



Sun Cluster 3.0 Installation Guide

Sun Microsystems, Inc.
901 San Antonio Road
Palo Alto, CA 94303-4900
U.S.A. 650-960-1300

Part Number 806-1419
November 2000, Revision A

Copyright Copyright 2000 Sun Microsystems, Inc. 901 San Antonio Road, Palo Alto, California 94303-4900 U.S.A. All rights reserved.

This product or document is protected by copyright and distributed under licenses restricting its use, copying, distribution, and decompilation. No part of this product or document may be reproduced in any form by any means without prior written authorization of Sun and its licensors, if any. Third-party software, including font technology, is copyrighted and licensed from Sun suppliers.

Parts of the product may be derived from Berkeley BSD systems, licensed from the University of California. UNIX is a registered trademark in the U.S. and other countries, exclusively licensed through X/Open Company, Ltd.

Sun, Sun Microsystems, the Sun logo, AnswerBook2, docs.sun.com, JumpStart, Solstice DiskSuite, Sun Enterprise SyMON, Sun Enterprise, Sun Management Center, OpenBoot, and Solaris are trademarks, registered trademarks, or service marks of Sun Microsystems, Inc. in the U.S. and other countries. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc.

The OPEN LOOK and Sun™ Graphical User Interface was developed by Sun Microsystems, Inc. for its users and licensees. Sun acknowledges the pioneering efforts of Xerox in researching and developing the concept of visual or graphical user interfaces for the computer industry. Sun holds a non-exclusive license from Xerox to the Xerox Graphical User Interface, which license also covers Sun's licensees who implement OPEN LOOK GUIs and otherwise comply with Sun's written license agreements.

RESTRICTED RIGHTS: Use, duplication, or disclosure by the U.S. Government is subject to restrictions of FAR 52.227-14(g)(2)(6/87) and FAR 52.227-19(6/87), or DFAR 252.227-7015(b)(6/95) and DFAR 227.7202-3(a).

DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID.

Copyright 2000 Sun Microsystems, Inc., 901 San Antonio Road, Palo Alto, Californie 94303 Etats-Unis. Tous droits réservés.

Ce produit ou document est protégé par un copyright et distribué avec des licences qui en restreignent l'utilisation, la copie, la distribution, et la décompilation. Aucune partie de ce produit ou document ne peut être reproduite sous aucune forme, par quelque moyen que ce soit, sans l'autorisation préalable et écrite de Sun et de ses bailleurs de licence, s'il y en a. Le logiciel détenu par des tiers, et qui comprend la technologie relative aux polices de caractères, est protégé par un copyright et licencié par des fournisseurs de Sun.

Des parties de ce produit pourront être dérivées des systèmes Berkeley BSD licenciés par l'Université de Californie. UNIX est une marque déposée aux Etats-Unis et dans d'autres pays et licenciée exclusivement par X/Open Company, Ltd.

Sun, Sun Microsystems, le logo Sun, AnswerBook2, docs.sun.com, JumpStart, Solstice DiskSuite, Sun Enterprise SyMON, Sun Enterprise, Sun Management Center, OpenBoot, et Solaris sont des marques de fabrique ou des marques déposées, ou marques de service, de Sun Microsystems, Inc. aux Etats-Unis et dans d'autres pays. Toutes les marques SPARC sont utilisées sous licence et sont des marques de fabrique ou des marques déposées de SPARC International, Inc. aux Etats-Unis et dans d'autres pays. Les produits portant les marques SPARC sont basés sur une architecture développée par Sun Microsystems, Inc.

L'interface d'utilisation graphique OPEN LOOK et Sun™ a été développée par Sun Microsystems, Inc. pour ses utilisateurs et licenciés. Sun reconnaît les efforts de pionniers de Xerox pour la recherche et le développement du concept des interfaces d'utilisation visuelle ou graphique pour l'industrie de l'informatique. Sun détient une licence non exclusive de Xerox sur l'interface d'utilisation graphique Xerox, cette licence couvrant également les licenciés de Sun qui mettent en place l'interface d'utilisation graphique OPEN LOOK et qui en outre se conforment aux licences écrites de Sun.

CETTE PUBLICATION EST FOURNIE "EN L'ETAT" ET AUCUNE GARANTIE, EXPRESSE OU IMPLICITE, N'EST ACCORDEE, Y COMPRIS DES GARANTIES CONCERNANT LA VALEUR MARCHANDE, L'APTITUDE DE LA PUBLICATION A REpondre A UNE UTILISATION PARTICULIERE, OU LE FAIT QU'ELLE NE SOIT PAS CONTREFAISANTE DE PRODUIT DE TIERS. CE DENI DE GARANTIE NE S'APPLIQUERAIT PAS, DANS LA MESURE OU IL SERAIT TENU JURIDIQUEMENT NUL ET NON AVENU.



Contents

Preface	7
1. Planning the Sun Cluster Configuration	11
Where to Find Sun Cluster Installation Tasks	11
Planning the Solaris Operating Environment	13
Guidelines for Selecting Your Solaris Installation Method	13
System Disk Partitions	13
Planning the Sun Cluster Environment	17
Licensing	17
Software Patches	17
IP Addresses	17
Sun Cluster Configurable Components	18
Planning the Global Devices and Cluster File Systems	22
Guidelines for Highly Available Global Devices and Cluster File Systems	23
Mount Information for Cluster File Systems	23
Planning Volume Management	23
Guidelines for Volume Manager software	24
Guidelines for Solstice DiskSuite	25
Guidelines for VERITAS Volume Manager	25

File-System Logging 26

Mirroring Guidelines 27

2. **Installing and Configuring Sun Cluster Software 31**

Installing the Software 32

▼ How to Install Cluster Control Panel Software on the Administrative Console 33

▼ How to Install the Solaris Operating Environment 36

▼ How to Install Sun Cluster Software and Establish New Cluster Nodes 38

▼ How to Use JumpStart to Install the Solaris Operating Environment and Establish New Cluster Nodes 42

▼ How to Configure the Name Service Switch 49

▼ How to Install Solstice DiskSuite Software 50

▼ How to Install VERITAS Volume Manager Software 52

▼ How to Set Up the Root User's Environment 53

▼ How to Install Data Service Software Packages 54

Configuring the Cluster 55

▼ How to Perform Post-Installation Setup 56

▼ How to Configure Volume Manager Software 58

▼ How to Add Cluster File Systems 59

▼ How to Configure Additional Public Network Adapters 63

▼ How to Configure Public Network Management (PNM) 64

▼ How to Change Private Hostnames 65

▼ How to Update Network Time Protocol (NTP) 66

Installing the Sun Cluster Module for Sun Management Center 67

Installation Requirements for Sun Management Center Software for Sun Cluster Monitoring 68

▼ How to Install the Sun Cluster Module for Sun Management Center 69

▼ How to Start Sun Management Center Software 70

▼ How to Add a Cluster Node as a Sun Management Center Agent Host Object 71

▼ How to Load the Sun Cluster Module 72

3.	Upgrading Sun Cluster Software	75
	Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Software	75
	Overview of Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Software	77
	▼ How to Shut Down the Cluster	78
	▼ How to Uninstall VERITAS Volume Manager Software	80
	▼ How to Upgrade the Solaris Operating Environment	81
	▼ How to Upgrade Cluster Software Packages	84
	▼ How to Update the Root User's Environment	87
	▼ How to Upgrade Data Service Software Packages	88
	▼ How to Finish Upgrading Cluster Software	90
	▼ How to Verify Cluster Membership	93
A.	Configuring Solstice DiskSuite Software	95
	Configuring Solstice DiskSuite for Sun Cluster Configurations	96
	▼ How to Set the Number of Metadevice Names and Disksets	97
	▼ How to Create Metadevice State Database Replicas	98
	Mirroring the Root Disk	99
	▼ How to Mirror the Root (/) File System	99
	▼ How to Mirror the Global Namespace	102
	▼ How to Mirror File Systems That Cannot Be Unmounted	105
	▼ How to Mirror User-Defined File Systems	109
	▼ How to Create a Diskset	111
	Adding Drives to a Diskset	113
	▼ How to Add Drives to a Diskset	113
	▼ How to Repartition Drives in a Diskset	115
	▼ How to Create an md.tab File	116
	▼ How to Activate Metadevices	118
	Mediators Overview	119

- ▼ How to Add Mediator Hosts 120
- ▼ How to Check the Status of Mediator Data 121
- ▼ How to Fix Bad Mediator Data 121
- Solstice DiskSuite Configuration Example 122
- B. Configuring VERITAS Volume Manager 125**
 - Configuring VxVM for Sun Cluster Configurations 125
 - ▼ How to Verify the Pseudo-Device Major Number 126
 - ▼ How to Change the Pseudo-Device Major Number 127
 - Setting Up a `rootdg` Disk Group Overview 128
 - ▼ How to Encapsulate the Root Disk 129
 - ▼ How to Create a Non-Root `rootdg` Disk Group 135
 - ▼ How to Create and Register a Shared Disk Group 136
 - ▼ How to Assign a New Minor Number to a Disk Device Group 137
 - ▼ How to Verify the Disk Groups 138

Preface

The *Sun Cluster 3.0 Installation Guide* contains guidelines for planning a Sun™ Cluster 3.0 configuration, and provides procedures for installing, upgrading, and configuring the Sun Cluster software.

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 environment and expertise with the volume manager software used with Sun Cluster.

Using UNIX Commands

This document contains information on commands specific to installing or upgrading a Sun Cluster configuration. This document might not contain complete information on basic UNIX® commands and procedures such as shutting down the system, booting the system, and configuring devices.

See one or more of the following for this information:

- AnswerBook2™ online documentation for the Solaris software environment
- Other software documentation that you received with your system
- Solaris operating environment man pages

Typographic Conventions

Typeface or Symbol	Meaning	Examples
AaBbCc123	The names of commands, files, and directories; on-screen computer output	Edit your <code>.login</code> file. Use <code>ls -a</code> to list all files. % You have mail.
AaBbCc123	What you type, contrasted with on-screen computer output	% su Password:
<i>AaBbCc123</i>	Book titles, new words or terms, words to be emphasized	Read Chapter 6 in the <i>User's Guide</i> . These are called <i>class</i> options. You <i>must</i> be superuser to do this.
	Command-line variable; replace with a real name or value	To delete a file, type <code>rm filename</code> .

Shell Prompts

Shell	Prompt
C shell	<i>machine_name%</i>
C shell superuser	<i>machine_name#</i>
Bourne shell and Korn shell	\$
Bourne shell and Korn shell superuser	#

Related Documentation

Application	Title	Part Number
Hardware	<i>Sun Cluster 3.0 Hardware Guide</i>	806-1420
Data Services	<i>Sun Cluster 3.0 Data Services Installation and Configuration Guide</i>	805-1421
API Development	<i>Sun Cluster 3.0 Data Services Developers' Guide</i>	805-1422
Administration	<i>Sun Cluster 3.0 System Administration Guide</i>	806-1423
Concepts	<i>Sun Cluster 3.0 Concepts</i>	806-1424
Error Messages	<i>Sun Cluster 3.0 Error Messages Manual</i>	806-1426
Release Notes	<i>Sun Cluster 3.0 Release Notes</i>	805-1428

Accessing Sun Documentation Online

The `docs.sun.com`SM web site enables you to access Sun technical documentation on the Web. You can browse the `docs.sun.com` archive or search for a specific book title or subject at:

<http://docs.sun.com>

Ordering Sun Documentation

Fatbrain.com, an Internet professional bookstore, stocks select product documentation from Sun Microsystems, Inc.

For a list of documents and how to order them, visit the Sun Documentation Center on Fatbrain.com at:

<http://www1.fatbrain.com/documentation/sun>

Getting Help

If you have problems installing or using Sun Cluster, contact your service provider and provide the following information:

- Your name and email address (if available)
- Your company name, address, and phone number
- The model number and serial number of your systems
- The release number of the operating environment (for example, Solaris 8)
- The release number of Sun Cluster (for example, Sun Cluster 3.0)

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

Command	Function
<code>prtconf -v</code>	Displays the size of the system memory and reports information about peripheral devices
<code>psrinfo -v</code>	Displays information about processors
<code>showrev --p</code>	Reports which patches are installed
<code>prtdiag -v</code>	Displays system diagnostic information
<code>/usr/cluster/bin/scinstall -pv</code>	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.

- “Where to Find Sun Cluster Installation Tasks” on page 11
- “Planning the Solaris Operating Environment” on page 13
- “Planning the Sun Cluster Environment” on page 17
- “Planning the Global Devices and Cluster File Systems” on page 22
- “Planning Volume Management” on page 23

Where to Find Sun Cluster Installation Tasks

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

TABLE 1-1 Location of Sun Cluster Software Installation Task Information

Task	For Instructions, Go To ...
Setting up cluster hardware	<i>Sun Cluster 3.0 Hardware Guide</i> Documentation shipped with your server and storage devices
Planning cluster software installation	Chapter 1 “Configuration Worksheets and Examples” in <i>Sun Cluster 3.0 Release Notes</i>
Installing cluster framework, volume manager, and data service software packages	Chapter 2
Configuring cluster framework and volume manager software	Chapter 2 Appendix A or Appendix B Volume manager documentation
Upgrading cluster framework, data services, and volume manager software	Chapter 3 Appendix A or Appendix B Volume manager documentation
Planning, installing, and configuring data services and resource groups	<i>Sun Cluster 3.0 Data Services Installation and Configuration Guide</i>
Using the API	<i>Sun Cluster 3.0 Data Services Developers' Guide</i>

Planning the Solaris Operating Environment

This section provides guidelines for planning Solaris software installation in a cluster configuration. For more information about Solaris software, refer to the Solaris installation documentation.

Guidelines for Selecting Your Solaris Installation Method

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

Refer to “How to Use JumpStart to Install the Solaris Operating Environment and Establish New Cluster Nodes” on page 42 for details about the custom JumpStart installation method. Refer to Solaris installation documentation for details about standard Solaris installation methods.

System Disk Partitions

Add this information to the “Local File System Layout Worksheet” in *Sun Cluster 3.0 Release Notes*.

When the Solaris operating environment is installed, ensure that the required Sun Cluster partitions are created, and that all partitions meet minimum space requirements.

- **swap** – Allocate at least 750 Mbytes or twice the physical memory, whichever is greater.
- `/globaldevices` – Create a 100-Mbyte file system that will be used by the `scinstall(1M)` utility for global devices.
- **Volume manager** – Create a 10-Mbyte partition for volume manager use on a slice at the end of the disk (slice 7). 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 operating environment.

Refer to the following guidelines for additional partition planning information.

Guidelines for the Root (/) File System

As with any other system running the Solaris operating environment, you can configure the root (/), /var, /usr, and /opt directories as separate file systems, or you can include all the directories in the root (/) file system. The following describes the software contents of the root (/), /var, /usr, and /opt directories in a Sun Cluster configuration. Consider this information when planning your partitioning scheme.

- root (/) – The Sun Cluster software itself occupies less than 40 Mbytes of space in the root (/) file system. Solstice DiskSuite™ software requires less than 5 Mbytes, and VxVM software requires less than 15 Mbytes. For best results, configure ample additional space and inode capacity for the creation of both block special devices and character special devices used by either Solstice DiskSuite or VxVM software, especially if a large number of shared disks are in the cluster. Therefore, add at least 100 Mbytes to the amount of space you would normally allocate for your root (/) file system.
- /var – The Sun Cluster software occupies a negligible amount of space in /var at installation time. However, set aside ample space for log files. Also, more messages might be logged on a clustered node than would be found on a typical standalone server. Therefore, allow at least 100 Mbytes for /var.
- /usr – Sun Cluster software occupies less than 25 Mbytes of space in /usr. Solstice DiskSuite and VxVM software each require less than 15 Mbytes.
- /opt – Sun Cluster framework software uses less than 2 Mbytes in /opt. However, each Sun Cluster data service might use between 1 Mbyte and 5 Mbytes. Solstice DiskSuite software does not use any space in /opt. 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 /opt. If you use Sun™ Management Center software (formerly named Sun Enterprise SyMON™) to monitor the cluster, you need an additional 25 Mbytes of space on each node to support the Sun Management Center agent and Sun Cluster module packages.

Guidelines for the swap Partition

The minimum size of the swap partition must be either 750 Mbytes or twice the amount of physical memory on the machine, whichever is greater. In addition, any third-party applications you install might also have swap requirements. Refer to third-party application documentation for any swap requirements.

Guidelines for the /globaldevices File System

Sun Cluster software requires that you set aside a special file system one of the local disks for use in managing global devices. This file system must be separate, as it will later be mounted as a cluster file system. Name this file system `/globaldevices`, which is the default name recognized by the `scinstall(1M)` command. The `scinstall(1M)` command later renames the file system `/global/.devices/node@nodeid`, where *nodeid* represents the number assigned to a node when it becomes a cluster member, and the original `/globaldevices` mount point is removed. The `/globaldevices` file system must have ample space and inode capacity for creating both block special devices and character special devices, especially if a large number of disks are in the cluster. A file system size of 100 Mbytes should be more than enough for most cluster configurations.

Volume Manager Requirements

If you use Solstice DiskSuite software, you must set aside a slice on the root disk for use in creating the replica database. Specifically, set aside a slice for this purpose on each local disk. But, if you only have one local disk on a node, you might need to create three replica databases in the same slice for Solstice DiskSuite software to function properly. Refer to the Solstice DiskSuite documentation for more information.

If you use VxVM and you intend to encapsulate the root disk, you need two unused slices available for use by VxVM, as well as some additional unassigned free space at either the beginning or end of the disk. Refer to the VxVM documentation for more information about encapsulation.

Example—Sample File-System Allocations

Table 1-2 shows a partitioning scheme for a cluster node that has less than 750 Mbytes of physical memory. This scheme will be installed with the Solaris operating environment End User System Support software group, Sun Cluster software, and the Sun Cluster HA for NFS data service. The last slice on the disk, slice 7, has been allocated with a small amount of space for volume manager use.

This layout allows for the use of either Solstice DiskSuite software or VxVM. If you use Solstice DiskSuite software, you use slice 7 for the replica database. If you use VxVM, you can later free slice 7 by assigning it a zero length. This layout frees two slices, 4 and 7, and it provides for unused space at the end of the disk.

TABLE 1-2 Sample File-System Allocation

Slices	Allocation (in Mbytes)	Description
0	1168	441 Mbytes for Solaris operating environment software. 100 Mbytes extra for root (/). 100 Mbytes extra for /var. 25 Mbytes for Sun Cluster software. 55 Mbytes for volume manager software. 1 Mbyte for Sun Cluster HA for NFS software. 25 Mbytes for the Sun Management Center agent and Sun Cluster module agent packages. 421 Mbytes (the remaining free space on the disk) for possible future use by database and application software.
\$wap	750	Minimum size when physical memory is less than 750 Mbytes.
0verlap	2028	The entire disk.
3 globaldevices	100	The Sun Cluster software later assigns this slice a different mount point and mounts it as a cluster file system.
unused	-	Available as a free slice for encapsulating the root disk under VxVM.
unused	-	
unused	-	
volume manager	10	If Solstice DiskSuite software, used for replica database. If VxVM, later free the slice and some space at the end of the disk.

Planning the Sun Cluster Environment

This section provides guidelines for planning and preparing for Sun Cluster software installation. For detailed information about Sun Cluster components, refer to *Sun Cluster 3.0 Concepts*.

Licensing

Ensure that you have any necessary license certificates available 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, refer to the installation documentation for those products.

Software Patches

After installing each software product, you must also install any required patches. For the current list of required patches, refer to *Sun Cluster 3.0 Release Notes* or consult your Enterprise Services representative or service provider. Refer to *Sun Cluster 3.0 System Administration Guide* for general guidelines and procedures for applying patches.

IP Addresses

You must set up a number of IP addresses for various Sun Cluster components, depending on your cluster configuration. Each node 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 IP addresses assigned to them. Add these IP addresses to any naming services used. Also add these IP addresses to the local `/etc/inet/hosts` file on each cluster node after Sun Cluster software is installed.

TABLE 1-3 Sun Cluster Components That Use IP Addresses

Component	IP Addresses Needed
Administrative console	1 per subnet
Cluster nodes	1 per node, per subnet
Terminal concentrator or System Service Processor	1
Logical addresses	1 per logical host resource, per subnet

Terminal Concentrator or System Service Processor

A terminal concentrator communicates between the administrative console and the cluster node consoles. Sun Enterprise™ E10000 servers use a System Service Processor (SSP) instead of a terminal concentrator. For more information about console access, refer to *Sun Cluster 3.0 Concepts*.

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. Refer to *Sun Cluster 3.0 Data Services Installation and Configuration Guide* for information and worksheets for planning resource groups. For more information about data services and resources, also refer to *Sun Cluster 3.0 Concepts*.

Sun Cluster Configurable Components

This section provides guidelines for the Sun Cluster components that you configure during installation.

Cluster Name

Add this planning information to the “Cluster and Node Names Worksheet” in *Sun Cluster 3.0 Release Notes*.

You specify a name for the cluster during Sun Cluster installation. The cluster name should be unique throughout the enterprise.

Node Names

Add this planning information to the “Cluster and Node Names Worksheet” in *Sun Cluster 3.0 Release Notes*. Information for most other worksheets is grouped by node name.

The node name is the name you assign to a machine during installation of the Solaris operating environment. During Sun Cluster installation, you specify the names of all nodes that you are installing as a cluster.

Private Network

Add this planning information to the “Cluster and Node Names Worksheet” in *Sun Cluster 3.0 Release Notes*.

Sun Cluster software uses the private network for internal communication between nodes. Sun Cluster requires at least two connections to the cluster interconnect on the private network. You specify the private network address and netmask when installing Sun Cluster software on the first node of the cluster. You can choose to accept the default private network address (172.16.0.0) and netmask (255.255.0.0), or type different choices if the default network address is already in use elsewhere in the enterprise.

Note - After you have successfully installed the node as a cluster member, you cannot change the private network address and netmask.

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

- Use zeroes for the last two octets of the address
 - Follow the guidelines in RFC 1597 for network address assignments
- Refer to *TCP/IP and Data Communications Administration Guide* for instructions on obtaining copies of RFCs.

If you specify a netmask other than the default, it must meet the following requirements.

- Minimally mask all bits given in the private network address
- Have no “holes”

Cluster Interconnect

Add this planning information to the “Cluster Interconnect Worksheet” in *Sun Cluster 3.0 Release Notes*.

The cluster interconnect provides the hardware pathway for private network communication between cluster nodes. Each interconnect consists of a cable between

two transport adapters, a transport adapter and a transport junction, or two transport junctions. During Sun Cluster installation, you specify the following configuration information for two cluster interconnects.

- **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-node cluster, you also specify whether your interconnect is direct connected (adapter to adapter) or uses a transport junction.
- **Transport junctions** – If transport junctions, such as a network switch, are used, specify the transport junction name for each interconnect. The default name is `switchN`, where `N` is a number automatically assigned during installation. Also specify the junction port name, or accept the default name. The default port name is the same as the node ID of the node hosting the adapter end of the cable. However, you cannot use the default port name for certain adapter types, such as SCI.

Note - Clusters with three or more nodes *must* use transport junctions. Direct connection between cluster nodes is supported only for two-node clusters.

You can configure additional private network connections after installation by using the `scsetup(1M)` utility.

For more information about the cluster interconnect, refer to *Sun Cluster 3.0 Concepts*.

Private Hostnames

Add this planning information to the “Cluster and Node Names Worksheet” in *Sun Cluster 3.0 Release Notes*.

The private hostname is the name used for inter-node communication over the private network interface. Private hostnames are automatically created during Sun Cluster installation, and follow the naming convention `clusternodenodeid-priv`, where `nodeid` is the numeral of the internal node ID. This node ID number is automatically assigned during Sun Cluster installation to each node when it becomes a cluster member. After installation, you can rename private hostnames by using the `scsetup(1M)` utility.

Public Networks

Add this planning information to the “Public Networks Worksheet” in *Sun Cluster 3.0 Release Notes*.

Public networks communicate outside the cluster. Consider the following points when planning your public network configuration.

- Public networks and the private network (cluster interconnect) must use separate adapters.
- You must have at least one public network that is connected to all cluster nodes.
- You can have as many additional public network connections as your hardware configuration allows.

See also “NAFO Groups” on page 21 for guidelines on planning public network adapter backup groups. For more information about public network interfaces, refer to *Sun Cluster 3.0 Concepts*.

Disk Device Groups

Add this planning information to the “Disk Device Group Configurations Worksheet” in *Sun Cluster 3.0 Release Notes*.

You must configure all volume manager disk groups as Sun Cluster disk device groups. This configuration enables multihost disks to be hosted by a secondary node if the primary node fails. Consider the following points when planning disk device groups.

- **Failover** – You can configure multiported disks and properly-configured volume manager devices as failover devices. Proper configuration of a volume manager device includes multiported disks and correct setup of the volume manager itself so that the exported device can be hosted by multiple nodes. You cannot configure tape drives, CD-ROMs, or single-ported disks as failover devices.
- **Mirroring** – You must mirror the disks to protect the data from disk failure. Refer to your volume manager documentation for instructions on mirroring.

For more information about disk device groups, refer to *Sun Cluster 3.0 Concepts*.

NAFO Groups

Add this planning information to the “Public Networks Worksheet” in *Sun Cluster 3.0 Release Notes*.

A Network Adapter Failover (NAFO) group provides public network adapter monitoring and failover, and is the foundation for a network address resource. If the active adapter of a NAFO group that is configured with two or more adapters fails, all of its addresses fail over to another adapter in the NAFO group. In this way, the active NAFO group adapter maintains public network connectivity to the subnet to which the adapters in the NAFO group connect.

Consider the following points when planning your NAFO groups.

- Each public network adapter must belong to a NAFO group.

- Each node can have only one NAFO group per subnet.
- Only one adapter in a given NAFO group can have a hostname association, in the form of an `/etc/hostname.adapter` file.
- NAFO group naming convention is `nafoN`, where *N* is the number you supply when you create the NAFO group.

For more information about Network Adapter Failover, refer to *Sun Cluster 3.0 Concepts*.

Quorum Devices

Sun Cluster configurations use quorum devices to maintain data and resource integrity. If the cluster temporarily loses connection to a node, the quorum device prevents amnesia or split-brain problems when the cluster node attempts to rejoin the cluster. You assign quorum devices by using the `scsetup(1M)` utility.

Consider the following points when planning quorum devices.

- **Minimum** – A two-node cluster must have at least one shared disk assigned as a quorum device. For other topologies, quorum devices are optional.
- **Odd number rule** – If more than one quorum device is configured in a two-node cluster, or in a pair of nodes directly connected to the quorum device, configure an odd number of quorum devices so that the quorum devices have completely independent failure pathways.
- **Connection** – A quorum device cannot be connected to more than two nodes.

For more information about quorum, refer to *Sun Cluster 3.0 Concepts*.

Planning the Global Devices and Cluster File Systems

This section provides guidelines for planning global devices and cluster file systems. For more information about global devices and cluster files systems, refer to *Sun Cluster 3.0 Concepts*.

Guidelines for Highly Available Global Devices and Cluster File Systems

Sun Cluster does not require any specific disk layout or file system size. Consider the following points when planning your global device and cluster file system layout.

- **Mirroring** – All global devices must be mirrored to be considered highly available.
- **Disks** – When mirroring, lay out disks so that they are mirrored across disk expansion units.
- **Availability** – A global device must have a physical connection to more than one node in the cluster to be considered highly available. This configuration can tolerate a single-node failure. A global device with only one physical connection is supported, but it is inaccessible from other nodes if the node with the connection is down.

Mount Information for Cluster File Systems

Consider the following points when planning mount points for cluster file systems.

- **Mount point location** – Create mount points in the `/global` directory, unless prohibited by other software products. Using a `/global` directory enables you to easily distinguish cluster file systems, which are globally available, from local file systems.
- **Nesting mount points** – Normally, you should not nest the mount points for cluster file systems. For example, do not set up one file system mounted on `/global/a` and another file system mounted on `/global/a/b`. Ignoring this rule can cause availability and node boot order problems, because the parent mount point might not be present. The only exception to this rule is if the devices for the two file systems have the same physical node connectivity (for example, different slices on the same disk).

Planning Volume Management

This section provides guidelines for planning volume management of your cluster configuration.

Sun Cluster uses volume manager software to group disks into disk device groups that can then be administered as one unit. Sun Cluster supports Solstice DiskSuite

software and VERITAS Volume Manager (VxVM). You can use only one volume manager within a single cluster configuration. Refer to your volume manager documentation and to either Appendix A or Appendix B for instructions on configuring the volume manager software. For more information about volume management in a cluster configuration, refer to *Sun Cluster 3.0 Concepts*.

Add this planning information to the “Disk Device Group Configurations Worksheet” and the “Volume Manager Configurations Worksheet” in *Sun Cluster 3.0 Release Notes*, and to the “Metadevices Worksheet (Solstice DiskSuite)” in *Sun Cluster 3.0 Release Notes*, if applicable.

Guidelines for Volume Manager software

Consider the following general guidelines when configuring your disks.

- **Mirrored multihost disks** – You must mirror all multihost disks across disk expansion units. See “Mirroring Multihost Disks” on page 27 for guidelines on mirroring multihost disks.
- **Mirrored root** – Mirroring the root disk ensures high availability, but such mirroring is not required. See “Mirroring Guidelines” on page 27 for guidelines on deciding whether to mirror the root disk.
- **Unique naming** – On any cluster node, if a local Solstice DiskSuite metadevice or VxVM volume is used as the device on which the `/global/.devices/node@nodeid` file system is mounted, the name of that metadevice or volume must be unique throughout the cluster.
- **Node lists** – To ensure high availability of a disk 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 disk device group, make the scalable resource group’s node list a superset of the disk device group’s node list. Refer to the resource group planning information in *Sun Cluster 3.0 Data Services Installation and Configuration Guide* for information about node lists.
- **Multiported disks** – You must connect, or port, all disks used to construct a device group within the cluster to all of the nodes configured in the node list for that device group. Solstice DiskSuite software is able to automatically check for this at the time that disks are added to a diskset. However, configured VxVM disk groups do not have an association to any particular set of nodes. In addition, when you use the clustering software to register Solstice DiskSuite disksets, VxVM disk groups, or individual sets of global devices as global device groups, you can perform only limited connectivity checking.
- **Hot spare disks** – You can use hot spare disks to increase availability, but they are not required.

Refer to your volume manager documentation for disk layout recommendations and any additional restrictions.

Guidelines for Solstice DiskSuite

Consider the following points when planning Solstice DiskSuite configurations.

- **Mediators** – Each diskset configured with exactly two disk strings and mastered by exactly two nodes must have Solstice DiskSuite mediators configured for the diskset. A *disk string* consists of a disk enclosure, its physical disks, cables from the enclosure to the node(s), and the interface adapter cards. You must configure each diskset with exactly two nodes acting as mediator hosts. You must use the same two nodes for all disksets requiring mediators and those two nodes must master those disksets. Mediators cannot be configured for disksets that do not meet the two-string and two-host requirements. See the `mediator(7)` man page for details.
- `/kernel/drv/md.conf` **settings** – All metadevices used by each diskset are created in advance, at reconfiguration boot time, based on configuration parameters found in the `/kernel/drv/md.conf` file. The fields in the `md.conf` file are described in the Solstice DiskSuite documentation. You must modify the `nmd` and `md_nsets` fields as follows to support a Sun Cluster configuration.
 - `nmd` – The `nmd` field defines the number of metadevices created for each diskset. You must set the value of `nmd` to the predicted largest number of metadevices used by any one of the disksets in the cluster. For example, if a cluster uses 10 metadevices in its first 15 disksets, but 1000 metadevices in the 16th diskset, you must set the value of `nmd` to at least 1000. The maximum number of metadevices allowed per diskset is 8192.
 - `md_nsets` – The `md_nsets` field defines the total number of disksets that can be created for a system to meet the needs of the entire cluster. You must set the value of `md_nsets` to the expected number of disksets in the cluster, plus one to allow Solstice DiskSuite software to manage the private disks on the local host (that is, those metadevices that are not in the local diskset). The maximum number of disksets allowed per cluster is 32.

Set these fields at installation time to allow for all predicted future expansion of the cluster. Increasing these values after the cluster is in production is time consuming because it requires a reconfiguration reboot for each node. Raising these values later also increases the possibility of inadequate space allocation in the root (`/`) file system to create all of the requested devices.



Caution - All cluster nodes must have identical `/kernel/drv/md.conf` files, regardless of the number of disksets served by each node. Failure to follow this guideline can result in serious Solstice DiskSuite errors and possible loss of data.

Guidelines for VERITAS Volume Manager

Consider the following points when planning VERITAS Volume Manager (VxVM) configurations.

- **Root disk group** – You must create a default root disk group (`rootdg`) on each node. The `rootdg` disk group can be created on the following disks.
 - The root disk, which must be encapsulated
 - One or more local non-root disks, which can be encapsulated or initialized
 - A combination of root and local non-root disks

The `rootdg` disk group must be local to the node.

- **Encapsulation** – Disks to be encapsulated must have two disk-slice table entries free.
- **Number of volumes** – Estimate the maximum number of volumes any given disk device group will use at the time the disk 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 disk device group volumes. No two disk device groups can have overlapping minor number assignments.
- **Dirty Region Logging** – Using Dirty Region Logging (DRL) is highly recommended but not required. Using DRL decreases volume recovery time after a node failure. Using DRL might decrease I/O throughput.

File-System Logging

Logging is required for cluster file systems. Sun Cluster supports the following logging file systems.

- Solstice DiskSuite trans-metadevice UNIX file system (UFS) logging
- Solaris UFS logging

For information about Solstice DiskSuite trans-metadevice UFS logging, refer to your Solstice DiskSuite documentation. For information about Solaris UFS logging, refer to the `mount_ufs(1M)` man page and *Solaris Transition Guide*.

The following table lists the logging file systems supported by each volume manager.

TABLE 1-4 Supported File-System Logging Matrix

Volume Manager	Supported File-System Logging
Solstice DiskSuite	Solstice DiskSuite trans-metadevice UFS logging, Solaris UFS logging
VERITAS Volume Manager	Solaris UFS logging

Consider the following points when choosing between Solaris UFS logging and Solstice DiskSuite trans-metadevice UFS logging for your Solstice DiskSuite volume manager.

- **Solaris UFS log size** – Solaris UFS logging always allocates the log using free space on the UFS file system, and depending on the size of the file system.
 - On file systems less than 1 Gbyte, the log occupies 1 Mbyte.
 - On file systems 1 Gbyte or greater, the log occupies 1 Mbyte per Gbyte on the file system, up to a maximum of 64 Mbytes.
- **Log metadevice** – With Solstice DiskSuite trans-metadevice UFS logging, the trans device used for logging creates a metadevice. The log is yet another metadevice that you can mirror and stripe. Furthermore, you can create a maximum 1-Tbyte logging file system with the Solstice DiskSuite software.

Mirroring Guidelines

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

Mirroring Multihost Disks

Mirroring all multihost disks in a Sun Cluster configuration enables the configuration to tolerate single-disk failures. Sun Cluster software requires that you mirror all multihost disks across disk expansion units.

Consider the following points when mirroring multihost disks.

- **Separate disk expansion units** – Each submirror of a given mirror or plex should reside in a different multihost disk expansion unit.
- **Disk space** – Mirroring doubles the amount of necessary disk space.

- **Three-way mirroring** – Solstice DiskSuite software and VERITAS Volume Manager (VxVM) support three-way mirroring. However, Sun Cluster requires only two-way mirroring.
- **Number of metadevices** – Under Solstice DiskSuite software, mirrors consist of other metadevices such as concatenations or stripes. Large configurations might contain a large number of metadevices. For example, seven metadevices are created for each logging UFS file system.
- **Differing disk sizes** – If you mirror to a disk of a different size, your mirror capacity is limited to the size of the smallest submirror or plex.

For more information about multihost disks, refer to *Sun Cluster 3.0 Concepts*.

Mirroring the Root Disk

For maximum availability, you should mirror root (/), /usr, /var, /opt, and swap on the local disks. Under VxVM, you encapsulate the root disk and mirror the generated subdisks. However, mirroring the root disk is not a requirement of Sun Cluster.

Before deciding whether to mirror the root disk, consider the risks, complexity, cost, and service time for the various alternatives concerning the root disk. There is no single mirroring strategy that works for all configurations. You might want to consider your local Enterprise Services representative's preferred solution when deciding whether to mirror root.

Refer to your volume manager documentation and to either Appendix A or Appendix B for instructions on mirroring the root disk.

Consider the following issues when deciding whether to mirror the root disk.

- **Complexity** – Mirroring the root disk adds complexity to system administration and complicates booting in single-user mode.
- **Backups** – Regardless of whether or not 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** – Under Solstice DiskSuite software, in failure scenarios in which metadevice state database quorum is lost, you cannot reboot the system until maintenance is performed. Refer to the Solstice DiskSuite documentation for information about the metadevice state database and state database replicas.
- **Separate controllers** – Highest availability includes mirroring the root disk on a separate controller.
- **Boot disk** – You can set up the mirror to be a bootable root disk so that you can boot from the mirror if the primary boot disk fails.

- **Secondary root disk** – With a mirrored root disk, the primary root disk can fail but work can continue on the secondary (mirror) root disk. At a later point, the primary root disk might return to service (perhaps after a power cycle or transient I/O errors) and subsequent boots are performed by using the primary root disk specified in the OpenBoot™ PROM `boot-device` field. In this situation no manual repair task occurs, but the drive starts working well enough to boot. Note that a Solstice DiskSuite 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 (causing a stale submirror). For example, changes to the `/etc/system` file would be lost. Some Solstice DiskSuite 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 it is booting from a mirror or an underlying physical device, and the mirroring becomes active partway through the boot process (after the metadevices are loaded). Before this point, the system is vulnerable to stale submirror problems.

Installing and Configuring Sun Cluster Software

This chapter provides step-by-step procedures for installing and configuring your cluster.

The following step-by-step instructions are in this chapter.

- “How to Install Cluster Control Panel Software on the Administrative Console” on page 33
- “How to Install the Solaris Operating Environment” on page 36
- “How to Install Sun Cluster Software and Establish New Cluster Nodes” on page 38
- “How to Use JumpStart to Install the Solaris Operating Environment and Establish New Cluster Nodes” on page 42
- “How to Configure the Name Service Switch” on page 49
- “How to Install Solstice DiskSuite Software” on page 50
- “How to Install VERITAS Volume Manager Software” on page 52
- “How to Set Up the Root User’s Environment” on page 53
- “How to Install Data Service Software Packages” on page 54
- “How to Perform Post-Installation Setup” on page 56
- “How to Configure Volume Manager Software” on page 58
- “How to Add Cluster File Systems” on page 59
- “How to Configure Additional Public Network Adapters” on page 63
- “How to Configure Public Network Management (PNM)” on page 64
- “How to Change Private Hostnames” on page 65
- “How to Update Network Time Protocol (NTP)” on page 66

- “How to Install the Sun Cluster Module for Sun Management Center” on page 69
- “How to Start Sun Management Center Software” on page 70
- “How to Add a Cluster Node as a Sun Management Center Agent Host Object” on page 71
- “How to Load the Sun Cluster Module” on page 72

Installing the Software

Before you begin, read the following manuals for information that will help you plan your cluster configuration and prepare your installation strategy.

- *Sun Cluster 3.0 Concepts*—overview of the Sun Cluster 3.0 product
- *Sun Cluster 3.0 Release Notes*—late-breaking information
- This entire manual

The following table lists the tasks you perform to install the software.

TABLE 2-1 Task Map: Installing the Software

Task	For Instructions, Go To ...
Plan the layout of your cluster configuration.	Chapter 1 and “Configuration Worksheets and Examples” in <i>Sun Cluster 3.0 Release Notes</i>
<i>(Optional)</i> Install the Cluster Control Panel (CCP) software on the administrative console.	“How to Install Cluster Control Panel Software on the Administrative Console” on page 33
Install the Solaris operating environment and Sun Cluster software using one of two methods.	
<i>Method 1</i> – Install Solaris software, then install the Sun Cluster software by using the <code>scinstall</code> utility.	“How to Install the Solaris Operating Environment” on page 36 and “How to Install Sun Cluster Software and Establish New Cluster Nodes” on page 38
<i>Method 2</i> – Install Solaris software and Sun Cluster software in one operation by using the <code>scinstall</code> utility custom JumpStart option.	“How to Use JumpStart to Install the Solaris Operating Environment and Establish New Cluster Nodes” on page 42

TABLE 2-1 Task Map: Installing the Software (continued)

Task	For Instructions, Go To ...
Configure the name service look-up order.	“How to Configure the Name Service Switch” on page 49
Install volume manager software.	
Install Solstice DiskSuite software.	“How to Install Solstice DiskSuite Software” on page 50 and Solstice DiskSuite documentation
Install VERITAS Volume Manager software.	“How to Install VERITAS Volume Manager Software” on page 52 and VERITAS Volume Manager documentation
Set up directory paths.	“How to Set Up the Root User’s Environment” on page 53
Install data service software packages.	“How to Install Data Service Software Packages” on page 54
Configure the cluster.	“Configuring the Cluster” on page 55

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

This procedure describes how to install the Cluster Control Panel (CCP) software on the administrative console. The CCP provides a launchpad for the `cconsole(1M)`, `ctelnet(1M)`, and `crlogin(1M)` tools. Each of these tools provides a multiple-window connection to a set of nodes, plus a common window that sends input to all nodes at one time.

You can use any desktop machine running the Solaris 8 operating environment as an administrative console. In addition, you can also use the administrative console as a Sun Management Center console and/or server, and as an AnswerBook server. Refer to Sun Management Center documentation for information on installing Sun Management Center software. Refer to *Sun Cluster 3.0 Release Notes* for information on installing an AnswerBook server.

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.

- 1. Ensure that the Solaris 8 operating environment and any Solaris patches are installed on the administrative console.**

All platforms require Solaris 8 with at least the End User System Support software group.

- 2. If you are installing from the CD-ROM, insert the Sun Cluster 3.0 CD-ROM into the CD-ROM drive of the administrative console.**

If the Volume Management daemon `vold(1M)` is running and configured to manage CD-ROM devices, it automatically mounts the CD-ROM on the `/cdrom/suncluster_3_0` directory.

- 3. Change to the `/cdrom_image/suncluster_3_0/SunCluster_3.0/Packages` directory.**

```
# cd /cdrom_image/suncluster_3_0/SunCluster_3.0/Packages
```

- 4. Install the `SUNWcccon` package.**

```
# pkgadd -d . SUNWcccon
```

- 5. Install the `SUNWscman` package.**

```
# pkgadd -d . SUNWscman
```

Installing the `SUNWscman` package on the administrative console enables you to view Sun Cluster man pages from the administrative console prior to installing Sun Cluster software on the cluster nodes.

- 6. If you installed from a CD-ROM, eject the CD-ROM.**

- 7. Create an `/etc/clusters` file.**

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

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

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

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

Add the physical node name of each cluster node, the terminal concentrator (TC) or System Service Processor (SSP) name, and the serial port numbers to the file.

Note - Use the `telnet(1)` port numbers, not the physical port numbers, for the serial port numbers in the `/etc/serialports` file. Determine the serial port number by adding 5000 to the physical port number. For example, if a physical port number is 6, the serial port number should be 5006.

```
# vi /etc/serialports
node1 TC_hostname 500n
node2 TC_hostname 500n
```

See the `/opt/SUNWcluster/bin/serialports(4)` man page for details and special considerations for the Sun Enterprise E10000 server.

9. For convenience, add the `/opt/SUNWcluster/bin` directory to the `PATH` and the `/opt/SUNWcluster/man` directory to the `MANPATH` on the administrative console.

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

10. Start the CCP utility.

```
# /opt/SUNWcluster/bin/ccp clustername
```

Refer to the procedure “How to Remotely Log In to Sun Cluster” in *Sun Cluster 3.0 System Administration Guide* and the `/opt/SUNWcluster/bin/ccp(1M)` man page for information about using the CCP.

Where to Go From Here

To install Solaris software, go to “How to Install the Solaris Operating Environment” on page 36. To use the `scinstall` custom JumpStart option to install Solaris and Sun Cluster software, go to “How to Use JumpStart to Install the Solaris Operating Environment and Establish New Cluster Nodes” on page 42.

▼ How to Install the Solaris Operating Environment

If you are not using the `scinstall(1M)` custom JumpStart installation method to install software, perform this task on each node in the cluster.

1. **Ensure that the hardware setup is complete and connections are verified before installing Solaris software.**

Refer to *Sun Cluster 3.0 Hardware Guide* and your server and storage device documentation for details.

2. **On each node of the cluster, determine whether the `local-mac-address` variable is correctly set to `false`.**

```
# /usr/sbin/eeprom local-mac-address?
```

- If the command returns `local-mac-address=false`, the variable setting is correct. Proceed to Step 3 on page 43.
- If the command returns `local-mac-address=true`, change the setting to `false`.

```
# /usr/sbin/eeprom local-mac-address?=false
```

The new setting becomes effective at the next system reboot.

3. **Have available your completed “Local File System Layout Worksheet” from *Sun Cluster 3.0 Release Notes*.**

4. **Update naming services.**

Add address-to-name mappings for all public hostnames and logical addresses to any naming services (such as NIS, NIS+, or DNS) used by clients for access to cluster services. See “IP Addresses” on page 17 for planning guidelines.

You also add these addresses to the local `/etc/inet/hosts` file on each node during the procedure “How to Configure the Name Service Switch” on page 49.

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

If the Cluster Control Panel is installed and configured on your administrative console, you can use the `cconsole(1M)` utility to display the individual console screens. Otherwise, you must connect to the consoles of each node individually.

To save time, you can install the Solaris operating environment on each node at the same time. Use the `cconsole` utility to install all nodes at once.

6. Are you installing a new node to an existing cluster?

- If no, proceed to Step 7 on page 37.
- If yes, perform the following steps to create a mount point on the new node for each cluster file system in the cluster.
- a. **From another, active node of the cluster, display the names of all cluster file systems.**

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

b. On the node you are adding to the cluster, create a mount point for each cluster file system in the cluster.

```
% mkdir -p mountpoint
```

For example, if a file system name returned by the `mount` command was `/global/dg-schost-1`, run `mkdir -p /global/dg-schost-1` on the node being added to the cluster.

7. Install the Solaris operating environment as instructed in the Solaris installation documentation.

Note - You must install all nodes in a cluster with the same version of the Solaris operating environment.

You can use any method normally used for installing the Solaris operating environment to install the software on new nodes being installed into a clustered environment. These methods include the Solaris interactive installation program, Solaris JumpStart, and Solaris Web Start.

During installation, do the following.

- Install at least the End User System Support software group. You might need to install other Solaris software packages which are not part of the End User System Support software group, for example, the Apache HTTP server packages. Third-party software, such as Oracle, might also require additional Solaris packages. Refer to third-party documentation for any Solaris software requirements.

Note - Sun Enterprise E10000 servers require the Entire Distribution + OEM software group.

- Create a file system of at least 100 MBytes with its mount point set as `/globaldevices`, as well as any file-system partitions needed to support your volume manager software. Refer to “System Disk Partitions” on page 13 for partitioning guidelines to support Sun Cluster software.

Note - The `/globaldevices` file system is required for Sun Cluster software installation to succeed.

- Answer `no` when asked if you want automatic power-saving shutdown. You must disable automatic shutdown in Sun Cluster configurations. Refer to the `pmconfig(1M)` and `power.conf(4)` man pages for more information.
- For ease of administration, set the same root password on each node.

Note - The Solaris interface groups feature is disabled by default during Solaris software installation. Interface groups are not supported in a Sun Cluster configuration and should not be enabled. Refer to the `ifconfig(1M)` man page for more information about Solaris interface groups.

8. Install any Solaris software patches.

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

9. Install any hardware-related patches and download any needed firmware contained in the hardware patches.

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

Where to Go From Here

To install Sun Cluster software on your cluster nodes, go to “How to Install Sun Cluster Software and Establish New Cluster Nodes” on page 38.

▼ How to Install Sun Cluster Software and Establish New Cluster Nodes

After installing the Solaris operating environment, perform this task on each node of the cluster.

Note - If you used the `scinstall(1M)` custom JumpStart method to install software, the Sun Cluster software is already installed. Proceed to “How to Configure the Name Service Switch” on page 49.

1. Have available the following completed configuration planning worksheets from *Sun Cluster 3.0 Release Notes*.

- “Cluster and Node Names Worksheet”
- “Cluster Interconnect Worksheet”

See Chapter 1 for planning guidelines.

2. Become superuser on the cluster node.

3. If you are installing from the CD-ROM, insert the Sun Cluster 3.0 CD-ROM into the CD-ROM drive of the node you want to install and configure.

If the Volume Management daemon `vold(1M)` is running and configured to manage CD-ROM devices, it automatically mounts the CD-ROM on the `/cdrom/suncluster_3_0` directory.

4. Change to the `/cdrom_image/suncluster_3_0/SunCluster_3.0/Tools` directory.

```
# cd /cdrom_image/suncluster_3_0/SunCluster_3.0/Tools
```

5. Start the `scinstall(1M)` utility.

```
# ./scinstall
```

Follow these guidelines while using the interactive `scinstall` utility.

- Interactive `scinstall` enables you to type ahead. Therefore, do not press Return more than once if the next menu screen does not appear immediately.
- Unless otherwise noted, pressing Control-D will return you either to the start of a series of related questions or to the Main Menu.
- Your session answers are stored as defaults for the next time you run this menu option.
- Until the node has successfully booted in cluster mode, you can rerun `scinstall` and change the configuration information as needed. However, if bad configuration data for the node has been pushed over to the established portion of the cluster, you might first need to remove the bad information. To do this, log in to one of the active cluster nodes, then use the `scsetup(1M)` utility to remove the bad adapter, junction, or cable information.

6. To install the first node and establish the new cluster, type 1 (Establish a new cluster).

Follow the prompts to install Sun Cluster software, using the information from your configuration planning worksheets. You will be asked for the following information.

- Cluster name
- Names of the other nodes that will become part of this cluster
- Node authentication
- Private network address and netmask—You cannot change the private network address after the cluster has successfully formed
- Cluster interconnect (transport adapters and transport junctions)—You can configure no more than two adapters by using the `scinstall` command, but you can configure more adapters later by using the `scsetup` utility
- Global devices file-system name
- Automatic reboot—Do not choose automatic reboot if you have Sun Cluster software patches to install

When you finish answering the prompts, the `scinstall` command generated from your input is displayed for confirmation. If you choose not to accept the command, the `scinstall` utility returns you to the Main Menu. From there you can rerun menu option 1 and provide different answers. Your previous entries are displayed as the defaults.

Note - Unless you have installed your own `/etc/inet/ntp.conf` file, the `scinstall` command installs a default `ntp.conf` file for you. Because the default file is shipped with references to eight nodes, 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 Update Network Time Protocol (NTP)” on page 66 for information on how to suppress these messages under otherwise normal cluster conditions.

7. To install the second node of the cluster, type 2 (Add this machine as a node).

You can start this step while the first node is still being installed.

Follow the prompts to install Sun Cluster software, using the information from your configuration planning worksheets. You will be asked for the following information.

- Name of an existing cluster node, referred to as the *sponsor node*
- Cluster name
- Cluster interconnect (transport adapters and transport junctions)
- Global devices file-system name
- Automatic reboot—Do not choose automatic reboot if you have Sun Cluster software patches to install

When you finish answering the prompts, the `scinstall` command generated from your input is displayed for confirmation. If you choose not to accept the command, the `scinstall` utility returns you to the Main Menu. From there you can rerun menu option 2 and provide different answers. Your previous answers are displayed as the defaults.

If you choose to continue installation and the sponsor node is not yet established, `scinstall` waits for the sponsor node to become available.

8. Repeat Step 7 on page 40 on each additional node until all nodes are fully configured.

You do not need to wait for the second node to complete installation before beginning installation on additional nodes.

9. Install any Sun Cluster software patches.

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

10. If you installed Sun Cluster software patches, shut down the cluster, then reboot each node in the cluster.

Before rebooting the first node of the cluster, shut down the cluster by using the `scshutdown` command. Until the cluster nodes are removed from install mode, only the first node, which establishes the cluster (the sponsor node), has a quorum vote. In an established cluster which is still in install mode, if the cluster is not shut down before the first node is rebooted, the remaining cluster nodes cannot obtain quorum and the entire cluster shuts down.

Cluster nodes remain in install mode until the first time you run the `scsetup(1M)` command, during the procedure “How to Perform Post-Installation Setup” on page 56.

Example—Installing Sun Cluster Software

The following example shows the progress messages displayed as `scinstall` installation tasks are completed on the node `phys-schost-1`, which is the first node to be installed in the cluster.

```
** Installing SunCluster 3.0 **
SUNWscr....done.
SUNWscdev....done.
SUNWscu....done.
SUNWscman....done.
SUNWscsal....done.
SUNWscsam....done.
SUNWscrmop.done.
SUNWsci....done.
SUNWscid....done.
SUNWscidx....done.
SUNWscvm....done.
SUNWmdm....done.

Initializing cluster name to "scluster" ... done
Initializing authentication options ... done
Initializing configuration for adapter "hme2" ... done
```

(continued)

```

Initializing configuration for adapter "hme4" ... done
Initializing configuration for junction "switch1" ... done
Initializing configuration for junction "switch2" ... done
Initializing configuration for cable ... done
Initializing configuration for cable ... done
Setting the node ID for "phys-schost-1" ... done (id=1)

Checking for global devices global file system ... done
Checking device to use for global devices file system ... done
Updating vfstab ... done

Verifying that NTP is configured ... done
Installing a default NTP configuration ... done
Please complete the NTP configuration after scinstall has finished.

Verifying that "cluster" is set for "hosts" in nsswitch.conf ... done
Adding the "cluster" switch to "hosts" in nsswitch.conf ... done

Verifying that "cluster" is set for "netmasks" in nsswitch.conf ... done
Adding the "cluster" switch to "netmasks" in nsswitch.conf ... done

Verifying that power management is NOT configured ... done
Unconfiguring power management ... done
/etc/power.conf has been renamed to /etc/power.conf.060199105132
Power management is incompatible with the HA goals of the cluster.
Please do not attempt to re-configure power management.

Ensure routing is disabled ... done
Network routing has been disabled on this node by creating /etc/notrouter.
Having a cluster node act as a router is not supported by Sun Cluster.
Please do not re-enable network routing.

Log file - /var/cluster/logs/install/scinstall.log.276

Rebooting ...

```

Where to Go From Here

To set up the name service look-up order, go to “How to Configure the Name Service Switch” on page 49.

▼ How to Use JumpStart to Install the Solaris Operating Environment and Establish New Cluster Nodes

Perform this procedure to use the custom JumpStart installation method. This method installs the Solaris operating environment and Sun Cluster software on all cluster nodes in a single operation.

1. Ensure that the hardware setup is complete and connections are verified before installing Solaris software.

Refer to *Sun Cluster 3.0 Hardware Guide* and your server and storage device documentation for details on setting up the hardware.

2. On each node of the cluster, determine whether the `local-mac-address` variable is correctly set to `false`.

```
# /usr/sbin/eeprom local-mac-address?
```

- If the command returns `local-mac-address=false`, the variable setting is correct. Proceed to Step 3 on page 43.
- If the command returns `local-mac-address=true`, change the setting to `false`.

```
# /usr/sbin/eeprom local-mac-address?=false
```

The new setting becomes effective at the next system reboot.

3. Have available the following information.

- The Ethernet address of each cluster node
- The following completed configuration planning worksheets from *Sun Cluster 3.0 Release Notes*.
 - “Local File System Layout Worksheet”
 - “Cluster and Node Names Worksheet”
 - “Cluster Interconnect Worksheet”

See Chapter 1 for planning guidelines.

4. Update naming services.

Add address-to-name mappings for all public hostnames and logical addresses, as well as the IP address and hostname of the JumpStart server, to any naming services (such as NIS, NIS+, or DNS) used by clients for access to cluster services. See “IP Addresses” on page 17 for planning guidelines. You also add these addresses to the local `/etc/inet/hosts` file on each node during the procedure “How to Configure the Name Service Switch” on page 49.

Note - If you do not use a name service, create `jumpstart-dir/autoscinstall.d/nodes/nodename/archive/etc/inet/hosts` files on the JumpStart install server, one file for each node of the cluster, where *nodename* is the name of a node of the cluster. Add the address-to-name mappings there.

5. As superuser, set up the JumpStart install server for Solaris operating environment installation.

Refer to the `setup_install_server(1M)` and `add_install_client(1M)` man pages and *Solaris Advanced Installation Guide* for instructions on setting up a JumpStart install server.

When setting up the install server, ensure that the following requirements are met.

- The install server is on the same subnet as the cluster nodes, but is not itself a cluster node.
- The install server installs the release of the Solaris operating environment required by the Sun Cluster software.
- A custom JumpStart directory exists for JumpStart installation of Sun Cluster. This `jumpstart-dir` directory must contain a copy of the `check(1M)` utility and be NFS exported for reading by the JumpStart install server.
- Each new cluster node is configured as a custom JumpStart install client using the custom JumpStart directory set up for Sun Cluster installation.

6. Create a directory on the JumpStart install server to hold your copies of the Sun Cluster and Sun Cluster data services CD-ROMs.

In the following example, the `/export/suncluster` directory is created for this purpose.

```
# mkdir -m 755 /export/suncluster
```

7. Copy the Sun Cluster CD-ROM to the JumpStart install server.

a. Insert the Sun Cluster 3.0 CD-ROM into the CD-ROM drive on the JumpStart install server.

If the Volume Management daemon `vold(1M)` is running and configured to manage CD-ROM devices, it automatically mounts the CD-ROM on the `/cdrom/suncluster_3_0` directory.

b. Change to the `/cdrom_image/suncluster_3_0/SunCluster_3.0/Tools` directory.

```
# cd /cdrom_image/suncluster_3_0/SunCluster_3.0/Tools
```

c. Copy the CD-ROM to a new directory on the JumpStart install server.

The `scinstall` command creates the new install directory as it copies the CD-ROM files. The install directory name `/export/suncluster/sc30` is used here as an example.

```
# ./scinstall -a /export/suncluster/sc30
```

d. Eject the CD-ROM.

```
# cd /  
# eject cdrom
```

e. Ensure that the Sun Cluster 3.0 CD-ROM image on the JumpStart install server is NFS exported for reading by the JumpStart install server.

Refer to *NFS Administration Guide* and the `share(1M)` and `dfstab(4)` man pages for more information about automatic file sharing.

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

The path `/export/suncluster/sc30` is used here as an example of the install directory you created.

```
# cd /export/suncluster/sc30/SunCluster_3.0/Tools  
# ./scinstall
```

Follow these guidelines while using the interactive `scinstall` utility.

- Interactive `scinstall` enables you to type ahead. Therefore, do not press Return more than once if the next menu screen does not appear immediately.
- Unless otherwise noted, pressing Control-D returns you either to the start of a series of related questions or to the Main Menu.
- Your session answers are stored as defaults for the next time you run this menu option.

9. To choose JumpStart installation, type 3 (Configure a cluster to be JumpStarted from this install server).

Note - If option 3 does not have an asterisk in front, this omission indicates the option is disabled because JumpStart setup is not complete or has an error. Exit the `scinstall` utility, correct JumpStart setup, then restart the `scinstall` utility.

Follow the prompts to specify Sun Cluster configuration information.

- JumpStart directory name

- Cluster name
- Cluster node names
- Node authentication
- Private network address and netmask—You cannot change the private network address after the cluster has successfully formed
- Cluster interconnect (transport adapters and transport junctions)—You can configure no more than two adapters by using the `scinstall` command, but you can configure additional adapters later by using the `scsetup` utility
- Global devices file-system name
- Automatic reboot—Do not choose automatic reboot if you have Sun Cluster software patches to install

When finished, the `scinstall` commands generated from your input are displayed for confirmation. If you choose not to accept one of them, the `scinstall` utility returns you to the Main Menu. From there you can rerun menu option 3 and provide different answers. Your previous entries are displayed as the defaults.

10. If necessary, make adjustments to the default `class` file, or profile, created by `scinstall`.

The `scinstall` command creates the following `autoscinstall.class` default class file in the `jumpstart-dir/autoscinstall.d/3.0` directory.

```
install_type      initial_install
system_type       standalone
partitioning      explicit
filesystems       rootdisk.s0 free /
filesystems       rootdisk.s1 750 swap
filesystems       rootdisk.s3 100 /globaldevices
filesystems       rootdisk.s7 10
cluster           SUNWCuser      add
package           SUNWman        add
```

Note - The default `class` file installs the End User System Support software group (SUNWCuser) of Solaris software. For Sun Enterprise E10000 servers, you must install the Entire Distribution + OEM software group. Also, some third-party software, such as Oracle, might require additional Solaris packages. Refer to third-party documentation for any Solaris software requirements.

You can change the profile 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 minimum file-system allocation requirements are met, no restrictions are imposed on changes to the Solaris operating environment install profile. Refer to “System Disk Partitions” on page 13 for partitioning guidelines and requirements to support Sun Cluster 3.0 software.

11. Are you installing a new node to an existing cluster?

- If no, proceed to Step 12 on page 47.
 - If yes, perform the following steps to create a mount point on the new node for each cluster file system in the cluster.
- a. **From another, active node of the cluster, display the names of all cluster file systems.**

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

- b. **On the node you are adding to the cluster, create a mount point for each cluster file system in the cluster.**

```
% mkdir -p mountpoint
```

For example, if a file system name returned by the `mount` command is `/global/dg-schost-1`, run `mkdir -p /global/dg-schost-1` on the node being added to the cluster.

12. Set up Solaris patch directories.

- a. **Create `jumpstart-dir/autoscinstall.d/nodes/nodename/patches` directories on the JumpStart install server, one directory for each node in the cluster, where `nodename` is the name of a cluster node.**

```
# mkdir jumpstart-dir/autoscinstall.d/nodes/nodename/patches
```

- b. **Place copies of any Solaris patches into each of these directories. Also place copies of any hardware-related patches that must be installed after Solaris software is installed into each of these directories.**

13. If you do not use a name service, set up files to contain the necessary hostname information.

- a. **On the JumpStart install server, create files named `jumpstart-dir/autoscinstall.d/nodes/nodename/archive/etc/inet/hosts`.**

Create one file for each node, where *nodename* is the name of a cluster node.

b. Add the following entries into each file.

- IP address and hostname of the NFS server that holds a copy of the Sun Cluster CD-ROM image. This could be the JumpStart install server or another machine.
- IP address and hostname of each node in the cluster.

14. Add your own post-installation finish script.

You can add your own finish script, which is run after the standard finish script installed by the `scinstall` command.

- a. Name your finish script `finish`.
- b. Copy your finish script to the `jumpstart-dir/autosinstall.d/nodes/nodename` directory, one directory for each node in the cluster.

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

If `cconsole(1M)` is installed and configured on your administrative console, you can use it to display the individual console screens. Otherwise, you must connect to the consoles of each node individually.

16. From the `ok` PROM prompt on the console of each node, type the `boot net - install` command to begin the network JumpStart installation of each node.

Note - The dash (-) in the command must be surrounded by a space on each side.

```
ok boot net - install
```

Note - Unless you have installed your own `ntp.conf` file in the `/etc/inet` directory, the `scinstall` command installs a default `ntp.conf` file for you. Because the default file is shipped with references to eight nodes, 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 Update Network Time Protocol (NTP)” on page 66 for information on how to suppress these messages under otherwise normal cluster conditions.

When the installation is successfully completed, each node is fully installed as a new cluster node.

Note - The Solaris interface groups feature is disabled by default during Solaris software installation. Interface groups are not supported in a Sun Cluster configuration and should not be enabled. Refer to the `ifconfig(1M)` man page for more information about Solaris interface groups.

17. Install any Sun Cluster software patches.

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

18. If you installed Sun Cluster software patches, shut down the cluster, then reboot each node in the cluster.

Before rebooting the first node of the cluster, shut down the cluster by using the `scshutdown` command. Until the cluster nodes are removed from install mode, only the first node, which establishes the cluster (the sponsor node), has a quorum vote. In an established cluster which is still in install mode, if the cluster is not shut down before the first node is rebooted, the remaining cluster nodes cannot obtain quorum and the entire cluster shuts down.

Cluster nodes remain in install mode until the first time you run the `scsetup(1M)` command, during the procedure “How to Perform Post-Installation Setup” on page 56.

Where to Go From Here

To set up the name service look-up order, go to “How to Configure the Name Service Switch” on page 49.

▼ How to Configure the Name Service Switch

Perform this task on each node in the cluster.

1. **Become superuser on the cluster node.**
2. **Edit the `/etc/nsswitch.conf` file.**
 - a. **Verify that `cluster` is the first source look-up for the `hosts` and `netmasks` database entries.**

This order is necessary for Sun Cluster software to function properly. The `scinstall(1M)` command adds `cluster` to these entries during installation.
 - b. **For the `hosts` and `netmasks` database entries, follow `cluster` with files.**

- c. For all other database entries, place files first in look-up order.

Note - Performing Step 2 on page 49 and Step 2 on page 50 can increase availability to data services if the naming service becomes unavailable.

The following example shows partial contents of an `/etc/nsswitch.conf` file. The look-up order for the `hosts` and `netmasks` database entries is first `cluster`, then `files`. The look-up order for other entries begins with `files`.

```
# vi /etc/nsswitch.conf
...
passwd:      files nis
group:       files nis
...
hosts:       cluster files nis
...
netmasks:   cluster files nis
...
```

3. Update the `/etc/inet/hosts` file with all public hostnames and logical addresses for the cluster.

Where to Go From Here

To install Solstice DiskSuite volume manager software, go to “How to Install Solstice DiskSuite Software” on page 50. To install VERITAS Volume Manager volume manager software, go to “How to Install VERITAS Volume Manager Software” on page 52.

▼ How to Install Solstice DiskSuite Software

Perform this task on each node in the cluster.

1. Become superuser on the cluster node.
2. If you are installing from the CD-ROM, insert the Solaris 8 Software 2 of 2 CD-ROM into the CD-ROM drive on the node.

Note - Solstice DiskSuite software packages are now located on the Solaris 8 software CD-ROM.

This step assumes that the Volume Management daemon `vold(1M)` is running and configured to manage CD-ROM devices.

3. Install the Solstice DiskSuite software packages.

Note - If you have Solstice DiskSuite software patches to install, do not reboot after installing the Solstice DiskSuite software.

Install software packages in the order shown in the following example.

```
# cd /cdrom_image/sol_8_sparc_2/solaris_8/EA/products/DiskSuite_4.2.1/sparc/Packages
# pkgadd -d . SUNWmdr SUNWmdu [SUNWmdx] optional-pkgs
```

The `SUNWmdr` and `SUNWmdu` packages are required for all Solstice DiskSuite installations. The `SUNWmdx` package is also required for the 64-bit Solstice DiskSuite installation. Refer to your Solstice DiskSuite installation documentation for information about optional software packages.

4. If you installed from a CD-ROM, eject the CD-ROM.

5. If not already installed, install any Solstice DiskSuite patches.

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

6. Manually populate the global device namespace for Solstice DiskSuite by running the `/usr/cluster/bin/scgdevs` command.

7. If you installed Solstice DiskSuite software patches, shut down the cluster, then reboot each node in the cluster.

Before rebooting the first node of the cluster, shut down the cluster by using the `scshutdown` command. Until the cluster nodes are removed from install mode, only the first node, which establishes the cluster (the sponsor node), has a quorum vote. In an established cluster which is still in install mode, if the cluster is not shut down before the first node is rebooted, the remaining cluster nodes cannot obtain quorum and the entire cluster shuts down.

Cluster nodes remain in install mode until the first time you run the `scsetup(1M)` command, during the procedure “How to Perform Post-Installation Setup” on page 56.

Refer to your Solstice DiskSuite installation documentation for complete information about installing Solstice DiskSuite software.

Where to Go From Here

To set up your root user's environment, go to "How to Set Up the Root User's Environment" on page 53.

▼ How to Install VERITAS Volume Manager Software

Perform this task on each node in the cluster.

1. **Become superuser on the cluster node.**
2. **Disable Dynamic Multipathing (DMP).**

```
# mkdir /dev/vx
# ln -s /dev/dsk /dev/vx/dmp
# ln -s /dev/rdisk /dev/vx/rdmp
```

3. **Insert the VxVM CD-ROM into the CD-ROM drive on the node.**
4. **Install the VxVM software packages.**

Note - If you have VxVM software patches to install, do not reboot after installing the VxVM software.

```
# cd /cdrom_image/volume_manager_3_0_4_solaris/pkg
# pkgadd -d . VRTSvxvm VRTSvmdev VRTSvmman
```

List `VRTSvxvm` first in the `pkgadd(1M)` command and `VRTSvmdev` second. Refer to your VxVM installation documentation for descriptions of the other VxVM software packages.

Note - The `VRTSvxvm` and `VRTSvmdev` packages are required for all VxVM installations.

5. Eject the CD-ROM.

6. Install any VxVM patches.

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

7. If you installed VxVM software patches, shut down the cluster, then reboot each node in the cluster.

Before rebooting the first node of the cluster, shut down the cluster by using the `scshutdown` command. Until the cluster nodes are removed from install mode, only the first node, which establishes the cluster (the sponsor node), has a quorum vote. In an established cluster which is still in install mode, if the cluster is not shut down before the first node is rebooted, the remaining cluster nodes cannot obtain quorum and the entire cluster shuts down.

Cluster nodes remain in install mode until the first time you run the `scsetup(1M)` command, during the procedure “How to Perform Post-Installation Setup” on page 56.

Refer to your VxVM installation documentation for complete information about installing VxVM software.

Where to Go From Here

To set up your root user’s environment, go to “How to Set Up the Root User’s Environment” on page 53.

▼ How to Set Up the Root User’s Environment

Perform these tasks on each node in the cluster.

1. Become superuser on the cluster node.

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

For VERITAS Volume Manager, also set your `PATH` to include `/etc/vx/bin`. If you installed the `VRTSvmsa` package, also add `/opt/VRTSvmsa/bin` to your `PATH`.

3. Set the `MANPATH` to include `/usr/cluster/man`. Also include the volume manager-specific paths.

- For Solstice DiskSuite software, set your MANPATH to include `/usr/share/man`.
 - For VERITAS Volume Manager, set your MANPATH to include `/opt/VRTSvxvm/man`. If you installed the `VRTSvmsa` package, also add `/opt/VRTSvmsa/man` to your MANPATH.
4. **For ease of administration, set the same root password on each node, if you have not already done so.**

Where to Go From Here

To install data service software packages, go to “How to Install Data Service Software Packages” on page 54.

▼ How to Install Data Service Software Packages

Perform this task on each cluster node.

Note - You must install the same set of data service packages on each node, even if a node is not expected to host resources for an installed data service.

1. **Become superuser on the cluster node.**
2. **If you are installing from the CD-ROM, insert the Data Services CD-ROM into the CD-ROM drive on the node.**
3. **Start the `scinstall(1M)` utility.**

```
# scinstall
```

Follow these guidelines while using the interactive `scinstall` utility.

- Interactive `scinstall` enables you to type ahead. Therefore, do not press Return more than once if the next menu screen does not appear immediately.
 - Unless otherwise noted, pressing Control-D returns you either to the start of a series of related questions or to the Main Menu.
4. **To add data services, type 4 (Add support for a new data service to this cluster node).**
Follow the prompts to select all data services you want to install.
 5. **If you installed from a CD-ROM, eject the CD-ROM.**
 6. **Install any Sun Cluster data service patches.**

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

Note - You do not have to reboot after installing Sun Cluster data service patches, unless specified by the patch special instructions. If a patch instruction requires that you reboot, before rebooting the first node of the cluster, shut down the cluster by using the `scshutdown` command. Until the cluster nodes are removed from install mode, only the first node, which establishes the cluster (the sponsor node), has a quorum vote. In an established cluster which is still in install mode, if the cluster is not shut down before the first node is rebooted, the remaining cluster nodes cannot obtain quorum and the entire cluster shuts down. Cluster nodes remain in install mode until the first time you run the `scsetup(1M)` command, during the procedure “How to Perform Post-Installation Setup” on page 56.

Where to Go From Here

For post-installation setup and configuration tasks, see “Configuring the Cluster” on page 55.

Configuring the Cluster

The following table lists the tasks to perform to configure your cluster.

TABLE 2-2 Task Map: Configuring the Cluster

Task	For Instructions, Go To ...
Perform post-installation setup	“How to Perform Post-Installation Setup” on page 56
Configure the Solstice DiskSuite or VERITAS Volume Manager volume manager and device groups.	“How to Configure Volume Manager Software” on page 58, and volume manager documentation
Create and mount cluster file systems.	“How to Add Cluster File Systems” on page 59

TABLE 2-2 Task Map: Configuring the Cluster *(continued)*

Task	For Instructions, Go To ...
<i>(Optional)</i> Configure additional public network adapters.	“How to Configure Additional Public Network Adapters” on page 63
Configure Public Network Management (PNM) and set up NAFO groups.	“How to Configure Public Network Management (PNM)” on page 64
<i>(Optional)</i> Change a node’s private hostname.	“How to Change Private Hostnames” on page 65
Edit the <code>/etc/inet/ntp.conf</code> file to update node name entries.	“How to Update Network Time Protocol (NTP)” on page 66
<i>(Optional)</i> Install the Sun Cluster module to Sun Management Center software.	“Installation Requirements for Sun Management Center Software for Sun Cluster Monitoring” on page 68 and Sun Management Center documentation
Install third-party applications and configure the applications, data services, and resource groups.	<i>Sun Cluster 3.0 Data Services Installation and Configuration Guide</i> , and third-party application documentation

▼ How to Perform Post-Installation Setup

Perform this procedure one time only, after the cluster is fully formed.

1. Verify that all nodes have joined the cluster.

- a. **From one node, display a list of cluster nodes to verify that all nodes have joined the cluster.**

You do not need to be logged in as superuser to run this command.

```
% scstat -n
```

Output resembles the following.

```
-- Cluster Nodes --
      Node name      Status
      -
      -
```

(continued)

```
Cluster node:  phys-schost-1  Online
Cluster node:  phys-schost-2  Online
```

- b. On each node, display a list of all the devices that the system checks to verify their connectivity to the cluster nodes.**

You do not need to be logged in as superuser to run this command.

```
% scdidadm -L
```

The list on each node should be the same. Output resembles the following.

```
1      phys-schost-1:/dev/rdisk/c0t0d0 /dev/did/rdsk/d1
2      phys-schost-1:/dev/rdisk/c1t1d0 /dev/did/rdsk/d2
2      phys-schost-2:/dev/rdisk/c1t1d0 /dev/did/rdsk/d2
3      phys-schost-1:/dev/rdisk/c1t2d0 /dev/did/rdsk/d3
3      phys-schost-2:/dev/rdisk/c1t2d0 /dev/did/rdsk/d3
...

```

- c. Identify from the `scdidadm` output the global device ID (DID) name of each shared disk you will configure as a quorum device.**

For example, the output in the previous substep shows that global device `d2` is shared by `phys-schost-1` and `phys-schost-2`. You need this information in Step 4 on page 58. Refer to “Quorum Devices” on page 22 for further information about planning quorum devices.

2. Become superuser on one node of the cluster.
3. Start the `scsetup(1M)` utility.

```
# scsetup
```

The Initial Cluster Setup screen is displayed.

Note - If the Main Menu is displayed instead, this procedure has already been successfully performed.

4. Respond to the prompts.

- a. **At the prompt** Do you want to add any quorum disks?, **configure at least one shared quorum device if your cluster is a two-node cluster.**

A two-node cluster remains in install mode until a shared quorum device is configured. After the `scsetup` utility configures the quorum device, the message `Command completed successfully` is displayed. If your cluster has three or more nodes, configuring a quorum device is optional.

- b. **At the prompt** Is it okay to reset "installmode"?, **answer** Yes.

After the `scsetup` utility sets quorum configurations and vote counts for the cluster, the message `Cluster initialization is complete` is displayed and the utility returns you to the Main Menu.

Note - If the quorum setup process is interrupted or fails to complete successfully, rerun Step 3 on page 57 and Step 4 on page 58.

5. From any node, verify that cluster install mode is disabled.

```
# scconf -p | grep "Cluster install mode:"
Cluster install mode:                                disabled
```

Where to Go From Here

To configure volume manager software, go to “How to Configure Volume Manager Software” on page 58.

▼ How to Configure Volume Manager Software

1. Have available the following information.

- Mappings of your storage disk drives
- The following completed configuration planning worksheets from *Sun Cluster 3.0 Release Notes*.
 - “Local File System Layout Worksheet”
 - “Disk Device Group Configurations Worksheet”
 - “Volume Manager Configurations Worksheet”
 - “Metadevices Worksheet (Solstice DiskSuite)”

See Chapter 1 for planning guidelines.

2. Follow the appropriate configuration procedures for your volume manager.

Volume Manager	Documentation
Solstice DiskSuite	Appendix A Solstice DiskSuite documentation
VERITAS Volume Manager	Appendix B VERITAS Volume Manager documentation

Where to Go From Here

After configuring your volume manager, to create a cluster file system, go to “How to Add Cluster File Systems” on page 59.

▼ How to Add Cluster File Systems

Perform this task for each cluster file system you add.



Caution - Creating a file system destroys any data on the disks. Be sure you have specified the correct disk device name. If you specify the wrong device name, you erase its contents when the new file system is created.

1. Become superuser on any node in the cluster.

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

2. Create a file system by using the `newfs(1M)` command.

```
# 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.

TABLE 2-3 Sample Raw Disk Device Names

Volume Manager	Sample Disk Device Name	Description
Solstice DiskSuite	/dev/md/oracle/rdisk/d1	Raw disk device d1 within the oracle diskset
VERITAS Volume Manager	/dev/vx/rdisk/oradg/vol01	Raw disk device vol01 within the oradg disk group
None	/dev/global/rdisk/d1s3	Raw disk device d1s3

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 will not be accessed on that node.

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

- device-group*** Name of the directory that corresponds to the name of the device group that contains the device
- mount-point*** Name of the directory on which to mount the cluster file system

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.

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

Note - The `syncdir` mount option is not required for cluster file systems. If you specify `syncdir`, you are guaranteed POSIX-compliant file system behavior. If you do not, you will have the same behavior that is seen with UFS file systems. Not specifying `syncdir` can significantly improve performance of writes that allocate disk blocks, such as when appending data to a file. However, in some cases, without `syncdir` you would not discover an out-of-space condition until you close a file. The cases in which you could have problems if you do not specify `syncdir` are rare. With `syncdir` (and POSIX behavior), the out-of-space condition would be discovered before the close.

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

b. Use the following required mount options.

- If you are using Solaris UFS logging, use the `global,logging` mount options.
- If a cluster file system uses a Solstice DiskSuite trans metadvice, use the `global` mount option (do not use the `logging` mount option). Refer to Solstice DiskSuite documentation for information about setting up trans metadvicees.

Note - Logging is required for all cluster file systems.

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

d. 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`.

e. Make sure the entries in each node's `/etc/vfstab` file list devices in the same order.

Refer to the `vfstab(4)` man page for details.

5. On any node in the cluster, verify that mount points exist and `/etc/vfstab` file entries are correct on all nodes of the cluster.

```
# sccheck
```

If no errors occur, nothing is returned.

6. From any node in the cluster, mount the cluster file system.

```
# mount /global/device-group/mount-point
```

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

You can use either the `df(1M)` or `mount(1M)` command to list mounted file systems.

Example—Creating a Cluster File System

The following example creates a UFS cluster file system on the Solstice DiskSuite metadevice `/dev/md/oracle/rdisk/d1`.

```
# newfs /dev/md/oracle/rdisk/d1
...

(on each node:)
# mkdir -p /global/oracle/d1
# vi /etc/vfstab
#device      device      mount  FS      fsck    mount  mount
#to mount    to fsck     point  type    pass   at boot options
#
/dev/md/oracle/dsk/d1 /dev/md/oracle/rdisk/d1 /global/oracle/d1 ufs 2 yes global,logging
(save and exit)

(on one node:)
# sccheck
# mount /global/oracle/d1
# mount
...
/global/oracle/d1 on /dev/md/oracle/dsk/d1 read/write/setuid/global/logging/
largefiles on Sun Oct 3 08:56:16 1999
```

Where to Go From Here

If your cluster nodes are connected to more than one public subnet, to configure additional public network adapters, go to “How to Configure Additional Public Network Adapters” on page 63.

Otherwise, to configure PNM and set up NAFO groups, go to “How to Configure Public Network Management (PNM)” on page 64.

▼ How to Configure Additional Public Network Adapters

If the nodes in the cluster are connected to more than one public subnet, you can configure additional public network adapters for the secondary subnets. However, configuring secondary subnets is not required.

Note - Configure only public network adapters, not private network adapters.

1. **Have available your completed “Public Networks Worksheet” from *Sun Cluster 3.0 Release Notes*.**
2. **Become superuser on the node being configured for additional public network adapters.**
3. **Create a file named `/etc/hostname.adapter`, where *adapter* is the adapter name.**

Note - In each NAFO group, an `/etc/hostname.adapter` file should exist for only one adapter in the group.

4. **Type the hostname of the public network adapter IP address in the `/etc/hostname.adapter` file.**

For example, the following shows the file `/etc/hostname.hme3`, created for the adapter `hme3`, which contains the hostname `phys-schost-1`.

```
# vi /etc/hostname.hme3
phys-schost-1
```

5. **On each cluster node, ensure that the `/etc/inet/hosts` file contains the IP address and corresponding hostname assigned to the public network adapter.**

For example, the following shows the entry for `phys-schost-1`.

```
# vi /etc/inet/hosts
...
192.29.75.101 phys-schost-1
...
```

Note - If you use a naming service, this information should also exist in the naming service database.

6. On each cluster node, turn on the adapter.

```
# ifconfig adapter plumb
# ifconfig adapter hostname netmask + broadcast + -trailers up
```

7. Verify that the adapter is configured correctly.

```
# ifconfig adapter
```

The output should contain the correct IP address for the adapter.

Where to Go From Here

Each public network adapter to be managed by the Resource Group Manager (RGM) must belong to a NAFO group. To configure PNM and set up NAFO groups, go to “How to Configure Public Network Management (PNM)” on page 64.

▼ How to Configure Public Network Management (PNM)

Perform this task on each node of the cluster.

Note - All public network adapters *must* belong to a Network Adapter Failover (NAFO) group. Also, each node can have only one NAFO group per subnet.

- 1. Have available your completed “Public Networks Worksheet” from *Sun Cluster 3.0 Release Notes*.**
- 2. Become superuser on the node being configured for a NAFO group.**
- 3. Create the NAFO group.**

```
# pnmset -c nafo_group -o create adapter [adapter ...]
```

-c *nafo_group* Configures the NAFO group *nafo_group*

- o create **adapter** Creates a new NAFO group containing one or more public network adapters

Refer to the `pnmset(1M)` man page for more information.

4. Verify the status of the NAFO group.

```
# pnmstat -l
```

Refer to the `pnmstat(1M)` man page for more information.

Example—Configuring PNM

The following example creates NAFO group `nafo0`, which uses public network adapters `qfe1` and `qfe5`.

```
# pnmset -c nafo0 -o create qfe1 qfe5
# pnmstat -l
group  adapters      status  fo_time  act_adp
nafo0  qfe1:qfe5        OK      NEVER    qfe5
nafo1  qfe6              OK      NEVER    qfe6
```

Where to Go From Here

If you want to change any private hostnames, go to “How to Change Private Hostnames” on page 65. Otherwise, to update the `/etc/inet/ntp.conf` file, go to “How to Update Network Time Protocol (NTP)” on page 66.

▼ How to Change Private Hostnames

Perform this task if you do not want to use the default private hostnames (`clusternode $nodeid$ -priv`) assigned during Sun Cluster software installation.

Note - This procedure should *not* be performed after applications and data services have been configured and started. Otherwise, an application or data service might continue using the old private hostname after it has been renamed, causing hostname conflicts. If any applications or data services are running, stop them before performing this procedure.

1. Become superuser on a node in the cluster.
2. Start the `scsetup(1M)` utility.

```
# scsetup
```

3. **To work with private hostnames, type 4 (Private hostnames).**
4. **To change a private hostname, type 1 (Change a private hostname).**
Follow the prompts to change the private hostname. Repeat for each private hostname you want to change.
5. **Verify the new private hostnames.**

```
# scconf -pv | grep "private hostname"
(phys-schost-1) Node private hostname:      phys-schost-1-priv
(phys-schost-3) Node private hostname:      phys-schost-3-priv
(phys-schost-2) Node private hostname:      phys-schost-2-priv
```

Where to Go From Here

To update the `/etc/inet/ntp.conf` file, go to “How to Update Network Time Protocol (NTP)” on page 66.

▼ How to Update Network Time Protocol (NTP)

Perform this task on each node.

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

2. **Edit the `/etc/inet/ntp.conf` file.**

The `scinstall(1M)` command copies a template file, `ntp.cluster`, to `/etc/inet/ntp.conf` as part of standard cluster installation. But if an `ntp.conf` file already exists before Sun Cluster software is installed, that existing file remains unchanged. If cluster packages are installed by using other means, such as direct use of `pkgadd(1M)`, you need to configure NTP.

- a. **Remove all entries for private hostnames that are not used by the cluster.**

If the `ntp.conf` file contains non-existent private hostnames, when a node is rebooted, error messages are generated on the node's attempts to contact those private hostnames.

- b. **If you changed any private hostnames after Sun Cluster software installation, update each file entry with the new private hostname.**

- c. **If necessary, make other modifications to meet your NTP requirements.**

The primary requirement when configuring NTP, or any time synchronization facility, within the cluster is that all cluster nodes be synchronized to the same time. Consider accuracy of time on individual nodes secondary to the

synchronization of time among nodes. You are free to configure NTP as best meets your individual needs, as long as this basic requirement for synchronization is met.

Refer to *Sun Cluster 3.0 Concepts* for further information about cluster time and to the `ntp.cluster` template for guidelines on configuring NTP for a Sun Cluster configuration.

3. Restart the NTP daemon.

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

Where to Go From Here

If you want to use the Sun Management Center product to configure resource groups or monitor the cluster, go to “Installation Requirements for Sun Management Center Software for Sun Cluster Monitoring” on page 68.

Otherwise, to install third-party applications, refer to the documentation supplied with the application software and to *Sun Cluster 3.0 Data Services Installation and Configuration Guide*. To register resource types, set up resource groups, and configure data services, refer to *Sun Cluster 3.0 Data Services Installation and Configuration Guide*.

Installing the Sun Cluster Module for Sun Management Center

The following table lists the tasks to perform to install the Sun Cluster module software for Sun Management Center.

TABLE 2-4 Task Map: Installing the Sun Cluster Module for Sun Management Center

Task	For Instructions, Go To...
Install Sun Management Center server, help server, agent, and console packages.	Sun Management Center documentation "Installation Requirements for Sun Management Center Software for Sun Cluster Monitoring" on page 68
Install Sun Cluster module packages.	"How to Install the Sun Cluster Module for Sun Management Center" on page 69
Start Sun Management Center server, console, and agent processes.	"How to Start Sun Management Center Software" on page 70
Add each cluster node as a Sun Management Center agent host object.	"How to Add a Cluster Node as a Sun Management Center Agent Host Object" on page 71
Load the Sun Cluster module to begin monitoring of the cluster.	"How to Load the Sun Cluster Module" on page 72

Installation Requirements for Sun Management Center Software for Sun Cluster Monitoring

The Sun Cluster module for the Sun Management Center product (formerly Sun Enterprise SyMON) is used to configure resource groups and monitor clusters. Perform the following tasks before installing 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 packages** – You must install the Sun Management Center server, help server, and console packages on non-cluster nodes. If you have an administrative console or other dedicated machine, you can realize improved performance by running the console on the administrative console and the server on a separate machine. You must install the Sun Management Center agent package on each cluster node.

Follow procedures in the Sun Management Center documentation to install the Sun Management Center packages.

- **Simple Network Management Protocol (SNMP) port** – When installing the Sun Management Center product on the agent, 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 you choose for reference later when configuring the cluster for monitoring.

▼ How to Install the Sun Cluster Module for Sun Management Center

Perform this procedure to install the Sun Cluster module console, server, and help server packages.

Note - The Sun Cluster module agent packages (SUNWscsa1 and SUNWscsam) were added to cluster nodes during Sun Cluster software installation.

1. Ensure that the Sun Management Center core packages are installed.

This step includes installing Sun Management Center agent packages on each cluster node. Refer to Sun Management Center documentation for installation instructions.

2. On the administrative console, install the Sun Cluster module console package.

a. **Become superuser.**

b. **If you are installing from the CD-ROM, insert the Sun Cluster module CD-ROM into the CD-ROM drive.**

c. **Change to the `/cdrom_image/SunCluster_3.0/Packages` directory.**

d. **Install the Sun Cluster module console package.**

```
# pkgadd -d . SUNWscscn
```

e. **If you installed from a CD-ROM, eject the CD-ROM.**

3. On the server machine, install the Sun Cluster module server package

SUNWscssv.

Use the same procedure as in Step 2 on page 69.

4. **On the help server machine, install the Sun Cluster module help server package** `SUNWscshl`.

Use the same procedure as in Step 2 on page 69.

5. **Install any Sun Cluster module patches.**

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

Where to Go From Here

To start the Sun Management Center software, go to “How to Start Sun Management Center Software” on page 70.

▼ How to Start Sun Management Center Software

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.**

```
# /opt/SUNWsymon/sbin/es-start -s
```

2. **As superuser, on each Sun Management Center agent machine (cluster node), start the Sun Management Center agent process.**

```
# /opt/SUNWsymon/sbin/es-start -a
```

3. **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.

```
% /opt/SUNWsymon/sbin/es-start -c
```

4. **Type your login name, password, and server hostname and click Login.**

Where to Go From Here

To add cluster nodes as monitored host objects, go to “How to Add a Cluster Node as a Sun Management Center Agent Host Object” on page 71.

▼ How to Add a Cluster Node as a Sun Management Center Agent Host Object

Perform this procedure to create a Sun Management Center agent host object for a cluster node.

Note - 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.

1. **From the Sun Management Center main window, select a domain from the Sun Management Center Administrative Domains pull-down list.**

This domain will contain the Sun Management Center agent host object you are creating. 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 one.

Refer to your Sun Management Center documentation for information about creating Sun Management Center domains.

2. **Select Edit>Create an Object from the pull-down menu.**
3. **Select the Node tab.**
4. **From the Monitor via pull-down list, select Sun Management Center Agent - Host.**
5. **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.

6. **In the Port text field, type the port number you chose during Sun Management Center agent installation.**
7. **Click OK.**

A Sun Management Center agent host object is created in the domain.

Where to Go From Here

To load the Sun Cluster module, go to “How to Load the Sun Cluster Module” on page 72.

▼ How to Load the Sun Cluster Module

Perform this procedure to start cluster monitoring.

- 1. From the Sun Management Center main window, double-click the agent host object for a cluster node.**

The agent host object is shown in two places. You can double-click either one. The Details window of the host object is then displayed.

- 2. Select the icon at the root (top) of the hierarchy to highlight it.**

This icon is labeled with the cluster node name.

- 3. Select Module>Load Module from the pull-down menu.**

The Load Module window lists each available Sun Management Center module and whether it is currently loaded.

- 4. Select Sun Cluster: Not loaded (usually at the bottom of the list) and click OK.**

The Module Loader window shows the current parameter information for the selected module.

- 5. Click OK.**

After a few moments the module is loaded and a Sun Cluster icon is displayed in the Details window.

- 6. In the Details window 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.

Refer to the Sun Cluster module online help for information about using Sun Cluster module features.

- To view online help for a specific Sun Cluster module item, place the cursor over the item, click the right mouse button, and select Help from the pop-up menu.
- To access the home page for the Sun Cluster module online help, place the cursor over the Cluster Info icon, click the right mouse button, and select Help from the pop-up menu.
- To directly access the home page for the Sun Cluster module online help, click the Sun Management Center Help button to launch the help browser, then go to the URL `file:/opt/SUNWsymon/lib/locale/C/help/main.top.html`.

Note - The Help button in the Sun Management Center browser accesses Sun Management Center online help, not topics specific to the Sun Cluster module.

Refer to Sun Management Center online help and your Sun Management Center documentation for information about using the Sun Management Center product.

Where to Go From Here

To install third-party applications, refer to the documentation supplied with the application software and to *Sun Cluster 3.0 Data Services Installation and Configuration Guide*. To register resource types, set up resource groups, and configure data services, refer to *Sun Cluster 3.0 Data Services Installation and Configuration Guide*.

Upgrading Sun Cluster Software

This chapter provides step-by-step procedures for upgrading a two-node Sun Cluster 2.2 configuration to Sun Cluster 3.0 software.

The following step-by-step instructions are in this chapter.

- “How to Shut Down the Cluster” on page 78
- “How to Uninstall VERITAS Volume Manager Software” on page 80
- “How to Upgrade the Solaris Operating Environment” on page 81
- “How to Upgrade Cluster Software Packages” on page 84
- “How to Update the Root User’s Environment” on page 87
- “How to Upgrade Data Service Software Packages” on page 88
- “How to Finish Upgrading Cluster Software” on page 90
- “How to Verify Cluster Membership” on page 93

For overview information about planning your Sun Cluster configuration, see Chapter 1. For a high-level description of the related procedures in this chapter, see “Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Software” on page 75.

Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Software

Perform the following tasks to upgrade your two-node cluster from Sun Cluster 2.2 to Sun Cluster 3.0 software.

TABLE 3-1 Task Map: Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Software

Task	For Instructions, Go To ...
Read upgrade conditions and restrictions, and plan a root disk partitioning scheme to support Sun Cluster 3.0 software.	“Overview of Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Software” on page 77
Take the cluster out of production.	“How to Shut Down the Cluster” on page 78
If your cluster uses VERITAS Volume Manager (VxVM), deport disk groups and remove VxVM software packages.	“How to Uninstall VERITAS Volume Manager Software” on page 80
Upgrade to the Solaris 8 operating environment if necessary, add a new <code>/globaldevices</code> file system, and change file system allocations to support Sun Cluster 3.0 software. If your cluster uses Solstice DiskSuite software, remove mediators and upgrade Solstice DiskSuite software.	“How to Upgrade the Solaris Operating Environment” on page 81
Upgrade to Sun Cluster 3.0 framework software. If your cluster uses Solstice DiskSuite software, recreate mediators.	“How to Upgrade Cluster Software Packages” on page 84
Update the <code>PATH</code> and <code>MANPATH</code> .	“How to Update the Root User’s Environment” on page 87
Upgrade to Sun Cluster 3.0 data services software. If necessary, upgrade third-party applications.	“How to Upgrade Data Service Software Packages” on page 88
Assign a quorum device, finish the cluster software upgrade, and start device groups and data services. If your cluster uses VERITAS Volume Manager (VxVM), reinstall VxVM software packages and import and register disk groups. If your cluster uses Solstice DiskSuite software, restore mediators.	“How to Finish Upgrading Cluster Software” on page 90
Verify that all nodes have joined the cluster.	“How to Verify Cluster Membership” on page 93

Overview of Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Software

This section provides conditions, restrictions, and planning guidelines for upgrading from Sun Cluster 2.2 to Sun Cluster 3.0 software.

Conditions and Restrictions

The following conditions must be met to upgrade from Sun Cluster 2.2 to Sun Cluster 3.0 software.

- The cluster must have exactly two nodes and be a supported configuration for Sun Cluster 3.0 software. The upgrade does not support clusters of three or more nodes.
- Only Ethernet adapters are supported. Transport adapters must have a transmission rate of 100 Mbit/sec or greater.
- All cluster hardware must be stable and working properly.
- All third-party applications must be functioning properly.
- The cluster must be running on, or upgraded to, the Solaris 8 operating environment.
- You must upgrade all Sun Cluster software, framework, and data services at the same time.
- Sun Cluster 3.0 software does not support upgrading directly to Sun Cluster 3.0 software from Solstice HA 1.3, Sun Cluster 2.0, or Sun Cluster 2.1 software.
- Sun Cluster 3.0 software does not support converting from one volume manager product to another during upgrade.
- The upgrade from Sun Cluster 2.2 to Sun Cluster 3.0 software cannot be reversed after the `scinstall(1M)` command has been started on a node, even if the command does not complete successfully. To restart a failed upgrade, you must first reinstall Sun Cluster 2.2 software on the node.

Planning the Upgrade

To support Sun Cluster 3.0 software, you probably need to change your current system disk layout. Consider the following when planning your new partitioning scheme.

- **Global devices namespace** – On each node you must create a file system of at least 100 MBytes and set its mount point as `/globaldevices`. This file system will be converted during upgrade to the appropriate global device namespace. If necessary, you can remove some of the swap space for this purpose, or use an external disk that is not shared with any other node.

- **Mirrored root** – If your root disks are mirrored, you must unmirror them before modifying partitions. The mirror can be used to recover the original configuration if the upgrade procedure fails. Refer to your volume manager documentation for information.
- **Root (/) file system allocation** – If you need to upgrade your configuration to the Solaris 8 operating environment, you probably need to increase the size of your root (/) partition on the root disks of all Sun Cluster nodes.

Refer to “System Disk Partitions” on page 13 for more information about disk space requirements to support Sun Cluster 3.0 software.

▼ How to Shut Down the Cluster

Before upgrading the software, take the cluster out of production.

1. **Have available the CD-ROMs, documentation, and patches for all the software products you are upgrading.**
 - Solaris 8 operating environment
 - Solstice DiskSuite software or VERITAS Volume Manager
 - Sun Cluster 3.0 framework
 - Sun Cluster 3.0 data services
 - Third-party applications

Solstice DiskSuite software and documentation are now part of the Solaris 8 product.

Note - These procedures assume you are installing from CD-ROMs. If you are installing from a network, ensure that the CD-ROM image for each software product is loaded on the network.

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

2. **Notify users that the cluster will be down.**
3. **Become superuser on each node of the cluster.**
4. **Search the `/var/adm/messages` log for unresolved error or warning messages.**
Correct any problems.
5. **Verify that no logical hosts are in the maintenance state.**
 - a. **Become superuser on a node of the cluster.**

- b. Use the `hastat(1M)` command to display the status of the cluster.

```
# hastat
HIGH AVAILABILITY CONFIGURATION AND STATUS
-----
...
LOGICAL HOSTS IN MAINTENANCE STATE
```

If the screen output displays `NONE`, no logical hosts are in the maintenance state. Proceed to Step 6 on page 79.

- c. If a logical host is in the maintenance state, use the `haswitch(1M)` command to perform switchover.

```
# haswitch hostname logicalhostname
```

hostname Specifies the name of the node that is to own the logical host

logicalhostname Specifies the name of the logical host

- d. Run the `hastat` command to verify the switchover completed successfully.

6. Ensure that the size of each logical host administrative file system is at least 10 Mbytes.

```
# df -k /logicalhostname
```

A logical host administrative file system that is not the required minimum size of 10 Mbytes will not be mountable after upgrade to Sun Cluster 3.0. If a logical host administrative file system is smaller than 10 Mbytes, follow your volume manager documentation procedure for growing this file system.

7. **Back up your system.**

Ensure that all users are logged off the system before you back it up.

8. **Stop the Sun Cluster 2.2 software on each node of the cluster.**

```
# scadmin stopnode
```

9. Run the `hastat` command to verify that no nodes are in the cluster.

Where to Go From Here

If your cluster uses VERITAS Volume Manager, go to “How to Uninstall VERITAS Volume Manager Software” on page 80. If your cluster uses Solstice DiskSuite software, to upgrade or prepare the Solaris operating environment to support Sun Cluster 3.0 software, go to “How to Upgrade the Solaris Operating Environment” on page 81.

▼ How to Uninstall VERITAS Volume Manager Software

If your cluster uses VERITAS Volume Manager (VxVM), perform this procedure on each node of the cluster to uninstall the VxVM software. Existing disk groups are retained and automatically reimported after you have upgraded all software.

Note - To upgrade to Sun Cluster 3.0 software, you must remove VxVM software and later reinstall it, regardless of whether you have the latest version of VxVM installed.

1. **Become superuser on a cluster node.**
2. **Deport all VxVM disk groups.**
Refer to your VxVM documentation for procedures.

Note - Ensure that disks containing data to be preserved are not used for other purposes during the upgrade.

3. **Unencapsulate the root disk, if it is encapsulated.**
Refer to your VxVM documentation for procedures.
4. **Shut down VxVM.**
Refer to your VxVM documentation for procedures.
5. **Remove all installed VxVM software packages.**
Refer to your VxVM documentation for procedures.
6. **Remove the VxVM device namespace.**

```
# rm -rf /dev/vx
```

Where to Go From Here

To upgrade or prepare the Solaris operating environment to support Sun Cluster 3.0 software, go to “How to Upgrade the Solaris Operating Environment” on page 81.

▼ How to Upgrade the Solaris Operating Environment

Perform this procedure on each node in the cluster to upgrade or prepare the Solaris operating environment to support Sun Cluster 3.0 software.

1. **Become superuser on the cluster node.**
2. **If your volume manager is Solstice DiskSuite and you are using mediators, unconfigure mediators.**
 - a. **Run the following command to verify that no mediator data problems exist.**

```
# medstat -s setname
```

-s *setname* Specifies the diskset name

If the value in the Status field is `Bad`, repair the affected mediator host by following the procedure “How to Fix Bad Mediator Data” on page 121. See the `medstat(1M)` man page for more information.

- b. **List all mediators.**

Use this information to determine which node, if any, has ownership of the diskset from which you will remove mediators.

```
# metaset -s setname
```

Save this information for when you restore the mediators during the procedure “How to Upgrade Cluster Software Packages” on page 84.

- c. **If no node has ownership, take ownership of the diskset.**

```
# metaset -s setname -t
```

-t Takes ownership of the diskset

- d. **Unconfigure all mediators.**

```
# metaset -s setname -d -m mediator_host_list
```

-s *setname* Specifies the diskset name

-d Deletes from the diskset

-m *mediator_host_list* Specifies the name of the node to remove as a mediator host for the diskset

Refer to the `mediator(7)` man page for further information about mediator-specific options to the `metaset` command.

e. Remove the mediator software.

```
# pkgrm SUNWmdm
```

3. Does your configuration currently run Solaris 8 software?

- If no, go to Step 4 on page 82.
- If yes,
 - a. **Create a file system of at least 100 Mbytes and set its mount point as /globaldevices.**

Note - The /globaldevices file system is necessary for Sun Cluster 3.0 software installation to succeed.

b. Reallocate space in other partitions as needed to support Sun Cluster 3.0 software.

Refer to “System Disk Partitions” on page 13 for guidelines.

c. Go to Step 6 on page 83.

4. Determine which procedure to use to upgrade to Solaris 8 software.

Volume Manager	Procedure to Use	For Instructions, Go To ...
Solstice DiskSuite	Upgrading both Solaris and Solstice DiskSuite software	Solstice DiskSuite installation documentation
VxVM	Performing a standard Solaris software installation	Solaris 8 installation documentation

5. Upgrade to Solaris 8 software, following the procedure you selected in Step 4 on page 82.

During installation, make the following changes to the root disk partitioning scheme.

- Create a file system of at least 100 Mbytes and set its mount point as `/globaldevices`. The `/globaldevices` file system is necessary for Sun Cluster 3.0 software installation to succeed.
- Reallocate space in other partitions as needed to support Sun Cluster 3.0 software.

Refer to “System Disk Partitions” on page 13 for partitioning guidelines.

Note - The Solaris interface groups feature is disabled by default during Solaris software installation. Interface groups are not supported in a Sun Cluster configuration and should not be enabled. Refer to the `ifconfig(1M)` man page for more information about Solaris interface groups.

6. Install any Solaris software patches.

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

7. Install any hardware-related patches.

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

8. For Solstice DiskSuite software, install any Solstice DiskSuite software patches.

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

Example—Unconfiguring Mediators

The following example shows the mediator host `phys-schost-1` unconfigured from the Solstice DiskSuite diskset `schost-1` before the upgrade to Solaris 8 software.

```
(Check mediator status:)
# medstat -s schost-1

(List all mediators:)
# metaset -s schost-1

(Unconfigure the mediator:)
# metaset -s schost-1 -d -m phys-schost-1

(Remove mediator software:)
# pkgrm SUNWmdm
```

(continued)

```
(Begin software upgrade)
```

Where to Go From Here

To upgrade to Sun Cluster 3.0 software, go to “How to Upgrade Cluster Software Packages” on page 84.

▼ How to Upgrade Cluster Software Packages

Perform this procedure on each node. You can perform this procedure on both nodes simultaneously if you have two copies of the Sun Cluster 3.0 framework CD-ROM.

Note - The `scinstall(1M)` upgrade command is divided into a two-step process: the `-u begin` option and the `-u finish` option. This procedure runs the `begin` option. The `finish` option is run in “How to Finish Upgrading Cluster Software” on page 90.

1. **Become superuser on a cluster node.**
2. **If your volume manager is Solstice DiskSuite, install the latest Solstice DiskSuite mediator package (SUNWmdm) on each node.**
 - a. **If you are installing from the CD-ROM, insert the Sun Cluster 3.0 CD-ROM into the CD-ROM drive on a node.**
If the volume daemon `vold(1M)` is running and configured to manage CD-ROM devices, it automatically mounts the CD-ROM on the `/cdrom/suncluster_3_0` directory.

- b. **Change to the `/cdrom_image/suncluster_3_0/SunCluster_3.0/Packages` directory.**

```
# cd /cdrom_image/suncluster_3_0/SunCluster_3.0/Packages
```

- c. **Add the SUNWmdm package.**

```
# pkgadd -d . SUNWmdm
```

- d. **Reboot the node.**

```
# shutdown -g 0 -y -i 6
```

e. Repeat on the other node.

3. Reconfigure mediators.

a. Determine which node has ownership of the diskset to which you will add the mediator hosts.

```
# metaset -s setname
```

-s *setname* Specifies the diskset name

b. If no node has ownership, take ownership of the diskset.

```
# metaset -s setname -t
```

-t Takes ownership of the diskset

c. Recreate the mediators.

```
# metaset -s setname -a -m mediator_host_list
```

-a Adds to the diskset

-m *mediator_host_list* Specifies the names of the nodes to add as mediator hosts for the diskset

d. Repeat for each diskset.

4. On each node, begin upgrade to Sun Cluster 3.0 software.

a. On one node, change to the `/cdrom_image/suncluster_3_0/SunCluster_3.0/Tools` directory.

```
# cd /cdrom_image/suncluster_3_0/SunCluster_3.0/Tools
```

b. Upgrade the cluster software framework.

To Upgrade The ...	Use This Command ...
First node	<code>./scinstall -u begin -F</code>
Second node	<code>./scinstall -u begin -N <i>clusternode1</i></code>

- F Specifies that this is the first node in the cluster that will be upgraded
- N *clusternode1* Specifies the name of the first node in the cluster that will be upgraded, *not* the name of the second node to be upgraded

Refer to the `scinstall(1M)` man page for more information.

c. Reboot the node.

```
# shutdown -g 0 -y -i 6
```

When the first node reboots into cluster mode, it establishes the cluster. The second node waits if necessary for the cluster to be established before completing its own processes and joining the cluster.

d. Repeat on the other cluster node.

5. On each node, install any Sun Cluster patches.

Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

Example—Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Software – Begin Process

The following example shows the beginning process of upgrading a two-node cluster from Sun Cluster 2.2 to Sun Cluster 3.0 software. The cluster node names are `phys-schost-1`, the sponsor node, and `phys-schost-2`, which joins the cluster that `phys-schost-1` established. The volume manager is Solstice DiskSuite and both nodes are used as mediator hosts for the diskset `schost-1`.

```
(Install the latest Solstice DiskSuite mediator package on each node:)
# cd /cdrom/suncluster_3_0/SunCluster_3.0/Packages
# pkgadd -d . SUNWmdm
```

(continued)

```

(Restore the mediator:)
# metaset -s schost-1 -t
# metaset -s schost-1 -a -m phys-schost-1 phys-schost-2

(Begin upgrade on the first node:)
phys-schost-1# cd /cdrom/suncluster_3_0/SunCluster_3.0/Tools
phys-schost-1# ./scinstall -u begin -F

(Begin upgrade on the second node:)
phys-schost-2# cd /cdrom/suncluster_3_0/SunCluster_3.0/Tools
phys-schost-2# ./scinstall -u begin -N phys-schost-1

(Reboot each node:)
# shutdown -g 0 -y -i 6

```

Where to Go From Here

To update the directory paths, go to “How to Update the Root User’s Environment” on page 87.

▼ How to Update the Root User’s Environment

Perform the following tasks on each node of the cluster.

1. **Set the PATH to include /usr/sbin and /usr/cluster/bin.**
For VERITAS Volume Manager, also set your PATH to include /etc/vx/bin. If you installed the VRTSvmsa package, also add /opt/VRTSvmsa/bin to your PATH.
2. **Set the MANPATH to include /usr/cluster/man. Also include the volume manager-specific paths.**
 - For Solstice DiskSuite software, set your MANPATH to include /usr/share/man.
 - For VERITAS Volume Manager, set your MANPATH to include /opt/VRTSvxvm/man. If you installed the VRTSvmsa package, also add /opt/VRTSvmsa/man to your MANPATH.
3. **For ease of administration, set the same root password on each node.**
4. **Start a new shell to activate the environment changes.**

Where to Go From Here

To upgrade to Sun Cluster 3.0 data service software, go to “How to Upgrade Data Service Software Packages” on page 88.

▼ How to Upgrade Data Service Software Packages

Perform this procedure on each cluster node.

1. **Become superuser on a node of the cluster.**
2. **Upgrade applications and apply application patches as needed.**
Refer to your application documentation for installation instructions.

Note - If the applications are stored on shared disks, you must master the relevant disk groups and manually mount the relevant file systems before you upgrade the application.

3. **Add data services.**
 - a. **Insert the Sun Cluster 3.0 data services CD-ROM into the CD-ROM drive on the node.**
 - b. **Enter the `scinstall(1M)` utility.**

```
# scinstall
```

Follow these guidelines while using the interactive `scinstall` utility.

- Interactive `scinstall` enables you to type ahead. Therefore, do not press Return more than once if the next menu screen does not appear immediately.
 - Unless otherwise noted, pressing Control-D returns you either to the start of a series of related questions or to the Main Menu.
- c. **To add data services, type 4 (Add support for a new data service to this cluster node).**
Follow the prompts to add data services.
 - d. **Eject the CD-ROM.**

4. **Install any Sun Cluster data service patches.**
Refer to *Sun Cluster 3.0 Release Notes* for the location of patches and installation instructions.

5. Repeat Step 1 on page 88 through Step 4 on page 88 on the other node of the cluster.

6. Shut down the second node to be upgraded to Sun Cluster 3.0 software.

Leave the second node shut down until after the first, or sponsor, node has been rebooted.

```
phys-schost-2# shutdown -g 0 -y -i 0
```

7. Reboot the first, sponsor node of the cluster.

Ensure that the second node has been shut down before rebooting the first node. Otherwise, rebooting the first node while the second node is still up will panic the second node because quorum votes have not yet been assigned.

```
phys-schost-1# shutdown -g 0 -y -i 6
```

8. After the first node has completed booting, boot the second node.

```
phys-schost-2# boot
```

9. After both nodes have rebooted, verify from either node that both nodes are cluster members.

```
# scstat -n
Node
Node Name:          phys-schost-1
Status:             Online

Node Name:          phys-schost-2
Status:             Online
```

See the `scstat(1M)` man page for more information about displaying cluster status.

Where to Go From Here

To assign a quorum device and finish the upgrade, go to “How to Finish Upgrading Cluster Software” on page 90.

▼ How to Finish Upgrading Cluster Software

This procedure finishes the `scinstall(1M)` upgrade process begun in “How to Upgrade Cluster Software Packages” on page 84. Perform these steps on each node of the cluster.

1. Become superuser on each node of the cluster.

2. Choose a shared disk to be the quorum device.

You can use any disk shared by both nodes as a quorum device. From either node, use the `scdidadm(1M)` command to determine the shared disk's device ID (DID) name. You specify this device name in Step 5 on page 90, in the `-q globaldev=devicename` option to `scinstall`.

```
# scdidadm -L
```

3. If your volume manager is VxVM, reinstall the VxVM software on each node of the cluster.

Note - Whenever you must reboot, you must shut down the second node of the cluster before rebooting the first, or sponsor, node. After rebooting the first node, then bring the second node back up. Otherwise, rebooting the first node while the second node is still up will panic the second node because quorum votes have not yet been assigned.

a. Install VxVM software, including any patches.

Follow the procedures in “How to Install VERITAS Volume Manager Software” on page 52.

b. Configure VxVM.

Follow the procedures listed in “Configuring VxVM for Sun Cluster Configurations” on page 125.

4. Insert the Sun Cluster 3.0 data services CD-ROM into the CD-ROM drive on the node.

This step assumes that the volume daemon `vold(1M)` is running and configured to manage CD-ROM devices.

5. Finish the cluster software upgrade on that node.

```
# scinstall -u finish -q globaldev=devicename \  
-d /cdrom_image/scdataservices_3_0 -s svc[,svc]
```

- | | |
|--|--|
| -q <code>globaldev=devicename</code> | Specifies the name of the quorum device |
| -d <code>/cdrom_image/scdataservices_3_0</code> | Specifies the directory location of the CD-ROM image |
| -s <code>srvc</code> | Specifies the name of the data service to configure |

Note - An error message similar to the following might be generated. You can safely ignore it.

```
** Installing Sun Cluster - Highly Available NFS Server **
Skipping "SUNWscnfs" - already installed
```

6. Eject the CD-ROM.

7. Repeat Step 4 on page 90 through Step 6 on page 91 on the other node.

When completed on both nodes, the cluster is removed from install mode and all quorum votes are assigned.

8. If your volume manager is Solstice DiskSuite, from either node bring pre-existing disk device groups online.

```
# scswitch -z -D disk-device-group -h node
```

- | | |
|--|--|
| -z | Performs the switch |
| -D <code>disk-device-group</code> | Specifies the name of the disk device group, which for Solstice DiskSuite software is the same as the diskset name |
| -h <code>node</code> | Specifies the name of the cluster node that serves as the primary of the disk device group |

9. From either node, bring pre-existing data service resource groups online.

At this point, Sun Cluster 2.2 logical hosts are converted to Sun Cluster 3.0 resource groups, and the names of logical hosts are appended with the suffix `-lh`. For example, a logical host named `lhost-1` is upgraded to a resource group named `lhost-1-lh`. Use these converted resource group names in the following command.

```
# scswitch -z -g resource-group -h node
```

`-g resource-group` Specifies the name of the resource group to bring online

You can use the `scrgadm -p` command to display a list of all resource types and resource groups in the cluster. The `scrgadm -pv` command displays this list with more detail.

10. If you are using the Sun Management Center product to monitor your Sun Cluster configuration, install the Sun Cluster module for the Sun Management Center product.

a. Ensure that you are using the most recent version of Sun Management Center software (formerly Sun Enterprise SyMON).

Refer to the Sun Management Center documentation for installation or upgrade procedures.

b. Follow guidelines and procedures in “Installation Requirements for Sun Management Center Software for Sun Cluster Monitoring” on page 68 to install the Sun Cluster module packages.

Example—Upgrading From Sun Cluster 2.2 to Sun Cluster 3.0 Software – Finish Process

The following example shows the finish process of upgrading a two-node cluster upgraded from Sun Cluster 2.2 to Sun Cluster 3.0 software. The cluster node names are `phys-schost-1` and `phys-schost-2`, the device group names are `dg-schost-1` and `dg-schost-2`, and the data service resource group names are `lh-schost-1` and `lh-schost-2`.

```
(Determine the DID of the shared quorum device:)
phys-schost-1# scdidadm -L

(Finish upgrade on each node:)
phys-schost-1# scinstall -u finish -q globaldev=d1 \
-d /cdrom/suncluster_3_0 -s nfs
phys-schost-2# scinstall -u finish -q globaldev=d1 \
-d /cdrom/suncluster_3_0 -s nfs
```

(continued)

```
(Bring device groups and data service resource groups on each node online:)
phys-schost-1# scswitch -z -D dg-schost-1 -h phys-schost-1
phys-schost-1# scswitch -z -g lh-schost-1 -h phys-schost-1
phys-schost-1# scswitch -z -D dg-schost-2 -h phys-schost-2
phys-schost-1# scswitch -z -g lh-schost-2 -h phys-schost-2
```

Where to Go From Here

To verify that all nodes have joined the cluster, go to “How to Verify Cluster Membership” on page 93.

▼ How to Verify Cluster Membership

Perform this procedure to verify that all nodes have joined the cluster.

1. Become superuser on any node in the cluster.

2. Display cluster status.

Verify that cluster nodes are online and that the quorum device, device groups, and data services resource groups are configured and online.

```
# scstat
```

See the `scstat(1M)` man page for more information about displaying cluster status.

3. On each node, display a list of all devices the system checks to verify their connectivity to the cluster nodes.

The output on each node should be the same.

```
# scdidadm -L
```

The cluster upgrade is complete. You can now return the cluster to production.

Configuring Solstice DiskSuite Software

Configure your local and multihost disks for Solstice DiskSuite software by using the procedures in this appendix, along with the information in Chapter 1. Refer to your Solstice DiskSuite documentation for additional details.

The following procedures are in this appendix.

- “How to Set the Number of Metadevice Names and Disksets” on page 97
- “How to Create Metadevice State Database Replicas” on page 98
- “How to Mirror the Root (/) File System” on page 99
- “How to Mirror the Global Namespace” on page 102
- “How to Mirror File Systems That Cannot Be Unmounted” on page 105
- “How to Mirror User-Defined File Systems” on page 109
- “How to Create a Diskset” on page 111
- “How to Add Drives to a Diskset” on page 113
- “How to Repartition Drives in a Diskset” on page 115
- “How to Create an `md.tab` File” on page 116
- “How to Activate Metadevices” on page 118
- “How to Add Mediator Hosts” on page 120
- “How to Check the Status of Mediator Data” on page 121
- “How to Fix Bad Mediator Data” on page 121

Configuring Solstice DiskSuite for Sun Cluster Configurations

The following table lists the tasks to perform to configure Solstice DiskSuite software for Sun Cluster configurations.

TABLE A-1 Task Map: Configuring Solstice DiskSuite for Sun Cluster Configurations

Task	For Instructions, Go To ...
Plan the layout of your Solstice DiskSuite configuration.	“Planning Volume Management” on page 23 “Solstice DiskSuite Configuration Example” on page 122
Calculate the number of metadvice names and disksets needed for your configuration, and modify the <code>/kernel/drv/md.conf</code> file.	“How to Set the Number of Metadvice Names and Disksets” on page 97
Create metadvice state database replicas on the local disks.	“How to Create Metadvice State Database Replicas” on page 98
<i>Optional.</i> Mirror file systems on the root disk.	“Mirroring the Root Disk” on page 99
Create disksets using the <code>metaset</code> command.	“How to Create a Diskset” on page 111
Add disk drives to the disksets.	“How to Add Drives to a Diskset” on page 113
Repartition drives in a diskset to allocate space to slices 1 through 6.	“How to Repartition Drives in a Diskset” on page 115
List device ID pseudo-driver mappings and define metadevices in the <code>/etc/lvm/md.tab</code> files.	“How to Create an <code>md.tab</code> File” on page 116
Initialize the <code>md.tab</code> files.	“How to Activate Metadevices” on page 118

TABLE A-1 Task Map: Configuring Solstice DiskSuite for Sun Cluster Configurations (continued)

Task	For Instructions, Go To ...
For dual-string configurations, configure mediator hosts, check the status of mediator data, and, if necessary, fix bad mediator data.	“How to Add Mediator Hosts” on page 120 “How to Check the Status of Mediator Data” on page 121 “How to Fix Bad Mediator Data” on page 121
Create and mount cluster file systems.	“How to Add Cluster File Systems” on page 59

▼ How to Set the Number of Metadevice Names and Disksets

This procedure describes how to calculate the quantity of metadevice names needed for your configuration, and how to modify the `/kernel/drv/md.conf` file.

Tip - The default number of metadevice names is 128. Many configurations need more than the default. Increasing this number before implementing a configuration saves administration time later.

1. Calculate the quantity of metadevice names needed by determining the largest of the metadevice names to be used in each diskset.

This requirement is based on the metadevice name *value* rather than on the *actual quantity*. For example, if your metadevice names range from `d950` to `d1000`, Solstice DiskSuite software requires 1000 names, not 50.

The cluster can have a maximum of 8192 metadevice names per diskset.

2. Calculate the total expected number of disksets in the cluster, then add one for private disk management.

The cluster can have a maximum of 32 disksets.

3. 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 disksets served by each node. Failure to follow this guideline can result in serious Solstice DiskSuite errors and possible loss of data.

- a. If the calculated quantity of metadevice names exceeds 128, set the `nmd` field to the largest metadevice name value used in a diskset.
- b. Set the `md_nsets` field to the quantity calculated in Step 2 on page 97.

Changes to the `/kernel/drv/md.conf` file become operative after a reconfiguration reboot is performed.

Where to Go From Here

To create local replicas, go to “How to Create Metadevice State Database Replicas” on page 98.

▼ How to Create Metadevice State Database Replicas

Perform this procedure on each node in the cluster.

1. **Become superuser on the cluster node.**
2. **Create replicas on one or more local disks for each cluster node by using the `metadb` command.**
Refer to the `metadb(1M)` man page and your Solstice DiskSuite documentation for details.

Tip - To provide protection of metadevice state data, which is necessary to run Solstice DiskSuite software, create at least three replicas for each node. Also, placing replicas on more than one disk provides protection if one of the disks fails.

3. **Verify the replicas.**

```
# metadb
```

Example—Creating Metadevice State Database Replicas

The following example shows three metadevice state database replicas, each created on a different disk.

```
# metadb -af c0t0d0s7 c0t1d0s7 c1t0d0s7
# metadb
flags          first blk      block count
```

(continued)

a	u	16	1034	/dev/dsk/c0t0d0s7
a	u	1050	1034	/dev/dsk/c0t1d0s7
a	u	2084	1034	/dev/dsk/c1t0d0s7

Where to Go From Here

If you want to mirror file systems on the root disk, go to “Mirroring the Root Disk” on page 99. Otherwise, to create Solstice DiskSuite disksets, go to “How to Create a Diskset” on page 111.

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 99
- “How to Mirror the Global Namespace” on page 102
- “How to Mirror File Systems That Cannot Be Unmounted” on page 105
- “How to Mirror User-Defined File Systems” on page 109

Note - Some of the steps in these mirroring procedures can cause an error message similar to the following, which is harmless and can be ignored.

```
metainit: dg-schost-1: d1s0: not a metadvice
```



Caution - For local disk mirroring, do not use /dev/global as the path when specifying the disk name. Specifying this path for anything other than cluster file systems prevents the system from booting.

▼ How to Mirror the Root (/) File System

Use this procedure to mirror the root (/) file system.

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

2. Use the `metainit(1M)` command to put the root slice in a single-slice (one-way) concatenation.

```
# metainit -f submirror1 1 1 root-disk-slice
```

3. Create a second concatenation.

```
# metainit -f submirror2 1 1 submirror-disk-slice
```

4. Create a one-way mirror with one submirror.

Note - The metadvice name for the mirror *must* be unique throughout the cluster.

```
# metainit mirror -m submirror1
```

5. Run the `metaroot(1M)` command.

This command edits the `/etc/vfstab` and `/etc/system` files so the system can be booted with the root (`/`) file system on a metadvice.

```
# metaroot mirror
```

6. Run the `lockfs(1M)` command.

This command flushes all transactions out of the log and writes the transactions to the master file system on all mounted UFS file systems.

```
# lockfs -fa
```

7. Evacuate any resource groups or device groups from the node.

```
# scswitch -s -h node
```

<code>-s</code>	Evacuates all resource groups and device groups
<code>-h <i>node</i></code>	Specifies the name of the node from which to evacuate resource or device groups

8. Reboot the node.

```
# shutdown -g 0 -y -i 6
```

9. Use the `metattach(1M)` command to attach the second submirror to the mirror.

```
# metattach mirror submirror2
```

10. If the disk used to mirror the root disk is physically connected to more than one node (multiported), enable the `localonly` property of the raw disk device group for the disk used to mirror the root disk.

You must enable the `localonly` property to prevent unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

a. If necessary, use the `scdidadm -L` command to display the full device ID (DID) pseudo-driver name of the raw disk device group.

In the following example, the raw disk device group name `dsk/d2` is part of the third column of output, which is the full DID pseudo-driver name.

```
# scdidadm -L
...
1          phys-schost-3:/dev/rdisk/ctl1d0    /dev/did/rdisk/d2
# sccnf -c -D name=dsk/d2,localonly=true
```

For more information about the `localonly` property, refer to the `scconf_dg_rawdisk(1M)` man page.

b. Use the `scconf(1M)` command to enable the `localonly` property.

```
# scconf -c -D name=rawdisk_groupname,localonly=true
```

`-D name=rawdisk_groupname` Specifies the name of the raw disk device group

11. Record the alternate boot path for possible future use.

```
# ls -l /dev/rdisk/root-disk-slice
```

12. Repeat Step 1 on page 99 through Step 11 on page 101 on each remaining node of the cluster.

Ensure that each metadvice name for a mirror is unique throughout the cluster.

Example—Mirroring the Root (/) File System

The following example shows creation of mirror d0 on the node `phys-schost-1`, which consists of submirror d10 on partition `c0t0d0s0` and submirror d20 on partition `c2t2d0s0`. Disk `c2t2d0` is a multiported disk, so the `localonly` property is enabled.

```
(Create the mirror:)
# metainit -f d10 1 1 c0t0d0s0
d11: Concat/Stripe is setup
# metainit -f d20 1 1 c2t2d0s0
d12: Concat/Stripe is setup
# metainit d0 -m d10
d10: Mirror is setup
# metaroot d0
# lockfs -fa

(Reboot the node:)
# scswitch -S -h phys-schost-1
# shutdown -g 0 -y -i 6

(Attach the second submirror:)
# metattach d0 d20
d0: Submirror d20 is attached

(Enable the localonly property of the mirrored disk's raw disk device group:)
# sconfig -c -D name=dsk/d2,localonly=true

(Record the alternate boot path:)
# ls -l /dev/rdisk/c2t2d0s0
lrwxrwxrwx 1 root      root           57 Apr 25 20:11 /dev/rdisk/c2t2d0s0 ->
../../devices/node@1/pci@1f,0/pci@1/scsi@3,1/disk@2,0:a,raw
```

Where to Go From Here

To mirror the global namespace, `/global/.devices/node@nodeid`, go to “How to Mirror the Global Namespace” on page 102.

▼ How to Mirror the Global Namespace

Use this procedure to mirror the global namespace, `/global/.devices/node@nodeid`.

1. Become superuser on a node of the cluster.
2. Put the global namespace slice in a single slice (one-way) concatenation.

```
# metainit -f submirror1 1 1 disk-slice
```

3. Create a second concatenation.

```
# metainit -f submirror2 1 1 submirror-disk-slice
```

4. Create a one-way mirror with one submirror.

Note - The metadvice name for the mirror *must* be unique throughout the cluster.

```
# metainit mirror -m submirror1
```

5. Attach the second submirror to the mirror.

This attachment starts a sync of the submirrors.

```
# 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.

```
# 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 on page 102 through Step 6 on page 103 on each remaining node of the cluster.

Ensure that each metadvice name for a mirror is unique throughout the cluster.

8. Wait for the sync of the mirrors, started in Step 5 on page 103, to complete.

Use the `metastat(1M)` command to view mirror status.

```
# metastat mirror
```

9. If the disk used to mirror the global namespace is physically connected to more than one node (multiported), enable the `localonly` property of the raw disk device group for the disk used to mirror the global namespace.

You must enable the `localonly` property to prevent unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

- a. If necessary, use the `scdidadm -L` command to display the full device ID (DID) pseudo-driver name of the raw disk device group.

In the following example, the raw disk device group name `dsk/d2` is part of the third column of output, which is the full DID pseudo-driver name.

```
# scdidadm -L
...
1          phys-schost-3:/dev/rdisk/clt1d0      /dev/did/rdsk/d2
# scconf -c -D name=dsk/d2,localonly=true
```

For more information about the `localonly` property, refer to the `scconf_dg_rawdisk(1M)` man page.

- b. Use the `scconf(1M)` command to enable the `localonly` property.

```
# scconf -c -D name=rawdisk_groupname,localonly=true
```

`-D name=rawdisk_groupname` Specifies the name of the raw disk device group

Example—Mirroring the Global 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`. Disk `c2t2d0` is a multiported disk, so the `localonly` property is enabled.

```
(Create the mirror:)
# metainit -f d111 1 1 c0t0d0s3
d111: Concat/Stripe is setup
# metainit -f d121 1 1 c2t2d0s3
d121: Concat/Stripe is setup
# metainit d101 -m d111
d101: Mirror is setup
# metattach d101 d121
d101: Submirror d121 is attached
```

(continued)

```

(Edit the /etc/vfstab file:)
# 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/rdisk/d101 /global/.devices/node@1 ufs 2 no global

(View the sync status:)
# metastat d101
d101: Mirror
    Submirror 0: d111
        State: Okay
    Submirror 1: d121
        State: Resyncing
    Resync in progress: 15 % done
...

(Identify the DID name of the mirrored disk's raw disk device group:)
# sctidadm -L
...
1      phys-schost-3:/dev/rdsk/c2t2d0      /dev/did/rdsk/d2

(Enable the localonly property of the mirrored disk's raw disk device group:)
# sccnf -c -D name=dsk/d2,localonly=true

```

Where to Go From Here

To mirror file systems that cannot be unmounted during normal system usage, such as `/usr`, `/opt`, or `swap`, go to “How to Mirror File Systems That Cannot Be Unmounted” on page 105. To mirror user-defined file systems, go to “How to Mirror User-Defined File Systems” on page 109.

Otherwise, to create a diskset, go to “How to Create a Diskset” on page 111.

▼ How to Mirror File Systems That Cannot Be Unmounted

Use this procedure to mirror file systems that cannot be unmounted during normal system usage, such as `/usr`, `/opt`, or `swap`.

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

- Put the slice on which an unmountable file system resides in a single slice (one-way) concatenation.

```
# metainit -f submirror1 1 1 disk-slice
```

- Create a second concatenation.

```
# metainit -f submirror2 1 1 submirror-disk-slice
```

- Create a one-way mirror with one submirror.

Note - The metadvice name for the mirror does *not* need to be unique throughout the cluster.

```
# metainit mirror -m submirror1
```

- Repeat Step 1 on page 105 through Step 4 on page 106 for each unmountable file system to be mirrored.
- 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.

```
# 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/rdisk/mirror /file-system ufs    2      no     global
```

- Evacuate any resource groups or device groups from the node.

```
# scswitch -s -h node
```

-S Evacuates all resource groups and device groups

-h *node* Specifies the name of the node from which to evacuate resource or device groups

8. Reboot the node.

```
# shutdown -g 0 -y -i 6
```

9. Attach the second submirror to each mirror.

This attachment starts a sync of the submirrors.

```
# metattach mirror submirror2
```

10. Wait for the sync of the mirrors, started in Step 9 on page 107, to complete.

Use the `metastat(1M)` command to view mirror status.

```
# metastat mirror
```

11. If the disk used to mirror the unmountable file system is physically connected to more than one node (multiported), enable the `localonly` property of the raw disk device group for the disk used to mirror the unmountable file system.

You must enable the `localonly` property to prevent unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

a. If necessary, use the `sddidadm -L` command to display the full device ID (DID) pseudo-driver name of the raw disk device group.

In the following example, the raw disk device group name `dsk/d2` is part of the third column of output, which is the full DID pseudo-driver name.

```
# sddidadm -L
...
1          phys-schost-3:/dev/rdisk/c1t1d0    /dev/did/rdsk/d2
# sccnf -c -D name=dsk/d2,localonly=true
```

For more information about the `localonly` property, refer to the `sccnf_dg_rawdisk(1M)` man page.

b. Use the `sccnf(1M)` command to enable the `localonly` property.

```
# sccnf -c -D name=rawdisk_groupname,localonly=true
```

`-D name=rawdisk_groupname` Specifies the name of the raw disk device group

Example—Mirroring Unmountable File Systems

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`. Disk `c2t2d0` is a multiported disk, so the `localonly` property is enabled.

```
(Create the mirror:)  
# metainit -f d11 1 1 c0t0d0s1  
d11: Concat/Stripe is setup  
# metainit -f d21 1 1 c2t2d0s1  
d21: Concat/Stripe is setup  
# metainit d1 -m d11  
d1: Mirror is setup  
  
(Edit the /etc/vfstab file:)  
# 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/rdisk/d1 /usr           ufs     2      no     global  
  
(Reboot the node:)  
# scswitch -S -h phys-schost-1  
# shutdown -g 0 -y -i 6  
  
(Attach the second submirror:)  
# metattach d1 d21  
d1: Submirror d21 is attached  
  
(View the sync status:)  
# metastat d1  
d1: Mirror  
    Submirror 0: d11  
    State: Okay  
    Submirror 1: d21  
    State: Resyncing  
    Resync in progress: 15 % done  
...  
  
(Identify the DID name of the mirrored disk's raw disk device group:)  
# scdidadm -L  
...  
1          phys-schost-3:/dev/rdisk/c2t2d0    /dev/did/rdisk/d2  
  
(Enable the localonly property of the mirrored disk's raw disk device group:)  
# scconf -c -D name=dsk/d2,localonly=true
```

Where to Go From Here

To mirror user-defined file systems, go to “How to Mirror User-Defined File Systems” on page 109. Otherwise, to create a diskset, go to “How to Create a Diskset” on page 111.

▼ How to Mirror User-Defined File Systems

Use this procedure to mirror user-defined file systems. In this procedure, nodes do not need to be rebooted.

1. **Become superuser on a node of the cluster.**
2. **Put the slice on which a user-defined file system resides in a single slice (one-way) concatenation.**

```
# metainit -f submirror1 1 1 disk-slice
```

3. **Create a second concatenation.**

```
# metainit -f submirror2 1 1 submirror-disk-slice
```

4. **Create a one-way mirror with one submirror.**

Note - The metadvice name for the mirror does *not* need to be unique throughout the cluster.

```
# metainit mirror -m submirror1
```

5. **Repeat Step 1 on page 109 through Step 4 on page 109 for each user-defined file system to be mirrored.**
6. **On each node, edit the `/etc/vfstab` file entry for each user-defined file system you mirrored.**

Replace the names in the device to mount and device to fsck columns with the mirror name.

```
# 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/rdisk/mirror /file-system ufs      2        no       global
```

7. **Attach the second submirror to the mirror.**

This attachment starts a sync of the submirrors.

```
# metattach mirror submirror2
```

8. Wait for the sync of the mirrors, started in Step 7 on page 109, to complete.

Use the `metastat(1M)` command to view mirror status.

```
# metastat mirror
```

9. If the disk used to mirror the user-defined file system is physically connected to more than one node (multiported), enable the `localonly` property of the raw disk device group for the disk used to mirror the user-defined file system.

You must enable the `localonly` property to prevent unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

a. If necessary, use the `scdidadm -L` command to display the full device ID (DID) pseudo-driver name of the raw disk device group.

In the following example, the raw disk device group name `dsk/d4` is part of the third column of output, which is the full DID pseudo-driver name.

```
# scdidadm -L
...
1          phys-schost-3:/dev/rdisk/c1t1d0    /dev/did/rdsk/d2
# scconf -c -D name=dsk/d2,localonly=true
```

For more information about the `localonly` property, refer to the `scconf_dg_rawdisk(1M)` man page.

b. Use the `scconf(1M)` command to enable the `localonly` property.

```
# scconf -c -D name=rawdisk_groupname,localonly=true
```

`-D name=rawdisk_groupname` Specifies the name of the raw disk device group

Example—Mirroring User-Defined File Systems

The following example shows creation of mirror `d4` to mirror `/home`, 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 `/home` is

updated to use the mirror name d4. Disk c2t2d0 is a multiported disk, so the localonly property is enabled.

```
(Create the mirror:)
# metainit -f d14 1 1 c0t0d0s4
d14: Concat/Stripe is setup
# metainit -f d24 1 1 c2t2d0s4
d24: Concat/Stripe is setup
# metainit d4 -m d14
d4: Mirror is setup

(Edit the /etc/vfstab file:)
# 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/rdisk/d4 /home      ufs      2        no       global

(Attach the second submirror:)
# metattach d4 d24
d4: Submirror d24 is attached

(View the sync status:)
# metastat d4
d4: Mirror
    Submirror 0: d14
        State: Okay
    Submirror 1: d24
        State: Resyncing
    Resync in progress: 15 % done
...

(Identify the DID name of the mirrored disk's raw disk device group:)
# scdidadm -L
...
1      phys-schost-3:/dev/rdsk/c2t2d0      /dev/did/rdsk/d2

(Enable the localonly property of the mirrored disk's raw disk device group:)
# scconf -c -D name=dsk/d2,localonly=true
```

Where to Go From Here

To create a diskset, go to “How to Create a Diskset” on page 111.

▼ How to Create a Diskset

Perform this procedure for each diskset in the cluster.

- If configured with exactly two disk strings, the diskset must connect to exactly two nodes and use exactly two mediator hosts, which must be the same two hosts used for the diskset. Refer to “Mediators Overview” on page 119 for details on setting up mediators.
 - If configured with more than two disk strings, ensure that for any two disk strings S1 and S2, the sum of the number of disks on those strings exceeds the number of disks on the third string S3. Stated as a formula, the requirement is that $\text{count}(S1) + \text{count}(S2) > \text{count}(S3)$.
1. **If the `.rhosts` file is not used to list the cluster node names, ensure that `root` is a member of group 14.**

```
# vi /etc/group
...
sysadmin::14:root
...
```

2. **Make sure the local metadvice state database replicas exist.**
If necessary, refer to the procedure “How to Create Metadvice State Database Replicas” on page 98.
3. **Become superuser on the cluster node that will master the diskset.**
4. **Create the diskset.**
This command also registers the diskset as a Sun Cluster disk device group.

```
# metaset -s setname -a -h nodename1 nodename2
```

- s *setname*** Specifies the diskset name
- a** Adds (creates) the diskset
- h *nodename1*** Specifies the name of the primary node to master the diskset
- nodename2*** Specifies the name of the secondary node to master the diskset

5. **Verify the status of the new diskset.**

```
# metaset -s setname
```

Example—Creating a Diskset

The following command creates two disksets, `dg-schost-1` and `dg-schost-2`, assigning the nodes `phys-schost-1` and `phys-schost-2` as the potential primaries.

```
# metaset -s dg-schost-1 -a -h phys-schost-1 phys-schost-2
# metaset -s dg-schost-2 -a -h phys-schost-1 phys-schost-2
```

Where to Go From Here

To add drives to the diskset, go to “How to Add Drives to a Diskset” on page 113.

Adding Drives to a Diskset

When a disk drive is added to a diskset, Solstice DiskSuite repartitions it as follows so that the metadevice state database for the diskset can be placed on the drive.

- A small portion of each drive is reserved in slice 7 for use by Solstice DiskSuite software. The remainder of the space on each drive is placed into slice 0.
- Drives are repartitioned when they are added to the diskset only if slice 7 is not set up correctly.
- Any existing data on the disks is lost by the repartitioning.
- If slice 7 starts at cylinder 0, and the disk is large enough to contain a state database replica, the disk is not repartitioned.

▼ How to Add Drives to a Diskset

1. **Become superuser on the node.**
2. **Ensure that the diskset has been created.**
If necessary, refer to “How to Create a Diskset” on page 111.
3. **List the device ID (DID) mappings.**

```
# scdidadm -L
```

- Choose drives that are shared by the cluster nodes that will master or potentially master the diskset.
- Use the full DID pseudo-driver names when adding drives to a diskset.

The first column of output is the DID instance number, the second column is the full path (physical path), and the third column is the full DID pseudo-driver name (pseudo path). A shared drive has more than one entry for the same DID instance number.

In the following example, the entries for DID instance number 2 indicate a drive that is shared by `phys-schost-1` and `phys-schost-2`, and the full DID name is `/dev/did/rdisk/d2`.

```

1      phys-schost-1:/dev/rdisk/c0t0d0 /dev/did/rdisk/d1
2      phys-schost-1:/dev/rdisk/clt1d0 /dev/did/rdisk/d2
2      phys-schost-2:/dev/rdisk/clt1d0 /dev/did/rdisk/d2
3      phys-schost-1:/dev/rdisk/clt2d0 /dev/did/rdisk/d3
3      phys-schost-2:/dev/rdisk/clt2d0 /dev/did/rdisk/d3
...

```

4. Take ownership of the diskset.

```
# metaset -s setname -t
```

-s *setname* Specifies the diskset name

-t Takes ownership of the diskset

5. Add the drives to the diskset.

Use the full DID pseudo-driver name.

Note - Do *not* use the lower-level device name (`cNtXdY`) when adding a drive to a diskset. 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.

```
# metaset -s setname -a DIDname
```

-a *DIDname* Adds the disk drive to the diskset

6. Verify the status of the diskset and drives.

```
# metaset -s setname
```

Example—Adding Drives to a Diskset

The `metaset` command adds the disk drives `/dev/did/dsk/d1` and `/dev/did/dsk/d2` to the diskset `dg-schost-1`.

```
# metaset -s dg-schost-1 -a /dev/did/dsk/d1 /dev/did/dsk/d2
```

Where to Go From Here

To repartition drives for use in metadevices, go to “How to Repartition Drives in a Diskset” on page 115.

▼ How to Repartition Drives in a Diskset

The `metaset(1M)` command repartitions drives in a diskset so that a small portion of each drive is reserved in slice 7 for use by Solstice DiskSuite software. The remainder of the space on each drive is placed into slice 0. To make more effective use of the disk, use this procedure to modify the disk layout. Allocating space to slices 1 through 6 enables you to use these slices when setting up metadevices.

1. **Become superuser on the cluster node.**
2. **Use the `format(1M)` command to change the disk partitioning for each drive in the diskset.**

When you repartition a drive, you must meet the following conditions to prevent the `metaset(1M)` command from repartitioning the disk.

- Create a partition 7 starting at cylinder 0 that is large enough to hold a state database replica (approximately 2 Mbytes).
- Set the `Flag` field in slice 7 to `V_UNMT` (unmountable) and do not set it to read-only.
- Do not allow slice 7 to overlap any other slice on the disk.

Refer to the `format(1M)` man page for details.

Where to Go From Here

To define metadevices by using an `md.tab` file, go to “How to Create an `md.tab` File” on page 116.

▼ How to Create an `md.tab` File

Create an `/etc/lvm/md.tab` file for each node in the cluster.

Note - When using Solstice DiskSuite software, ensure that local metadevices have names that are distinct from the device ID (DID) names used to form disksets. For example, if the DID name `/dev/did/dsk/d3` is used in a diskset, do not use the name `/dev/md/dsk/d3` in configuring local metadevices.

1. Become superuser on the cluster node.

2. List the DID mappings for reference when creating your `md.tab` file.

Use the full DID pseudo-driver names in the `md.tab` file in place of the lower-level device names (`cNtXdY`).

```
# sccidadm -L
```

In the following example, the first column of output is the DID instance number, the second column is the full path (physical path), and the third column is the full DID pseudo-driver name (pseudo path).

```
1      phys-schost-1:/dev/rdisk/c0t0d0 /dev/did/rdsk/d1
2      phys-schost-1:/dev/rdisk/clt1d0 /dev/did/rdsk/d2
2      phys-schost-2:/dev/rdisk/clt1d0 /dev/did/rdsk/d2
3      phys-schost-1:/dev/rdisk/clt2d0 /dev/did/rdsk/d3
3      phys-schost-2:/dev/rdisk/clt2d0 /dev/did/rdsk/d3
...
```

3. Create an `/etc/lvm/md.tab` file and edit it by hand with your preferred text editor.

Refer to the Solstice DiskSuite documentation for details on creating an `md.tab` file.

Example—Sample `md.tab` File

The following sample `md.tab` file defines the metadevices for the diskset named `dg-schost-1`. The ordering of lines in the `md.tab` file is not important.

```
dg-schost-1/d0 -t dg-schost-1/d1 dg-schost-1/d4
dg-schost-1/d1 -m dg-schost-1/d2
dg-schost-1/d2 1 1 /dev/did/rdisk/d1s4
dg-schost-1/d3 1 1 /dev/did/rdisk/d55s4
dg-schost-1/d4 -m dg-schost-1/d5
dg-schost-1/d5 1 1 /dev/did/rdisk/d3s5
dg-schost-1/d6 1 1 /dev/did/rdisk/d57s5
```

The sample `md.tab` file is constructed as follows.

- The first line defines the trans metadevice `d0` to consist of a master (UFS) metadevice `d1` and a log device `d4`. The `-t` signifies this is a trans metadevice. The master and log devices are specified by their position after the `-t` flag.

```
dg-schost-1/d0 -t dg-schost-1/d1 dg-schost-1/d4
```

- The second line defines the master device as a mirror of the metadevices. The `-m` in this definition signifies a mirror device, and one of the submirrors, `d2`, is associated with the mirror device, `d1`.

```
dg-schost-1/d1 -m dg-schost-1/d2
```

- The fifth line similarly defines the log device, `d4`, as a mirror of metadevices.

```
dg-schost-1/d4 -m dg-schost-1/d5
```

- The third line defines the first submirror of the master device, `d2`, as a one-way stripe.

```
dg-schost-1/d2 1 1 /dev/did/rdisk/d1s4
```

- The fourth line defines the second submirror of the master device, `d3`.

```
dg-schost-1/d3 1 1 /dev/did/rdisk/d55s4
```

- Finally, the log device submirrors, `d5` and `d6`, are defined. In this example, simple metadevices for each submirror are created.

```
dg-schost-1/d5 1 1 /dev/did/rdisk/d3s5
dg-schost-1/d6 1 1 /dev/did/rdisk/d57s5
```

Note - If you have existing data on the disks that will be used for the submirrors, you must back up the data before metadevice setup and restore it onto the mirror.

Where to Go From Here

To activate the metadevices defined in the `md.tab` files, go to “How to Activate Metadevices” on page 118.

▼ How to Activate Metadevices

1. Ensure that you meet the following requirements.

- You have ownership of the diskset on the node where the command will be executed.
- The `md.tab` files are located in the `/etc/lvm` directory.

2. Become superuser on the cluster node.

3. Take ownership of the diskset.

```
# metaset -s setname -t
```

`-s setname` Specifies the diskset name

`-t` Takes ownership of the diskset

4. Activate the diskset’s metadevices, which are defined in the `md.tab` file.

```
# metainit -s setname -a
```

`-a` Activates all metadevices in the `md.tab` file

5. For each master and log device, attach the second submirror (*submirror2*).

When the metadevices in the `md.tab` file are activated, only the first submirror (*submirror1*) of the master and log devices is attached, so *submirror2* must be attached by hand.

```
# metattach mirror submirror2
```

6. Repeat Step 4 on page 118 and Step 5 on page 118 for each diskset in the cluster.

If necessary, run the `metainit(1M)` command from another node that has connectivity to the disks. This step is required for cluster-pair topologies, where the disks are not accessible by all nodes.

7. Check the status of the metadevices.

```
# metastat -s setname
```

Example—Activating Metadevices in the `md.tab` File

In the following example, all metadevices defined in the `md.tab` file for diskset `dg-schost-1` are activated. Then the second submirrors of master device `dg-schost-1/d1` and log device `dg-schost-1/d4` are activated.

```
# metainit -s dg-schost-1 -a
# metattach dg-schost-1/d1 dg-schost-1/d3
# metattach dg-schost-1/d4 dg-schost-1/d6
```

Where to Go From Here

If your cluster contains disksets configured with exactly two disk enclosures and two nodes, those disksets require mediators. To add mediator hosts, go to “Mediators Overview” on page 119. Otherwise, to create a cluster file system, go to “How to Add Cluster File Systems” on page 59.

Mediators Overview

A mediator, or mediator host, is a cluster node that stores mediator data. Mediator data provides information on the location of other mediators and contains a commit count that is identical to the commit count 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.

Mediators are required for all Solstice DiskSuite disksets configured with exactly two disk strings and two cluster nodes. A *disk string* consists of a disk enclosure, its physical disks, cables from the enclosure to the node(s), and the interface adapter cards. 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. The following rules apply to dual-string configurations using mediators.

- Disksets must be configured with exactly two mediator hosts, and those two mediator hosts must be the same two cluster nodes used for the diskset.
- A diskset cannot have more than two mediator hosts.
- Mediators cannot be configured for disksets that do not meet the two-string and two-host criteria.

These rules do not require that the entire cluster have exactly two nodes. Rather, they only require that those disksets 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 mediators.

1. **Become superuser on the node currently mastering the diskset to which you want to add mediator hosts.**
2. **Run the `metaset(1M)` command to add each node with connectivity to the diskset as a mediator host for that diskset.**

```
# metaset -s setname -a -m mediator_host_list
```

-s <i>setname</i>	Specifies the diskset name
-a	Adds to the diskset
-m <i>mediator_host_list</i>	Specifies the name of the node to add as a mediator host for the diskset

See the `mediator(7)` man page for details about mediator-specific options to the `metaset` command.

Example—Adding Mediator Hosts

The following example adds the nodes `phys-schost-1` and `phys-schost-2` as mediator hosts for the diskset `dg-schost-1`. Both commands are run from the node `phys-schost-1`.

```
# metaset -s dg-schost-1 -a -m phys-schost-1  
# metaset -s dg-schost-1 -a -m phys-schost-2
```

Where to Go From Here

To check the status of mediator data, go to “How to Check the Status of Mediator Data” on page 121.

▼ How to Check the Status of Mediator Data

1. Run the `medstat` command.

```
# medstat -s setname
```

`-s setname` Specifies the diskset name

If the value in the Status field is `Bad`, repair the affected mediator host by following the procedure “How to Fix Bad Mediator Data” on page 121.

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

Where to Go From Here

If the mediator data for one of the mediator hosts is bad, go to “How to Fix Bad Mediator Data” on page 121 to fix the problem. Otherwise, to create a cluster file system, go to “How to Add Cluster File Systems” on page 59.

▼ How to Fix Bad Mediator Data

Perform this procedure to repair bad mediator data. To determine the status of mediator data, follow the procedure “How to Check the Status of Mediator Data” on page 121.

1. Become superuser on the node that owns the affected diskset.
2. Remove the mediator host(s) with bad mediator data from all affected disksets.

```
# metaset -s setname -d -m mediator_host_list
```

`-s setname` Specifies the diskset name

`-d` Deletes from the diskset

`-m mediator_host_list` Specifies the name of the node to remove as a mediator host for the diskset

3. Restore the mediator host.

```
# metaset -s setname -a -m mediator_host_list
```

`-a` Adds to the diskset

`-m mediator_host_list` Specifies the name of the node to add as a mediator host for the diskset

See the `mediator(7)` man page for details about mediator-specific options to the `metaset` command.

Where to Go From Here

To create a cluster file system, go to “How to Add Cluster File Systems” on page 59.

Solstice DiskSuite Configuration Example

The following example helps to explain the process for determining the number of disks to place in each diskset when using Solstice DiskSuite software. This example assumes you are using three storage devices. In this example, existing applications are running over NFS (two file systems of 5 Gbytes each) and two Oracle databases (one 5 Gbytes and one 10 Gbytes).

The following table shows the calculations used to determine the number of drives needed in the sample configuration. If you have three storage devices, you would need 28 drives that would be divided as evenly as possible among each of the three storage devices. Note that the 5-Gbyte file systems were given an additional Gbyte of disk space because the number of disks needed was rounded up.

TABLE A-2 Determining Drives Needed for a Configuration

Use	Data	Disk Storage Needed	Drives Needed
<code>nfs1</code>	5 Gbytes	3x2.1 Gbyte disks * 2 (Mirror)	6
<code>nfs2</code>	5 Gbytes	3x2.1 Gbyte disks * 2 (Mirror)	6
<code>oracle1</code>	5 Gbytes	3x2.1 Gbyte disks * 2 (Mirror)	6
<code>oracle2</code>	10 Gbytes	5x2.1 Gbyte disks * 2 (Mirror)	10

The following table shows the allocation of drives among the two disksets and four data services.

TABLE A-3 Division of Disksets

Diskset	Data Services	Disks	Storage Device 1	Storage Device 2	Storage Device 3
dg-schost-1	nfs1/ oracle1	12	4	4	4
dg-schost-2	nfs2/ oracle2	16	5	6	5

Initially, four disks on each storage device (a total of 12 disks) are assigned to dg-schost-1, and five or six disks on each (a total of 16) are assigned to dg-schost-2.

No hot spare disks are assigned to either diskset. A minimum of one hot spare disk per storage device per diskset enables one drive to be hot spared (restoring full two-way mirroring).

Configuring VERITAS Volume Manager

Configure your local and multihost disks for VERITAS Volume Manager (VxVM) by using the procedures in this appendix, along with the information in Chapter 1. Refer to your VxVM documentation for additional details.

The following procedures are in this appendix.

- “How to Verify the Pseudo-Device Major Number” on page 126
- “How to Change the Pseudo-Device Major Number” on page 127
- “How to Encapsulate the Root Disk” on page 129
- “How to Create a Non-Root `rootdg` Disk Group” on page 135
- “How to Create and Register a Shared Disk Group” on page 136
- “How to Assign a New Minor Number to a Disk Device Group” on page 137
- “How to Verify the Disk Groups” on page 138

Configuring VxVM for Sun Cluster Configurations

The following table lists the tasks to perform to configure VxVM for Sun Cluster configurations.

TABLE B-1 Task Map: Configuring VxVM for Sun Cluster Configurations

Task	For Instructions, Go To ...
Plan the layout of your VxVM configuration.	“Planning Volume Management” on page 23
Verify that the pseudo-device major number is the same on each node.	“How to Verify the Pseudo-Device Major Number” on page 126
If necessary, change a node’s pseudo-device major number.	“How to Change the Pseudo-Device Major Number” on page 127
Create the root disk group (<code>rootdg</code>).	“Setting Up a <code>rootdg</code> Disk Group Overview” on page 128
Create shared disk groups and volumes.	“How to Create and Register a Shared Disk Group” on page 136
If necessary, resolve minor number conflicts between disk device groups by assigning a new minor number.	“How to Assign a New Minor Number to a Disk Device Group” on page 137
Verify the shared disk groups and volumes.	“How to Verify the Disk Groups” on page 138
Create and mount cluster file systems.	“How to Add Cluster File Systems” on page 59

▼ How to Verify the Pseudo-Device Major Number

The `vxio` driver must have identical pseudo-device major numbers on all cluster nodes. You can find this number in the `/etc/name_to_major` file after you complete the installation. Use the following procedures to verify the pseudo-device major numbers.

- 1. Become superuser on a node in the cluster.**
- 2. On each cluster node, view the pseudo-device major number.**

For example, type the following.

```
# grep vxio /etc/name_to_major
vxio 45
```

3. Compare the pseudo-device major numbers of all the nodes.

The major numbers should be identical on each node. If numbers vary, you must change the major numbers that are different.

Where to Go From Here

To change a node's pseudo-device major number, go to "How to Change the Pseudo-Device Major Number" on page 127.

Otherwise, to set up the root disk group (`rootdg`), go to "Setting Up a `rootdg` Disk Group Overview" on page 128.

▼ How to Change the Pseudo-Device Major Number

Perform this procedure if the pseudo-device major number is not the same for each node of the cluster.

- 1. Become superuser on a node that has the major number you want to change.**
- 2. Edit the `/etc/name_to_major` file to make the number identical on all nodes.**

Be sure that the number is unique in the `/etc/name_to_major` file for each node. A quick way to check for uniqueness is to find, by visual inspection, the maximum number assigned on each node in the `/etc/name_to_major` file, compute the maximum of these numbers, add one, then assign the sum to the `vxio` driver.
- 3. Reboot the node.**
 - a. Use the `scswitch(1M)` command to evacuate any resource groups or device groups from the node.**

```
# scswitch -s -h node
```

<code>-S</code>	Evacuates all resource groups and device groups.
<code>-h <i>node</i></code>	Specifies the name of the node from which to evacuate resource or device groups.

- b. Use the `shutdown(1M)` command to reboot the node.**

```
# shutdown -g 0 -y -i 6
```

4. If the system reports disk group errors and the cluster does not start, you might need to perform the following steps.
 - a. Become superuser on the node.
 - b. Use the `vxedit(1M)` command to change the `failing` field to `off` for affected subdisks.
Refer to the `vxedit(1M)` man page for more information.
 - c. Make sure all volumes are enabled and active.

Where to Go From Here

To set up the root disk group (`rootdg`), go to “Setting Up a `rootdg` Disk Group Overview” on page 128.

Setting Up a `rootdg` Disk Group Overview

Each cluster node requires the creation of a `rootdg` disk group. This disk group is used by VxVM to store configuration information, and has the following restrictions.

- Access to a node's `rootdg` disk group must be restricted to only that node.
- Remote nodes should never access data stored in another node's `rootdg`.
- Do not use the `scconf(1M)` command to register the `rootdg` disk group as a shared disk group.
- Whenever possible, configure the `rootdg` for each node on a non-shared disk.

Sun Cluster software supports the following methods for configuring the `rootdg` disk group.

- **Encapsulation of the node's root disk** – This method enables the root disk to be mirrored, providing a boot alternative if the root disk is corrupted or damaged. Root disk encapsulation requires two free disk slices as well as free cylinders, preferably at the beginning or the end of the disk.
- **Use of local, non-root disks** – VxVM can initialize or encapsulate local disks for use as `rootdg`, essentially creating a disk group that is local to that node (neither globally accessible nor highly available). As with the root disk, encapsulation of a

non-root disk requires two free disk slices as well as free cylinders at the beginning or the end of the disk.

Refer to your VxVM installation documentation for more information.

▼ How to Encapsulate the Root Disk

Use this procedure to create a `rootdg` disk group by encapsulating the root disk.

1. **Have available the VERITAS Volume Manager (VxVM) license keys.**
2. **Become superuser on a node in the cluster.**
3. **Ensure that the root disk has at least two slices with 0 cylinders and one or more free cylinders at the end or beginning of the disk.**

If necessary, use the `format(1M)` command to assign 0 cylinders to each VxVM slice. If slice 7 was reserved for volume manager use, formatting slice 7 also frees the extra space needed at the end of the disk.

4. **Start the `vxinstall(1M)` utility.**

```
# vxinstall
```

When prompted, make the following choices or entries.

- Choose Custom Installation.
- Encapsulate the root disk.
- Choose a name for the root disk that is unique throughout the cluster. A simple way to name the root disk is to add an extra letter to the default name.
For example, if the default name given is `rootdisk`, name the root disk `rootdiska` on one node, `rootdiskb` on the next node, and so forth.
- Do not add any other disks to the `rootdg` disk group.
- For any other controllers, choose 4 (Leave these disks alone).
- Do not accept shutdown and reboot.

Refer to the VxVM installation documentation for details.

Note - Because Dynamic Multipathing (DMP) is disabled, an error message similar to the following might be generated. You can safely ignore it.

```
vxvm:vxdmadm: ERROR: vxdmp module is not loaded on the system. Command invalid.
```

5. Edit the `/etc/vfstab` file device names for the `/global/.devices/node@nodeid` file system.

Note - You need to make this modification in order for VxVM to recognize that the `/global/.devices/node@nodeid` file system is on the root disk.

Replace the existing device names with the names used in the `/globaldevices` entry, which is commented out. For example, consider the following `/etc/vfstab` file entries for `/globaldevices` and `/global/.devices/node@2`.

```
#device          device          mount          FS   fsck  mount  mount
#to mount        to fsck         point          type pass  at boot options
...
#/dev/dsk/clt3d0s3 /dev/rdisk/clt3d0s3 /globaldevices ufs 2   yes   -
...
/dev/did/dsk/d4s3  /dev/did/rdsk/d4s3 /global/.devices/node@2 ufs 2 no global
```

You would change the `/global/.devices/node@2` entry to the following.

```
#device          device          mount          FS   fsck  mount  mount
#to mount        to fsck         point          type pass  at boot options
...
#/dev/dsk/clt3d0s3 /dev/rdisk/clt3d0s3 /globaldevices ufs 2   yes   -
...
/dev/dsk/clt3d0s3  /dev/rdisk/clt3d0s3 /global/.devices/node@2 ufs 2 no global
```

6. Repeat Step 2 on page 129 through Step 5 on page 130 on each node of the cluster.
7. From one node, use the `scshutdowm(1M)` command to shut down the cluster.

```
# scshutdowm
```

8. Reboot each node in *non-cluster* mode.
 - a. Run the following command on each node to reboot in *non-cluster* mode.

```
ok boot -x
```

Note - Do *not* reboot the node in cluster mode.

b. If a node displays a message similar to the following, press Control-D to continue the boot.

Ignore the instruction to run `fsck` manually. Instead, press Control-D to continue with the boot and complete the remaining root disk encapsulation procedures.

```
WARNING - Unable to repair the /global/.devices/node@1 filesystem.
Run fsck manually (fsck -F ufs /dev/vx/rdisk/rootdisk3vola). Exit
the shell when done to continue the boot process.
```

```
Type control-d to proceed with normal startup,
(or give root password for system maintenance):
```

The `/global/.devices/node@nodeid` file system still requires additional changes before the cluster can globally mount it on each node. Because of this requirement, all but one node will fail to mount the `/global/.devices/node@nodeid` file system during this reboot, resulting in a warning message.

VxVM encapsulates the root disk and updates the `/etc/vfstab` entries.

9. Unmount the `/global/.devices/node@nodeid` file system that successfully mounted in Step 8 on page 130.

```
# umount /global/.devices/node@nodeid
```

Unmounting this file system enables you to reminor the disk group during Step 10 on page 131 without needing to reboot the node twice to initialize the change. This file system is automatically remounted when you reboot during Step 14 on page 134.

10. Re-minor the `rootdg` disk group on each node of the cluster.

Specify a `rootdg` minor number that is unique throughout the cluster and smaller than 1000 to prevent minor number conflicts with shared disk groups. An effective re-minoring scheme is to assign 100 on the first node, 200 on the second, and so forth.

```
# vxvg remminor rootdg n
```

n Specifies the `rootdg` minor number

After executing this command, warning messages similar to the following might be displayed. You can safely ignore them.

```
vxvm:vxvg: WARNING: Volume swapvol: Device is open, will renumber on reboot
```

The new minor number is applied to the root disk volumes. The swap volume is renumbered after you reboot.

```
# ls -l /dev/vx/dsk/rootdg
total 0
brw----- 1 root      root      55,100 Apr  4 10:48 rootdiska3vol
brw----- 1 root      root      55,101 Apr  4 10:48 rootdiska7vol
brw----- 1 root      root      55,  0 Mar 30 16:37 rootvol
brw----- 1 root      root      55,  7 Mar 30 16:37 swapvol
```

11. On each node of the cluster, if the /usr file system is not collocated with the root (/) file system on the root disk, manually update the device nodes for the /usr volume.

a. Remove existing /usr device nodes.

```
# rm /dev/vx/dsk/usr
# rm /dev/vx/dsk/rootdg/usr
# rm /dev/vx/rdsk/usr
# rm /dev/vx/rdsk/rootdg/usr
```

b. Determine the new minor number assigned to the /usr file system.

```
# vxprint -l -v usrvol
Disk group: rootdg Volume:  usrvol
...
device:   minor=102 bdev=55/102 cdev=55/102 path=/dev/vx/dsk/rootdg/
usrvol
```

c. Create new /usr device nodes by using the new minor number.

```
# mknod /dev/vx/dsk/usr b major_number new-minor-number
# mknod /dev/vx/dsk/rootdg/usr b major_number new-minor-number
# mknod /dev/vx/rdisk/usr c major_number new-minor-number
# mknod /dev/vx/rdisk/rootdg/usr c major_number new-minor-number
```

12. On each node of the cluster, if the /var file system is not collocated with the root (/) file system on the root disk, manually update the device nodes for the /var volume.

a. Remove existing /var device nodes.

```
# rm /dev/vx/dsk/var
# rm /dev/vx/dsk/rootdg/var
# rm /dev/vx/rdisk/var
# rm /dev/vx/rdisk/rootdg/var
```

b. Determine the new minor number assigned to the /var file system.

```
# vxprint -l -v usrvol
Disk group: rootdg Volume:  usrvol
...
device:  minor=103 bdev=55/102 cdev=55/102 path=/dev/vx/dsk/rootdg/
usrvol
```

c. Create new /var device nodes by using the new minor number.

```
# mknod b /dev/vx/dsk/var major_number new-minor-number
# mknod b /dev/vx/dsk/rootdg/var major_number new-minor-number
# mknod c /dev/vx/rdisk/var major_number new-minor-number
# mknod c /dev/vx/rdisk/rootdg/var major_number new-minor-number
```

13. From one node, shut down the cluster.

```
# scshutdwn
```

14. Reboot each node into cluster mode.

```
ok boot
```

15. Mirror the root disk on each node of the cluster.

Refer to your VxVM documentation for instructions on mirroring root.

16. If you mirrored the root disk, on each node of the cluster enable the `localonly` property of the raw disk device group associated with the disk used to mirror that node's root disk.

For each node, configure a different raw disk device group, which will be used exclusively by that node to mirror the root disk. You must enable the `localonly` property to prevent unintentional fencing of a node from its boot device if the boot device is connected to multiple nodes.

```
# scconf -c -D name=rawdsk_groupname,localonly=true
```

-D name=rawdsk_groupname Specifies the cluster-unique name of the raw disk device group

Use the `scdidadm -L` command to display the full device ID (DID) pseudo-driver name of the raw disk device group. In the following example, the raw disk device group name `dsk/d1` is extracted from the third column of output, which is the full DID pseudo-driver name. The `scconf` command then configures the `dsk/d1` raw disk device group to be used exclusively by the node `phys-schost-3` to mirror its root disk.

```
# scdidadm -L
...
1      phys-schost-3:/dev/rdsk/c0t0d0      /dev/did/rdsk/d1
phys-schost-3# scconf -c -D name=dsk/d1,localonly=true
```

For more information about the `localonly` property, refer to the `scconf_dg_rawdisk(1M)` man page.

Where to Go From Here

To create shared disk groups, go to “How to Create and Register a Shared Disk Group” on page 136.

▼ How to Create a Non-Root `rootdg` Disk Group

Use this procedure to create a `rootdg` disk group by encapsulating or initializing local non-root disks.

1. **Have available the VERITAS Volume Manager (VxVM) license keys.**
2. **Become superuser on the node.**
3. **If the disks will 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.
4. **Start the `vxinstall(1M)` utility.**

```
# vxinstall
```

When prompted, make the following choices or entries.

- Choose Custom Installation.
- Do not encapsulate the root disk.
- Choose any disks you want added to the `rootdg` disk group.
- Do not accept automatic reboot.

5. **Evacuate any resource groups or device groups from the node.**

```
# scswitch -s -h node
```

<code>-s</code>	Evacuates all resource groups and device groups.
<code>-h <i>node</i></code>	Specifies the name of the node from which to evacuate resource or device groups.

6. **Reboot the node.**

```
# shutdown -g 0 -y -i 6
```

Where to Go From Here

To create shared disk groups, go to “How to Create and Register a Shared Disk Group” on page 136.

▼ How to Create and Register a Shared Disk Group

Use this procedure to create your VxVM disk groups and volumes.

Run this procedure from a node that is physically connected to the disks that make up the disk group being added.

Note - After the disk group has been registered with the cluster as a disk device group, you should never import or deport VxVM disk groups by using VxVM commands. The Sun Cluster software can handle all cases where disk groups need to be imported or deported. Refer to *Sun Cluster 3.0 System Administration Guide* for procedures on managing Sun Cluster disk device groups.

1. Have available the following information.

- Mappings of your storage disk drives. Refer to the *Sun Cluster 3.0 Hardware Guide* chapter on performing an initial installation for your storage device.
- The following completed configuration planning worksheets from *Sun Cluster 3.0 Release Notes*.
 - “Local File System Layout Worksheet”
 - “Disk Device Group Configurations Worksheet”
 - “Volume Manager Configurations Worksheet”

See Chapter 1 for planning guidelines.

2. Become superuser on the node that will have ownership of the disk group.

3. Create the VxVM disk group and volume.

Use your preferred method to create the disk group and volume.

Note - You can use Dirty Region Logging (DRL) to decrease volume recovery time in the event of a node failure. However, using DRL might decrease I/O throughput.

See the VERITAS Volume Manager documentation for the procedures to complete this step.

4. Register the disk group as a Sun Cluster disk device group.

a. Start the `scsetup(1M)` utility.

```
# scsetup
```

b. To work with disk device groups, type 3 (Device groups).

c. To register a disk device group, type 1 (Register a VxVM disk group).

Follow the instructions and type the VxVM disk device group to be registered as a Sun Cluster disk device group.

If you encounter the following error while attempting to register the disk device group, use the procedure “How to Assign a New Minor Number to a Disk Device Group” on page 137. This procedure enables you to assign a new minor number that does not conflict with a minor number used by existing disk device groups.

```
scconf: Failed to add device group - in use
```

d. When finished, type `q` (Quit) to leave the `scsetup` utility.

5. **Verify that the disk device group has been registered.**

Look for the disk device information for the new disk displayed by the following command.

```
# scconf -pv | egrep disk-device-group
```

Note - If you change any configuration information for a VxVM disk group or volume, re-register the Sun Cluster disk device group. Re-registering the disk device group ensures that the global namespace is in the correct state. Refer to *Sun Cluster 3.0 System Administration Guide* for procedures for re-registering a disk device group.

Where to Go From Here

To verify your VxVM disk groups and volumes, go to “How to Verify the Disk Groups” on page 138.

▼ How to Assign a New Minor Number to a Disk Device Group

If registering a disk device group fails because of a minor number conflict with another disk group, the new disk group must be assigned a new, unused minor number. After assigning the new minor number, you then re-register the disk group as a Sun Cluster disk device group.

1. **Become superuser on a node of the cluster.**
2. **Determine the minor numbers in use.**

```
# ls -l /global/.devices/node@1/dev/vx/dsk/*
```

3. Choose any other multiple of 1000 that is not in use to be the base minor number for the new disk group.
4. Assign the new base minor number to the disk group.

```
# vxdg reminor diskgroup base_minor_number
```

5. Return to Step 4 on page 136 of “How to Create and Register a Shared Disk Group” on page 136 to register the disk group as a Sun Cluster disk device group.

Example—How to Assign a New Minor Number to a Disk Device Group

This example shows the minor numbers 16000-16002 and 4000-4001 being used. The `vxdg reminor` command is used to re-minor the new disk device group to use the base minor number 5000.

```
# 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
# vxdg reminor dg3 5000
```

Where to Go From Here

You must register the disk group as a Sun Cluster disk device group. Go to Step 4 on page 136 of “How to Create and Register a Shared Disk Group” on page 136.

▼ How to Verify the Disk Groups

Perform this procedure on each node of the cluster.

1. Verify that only the local disks are included in the root disk group (`rootdg`), and shared disk groups are imported on the current primary node only.

```
# vxdisk list
```

2. Verify that all volumes have been started.

```
# vxprint
```

3. Verify that all shared disk groups have been registered as Sun Cluster disk device groups and are online.

```
# scstat -D
```

Where to Go From Here

To configure cluster file systems, go to “How to Add Cluster File Systems” on page 59.