



Sun™ Cluster 3.0 and Sun StorEdge™ Software Integration Guide

For The Sun StorEdge Availability Suite Software

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054 U.S.A.
650-960-1300

Part No. 816-5127-11
May 2003, [Revision A](#)

Send comments about this document to: docfeedback@sun.com

Copyright 2003 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, California 95054, U.S.A. All rights reserved.

Sun Microsystems, Inc. has intellectual property rights relating to technology that is described in this document. In particular, and without limitation, these intellectual property rights may include one or more of the U.S. patents listed at <http://www.sun.com/patents> and one or more additional patents or pending patent applications in the U.S. and in other countries.

This document and the product to which it pertains are distributed under licenses restricting their use, copying, distribution, and decompilation. No part of the product or of this 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 in other countries, exclusively licensed through X/Open Company, Ltd.

Sun, Sun Microsystems, the Sun logo, Sun StorEdge, Sun Fire, AnswerBook2, docs.sun.com, and Solaris are trademarks or registered trademarks of Sun Microsystems, Inc. in the U.S. and in other countries.

All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. and in 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.

U.S. Government Rights—Commercial use. Government users are subject to the Sun Microsystems, Inc. standard license agreement and applicable provisions of the FAR and its supplements.

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 2003 Sun Microsystems, Inc., 4150 Network Circle, Santa Clara, California 95054, Etats-Unis. Tous droits réservés.

Sun Microsystems, Inc. a les droits de propriété intellectuelle relatants à la technologie qui est décrit dans ce document. En particulier, et sans la limitation, ces droits de propriété intellectuelle peuvent inclure un ou plus des brevets américains énumérés à <http://www.sun.com/patents> et un ou les brevets plus supplémentaires ou les applications de brevet en attente dans les Etats-Unis et dans les autres pays.

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, Sun StorEdge, Sun Fire, AnswerBook2, docs.sun.com, et Solaris sont des marques de fabrique ou des marques déposées 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 licenciées de Sun qui mettent en place l'interface d'utilisation graphique OPEN LOOK et qui en outre se conforment aux licences écrites de Sun.

LA DOCUMENTATION EST FOURNIE "EN L'ÉTAT" ET TOUTES AUTRES CONDITIONS, DECLARATIONS ET GARANTIES EXPRESSES OU TACITES SONT FORMELLEMENT EXCLUES, DANS LA MESURE AUTORISEE PAR LA LOI APPLICABLE, Y COMPRIS NOTAMMENT TOUTE GARANTIE IMPLICITE RELATIVE A LA QUALITE MARCHANDE, A L'APTITUDE A UNE UTILISATION PARTICULIERE OU A L'ABSENCE DE CONTREFAÇON.

Contents

Preface ix

1. Overview 1

Terminology Used in This Guide 2

Supported Software and Hardware 3

Using the Sun StorEdge Availability Suite Software in a Sun Cluster
Environment 4

Global and Local Use of the Sun StorEdge Availability Suite Software 5

Volumes Eligible for Use with the Remote Mirror and Point-in-time Copy
Software 6

Sun StorEdge Fast Write Cache 7

2. Installing and Configuring The Sun StorEdge Availability Suite Software 9

Shutting Down Nodes After Installing the Sun StorEdge Availability Suite
Software 10

Task Summary: Steps to Configure the Sun Cluster 3.0 Software and Install the
Sun StorEdge Availability Suite Software 11

Disk Device Groups and the Sun StorEdge Availability Suite Software 12

Requirements for the Configuration Location 13

Running the `probe_script` Validation Script 15

Syntax 15

`probe_script [-h | -j]` 15

▼ To Run the <code>probe_script</code> File	16
Installing the Software	18
Installation Script Syntax	18
<code>install.sh [-j] {-a -p -r}</code>	18
▼ To Install the Software	19
Editing the Sun StorEdge Availability Suite Bitmap Parameter Files	21
The <code>/usr/kernel/drv/rdc.conf</code> File	22
▼ To Edit the <code>rdc.conf</code> File	22
The <code>/usr/kernel/drv/ii.conf</code> File	23
▼ To Edit the <code>ii.conf</code> File	23
Shutting Down and Restarting the Nodes	24
▼ To Shut Down and Restart a Node	24
Supported Configurations for The Remote Mirror Software	25
The Remote Mirror Primary Host is Configured On a Cluster Node	25
The Remote Mirror Secondary Host is Configured On a Cluster Node	26
The Remote Mirror Primary and Secondary Hosts are Configured On a Cluster Node	26
Supported Configurations for the Point-in-time Copy Software	27
Configuration Rules for the Point-in-time Copy and Remote Mirror Software	27
Rules For the Point-in-time Copy Software	27
Rules For the Remote Mirror Software	28
Configuring the Sun Cluster 3.0 Environment for Use with the Sun StorEdge Availability Suite Software	29
▼ To Configure the Sun Cluster Environment for HAStorage	30
▼ To Configure the Sun Cluster Environment for HAStoragePlus	33
3. Using the Sun StorEdge Availability Suite <code>iiadm</code> and <code>sndradm</code> Commands	37
Configuring Remote Mirror and Point-in-time Copy Volume Sets	38
Using the Remote Mirror Software With the Point-in-Time Copy Software	39
Mounting and Replicating Global Volume File Systems	40

The <code>iiadm</code> and <code>sndradm</code> Command Syntax	41
Global Device Command Syntax	41
Local Device Command Syntax	43
The <code>C tag</code> and <code>-C tag</code> Options	45
Which Host Do I Issue Commands From?	46
Putting All Cluster Volume Sets in an I/O Group	48
Preserving Point-in-time Copy Volume Data During a System Failure or Failover	50
A Directory Named <code>._</code> is Created After Mounting a Secondary Volume	51

Tables

TABLE 1-1	Supported Software and Hardware	3
TABLE 1-2	Cluster Terminology and Status	4
TABLE 2-2	Configuration Location Requirements and Considerations	14
TABLE 2-3		29
TABLE 2-4		30
TABLE 2-5		31
TABLE 2-6		32
TABLE 2-7		33
TABLE 2-8		34
TABLE 2-9		35
TABLE 3-1	Which Host to Issue Remote Mirror Commands From	46

Preface

The *Sun Cluster 3.0 and Sun StorEdge Software Integration Guide* describes how to integrate the Sun StorEdge™ Availability Suite 3.1 remote mirror and point-in-time copy software products in Sun™ Cluster 3.0 Update 1 and Update 2 environments.

This guide is intended for system administrators who have experience with the Solaris™ operating environment, Sun Cluster software, and related disk storage systems.

Note – The remote mirror software was previously known as the Sun StorEdge Network Data Replicator (Sun SNDR) software. The point-in-time copy software was previously known as the Sun StorEdge Instant Image software. Previous versions are also referred to as the data services software.

Sun Cluster 3.0 Update 1 is also known as the Sun Cluster 3.0 07/01 release. Sun Cluster 3.0 Update 2 is also known as the 12/01 release.



Caution – Do not install the Sun StorEdge Availability Suite 3.1 software on a system running the initial release of the Sun Cluster 3.0 software.

Before You Read This Book

Note – Before you install the Sun StorEdge Availability Suite software as described in the installation and release documentation in [“Related Documentation” on page xii](#), see [Chapter 2](#).

To fully use the information in this document, you must have thorough knowledge of the topics discussed in the books in [“Related Documentation” on page xii](#).

How This Book Is Organized

[Chapter 1](#) is an overview of the Sun Cluster and Sun StorEdge Availability Suite software integration.

[Chapter 2](#) describes installing and configuring the Sun StorEdge Availability Suite software for use in a Sun Cluster environment.

[Chapter 3](#) describes using the Sun StorEdge Availability Suite software commands in a Sun Cluster environment.

Using UNIX Commands

This document may not contain information on basic UNIX® commands and procedures such as shutting down the system, booting the system, and configuring devices.

See the software documentation that you received with your system.

Typographic Conventions

Typeface	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, when 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> .
[]	In syntax, brackets indicate that an argument is optional.	<code>scmadm [-d sec] [-r n[:n][,n]...] [-z]</code>
{ arg arg }	In syntax, braces and pipes indicate that one of the arguments must be specified.	<code>sndradm -R b {p s}</code>
\	At the end of a command line, the \ (backslash) indicates that the command continues on the next line.	<code>scrgadm -a -L \</code> <code>-g <i>groupname</i>-stor-rg \</code> <code>-l <i>lhost1</i>, <i>lhost2</i> \</code> <code>-n <i>nafo0@node</i>, <i>nafo0@node</i></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 U1 Hardware Guide</i>	806-7070
	<i>Sun Cluster 3.0 12/01 Hardware Guide</i>	816-2023
Software Installation	<i>Sun Cluster 3.0 U1 Installation Guide</i>	806-7069
	<i>Sun Cluster 3.0 12/01 Software Installation Guide</i>	816-2022
	<i>Sun StorEdge Availability Suite 3.1 Remote Mirror Software Installation Guide</i>	816-4413
	<i>Sun StorEdge Availability Suite 3.1 Point-in-time Copy Software Installation Guide</i>	816-4312
Data Services	<i>Sun Cluster 3.0 U1 Data Services Installation and Configuration Guide</i>	806-7071
	<i>Sun Cluster 3.0 12/01 Data Services Installation and Configuration Guide</i>	816-2024
Concepts	<i>Sun Cluster 3.0 U1 Concepts</i>	806-7074
	<i>Sun Cluster 3.0 12/01 Concepts</i>	816-2027
Error Messages	<i>Sun Cluster 3.0 U1 Error Messages Manual</i>	806-7076
	<i>Sun Cluster 3.0 12/01 Error Messages Manual</i>	816-2028
Release Notes	<i>Sun Cluster 3.0 U1 Release Notes</i>	806-7078
	<i>Sun Cluster 3.0 12/01 Release Notes</i>	816-2029
	<i>Sun Cluster 3.0 U1 Release Notes Supplement</i>	806-7079
	<i>Sun Cluster 3.0 12/01 Release Notes Supplement</i>	816-3753
	<i>Sun StorEdge Availability Suite 3.1 Remote Mirror Software Release Notes</i>	816-4314
	<i>Sun StorEdge Availability Suite 3.1 Point-in-time Copy Software Release Notes</i>	816-4314
System Administration	<i>Sun Cluster 3.0 U1 System Administration Guide</i>	806-7073

Application	Title	Part Number
	<i>Sun Cluster 3.0 12/01 System Administration Guide</i>	816-2026
	<i>Sun StorEdge Availability Suite 3.1 Remote Mirror Software Administration and Operations Guide</i>	816-4315
	<i>Sun StorEdge Availability Suite 3.1 Point-in-time Copy Software Administration and Operations Guide</i>	816-4313

Accessing Sun Documentation Online

A broad selection of Sun system documentation is located at:

<http://www.sun.com/products-n-solutions/hardware/docs>

A complete set of Solaris documentation and many other titles are located at:

<http://docs.sun.com>

For the latest version of released storage software documentation, go to:

<http://www.sun.com/products-n-solutions/hardware/docs/Software/>

Sun Welcomes Your Comments

Sun is interested in improving its documentation and welcomes your comments and suggestions. You can email your comments to Sun at:

docfeedback@sun.com

Please include the part number (816-5127-11) of your document in the subject line of your email.

Overview

Note – This guide assumes that you have already installed the volume manager software and the Sun Cluster 3.0 Update 1 (07/01) or Update 2 (12/01) software on each node in your cluster.

The Sun Cluster and Sun StorEdge Availability Suite 3.1 software combine to provide a highly available environment for cluster storage. The remote mirror software is a data replication application that provides access to data as part of business continuance and disaster recovery plans. The point-in-time copy software is a point-in-time snapshot copy application that enables you to create copies of application or test data.

The topics in this chapter include:

- [“Terminology Used in This Guide” on page 2](#)
- [“Supported Software and Hardware” on page 3](#)
- [“Using the Sun StorEdge Availability Suite Software in a Sun Cluster Environment” on page 4](#)
- [“Sun StorEdge Fast Write Cache” on page 7](#)

Terminology Used in This Guide

Data service

Highly Available (HA) applications within the Sun Cluster environment are also known as data services. The term *data service* is used to describe a third-party application that has been configured to run on a cluster rather than on a single server. A data service includes the application software and Sun Cluster software that starts, stops, and monitors the application.

Primary and secondary hosts and nodes

In this guide and the remote mirror software documentation, the terms *primary host* and *secondary host* are used as follows.

- The primary and secondary hosts are physically-separate servers running the remote mirror software. The primary host contains the primary volume and bitmap volume to be initially replicated to a remote server called a secondary host. The secondary hosts contains the secondary volume and bitmap volume

The terms *primary node* and *secondary node* refers to cluster nodes with respect to device group mastering in a cluster.

Supported Software and Hardware

Note – You cannot use the Sun StorEdge Fast Write Cache (FWC) product (all versions) in any Sun Cluster environment because cached data is inaccessible from other machines in a cluster. To compensate, you can use a Sun caching array. See [“Sun StorEdge Fast Write Cache” on page 7](#).

TABLE 1-1 Supported Software and Hardware

Operating Environment Software	Solaris™ 8, all releases that are supported by the Sun Cluster 3.0 Update 1 and Update 2 software
Sun Cluster Software	Sun Cluster 3.0 Update 1 software (also known as the 07/01 release) Sun Cluster 3.0 Update 2 software (also known as the 12/01 software)
Volume Manager Software	Solaris Disk Suite 4.2.1, Solaris Volume Manager VERITAS Volume Manager (VxVM) 3.1
Sun StorEdge Software	Sun StorEdge Availability Suite 3.1 remote mirror and point-in-time copy software.
Supported Cluster Configuration	The Sun Cluster 3.0 (07/01 and 12/01) and Sun StorEdge Availability Suite software are supported in a two-node cluster environment only.
Disk Space Requirements	The Sun StorEdge Availability Suite software requires approximately 11 Mbytes of disk space. <ul style="list-style-type: none">• The remote mirror software requires approximately 1.4 Mbytes• The point-in-time copy software requires approximately 1 Mbyte• The configuration location requires 5.5 Mbytes• Supporting core packages require approximately 3 Mbytes
Hardware	<p>A CD-ROM drive connected to the host server where the Sun software is to be installed.</p> <p>The remote mirror and point-in-time copy software is supported on servers using the Solaris operating environment and any network interface card supported by Sun. Hosts include but are not limited to:</p> <ul style="list-style-type: none">• Sun Enterprise™ 220R, 250, 420R, and 450 servers• Sun Enterprise 3500, 4500, 5500, 6500, and 10000 servers• Sun Fire 3800, 4800, 4810, 6800, and 15K servers

Using the Sun StorEdge Availability Suite Software in a Sun Cluster Environment

To use cluster failover features with the Sun StorEdge Availability Suite software, your software environment requires the Sun Cluster Update 1 (07/01 release) or Update 2 (12/01) software.

In this environment, the Sun StorEdge Availability Suite software is *cluster aware*. See TABLE 1-2.

The `sndradm` and `iiadm` commands are used to control the Sun StorEdge Availability Suite software. You can use the command options `C tag` and `-C tag` in a cluster environment only. If you accidentally use these options in a noncluster environment, the specified operation does not execute. See [Chapter 3](#) in this guide for more information.

TABLE 1-2 Cluster Terminology and Status

Term	Definition	Sun StorEdge Availability Suite Status
Cluster aware	A software product is Sun Cluster aware if it can coexist with the Sun Cluster environment and fails over and fails back as the logical host containing the software product fails over and fails back. A Sun Cluster aware product can then be made highly available by utilizing the High Availability framework that Sun Cluster provides.	The Sun StorEdge Availability Suite 3.1 software is cluster-aware in a two-node, Sun Cluster 3.0 Update 1 and Update 2 software environment.
Cluster tolerant or coexistent	A software product is Sun Cluster tolerant if it can coexist with the Sun Cluster environment and does not interfere with the Sun Cluster software and applications running in this environment. A product that is cluster tolerant is not expected to fail over or fail back when a Sun Cluster logical host fails over and fails back.	The Sun StorEdge Availability Suite 3.1 software <i>is not cluster-tolerant</i> in the initial release of the Sun Cluster 3.0 software.

Global and Local Use of the Sun StorEdge Availability Suite Software

Note – See [“Configuration Rules for the Point-in-time Copy and Remote Mirror Software”](#) on page 27.

The Sun StorEdge Availability Suite software can use volumes that are local or global devices. Global devices are those Sun StorEdge Availability Suite software or other volumes accessible from any cluster node and which will fail over under the control of the Sun Cluster framework. Local devices are volumes that are local to the individual node (host machine), not defined in a disk device or resource group, and not managed within a cluster file system. Local devices do not fail over and switch back.

To access local devices, use the `C local` or `-C local` options as part of the `sndradm` and `iiadm` commands. To access global devices, use the command options `C tag` and `-C tag`. (Typically, you do not need to specify the `-C tag` option as `iiadm` and `sndradm` automatically detect the disk device group.)

See [Chapter 3](#) in this guide and the Sun StorEdge Availability Suite administration and operations guides listed in [“Related Documentation”](#) on page xii.

Volumes Eligible for Use with the Remote Mirror and Point-in-time Copy Software

You can replicate the following critical volumes using the remote mirror software:

- Database and database management system (DBMS) logs (the total database or online DBMS log)
- Access control files

You can exclude volumes from replication if they can be reconstructed at the recovery site or if they seldom change:

- Temporary volumes (such as those used in sort operations)
- Spool files
- Paging volumes

When selecting a volume to be used in the volume set (including the configuration location), ensure that volume does not contain disk label private areas (for example, slice 2 on a Solaris operating environment-formatted volume). The disk label region is contained in the first sectors of cylinder 0 of a disk.

The point-in-time copy software supports all Sun-supported storage. It works independently of the underlying data reliability software (for example, RAID-1, RAID-5, or volume manager). Additionally, you can use it as a tool when migrating data to and from differing storage types.

Typical uses for the point-in-time copy software include:

- Backup of live application data
- Load data warehouses and fast resynchronization of data warehouses at predefined intervals
- Application development and test on a point-in-time snapshot of live data
- Migrate data across different types of storage platforms and volumes
- Hot back up of application data from frequent point-in-time snapshots

Sun StorEdge Fast Write Cache

Note – The Sun StorEdge Fast Write Cache product, all versions, is not supported in any cluster environment.

All versions of the Sun StorEdge Fast Write Cache product, including the SUNWnvm packages included with the Sun StorEdge Availability Suite versions software 3.0 and 3.0.1, are not supported when in a Sun Cluster 3.0 environment because cached data is inaccessible from other machines in a cluster. To compensate, you can use a Sun caching array.

Installing and Configuring The Sun StorEdge Availability Suite Software

Note – This guide assumes that you have already installed the volume manager software and the Sun Cluster 3.0 Update 1 (07/01) or Update 2 (12/01) software on each node in your cluster.



Caution – Do not install the Sun StorEdge Availability Suite 3.1 software on a system running the initial release of the Sun Cluster 3.0 software.

The Sun StorEdge Availability Suite installation guides listed in [“Related Documentation” on page xii](#) describe how to install the Sun StorEdge Availability Suite software in a noncluster environment. The installation steps to install this software in a Sun Cluster 3.0 environment are generally the same as described in the installation guides. This chapter describes the differences when you install the software in a Sun Cluster 3.0 environment.

The topics in this chapter include:

- [“Shutting Down Nodes After Installing the Sun StorEdge Availability Suite Software” on page 10](#)
- [“Task Summary: Steps to Configure the Sun Cluster 3.0 Software and Install the Sun StorEdge Availability Suite Software” on page 11](#)
- [“Disk Device Groups and the Sun StorEdge Availability Suite Software” on page 12](#)
- [“Requirements for the Configuration Location” on page 13](#)
- [“Running the probe_script Validation Script” on page 15](#)
- [“Installing the Software” on page 18](#)
- [“Editing the Sun StorEdge Availability Suite Bitmap Parameter Files” on page 21](#)
- [“Shutting Down and Restarting the Nodes” on page 24](#)

- [“Supported Configurations for The Remote Mirror Software” on page 25](#)
- [“Supported Configurations for the Point-in-time Copy Software” on page 27](#)
- [“Configuration Rules for the Point-in-time Copy and Remote Mirror Software” on page 27](#)
- [“Configuring the Sun Cluster 3.0 Environment for Use with the Sun StorEdge Availability Suite Software” on page 29](#)

Shutting Down Nodes After Installing the Sun StorEdge Availability Suite Software

Because the Sun StorEdge Availability Suite installation process requires you to shut down and restart each node in the cluster, *make sure that you install the Sun StorEdge Availability Suite software during your normal maintenance window.*

As a result of this shut down and restart, you might experience a panic condition on the node you are restarting. The node panic is expected behavior in the cluster and is part of the cluster software’s *failfast mechanism*. The *Sun Cluster 3.0 Concepts* manual describes this mechanism and the Cluster Membership Monitor (CMM).

See [“Shutting Down and Restarting the Nodes” on page 24.](#)

Task Summary: Steps to Configure the Sun Cluster 3.0 Software and Install the Sun StorEdge Availability Suite Software

The recommended installation order in a cluster environment is as follows for each node. See TABLE 2-1.

1. Install the volume manager software.
2. Install the Sun Cluster 3.0 software.
3. Install the Sun StorEdge Availability Suite software.

TABLE 2-1 Installation and Configuration Steps

Installation Steps	For More Information, See...
1. Select a configuration location.	“Requirements for the Configuration Location” on page 13
2. Run the <code>probe_script</code> validation script.	“Running the <code>probe_script</code> Validation Script” on page 15
3. Install the Sun StorEdge Availability Suite core, remote mirror, and point-in-time copy software on a cluster node.	Sun StorEdge Availability Suite installation guides listed in “Related Documentation” on page xii .
	“Supported Software and Hardware” on page 3 .
4. Edit the <code>/usr/kernel/drv/rdc.conf</code> or <code>/usr/kernel/drv/ii.conf</code> files, if necessary.	“Editing the Sun StorEdge Availability Suite Bitmap Parameter Files” on page 21
5. Shut down and restart the node.	“Shutting Down and Restarting the Nodes” on page 24
6. Repeat Step 3 through Step 5 for each additional cluster node.	
7. Configure the Sun Cluster software for use with the Sun StorEdge Availability Suite software.	“Supported Configurations for The Remote Mirror Software” on page 25
	“Configuration Rules for the Point-in-time Copy and Remote Mirror Software” on page 27
	“Configuring the Sun Cluster 3.0 Environment for Use with the Sun StorEdge Availability Suite Software” on page 29

Disk Device Groups and the Sun StorEdge Availability Suite Software

The Solstice Disk Suite (SDS) and VERITAS Volume Manager (VxVM) can arrange disk devices into a group to be mastered by a cluster node. You can then configure these disk device groups to fail over to another cluster node, as described in [“Configuring the Sun Cluster 3.0 Environment for Use with the Sun StorEdge Availability Suite Software” on page 29.](#)

The SDS and VxVM device paths contain the disk device group. When operating in a Sun Cluster 3.0 environment, the Sun StorEdge Availability Suite commands `sndradm` and `iiadm` automatically detect and use the disk device group as configured in [“Configuring the Sun Cluster 3.0 Environment for Use with the Sun StorEdge Availability Suite Software” on page 29.](#)

You can also use the `sndradm` and `iiadm` commands to select specified disk device groups or to operate on a volume set as a local node-only configuration entry. See [“The `iiadm` and `sndradm` Command Syntax” on page 41.](#)

Requirements for the Configuration Location



Caution – When selecting a volume to be used as the configuration location, ensure that volume does not contain disk label private areas (for example, slice 2 on a Solaris operating environment-formatted volume). The disk label region is contained in the first sectors of cylinder 0 of a disk.

When you install the Sun StorEdge Availability Suite software on the first cluster node, the installation process asks you to specify a block device for the single configuration location used by all Sun StorEdge Availability Suite software you plan to install.

In a Sun Cluster 3.0 environment, the Sun StorEdge Availability Suite configuration location must be available to all nodes running the Sun StorEdge Availability Suite software.

See TABLE 2-2 for the requirements for this configuration location.

The `scdidadm -L` command shows the local and shared disks by device ID.

TABLE 2-2 Configuration Location Requirements and Considerations

Item	Requirement or Consideration
Cluster environment	<p>If you are installing the software in a cluster environment, your configuration location must be a block device and it must exist in the directory <code>/dev/did</code>.</p> <p>The location:</p> <ul style="list-style-type: none"> • Must be on a shared disk; it cannot be on the quorum disk or any disk that might be locked during the boot process. It cannot be on a global or shared file system. • Cannot be stored on the same disk as the replicated volumes. • The shared disk volume cannot be under volume manager control, like a disk device group that is mastered to a particular node. • The shared disk volume must be under device ID (DID) control only. <p>An optional remote mirror volume set file (specified by the <code>sndradm -f volset-file</code> command) is not the same as a configuration location. A configuration location contains information about <i>all</i> devices used by the Sun StorEdge Availability Suite software.</p>
Availability	<ul style="list-style-type: none"> • If the location is a block device, it cannot be the same location as the current boot device. • The location must be writable by the superuser. • The location is available or persistent at system startup and reboots. • The location does not exist on an invalid file system type such as <code>cachefs</code>, <code>tmpfs</code>, <code>nfs</code>, <code>procfs</code>, <code>hsfs</code>, <code>autofs</code>, <code>fdfs</code>, and <code>mntfs</code>. • The location does not exist on a reserved mount point such as <code>/cdrom</code>, <code>/tmp</code>, <code>/proc</code>, <code>/mnt</code>, <code>/net</code>, <code>/floppy</code>, and <code>/vol</code>.
Disk space	The configuration location requires 5.5 Mbytes of disk space.

Running the probe_script Validation Script

Run the `probe_script` validation script before you install the version 3.1 software. The script does the following:

- Verifies that you are logged in as the superuser (root) user
- Checks that you have the correct minimum required version of the Solaris OE installed
- Lists any installed packages that you must remove and the order in which to remove them. Use the `pkgrm(1M)` program to remove these packages.

Syntax

The `probe_script` file on the product CD has the following syntax.

Note – Run the script with no options where the root installation path is the standard root slice (/).

`probe_script [-h | -j]`

where:

-h	Displays syntax usage.
-j	Uses standard script checking but also checks systems where the root installation path is a path other than the standard root slice (/). For example, use this option when root is located on a remotely-mounted device or when older packages might be located on a remotely-mounted device.

▼ To Run the probe_script File

1. Log on as the superuser user.
2. Insert the CD into the CD-ROM drive that is connected to your system.
3. Start the Volume Manager daemon `vold(1M)` (if needed) and perform one of the following:
 - Where the package installation path is the normal root slice (/):

```
# /etc/init.d/volmgt start
# cd /cdrom/cdrom0
# ./probe_script
```

- Where the package installation root path is located on a remotely-mounted device or when older packages might be located on a remote-mounted device:

```
# /etc/init.d/volmgt start
# cd /cdrom/cdrom0
# ./probe_script -j
```

If you typed `probe_script -j`, the script prompts you as follows. Otherwise, skip to [Step 5](#).

```
What is the root_path for this package installation? [ / ]
```

4. If you typed `probe_script -j`, perform one of the following:
 - Press Return to accept the default root path (/).
 - Type the full path where the root slice is mounted.

For either option in [Step 3](#), the script begins checking your system.

- **If you are not the superuser user or are not running the minimum required Solaris OE version**, the script displays messages stating:

```
WARNING : You're currently not the root user
You must be root when you execute the installation scripts.
WARNING: The version of Solaris currently running is not a supported version for
this installation
Supported versions include: 5.6, 5.7, 5.8, 5.9
Exiting...
```

- **If the script detects that previous versions of the Sun SNDR or Instant Image software are currently installed on your system**, note the packages to remove as listed by the script. For example:

```
WARNING:
Pre-existing Sun StorEdge software packages that must be uninstalled:
SUNWii SUNWrdu SUNWrdr SUNWspvu SUNWspvr SUNWscmu SUNWscmr

Installation cannot continue unless these packages are removed.

Please use pkgrm to uninstall these packages in the order they appear.
This list can be found in /tmp/pkgrm1ist.04_25_02_14:44:06
```

After the script executes successfully, the system displays a ready-to-install message and exits.

5. To install the software, go to [“Installing the Software” on page 18](#).

Installing the Software

Install the remote mirror software on the primary and secondary host machines. This process also installs the Sun StorEdge Availability Suite core and point-in-time copy software.

Note – Install the software on the primary hosts first.

Installation Script Syntax

The `install.sh` installation script on the product CD has the following syntax. You can install all Sun StorEdge software or just individual packages.

Note – Each option also installs the core software, required for all products. The script checks to see if the core software is already installed. If it is not, it is installed.

```
install.sh [-j] {-a | -p | -r}
```

where:

-j	Install the packages where the root installation path is a path other than the standard root slice (/). For example, use this option when root is located on a remotely-mounted device and you want to install the packages on a remotely-mounted device.
-a	Install the core, remote mirror, and point-in-time copy software located on the CD.
-p	Install the core and point-in-time software.
-r	Install the core and remote mirror software.

▼ To Install the Software

Note – See [“Installation Script Syntax”](#) on page 18.

1. Log on as superuser.

- Install the software on the primary host first.
- For a new installation, you can install this software in single-user or multiuser state.
- For an upgrade installation, install this software in single-user mode to avoid volume data corruption.

2. Insert the CD into the CD-ROM drive that is connected to your system.

It might already be in your drive if you ran the [probe_script](#) script.

3. Start the Volume Manager daemon `vold(1M)` (if needed).

```
# /etc/init.d/volmgt start
```

Note – You only need to start the Volume Manager daemon once. Do not start the daemon again.

4. Install the core, point-in-time copy, and remote mirror software.

```
# cd /cdrom/cdrom0
# ./install.sh_-a
```

The script prompts you as follows:

```
Attention!
By continuing with this installation, you acknowledge you have read installation
documentation and the "probe_script" has been run on this system
Do you want to continue [y,n,?]
```

5. Type Y to continue or N to cancel the installation.

- The core software package installation starts and displays the following message.

```
Enter database configuration location:
```

Note – The location size must be 5.5 Mbytes.

6. Type a block device for the single configuration location used by all Sun StorEdge software you plan to install.

For example: `/dev/dsk/clt1d0s7` or `/config`

When the core software installation finishes, the remote mirror software installation starts.

When the remote mirror and point-in-time software installation finishes, the `install.sh` script displays an “installation complete” message.

7. Eject the CD.

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

8. Perform any-post installation steps for the software as described in “[Editing the Sun StorEdge Availability Suite Bitmap Parameter Files](#)” on page 21 and the Sun StorEdge Availability Suite installation guides listed in “[Related Documentation](#)” on page xii.

Note – Ensure that you place the names and IP addresses of all machines you plan to use with the remote mirror software in the `/etc/hosts` file. Make sure you include the logical host names and IP addresses of the logical hosts you plan to use with the remote mirror software in the `/etc/hosts` file.

Edit this file on each machine where you are installing and running the remote mirror software.

9. Shut down and restart this node.

See “[Shutting Down Nodes After Installing the Sun StorEdge Availability Suite Software](#)” on page 10 and “[Shutting Down and Restarting the Nodes](#)” on page 24.

10. Log on as the root user at the next cluster node where you are installing the software and repeat these installation steps.

11. See “[Mounting and Replicating Global Volume File Systems](#)” on page 40 for information about global file systems.

Editing the Sun StorEdge Availability Suite Bitmap Parameter Files

Note – After editing the files described in this section and performing other post-installation configuration procedures, shutdown and restart your node as described in [“Shutting Down and Restarting the Nodes” on page 24](#).

Bitmap volumes are used by the remote mirror and point-in-time copy software to track differences between volumes and provide information for volume updates. The Sun StorEdge software documentation listed in [“Related Documentation” on page xii](#) describes the bitmap size and other requirements.

In a Sun Cluster 3.0 environment, a bitmap must reside only on a volume. The bitmap volume in this case must be part of the same disk device group or cluster resource group as the corresponding primary host or secondary hosts data volume.

The remote mirror and point-in-time copy software include two configuration files that determine how bitmap volumes are written to and saved:

- remote mirror - /usr/kernel/drv/rdc.conf
- point-in-time copy - /usr/kernel/drv/ii.conf



Caution – The Sun StorEdge Availability Suite 3.1 remote mirror and point-in-time copy software do not support bitmap files. The software uses regular raw devices to store bitmaps. These raw devices should be located on a disk separate from the disk that contains your data.

The /usr/kernel/drv/rdc.conf File

The /usr/kernel/drv/rdc.conf file contains one field that sets the remote mirror bitmap save mode:

- `rdc_bitmap_mode` - modify to change how the bitmap volume is saved during a shut down or system crash. In a Sun Cluster 3.0 environment, set this to 1.

A bitmap maintained on disk can persist across a system crash when this field is set to 1.

▼ To Edit the rdc.conf File

1. Open the /usr/kernel/drv/rdc.conf file using a text editor such as `vi(1)`.
2. In a Sun Cluster 3.0 environment, set the bitmap mode to 1.

For example:

```
# rdc_bitmap_mode
# - Sets the mode of the RDC bitmap operation, acceptable values are:
#   0 - autodetect bitmap mode depending on the state of SDBC (default).
#   1 - force bitmap writes for every write operation, so an update resync
#       can be performed after a crash or reboot.
#   2 - only write the bitmap on shutdown, so a full resync is
#       required after a crash, but an update resync is required after
#       a reboot.
#
rdc_bitmap_mode=1;
```

3. Save and exit the file.
4. Shut down and restart your server as described in [“Shutting Down and Restarting the Nodes”](#) on page 24.

The /usr/kernel/drv/ii.conf File

The /usr/kernel/drv/ii.conf file contains one setting that sets the point-in-time copy bitmap save mode:

- `ii_bitmap` - modify to change how the bitmap volume is saved during a shut down or system crash. In a Sun Cluster 3.0 environment, set this to 1.

A bitmap maintained on disk can persist across a system crash when this field is set to 1.

▼ To Edit the ii.conf File

1. Open the /usr/kernel/drv/ii.conf file using a text editor such as `vi`(1).
2. In a Sun Cluster 3.0 environment, set the bitmap mode to 1.

For example:

```
# bitmap volume storage strategy:
# 0 indicates kernel memory loaded from bitmap volume when shadow is resumed
#   and saved to bitmap volume when shadow is suspended.
# 1 indicates permanent SDBC storage, bitmap volume is updated directly as
#   bits are changed.
# 2 indicates that if FWC is present strategy 1 is used, otherwise strategy 0.
ii_bitmap=1;
```

3. Save and exit the file.
4. Shut down and restart your server as described in [“Shutting Down and Restarting the Nodes”](#) on page 24.

Shutting Down and Restarting the Nodes



Caution – Because the Sun StorEdge Availability Suite installation process requires you to shut down and restart each node in the cluster, *make sure that you install the Sun StorEdge Availability Suite software and any related patches during your normal maintenance window.*

As a result of this shutdown and restart, you might experience a panic condition on the node you are restarting. The node panic is expected behavior in the cluster and is part of the cluster software's *failfast mechanism*. The *Sun Cluster 3.0 Concepts* manual describes this mechanism and the Cluster Membership Monitor (CMM).

After performing the steps listed in [“Task Summary: Steps to Configure the Sun Cluster 3.0 Software and Install the Sun StorEdge Availability Suite Software”](#) on page 11, shutdown and restart each node.

Note – The `shutdown(1M)` command shuts down a single node or machine; the `scshutdown(1M)` command shuts down all nodes in a cluster. To shut down a single node, use the `scswitch(1M)` command as describes in the *Sun Cluster 3.0 System Administration Guide*.

▼ To Shut Down and Restart a Node

- Shut down and restart your node as follows:

```
# scswitch -S -h nodelist
# /etc/shutdown -y -g0 -i 6
```

-S

Evacuates all device and resource groups from the node.

-h *node1,node2*

Specifies the cluster nodes that can master this resource group. If you do not specify these nodes, it defaults to all the nodes in the cluster.

Supported Configurations for The Remote Mirror Software

Note – Place the names and IP addresses of all machines you plan to use with the remote mirror software in the `/etc/hosts` file. Make sure you include the logical host names and IP addresses of the logical hosts you plan to use with the remote mirror software in the `/etc/hosts` file. Edit this file on each machine where you are installing and running the remote mirror software.

Three configurations for the remote mirror software are supported:

- [“The Remote Mirror Primary Host is Configured On a Cluster Node” on page 25](#)
- [“The Remote Mirror Secondary Host is Configured On a Cluster Node” on page 26](#)
- [“The Remote Mirror Primary and Secondary Hosts are Configured On a Cluster Node” on page 26](#)

The Remote Mirror Primary Host is Configured On a Cluster Node

In this configuration, the remote mirror primary host is the logical host you created in the remote mirror resource group for the remote mirror disk group using the `scrgadm` command; for example, see [Step 6 in “To Configure the Sun Cluster Environment for HAStorage” on page 30](#).

If you have configured the remote mirror autosynchronization feature on the primary host, the remote mirror software starts an update resynchronization from the primary host for all affected remote mirror volume sets following a switchover or failover event (if the autosynchronization feature is enabled for those volume sets). This operation is performed after the resource group and network switchover operation is complete. See the `sndradm` man page and the *Sun StorEdge Availability Suite 3.1 Remote Mirror Software Administration and Operations Guide* for a description of the `sndradm -a` command to set the autosynchronization feature.

The Remote Mirror Secondary Host is Configured On a Cluster Node

In this configuration, the remote mirror secondary host is the logical host you created in the remote mirror resource group for the remote mirror disk group using the `scrgadm` command; for example, see [Step 6 in “To Configure the Sun Cluster Environment for HAStorage” on page 30](#).

Operations such as update resynchronizations occur and are issued from the primary host machine. Following a switchover (or failover) event, the remote mirror software attempts to start an update resynchronization for all affected remote mirror volume sets (if the autosynchronization feature is enabled for those volume sets). However, the remote mirror secondary host in a remote mirror volume set cannot initiate an update resynchronization.

This operation is performed after the resource group and network switchover operation is complete. In this case, the remote mirror secondary host switchover appears to be a short network outage to the remote mirror primary host.

If you have configured the remote mirror autosynchronization feature on the primary host, the `sndrsyncd` synchronization daemon attempts to resynchronize the volume sets if the system reboots or link failures occur. See the `sndradm` man page and the *Sun StorEdge Availability Suite 3.1 Remote Mirror Software Administration and Operations Guide* for a description of the `sndradm -a` command to set the autosynchronization feature.

If this feature is disabled (its default setting) and volume sets are logging but not replicating, perform the updates manually using the `sndradm` command.

The Remote Mirror Primary and Secondary Hosts are Configured On a Cluster Node

Remote mirror replication within the cluster is not supported; that is, when the primary and secondary hosts reside in the same cluster and the primary, secondary, and bitmap volumes in a volume set reside in the same disk device group. See [“Configuration Rules for the Point-in-time Copy and Remote Mirror Software” on page 27](#).

However, if the remote mirror primary and secondary hosts are configured in different clusters, see [“The Remote Mirror Primary Host is Configured On a Cluster Node” on page 25](#) and [“The Remote Mirror Secondary Host is Configured On a Cluster Node” on page 26](#) for operating considerations.

Supported Configurations for the Point-in-time Copy Software

See [“Configuration Rules for the Point-in-time Copy and Remote Mirror Software” on page 27.](#)

Configuration Rules for the Point-in-time Copy and Remote Mirror Software



Caution – Make sure that only one system administrator or root user at a time is creating and configuring volume sets to help avoid corrupting the Sun StorEdge Availability Suite configuration. See [“Configuring Remote Mirror and Point-in-time Copy Volume Sets” on page 38.](#)

The information in this guide, including this section, only applies to the remote mirror and point-in-time copy Version 3.1 software installed in a Sun Cluster 3.0 environment.

Rules For the Point-in-time Copy Software

- **All point-in-time copy volume set components must reside in the same disk device group.** (A point-in-time copy volume set includes the master, shadow, bitmap, and optional overflow volumes.)
- **With the point-in-time copy software, you can use more than one disk device group for cluster switchover and failover,** but each component in the volume set must reside in the same disk device group.

For example, you cannot have a master volume with a disk device group name of `ii-group` and a shadow volume with a disk device group name of `ii-group2` in the same volume set.

- **If a Solaris operating environment failure or Sun Cluster failover occurs during a point-in-time copy or update operation to the master volume**, specifically where the shadow volume is copying (`iiadm -c m`) or updating (`iiadm -u m`) data to the master volume, the master volume might be in an inconsistent state (that is, the copy or update operation might be incomplete). [“Preserving Point-in-time Copy Volume Data During a System Failure or Failover”](#) on page 50 describes how to avoid this situation.

Rules For the Remote Mirror Software

- **The primary volume and its bitmap volume or the secondary volume and its bitmap volume in a remote mirror volume set must reside in the same disk device group per node.** (A remote mirror volume set also includes information about primary and secondary hosts and operating mode.)
For example, you cannot have a primary volume with a disk device group name of `sndrdg` and a primary bitmap volume with a disk device group name of `sndrdg2` in the same remote mirror volume set.
- **With the remote mirror software, you can use more than one disk device group for cluster switchover and failover**, but each primary or secondary disk device component in the cluster node’s volume set must reside in the same disk device group.
- **The remote mirror software also requires a resource group containing the disk device group and logical failover host.** The disk device group is used to create a lightweight resource group containing the disk and a logical failover host.
The remote mirror software requires that the `SUNW.HAStorage` resource is configured in the same resource group as the logical host, as described in the procedures in [“To Configure the Sun Cluster Environment for HAStorage”](#) on page 30.
The resource group name you specify consists of the disk device group name appended with `-stor-rg`. For example, if the group name is `sndrdg`, then the resource group name would be:

`sndrdg-stor-rg`
- **Remote mirror replication within the cluster is not supported.** An example is when the primary host is cluster node 1 and the secondary host is cluster node 2 in the cluster and the primary, secondary, and bitmap volumes in a volume set reside in the same disk device group.
Typically, the remote mirror primary host is part of one cluster configuration, while the replicating secondary host might or might not be part of a different cluster.

Configuring the Sun Cluster 3.0 Environment for Use with the Sun StorEdge Availability Suite Software

Note – The Sun StorEdge Availability Suite software is supported *only* in a two-node Sun Cluster 3.0 Update 1 (or later) Runtime Environment.

The procedures in this section describe how to configure the Sun Cluster software for use with the remote mirror and point-in-time copy software. The *Sun Cluster 3.0 Data Installation and Configuration Guide* contains more information about configuring and administering Sun Cluster data services. See the `scrgadm(1M)` and `scswitch(1M)` man pages for more information.

Note – HASToragePlus is available to users of Sun StorEdge Availability Suite 3.2 software. HASToragePlus is available to users of Sun StorEdge Availability Suite 3.1 software only after the installation of patches #113054-05 and #113057-05 to the version 3.1 software.

Note – HASToragePlus is only available in the Sun Cluster 3.0 Update 2 (or later) Runtime Environment.

The general configuration steps are:

TABLE 2-3

-
1. Log on to any node in the cluster.
 2. Configure a disk device group using your volume manager.
 3. Register the `SUNW.HASTorage` or `SUNW.HASToragePlus` resource type.
 4. Create a resource group.
 5. Add `SUNW.HASTorage` or `SUNW.HASToragePlus` to the disk device group.
 6. (remote mirror step only) Add a logical failover host to the resource group.
 7. Enable the resource group and bring it online.
-

- To configure the software for use with HASTorage, see [“To Configure the Sun Cluster Environment for HASTorage” on page 30](#).

- To configure the software for use with HAStoragePlus, see [“To Configure the Sun Cluster Environment for HAStoragePlus” on page 33.](#)

When you complete the selected procedure, the resource group is configured and ready to use.

▼ To Configure the Sun Cluster Environment for HAStorage



Caution – You must adhere to the naming conventions and configuration rules specified in this procedure. Failure to do so results in an unsupported configuration and may lead to cluster hangs and panics.

1. Log on as the root user on any node in the cluster.

2. Configure a disk device group using your volume manager software.

See the documentation that came with your volume manager software. Also you might check the currently configured groups before configuring a new disk device group. For example, use the `metaset(1M)`, `vxdg`, or `vxprint` commands, depending on your volume manager software.

3. Register `SUNW.HAStorage` as a resource type.

```
# scrgadm -a -t SUNW.HAStorage
```

4. Create a resource group for the *devicegroup*.

```
# scrgadm -a -g devicegroup-stor-rg -h node1,node2
```

TABLE 2-4

<i>devicegroup</i>	is the required disk device group name.
<code>-h node1,node2</code>	specifies the cluster nodes that can master this resource group. If you do not specify these nodes, it defaults to all the nodes in the cluster.



Caution – Do not add additional resources, other than HAStorage and a logical host, to this lightweight resource group. Failure to follow this rule might cause the Sun StorEdge Availability Suite software to not properly failover or switchover.

5. Add a SUNW.HAStorage resource to the resource group.

```
# scrgadm -a -j devicegroup-stor -g devicegroup-stor-rg \  
-t SUNW.HAStorage \  
-x ServicePaths=devicegroup -x AffinityOn=True
```

TABLE 2-5

<i>devicegroup</i>	is the disk device group name.
-x ServicePaths=	specifies the extension property that the Sun StorEdge Availability Suite software relies on. In this case, use the disk device <i>devicegroup</i> .
-x AffinityOn=True	specifies that the SUNW.HAStorage resource needs to perform an affinity switchover for the global devices and cluster file systems defined in -x ServicePaths.

It also enforces co-location of resource groups and disk device groups on the same node, thus enhancing the performance of disk-intensive data services.

If the device group is switched to another node while the SUNW.HAstorage resource is online, AffinityOn has no effect and the resource group does not migrate along with the device group. On the other hand, if the resource group is switched to another node, AffinityOn being set to True causes the device group to follow the resource group to the new node.

6. Add a logical hostname resource to the resource group.

Note – Perform this step for the remote mirror volumes only. This step is not needed for point-in-time copy volumes.

```
# scrgadm -a -L [-j lhost-stor] -g devicegroup-stor-rg \  
-l lhost1,lhost2,...lhostN \  
-n nafo0@node,nafo0@node
```

TABLE 2-6

<code>-j lhost-stor</code>	is the optional resource <i>lhost-stor</i> . If you do not specify this option and resource, the name defaults to the first logical hostname specified in the <code>-l</code> option.
<code>-l lhost1,lhost2,...lhostN</code>	specifies a comma-separated list of UNIX hostnames (logical hostnames) by which clients communicate with the Sun StorEdge Availability Suite software in the resource group.
<code>-n nafo0@node,nafo0@node</code>	specifies the comma-separated list of Network Adapter Failover (NAFO) groups on each node.
	<i>node</i> can be a node name or ID. You can display the node ID using the <code>scconf -p</code> command.

7. Enable the resources in the resource group, manage the resource group, and bring the resource group online.

```
# scswitch -Z -g devicegroup-stor-rg
```

8. Verify that the resource is online.

a. Run the following command on any cluster node.

```
# scstat -g
```

b. Look for the resource group state field to determine if the resource group is online on the nodes specified in the node list.

▼ To Configure the Sun Cluster Environment for HAStoragePlus



Caution – You must adhere to the naming conventions and configuration rules specified in this procedure. Failure to do so results in an unsupported configuration and may lead to cluster hangs and panics.

1. Log on as the root user on any node in the cluster.

2. Configure a disk device group using your volume manager software.

See the documentation that came with your volume manager software. Also you might check the currently configured groups before configuring a new disk device group. For example, use the `metaset(1M)`, `vxdg`, or `vxprint` commands, depending on your volume manager software.

3. Register `SUNW.HAStoragePlus` as a resource type.

```
# scrgadm -a -t SUNW.HAStoragePlus
```

4. Create a resource group for the *devicegroup*.

```
# scrgadm -a -g devicegroup-stor-rg -h node1,node2
```

TABLE 2-7

<i>devicegroup</i>	is the required disk device group name.
<code>-h node1,node2</code>	specifies the cluster nodes that can master this resource group. If you do not specify these nodes, it defaults to all the nodes in the cluster.



Caution – Do not add additional resources, other than HAStoragePlus and a logical host, to this lightweight resource group. Failure to follow this rule might cause the Sun StorEdge Availability Suite software to not properly failover or switchover.

5. Add a SUNW.HAStoragePlus resource to the resource group.

```
# scrgadm -a -j devicegroup-stor -g devicegroup-stor-rg \  
-t SUNW.HAStoragePlus \  
-x GlobalDevicePaths=devicegroup -x AffinityOn=True
```

TABLE 2-8

<i>devicegroup</i>	is the disk device group name.
-x GlobalDevicePaths=	specifies the extension property that the Sun StorEdge Availability Suite software relies on. In this case, use the disk device <i>devicegroup</i> .
-x AffinityOn=True	specifies that the SUNW.HAStoragePlus resource needs to perform an affinity switchover for the global devices and cluster file systems defined in -x GlobalDevicePaths.

It also enforces co-location of resource groups and disk device groups on the same node, thus enhancing the performance of disk-intensive data services.

If the device group is switched to another node while the SUNW.HAStoragePlus resource is online, AffinityOn has no effect and the resource group does not migrate along with the device group. On the other hand, if the resource group is switched to another node, AffinityOn being set to True causes the device group to follow the resource group to the new node.

6. Add a logical hostname resource to the resource group.

Note – Perform this step for the remote mirror volumes only. This step is not needed for point-in-time copy volumes.

```
# scrgadm -a -L [-j lhost-stor] -g devicegroup-stor-rg \  
-l lhost1,lhost2,...lhostN \  
-n nafo0@node,nafo0@node
```

TABLE 2-9

<code>-j lhost-stor</code>	is the optional resource <i>lhost-stor</i> . If you do not specify this option and resource, the name defaults to the first logical hostname specified in the <code>-l</code> option.
<code>-l lhost1,lhost2,...lhostN</code>	specifies a comma-separated list of UNIX hostnames (logical hostnames) by which clients communicate with the Sun StorEdge Availability Suite software in the resource group.
<code>-n nafo0@node,nafo0@node</code>	specifies the comma-separated list of Network Adapter Failover (NAFO) groups on each node.
	<i>node</i> can be a node name or ID. You can display the node ID using the <code>scconf -p</code> command.

7. Enable the resources in the resource group, manage the resource group, and bring the resource group online.

```
# scswitch -Z -g devicegroup-stor-rg
```

8. Verify that the resource is online.

a. Run the following command on any cluster node.

```
# scstat -g
```

b. Look for the resource group state field to determine if the resource group is online on the nodes specified in the node list.

c.

Using the Sun StorEdge Availability Suite `iiadm` and `sndradm` Commands

This chapter describes using the Sun StorEdge Availability Suite commands `iiadm` and `sndradm` in a Sun Cluster 3.0 environment. The topics in this chapter include:

- “Configuring Remote Mirror and Point-in-time Copy Volume Sets” on page 38
- “The `iiadm` and `sndradm` Command Syntax” on page 41
- “Putting All Cluster Volume Sets in an I/O Group” on page 48
- “Preserving Point-in-time Copy Volume Data During a System Failure or Failover” on page 50
- “A Directory Named `._` is Created After Mounting a Secondary Volume” on page 51

Configuring Remote Mirror and Point-in-time Copy Volume Sets

The remote mirror and point-in-time copy software administration and operations guides listed in [“Related Documentation” on page xii](#) describe how to configure volume sets.



Caution – In a clustered environment, only one system administrator or root user at a time is allowed to create and configure Sun StorEdge volume sets. This restriction helps avoid creating an inconsistent Sun StorEdge Availability Suite volume set configuration.

Two administrators should not be writing to the Sun StorEdge Availability Suite software configuration at the same time. The operations that access the configuration include but are not limited to:

- Creating and deleting volume sets
- Adding and removing volume sets from I/O groups
- Assigning new bitmap volumes to a volume set
- Updating the disk device group or resource name
- Any operation that changes the Sun StorEdge Availability Suite software and related volume set configuration

Using the Remote Mirror Software With the Point-in-Time Copy Software

Note – You must place the related remote mirror volume set in logging mode for the point-in-time copy software to successfully perform an enable, copy, update, or reset operation on a remote mirror volume. If the volume set is not in logging mode, the point-in-time copy operation fails and the remote mirror software reports that the operation is denied

To help ensure a high level of data integrity on both sites during normal operations or during data recovery, use the remote mirror software with the Sun StorEdge Availability Suite 3.1 point-in-time copy software.

Use the point-in-time copy software just before you perform a resynchronization, to help ensure that a consistent copy of data exists. If a failure occurs during the resynchronization, you have the point-in-time data copy to roll back to and you can start again when its safe to do so.

During the resynchronization process of updating the local and remote sites, the data on a secondary remote mirror volume is temporarily inconsistent with the primary volume. The secondary volume cannot be relied on for data recovery. Consistency is restored when the resynchronization is complete. To help ensure data integrity, use point-in-time copy software regularly to create a point-in-time copy of data at both sites.

Mounting and Replicating Global Volume File Systems

If a volume contains a file system and you wish to replicate the file system using the Sun StorEdge Availability Suite software, you must create and mount a related global file system on all cluster nodes. These steps ensure that file system is available to all nodes and hosts when you copy or update the volume sets.

Note – See the *Sun Cluster 3.0 U1 System Administration Guide* or *Sun Cluster 3.0 12/01 System Administration Guide* for information about administering cluster file systems, including creating and mounting global file systems. See also the `mount(1M)` and `mount_ufs(1M)` commands.

For example:

1. **Create the file systems on the appropriate diskset metadevices or disk group volumes.**

```
# newfs raw-disk-device
```

For example, using the VERITAS Volume Manager, you might specify *raw-disk-device* as `/dev/vx/rdisk/sndrdg/vol01`.

2. **On each node, create a mount point directory for the file system.**

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

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

3. **On each node, add an entry to the `/etc/vfstab` file for the mount point and use the `global` mount option.**
4. **On a cluster node, use `sccheck(1M)` to verify the mount points and other entries.**
5. **From any node in the cluster, mount the file system.**

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

6. **Verify that the file system is mounted using the `mount` command with no options.**

The `iiadm` and `sndradm` Command Syntax

The Sun StorEdge Availability Suite software can use volumes that are global or local devices.

- *Global devices* are Sun StorEdge Availability Suite or other volumes accessible from any cluster node and which fail over and switch back under the control of the Sun Cluster framework.
- *Local devices* are Sun StorEdge Availability Suite software volumes that are local to the individual node (host machine), not defined in a disk or resource group, and not managed within a cluster file system. Local devices do not fail over and switch back.

The Sun StorEdge Availability Suite administrator guides listed in “[Related Documentation](#)” on page xii describe the full command syntax and options for `iiadm` and `sndradm`.

Global Device Command Syntax

Note – During the initial enable of the remote mirror or point-in-time copy volume sets, you can optionally specify the global device disk group with the `-C tag` cluster option when you use the `iiadm` or `sndradm` commands. As this section shows, however, you do not have to use the `-C tag` cluster option. Also see “[The C tag and -C tag Options](#)” on page 45.

The Sun StorEdge Availability Suite software automatically derives the disk device group name from the volume path when you first enable volume sets. During this initial enable operation, the remote mirror and point-in-time copy software creates a configuration entry for each volume set. Part of the entry is the disk device group name for use in a cluster.

The remote mirror software shows this name as `C tag`, where `tag` is the disk device group name. point-in-time copy shows this name as `Cluster tag: tag`.

Remote Mirror Example

When you enable this remote mirror volume set where `host1` is a logical failover host name:

```
# sndradm -e host1 /dev/vx/rdsk/sndrdg/datavol /dev/vx/rdsk/sndrdg/datavolbm1 \  
host2 /dev/rdsk/clt3d0s0 /dev/rdsk/clt2d0s4 ip sync
```

the corresponding configuration information as shown by the `sndradm -i` command is:

```
# sndradm -i  
  
host1 /dev/vx/rdsk/sndrdg/datavol /dev/vx/rdsk/sndrdg/datavolbm1 \  
host2 /dev/rdsk/clt3d0s0 /dev/rdsk/clt2d0s4 ip sync \  
C sndrdg
```

The `C` portion of the entry shows a disk device group name `sndrdg`.

Point-in-time Copy Example

When you enable a point-in-time copy volume set on a cluster node (logical failover host):

```
# iiadm -e ind /dev/vx/rdsk/iidg/clt3d0s0 /dev/vx/rdsk/iidg/clt3d0s4 \  
/dev/vx/rdsk/iidg/clt2d0s5
```

the corresponding configuration as shown by `iiadm -i` command is:

```
# iiadm -i  
  
/dev/vx/rdsk/iidg/clt3s0d0: (master volume)  
/dev/vx/rdsk/iidg/clt3d0s4: (shadow volume)  
/dev/vx/rdsk/iidg/clt2d0s5: (bitmap volume)  
Cluster tag: iidg  
Independent copy  
Volume size: 208278  
Percent of bitmap set: 0
```

The `Cluster tag` entry shows the derived disk device group name `iidg`.

Local Device Command Syntax

Note – Enabling a local disk device group named `local` prevents you from configuring a cluster disk device group named `local`.

- When you enable a point-in-time copy volume set, use the `-C local` option to specify that the volume set's disk device group name is `local`:
- When you enable a remote mirror volume set, use the `C local` option as part of the *vol-set* volume set definition:

```
iiadm -C local -e {dep | ind} master shadow bitmap
```

```
sndradm -e vol-set
```

where *vol-set* is:

```
phost pdev pbitmap shost sdev sbitmap ip {sync | async} [g io-groupname][C local]
```

The `local` disk device group is local to the individual cluster node and is not defined in a cluster disk or resource group. Local devices do not fail over and switch back. This initial configuration is similar to using the Sun StorEdge Availability Suite software in a nonclustered environment.

When you enable a volume set with the `local` disk device group, its configuration entry includes the name of its host machine.



Caution – Volumes and bitmaps used in a local remote mirror volume set cannot reside in a shared disk device group or metaset.

Point-in-time Copy Example

When you enable this point-in-time copy volume set where `local` indicates a disk device group:

```
# iiadm -C local -e ind /dev/rdisk/clt90d0s5 /dev/rdisk/clt90d0s6 \  
/dev/rdisk/clt90d0s7
```

the corresponding configuration as shown by `iiadm -i` command is:

```
# iiadm -i  
  
/dev/rdisk/iidg/clt90d0s5: (master volume)  
/dev/rdisk/iidg/clt90d0s6: (shadow volume)  
/dev/rdisk/iidg/clt90d0s7: (bitmap volume)  
Cluster tag: localhost (local)  
Independent copy  
Volume size: 208278  
Percent of bitmap set: 0
```

where *localhost* is the local host name as returned by the `hostname(1)` command.

The corresponding configuration information as shown by the `dscfg -l` command is:

```
# dscfg -l | grep /dev/rdisk/clt3d0s0  
  
ii: /dev/rdisk/clt90d0s5 /dev/rdisk/clt90d0s6 /dev/rdisk/clt90d0s7 I - 1.localhost - -
```

The `C tag` and `-C tag` Options

`C tag` is displayed as part of a volume set's configuration information as shown in [“Global Device Command Syntax” on page 41](#).

Typically, the Sun StorEdge Availability Suite software derives the disk device group name from the volume path and does not require the `-C tag` option.

Use the `-C tag` option and `C tag` volume set option to execute the `iiadm` and `sndradm` commands on the enabled volume sets in the disk device group name `tag`, when the disk device group name is not indicated by the volume path. The commands are not executed on any other volume sets in your configuration; `-C tag` excludes those volume sets not contained in the `tag` disk device group from the specified operation.

You can use more than one disk device group for cluster switchover and failover; ensure that your configuration follows the [“Configuration Rules for the Point-in-time Copy and Remote Mirror Software” on page 27](#).

Example

This command makes a point-in-time copy volume set in the `iigrp2` disk device group wait for all copy or update operations to finish before you can issue other point-in-time copy commands.

```
# iiadm -w /dev/vx/rdisk/iigrp2/nfsvol-shadow -C iigrp2
```

Which Host Do I Issue Commands From?

The Sun StorEdge Availability Suite software requires that you issue the `iiadm` or `sndradm` commands from the node that is the current primary host for the disk device group that the command applies to.

In a clustered environment, you can issue the command from the node mastering the disk device group you specified in [Step 2](#) in “[To Configure the Sun Cluster Environment for HAStorage](#)” on page 30.

When you enable the remote mirror software for the first time, issue the `sndradm enable` command from the primary and secondary hosts. See TABLE 3-1.

TABLE 3-1 Which Host to Issue Remote Mirror Commands From

Task	Where Command Is Issued	Comments
Assign a new bitmap to a volume set.	Primary and secondary host	Perform this command first on the host where the new bitmap resides and is being assigned, and then perform it on the other host.
Disable the remote mirror software.	Primary or secondary host	You can disable on one host, leave the other host enabled, and then re-enable the disabled host. Perform this operation on both hosts if you are deleting a volume set.
Enable the remote mirror software.	Primary and secondary host	When enabling the remote mirror software for the first time, issue the command from both hosts.
Full forward or reverse synchronization (copy).	Primary host	Ensure that both hosts are enabled.
Forward or reverse synchronization (update).	Primary host	Ensure that both hosts are enabled.

TABLE 3-1 Which Host to Issue Remote Mirror Commands From *(Continued)*

Task	Where Command Is Issued	Comments
Log.	Primary host	Perform on the primary host only if a synchronization is in progress.
		Perform on the secondary host if the primary host failed.
	Primary or secondary host	Perform on either host if no synchronization is in progress.
Toggle the autosynchronization state.	Primary host	
Update an I/O group.	Primary and secondary hosts	

Putting All Cluster Volume Sets in an I/O Group

Note – Placing volume sets in an I/O group does not affect the cluster operations of all volume sets configured in disk device and resource groups.



Caution – Do not reverse synchronize the primary volume from more than one secondary volume or host at a time. You can group one-to-many sets that share a common primary volume into a single I/O group to forward synchronize all sets simultaneously instead of issuing a separate command for each set.

You cannot use this technique to reverse synchronize volume sets, however. In this case, you must issue a separate command for each set and reverse update the primary volume by using a specific secondary volume.

The remote mirror and point-in-time copy software enables you to assign volume sets to I/O groups. Instead of issuing one command for each volume set, you can:

- Assign specific volume sets to an I/O group
- Issue one command specifying the I/O group
- Perform operations on those volume sets only

Like the `-C tag` and `C tag` options, the I/O group name excludes all other enabled volume sets from operations you specify.

In a clustered environment, you can assign some or all volume sets in a specific disk device group to an I/O group when you enable each volume set. If you have created more than one disk device group, you can also assign volume sets from different disk device groups into one I/O group.

Example

1. Enable three point-in-time copy volume sets and place them in an I/O group named `cluster1`.

```
# iiadm -g cluster1 -e ind /dev/rdisk/iigrp2/clt3d0s0 \  
/dev/rdisk/iigrp2/clt3d0s4 /dev/rdisk/iigrp2/clt2d0s5  
  
# iiadm -g cluster1 -e dep /dev/rdisk/iigrp2/clt4d0s0 \  
/dev/rdisk/iigrp2/clt4d0s4 /dev/rdisk/iigrp2/clt3d0s5  
  
# iiadm -g cluster1 -e ind /dev/rdisk/iigrp2/clt5d0s0 \  
/dev/rdisk/iigrp2/clt5d0s4 /dev/rdisk/iigrp2/clt4d0s5
```

2. Wait for any disk write operations to complete before issuing another command.

```
# iiadm -g cluster1 -w
```

3. Allow your applications to write to the master volumes.
4. Update the shadow volumes.

```
# iiadm -g cluster1 -u s
```

Preserving Point-in-time Copy Volume Data During a System Failure or Failover

If a Solaris operating environment system failure or Sun Cluster failover occurs during a point-in-time copy or update operation to the master volume, specifically where the shadow volume is copying (`iiadm -c m`) or updating (`iiadm -u m`) data to the master volume, the master volume might be in an inconsistent state (that is, the copy or update operation might be incomplete).

To avoid or reduce the risk of inconsistent data if a system failover occurs during such a copy or update operation, perform the following before performing the shadow volume-to-master volume copy or update operation:

1. Create a second independent shadow volume copy of the master volume by issuing an `iiadm -e ind` command.

This operation results in a full shadow volume copy of the master volume data.

2. Ensure that all copy or update operations to this second shadow volume are finished by issuing a wait command (`iiadm -w shadowvol`) after issuing the `iiadm -e ind` command.

You can now perform the copy or update operation from the original shadow volume to the master volume. If a system failure or failover occurs during this operation, you at least have a known good copy of your original master volume data. When this operation is complete, you can keep the second shadow volume under point-in-time copy control or return it to your storage pool.

A Directory Named `._` is Created After Mounting a Secondary Volume

After you synchronize the primary and secondary remote mirror volumes, you might notice a directory named `._` if you perform a directory listing. This directory is created by the cluster file system. For example:

```
secondary_hostname# ls -a
.
..
._
.profile
bin
classes
[and so on]
```

You can ignore this directory or delete it. When you unmount the cluster file system, the directory disappears.

Index

A

a, 2

B

bitmap volume
 parameter files, 21
bitmap volumes
 ii.conf file, 23
 rdc.conf file, 22

C

-C local, 43
commands
 iiadm and sndradm, 37
configuration
 location, 13
 procedures, 29
 rules, 27
 supported for point-in-time copy software, 27
 supported for remote mirror software, 25
configuration location
 disk space required for, 14
 specifying, 20
configuring the Sun StorEdge software, 9, 29

D

data service

defined, 2

disk device groups, 12

F

Fast Write Cache, 7

G

global devices, 5
 command syntax, 41
grouping volume sets, 48

H

High Availability applications
 see data services, 2

I

I/O group
 grouping volume sets in a cluster, 48
ii.conf file, 23
installation
 remote mirror software, 19
 Sun StorEdge core services software, 19

L

local devices, 5
 command syntax, 43

P

point-in-time copy
 configuration rules, 27
 ii.conf file, 23
point-in-time copy software
 data in a system failover, 50
 I/O groups, 48
 iiadm command, 4, 41
 supported configurations, 27
 with the remote mirror software, 39
probe_script
 see validation script, 15

R

rdc.conf file, 22
reboot
 shutdown and restart node, 24
remote mirror software
 rdc.conf file, 22
Remote mirror software
 supported configurations, 25
remote mirror software
 configuration rules, 28
 I/O groups, 48
 sndradm command, 4, 41
requirements
 bitmap volumes, 21

S

shutdown and restart node, 24
software
 installing, 19
Sun Cluster
 configuration rules with Sun StorEdge
 software, 27
Sun StorEdge
 installation, 18

validation script, 15

Sun StorEdge software
 bitmap volumes, 21
 configuration
 location, 13
 rules, 27
 configuring, 9
 Fast Write Cache
 SUNWnvm, 7
 iiadm command, 5
 installing, 9
 reboot node, 24
 sndradm command, 5
SUNWnvm, 4
supported configurations, 25, 27
syntax
 global device command, 41
 iiadm and sndradm commands, 41
 local device command, 43

V

validation script, 15
volume sets
 I/O group, 48