer to *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

You can configure a Solaris 10 non-global zone, simply referred to as a zone, on a cluster node while the node is booted in either cluster mode or in noncluster mode.

- If you create a zone while the node is booted in noncluster mode, the cluster software discovers the zone when the node joins the cluster.

- If you create or remove a zone while the node is in cluster mode, the cluster software dynamically changes its list of zones that can master resource groups.

Before You Begin   Perform the following tasks:

- Plan your non-global zone configuration. Observe the requirements and restrictions in "Guidelines for Non-Global Zones in a Global Cluster" on page 22.

- Have available the following information:

  - The total number of non-global zones that you will create.

  - The public adapter and public IP address that each zone will use.

  - The zone path for each zone. This path must be a local file system, not a cluster file system or a highly available local file system.

  - One or more devices that should appear in each zone.

  - (Optional) The name that you will assign each zone.

- If you will assign the zone a private IP address, ensure that the cluster IP address range can support the additional private IP addresses that you will configure. Use the `cluster show-netprops` command to display the current private-network configuration.

  If the current IP address range is not sufficient to support the additional private IP addresses that you will configure, follow the procedures in "How to Change the Private Network Configuration When Adding Nodes or Private Networks" on page 120 to reconfigure the private IP-address range.

For additional information, see "Zone Components" in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

1   Become superuser on the global-cluster node where you are creating the non-voting node.

You must be working in the global zone.

2   For the Solaris 10 OS, 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 becomes online before you proceed to the next step.

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

3   Configure, install, and boot the new zone.

Note – You must set the autoboot property to true to support resource-group functionality in the non-voting node on the global cluster.

Follow procedures in the Solaris documentation:

a.   Perform procedures in Chapter 18, "Planning and Configuring Non-Global Zones (Tasks)," in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

b.   Perform procedures in "Installing and Booting Zones" in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

c.   Perform procedures in "How to Boot a Zone" in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

4   Verify that the zone is in the ready **state.**

```
phys-schost# zoneadm list -v
ID  NAME     STATUS      PATH
 0  global   running     /
 1  my-zone  ready       /zone-path
```

5   For a whole-root zone with the ip-type **property set to** exclusive**: If the zone might host a logical-hostname resource, configure a file system resource that mounts the method directory from the global zone.**

```
phys-schost# zonecfg -z sczone
zonecfg:sczone> add fs
zonecfg:sczone:fs> set dir=/usr/cluster/lib/rgm
zonecfg:sczone:fs> set special=/usr/cluster/lib/rgm
zonecfg:sczone:fs> set type=lofs
zonecfg:sczone:fs> end
zonecfg:sczone> exit
```

6 (Optional) For a shared-IP zone, assign a private IP address and a private hostname to the zone.

The following command chooses and assigns an available IP address from the cluster's private IP-address range. The command also assigns the specified private hostname, or host alias, to the zone and maps it to the assigned private IP address.

```
phys-schost# clnode set -p zprivatehostname=hostalias node:zone
```

-p	Specifies a property.
zprivatehostname=*hostalias*	Specifies the zone private hostname, or host alias.
node	The name of the node.
zone	The name of the global-cluster non-voting node.

7 Perform the initial internal zone configuration.

Follow the procedures in "Performing the Initial Internal Zone Configuration" in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*. Choose either of the following methods:

- Log in to the zone.
- Use an /etc/sysidcfg file.

8 In the non-voting node, modify the nsswitch.conf **file.**

These changes enable the zone to resolve searches for cluster-specific hostnames and IP addresses.

a. Log in to the zone.

```
phys-schost# zlogin -c zonename
```

b. Open the /etc/nsswitch.conf **file for editing.**

```
sczone# vi /etc/nsswitch.conf
```

c. Add the cluster **switch to the beginning of the lookups for the** hosts **and** netmasks **entries, followed by the** files **switch.**

The modified entries should appear similar to the following:

```
...
hosts:      cluster files nis [NOTFOUND=return]
...
netmasks:   cluster files nis [NOTFOUND=return]
...
```

d. For all other entries, ensure that the files **switch is the first switch that is listed in the entry.**

e. Exit the zone.

9   **If you created an exclusive-IP zone, configure IPMP groups in each** /etc/hostname.*interface* **file that is on the zone.**

You must configure an IPMP group for each public-network adapter that is used for data-service traffic in the zone. This information is not inherited from the global zone. See "Public Networks" on page 27 for more information about configuring IPMP groups in a cluster.

10   **Set up name-to-address mappings for all logical hostname resources that are used by the zone.**

a.   **Add name-to-address mappings to the** /etc/inet/hosts **file on the zone.**

This information is not inherited from the global zone.

b.   **If you use a name server, add the name-to-address mappings.**

Next Steps   To install an application in a non-global zone, use the same procedure as for a stand-alone system. See your application's installation documentation for procedures to install the software in a non-global zone. Also see "Adding and Removing Packages and Patches on a Solaris System With Zones Installed (Task Map)" in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

To install and configure a data service in a non-global zone, see the Sun Cluster manual for the individual data service.

Configuring a Zone Cluster

This section provide procedures to configure a cluster of non-global zones.

Overview of the clzonecluster **Utility**

The clzonecluster utility creates, modifies, and removes a zone cluster. The clzonecluster utility actively manages a zone cluster. For example, the clzonecluster utility both boots and halts a zone cluster. Progress messages for the clzonecluster utility are output to the console, but are not saved in a log file.

The utility operates in the following levels of scope, similar to the zonecfg utility:

- The cluster scope affects the entire zone cluster.

- The node scope affects only the one zone-cluster node that is specified.

- The resource scope affects either a specific node or the entire zone cluster, depending on which scope you enter the resource scope from. Most resources can only be entered from the node scope. The scope is identified by the following prompts:

```
clzc:zoneclustername:resource>          cluster-wide setting
clzc:zoneclustername:node:resource>      node-specific setting
```

You can specify any Solaris zones resource parameter, as well as parameters that are specific to zone clusters, by using the clzonecluster utility. For information about parameters that you can set in a zone cluster, see the clzonecluster(1CL)man page. Additional information about 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 39.

- 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 and Property Types" in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*.

  - The name of each node in the global cluster on which to create a zone-cluster node.

  - The zone public hostname, or host alias, that you assign to each zone-cluster node.

  - The public-network IP address that each zone-cluster node uses.

  - The name of the public-network adapter that each zone-cluster node uses to connect to the public network.

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.

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

```
phys-schost-1# clzonecluster configure zoneclustername
clzc:zoneclustername> create

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

    Add the first node and specify node-specific settings
clzc:zoneclustername> add node
clzc:zoneclustername:node> set physical-host=baseclusternode1
clzc:zoneclustername:node> set hostname=hostname1
clzc:zoneclustername:node> add net
clzc:zoneclustername:node:net> set address=public_netaddr
clzc:zoneclustername:node:net> set physical=adapter
clzc:zoneclustername:node:net> end
clzc:zoneclustername:node> end

    Set the root password globally for all nodes in the zone cluster
clzc:zoneclustername> add sysid
clzc:zoneclustername:sysid> set root_password=encrypted_password
clzc:zoneclustername:sysid> end

    Save the configuration and exit the utility
clzc:zoneclustername> commit
clzc:zoneclustername> exit
```

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

```
phys-schost-1# clzonecluster configure zoneclustername
clzc:zoneclustername> add node
clzc:zoneclustername:node> set physical-host=baseclusternode2
clzc:zoneclustername:node> set hostname=hostname2
clzc:zoneclustername:node> add net
clzc:zoneclustername:node:net> set address=public_netaddr
clzc:zoneclustername:node:net> set physical=adapter
clzc:zoneclustername:node:net> end
clzc:zoneclustername:node> end
clzc:zoneclustername> exit
```

5 Verify the zone cluster configuration.

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

```
phys-schost-1# clzonecluster verify zoneclustername
phys-schost-1# clzonecluster status zoneclustername
=== Zone Clusters ===

--- Zone Cluster Status ---

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

6 Install and boot the cluster.

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

     Installation of the zone cluster might take several minutes
phys-schost-1# clzonecluster boot zoneclustername
Waiting for zone boot commands to complete on all the nodes of
the zone cluster "zoneclustername"...
```

Example 6–2 Configuration File to Create a Zone Cluster

The following example shows the contents of a command file that can be used with the clzonecluster utility to create a zone cluster. The file contains the series of clzonecluster commands that you would input manually.

In the following configuration, the zone cluster sczone is created on the global-cluster node phys-schost-1. The zone cluster uses /zones/sczone as the zone path. The first node of the zone cluster is assigned the hostname zc-host-1 and uses the public-network address

172.16.0.1 and the bge0 adapter. The second node of the zone cluster is created on the global-cluster node phys-schost-2. This second zone-cluster node is assigned the hostname zc-host-2 and uses the public-network address 172.16.0.2 and the bge1 adapter.

```
create
set zonepath=/zones/sczone
add node
set physical-host=phys-schost-1
set hostname=zc-host-1
add net
set address=172.16.0.1
set physical=bge0
end
end
add sysid
set root_password=encrypted_password
end
add node
set physical-host=phys-schost-2
set hostname=zc-host-2
add net
set address=172.16.0.2
set physical=bge1
end
end
commit
exit
```

Example 6–3   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 To add the use of a file system to the zone cluster, go to "Adding File Systems to a Zone Cluster" on page 218.

To add the use of global storage devices to the zone cluster, go to "Adding Storage Devices to a Zone Cluster" on page 222.

Adding File Systems to a Zone Cluster

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

After a file system is added to a zone cluster and brought online, the file system is visible on from within that zone cluster.

Note – You cannot use the clzonecluster command to add a local file system, which is mounted on a single global-cluster node, to a zone cluster. Instead, use the zonecfg command as you normally would in a stand-alone system. The local file system would not be under cluster control.

You cannot add a cluster file system to a zone cluster.

The following procedures are in this section:

- "How to Add a Highly Available Local File System to a Zone Cluster" on page 218
- "How to Add a ZFS Storage Pool to a Zone Cluster" on page 220
- "How to Add a QFS Shared File System to a Zone Cluster" on page 221

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

Perform this procedure to add a highly available 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 220.

---

1 On the global cluster, configure the highly available local file system that you want to use in the zone cluster.

See "Enabling Highly Available Local File Systems" in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

2 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.

3 Display the /etc/vfstab **entry for the file system that you want to mount on the zone cluster.**

```
phys-schost# vi /etc/vfstab
```

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

```
phys-schost# clzonecluster configure zoneclustername
clzc:zoneclustername> add fs
clzc:zoneclustername:fs> set dir=mountpoint
clzc:zoneclustername:fs> set special=disk-device-name
clzc:zoneclustername:fs> set raw=raw-disk-device-name
clzc:zoneclustername:fs> set type=FS-type
clzc:zoneclustername:fs> end
clzc:zoneclustername> exit
```

dir=*mountpoint*   Specifies the file system mount point

special=*disk-device-name*   Specifies the name of the disk device

raw=*raw-disk-device-name*   Specifies the name of the raw disk device

type=*FS-type*   Specifies the type of file system

5 Verify the addition of the file system.

```
phys-schost# clzonecluster show -v zoneclustername
```

Example 6–4 Adding a Highly Available Local File System to a Zone Cluster

This example adds the highly available local file system /global/oracle/d1 for use by the sczone zone cluster.

```
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/oracle/d1 ufs 5 no logging

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> end
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:                                []
...
```

▼ How to Add a ZFS Storage Pool to a Zone Cluster

Perform this procedure to add a ZFS storage pool for use by a zone cluster.

1 Configure the ZFS storage pool on the global cluster.

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

See *Solaris ZFS Administration Guide* for procedures to create a ZFS pool..

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

3 Add the pool to the zone-cluster configuration.

```
phys-schost# clzonecluster configure zoneclustername
clzc:zoneclustername> add dataset
clzc:zoneclustername:dataset> set name=ZFSpoolname
clzc:zoneclustername:dataset> end
clzc:zoneclustername> exit
```

4 Verify the addition of the file system.

```
phys-schost# clzonecluster show -v zoneclustername
```

Example 6–5 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> exit

phys-schost-1# clzonecluster show -v sczone
...
  Resource Name:                                    dataset
    name:                                           zpool1
...
```

▼ How to Add a QFS Shared File System to a Zone Cluster

Perform this procedure to add a Sun StorageTek QFS shared file system for use by a zone cluster.

Note – At this time, QFS shared file systems are only supported for use in clusters that are configured with Oracle Real Application Clusters (RAC). On clusters that are not configured with Oracle RAC, you can use a single-machine QFS file system that is configured as a highly available local file system.

---

1 On the global cluster, configure the QFS shared file system that you want to use in the zone cluster.

Follow procedures in "Tasks for Configuring the Sun StorEdge QFS Shared File System for Oracle Files" in *Sun Cluster Data Service for Oracle RAC Guide for Solaris OS*.

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

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

3 Display the /etc/vfstab **entry for the file system that you want to mount on the zone cluster.**

You will use information from the entry to specify the file system to the zone-cluster configuration.

```
phys-schost# vi /etc/vfstab
```

4 Add the file system to the zone cluster configuration.

```
phys-schost# clzonecluster configure zoneclustername
clzc:zoneclustername> add fs
clzc:zoneclustername:fs> set dir=mountpoint
clzc:zoneclustername:fs> set special=QFSfilesystemname
clzc:zoneclustername:fs> set type=samfs
clzc:zoneclustername:fs> end
clzc:zoneclustername> exit
```

5   Verify the addition of the file system.

```
phys-schost# clzonecluster show -v zoneclustername
```

Example 6–6   Adding a QFS Shared File System to a Zone Cluster

The following example shows the QFS shared file system Data-cz1 added to the zone cluster
sczone. From the global cluster, the mount point of the file system is
/zones/sczone/root/db_qfs/Data1, where /zones/sczone/root/ is the zone's root path.
From the zone-cluster node, the mount point of the file system is db_qfs/Data1.

```
phys-schost-1# vi /etc/vfstab
#device          device          mount   FS      fsck    mount   mount
#to mount        to fsck         point   type    pass    at boot options
#
Data-cz1         -               /zones/sczone/root/db_qfs/Data1 samfs - no shared,notrace

phys-schost-1# clzonecluster configure sczone
clzc:sczone> add fs
clzc:sczone:fs> set dir=/db_qfs/Data1
clzc:sczone:fs> set special=Data-cz1
clzc:sczone:fs> set type=samfs
clzc:sczone:fs> end
clzc:sczone> exit

phys-schost-1# clzonecluster show -v sczone
...
  Resource Name:                        fs
    dir:                                /db_qfs/Data1
    special:                            Data-cz1
    raw:
    type:                               samfs
    options:                            []
...
```

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.

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/setnumber
```

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 zoneclustername
clzc:zoneclustername> add device
clzc:zoneclustername:device> set match=/dev/md/diskset/*dsk/metadevice
clzc:zoneclustername:device> end
clzc:zoneclustername> add device
clzc:zoneclustername:device> set match=/dev/md/shared/setnumber/*dsk/metadevice
clzc:zoneclustername:device> end
clzc:zoneclustername:> 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 zoneclustername
```

Example 6–7   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:> 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/setnumber
```

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 zoneclustername
clzc:zoneclustername> add device
clzc:zoneclustername:device> set match=/dev/md/diskset/*dsk/*
clzc:zoneclustername:device> end
clzc:zoneclustername> add device
clzc:zoneclustername:device> set match=/dev/md/shared/setnumber/*dsk/*
clzc:zoneclustername:device> end
clzc:zoneclustername:> 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 zoneclustername
```

Example 6–8 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:> 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 zoneclustername
clzc:zoneclustername> add device
clzc:zoneclustername:device> set match=/dev/did/*dsk/dNs*
clzc:zoneclustername:device> end
clzc:zoneclustername:> 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 zoneclustername
```

Example 6–9 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:> exit

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

▼ How to Add a Raw-Disk Device to a Zone Cluster

● **Use the** zonecfg **command to export raw-disk devices (cNtXdYsZ) to a zone-cluster node, 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 "How to Import Raw and Block Devices by Using zonecfg" in *System Administration Guide: Solaris Containers-Resource Management and Solaris Zones*for more information about exporting raw-disk devices to a non-global zone.

Installing the Sun Cluster Module to Sun Management Center

This chapter provides guidelines and procedures to install the Sun Cluster module to the Sun Management Center graphical user interface (GUI).

SPARC: Installing the Sun Cluster Module for Sun Management Center

This section provides information and procedures to install software for the Sun Cluster module to Sun Management Center.

The Sun Cluster module for Sun Management Center enables you to use Sun Management Center to monitor the cluster. The following table lists the tasks to perform to install the Sun Cluster–module software for Sun Management Center. Complete the procedures in the order that is indicated.

TABLE 7–1   Task Map: Installing the Sun Cluster Module for Sun Management Center

Task	Instructions
Install Sun Management Center server, agent, and console packages.	Sun Management Center documentation "SPARC: Installation Requirements for Sun Cluster Monitoring" on page 228
Install Sun Cluster–module packages.	"SPARC: How to Install the Sun Cluster Module for Sun Management Center" on page 229
Start Sun Management Center server, console, and agent processes.	"SPARC: How to Start Sun Management Center" on page 230
Add each cluster node as a Sun Management Center agent host object.	"SPARC: How to Add a Global-Cluster Node as a Sun Management Center Agent Host Object" on page 231

TABLE 7–1   Task Map: Installing the Sun Cluster Module for Sun Management Center   *(Continued)*

Task	Instructions
Load the Sun Cluster module to begin to monitor the cluster.	"SPARC: How to Load the Sun Cluster Module" on page 232

SPARC: Installation Requirements for Sun Cluster Monitoring

The Sun Cluster module for Sun Management Center is used to monitor a Sun Cluster configuration. Perform the following tasks before you install the Sun Cluster module packages.

- **Space requirements** – Ensure that 25 Mbytes of space is available on each cluster node for Sun Cluster–module packages.

- **Sun Management Center installation** – Follow procedures in your Sun Management Center installation documentation to install Sun Management Center software.

  The following are additional requirements for a Sun Cluster configuration:

  - Install the Sun Management Center agent package on each cluster node.

  - When you install Sun Management Center on an agent machine (cluster node), choose whether to use the default of 161 for the agent (SNMP) communication port or another number. This port number enables the server to communicate with this agent. Record the port number that you choose for reference later when you configure the cluster nodes for monitoring.

    See your Sun Management Center installation documentation for information about choosing an SNMP port number.

  - If you have an administrative console or other dedicated machine, you can run the console process on the administrative console and the server process on a separate machine. This installation approach improves Sun Management Center performance.

  - For best results, install the Sun Management Center server and console packages on noncluster machines.

  - If you choose to install the server or console packages on a cluster node, you might experience the following adverse performance impacts:

    - Cluster performance might be noticeably slowed due to the increased load from Sun Management Center processes. This is especially likely when the Sun Management Center server runs on a cluster node.

    - Sun Management Center is not made highly available when you install the server on a cluster node. Sun Management Center service will fail if the node goes down, such as during a failover to another node.

■ **Web browser** – Ensure that the web browser that you use to connect to Sun Management Center is supported by Sun Management Center. Certain features might not be available on unsupported web browsers. See your Sun Management Center documentation for information about supported web browsers and any configuration requirements.

▼ SPARC: How to Install the Sun Cluster Module for Sun Management Center

Perform this procedure to install the Sun Cluster–module server package.

Note – The Sun Cluster-module agent packages, SUNWscsal and SUNWscsam, are already added to global-cluster nodes during Sun Cluster software installation.

Before You Begin Ensure that all Sun Management Center core packages are installed on the appropriate machines. This task includes installing Sun Management Center agent packages on each cluster node. See your Sun Management Center documentation for installation instructions.

1 **On the Sun Management Center server machine, install the Sun Cluster–module server package** SUNWscssv**.**

   a. **Become superuser.**

   b. **Load the Sun Java Availability Suite DVD-ROM into the DVD-ROM drive.**

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

   c. **Change to the** Solaris_sparc/Product/sun_cluster/Solaris_*ver*/Packages/ **directory, where** *ver* **is** 9 **for Solaris 9 or** 10 **for Solaris 10.**

   phys-schost# **cd /cdrom/cdrom0/Solaris_sparc/Product/sun_cluster/Solaris_*ver*/Packages/**

   d. **Install the Sun Cluster–module server package.**

   ```
   phys-schost# pkgadd -d . SUNWscssv
   ```

   e. **Unload the Sun Java Availability Suite DVD-ROM from the DVD-ROM drive.**

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

      ii. **Eject the DVD-ROM.**

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

2   Install any Sun Cluster–module patches.

See "Patches and Required Firmware Levels" in *Sun Cluster Release Notes* for the location of patches and installation instructions.

Next Steps   Start Sun Management Center. Go to "SPARC: How to Start Sun Management Center" on page 230.

▼ SPARC: How to Start Sun Management Center

Perform this procedure to start the Sun Management Center server, agent, and console processes.

1   As superuser, on the Sun Management Center server machine, start the Sun Management Center server process.

The *install-dir* is the directory on which you installed the Sun Management Center software. The default directory is /opt.

```
server# /install-dir/SUNWsymon/sbin/es-start -S
```

2   As superuser, on each Sun Management Center agent machine (cluster node), start the Sun Management Center agent process.

```
phys-schost# /install-dir/SUNWsymon/sbin/es-start -a
```

3   On each Sun Management Center agent machine (cluster node), ensure that the scsymon_srv **daemon is running.**

```
phys-schost# ps -ef | grep scsymon_srv
```

If any cluster node is not already running the scsymon_srv daemon, start the daemon on that node.

```
phys-schost# /usr/cluster/lib/scsymon/scsymon_srv
```

4   On the Sun Management Center console machine (administrative console), start the Sun Management Center console.

You do not need to be superuser to start the console process.

```
adminconsole% /install-dir/SUNWsymon/sbin/es-start -c
```

Next Steps   Add a cluster node as a monitored host object. Go to "SPARC: How to Add a Global-Cluster Node as a Sun Management Center Agent Host Object" on page 231.

▼ SPARC: How to Add a Global-Cluster Node as a Sun Management Center Agent Host Object

Perform this procedure to create a Sun Management Center agent host object for a global-cluster node.

1 Log in to Sun Management Center.

See your Sun Management Center documentation.

2 From the Sun Management Center main window, select a domain from the Sun Management Center Administrative Domains pull-down list.

This domain contains the Sun Management Center agent host object that you create. During Sun Management Center software installation, a Default Domain was automatically created for you. You can use this domain, select another existing domain, or create a new domain.

See your Sun Management Center documentation for information about how to create Sun Management Center domains.

3 Choose Edit⇒Create an Object from the pull-down menu.

4 Click the Node tab.

5 From the Monitor Via pull-down list, select Sun Management Center Agent - Host.

6 Fill in the name of the cluster node, for example, `phys-schost-1`**, in the Node Label and Hostname text fields.**

Leave the IP text field blank. The Description text field is optional.

7 In the Port text field, type the port number that you chose when you installed the Sun Management Center agent machine.

8 Click OK.

A Sun Management Center agent host object is created in the domain.

Next Steps Load the Sun Cluster module. Go to "SPARC: How to Load the Sun Cluster Module" on page 232.

Troubleshooting You need only one cluster node host object to use Sun Cluster-module monitoring and configuration functions for the entire cluster. However, if that cluster node becomes unavailable, connection to the cluster through that host object also becomes unavailable. Then you need another cluster-node host object to reconnect to the cluster.

▼ SPARC: How to Load the Sun Cluster Module

Perform this procedure to start cluster monitoring.

1   In the Sun Management Center main window, right click the icon of a cluster node.
The pull-down menu is displayed.

2   Choose Load Module.
The Load Module window lists each available Sun Management Center module and whether the module is currently loaded.

3   Choose Sun Cluster: Not Loaded and click OK.
The Module Loader window shows the current parameter information for the selected module.

4   Click OK.
After a few moments, the module is loaded. A Sun Cluster icon is then displayed in the Details window.

5   Verify that the Sun Cluster module is loaded.
Under the Operating System category, expand the Sun Cluster subtree in either of the following ways:

   ■   **In the tree hierarchy on the left side of the window, place the cursor over the Sun Cluster module icon and single-click the left mouse button.**

   ■   **In the topology view on the right side of the window, place the cursor over the Sun Cluster module icon and double-click the left mouse button.**

See Also   See your Sun Management Center documentation for information about how to use Sun Management Center.

Next Steps   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 *Sun Cluster Data Services Planning and Administration Guide for Solaris OS*.

8

Uninstalling Software From the Cluster

This chapter provides procedures for uninstalling or removing certain software from a Sun Cluster configuration. The following procedures are in this chapter:

Uninstalling the Software

This section provides procedures to uninstall or remove certain software products from a global cluster.

▼ How to Unconfigure Sun 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 147, do not perform this procedure. Instead, go to "How to Uninstall Sun Cluster Software From a Cluster Node" in *Sun Cluster System Administration Guide for Solaris OS*.

Before You Begin   Attempt to rerun cluster configuration of the node. You can correct certain cluster node configuration failures by repeating Sun 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 the Authorized Node List" in *Sun Cluster System Administration Guide for Solaris OS* for procedures.

2   On a node that you intend to unconfigure, become superuser.

3   Shut down the node.

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

4   Reboot the node into noncluster mode.

- **On SPARC based systems, perform the following command:**

```
ok boot -x
```

- **On x86 based systems, perform the following commands:**

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

The GRUB menu appears similar to the following:

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

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

For more information about GRUB based booting, see Chapter 11, "GRUB Based Booting (Tasks)," in *System Administration Guide: Basic Administration*.

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

The GRUB boot parameters screen appears similar to the following:

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

c. Add - x to the command to specify that the system boot into noncluster mode.

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

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

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

The screen displays the edited command.

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

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

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

5 **Change to a directory, such as the root (/) directory, that does not contain any files that are delivered by the Sun Cluster packages.**

```
phys-schost# cd /
```

6 **Remove the node from the cluster configuration.**

```
phys-schost# /usr/cluster/bin/clnode remove
```

The node is removed from the cluster configuration but Sun Cluster software is not removed from the node.

See the clnode(1CL) man page for more information.

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

8 **(Optional) Uninstall Sun Cluster framework and data service software packages.**

> **Note –** If you do not need to remove or reinstall Sun Cluster software, you can skip this step.

This step also removes the Sun Cluster entry from the Sun Java Enterprise System (Java ES) product registry. If the Java ES product registry contains a record that Sun Cluster software is installed, the Java ES installer will show the Sun Cluster component grayed out and will not permit reinstallation.

     a. Start the uninstall program.

     Run the following command, where *ver* is the version of the Java ES distribution from which you installed Sun Cluster software.

```
phys-schost# /var/sadm/prod/SUNWentsysver/uninstall
```

     b. Follow the onscreen instructions to select the Sun Cluster components to uninstall.

> **Note –** If Sun Cluster Geographic Edition software is installed, you must uninstall it as well.

     For more information about using the uninstall program, see Chapter 8, "Uninstalling," in *Sun Java Enterprise System 5 Installation Guide for UNIX*. *Sun Java Enterprise System 5 Installation Guide for UNIX*

Next Steps To reinstall or reconfigure Sun 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 *Sun Cluster 3.1 - 3.2 Hardware Administration Manual for Solaris OS* and the removal procedure in the Sun Cluster Hardware Administration Collection manual for your storage array.

▼ How to Remove Sun Cluster Information From a JumpStart Install Server

Perform this procedure to remove Sun Cluster information from a JumpStart install server that was used to install and configure a cluster. You can remove information for one or more individual nodes or for one or more entire clusters. For more information about the JumpStart feature, see *Solaris 10 5/08 Installation Guide: Custom JumpStart and Advanced Installations*.

1 Become superuser on the JumpStart install server.

2 Change to the JumpStart directory that you used to install Sun Cluster software.

```
installserver# cd jumpstart-dir
```

3 Remove from the rules **file any entries that were created by the** scinstall **command that contain the name of a node that you want to remove.**

The Sun Cluster entries reference autostinstall.class or autoscinstall.finish or both. Entries appear similar to the following, where *release* is the version of Sun Cluster software:

```
hostname phys-schost-1 - autoscinstall.d/release/autoscinstall.class \
autoscinstall.d/release/autoscinstall.finish
```

4 Regenerate the rules.ok **file.**

Run the check command that is located in the *jumpstart-dir/* directory to regenerate the rules.ok file.

```
installserver# ./check
```

5 Remove from the appropriate clusters/*clustername/* **directory the symbolic link for each node that you want to remove.**

- To remove the symbolic link for one or more nodes in a cluster, remove the link that is named for each node to remove.

  ```
  installserver# rm -f autoscinstall.d/clusters/clustername/nodename
  ```

- To remove the symbolic links for an entire cluster, recursively remove the directory that is named for the cluster to remove.

  installserver# **rm -rf autoscinstall.d/clusters/***clustername*

- To remove the symbolic links for all clusters, recursively remove the clusters/ directory.

  installserver# **rm -rf autoscinstall.d/clusters**

6 **Remove from the** autoscinstall.d/ **directory the node configuration directory that is named for each node that you want to remove.**

If you are removing information for an entire cluster, remove the directory for each node in the cluster.

- To remove information for one or more nodes in a cluster, recursively remove the directory for each node.

  installserver# **rm -rf autoscinstall.d/nodes/***nodename*

- To remove all entries for all clusters, recursively remove the autoscinstall.d directory.

  installserver# **rm -rf autoscinstall.d**

7 **Remove the** .autoscinstall.log.3 **file.**

installserver# **rm .autoscinstall.log.3**

8 **(Optional) If you used a flash archive to JumpStart install the cluster, remove the flash archive if you no longer need the file.**

installserver# **rm** *filename.***flar**

Next Steps   If you intend to use custom JumpStart to reinstall a cluster from which you removed information for one or more nodes that were removed from that cluster, you must rerun interactive scinstall to update the cluster node list. See "How to Install Solaris and Sun Cluster Software (JumpStart)" on page 100.

▼ How to Uninstall SunPlex Manager Software

Use this procedure to uninstall SunPlex Manager software that was installed by the Java ES installer utility up to and including the Sun Java Enterprise System 2005Q4 distribution or by any other installation method.

To remove Sun Cluster Manager software that was installed with the Sun Java Enterprise System 5 or compatible distribution of the Java ES installer utility, instead use the Java ES

uninstall utility to remove these packages. For more information, see Chapter 8, "Uninstalling," in *Sun Java Enterprise System 5 Installation Guide for UNIX*.

Note – Do not uninstall SunPlex Manager or Sun Cluster Manager or its shared components if you intend to use the graphical user interface (GUI) to Sun Cluster Geographic Edition software or to Sun Cluster features that are available through the GUI. These features include the data-service configuration wizards or system resource monitoring.

However, if you use the command-line interface to administer these features, it is safe to uninstall SunPlex Manager or Sun Cluster Manager software.

Perform this procedure on each node in the cluster to uninstall SunPlex Manager software and related Sun Java Enterprise System (Java ES) shared components.

Note – SunPlex Manager software must be installed on all cluster nodes or on none.

---

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

2 **Remove the SunPlex Manager software packages.**

    phys-schost# **pkgrm SUNWscspm SUNWscspmu SUNWscspmr**

3 **(Optional) Remove Sun Java Web Console software packages, if you have no other need for them.**

    phys-schost# **pkgrm SUNWmctag SUNWmconr SUNWmcon SUNWmcos SUNWmcosx**

4 **(Optional) If you removed Sun Java Web Console packages, remove Apache Tomcat and Java Studio Enterprise Web Application Framework (Java ATO) software packages, if you have no other need for them.**

Remove the packages listed below for each additional product that you want to uninstall, in the order that the set of packages is listed.

    phys-schost# **pkgrm** *packages*

Product	Package Names
Apache Tomcat	SUNWtcatu
Java ATO	SUNWjato SUNWjatodmo SUNWjatodoc

▼ How to Uninstall the SUNWscrdt **Package**

Perform this procedure on each node in the cluster.

Before You Begin   Verify that no applications are using the RSMRDT driver before performing this procedure.

1   **Become superuser on the node to which you want to uninstall the** SUNWscrdt **package.**

2   **Uninstall the** SUNWscrdt **package.**

```
phys-schost# pkgrm SUNWscrdt
```

▼ How to Unload the RSMRDT Driver Manually

If the driver remains loaded in memory after completing "How to Uninstall the SUNWscrdt Package" on page 240, perform this procedure to unload the driver manually.

1   **Start the** adb **utility.**

```
phys-schost# adb -kw
```

2   **Set the kernel variable** clifrsmrdt_modunload_ok **to** 1**.**

```
physmem NNNN
clifrsmrdt_modunload_ok/W 1
```

3   **Exit the** adb **utility by pressing Control-D.**

4   **Find the** clif_rsmrdt **and** rsmrdt **module IDs.**

```
phys-schost# modinfo | grep rdt
```

5   **Unload the** clif_rsmrdt **module.**

You must unload the clif_rsmrdt module before you unload the rsmrdt module.

```
phys-schost# modunload -i clif_rsmrdt_id
```

clif_rsmrdt_id   Specifies the numeric ID for the module that is being unloaded.

6   **Unload the** rsmrdt **module.**

```
phys-schost# modunload -i rsmrdt_id
```

rsmrdt_id   Specifies the numeric ID for the module that is being unloaded.

7   **Verify that the module was successfully unloaded.**

```
phys-schost# modinfo | grep rdt
```

Example 8–1   Unloading the RSMRDT Driver

The following example shows the console output after the RSMRDT driver is manually unloaded.

```
phys-schost# adb -kw
physmem fc54
clifrsmrdt_modunload_ok/W 1
clifrsmrdt_modunload_ok: 0x0 = 0x1
^D
phys-schost# modinfo | grep rsm
 88 f064a5cb 974 - 1 rsmops (RSMOPS module 1.1)
 93 f08e07d4 b95 - 1 clif_rsmrdt (CLUSTER-RSMRDT Interface module)
 94 f0d3d000 13db0 194 1 rsmrdt (Reliable Datagram Transport dri)
phys-schost# modunload -i 93
phys-schost# modunload -i 94
phys-schost# modinfo | grep rsm
 88 f064a5cb 974 - 1 rsmops (RSMOPS module 1.1)
```

Troubleshooting   If the modunload command fails, applications are probably still using the driver. Terminate the applications before you run modunload again.

▼ How to Remove Quorum Server Software

Before You Begin   Before you uninstall Quorum Server software, ensure that you have completed the following tasks:

- On each cluster that uses the quorum server, remove the quorum server as a quorum device. Perform the steps in "How to Remove a Quorum Device" in *Sun Cluster System Administration Guide for Solaris OS*.

  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 *Sun Cluster System Administration Guide for Solaris OS*.

- On each quorum-server host computer, stop the quorum server by following the steps in "How to Stop a Quorum Server" in *Sun Cluster System Administration Guide for Solaris OS*.

1. **(Optional) To use the interactive graphical interface to uninstall the Java Enterprise System components, ensure that the display environment of the host server to uninstall is set to display the GUI.**

   ```
   % xhost +
   % setenv DISPLAY nodename:0.0
   ```

2. **Become superuser on the quorum server host computer to uninstall.**

3. **Navigate to the directory where the uninstaller is located:**

   ```
   phys-schost# cd /var/sadm/prod/SUNWentsysver
   ```

   ver   The version of Java Enterprise System that is installed on your system.

4. **Start the installation wizard.**

   ```
   phys-schost# ./uninstall
   ```

5. **Follow instructions on the screen to uninstall the Quorum Server software from the quorum-server host computer.**

   After removal is finished, you can view any available log. See the *Sun Java Enterprise System 2006Q4 Installation Guide for UNIX* for additional information about using the Java Enterprise System `installer` program.

6. **(Optional) On each quorum-server host computer, clean up or remove the quorum server directories.**

   By default, this directory is /var/scqsd.

▼ How to Unconfigure a Zone Cluster

Perform this procedure to remove a zone cluster.

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

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

2. **Take offline each resource group in the zone cluster and disable its resources.**

   Note – The following steps are performed from a global-cluster node. To instead perform these steps from a node of the zone cluster, log in to the zone-cluster node and omit "-Z *zonecluster*" from each command.

   ---

   a. **Take each resource offline.**

      ```
      phys-schost# clresource offline -Z zonecluster resource-group
      ```

b. List all enabled resources in the zone cluster.

```
phys-schost# clresource show -Z zonecluster -p Enabled
=== Resources ===

Resource:                                    resource
  Enabled{nodename1}:                        True
  Enabled{nodename2}:                        True
...
```

c. Identify those resources that depend on other resources.

```
phys-schost# clresource show -Z zonecluster -p resource_dependencies
=== Resources ===

Resource:                                    node
  Resource_dependencies:                     node
...
```

You must disable dependent resources first before you disable the resources that they depend on.

d. Disable each enabled resource in the cluster.

```
phys-schost# clresource disable -Z zonecluster resource
```

See the clresource(1CL) man page for more information.

e. Verify that all resources are disabled.

```
phys-schost# clresource show -Z zonecluster -p Enabled
=== Resources ===

Resource:                                    resource
  Enabled{nodename1}:                        False
  Enabled{nodename2}:                        False
...
```

f. Move each resource group to the unmanaged state.

```
phys-schost# clresourcegroup unmanage -Z zonecluster resource-group
```

g. Verify that all resources on all nodes are `Offline` **and that all resource groups are in the** `Unmanaged` **state.**

```
phys-schost# cluster status -Z zonecluster -t resource,resourcegroup
```

h. Delete all resource groups and their resources from the zone cluster.

```
phys-schost# clresourcegroup delete -F -Z zonecluster +
```

3 Halt the zone cluster.

phys-schost# **clzonecluster halt** *zoneclustername*

4 Uninstall the zone cluster.

phys-schost# **clzonecluster uninstall** *zoneclustername*

5 Unconfigure the zone cluster.

phys-schost# **clzonecluster delete** *zoneclustername*

A

Sun Cluster Installation and Configuration Worksheets

This appendix provides worksheets to plan various components of your cluster configuration and examples of completed worksheets for your reference. See "Installation and Configuration Worksheets" in *Sun Cluster Data Services Planning and Administration Guide for Solaris OS* for configuration worksheets for resources, resource types, and resource groups.

Installation and Configuration Worksheets

If necessary, make additional copies of a worksheet to accommodate all the components in your cluster configuration. Follow planning guidelines in Chapter 1, "Planning the Sun Cluster Configuration," to complete these worksheets. Then refer to your completed worksheets during cluster installation and configuration.

Note – The data used in the worksheet examples is intended as a guide only. The examples do not represent a complete configuration of a functional cluster.

The following table lists the planning worksheets and examples provided in this appendix, as well as the titles of sections in Chapter 1, "Planning the Sun Cluster Configuration," that contain related planning guidelines.

TABLE A–1   Cluster Installation Worksheets and Related Planning Guidelines

Worksheet	Example	Section Titles of Related Planning Guidelines
"Local File System Layout Worksheet" on page 247	"Example: Local File System Layout Worksheets, With and Without Mirrored Root" on page 248	"System Disk Partitions" on page 19 "Guidelines for Mirroring the Root Disk" on page 52
"Local Devices Worksheets" on page 249	"Example: Local Devices Worksheets" on page 250	- - -
"Device Group Configurations Worksheet" on page 251	"Example: Device Group Configurations Worksheet" on page 252	"Device Groups" on page 41 "Planning Volume Management" on page 46
"Volume-Manager Configurations Worksheet" on page 253	"Example: Volume-Manager Configurations Worksheet" on page 254	"Planning Volume Management" on page 46 Your volume manager documentation
"Volumes Worksheet (Solaris Volume Manager)" on page 255	"Example: Volumes Worksheet (Solaris Volume Manager)" on page 256	"Planning Volume Management" on page 46 *Solaris Volume Manager Administration Guide* (Solaris 9 or Solaris 10)

Local File System Layout Worksheet

Node name: _____

TABLE A–2 Local File Systems With Mirrored Root Worksheet

Volume Name	Component	Component	File System	Size
			/	
			swap	
			/globaldevices	

TABLE A–3 Local File Systems With Nonmirrored Root Worksheet

Device Name	File System	Size
	/	
	swap	
	/globaldevices	

Example: Local File System Layout Worksheets, With and Without Mirrored Root

Node name: `phys-schost-1`

TABLE A–4   Example: Local File Systems With Mirrored Root Worksheet

Volume Name	Component	Component	File System	Size
`d1`	`c0t0d0s0`	`c1t0d0s0`	/	`6.75 GB`
`d2`	`c0t0d0s1`	`c1t0d0s1`	swap	`750 MB`
`d3`	`c0t0d0s3`	`c1t0d0s3`	/globaldevices	`512 MB`
`d7`	`c0t0d0s7`	`c1t0d0s7`	`SVM replica`	`20 MB`

TABLE A–5   Example: Local File Systems With Nonmirrored Root Worksheet

Device Name	File System	Size
`c0t0d0s0`	/	`6.75 GB`
`c0t0d0s1`	swap	`750 MB`
`c0t0d0s3`	/globaldevices	`512 MB`
`c0t0d0s7`	`SVM replica`	`20 MB`

Local Devices Worksheets

Node name:_____

TABLE A–6   Local Disks Worksheet

Local Disk Name	Size

TABLE A–7   Other Local Devices Worksheet

Device Type	Name

Example: Local Devices Worksheets

Node name: `phys-schost-1`

TABLE A–8   Example: Local Disks Worksheet

Local Disk Name	Size
`c0t0d0`	`2G`
`c0t1d0`	`2G`
`c1t0d0`	`2G`
`c1t1d0`	`2G`

TABLE A–9   Example: Other Local Devices Worksheet

Device Type	Name
tape	`/dev/rmt/0`

Device Group Configurations Worksheet

Volume manager (circle one):

Solaris Volume Manager | VxVM

TABLE A–10   Device Groups Worksheet

Disk Group/ Disk Set Name	Node Names (indicate priority if ordered list)	Ordered priority? (circle one)	Failback? (circle one)
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No

Example: Device Group Configurations Worksheet

Volume manager (circle one):

Solaris Volume Manager

TABLE A–11   Example: Device Groups Configurations Worksheet

Disk Group/ Disk Set Name	Node Names (indicate priority if ordered list)	Ordered priority? (circle one)	Failback? (circle one)
dg-schost-1	1) phys-schost-1, 2) phys-schost-2	Yes	Yes
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No
		Yes | No	Yes | No

Volume-Manager Configurations Worksheet

Volume manager (circle one):

Solaris Volume Manager | VxVM

TABLE A–12   Volume-Manager Configurations Worksheet

Name	Type	Component	Component

Example: Volume-Manager Configurations Worksheet

Volume manager (circle one):

Solaris Volume Manager

TABLE A–13 Example: Volume-Manager Configurations Worksheet

Name	Type	Component	Component
dg-schost-1/d1	mirror	c0t0d0s4	c4t4d0s4
dg-schost-1/d4	mirror	c0t0d2s5	d4t4d2s5

Volumes Worksheet (Solaris Volume Manager)

TABLE A–14 Volumes Worksheet (Solaris Volume Manager)

File System	Volume	Mirror	Submirrors	Hot-Spare Pool	Physical Devices

Example: Volumes Worksheet (Solaris Volume Manager)

TABLE A–15 Example: Volumes Worksheet (Solaris Volume Manager)

File System	Volume	Mirror	Submirrors	Hot-Spare Pool	Physical Devices
/A	d10	d11	d12, d13	hsp000	c1t0d0s0, c2t0d1s0

Index

W

whole-root zones, requirement for exclusive-IP
   zones, 211

X

`xntpd.cluster` command, starting NTP, 152
`xntpd` command
   starting NTP, 152
   stopping NTP, 151

Z

ZFS storage pools
   adding to a zone cluster, 220-221
   restrictions
      adding quorum disks, 38
      global-devices file systems, 21
      root file systems, 20
zone clusters
   adding devices
      DID devices, 225-226
      disk sets, 224-225
      metadevices, 223-224
      raw-disk devices, 226
   adding file systems
      highly available local file systems, 218-220
      QFS, 221-222
      ZFS storage pools, 220-221
   adding to the private IP-address range, 33
   configuration file, 216
   configuring, 213-226
   creating, 214-218
   creating from a configuration file, 217
   halting, 244
   IP addresses, 26
   naming, 40
   overview, 213-214
   planning, 39-40
   unconfiguring, 242-244
zones, *See* Solaris zones