

Data ONTAP® 7.1

Cluster Installation and Administration Guide

Network Appliance, Inc.
495 East Java Drive
Sunnyvale, CA 94089 USA
Telephone: +1 (408) 822-6000
Fax: +1 (408) 822-4501
Support telephone: +1 (888) 4-NETAPP
Documentation comments: doccomments@netapp.com

Information Web: <http://www.netapp.com>

Part number 210-01386_C0
December 2005

Copyright and trademark information

Copyright information

Copyright © 1994–2005 Network Appliance, Inc. All rights reserved. Printed in the U.S.A.

No part of this document covered by copyright may be reproduced in any form or by any means—graphic, electronic, or mechanical, including photocopying, recording, taping, or storage in an electronic retrieval system—without prior written permission of the copyright owner.

Portions of this product are derived from the Berkeley Net2 release and the 4.4-Lite-2 release, which are copyrighted and publicly distributed by The Regents of the University of California.

Copyright © 1980–1995 The Regents of the University of California. All rights reserved.

Portions of this product are derived from NetBSD, which is copyrighted by Carnegie Mellon University.

Copyright © 1994, 1995 Carnegie Mellon University. All rights reserved. Author Chris G. Demetriou.

Permission to use, copy, modify, and distribute this software and its documentation is hereby granted, provided that both the copyright notice and its permission notice appear in all copies of the software, derivative works or modified versions, and any portions thereof, and that both notices appear in supporting documentation.

CARNEGIE MELLON ALLOWS FREE USE OF THIS SOFTWARE IN ITS “AS IS” CONDITION. CARNEGIE MELLON DISCLAIMS ANY LIABILITY OF ANY KIND FOR ANY DAMAGES WHATSOEVER RESULTING FROM THE USE OF THIS SOFTWARE.

Software derived from copyrighted material of The Regents of the University of California and Carnegie Mellon University is subject to the following license and disclaimer:

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notices, this list of conditions, and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notices, this list of conditions, and the following disclaimer in the documentation and/or other materials provided with the distribution.
3. All advertising materials mentioning features or use of this software must display the following acknowledgment:

This product includes software developed by the University of California, Berkeley and its contributors.
4. Neither the name of the University nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE REGENTS AND CONTRIBUTORS “AS IS” AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE REGENTS OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER

IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

This software contains materials from third parties licensed to Network Appliance Inc. which is sublicensed, and not sold, and title to such material is not passed to the end user. All rights reserved by the licensors. You shall not sublicense or permit timesharing, rental, facility management or service bureau usage of the Software.

Portions developed by the Apache Software Foundation (<http://www.apache.org/>). Copyright © 1999 The Apache Software Foundation.

Portions Copyright © 1995–1998, Jean-loup Gailly and Mark Adler

Portions Copyright © 2001, Sitraka Inc.

Portions Copyright © 2001, iAnywhere Solutions

Portions Copyright © 2001, i-net software GmbH

Portions Copyright © 1995 University of Southern California. All rights reserved.

Redistribution and use in source and binary forms are permitted provided that the above copyright notice and this paragraph are duplicated in all such forms and that any documentation, advertising materials, and other materials related to such distribution and use acknowledge that the software was developed by the University of Southern California, Information Sciences Institute. The name of the University may not be used to endorse or promote products derived from this software without specific prior written permission.

Portions of this product are derived from version 2.4.11 of the libxml2 library, which is copyrighted by the World Wide Web Consortium.

Network Appliance modified the libxml2 software on December 6, 2001, to enable it to compile cleanly on Windows, Solaris, and Linux. The changes have been sent to the maintainers of libxml2. The unmodified libxml2 software can be downloaded from <http://www.xmlsoft.org/>.

Copyright © 1994–2002 World Wide Web Consortium, (Massachusetts Institute of Technology, Institut National de Recherche en Informatique et en Automatique, Keio University). All Rights Reserved. <http://www.w3.org/Consortium/Legal/>

Software derived from copyrighted material of the World Wide Web Consortium is subject to the following license and disclaimer:

Permission to use, copy, modify, and distribute this software and its documentation, with or without modification, for any purpose and without fee or royalty is hereby granted, provided that you include the following on ALL copies of the software and documentation or portions thereof, including modifications, that you make:

The full text of this NOTICE in a location viewable to users of the redistributed or derivative work.

Any pre-existing intellectual property disclaimers, notices, or terms and conditions. If none exist, a short notice of the following form (hypertext is preferred, text is permitted) should be used within the body of any redistributed or derivative code: "Copyright © [\$date-of-software] World Wide Web Consortium, (Massachusetts Institute of Technology, Institut National de Recherche en Informatique et en Automatique, Keio University). All Rights Reserved. <http://www.w3.org/Consortium/Legal/>.

Notice of any changes or modifications to the W3C files, including the date changes were made.

THIS SOFTWARE AND DOCUMENTATION IS PROVIDED "AS IS," AND COPYRIGHT HOLDERS MAKE NO REPRESENTATIONS OR WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO, WARRANTIES OF MERCHANTABILITY OR FITNESS

FOR ANY PARTICULAR PURPOSE OR THAT THE USE OF THE SOFTWARE OR DOCUMENTATION WILL NOT INFRINGE ANY THIRD PARTY PATENTS, COPYRIGHTS, TRADEMARKS OR OTHER RIGHTS.

COPYRIGHT HOLDERS WILL NOT BE LIABLE FOR ANY DIRECT, INDIRECT, SPECIAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF ANY USE OF THE SOFTWARE OR DOCUMENTATION.

The name and trademarks of copyright holders may NOT be used in advertising or publicity pertaining to the software without specific, written prior permission. Title to copyright in this software and any associated documentation will at all times remain with copyright holders.

Software derived from copyrighted material of Network Appliance, Inc. is subject to the following license and disclaimer:

Network Appliance reserves the right to change any products described herein at any time, and without notice. Network Appliance assumes no responsibility or liability arising from the use of products described herein, except as expressly agreed to in writing by Network Appliance. The use or purchase of this product does not convey a license under any patent rights, trademark rights, or any other intellectual property rights of Network Appliance.

The product described in this manual may be protected by one or more U.S. patents, foreign patents, or pending applications.

RESTRICTED RIGHTS LEGEND: Use, duplication, or disclosure by the government is subject to restrictions as set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.277-7103 (October 1988) and FAR 52-227-19 (June 1987).

Trademark information

NetApp, the Network Appliance logo, the bolt design, NetApp—the Network Appliance Company, DataFabric, Data ONTAP, FAServer, FilerView, MultiStore, NearStore, NetCache, SecureShare, SnapManager, SnapMirror, SnapMover, SnapRestore, SnapVault, SyncMirror, and WAFL are registered trademarks of Network Appliance, Inc. in the United States, and/or other countries. gFiler, Network Appliance, SnapCopy, Snapshot, and The Evolution of Storage are trademarks of Network Appliance, Inc. in the United States and/or other countries and registered trademarks in some other countries. ApplianceWatch, BareMetal, Camera-to-Viewer, ComplianceClock, ComplianceJournal, ContentDirector, ContentFabric, EdgeFiler, FlexClone, FlexVol, FPolicy, HyperSAN, InfoFabric, LockVault, Manage ONTAP, NOW, NOW NetApp on the Web, ONTAPI, RAID-DP, RoboCache, RoboFiler, SecureAdmin, Serving Data by Design, SharedStorage, Simulate ONTAP, Smart SAN, SnapCache, SnapDirector, SnapDrive, SnapFilter, SnapLock, SnapMigrator, SnapSuite, SnapValidator, SohoFiler, vFiler, VFM, Virtual File Manager, VPolicy, and Web Filer are trademarks of Network Appliance, Inc. in the United States and other countries. NetApp Availability Assurance and NetApp ProTech Expert are service marks of Network Appliance, Inc. in the United States. Spinnaker Networks, the Spinnaker Networks logo, SpinAccess, SpinCluster, SpinFS, SpinHA, SpinMove, and SpinServer are registered trademarks of Spinnaker Networks, LLC in the United States and/or other countries. SpinAV, SpinManager, SpinMirror, SpinRestore, SpinShot, and SpinStor are trademarks of Spinnaker Networks, LLC in the United States and/or other countries.

Apple is a registered trademark and QuickTime is a trademark of Apple Computer, Inc. in the United States and/or other countries. Microsoft is a registered trademark and Windows Media is a trademark of Microsoft Corporation in the United States and/or other countries. RealAudio, RealNetworks, RealPlayer, RealSystem, RealText, and RealVideo are registered trademarks and RealMedia, RealProxy, and SureStream are trademarks of RealNetworks, Inc. in the United States and/or other countries.

All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such.

Network Appliance is a licensee of the CompactFlash and CF Logo trademarks.

Network Appliance NetCache is certified RealSystem compatible.

Table of Contents

	Safety Information (Sicherheitshinweise)	xi
	Preface	xiii
Chapter 1	Understanding clusters	1
	Understanding standard clusters	3
	Understanding mirrored clusters	7
	Understanding nonswitched MetroClusters	9
	Understanding Fabric Attached MetroClusters.	12
	Cluster failover and single-point-of-failure analysis	17
Chapter 2	Installing a Standard or Mirrored Cluster	27
	Cluster installation types	28
	Required documentation, tools, and equipment	29
	Preparing your equipment	32
	Cabling a FAS3000 series cluster	35
	Cabling Channel A	36
	Cabling Channel B	40
	Cabling the cluster adapters	44
	Cabling a FAS900 series cluster	45
	Cabling Channel A	47
	Cabling Channel B	52
	Cabling the cluster adapters	56
	Cabling an F800 series cluster	57
	Cabling Channel A	58
	Cabling Channel B	61
	Cabling the cluster adapters	64
Chapter 3	Installing a MetroCluster	67
	Required documentation, tools, and equipment	68
	Cabling a new Fabric Attached MetroCluster	72
	Stage 1: Planning the Fabric Attached MetroCluster	73

Stage 2: Configuring the local switches	74
Stage 3: Cabling the local node	78
Stage 4: Cabling the switches and nodes at the partner site	84
Cabling a nonswitched MetroCluster.	91
Converting a cluster into a Fabric Attached MetroCluster	93

Chapter 4

Configuring a Cluster	95
Stage 1: Bringing up the cluster for the first time	96
Stage 2: Running diagnostics.	101
Stage 3: Enabling licenses	103
Stage 4: Setting options and parameters	105
Stage 5: Configuring network interfaces	110
Stage 6: Checking and testing cluster configuration	112

Chapter 5

Cluster Management	115
How takeover and giveback work	116
Managing a cluster in normal mode	118
Configuring when takeover occurs	125
Managing a cluster in takeover mode	129
Managing an emulated node	130
Backing up and restoring from the emulated node	135
Managing giveback.	136
Configuring giveback	137
Performing a giveback	139
Enabling automatic giveback.	142
Configuring network interfaces for clusters	143
Understanding interface configurations for takeover	144
Configuring shared interfaces	148
Configuring dedicated and standby interfaces	151
Checking cluster configuration	154
Correcting takeover or giveback failures	155
Removing clustering	157
Removing F800 series clustering hardware	158
Removing FAS900 and FAS3000 series clustering hardware	160

Chapter 6	Disaster Recovery Using MetroCluster	.163
	Recognizing a disaster	.164
	Recovering from a disaster	.166
 Chapter 7	 Modifying Cluster Storage	 .177
	Adding disk shelves to a cluster	.178
	Installing a disk shelf in an existing mixed-shelf loop	.182
	Hot-adding a DS14, DS14mk2 FC, or DS14mk2 AT to an existing loop . . .	.187
	Upgrading or replacing LRC, ESH, ESH2, or AT-FCX modules	.193
	Understanding single and dual pathing in clusters	.196
	Upgrading to ESH or ESH2 modules	.199
	Hot-swapping an LRC, ESH, ESH2, or AT-FCX module	.203
	Hot-swap disk loop error messages on Fabric Attached MetroClusters . . .	.206
 Chapter 8	 Changing Clustered Hardware Nondisruptively	 .207
 Appendix A	 Communications Regulations.	 .211
	Regulatory notices	.212
	Declaration of Conformity	.214
 Appendix B	 Feature Update Record	 .215
	 Index	 .217

Safety Information (Sicherheitshinweise)

Safety rules

All products are Class 1 laser devices, except the NVRAM5 cluster media converter, which is Class 1M. You must follow these safety rules when working with this equipment:

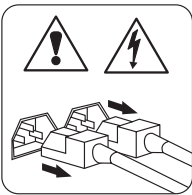
WARNING

Failure to follow these directions could result in bodily harm or death.

- ◆ When using an NVRAM5 cluster media converter, the storage system must be installed in a restricted access location.
- ◆ **Switzerland only—for FAS900, GF900, R200, and C6200 systems:** This equipment relies on fuses/circuit breakers in the building installation for overcurrent protection. Each power supply must receive power from a separately dedicated outlet with a 10A fuse/circuit breaker.
- ◆ When installing disk shelves and a storage system into a movable cabinet or rack, install from the bottom up for best stability.
- ◆ DC-based systems must be installed in a restricted access location and the two input power terminals for the DC power supply must be connected to separate isolated branch circuits.
- ◆ To reduce the risk of personal injury or equipment damage, allow internal components time to cool before touching them and ensure that the equipment is properly supported or braced when installing options.
- ◆ This equipment is designed for connection to a grounded outlet. The grounding type plug is an important safety feature. To avoid the risk of electrical shock or damage to the equipment, do not disable this feature.
- ◆ This equipment has one or more replaceable batteries. There is danger of explosion if the battery is incorrectly replaced. Replace the battery only with the same or equivalent type recommended by the manufacturer. Dispose of used batteries according to the manufacturer's instructions.

For units with multiple power cords

If your storage system or disk shelf has multiple power cords and you need to turn the unit off, heed the following warning:



WARNING

This unit has more than one power supply cord. To reduce the risk of electrical shock, disconnect all power supply cords before servicing.

Sicherheitsvorgaben

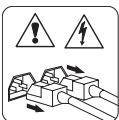
Alle Produkte sind Lasergeräte der Klasse 1, mit Ausnahme des NVRAM5 Cluster-Medienkonverters, der in Klasse 1M fällt. Beim Einsatz dieser Geräte sind die Sicherheitsvorschriften zu beachten:

Vorsicht

Nichtbeachtung dieser Vorschriften kann zu Verletzungen oder Tod führen.

- ◆ Bei der Verwendung eines NVRAM5 Cluster-Medienkonverters muss das Speichersystem an einem Standort mit beschränktem Zugriff installiert werden.
- ◆ **Nur für die Schweiz - Systeme FAS900, GF900, R200 und C6200:**
Diese Geräte erfordern den Festeinbau von Sicherungen zum Überstromschutz. Jeder Netzanschluss muss mit Strom aus getrennten, speziell für diesen Zweck vorgesehenen Steckdosen versorgt werden, die jeweils mit einer 10A-Sicherung geschützt sind.
- ◆ Werden die Plattenregale und das Speichersystem in einen beweglichen Schrank oder Turm eingebaut, ist wegen der höheren Stabilität der Einbau von unten nach oben vorzunehmen.
- ◆ Gleichstrom-Systeme müssen an Betriebsstätte mit beschränktem Zutritt installiert sein und die beiden Eingangsstromklemmen für das Gleichstrom-Netzteil müssen an separate und isolierte Abzweigungen angeschlossen sein.
- ◆ Zum Schutz vor Körperverletzung oder Sachschäden am Gerät lassen Sie die inneren Bauteile stets vor dem Berühren abkühlen. Sorgen Sie dafür, dass das Gerät richtig abgestützt ist oder fest aufrecht steht, bevor Sie neues Zubehör einbauen.
- ◆ Dieses Gerät ist für die Einspeisung aus einer geerdeten Netzverbindung ausgelegt. Der Netzstecker mit Erdungsvorrichtung ist ein wichtiger Sicherheitsschutz. Zum Schutz vor elektrischem Schlag oder Sachschäden am Gerät die Erdung nicht abschalten.
- ◆ Das Gerät ist mit einer oder mehreren auswechselbaren Batterien ausgestattet. Bei unsachgemäßem Auswechseln der Batterie besteht Explosionsgefahr. Batterien nur mit dem vom Hersteller empfohlenen Typ oder entsprechenden Typen ersetzen. Gebrauchte Batterien sind gemäß den Anweisungen des Herstellers zu entsorgen.

Für Geräte mit mehrfachen Netzan-schlussleitungen



Wenn Ihr Speichersystem oder Plattenregal über mehrere Stromkabel verfügt und Sie die Einheit ausschalten müssen, folgenden Warnhinweis beachten:

ACHTUNG

Gerät besitzt zwei Netzan-schlussleitungen. Vor Wartung alle Anschlüsse vom Netz trennen.

Preface

About this guide

This guide describes the following tasks and topics:

- ◆ Preparing to install an F800 series cluster, a FAS900 series cluster, or a FAS3000 series cluster
- ◆ Cabling a pair of F800 series, FAS900 series, or FAS3000 series storage systems into a cluster
- ◆ Configuring the cluster when you first bring it up
- ◆ Managing a cluster
- ◆ Performing disaster recovery with a MetroCluster
- ◆ Performing basic troubleshooting on the cluster

This guide does not cover basic system or network administration, which is covered in the *Data ONTAP™ System Administration Guide* and *Data ONTAP Network Administration Guide*.

Audience

This guide is for qualified system administrators and service personnel who are familiar with Network Appliance™ equipment and who need to

- ◆ Cable and configure two systems into a standard or mirrored cluster system
- ◆ Convert stand-alone systems into a standard or mirrored cluster system
- ◆ Cable and configure a new Fabric Attached or nonswitched MetroCluster
- ◆ Manage a cluster
- ◆ Troubleshoot a cluster

Terminology

This guide uses the following terms:

General storage system terminology:

- ◆ NetApp® storage products (filers, FAS storage systems, and NearStore® systems) are all storage systems. All these products are also sometimes called *filers*, *appliances*, or *storage appliances*.
- ◆ *FAS3000 series* refers to NetApp storage systems belonging to that series, such as the FAS3050c.
- ◆ *FAS900 series* refers to NetApp storage systems belonging to that series, such as the FAS980c.
- ◆ *F800 series* refers to NetApp storage systems belonging to that series, such as the F880c.

- ◆ *Storage system* refers to any F800 series, FAS900 series, or FAS3000 series storage system.

Cluster terminology:

- ◆ *Cluster* refers to two storage systems, or nodes, that are connected in such a way that they can take over operations for each other, either automatically or with operator intervention. One node is called the local node, and the other node is called the partner node.
- ◆ *Active-active controller configuration* is another term for cluster.
- ◆ *Standard cluster*, or *failover cluster*, refers to a cluster configured so that one node automatically takes over for its partner when the partner node becomes impaired.
- ◆ *Mirrored cluster* refers to a standard cluster where there are two copies, or plexes, of the data.
- ◆ *Fabric Attached MetroCluster* refers to a cluster running syncmirror_local and cluster_remote, where the nodes (storage systems) and disk shelves are attached to two pairs of Fibre Channel switches, where the nodes and disk shelves are not cabled together in standard cluster fashion, and where the nodes are separated by more than 500 meters and up to 30 kilometers.
- ◆ *Nonswitched MetroCluster* refers to a cluster running syncmirror_local and cluster_remote, where the nodes are separated without a switch by up to 500 meters.
- ◆ *Switched cluster* refers to a cluster that has the local and partner nodes connected to each other, then connected to clients through a switch.
- ◆ *Cluster failover*, also referred to as *controller failover* or *CFO*, refers to the technology that enables two storage systems to take over each other's data, thus improving data availability.
- ◆ *Node* refers to either the local or partner storage system.
- ◆ *Local cluster node* refers to the cluster node in the same area as the system administrator.
- ◆ *Remote cluster node* refers to the cluster node that is *not* in the same area as the system administrator.
- ◆ *Remote storage* refers to the storage that is accessible to the local cluster node, but is located at the location of the remote cluster node.

Clustering and storage hardware terminology:

- ◆ *FC HBA for Disk* or *FC HBA* refers to the Fibre Channel host bus adapter that connects the node to the switch or to the disks.
- ◆ *Disk shelf* refers to a Fibre Channel StorageShelf FC7, Fibre Channel StorageShelf FC8, Fibre Channel StorageShelf FC9, DiskShelf14, DS14mk2 FC, or DS14mk2 AT.

- ◆ The *LRC (Loop Resiliency Circuit)* module keeps the FC-AL loop intact during the addition and removal of disks within a disk shelf. It also contains the enclosure services processor, which communicates the environmental data of the disk shelf. The LRC reconditions the signal so that there is no accumulated error in the data signals. As the signal moves, waveforms might distort slightly. Over many hops these distortions would otherwise accumulate, causing high error rates. The LRC prevents this.
- ◆ The *ESH (Embedded Switching Hub)* module provides a means of managing an FC-AL loop in an intelligent manner, such that a single drive failure does not take down the loop. It also contains the enclosure services processor, which communicates the environmental data of the disk shelf.
- ◆ The *ESH2* module is a second-generation ESH module.
- ◆ *AT-FCX* is an enhanced FC-AL to Serial ATA bridge used in the DS14Mk2 AT disk shelf.

Command conventions

You can enter commands on the system console or from any client that can obtain access to the storage system using a Telnet session. In examples that illustrate commands executed on a UNIX® workstation, the command syntax and output might differ, depending on your version of UNIX.

Keyboard conventions

This guide uses capitalization and some abbreviations to refer to the keys on the keyboard. The keys on your keyboard might not be labeled exactly as they are in this guide.

What is in this guide...	What it means...
hyphen (-)	Used to separate individual keys. Example: Ctrl-D means holding down the Ctrl key while pressing the D key.
Enter	Used to refer to the key that generates a carriage return, although the key is named Return on some keyboards.
type	Used to mean pressing one or more keys on the keyboard.
enter	Used to mean pressing one or more keys and then pressing the Enter key.

Formatting conventions

The following table lists the kinds of formatting this guide uses to identify special information.

Formatting convention	Type of information
<i>Italic type</i>	◆ Words or characters that require special attention. ◆ Placeholders for information you must supply. For example, if the guide requires you to enter the arp -d hostname command, you enter the characters “arp -d” followed by the actual name of the host. ◆ Book titles in cross-references.
Monospaced font	◆ Command and daemon names. ◆ Information displayed on the system console or other computer monitors. ◆ Contents of files.
Bold monospaced font	Words or characters you type. What you type is always shown in lowercase letters, unless you must type it in uppercase letters for it to work properly.

Special messages

This guide contains special messages that are described as follows:

Note

A note contains important information that helps you install or operate the system efficiently.

Attention

An attention contains instructions that you must follow to avoid damage to the equipment, a system crash, or loss of data.

WARNING

A warning contains instructions that you must follow to avoid personal injury.

What a cluster is

A cluster is two storage systems (nodes) that are connected to each other either directly or through switches. You can configure the two nodes so that they share access to a single set of disks, subnets, and tape drives, or you can configure them to have two distinct sets of storage, each owned by one of the nodes.

Note

You may also see the term *active-active controller configuration*; this is equivalent to *cluster* as it is used in this document.

Although most storage systems have external controller units, such as the F800 series and the FAS900 series, there are some storage systems, such as the FAS200 series, that have internal head units. The FAS270c, for example, is a cluster consisting of two internal controller units that reside in and share the same disk shelf.

About this chapter

This chapter discusses the following four types of clusters:

- ◆ “[Understanding standard clusters](#)” on page 3
- ◆ “[Understanding mirrored clusters](#)” on page 7
- ◆ “[Understanding nonswitched MetroClusters](#)” on page 9
- ◆ “[Understanding Fabric Attached MetroClusters](#)” on page 12

In addition, this chapter discusses how standard and mirrored clusters help reduce single-points-of-failure (SPOFs). For more information, see “[Cluster failover and single-point-of-failure analysis](#)” on page 17.

Cluster type comparison

The following table outlines the differences between the different types of clusters, and when you might want to use each type.

Cluster type	Data duplication?	Distance between nodes	Failover possible after loss of entire node (including storage)?	Notes
Standard (failover) cluster	No	Up to 500 meters (see note)	No	Use this configuration to provide higher availability by protecting against many hardware single-points-of-failure.
Mirrored (failover) cluster	Yes	Up to 500 meters (see note)	No	Use this configuration to add increased data protection to the benefits of a standard cluster.
Nonswitched MetroCluster	Yes	Up to 500 meters (see note)	Yes	Use this configuration to provide data and hardware duplication to protect against a local disaster (for example, a power outage to one node).
Fabric Attached MetroCluster	Yes	Up to 30 kilometers	Yes	Use this configuration to provide data and hardware duplication to protect against a larger-scale disaster.

Note

DS14mk2 FC disk shelves using ESH or ESH2 modules, or DS14mk2 AT disk shelves using AT-FCX modules operating at 2-Gb Fibre Channel speeds can attain a maximum node separation distance of only 300 meters, not 500.

Understanding standard clusters

What standard clusters contain

A standard cluster contains two sets of Fibre Channel disk shelves, a local node, and a partner node (except the FAS270c, which does not require two sets of disk shelves). Both nodes must be the same model.

The nodes are connected to each other through an IB cluster adapter or NVRAM5 adapter, which allows one node to serve data to the disks of its failed partner node. Each node continually monitors its partner, mirroring the data for each other's NVRAM.

Benefits of clusters

Configuring storage systems in a cluster provides the following benefits:

- ◆ Fault tolerance

When one node fails or becomes impaired, a takeover occurs, and the partner node continues to serve the failed node's data.

- ◆ Nondisruptive software upgrades

When you halt one node and allow takeover, the partner node continues to serve data for the halted node while you upgrade the node you halted.

For more information about nondisruptive upgrades, see the *Upgrade Guide*.

- ◆ Nondisruptive storage system and disk maintenance

When you halt one node and allow takeover, the partner node continues to serve data for the halted node while you replace or repair hardware in the node you halted.

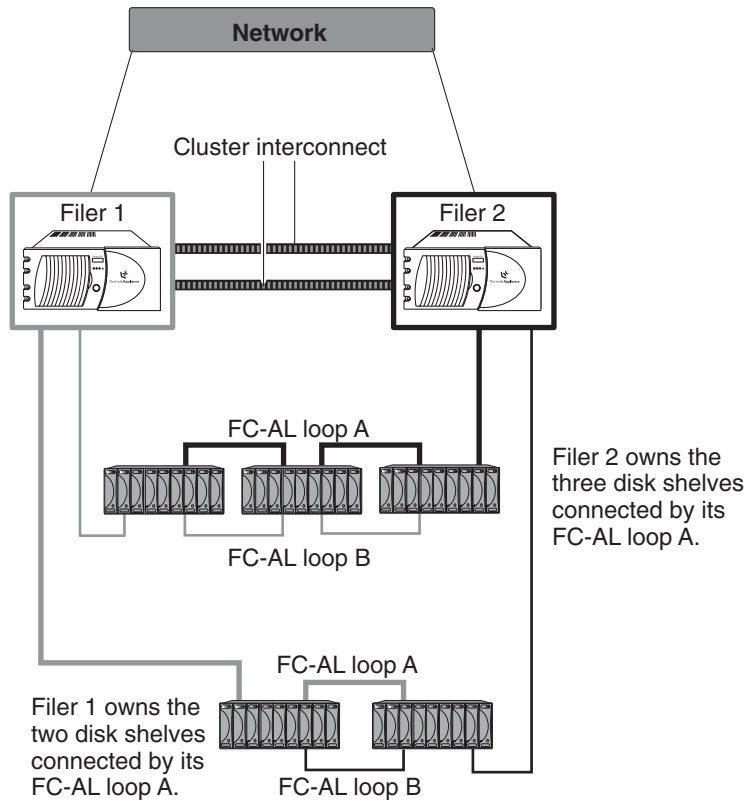
For more information about nondisruptive storage system maintenance, see “[Changing Clustered Hardware Nondisruptively](#)” on page 207.

How Data ONTAP works with standard clusters

In a standard cluster, Data ONTAP functions so that each node monitors the functioning of its partner through a heartbeat signal sent between the nodes. Data from the NVRAM of one node is mirrored by its partner, and each node can take over the partner's disks if the partner fails. Also, the nodes synchronize each other's time. For more information, see “[Cluster Management](#)” on page 115.

Cluster diagram

The following figure shows a common configuration of a standard cluster.



Note

The preceding diagram assumes that each node owns the disks attached to its own shelves, connected using the FC-AL A loop. If you are using a FAS270c (which contains two controller units) or the SnapMover® software, you have software disk ownership. In these cases, you need to establish disk ownership explicitly.

For more information about software disk ownership, see the Disk chapter of the *Storage Management Guide*. For more information about SnapMover, see the *MultiStore® Management Guide*.

Cluster characteristics

Storage systems in a cluster have the following characteristics:

- ◆ They are connected to each other through a cluster interconnect consisting of adapters and cable through which they do the following tasks:
 - ❖ Continually check whether the other node is functioning
 - ❖ Mirror log data for each other's NVRAM
 - ❖ Synchronize each other's time
- ◆ They use two or more FC-AL loops in which the following conditions apply:
 - ❖ Each node manages its own disks.
 - ❖ Each node in takeover mode manages its partner's disks.

Note

For the FAS 270c and systems using the SnapMover software, software disk ownership is in use, so disk ownership is established explicitly rather than by which shelf the disk is attached to.

For more information about software disk ownership, see the chapter on disk management in the *Storage Management Guide*. For more information about SnapMover, see the *Multistore Management Guide*.

- ◆ They own their spare disks and do not share them with the other node.
- ◆ They each have two mailbox disks on the root volume (four if the root volume is mirrored using the SyncMirror® feature). The mailbox disks are used to do the following tasks:
 - ❖ Maintain consistency between the pair
 - ❖ Continually check whether the other node is running or whether it has performed a takeover
 - ❖ Store cluster information that is not specific to any particular node
- ◆ They can reside on the same NT domain or on different domains.

Standard cluster setup requirements and restrictions

The following list specifies the requirements and restrictions you must be aware of when setting up a new cluster or converting an existing cluster:

- ◆ Architecture compatibility

Both nodes must be the same type and model.
- ◆ Storage system capacity

The number of disks must not exceed the maximum cluster capacity. To determine the maximum capacity, see the *System Configuration Guide* on the NetApp on the Web™ (NOW) site at Service & Support > Access Product Documentation > System Configuration Guide.

- ◆ Disks and disk shelf compatibility
The disk shelves must have two paths so that each node can attach to the shelf.
- ◆ Cluster interconnect adapters and cables must be installed.
- ◆ Nodes must be attached to the same network and the NICs must be configured correctly.
- ◆ Cluster software must be licensed and enabled on both nodes.
- ◆ The same system software, such as CIFS, NFS, or SyncMirror, must be licensed and enabled on both nodes.

Note

MultiStore software can be licensed on one node only; this does not prevent a successful takeover of the partner node.

Best practices when deploying clusters

You can do many things to ensure that clustered storage systems achieve maximum uptime:

- ◆ Make sure that the clustered controller units and shelves are on different power supplies/grids, so that a single power outage does not affect both controller units and shelves.
- ◆ Use VIFs (virtual interfaces) to provide redundancy and improve availability of network communication.
- ◆ Follow the documented procedures in the *Upgrade Guide* when upgrading clustered storage systems.
- ◆ Maintain consistent configuration between clustered nodes. An inconsistent clustered storage system configuration is often a culprit for failover problems.
- ◆ Test the failover capability routinely (for example, during planned maintenance) to ensure proper clustered storage system configuration.
- ◆ Make sure that clustered nodes have sufficient resources to adequately support workload during takeover mode.
- ◆ Use the Cluster Configuration Checker to help ensure that failovers are successful.

For more information, see “[Adding disk shelves to a cluster](#)” on page 178.

Understanding mirrored clusters

What mirrored clusters do

Mirrored clusters provide high availability through failover, just as standard clusters do. Additionally, mirrored clusters provide the capability to have two complete copies of all mirrored data. These copies are called plexes and are continually and synchronously updated every time Data ONTAP writes a mirrored aggregate. The plexes can be physically separated to protect against the loss of one set of disks.

Note

Mirrored clusters do *not* provide the capability to fail over to the partner node if one node is completely lost. For example, if power is lost to one entire node, including its storage, you cannot fail over to the other node. For this capability, use a MetroCluster.

Mirrored clusters use SyncMirror. For more information, see the *Data Protection Online Backup and Recovery Guide*.

The restrictions, requirements, and best practices for mirrored clusters are the same as for a standard cluster. For more information, see “[Failover event cause-and-effect table](#)” on page 20.

Advantages of mirrored clusters

Mirroring your data protects it from the following problems, both of which would cause data loss without mirroring:

- ◆ The failure or loss of two or more disks in a RAID4 aggregate
- ◆ The failure or loss of three or more disks in a RAID-DP™ aggregate

In addition, the failure of an FC-AL adapter, loop, or ESH module does not require a cluster failover in a mirrored cluster.

Similar to standard clusters, if either node in a mirrored cluster becomes impaired or cannot access its data, the other node can automatically serve the impaired node's data until the problem is corrected.

Setup requirements and restrictions

The restrictions and requirements for mirrored clusters include those for a standard cluster, as outlined in “[Failover event cause-and-effect table](#)” on page 20. In addition, the following list of restrictions applies when setting up a mirrored cluster:

- ◆ Disks in the same plex must be from the same pool, with disks in the opposite plex from the opposite pool.
- ◆ If software disk ownership isn’t used, you must observe the SyncMirror pool rules for direct-attached Fibre Channel disks.
For more information, see “[SyncMirror pool rules for direct-attached Fibre Channel disks](#)” on page 8.
- ◆ There must be sufficient spares in each pool to account for a disk failure.
- ◆ The following licenses must be enabled on the local and partner nodes:
 - ❖ cluster
 - ❖ syncmirror_local

SyncMirror pool rules for direct-attached Fibre Channel disks

The FAS3000 series storage system uses new slots for direct-attached Fibre Channel disks.

Note

This is valid only if software disk ownership is not being used.

Follow these SyncMirror pool rules for direct-attached Fibre Channel disks:

For FAS3000 series:

- ❖ pool 0 - Onboard adapters 0a, 0b, and adapters in PCI slots 1, 2
- ❖ pool 1 - Onboard adapters 0c, 0d, and adapters in PCI slots 3, 4

For all other storage systems:

- ❖ pool 0 - Adapters in PCI slots 1 - 7
- ❖ pool 1 - Adapters in PCI slots 8 - 11

Understanding nonswitched MetroClusters

What nonswitched MetroClusters do

Like mirrored clusters, nonswitched MetroClusters contain two complete copies of the specified data volumes or file systems that you indicated as being mirrored volumes or file systems in your cluster. These copies are called plexes and are continually and synchronously updated every time Data ONTAP writes data to the disks. Plexes are physically separated from each other across different groupings of disks.

Unlike mirrored clusters, MetroCluster provides the capability to force a failover when an entire node is destroyed or unavailable.

Configuration overview

The MetroCluster configuration employs SyncMirror and a cluster configured for mirrored volumes to build a system that provides data even after complete loss of one of the cluster nodes and the storage at that site. All mirrored volumes (both root volumes and data volumes) must be mirrored with one plex at each site to preserve data using MetroCluster.

Attention

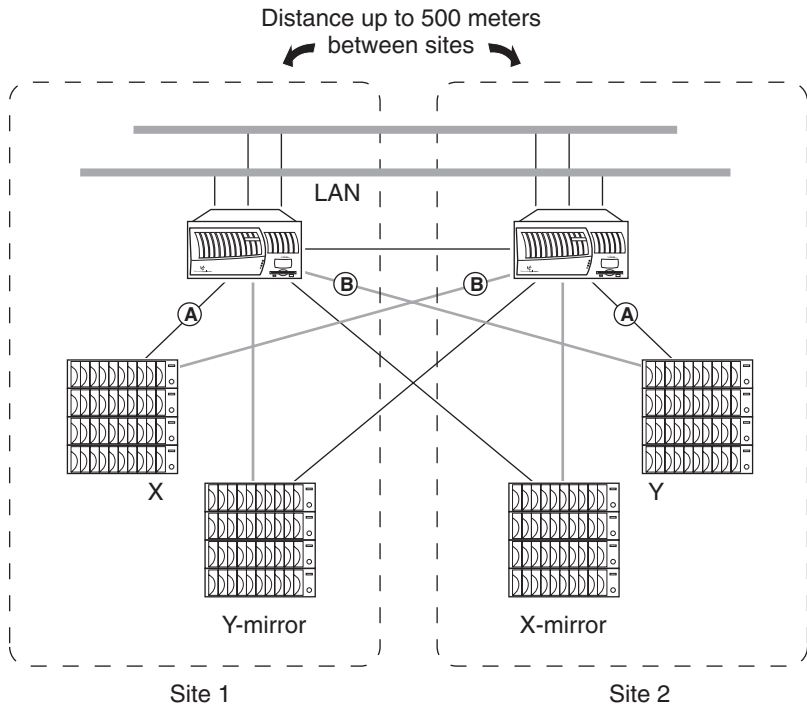
The MetroCluster configuration can preserve data only if volumes are mirrored. Unmirrored volumes are lost if the storage where they reside is destroyed.

See the *Data Protection Online Backup and Recovery Guide* for detailed information about using SyncMirror to mirror data.

The following figure illustrates the MetroCluster configuration.

Note

DS14mk2 FC disk shelves using ESH or ESH2 modules, or DS14mk2 AT disk shelves using AT-FCX modules operating at 2-Gb Fibre Channel speeds, can attain a maximum node separation distance of only 300 meters, not 500.



Advantages of nonswitched MetroClusters

MetroClusters provide the same advantages of mirroring your data as standard mirrored clusters do:

- ◆ Your data is protected in the event of a failure or loss of two or more disks in a RAID4 aggregate or three or more disks in a RAID-DP aggregate.
- ◆ The failure of an FC-AL adapter, loop, or ESH module does not require a cluster failover.

In addition, a MetroCluster enables you to initiate a failover if an entire site becomes lost or unavailable. If a disaster occurs at one of the node locations and destroys your data there, your data not only survives on the other node, but can be served by that node while you address the issue or rebuild the cluster.

For more information about recovering from a disaster using MetroCluster, see [“Disaster Recovery Using MetroCluster”](#) on page 163.

How Data ONTAP works with nonswitched MetroClusters

During configuration of the cluster, Data ONTAP identifies spare disks and divides them into separate groupings called pools. These pools of disks are physically separated from each other, allowing for high availability of mirrored volumes. When you add a mirrored volume or add disks to one side of a mirrored volume, Data ONTAP determines how much storage you need for the second half of the mirror, and dedicates that storage from a separate pool to the mirrored volume.

Setup requirements and restrictions

The following list specifies the requirements and restrictions you must be aware of when setting up a nonswitched MetroCluster.

- ◆ Architecture compatibility: both nodes must be the same type and model.
- ◆ Disks in the same plex must be from the same pool, with disks in the opposite plex from the opposite pool.
- ◆ The SyncMirror pool rules for direct-attached Fibre Channel disks must be observed.

For more information, see “[SyncMirror pool rules for direct-attached Fibre Channel disks](#)” on page 8.

- ◆ There must be sufficient spares in each pool to account for a disk failure.
- ◆ The following licenses must be enabled on the local and partner nodes:
 - ❖ cluster
 - ❖ syncmirror_local
 - ❖ cluster_remote

Note

See the “Clustered Failover with 500 Meter Distance” document on the NOW™ site for more information about hardware requirements for this configuration.

Understanding Fabric Attached MetroClusters

Configuration overview

The cluster configuration for distances between 500 meters and 30 kilometers uses four Brocade switches in a fabric configuration to connect the two cluster nodes.

Definition of a switch fabric

A Fabric Attached MetroCluster contains two switch fabrics. A switch fabric consists of a switch on the local half of the cluster connected to a switch on the remote half of the cluster. The two switches are connected to each other through long distance ISL (Inter Switch Link) cables.

For example, the first fabric in a MetroCluster begins from Switch 1 on the local side and is completed by connecting the ISL cable to the first switch on the remote side—Switch 3. The second fabric is created using Switch 2 on the local side, connected through a second ISL cable, to the second switch on the remote side—Switch 4.

Note

Only one ISL is supported per fabric.

How switches are divided

The Fibre Channel switches are divided into banks and quadrants. Each switch has two banks and four quadrants, dividing the switch into equal parts. For example, a Brocade® 3800 switch has 16 ports, divided into two banks of eight ports per bank.

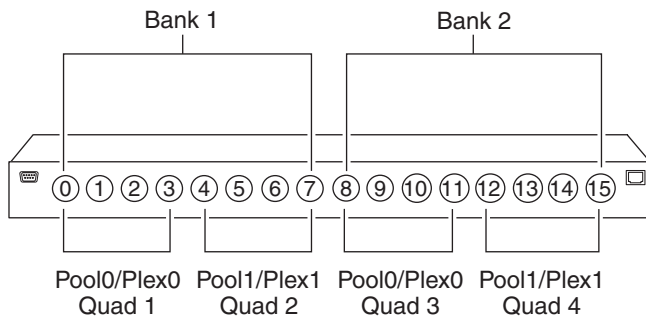
The banks are further subdivided into two quadrants of four ports each (two for the Brocade 3200 and 3250). These quadrants are also identified as pools, which function the same in a Fabric Attached MetroCluster as in a mirrored cluster disk pool. Quadrants 1 and 3 belong to pool 0, and quadrants 2 and 4 belong to pool 1. Each switch bank contains a pool 0 and a pool 1.

The following illustrations show a 16-port and an 8-port Brocade switch, divided into banks, quadrants, and pools.

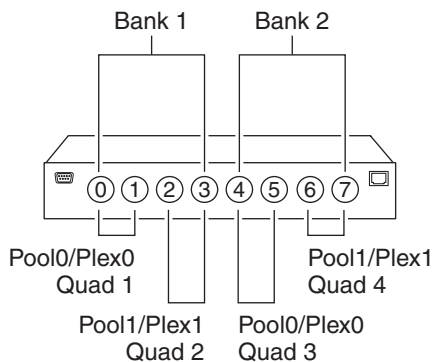
Note

Pools can only be pool 0 or pool 1. However, Data ONTAP can assign arbitrary numbers to plexes. For simplicity, plexes are shown in this document as plex0 and plex1, but they could have different numbers in your installation.

16-Port Switch



8-Port Switch



**Advantages of
Fabric Attached
MetroClusters**

Fabric Attached MetroClusters provide the same advantages as nonswitched MetroClusters by mirroring the data in separate locations. They also provide the following additional advantages:

- ◆ Expanded distance between halves of the cluster beyond 500 meters provides increased disaster protection.

- ◆ Disk shelves and nodes are not connected directly to each other, but are connected to a fabric with multiple data routes, ensuring no single point of failure.

How Data ONTAP works with Fabric Attached MetroClusters

Data ONTAP functions the same as it does in a nonswitched MetroCluster. For more information, see [“How Data ONTAP works with nonswitched MetroClusters”](#) on page 11.

Setup requirements and restrictions

The setup requirements for a Fabric Attached MetroCluster are the same as for nonswitched MetroClusters, with the following exceptions:

- ◆ Node requirements
 - ❖ The nodes must be an F825 or F880 cluster pair, or a FAS900 or FAS3000 series cluster pair, configured for mirrored volume use.
 - ❖ Each node must have one or two supported HBA kits installed and running. For Fabric Attached MetroClusters, only the QLogic 2342 (X2050A) FC HBA or the FAS3000 on-board ports are supported.
 - ❖ Each node requires a QLA2352 VI-MC cluster adapter; the slot position is dependent on the type of storage system.
- ◆ Disk and shelf requirements
 - ❖ Only DS14 and DS14Mk2 disk shelves are supported.
 - ❖ Only FC-type disks are supported; you cannot use ATA drives for Fabric Attached MetroCluster configurations.
 - ❖ You can connect a maximum of 336 drives to a Fabric Attached MetroCluster.

Note

You can not exceed the maximum storage capacity for each node, regardless of the number of disks. For information about the maximum storage capacity for storage systems, see the System Configuration Guide on the NOW site.

- ◆ Switch requirements
 - ❖ Each pair of nodes must use two switches to communicate with the clients.
 - ❖ Switches must be either Brocade® 3200, 3250, 3800, or 3850.
 - ❖ Mixed switch types can be used, but must be the same type on each side of the cluster.
 - ❖ Switches must be running the correct firmware version.

- ❖ Switches must have long distance Small Form-factor Pluggable (SFP) transceivers.

Note

You can download Brocade switch firmware from the NOW site at now.netapp.com.

- ◆ License requirements
 - ❖ `cluster`—Enables the two nodes to communicate and take over each other's functions if a failure occurs.
 - ❖ `syncmirror_local`—Enables mirrored volumes.
 - ❖ `cluster_remote`—Enables mirrored volumes over long distances and enables the local node to take over the functions of the remote node if a failure occurs.

Configuration limitations

The Fabric Attached MetroCluster configuration has the following limitations:

- ◆ You cannot use the switches in a MetroCluster configuration to connect Fibre Channel tape devices, or for FCP traffic of any kind. You can connect only storage systems and disk shelves to the switches in a MetroCluster configuration.
- ◆ You can connect a tape storage area network (SAN) to either of the nodes in a MetroCluster configuration, but the tape SAN must not use the switches used in the MetroCluster configuration.

Cluster failover and single-point-of-failure analysis

Benefits of cluster failover

You can use cluster failover, a high-availability data clustering solution, to further increase the uptime of storage systems. It protects against appliance controller failure by transferring the data service from the failed storage system to its partner storage system. In addition to appliance head failure, cluster failover can also protect against other hardware failures, such as network interface cards, Fibre Channel-Arbitrated Loops (FC-AL loops), and LRC, ESH, ESH2, or AT-FCX modules. Cluster failover is also an effective tool for reducing planned downtime of one of the nodes.

Note

You may also see the term *controller failover*; this is equivalent to the term *cluster failover* used in this document.

This section discusses the following topics:

- ◆ [“SPOF analysis for clusters”](#) on page 18
- ◆ [“Failover event cause-and-effect table”](#) on page 20

Single-point-of-failure definition

SPOF represents the failure of a single hardware component that can lead to loss of data access or potential loss of data. SPOF does not include multiple/rolling hardware errors, such as double disk failure, dual LRC, ESH, ESH2, or AT-FCX failure, etc.

All hardware components have demonstrated very good reliability with low failure rates. If a hardware component fails, such as an appliance controller or adapter, you can use cluster failover to provide continuous data availability and preserve data integrity for client applications and users.

SPOF analysis for clusters

The following table shows which hardware components are SPOFs, and how cluster failover can eliminate these SPOFs to improve data availability.

Hardware components	SPOF		How cluster failover eliminates SPOF
	Single storage system	Clustered storage system	
Storage system controller	Yes	No	If a storage system controller fails, the storage system automatically fails over to its partner node. The partner (takeover) node serves data for both of the nodes.
NVRAM	Yes	No	If an NVRAM adapter fails, the storage system automatically fails over to its partner node. The partner (takeover) node serves data for both of the nodes.
CPU fan	Yes	No	If the CPU fan fails, the storage system gracefully shuts down and automatically fails over to its partner node. The partner (takeover) node serves data for both of the nodes.
Multiple NICs with VIFS (virtual interfaces)	No	No	If one of the networking links within a VIF fails, the networking traffic is automatically sent over the remaining networking links on the storage system. No failover is needed in this situation. If all the NICs in a VIF fail, the node automatically fails over to its partner node, if failover is enabled for the VIF.
Single NIC	Yes	No	If a NIC fails, the node automatically fails over to its partner node, if failover is enabled for the NIC.
FC-AL adapter	Yes	No	If an FC-AL adapter for the primary loop fails, the partner node attempts a failover at the time of failure. If the FC-AL adapter for the secondary loop fails, the failover capability is disabled, but both nodes continue to serve data to their respective applications and users, with no impact or delay.
FC-AL cable (appliance-to-shelf, shelf-to-shelf)	Yes	No	If an FC-AL loop breaks, it could lead to a failover, depending on the shelf type. The clustered storage systems invoke the negotiated failover feature to determine which node is best for serving data, based on the disk shelf count.

Hardware components	SPOF		How cluster failover eliminates SPOF
	Single storage system	Clustered storage system	
LRC, ESH, ESH2, AT-FCX	Yes	No	If one of these components fails, it could lead to a failover. The clustered storage systems invoke the negotiated failover feature to determine which node is best for serving data, based on the disk shelf count.
Disk drive	No	No	If a disk fails, the storage system can reconstruct data from the RAID4 parity disk. No failover is needed in this situation.
Disk shelf (including backplane)	No	No	A disk shelf is a passive backplane with dual power supplies, dual fans, dual LRCs, and dual FC-AL loops. It is the single most reliable component in a storage system.
Power supply	No	No	Both the storage system head and disk shelf have dual power supplies. If one power supply fails, the second power supply automatically kicks in. No failover is needed in this situation. If both power supplies fail, the node automatically fails over to its partner node, which serves data for both nodes.
Fan (storage system or disk shelf)	No	No	Both the storage system head and disk shelf have multiple fans. If one fan fails, the second fan automatically provides cooling. No failover is needed in this situation. If both fans fail, the node automatically fails over to its partner node, which serves data for both nodes.
Cluster adapter	N/A	No	If a cluster adapter fails, the failover capability is disabled but both nodes continue to serve data to their respective applications and users.

Hardware components	SPOF		How cluster failover eliminates SPOF
	Single storage system	Clustered storage system	
Cluster interconnect cable	N/A	No	The cluster adapter supports dual cluster interconnect cables. If one cable fails, the cluster traffic (heartbeat and NVRAM data) are automatically sent over the second cable with no delay or interruption. If both cables fail, the failover capability is disabled but both nodes continue to serve data to their respective applications and users.

Failover event cause-and-effect table

Use this table to understand what happens when a failover event occurs, and how the various clustering configurations handle these events.

Event	Does the event trigger a failover?		Does the event prevent a future failover from occurring, or a failover from occurring successfully?		Is data still available on affected volume after event?		
	Cluster	Metro-Cluster	Cluster	Metro-Cluster	Single storage system	Cluster	Metro-Cluster
Single disk failure	No	No	No	No	Yes	Yes	Yes

Event	Does the event trigger a failover?		Does the event prevent a future failover from occurring, or a failover from occurring successfully?		Is data still available on affected volume after event?		
	Cluster	Metro-Cluster	Cluster	Metro-Cluster	Single storage system	Cluster	Metro-Cluster
Double disk failure (2 disks fail in same RAID group)	Yes, unless you are using Sync-Mirror or RAID-DP™, then no.	No	Maybe. If root volume has double disk failure or if the mailbox disks are affected, no failover is possible.	No	No, unless you are using RAID-DP protection or Sync-Mirror, then yes.	No, unless you are using RAID-DP or Sync-Mirror, then yes.	Yes, with no failover necessary.
Triple disk failure (3 disks fail in same RAID group)	Maybe. If Sync-Mirror is being used, there won't be a takeover; otherwise, yes.	No	Maybe. If root volume has triple disk failure, no failover is possible.	No	No	No	Yes, with no failover necessary.
Single HBA (initiator) failure, Loop A	Maybe. If Sync-Mirror is being used, then no; otherwise, yes.	No	Maybe. If root volume has double disk failure, no failover is possible.	No	Yes, if shelf is dual-pathed, or if Sync-Mirror is being used.	Yes, if Sync-Mirror is being used, or if failover succeeds.	Yes, with no failover necessary.

Event	Does the event trigger a failover?		Does the event prevent a future failover from occurring, or a failover from occurring successfully?		Is data still available on affected volume after event?		
	Cluster	Metro-Cluster	Cluster	Metro-Cluster	Single storage system	Cluster	Metro-Cluster
Single HBA (initiator) failure, Loop B	No	No	Yes, unless you are using SyncMirror and the mailbox disks aren't affected, then no.	No	Yes, if shelf is dual-pathed or mirrored.	Yes, with no failover necessary if Sync-Mirror is being used.	Yes, with no failover necessary.
Single HBA initiator failure, (both loops at same time)	Yes, unless the data is mirrored on a different (up) loop, then no takeover needed.	No	Maybe. If the data is mirrored and the mailbox disks aren't affected, then no; otherwise, yes.	No	No, unless the data was mirrored, then yes.	No failover needed if data was mirrored.	Yes, with no failover necessary.
LRC failure (Loop A)	Only if multidisk volume failure or open loop condition occurs.	No	Maybe. If root volume has double disk failure, no failover is possible because this impacts the mailbox disks.	No	Yes, if shelf is dual-pathed or mirrored.	Yes, if failover succeeds or if data is mirrored.	Yes, with no failover necessary.

Event	Does the event trigger a failover?		Does the event prevent a future failover from occurring, or a failover from occurring successfully?		Is data still available on affected volume after event?		
	Cluster	Metro-Cluster	Cluster	Metro-Cluster	Single storage system	Cluster	Metro-Cluster
LRC failure (Loop B)	No	No	Maybe. If the data is mirrored and the mailbox disks aren't affected, automatic failover happens.	No	Yes, if shelf is dual-pathed or the data is mirrored.	Yes	Yes
ESH or AT-FCX failure (Loop A)	Only if multidisk volume failure or open loop condition occurs, and data isn't mirrored.	No	Maybe. If root volume has double disk failure, no failover is possible.	No	No	Yes, if failover succeeds.	Yes, with no failover necessary.
ESH or AT-FCX failure (Loop B)	No	No	Maybe. If SyncMirror is used, then no; otherwise, yes.	No	Yes, if shelf is dual-pathed or data is mirrored.	Yes	Yes

Event	Does the event trigger a failover?		Does the event prevent a future failover from occurring, or a failover from occurring successfully?		Is data still available on affected volume after event?		
	Cluster	Metro-Cluster	Cluster	Metro-Cluster	Single storage system	Cluster	Metro-Cluster
Shelf (backplane) failure	Only if multidisk volume failure or open loop condition occurs, and data is mirrored.	No	Maybe. If root volume has double disk failure or if the mailboxes are affected, no failover is possible.	No	Maybe. If data is mirrored, then yes; otherwise, no.	Maybe. If data is mirrored, then yes; otherwise, no.	Yes, with no failover necessary.
Shelf, single power failure	No	No	No	No	Yes	Yes	Yes
Shelf, dual power failure	Only if multidisk volume failure or open loop condition occurs and data isn't mirrored.	No	Maybe. If root volume has double disk failure or if the mailbox disks are affected, no failover is possible.	No	Maybe. If data is mirrored, then yes; otherwise, no.	Maybe. If data is mirrored, then yes; otherwise, no.	Yes, with no failover necessary.
Head, single power failure	No	No	No	No	Yes	Yes	Yes
Head, dual power failure	Yes	Yes	Yes, until power is restored.	Yes, until power is restored.	No	Yes, if failover succeeds.	Yes, if failover succeeds.

Event	Does the event trigger a failover?		Does the event prevent a future failover from occurring, or a failover from occurring successfully?		Is data still available on affected volume after event?		
	Cluster	Metro-Cluster	Cluster	Metro-Cluster	Single storage system	Cluster	Metro-Cluster
Cluster interconnect failure (1 port)	No	No	No	No	n/a	Yes	Yes
Cluster interconnect failure (both ports)	No	No	Yes	No. Failover is possible.	n/a	Yes	Yes
Ethernet interface failure (primary, no VIF)	Yes, if set up to do so	Yes, if set up to do so	No	No	Yes	Yes	Yes
Ethernet interface failure (primary, VIF)	Yes, if set up to do so	Yes, if set up to do so	No	No	Yes	Yes	Yes
Ethernet interface failure (secondary, VIF)	Yes, if set up to do so	Yes, if set up to do so	No	No	Yes	Yes	Yes
Ethernet interface failure (VIF, all ports)	Yes, if set up to do so	Yes, if set up to do so	No	No	Yes	Yes	Yes

Event	Does the event trigger a failover?		Does the event prevent a future failover from occurring, or a failover from occurring successfully?		Is data still available on affected volume after event?		
	Cluster	Metro-Cluster	Cluster	Metro-Cluster	Single storage system	Cluster	Metro-Cluster
Tape interface failure	No	No	No	No	Yes	Yes	Yes
Heat exceeds permissible amount	Yes	Yes	No	No	No	No	No
Fan failures (disk shelves, storage system, or storage engine)	No	No	No	No	Yes	Yes	Yes
Reboot	No	No	No	No	Maybe. Depends on cause of reboot.	Maybe. Depends on cause of reboot.	Maybe. Depends on cause of reboot.
Panic	Yes	Yes	Yes	No	Maybe. Depends on cause of panic.	Maybe. Depends on cause of panic.	Maybe. Depends on cause of panic.

About this chapter

This chapter describes the tasks you need to do to install and cable a new standard or mirrored cluster in either NetApp system cabinets or in equipment racks.

Topics in this chapter

The following topics are covered in this chapter:

- ◆ [“Cluster installation types”](#) on page 28
- ◆ [“Required documentation, tools, and equipment”](#) on page 29
- ◆ [“Preparing your equipment”](#) on page 32
- ◆ [“Cabling a FAS3000 series cluster”](#) on page 35
- ◆ [“Cabling a FAS900 series cluster”](#) on page 45
- ◆ [“Cabling an F800 series cluster”](#) on page 57

Cluster installation types

Types of cluster installations

There are two types of cluster installations:

- ◆ Clusters in a system cabinet
- ◆ Clusters in an equipment rack

Clusters in a system cabinet

If you ordered a cluster in a system cabinet, it comes in one or more NetApp System Cabinets. The number of system cabinets you receive depends on how much storage you ordered. All internal adapters, such as networking adapters, Fibre Channel adapters, and other adapters, arrive preinstalled in the nodes.

Single system cabinet cluster: If the cluster you ordered has six or fewer disk shelves, it arrives in a single system cabinet. This system cabinet has both the Channel A and Channel B disk shelves cabled, and also has the cluster adapters cabled.

Multiple system cabinet cluster: If the cluster you ordered has more than six disk shelves, the cluster arrives in two or more system cabinets. You must complete the cabling by cabling the local node to the partner node's disk shelves and the partner node to the local node's disk shelves. You must also cable the nodes together by cabling the cluster adapters. If the cluster uses switches, you must install the switches, as described in the accompanying switch documentation. The system cabinets might also need to be connected to each other. See your System Cabinet Guide for information about connecting your system cabinets together for clustering.

Clusters in an equipment rack

Depending on the amount of storage you ordered, you need to install the equipment in one or more telco-style equipment racks. These racks can hold one or two nodes on the bottom and eight or more disk shelves. For information about how to install the disk shelves and nodes into the equipment racks, see the appropriate Fibre Channel disk shelf guide, hardware guide, or Quick Start guide that came with your equipment.

Required documentation, tools, and equipment

Required documentation

NetApp hardware and service documentation is not contained within a single guide. Instead, the field-replaceable units are documented as separate flyers at <http://now.netapp.com/>.

The following table lists and briefly describes the documentation you might need to reference when preparing a new cluster, or converting two stand-alone systems into a clustered system.

Manual name	Description
System Cabinet Guide	This guide describes how to install NetApp equipment into a system cabinet.
<i>Site Requirements Guide</i>	This guide describes the physical requirements your site must meet to install NetApp equipment.
The appropriate disk shelf guide	These guides describe how to cable a disk shelf to a NetApp storage system.
The appropriate hardware guide: ◆ <i>F800 Series Hardware Guide</i> ◆ <i>FAS900 Series Hardware and Service Guide</i> ◆ FAS3000 series documentation	These guides describe how to install the storage system, connect it to a network, and bring it up for the first time.
<i>NetApp Diagnostics Guide</i>	This guide describes the diagnostics tests that you can run on the storage system.
<i>Upgrade Guide</i>	This guide describes how to upgrade storage system and disk firmware, and how to upgrade storage system software.

Manual name	Description
<i>Data Protection Online Backup and Recovery Guide</i>	This guide describes, among other topics, SyncMirror technology, which is used for mirrored clusters.
<i>System Administration Guide</i>	This guide describes general storage system administration.
<i>Software Setup Guide</i>	This guide describes how to configure the software of a new storage system for the first time.

Required tools

The following list specifies the tools you need to install the cluster:

- ◆ #1 and #2 Phillips screwdrivers
- ◆ Hand level
- ◆ Marker

Required equipment

When you receive your cluster, you should receive the equipment listed in the following table. See the *System Configuration Guide* at <http://now.netapp.com> to confirm your storage system type, storage capacity, etc.

Required equipment	Standard or mirrored cluster
Storage system type	Two of the same type of FAS3000, FAS900, or F800 series storage systems.
Storage	See the <i>System Configuration Guide</i> .
Cluster interconnect adapter	IB cluster adapter (The NVRAM5 functions as the cluster interconnect adapter on FAS900 series and FAS3000 series storage systems.)
FC-AL or FC HBA (FC HBA for Disk) adapters	Minimum of two FC-AL adapters
Fibre Channel switches	N/A

Required equipment	Standard or mirrored cluster
SFP (Small Form Pluggable) modules	N/A
NVRAM5 cluster adapter media converter	Only if using fiber cabling.
Cables (provided with shipment unless otherwise noted)	◆ One optical cable per port on the dual Fibre Channel adapter or One HSSDC to DB-9 node-to-shelf cable per disk loop ◆ Multiple DB-9 shelf-to-shelf cables ◆ For systems using NVRAM5, two 4xIB copper cables, or two 4xIB optical cables Note_____ You must purchase longer optical cables separately for cabling distances greater than 30 meters. ◆ Two optical cables with media converters for the IB cluster adapter Note_____ Clusters that have FC9 disk shelves require one MIA (media interface adapter) for each input LRC on the first disk shelves only.

Preparing your equipment

Where to install a cluster

You can install a cluster in one or more system cabinets or in equipment racks. Install the equipment as described in the appropriate section:

- ◆ [“Preparing the system cabinets”](#)
- ◆ [“Preparing the equipment racks”](#) on page 34

Preparing the system cabinets

To prepare your system cabinets for clustering, complete the following steps.

Step	Action
1	Install the system cabinet or cabinets, as described in the System Cabinet Guide.

Step	Action	
2	If you have a cluster in...	Then...
	A single system cabinet	a. Connect the nodes to the network. b. Connect the system cabinet to an appropriate power source. c. Apply power to the system cabinet and configure the cluster, as described in Chapter 4, “Configuring a Cluster,” on page 95.
	Multiple system cabinets	a. Remove the front and rear doors and any side panels that need to be removed to connect the system cabinets together. b. Connect the system cabinets together, as described in the System Cabinet Guide. c. Connect the nodes to the network. d. Connect the system cabinets to an appropriate power source. e. Apply power to the system cabinets and configure the cluster, as described in Chapter 4, “Configuring a Cluster,” on page 95.
3	See one of the following sections to cable your cluster: ◆ “Cabling a FAS3000 series cluster” on page 35 ◆ “Cabling a FAS900 series cluster” on page 45 ◆ “Cabling an F800 series cluster” on page 57	

Preparing the equipment racks

To prepare your nodes for clustering in equipment racks, complete the following steps.

Step	Action
1	Install the nodes in the equipment rack, as described in the Fibre Channel disk shelf guide, hardware guide, or quick start guide that came with your equipment.
2	Install the disk shelves in the equipment rack, as described in the appropriate Fibre Channel disk shelf guide.
3	Label the dual-port FC-AL interfaces, where appropriate.
4	Connect the nodes to the network, as described in the <i>FAS3000 series Installation and Setup Instructions</i> , <i>FAS900 Series Hardware and Service Guide</i> , or the <i>F800 Series Hardware Guide</i> .
5	See one of the following sections to cable your cluster: ◆ “Cabling a FAS3000 series cluster” on page 35 ◆ “Cabling a FAS900 series cluster” on page 45 ◆ “Cabling an F800 series cluster” on page 57

Cabling a FAS3000 series cluster

Recommended slot locations for FAS3000 series

Check the *System Configuration Guide* at <http://now.netapp.com> to confirm the slot assignment information.

For detailed information

For detailed information about the cabling process, see the following topics:

- ◆ “[Cabling Channel A](#)” on page 36
- ◆ “[Cabling Channel B](#)” on page 40
- ◆ “[Cabling the cluster adapters](#)” on page 44

Cabling Channel A

Channel A loop cabling differences

The procedure for cabling the Channel A loops for a standard cluster and mirrored cluster are different. Follow the directions in the appropriate section:

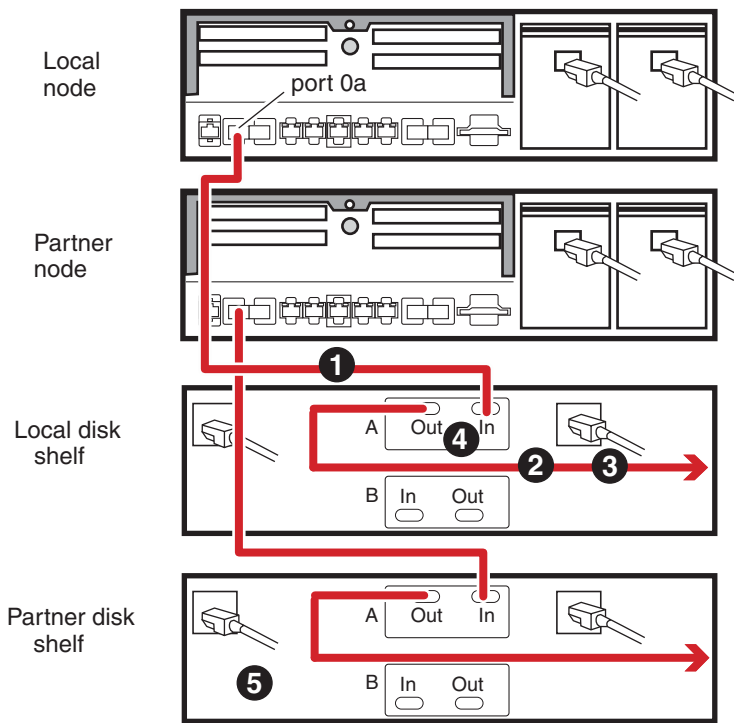
- ◆ “Cabling the Channel A loops in a standard cluster”
- ◆ “Cabling the Channel A loops in a mirrored cluster” on page 38

Note

A switched cluster can be cabled as either a standard or mirrored cluster.

Cabling the Channel A loops in a standard cluster

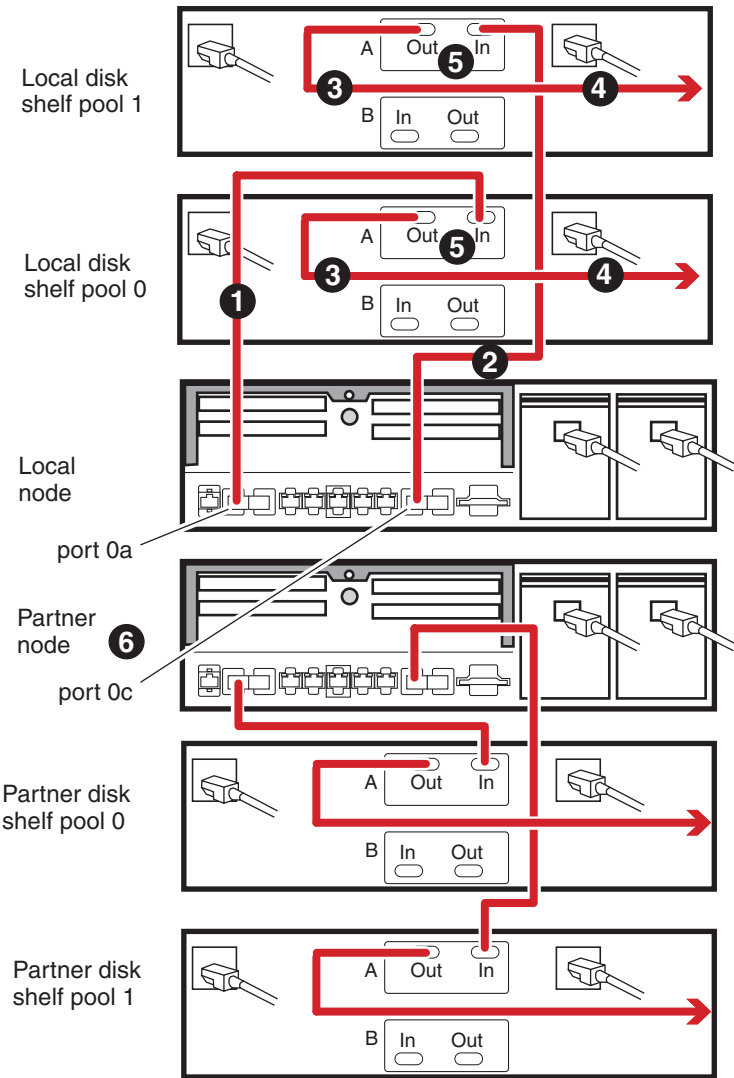
To cable the Channel A loops in a standard cluster, complete the following steps, using the diagram for reference. The circled numbers in the diagram correspond to the step numbers in the table.



Step	Action
1	Cable the dual-port FC-AL adapter in port 0a of the local node to the local disk shelf Channel A LRC, ESH, ESH2, or AT-FCX Input port.
2	Cable the local disk shelf Channel A LRC, ESH, ESH2, or AT-FCX Output port to the Channel A LRC, ESH, ESH2, or AT-FCX Input port of the next disk shelf in the loop. Note_____ The illustration shows only one disk shelf per loop. The number of disk shelves per loop might be different for your storage system. _____
3	Repeat Step 2, connecting output to input, for any remaining disk shelves in the loop.
4	Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the loop, then set the LRC or ESH terminate switches on the last disk shelf to On. Note_____ The ESH2 is self-terminating and does not have a terminate switch. The AT-FCX is self-terminating as long as no cable is plugged into the Out port of the last shelf. _____
5	Repeat Step 1 through Step 4 for the partner node.
6	Go to “ Cabling Channel B ” on page 40.

Cabling the Channel A loops in a mirrored cluster

To cable the Channel A loops in a mirrored cluster, complete the following steps, using the diagram for reference.



Step	Action
1	Cable the dual-port FC-AL adapter in port 0a of the local node to the first local disk shelf Channel A LRC, ESH, ESH2, or AT-FCX Input port of disk shelf pool 0.

Step	Action
2	Cable the dual-port FC-AL in port 0c of the local node to the first local disk shelf Channel A LRC, ESH, ESH2, or AT-FCX Input port of disk pool 1.
3	Cable the partner disk shelf Channel A LRC, ESH, ESH2, or AT-FCX Output port to the next disk shelf Channel A LRC, ESH, ESH2, or AT-FCX Input port in the loop for both disk pools. Note_____ The illustration shows only one disk shelf per disk pool. The number of disk shelves per pool might be different for your storage system. _____
4	Repeat Step 3, connecting Channel A output to input, for any remaining disk shelves in each disk pool.
5	Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the disk pool, then set the LRC or ESH terminate switches on the last disk shelf to On. Note_____ The ESH2 is self-terminating and does not have a terminate switch. The AT-FCX is self-terminating as long as no cable is plugged into the Out port of the last shelf. _____
6	Repeat Step 1 through Step 5 for the partner node.
7	Go to “ Cabling Channel B ” on page 40.

Cabling Channel B

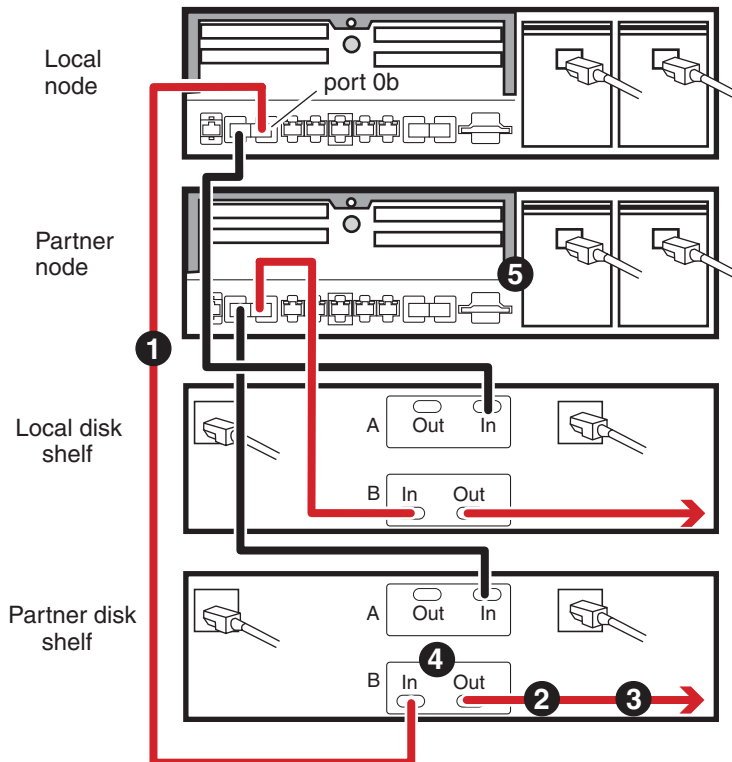
Channel B loop cabling differences

The procedure for cabling the Channel B loops for a standard cluster and mirrored cluster are different. Follow the directions in the appropriate section:

- ◆ “Cabling the Channel B loops in a standard cluster”
- ◆ “Cabling the Channel B loops in a mirrored cluster” on page 42

Cabling the Channel B loops in a standard cluster

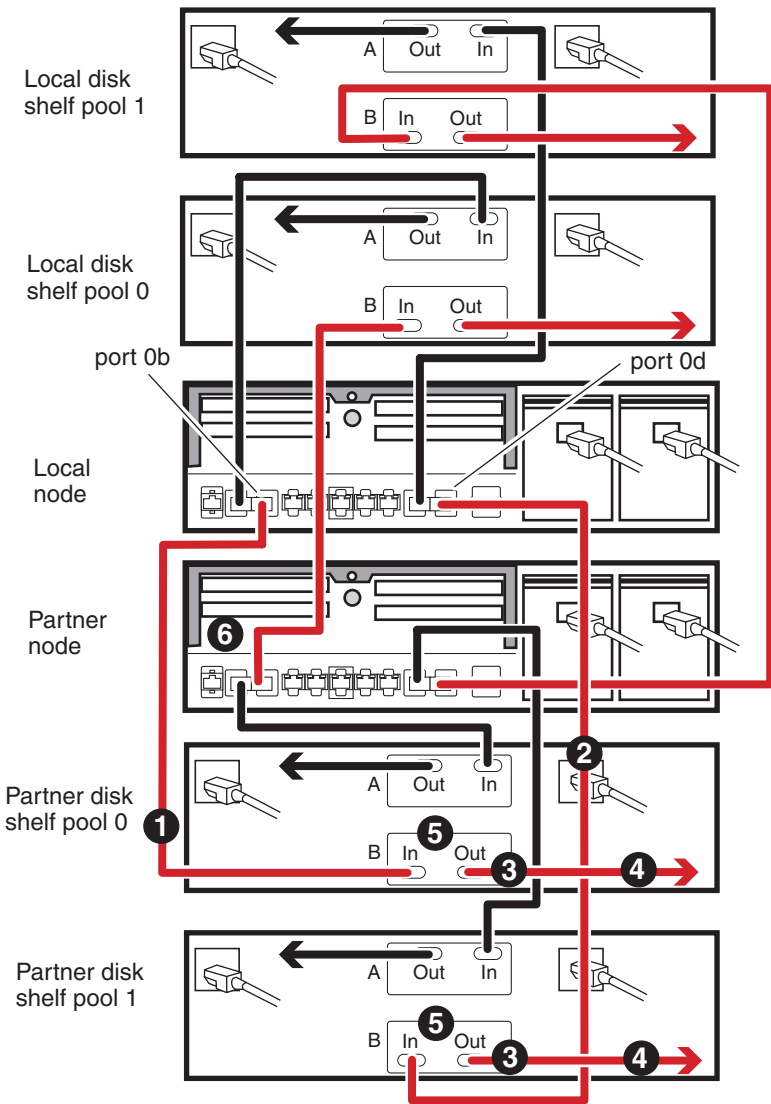
To cable the Channel B loops in a standard cluster, complete the following steps, using the diagram for reference.



Step	Action
1	Cable the dual-port FC-AL adapter in port 0b of the local node, to the partner node's first disk shelf Channel B LRC, ESH, ESH2, or AT-FCX Input port.
2	Cable the partner disk shelf 1 Channel B LRC, ESH, ESH2, or AT-FCX Output port to the next disk shelf of the partner node Channel B LRC, ESH, ESH2, or AT-FCX Input port. Note_____ The illustration shows only one disk shelf per disk pool. The number of disk shelves per pool might be different for your storage system.
3	Repeat Step 2, connecting Channel B output to input, for any remaining disk shelves in the loop.
4	Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the loop, then set the LRC or ESH terminate switches on the last disk shelf to On. Note_____ The ESH2 is self-terminating and does not have a terminate switch. The AT-FCX is self-terminating as long as no cable is plugged into the Out port of the last shelf.
5	Repeat Step 1 through Step 4 for the partner node and local disk shelves.
6	Go to “Cabling the cluster adapters” on page 44.

Cabling the Channel B loops in a mirrored cluster

To cable the Channel B loops in a mirrored cluster, complete the following steps, using the diagram for reference.



Step	Action
1	Cable the dual-port FC-AL adapter in port 0b of the local node to the first partner disk shelf Channel B LRC, ESH, ESH2, or AT-FCX Input port of disk shelf pool 0.

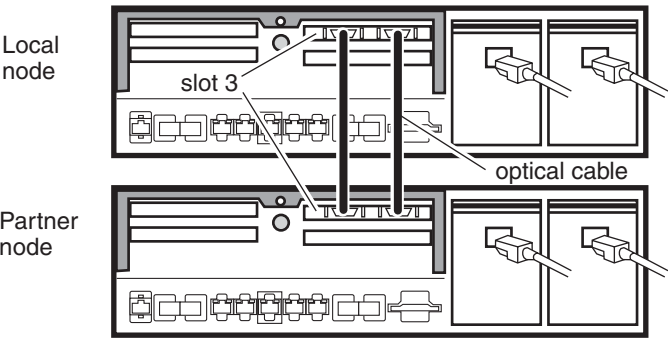
Step	Action
2	Cable the dual-port FC-AL in port 0d of the local node to the first partner disk shelf Channel B LRC, ESH, ESH2, or AT-FCX Input port of disk pool 1.
3	Cable the partner disk shelf Channel B LRC, ESH, ESH2, or AT-FCX Output port to the next disk shelf Channel B LRC, ESH, ESH2, or AT-FCX Input port in the loop for both disk pools. Note_____ The illustration shows only one disk shelf per disk pool. The number of disk shelves per pool might be different for your storage system. _____
4	Repeat Step 3, connecting Channel B output to input, for any remaining disk shelves in the disk pool.
5	Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the disk pool, then set the LRC or ESH terminate switches on the last disk shelf to On. Note_____ The ESH2 is self-terminating and does not have a terminate switch. The AT-FCX is self-terminating as long as no cable is plugged into the Out port of the last shelf. _____
6	Repeat Step 1 through Step 5 for the partner node.
7	Go to “ Cabling the cluster adapters ” on page 44.

Cabling a FAS3000 series cluster

Cabling the cluster adapters

Cabling the adapters

To cable the cluster adapters, complete the following steps, using the diagram for reference. NVRAM5 adapters, which function as the cluster interconnect adapters, must be in slot 3 of your storage system.



Step	Action
1	Plug one end of the optical cable into one of the local node’s cluster adapter ports, then plug the other end into the partner node’s corresponding cluster adapter port. Attention _____ You must not cross-cable the cluster interconnect adapter. Cable the local node ports only to the identical ports on the partner node.
2	Repeat Step 1 for the remaining port on the cluster adapter.
3	Gently pull on the cable ends to ensure that they are properly seated.
4	Go to Chapter 4, “ Configuring a Cluster ,” on page 95.

Cabling a FAS900 series cluster

Type of adapter used in the cluster The dual-port Fibre Channel (dual-port FC-AL) adapters have two ports per adapter. These ports support both the Channel A and Channel B disk shelf loops. The dual-port FC-AL adapter allows your storage system to support more storage than a single-port FC-AL adapter.

Minimum number of adapters needed for mirrored clusters If the cluster is mirrored, each node must have a minimum of two dual-port FC-AL adapters per node. The specific number of dual-port FC-AL adapters you need depends on the number of disk shelves you have in each loop. See the *System Configuration Guide* at <http://www.now.netapp.com> for more information about your hardware.

Recommended slot locations for FAS900 series The following table lists recommended slot locations for the dual-port FC-AL adapters in a single node of a FAS900 series cluster.

Attention —————
Check the *System Configuration Guide* at <http://now.netapp.com> to confirm the slot assignment information.

Cluster type	Slot number on the back of the node
Standard cluster—One dual-port FC-AL adapter	Slot 2
Standard cluster—Two dual-port FC-AL adapters per node	Slot 2, slot 3
Mirrored cluster—Two dual-port FC-AL adapters per node	Slot 8 (pool 1), slot 5 (pool 0)
Standard cluster—Four dual-port FC-AL adapters per node	Slot 2, slot 3, slot 4, slot 8
Mirrored cluster—Four dual-port FC-AL adapters per node	Slot 8 (pool 1), slot 5 (pool 0), slot 10 (pool 1), slot 3 (pool 0)

**For detailed
information**

For detailed information about the cabling process, see the following topics:

- ◆ [“Cabling Channel A”](#) on page 47
- ◆ [“Cabling Channel B”](#) on page 52
- ◆ [“Cabling the cluster adapters”](#) on page 56

Cabling Channel A

Channel A loop cabling differences

The procedure for cabling the Channel A loops for a standard cluster and mirrored cluster are different. Follow the directions in the appropriate section:

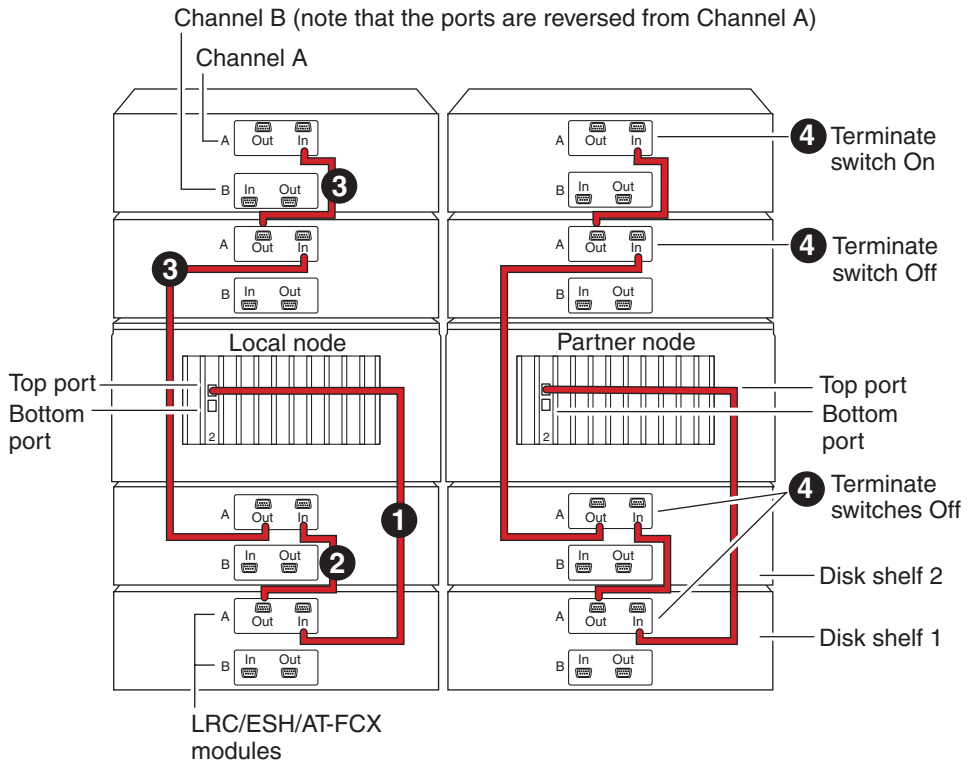
- ◆ “[Cabling the Channel A loops in a standard cluster](#)”
- ◆ “[Cabling the Channel A loops in a mirrored cluster](#)” on page 50

Note

A switched cluster can be cabled as either a standard or mirrored cluster.

Cabling the Channel A loops in a standard cluster

To cable the Channel A loops in a standard cluster, complete the following steps, using the diagram for reference. The circled numbers in the diagram correspond to the step numbers in the table.

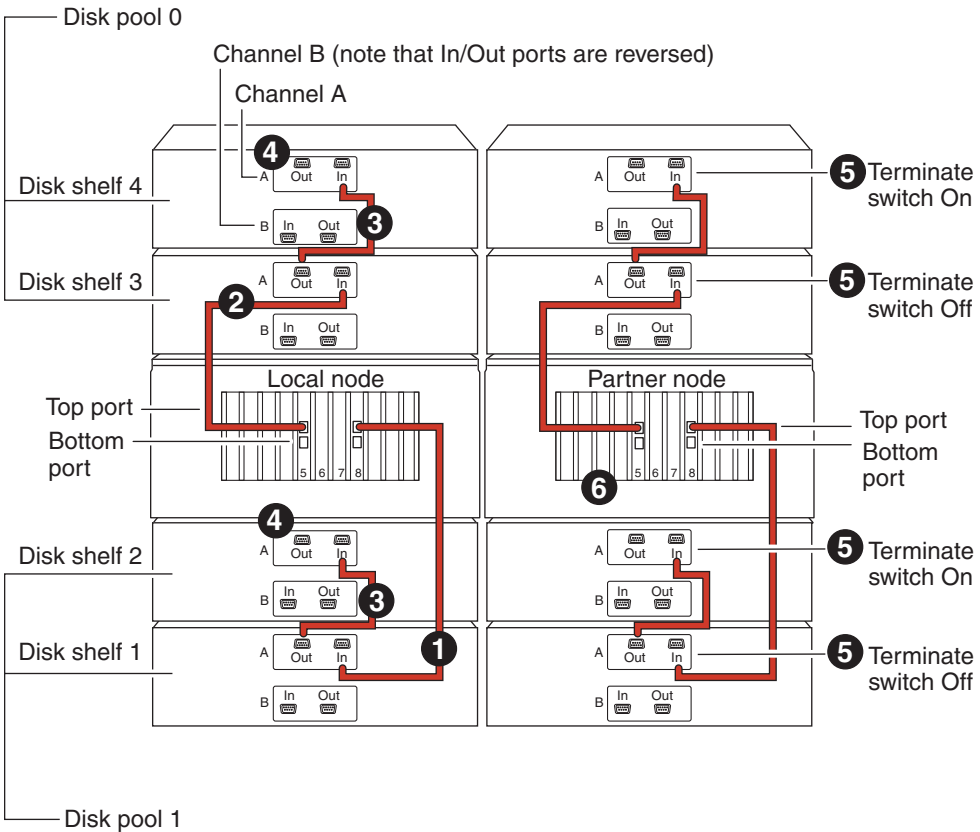


Step	Action
1	Cable the dual-port FC-AL adapter in slot 2, top port, to the disk shelf 1 Channel A LRC, ESH, ESH2, or AT-FCX Input port.
2	Cable the disk shelf 1 Channel A LRC, ESH, ESH2, or AT-FCX Output port to the disk shelf 2 Channel A LRC, ESH, ESH2, or AT-FCX Input port. Label this cable with a solid-colored label.
3	Repeat Step 2, connecting output to input, for the remaining disk shelves in the loop.

Step	Action
4	Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the loop, then set the LRC or ESH terminate switches on the last disk shelf to On. Note_____ The ESH2 is self-terminating and does not have a terminate switch. The AT-FCX is self-terminating as long as no cable is plugged into the Out port of the last shelf. _____
5	Repeat Step 1 through Step 4 for the partner node.
6	Go to “ Cabling Channel B ” on page 52.

Cabling the Channel A loops in a mirrored cluster

To cable the Channel A loops in a mirrored cluster, complete the following steps, using the diagram for reference.



Step	Action
1	Cable the dual-port FC-AL adapter in slot 8, top port, of the local node to the disk shelf 1 Channel A LRC, ESH, ESH2, or AT-FCX Input port of disk pool 1.
2	Cable the dual-port FC-AL in slot 5, top port, of the local node to the disk shelf 3 LRC, ESH, ESH2, or AT-FCX Input port of disk pool 0.
3	Cable the local disk shelf 1 Channel A LRC, ESH, ESH2, or AT-FCX Output port to the disk shelf 2 Channel A LRC, ESH, ESH2, or AT-FCX Input port for both disk pools. Label this cable with a solid-colored label.

Step	Action
4	Repeat Step 3, connecting Channel A output to input, for any remaining disk shelves in each disk pool. Note_____ The illustration shows two disk shelves per disk pool. The number of disk shelves per pool might be different for your storage system. _____
5	Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the disk pool, then set the LRC or ESH terminate switches on the last disk shelf to On. Note_____ The ESH2 is self-terminating and does not have a terminate switch. The AT-FCX is self-terminating as long as no cable is plugged into the Out port of the last shelf. _____
6	Repeat Step 1 through Step 5 for the partner node.
7	Go to “ Cabling Channel B ” on page 52.

Cabling Channel B

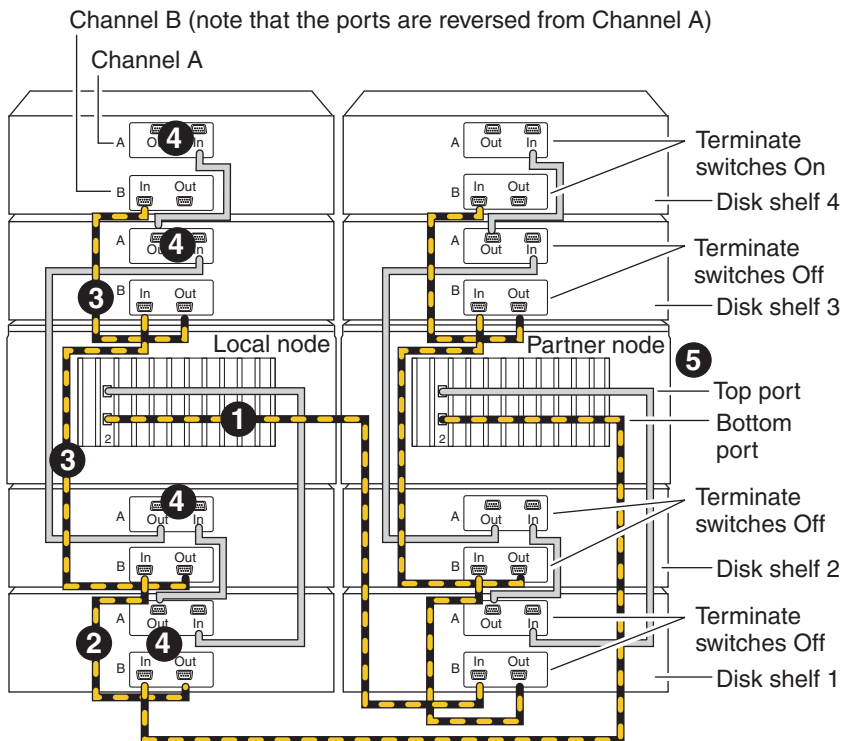
Channel B loop cabling differences

The procedure for cabling the Channel B loops for a standard cluster and mirrored cluster are different. Follow the directions in the appropriate section:

- ◆ “Cabling the Channel B loops in a standard cluster”
- ◆ “Cabling the Channel B loops in a mirrored cluster” on page 54

Cabling the Channel B loops in a standard cluster

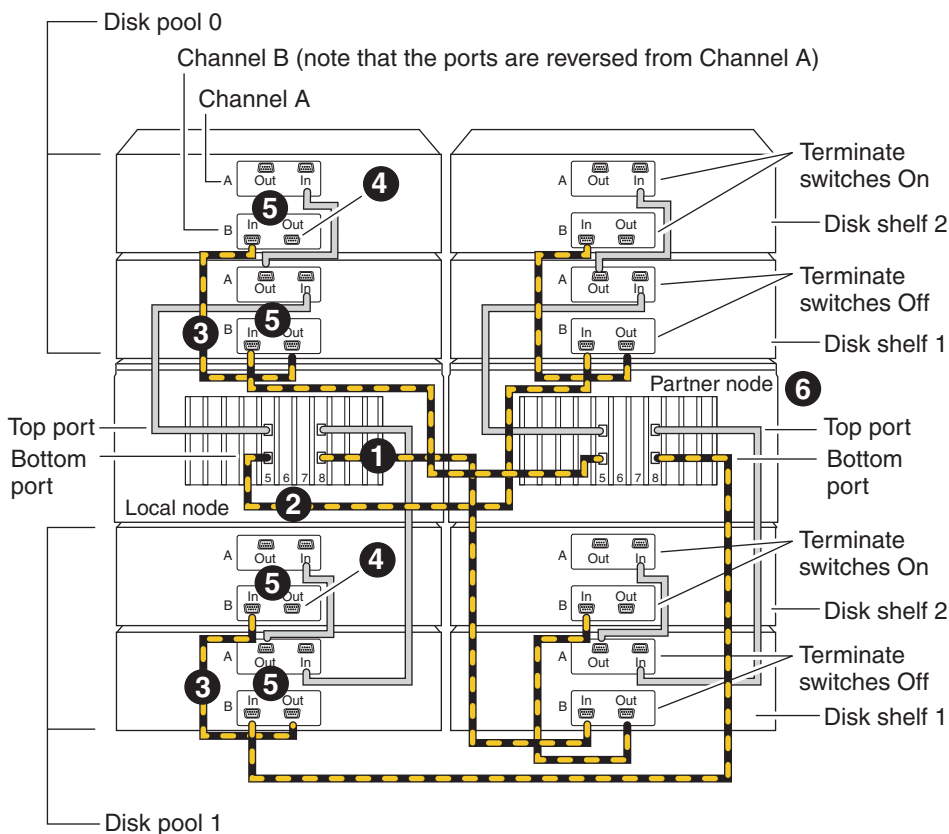
To cable the Channel B loops in a standard cluster, complete the following steps, using the diagram for reference.



Step	Action
1	Cable the dual-port FC-AL adapter in slot 2, bottom port, of the local node, to the partner disk shelf 1 Channel B LRC, ESH, ESH2, AT-FCX Input port.
2	Cable the local disk shelf 1 Channel B LRC, ESH, ESH2, or AT-FCX Output port to the local disk shelf 2 Channel B LRC, ESH, ESH2, or AT-FCX Input port. Label this cable with a striped label.
3	Repeat Step 2, connecting Channel B output to input, for any remaining disk shelves in the disk pool. Note_____ The illustration shows four disk shelves per disk pool. The number of disk shelves per pool might be different for your storage system. _____
4	Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the disk pool, then set the LRC or ESH terminate switches on the last disk shelf to On. Note_____ The ESH2 is self-terminating and does not have a terminate switch. The AT-FCX is self-terminating as long as no cable is plugged into the Out port of the last shelf. _____
5	Repeat Step 1 through Step 4 for the partner node.
6	Go to “ Cabling the cluster adapters ” on page 56.

Cabling the Channel B loops in a mirrored cluster

To cable the Channel B loops in a mirrored cluster, complete the following steps, using the diagram for reference.



Step	Action
1	Cable the dual-port FC-AL adapter in slot 8, bottom port, of the local node to the partner disk shelf 1 Channel B LRC, ESH, ESH2, or AT-FCX Input port of disk pool 1.
2	Cable the dual-port FC-AL adapter in slot 5, bottom port, of the local node to the partner disk shelf 1 Channel B LRC, ESH, ESH2, or AT-FCX Input port of disk pool 0.

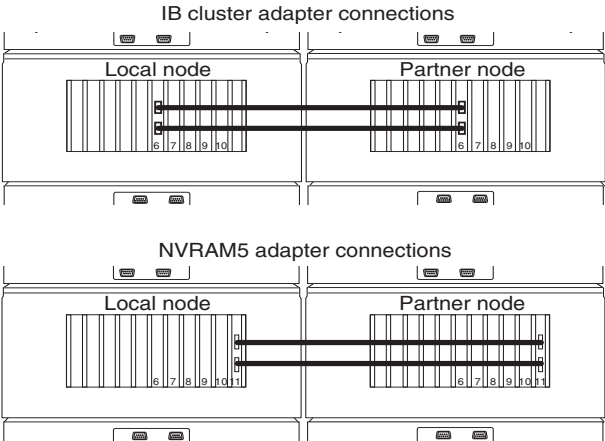
Step	Action
3	Cable the local disk shelf 1 Channel B LRC, ESH, ESH2, or AT-FCX Output port to the local disk shelf 2 Channel B LRC, ESH, ESH2, or AT-FCX Input port for both disk pools. Label this cable with a striped label.
4	Repeat Step 3, connecting Channel B output to input, for any remaining disk shelves in the local disk pools. Note_____ The illustration shows only two disk shelves per disk pool. The number of disk shelves per pool might be different for your storage system. _____
5	Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the local disk pools, then set the LRC or ESH terminate switches on the last disk shelf to On. Note_____ The ESH2 is self-terminating and does not have a terminate switch. The AT-FCX is self-terminating as long as no cable is plugged into the Out port of the last shelf. _____
6	Repeat Step 1 through Step 5 for the partner node.
7	Go to “ Cabling the cluster adapters ” on page 56.

Cabling a FAS900 series cluster

Cabling the cluster adapters

Cabling the adapters

To cable the cluster adapters, complete the following steps, using the diagram for reference. The IB cluster adapters are in slot 6. NVRAM5 adapters, which function as the cluster interconnect adapters, are in slot 11.



Step	Action
1	Plug one end of the optical cable into one of the local node’s cluster adapter ports, then plug the other end into the partner node’s corresponding cluster adapter port. Attention _____ You must not cross-cable the cluster interconnect adapter. Cable the local node ports only to the identical ports on the partner node. _____
2	Repeat Step 1 for the remaining port on the cluster adapter.
3	Gently pull on the cable ends to ensure that they are properly seated.
4	Go to Chapter 4, “ Configuring a Cluster ,” on page 95.

Cabling an F800 series cluster

Recommended slot locations

The following table lists recommended slot locations for F800 series clusters with FC-AL adapters or with dual Fibre Channel adapters.

Attention —
Check <http://now.netapp.com> to confirm the slot assignment information.

Cluster type	Channel A	Channel B
Clusters with two FC-AL adapters per node	Slot 8	Slot 7
Clusters with four FC-AL adapters per node	Slot 8 and slot 5	Slot 7 and slot 9
Clusters with a single dual Fibre Channel adapter per node	Slot 8 (Channel B is on the same adapter)	N/A
Clusters with multiple dual Fibre Channel adapters per node	Slot 8, slot 7, slot 5, and slot 9	N/A

For detailed information

- For detailed information about the cabling process, see the following topics:
- ◆ “[Cabling Channel A](#)” on page 58
 - ◆ “[Cabling Channel B](#)” on page 61
 - ◆ “[Cabling the cluster adapters](#)” on page 64

Cabling Channel A

Channel A loop cabling differences

Generally, the Channel A loop in a cluster is cabled by connecting the FC-AL adapter in the node to the LRC, ESH, or ESH2 Input port of the first disk shelf in the loop, followed by cabling the disk shelves together. This process differs only by the type of Fibre Channel adapter you have in your cluster and the type of disk shelves you have.

Note

If you are cabling a cluster in a system cabinet, skip Channel A cabling and go to [“Cabling the cluster adapters”](#) on page 64. The Channel A loop is already cabled for you.

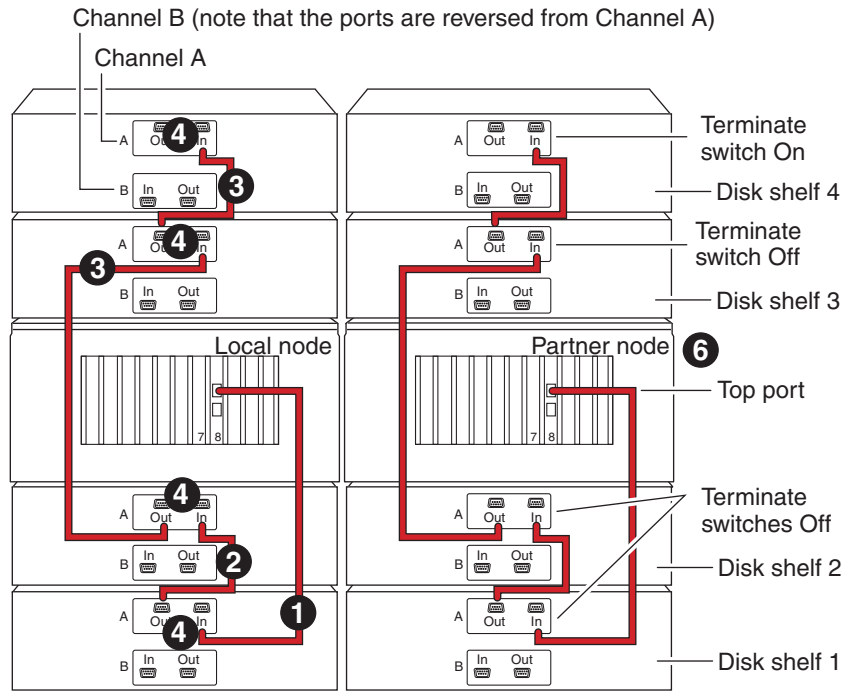
The following table lists exceptions to the general cabling process for the Channel A loop.

If you have...	Then...
Nodes using dual Fibre Channel adapters with any disk shelf type	◆ You use optical cables to connect the nodes to the first disk shelf in the Channel A loop. ◆ The first port (labeled Port 1 on the adapter) of the dual Fibre Channel adapter is reserved for the Channel A loop.
DS14 or DS14mk2 FC disk shelves with nodes using copper FC-AL adapters	You use the HSSDC to DB-9 cable to connect the nodes to the first disk shelf in the Channel A loop.
DS14 or DS14mk2 FC disk shelves with ESH, ESH2, or AT-FCX modules	All disk shelves must have ESH or ESH2 modules installed in them and you must not have mixed shelf-type loops. Use the supplied cables for your FC-AL adapters.
FC9 disk shelves with nodes using dual Fibre Channel adapters	Your first disk shelf LRC requires an MIA (media interface adapter) plugged into the Input port.

Cabling the Channel A loops

To cable the Channel A loops, complete the following steps, using the diagram for reference.

Note All diagrams in this book show DS14 disk shelves. If you have other types of disk shelves, the Input and Output ports might be in different locations.



Step	Action
1	Cable the FC-AL adapter in Slot 8, top port, of the local node, to the disk shelf 1 Channel A LRC, ESH, or ESH2 Input port. Note AT-FCX modules are not supported on F800 series storage systems
2	Cable the disk shelf 1 Channel A LRC, ESH, or ESH2 Output port to the disk shelf 2 Channel A LRC, ESH, or ESH2 Input port. Apply solid-colored labels to the cables.

Step	Action	
3	Repeat Step 2, connecting output to input, for the remaining disk shelves in the loop.	
4	Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the loop, then set the LRC or ESH terminate switches on the last disk shelf to On. Note _____The ESH2 is self-terminating and does not have a terminate switch.	
5	If you have...	Then...
	A mix of DS14 and DS14mk2 FC disk shelves with LRC modules	Set the 1/2 Gb switch to 1 Gb.
	A mix of DS14 and DS14mk2 FC disk shelves with ESH or ESH2 modules	Set the 1/2 Gb switch to 1 Gb.
	Only DS14mk2 FC disk shelves with ESH, or ESH2 modules	Set the 1/2 Gb switch to either 1 or 2 Gb.
6	Repeat Step 1 through Step 5 for the partner node.	
7	Go to “ Cabling Channel B ” on page 61.	

**Channel B loop
cabling differences**

Generally, the Channel B loop in a cluster is cabled by connecting the FC-AL adapter in the local node to the LRC, ESH, or ESH2 Input port of the first disk shelf in the partner node’s Channel B loop, followed by cabling the disk shelves together. This process differs only by the type of Fibre Channel adapter you have in your cluster and the type of disk shelves you have.

Note_____

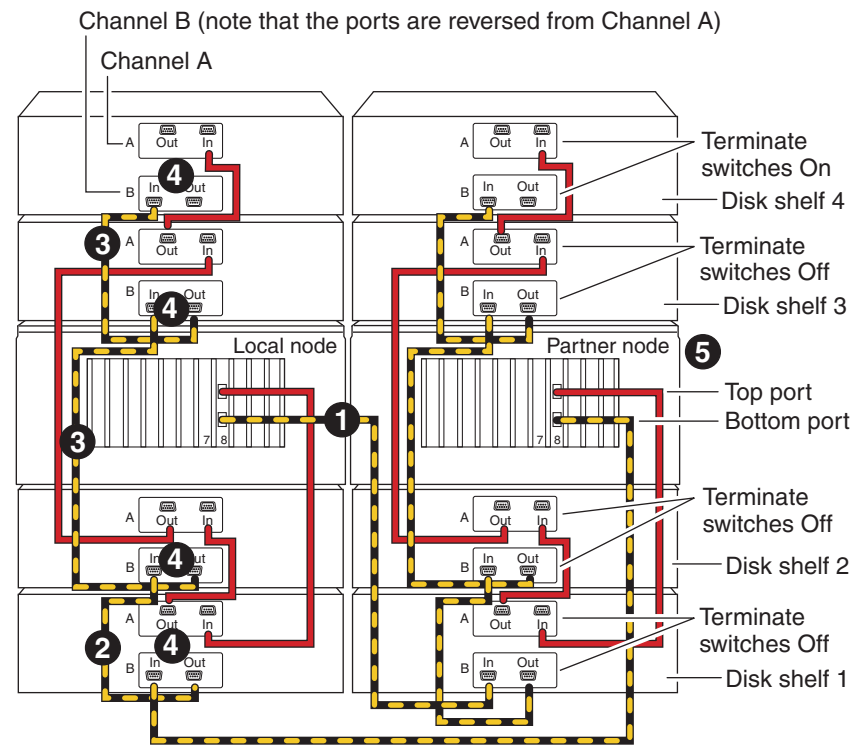
If you are cabling a cluster in a system cabinet, skip Channel B cabling and go to “[Cabling the cluster adapters](#)” on page 64. The Channel B loop is already cabled for you.

The following table lists exceptions to the general cabling process for the Channel B loop.

If you have...	Then...
Nodes using dual Fibre Channel adapters with any disk shelf type	◆ You use optical cables to connect the nodes to the first disk shelf in the Channel A loop. ◆ The second port (labeled Port 1 on the adapter) of the dual Fibre Channel adapter is reserved for the Channel B loop.
DS14 or DS14mk2 FC disk shelves with nodes using copper FC-AL adapters	You use the HSSDC to DB-9 cable to connect the nodes to the first disk shelf in the Channel A loop.
DS14 or DS14mk2 FC disk shelves with ESH or ESH2 modules	All disk shelves must have ESH or ESH2 modules installed in them and you must not have mixed shelf-type loops. Use the supplied cables for your FC-AL adapters.
FC9 disk shelves with nodes using dual Fibre Channel adapters	Your first disk shelf LRC requires an MIA (media interface adapter) plugged into the Input port.

Cabling the Channel B loops

To cable the Channel B loops, complete the following steps, using the diagram for reference.



Step	Action
1	Cable the FC-AL adapter in Slot 8, bottom port, of the local node, to the partner disk shelf 1 Channel B LRC, ESH, or ESH2 Input port.
2	Cable the local disk shelf 1 Channel B LRC, ESH, or ESH2 Output port to the local disk shelf 2 Channel B LRC, ESH, or ESH2 Input port. Label this cable with a striped label.
3	Repeat Step 2, connecting output to input, for the remaining disk shelves in the loop.

Step	Action
4	Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the disk pool, then set the LRC or ESH terminate switches on the last disk shelf to On. Note_____ The ESH2 is self-terminating and does not have a terminate switch. _____
5	Repeat Step 1 through Step 4 for the partner node.
6	Go to “ Cabling the cluster adapters ” on page 64.

Cabling the cluster adapters

Cluster cabling differences

Your cluster arrives with one of four different types of cluster adapters, depending on the configuration you are installing. The following table lists the different supported cluster adapters and the major differences between them.

See the *System Configuration Guide* at <http://now.netapp.com> for current information about supported cluster adapters for specific storage systems.

Cluster adapter type	General characteristics
Cluster interconnect adapter	◆ No longer shipped with new systems. ◆ Does not support mirrored clusters.
IB cluster adapter	◆ Uses an optical cable up to 500 meters long. ◆ Requires two SC/SC cables and GBICs for distances over 30 meters (purchased separately).
VI cluster adapter (Rev B)	◆ Uses an optical cable up to 500 meters long. ◆ Requires an LC/LC cable for distances over 30 meters (purchased separately).
VI-MC cluster adapter	◆ Connects directly to Fibre Channel switches. ◆ Requires two SC/SC cables (purchased separately).

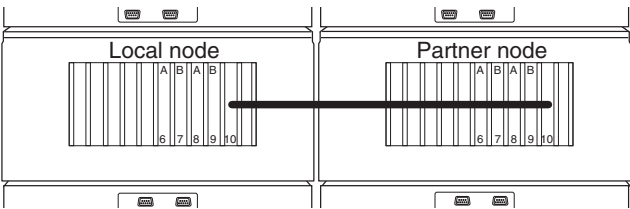
Cabling the IB cluster adapter

To cable the IB cluster adapters, complete the following steps.

Step	Action
1	Plug one end of the optical cable into one of the local node's IB cluster adapter ports, then plug the other end into the partner node's corresponding IB cluster adapter port. Attention ————— You must not cross-cable the cluster interconnect adapter. Cable the local node ports only to the identical ports on the partner node.
2	Repeat Step 1 for the remaining port on the cluster adapter.
3	Gently pull on the cable ends to ensure that they are properly seated.
4	Go to Chapter 4, “ Configuring a Cluster ,” on page 95.

Cabling the VI cluster adapter

To cable the VI cluster adapters, complete the following steps, using the diagram for reference.



Step	Action
1	Plug one end of the optical cable into the GBIC of the local node's VI cluster adapter, then plug the other end into the partner node's corresponding VI cluster adapter. Attention ————— You must not cross-cable the cluster interconnect adapter. Cable the local node ports only to the identical ports on the partner node.
2	Go to Chapter 4, “ Configuring a Cluster ,” on page 95.

About this chapter

This chapter contains instructions for cabling a new Fabric Attached or nonswitched MetroCluster. This chapter also describes how to convert a cluster into a Fabric Attached MetroCluster.

Topics in this chapter

This chapter discusses the following topics:

- ◆ [“Required documentation, tools, and equipment”](#) on page 68
- ◆ [“Cabling a new Fabric Attached MetroCluster”](#) on page 72
- ◆ [“Cabling a nonswitched MetroCluster”](#) on page 91
- ◆ [“Converting a cluster into a Fabric Attached MetroCluster”](#) on page 93

Required documentation, tools, and equipment

Required documentation

NetApp hardware and service documentation is not contained within a single guide. Instead, the field-replaceable units are documented as separate flyers at <http://now.netapp.com/>.

The following table lists and briefly describes the documentation you might need to reference when preparing a new cluster, or converting two stand-alone systems into a clustered system.

Manual name	Description
System Cabinet Guide	This guide describes how to install NetApp equipment into a system cabinet.
<i>Site Requirements Guide</i>	This guide describes the physical requirements your site must meet to install NetApp equipment.
The appropriate disk shelf guide	These guides describe how to cable a disk shelf to storage systems.
The appropriate hardware guide: ◆ <i>F800 Series Hardware Guide</i> ◆ <i>FAS900 Series Hardware and Service Guide</i> ◆ FAS3000 series documentation	These guides describe how to install the storage system, connect it to a network, and bring it up for the first time.
<i>NetApp Diagnostics Guide</i>	This guide describes the diagnostics tests that you can run on the storage system.
<i>Upgrade Guide</i>	This guide describes how to upgrade storage system and disk firmware, and how to upgrade storage system software.
<i>Data Protection Online Backup and Recovery Guide</i>	This guide describes, among other topics, SyncMirror technology.

Manual name	Description
<i>System Administration Guide</i>	This guide describes general storage system administration.
<i>Software Setup Guide</i>	This guide describes how to configure the software of a new storage system for the first time.
The appropriate Brocade manuals	These guides describe how to configure and maintain Brocade switches, which are used for Fabric Attached MetroClusters.

Required tools

The following list specifies the tools you need to install the cluster:

- ◆ #1 and #2 Phillips screwdrivers
- ◆ Hand level
- ◆ Marker

Required equipment

When you receive your cluster, you should receive the equipment listed in the following table. See the *System Configuration Guide* at <http://now.netapp.com> to confirm your storage system type, storage capacity, etc.

Required equipment	Nonswitched MetroCluster	Fabric Attached MetroCluster
Storage system model	Two of the same type of FAS3000, FAS900, or F800 series storage systems.	
Storage	See the <i>System Configuration Guide</i> .	
Cluster interconnect adapter	IB cluster adapter (The NVRAM5 functions as the cluster interconnect adapter on F800 series, FAS900 series, and FAS3000 series storage systems).	VI-MC cluster adapter

Required equipment	Nonswitched MetroCluster	Fabric Attached MetroCluster
FC-AL or FC HBA (FC HBA for Disk) adapters	Two or four Fibre Channel HBAs. Note_____ The ports on the Fibre Channel HBAs are labeled 1 and 2. However, the software refers to them as a and b. You see these labeling conventions in the user interface and system messages displayed on the console. _____	
Fibre Channel switches	N/A	Two pairs of Brocade 3800 or 3200 series switches. Note_____ Mixed switch types can be used, but switches must be the same type on each side of the cluster. _____
SFP (Small Form Pluggable) modules	N/A	Two per ISL—four total, long distance. One per switch port used, short distance.
NVRAM5 cluster adapter media converter	Only if using fiber cabling.	N/A

Required equipment	Nonswitched MetroCluster	Fabric Attached MetroCluster
Cables (provided with shipment unless otherwise noted)	◆ Four SC/LC FC-AL to disk shelf cables ◆ Two SC/LC IB cluster adapter cables ◆ Four SC/LC or LC/LC cables	◆ LC/LC node-to-switch cables ◆ SC/LC (for DS14) or LC/LC (for DS14mk2 FC) disk shelf-to-switch cables ◆ Two LC/LC ISL cables, not provided in the shipment ◆ HSSDC2-to-HSSDC2 or DB-9 to DB-9 shelf-to-shelf cables for DS14 disk shelves with ESH or ESH2 modules

Cabling a new Fabric Attached MetroCluster

What you need to know

A Fabric Attached MetroCluster works with LRC, ESH, and ESH2 modules only. AT-FCX modules are not supported with a Fabric Attached MetroCluster.

To cable a Fabric Attached MetroCluster, you must be familiar with clustering concepts, the Brocade command-line interface, and synchronous mirroring. You should also be familiar with the characteristics of Fabric Attached MetroClusters, as described in “[Understanding Fabric Attached MetroClusters](#)” on page 12. You must also have the following information:

- ◆ Unique domain IDs for the switches
- ◆ Ethernet IP address for both the switches and nodes

Note

The switches ship with a default IP address (10.77.77.77), which you can use if the switches won't be attached to a network.

- ◆ Ethernet subnetmask
- ◆ Gateway address

For detailed information

For details about setting up and cabling your Fabric Attached MetroCluster, see the following sections:

- ◆ “[Stage 1: Planning the Fabric Attached MetroCluster](#)” on page 73
- ◆ “[Stage 2: Configuring the local switches](#)” on page 74
- ◆ “[Stage 3: Cabling the local node](#)” on page 78
- ◆ “[Stage 4: Cabling the switches and nodes at the partner site](#)” on page 84

Stage 1: Planning the Fabric Attached MetroCluster

Fabric Attached MetroCluster worksheet

The Fabric Attached MetroCluster worksheet is a planning aid to record specific cabling information about your Fabric Attached MetroCluster. You must identify several pieces of information that you use during configuration procedures. Recording this information can reduce configuration errors.

Local node name: _____

Cluster adapter Port 1 is cabled to Switch 1 Port__

Switch 1 name: _____ (up to 12 characters)

Switch 1 IP address:_____

FC HBA 1, Port 1, is cabled to Switch 1, Port__

FC HBA 1, Port 2, is cabled to Switch 2, Port__

Pool 0 disk shelves:

Disk 1 is connected to port ____

Disk 2 is connected to port ____

Remote node name: _____

Cluster adapter Port 1 is cabled to Switch 3, port __

Switch 3 name: _____ (up to 12 characters)

Switch 3 IP address:_____

FC HBA 1, Port 1, is cabled to Switch 3, Port__

FC HBA 1, Port 2, is cabled to Switch 4, Port__

Pool 0 disk shelves:

Disk 1 connected to Port ____

Disk 2 connected to Port ____

Cluster adapter Port 2 is cabled to Switch 2 Port __

Switch 2 name: _____ (up to 12 characters)

Switch 2 IP address:_____

FC HBA 2, Port 1, is cabled to Switch 1, Port ____

FC HBA 2, Port 2, is cabled to Switch 2, Port ____

Pool 1 disk shelves:

Disk 3 is connected to Port ____

Disk 4 is connected to Port ____

Cluster adapter Port 2 is cabled to Switch 4, Port__

Switch 4 name: _____ (up to 12 characters)

Switch 4 IP address:_____

FC HBA 2, Port 1, is cabled to Switch 3, Port ____

FC HBA 2, Port 2, is cabled to Switch 4, Port ____

Pool 1 disk shelves:

Disk 3 connected to Port ____

Disk 4 connected to Port ____

Cabling a new Fabric Attached MetroCluster

Stage 2: Configuring the local switches

Configuring the switch

To configure the Brocade 3800, 3850, 3200, or 3250 switch for clustering, complete the following steps.

Step	Action
1	Log in to the switch, using the console port and hyperterm, and set the appropriate variables and IP addresses, as described in the appropriate Brocade switch hardware reference manual. Note _____ Use telnet for subsequent switch login sessions, using the switch IP address.
2	Check the switch firmware by entering the following command: version Result: The currently installed switch firmware is displayed.
3	Check the MetroCluster Fabric Switch download page on http://now.netapp.com for the currently supported version of the switch firmware. Note _____ To access this page, go to Fibre Channel Switch link on the Download Software page, selecting Brocade in the platform list.

Step	Action	
4	If the switch firmware is...	Then...
	Not the supported version	a. Download the switch firmware from http://now.netapp.com and install it, as described in the appropriate Brocade switch hardware reference manual. b. Reboot the switch.
	Current	Go to Step 5 .
5	Disable the switch by entering the following command: switchdisable	
6	Clear any existing configuration on the switch by entering the following command: cfgclear	
7	Disable the configuration by entering the following command: cfgdisable	
8	Save the configuration by entering the following command: cfgsave	
9	Ensure that the switch is configured with default settings by entering the following command: configdefault	
10	Set the switch parameters by entering the following command: configure You should set the following parameters only: ◆ Fabric parameters = y ◆ domain_id = XXX (See your system administrator for this value.) ◆ Disable device probing = 1	
11	Exit the configuration utility by entering the following command: ctrl-d	

Step	Action	
12	Reboot the switch so that the default settings take effect by entering the following command: fastboot	
13	Log in to the switch and disable the switch by entering the following command: switchdisable	
14	Set the switch port speeds to half duplex by entering the following command for each switch port into which a disk is plugged: portcfglport <port#>,0,0,1 Note _____You only need to run this command on the ports to which disks are attached. _____	
15	If the distance between switches is...	Then...
	Less than 10 kilometers	Set the ISL port distance on each ISL port to short distance by entering the following command: portCfgLongDistance ISL-port#, "LE"
	Between 10 and 30 kilometers	Set the ISL port distance on each ISL port to long distance by entering the following command: portCfgLongDistance ISL-port#, "L1"
16	Enable the switch by entering the following command: switchenable	
17	Set the switch name by entering the following command: switchname "Metro1_sw1" <i>Metro1_sw1</i> is the name given to the current switch. Use a meaningful name up to 12 characters long.	
18	Check the switch settings by entering the following command: configshow	

Step	Action
19	Repeat Step 1 through Step 18 for the second switch on the local node.
20	Go to “ Stage 3: Cabling the local node ” on page 78.

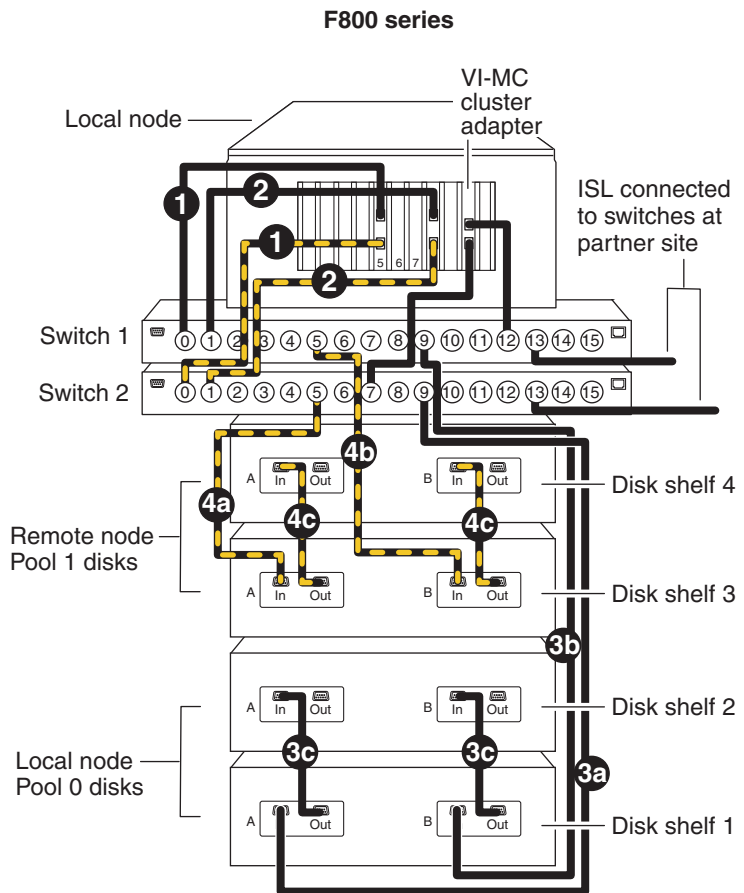
Stage 3: Cabling the local node

Cabling the local node and disk shelves

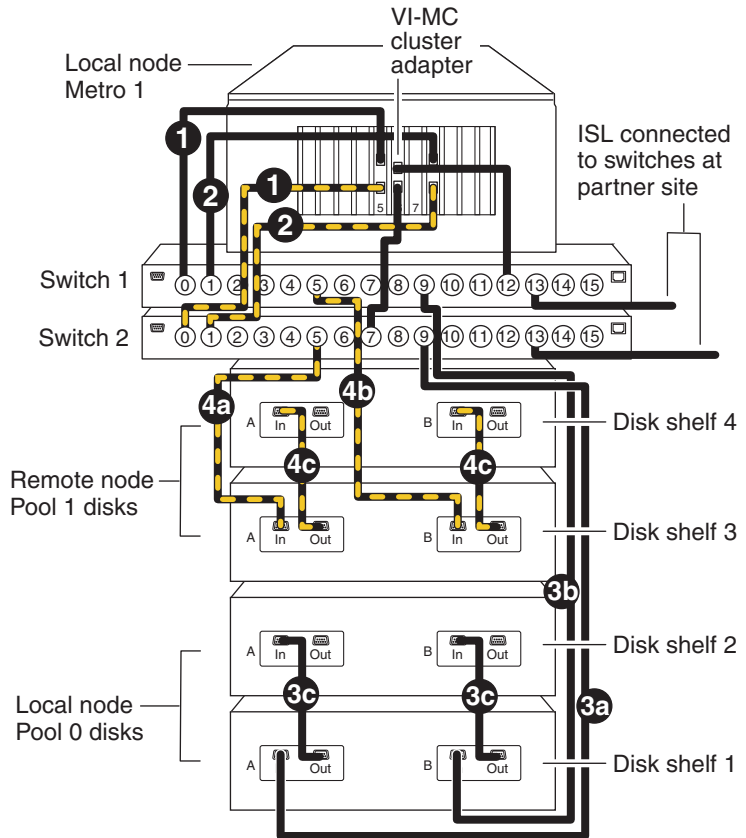
To cable the node, complete the following steps, using the applicable diagram for reference.

Note.

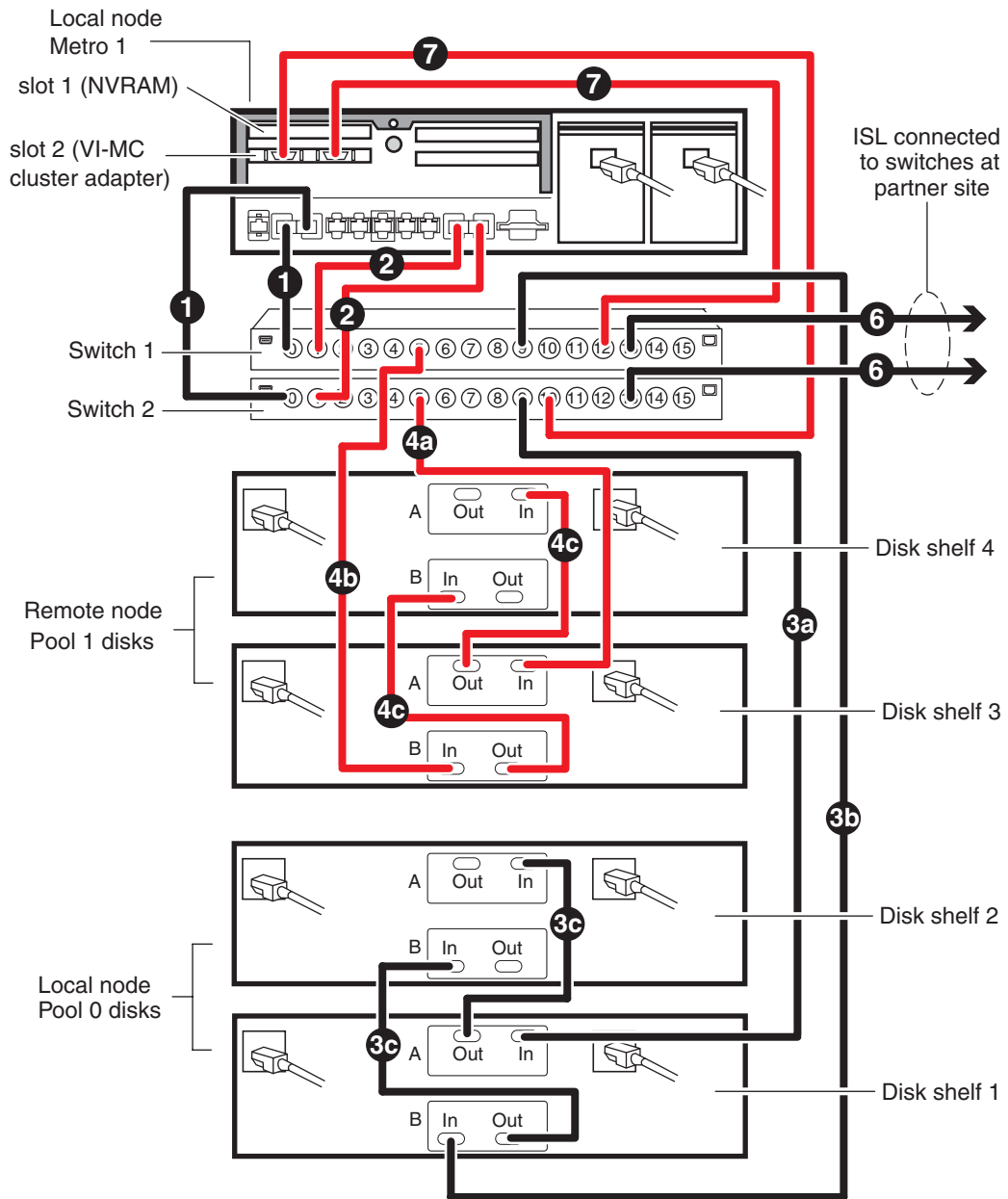
You can only cable a maximum of two disk shelves per loop.



FAS900 series (except FAS980)



FAS3000 series



Step	Action
1	Cable the first Fibre Channel HBA of the local node to Switch 1 and Switch 2. In the example, one port (0a) of the Fibre Channel HBA is cabled to Port 0 of Switch 1 and the second port (0b) is cabled to Port 0 of Switch 2.
2	Cable the second Fibre Channel HBA of the local node to Switch 1 and Switch 2. In the example, one port (0c) of the Fibre Channel HBA is cabled to port 1 of Switch 1 and the second port (0d) is cabled to port 1 of Switch 2.
3	Connect the local node pool 0 disks to the switches by completing the following substeps: - Connect the Input port of the LRC, ESH, or ESH2 module Channel A on disk shelf 1 to port 9 on Switch 2. - Connect the Input port of the LRC, ESH, or ESH2 module Channel B on disk shelf 1 to port 9 on Switch 1. - Connect disk shelf 1 to disk shelf 2 by connecting the Output ports of the LRC, ESH, or ESH2 module of disk shelf 1 to the Input ports of the corresponding LRC, ESH, or ESH2 module of the next disk shelf. - Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the disk pool, then set the LRC or ESH terminate switches on the last disk shelf to On. Note _____ ESH2 modules are self-terminating and therefore do not have a terminate switch. In the example, the pool 0 disk shelves are connected to switch port 9 on both Switch 1 and 2. However, you can use any open set of switch ports in this quadrant (limited to ports 8-11 in this configuration).

Step	Action																																										
4	Connect the remote node pool 1 disks to the switches by completing the following substeps: a. Connect the Input port of the LRC, ESH, or ESH2 module Channel A on disk shelf 3 to port 5 on Switch 2. b. Connect the Input port of the LRC, ESH, or ESH2 module Channel B on disk shelf 3 to port 5 on Switch 1. c. Connect disk shelf 3 to disk shelf 4 by connecting the Output ports of the LRC, ESH, or ESH2 module of disk shelf 3 to the Input ports of the corresponding LRC, ESH, or ESH2 module of the next disk shelf. d. Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the disk pool, then set the LRC or ESH terminate switches on the last disk shelf to On. Note _____ The ESH2 modules are self-terminating and therefore do not have a terminate switch. In the example, the pool 1 disk shelves are connected to switch port 5 on both Switch 1 and 2. However, you can use any open set of switch ports in this quadrant (limited to ports 4-7 in this configuration).																																										
5	Confirm that the disks are visible and have dual paths by entering the following command on the console: storage show disk -p Results: The output shows the disks connected to the switches, to what port they are connected, and to what disk and to which LRC, ESH, or ESH2 they belong, as shown in the following example: Example: ha15*> storage show disk -p <table><tr><th>PRIMARY</th><th>PORT</th><th>SECONDARY</th><th>PORT</th><th>SHELF</th><th>BAY</th></tr><tr><td>-----</td><td>----</td><td>-----</td><td>----</td><td>-----</td><td>----</td></tr><tr><td>switch1:4.40</td><td>A</td><td>switch2:4.40</td><td>B</td><td>5</td><td>0</td></tr><tr><td>switch2:4.41</td><td>B</td><td>switch1:4.41</td><td>A</td><td>5</td><td>1</td></tr><tr><td>.</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>.</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>switch1:8.52</td><td>B</td><td>switch2:8.52</td><td>A</td><td>6</td><td>4</td></tr></table>	PRIMARY	PORT	SECONDARY	PORT	SHELF	BAY	-----	----	-----	----	-----	----	switch1:4.40	A	switch2:4.40	B	5	0	switch2:4.41	B	switch1:4.41	A	5	1	.						.						switch1:8.52	B	switch2:8.52	A	6	4
PRIMARY	PORT	SECONDARY	PORT	SHELF	BAY																																						
-----	----	-----	----	-----	----																																						
switch1:4.40	A	switch2:4.40	B	5	0																																						
switch2:4.41	B	switch1:4.41	A	5	1																																						
.																																											
.																																											
switch1:8.52	B	switch2:8.52	A	6	4																																						

Step	Action
6	Connect one ISL cable to any port on each switch, making sure that the ports have the same number. In the example, Port 13 is used on both switches.
7	Connect the VI-MC cluster adapter to the switches, with one port connecting to Switch 1 and the second port to Switch 2. Note_____ There should be one VI-MC cluster adapter connection per switch. The switch port connections do not have to be identical; therefore, use any available switch ports. For a FAS3000 series MetroCluster, the NVRAM card cannot be used as a cluster adapter and must be moved to slot 1.
8	Boot the node.
9	Check to make sure that the node can see all disk shelves by entering the following command on the console: <pre>aggr status -r</pre> Result: All disks are displayed.
10	Go to the remote site and follow the instructions in “ Stage 4: Cabling the switches and nodes at the partner site ” on page 84.

Stage 4: Cabling the switches and nodes at the partner site

Cabling switches at the partner site

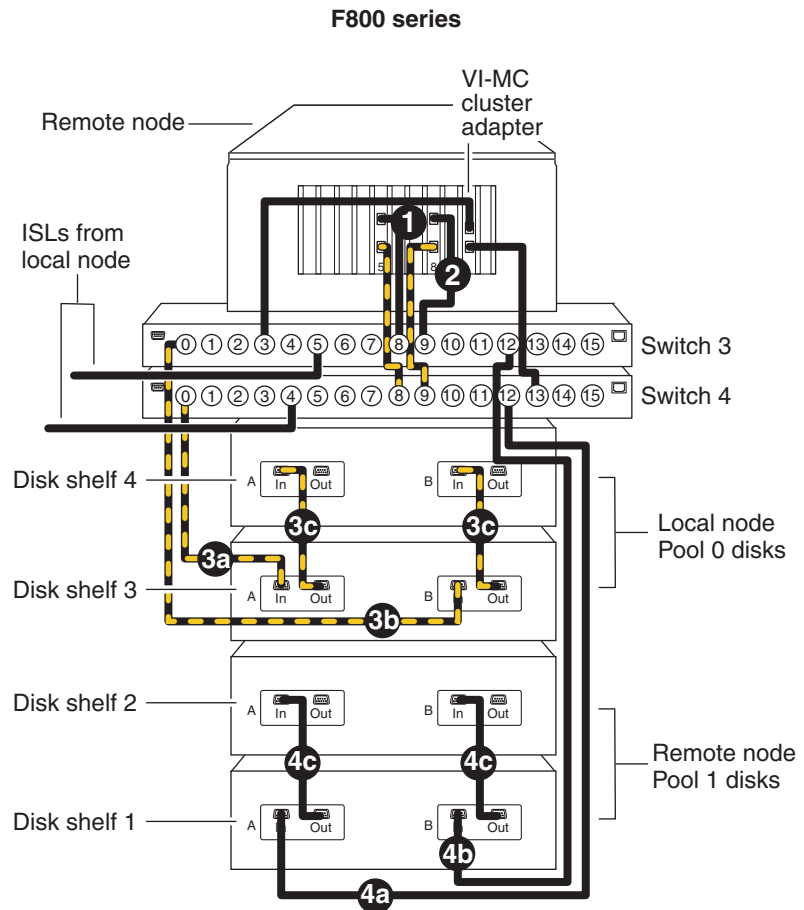
You need to cable the switches at the partner site so that they can work with the local switches.

To cable the partner site switches, complete the following steps.

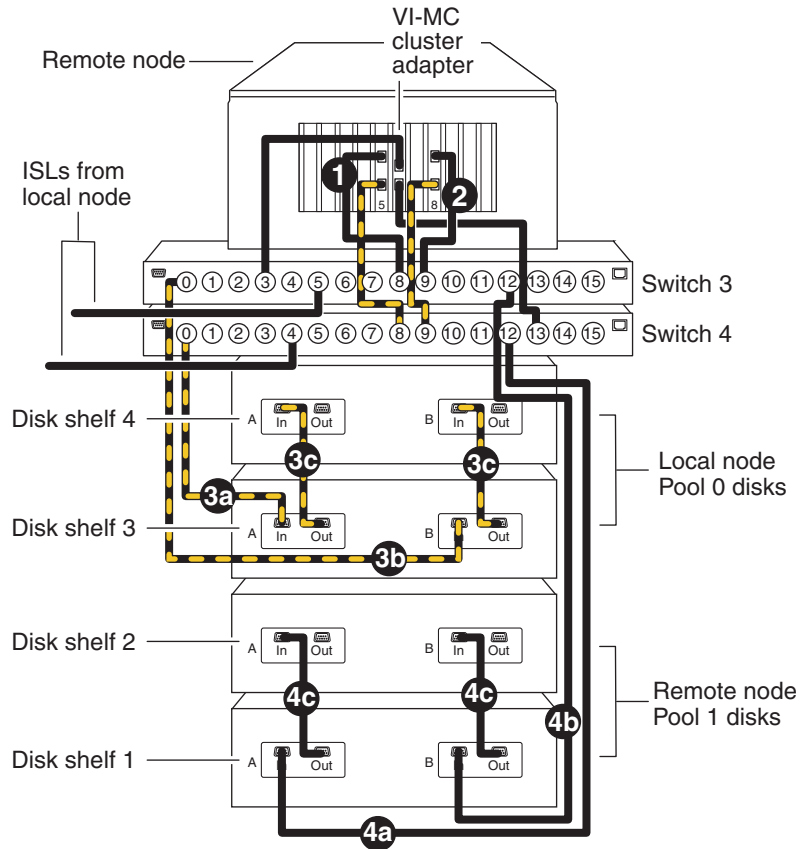
Step	Action
1	Log in to Switch 3 and set it up as described in “ Stage 2: Configuring the local switches ” on page 74.
2	Repeat Step 1 for Switch 4.

Cabling the remote node and disk shelves at the partner site

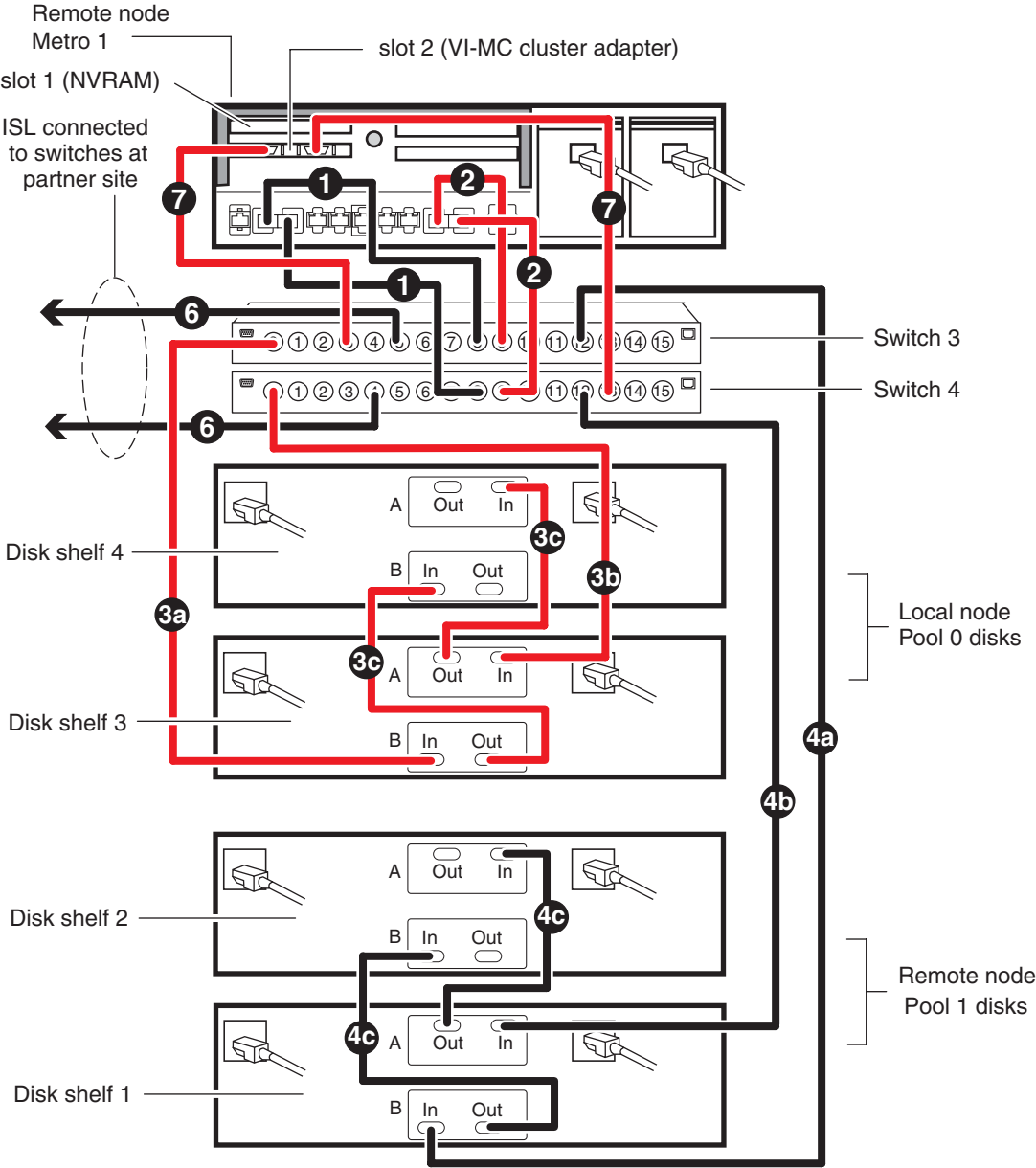
To cable the remote node, disk shelves, and switch, complete the following steps, using the applicable diagram for reference.



FAS900 series (except FAS980)



FAS3000 series



Step	Action
1	Cable the first Fibre Channel HBA in slot 5 of the remote node to Switch 3 and Switch 4. In the example, the top port of the Fibre Channel HBA is cabled to Port 8 of Switch 3 and the bottom port is cabled to Port 8 of Switch 4.
2	Cable the second Fibre Channel HBA in slot 8 of the remote node to Switch 3 and Switch 4. In the example, the top port of the Fibre Channel HBA is cabled to Port 9 of Switch 3 and the bottom port is cabled to Port 9 of Switch 4.
3	Connect the local node pool 0 disks to the switches by completing the following substeps: - Connect the Input port of the LRC, ESH, or ESH2 module Channel A of disk shelf 3, to port 0 on Switch 4. - Connect the Input port of the LRC, ESH, or ESH2 module Channel B of disk shelf 3, to port 0 on Switch 3. - Connect disk shelf 3 to disk shelf 4 by connecting the Output ports of the LRC, ESH, or ESH2 module of disk shelf 3 to the Input ports of the corresponding LRC, ESH, or ESH2 module of disk shelf 4. - Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the disk pool, then set the LRC or ESH terminate switches on the last disk shelf to On. Note _____ The ESH2 modules are self-terminating and therefore do not have a terminate switch. In the example, the pool 0 disk shelves are connected to switch port 0 on both Switch 3 and 4. However, you can use any open set of switch ports in this quadrant (limited to ports 0-3 in this configuration).

Step	Action
4	Connect the remote node pool 1 disks to the switches by completing the following substeps: a. Connect the Input port of the LRC, ESH, or ESH2 module Channel A on disk shelf 1 to port 12 on Switch 4. b. Connect the Input port of the LRC, ESH, or ESH2 module Channel B on disk shelf 1 to port 12 on Switch 3. c. Connect disk shelf 1 to disk shelf 2 by connecting the Output ports of the LRC, ESH, or ESH2 module of disk shelf 1 to the Input ports of the corresponding LRC, ESH, or ESH2 module of disk shelf 2. d. Set the terminate switches on the LRC or ESH modules to Off on all but the last disk shelf in the disk pool, then set the LRC or ESH terminate switches on the last disk shelf to On. Note _____ The ESH2 modules are self-terminating and therefore do not have a terminate switch. In the example, the pool 1 disk shelves are connected to switch port 12 on both Switch 3 and 4. However, you can use any open set of switch ports in this quadrant (limited to ports 12-15 in this configuration).

Step	Action																																																
5	Confirm that the disks are visible and have dual paths by entering the following command on the console: <pre>storage show disk -p</pre> Results: The output shows the disks connected to the switches, to what port they are connected, and to what disk and to which LRC, ESH, or ESH2 they belong, as shown in the following example: Example: <pre>ha16*> storage show disk -p</pre> <table><thead><tr><th>PRIMARY</th><th>PORT</th><th>SECONDARY</th><th>PORT</th><th>SHELF</th><th>BAY</th></tr><tr><th>-----</th><th>----</th><th>-----</th><th>----</th><th>-----</th><th>--</th></tr></thead><tbody><tr><td>switch3:0.40</td><td>A</td><td>switch4:0.40</td><td>B</td><td>5</td><td>0</td></tr><tr><td>switch4:0.41</td><td>B</td><td>switch3:0.41</td><td>A</td><td>5</td><td>1</td></tr><tr><td>.</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>.</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>.</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>switch3:12.52</td><td>B</td><td>switch4:12.52</td><td>A</td><td>6</td><td>4</td></tr></tbody></table>	PRIMARY	PORT	SECONDARY	PORT	SHELF	BAY	-----	----	-----	----	-----	--	switch3:0.40	A	switch4:0.40	B	5	0	switch4:0.41	B	switch3:0.41	A	5	1	.						.						.						switch3:12.52	B	switch4:12.52	A	6	4
PRIMARY	PORT	SECONDARY	PORT	SHELF	BAY																																												
-----	----	-----	----	-----	--																																												
switch3:0.40	A	switch4:0.40	B	5	0																																												
switch4:0.41	B	switch3:0.41	A	5	1																																												
.																																																	
.																																																	
.																																																	
switch3:12.52	B	switch4:12.52	A	6	4																																												
6	Connect the ISL from Switch 1 to Switch 3 and connect the ISL from Switch 2 to Switch 4.																																																
7	Connect the VI-MC cluster adapter to the switches, with the top port connecting to Switch 3 and the bottom port connecting to Switch 4. Note_____ The example shows the VI-MC cluster adapter connecting to switch ports 3 and 13, but you can use any available switch port. For a FAS3000 series MetroCluster, the NVRAM card cannot be used as a cluster adapter and must be moved to slot 1. _____																																																
8	Boot the node.																																																
9	Check to make sure that the node can see all disk shelves by entering the following command on the console: <pre>aggr status -r</pre>																																																
10	Configure the cluster, as described in Chapter 4, “ Configuring a Cluster ,” on page 95.																																																

Cabling a nonswitched MetroCluster

Cabling a nonswitched MetroCluster

To cable a nonswitched MetroCluster, complete the following steps.

Step	Action
1	Prepare your sites, as described in “ Preparing your equipment ” on page 32.

Step	Action	
2	If you are cabling for...	Then...
	A FAS3000 series	a. Cable the Channel A loops of the MetroCluster, as described in “Cabling the Channel A loops in a mirrored cluster” on page 38. b. Cable the Channel B loops of the MetroCluster, as described in “Cabling the Channel B loops in a mirrored cluster” on page 42. c. Cable the nodes together, as described in “Cabling the cluster adapters” on page 44.
	A FAS900 series	a. Cable the Channel A loops of the MetroCluster, as described in “Cabling the Channel A loops in a mirrored cluster” on page 50. b. Cable the Channel B loops of the MetroCluster, as described in “Cabling the Channel B loops in a mirrored cluster” on page 54. c. Cable the nodes together, as described in “Cabling the cluster adapters” on page 56.
	An F800 series	a. Cable the Channel A loops, as described in “Cabling the Channel A loops” on page 59. b. Cable the Channel B loops, as described in “Cabling the Channel B loops” on page 62. c. Cable the nodes together, as described in “Cabling the cluster adapters” on page 64.
3	Configure the cluster, as described in Chapter 4, “ Configuring a Cluster ,” on page 95.	

Converting a cluster into a Fabric Attached MetroCluster

Converting a cluster If you are upgrading an existing cluster to a MetroCluster configuration, you must upgrade disk firmware to the latest version. After upgrading disk firmware, you must power-cycle the affected disk drives before they work correctly in a MetroCluster fabric configuration. You can download the latest disk firmware from now.netapp.com.

To convert an existing cluster into a Fabric Attached MetroCluster, complete the following steps.

Step	Action	
1	Update Data ONTAP, storage system firmware, and disk firmware, as described in the <i>Data ONTAP Upgrade Guide</i> , making sure to shut down the nodes to the OK prompt.	
2	If you are converting...	Then...
	A FAS3000 system	Move the NVRAM5 adapter from slot 3 to slot 1, and add the VI-MC cluster adapter to slot 2.
	A FAS980 system	Move the NVRAM5 adapter from slot 11 to slot 10, and add the VI-MC cluster adapter to slot 11. For information about removing adapters, see “Installing a disk shelf in an existing mixed-shelf loop” on page 182.
	All other platforms	Update your hardware by replacing components, as described in “Installing a disk shelf in an existing mixed-shelf loop” on page 182. Note _____ You must have FC HBA for Disk adapters installed in each node, and each node must have VI-MC cluster adapters. _____

Step	Action	
3	Determine your switch and cluster configuration by completing the planning worksheet in “ Stage 1: Planning the Fabric Attached MetroCluster ” on page 73.	
4	Set up and configure the local switches, as described in “ Stage 2: Configuring the local switches ” on page 74.	
5	Cable the local node, as described in “ Stage 3: Cabling the local node ” on page 78.	
6	If you are converting a...	Then...
	Standard cluster	Set up mirroring and configure the local node, as described in “ Stage 3: Cabling the local node ” on page 78.
	Nonswitched MetroCluster	Go to Step 7 .
7	Transport the partner node, disk shelves, and switches to the remote location.	
8	Set up the remote node, disk shelves, and switch, as described in “ Stage 4: Cabling the switches and nodes at the partner site ” on page 84.	
9	Configure the cluster, as described in “ Configuring a Cluster ” on page 95.	

About this chapter

This chapter describes how to bring up a new standard or mirrored cluster for the first time. It describes how to enable licenses, set cluster options, configure networking for the cluster, and test the configuration.

Note

For information about configuring a cluster to use FCP, see “How Data ONTAP supports FCP with clustered systems” in the *Block Access Management Guide*.

Topics in this chapter

This chapter addresses the following topics:

- ◆ “[Stage 1: Bringing up the cluster for the first time](#)” on page 96
- ◆ “[Stage 2: Running diagnostics](#)” on page 101
- ◆ “[Stage 3: Enabling licenses](#)” on page 103
- ◆ “[Stage 4: Setting options and parameters](#)” on page 105
- ◆ “[Stage 5: Configuring network interfaces](#)” on page 110
- ◆ “[Stage 6: Checking and testing cluster configuration](#)” on page 112

Stage 1: Bringing up the cluster for the first time

What to do to bring up the cluster

After cabling the nodes and storage together, you must connect the cluster to the network, apply power to the system, and configure the nodes of the cluster.

Bringing up the new cluster

To bring up the cluster for the first time, complete the following steps.

Step	Action	
1	Connect both nodes to the network, as described in the hardware and service documentation.	
2	Turn on the power to the disk shelves and wait 30 seconds for the electronics to complete initialization.	
3	If you are configuring...	Then...
	Storage systems that were already in use	1. Turn on the power to both storage systems and wait for the system prompt. Attention _____ Press the Delete key to stop booting if the system starts to boot. 2. Proceed to “Stage 2: Running diagnostics” on page 101.
	New storage systems	Apply power; the storage systems enter the setup program automatically. Proceed to Step 4.

Step	Action
4	Use the <code>setup</code> program to configure both storage systems. For more information about running setup on a system set up for clustering, see “About running setup on clustered systems” on page 97. For more information about running setup, see the <i>Software Setup Guide</i>.

About running setup on clustered systems

When the `setup` program runs on a storage system that is set up for clustering, setup prompts you to answer some questions specific for clusters. The following list outlines some of the questions about your installation that you should think about before proceeding through the `setup` program:

- ◆ Do you want to configure virtual interfaces (VIFs) for your network interfaces?

For information about VIFs, see the *Network Management Guide*.

Note

You are advised to use VIFs with clusters to reduce SPOF (single-points-of-failure).

- ◆ How do you want to configure your interfaces for takeover?
 - ❖ For information about configuring network interfaces for clusters, see “[Configuring network interfaces for clusters](#)” on page 143.
 - ❖ For information about using setup to configure shared interfaces, see “[Using setup to configure shared interfaces](#)” on page 98.
 - ❖ For information about using setup to configure dedicated interfaces, see “[Using setup to configure dedicated interfaces](#)” on page 99.
 - ❖ For information about using setup to configure standby interfaces, see “[Using setup to configure standby interfaces](#)” on page 99.

Note

If you do not want to configure your network for clustering when you run setup for the first time, you can configure it later in the cluster configuration process, either by running setup again, or by using the `ifconfig` command and editing the `/etc/rc` file manually. However, you must provide at least one local IP address to exit setup.

Using setup to
configure shared
interfaces

To use setup to configure shared interfaces, complete the following steps.

Step	Action
1	Enter the IP address for the interface you are configuring. Example: Please enter the IP address for Network Interface e0 [] : <i>nnn.nn.nn.nnn</i> <i>nnn.nn.nn.nnn</i> is the local address for the node you are configuring.
2	Enter the netmask for the interface you are configuring, or press Return if the default value is correct. Example: Please enter the netmask for Network Interface e1 [255.255.0.0] :
3	Specify that this interface is to take over a partner IP address. Example: Should interface e1 take over a partner IP address during failover? [n] : y
4	Enter the IP address or interface name of the partner. Example: Please enter the IP address or interface name to be taken over by e1 [] : <i>nnn.nn.nn.nnn</i> Note_____ If the partner is a VIF, you must the interface name. _____

Using setup to configure dedicated interfaces

To use setup to configure dedicated interfaces, complete the following steps.

Step	Action
1	Enter the IP address for the interface you are configuring. Example: Please enter the IP address for Network Interface e0 [] : <i>nnn.nn.nn.nnn</i> <i>nnn.nn.nn.nnn</i> is the local address for the node you are configuring.
2	Enter the netmask for the interface you are configuring, or press Return if the default value is correct. Example: Please enter the netmask for Network Interface e1 [255.255.0.0] :
3	Specify that this interface does not take over a partner IP address. Example: Should interface e1 take over a partner IP address during failover? [n] : n

Using setup to configure standby interfaces

To use setup to configure standby interfaces, complete the following steps.

Step	Action
1	Do not enter an IP address for a standby interface, press Return. Example: Please enter the IP address for Network Interface e0 [] :
2	Specify that this interface is to take over a partner IP address. Example: Should interface e1 take over a partner IP address during failover? [n] : y

Step	Action
3	Enter the IP address or interface name of the partner. Example: Please enter the IP address or interface name to be taken over by e1 []: <i>nnn.nn.nn.nnn</i> Note If the partner is a VIF, you must use the interface name.

Stage 2: Running diagnostics

About diagnostics

Running system diagnostics helps you ensure that your storage systems are in good working order before configuring your cluster.

Running system diagnostics

To run system diagnostics on your storage systems, complete the following steps.

Step	Action	
1	If the nodes of your cluster...	Then...
	Use a CompactFlash™ card, and are CFE (Common Firmware Environment) based systems (for example, FAS200 series)	Run the diagnostics program from the CompactFlash card by entering the following command at the prompt: <code>boot_diags</code>
	Use a CompactFlash card, and are not CFE based (for example, F825, F880, FAS900 series)	Run the diagnostics program from the CompactFlash card by entering the following command at the prompt: <code>boot_diags</code>
	Do not use a CompactFlash card (for example, F810, F820, F840)	Run the diagnostics program from the motherboard flash by entering the following command at the prompt: <code>sys-diags</code>
2	Select the diagnostics tests that check all dual-port FC-AL adapters and the cluster adapter. For a description of these tests, see the Diagnostics Guide at http://now.netapp.com .	
3	Boot both nodes and observe the console output by entering the following command from each node's console: <code>boot</code>	

Step	Action
4	Check the status of the cluster by entering the following command on each node's console: <pre>cf status</pre> Result: The following message is displayed on the console, indicating that clustering is not activated: <pre>cluster monitor not initialized</pre>
5	Go to “ Stage 3: Enabling licenses ” on page 103.

Stage 3: Enabling licenses

Licenses you need to add

The cluster licenses you need to add depend on your cluster configuration. The following table outlines the required licenses for each cluster configuration.

Cluster configuration	Required licenses
Standard (CFO) cluster	◆ cluster
Mirrored cluster	◆ cluster ◆ snapmirror_local
MetroCluster	◆ cluster ◆ snapmirror_local ◆ cluster_remote

Note

For more information about adding licenses, see the *System Administration Guide*.

Enabling licenses on the cluster

To enable a license on the cluster, complete the following steps.

Step	Action
1	For each required license, enter the license and code on both node consoles, as shown in the following example: <code>license add xxxxxx</code> xxxxx is the license code you received for the feature.
2	Reboot both nodes using the following command: <code>reboot</code>
3	Enable clustering on each node by entering the following command on the local node console: <code>cf enable</code>

Step	Action
4	Verify that clustering is enabled by entering the following command on each node console: <pre>cf status</pre> Cluster enabled, filer2 is up.
5	Repeat Step 1 for any other licenses that you need to enable, using the license type and code for each licensed product installed on the cluster.
6	Go to “Stage 4: Setting options and parameters” on page 105.

Stage 4: Setting options and parameters

What the options do Options help you maintain various functions of your node, such as security, file access, and network communication. During takeover, the value of an option might be changed by the node doing the takeover. This can cause unexpected behavior during a takeover. To avoid unexpected behavior, specific option values must be the same on both the local and partner node.

Each option performs a specific task or has a specific function. See the `na_options(1)` man page for a description of available options and the function each performs.

Checking options and parameters with the cluster configuration checker

The easiest way to check options and parameters for your cluster is to run the Cluster Configuration Checker. For more information, see [“Adding disk shelves to a cluster”](#) on page 178.

After you correct any errors displayed by the Configuration Checker, proceed to [“Stage 5: Configuring network interfaces”](#) on page 110.

About managing options between the nodes

In a cluster, options are one of the following types:

- ◆ Options that must be the same on both two nodes for the cluster to function correctly
See [“Options that must be the same on both nodes”](#) on page 106.
- ◆ Options that might be overwritten on the node that is failing over
These options must be the same on both nodes to avoid losing system state after a failover. See [“Options that must be the same on both nodes”](#) on page 106.
- ◆ Options that are recommended to be the same on both nodes so that system behavior does not change during failover
See [“Options that are recommended to be the same”](#) on page 107.
- ◆ Options that can be different on each node

Note

You can find out whether an option must be the same on both nodes of a cluster from the comments that accompany the option value when you enter the `option` command. No comments indicates the option can be different on each node.

Options that must be the same on both nodes

The options that must be the same on both nodes for correct system operation or to avoid losing state after failover are listed in the following table.

Option name	Option name
auditlog.enable	raid.disk.copy.auto.enable
auditlog.max_file_size	ip.tcp.offload.enable
autologout.console.enable	ip.tcp.sack.enable
autologout.console.timeou	ip.v6.enable
autologout.telnet.enable	locking.grace_lease_seconds
autologout.telnet.timeout	lun.use_partner.cc.enable
autosupport.retry.count	ndmpd.preferred_interface
autosupport.retry.interval	pcnfsd.enable
cf.quickloop.enable	raid.media_scrub.enable
coredump.dump.attempts	raid.media_scrub.rate
disk.auto_assign	raid.media_scrub.spares.enable
disk.maint_center.allowed_entries	raid.mirror_read_plex_pref
disk.maint_center.enable	raid.reconstruct.perf_impact
disk.maint_center.max_disks	raid.reconstruct.wafliron.enable
disk.maint_center.spares_check	raid.resync.perf_impact
dns.cache.enable	raid.rpm.ata.enable
dns.domainname	raid.rpm.fcsl.enable
dns.enable	raid.scrub.perf_impact
dns.update.enable	raid.timeout
dns.update.ttl	raid.verify.perf_impact
ftpd.idle_timeout	rlm.setup
ftpd.max_connections	rmc.lan.smtp_ip
ftpd.max_connections_threshold	rmc.setup

Option name	Option name
httpd.log.format	snaplock.autocommit_period
httpd.timeout	snaplock.compliance.write_verify
httpd.timewait.enable	snapmirror.vbn_log_enable
ip.drd.ping_interval	snmp.access
ip.drd.ping_retry	snmp.enable
ip.fastpath.enable	tape.reservations
ip.icmp_ignore_redirect.enable	telnet.access
ip.match_any_ifaddr	telnet.distinct.enable
ip.max_vip	telnet.enable
ip.path_mtu_discovery.enable	trusted.hosts
ip.ping_throttle.alarm_interval	vol.copy.throttle
ip.ping_throttle.drop_level	wafl.group_cp
ip.rip.update_interval	wafl.maxdirsize
ip.tcp.newreno.enable	wafl.root_only_chown

Options that are recommended to be the same

The options that are recommended to be the same to avoid differences in system functionality during failover are listed in the following table.

Option name	Option name
cf.takeover.on_network_interface_fail.policy	timed.proto
timed.enable	timed.sched
timed.log	timed.servers
timed.max_skew	timed.window

Manually setting options

To manually set options, complete the following steps.

Step	Action
1	View and note the values of the options on the local and partner nodes, using the following command on each console: options Result: The current option settings for the node are displayed on the console. Output similar to the following is displayed: <pre>autosupport.doit DONT autosupport.enable off . . .</pre>
2	Verify that the options that have comments in parentheses are set to the same value for both nodes. The comments are as follows: ◆ Value might be overwritten in takeover ◆ Same value required in local+partner ◆ Same value in local+partner recommended
3	Correct any mismatched options using the following command: options option_name option_value Note_____ See the na_options man page for more information about the options. _____

What the parameters do

The parameters help keep each node synchronized and help maintain routing of processes throughout the system.

Parameters that must be the same

The parameters listed in the following table must be the same so that takeover is smooth and data is transferred between the nodes correctly.

Parameter	Setting for...
date	date, rdate

Parameter	Setting for...
NDMP (on or off)	ndmp (on or off)
route table published	route
routed enabled	routed (on or off)
Time zone	timezone

Stage 5: Configuring network interfaces

When this stage is necessary

If you configured your interfaces using setup when you first applied power to your storage systems, you do not need to configure them again. If you want to configure your interfaces for automatic takeover, complete the steps in [“Configuring dedicated and standby interfaces”](#) on page 151. Otherwise, you can proceed to [“Stage 6: Checking and testing cluster configuration”](#) on page 112.

Note

For information about configuring a cluster to use FCP, see “How Data ONTAP supports FCP with clustered systems” in the *Block Access Management Guide*.

What the networking interfaces do

When a node in a cluster fails, the surviving node must be able to assume the identity of the failed node on the network. Networking interfaces allow individual nodes in the cluster to maintain communication with the network if the partner fails. See the *Data ONTAP Network Management Guide* for a description of available options and the function each performs.

Note

Be sure you read and understand the material in [“Configuring network interfaces for clusters”](#) on page 143 before configuring your cluster’s network interfaces.

You should always use multiple NICs with VIFS to improve networking availability for both single storage systems and clustered storage systems.

Network connection requirements

Both nodes in the cluster must have interfaces that access the same collection of networks and subnetworks.

Information needed for configuration

You must gather the following information before configuring the interfaces:

- ◆ The IP address for both the local node and partner node
- ◆ The netmask for both the local node and partner node

Configuring networking for the cluster

To configure the networking interfaces on a cluster, complete the following steps.

Step	Action
1	Determine whether to use shared interfaces or dedicated and standby interfaces. For more information, see “ Configuring network interfaces for clusters ” on page 143.
2	Configure your interfaces on one node using the procedures outlined in “ Configuring shared interfaces ” on page 148 or “ Configuring dedicated and standby interfaces ” on page 151.
3	Repeat Step 2 on the other node.
4	If you want to configure automatic takeover for any of your interfaces or VIFs, complete the steps outlined in “ Configuring dedicated and standby interfaces ” on page 151.

Stage 6: Checking and testing cluster configuration

Testing takeover and giveback

To test takeover and giveback, complete the following steps.

Step	Action
1	Check the cabling on the cluster interconnect cables to make sure that they are secure.
2	Verify that you can create and retrieve files on both nodes for each licensed protocol.
3	Enter the following command from the local node console: <pre>cf takeover</pre> Result: The local node takes over the partner node and gives the following output: <pre>takeover completed</pre>

Step	Action
4	Test communication between the local node and partner node. For example, you can use the <code>fcstat device_map</code> command to ensure that one node can access the other node's disks. Example: <pre>filer-1> fcstat device_map Loop Map for channel 8: Translated Map: Port Count 32 7 16 17 18 19 32 33 34 35 48 49 50 51 24 26 27 57 61 42 43 44 45 28 29 30 52 53 54 37 38 21 22 Shelf mapping: Shelf 2: 22 21 XXX 19 18 17 16 Shelf 3: 30 29 28 27 26 XXX 24 Shelf 4: 38 37 XXX 35 34 33 32 Shelf 5: XXX 45 44 43 42 XXX XXX Shelf 6: 54 53 52 51 50 49 48 Loop Map for channel 5: Translated Map: Port Count 22 7 38 35 22 20 19 30 29 28 27 14 41 42 43 26 25 24 18 16 34 33 32 Shelf mapping: Shelf 2: 22 XXX 20 19 18 XXX 16 Shelf 3: 30 29 28 27 26 25 24 Shelf 4: 38 XXX XXX 35 34 33 32 Shelf 5: XXX XXX XXX 43 42 41 XXX Loop Map for channel 7a: Translated Map: Port Count 10 7 40 42 43 44 45 46 58 59 60 Shelf mapping: Shelf 5: 46 45 44 43 42 XXX 40 Shelf 7: XXX XXX 60 59 58 XXX XXX Loop Map for channel 7b: Translated Map: Port Count 13 7 16 17 18 19 20 32 33 34 35 36 37 38 Shelf mapping: Shelf 2: XXX XXX 20 19 18 17 16 Shelf 4: 38 37 36 35 34 33 32</pre>

Step	Action	
5	Give back the partner node by entering the following command: <pre>cf giveback</pre> Result: The local node releases the partner node, which reboots and resumes normal operation. The following message is displayed on the console when the process is complete: <pre>giveback completed</pre>	
6	If takeover and giveback...	Then...
	Is completed successfully	Repeat Step 2 through Step 5 on the partner node.
	Fails	Go to “ Correcting takeover or giveback failures ” on page 155.

About this chapter

This chapter discusses the different modes clusters can be in and some common management tasks for clusters.

Topics in this chapter

This chapter discusses the following topics:

- ◆ [“How takeover and giveback work”](#) on page 116
- ◆ [“Managing a cluster in normal mode”](#) on page 118
- ◆ [“Configuring when takeover occurs”](#) on page 125
- ◆ [“Managing a cluster in takeover mode”](#) on page 129
- ◆ [“Managing an emulated node”](#) on page 130
- ◆ [“Backing up and restoring from the emulated node”](#) on page 135
- ◆ [“Managing giveback”](#) on page 136
- ◆ [“Configuring network interfaces for clusters”](#) on page 143
- ◆ [“Checking cluster configuration”](#) on page 154
- ◆ [“Correcting takeover or giveback failures”](#) on page 155
- ◆ [“Removing clustering”](#) on page 157

How takeover and giveback work

When takeovers occur

The conditions under which takeovers occur depend on how you configure the cluster. For more information about configuring a cluster for takeover, see [“Configuring when takeover occurs”](#) on page 125. Takeovers can be initiated when one of the following conditions occur:

- ◆ A node is in a cluster that is configured for immediate takeover on panic and undergoes a software or system failure that leads to a panic.
- ◆ A node that is in a cluster undergoes a system failure (for example, a loss of power) and cannot reboot.

Note

If the storage for a node also loses power at the same time, a standard takeover is not possible. For MetroClusters, you can initiate a forced takeover in this situation.

- ◆ There is a mismatch between the disks that one node can see and the disks that the other node can see.
- ◆ One or more network interfaces that are configured to support failover become unavailable.
- ◆ A node cannot send heartbeat messages to its partner. This could happen if the node experienced a hardware or software failure that did not result in a panic but still prevented it from functioning correctly.
- ◆ You halt one of the clustered nodes without using the `-f` flag.
- ◆ You initiate a takeover manually.

What happens during takeover

When a takeover occurs, the unimpaired partner node in the cluster takes over the functions and disk drives of the failed node by creating an emulated storage system that

- ◆ Assumes the identity of the failed node
- ◆ Accesses the failed node's disks and serves its data to clients

The partner node maintains its own identity and its own primary functions, but also handles the added functionality of the failed node through the emulated node.

Note

When a takeover occurs, existing CIFS sessions are terminated. A graceful shutdown of the CIFS sessions is not possible, and some data loss could occur for CIFS users.

What happens after takeover

After a takeover occurs, you view the surviving partner as having two identities, its own and its partner's, that exist simultaneously on the same storage system. Each identity can access only the appropriate volumes and networks. You can send commands or log in to either storage system by using the `rsh` command, allowing remote scripts that invoke storage system commands through a Remote Shell connection to continue to operate normally.

Access with `rsh`: Commands sent to the failed node through a Remote Shell connection are serviced by the partner node, as are `rsh` command login requests.

Access with Telnet: If you log in to a failed node through a Telnet session, you see a message alerting you that your storage system failed and to log in to the partner node instead. If you are logged in to the partner node, you can access the failed node or its resources from the partner node by using the `partner` command.

What happens during a giveback

When the failed node is functioning again

- ◆ You initiate a `giveback` command that terminates the emulated node on the partner
- ◆ The failed node resumes normal operation, serving its own data
- ◆ The cluster resumes normal operation, with each node ready to take over for its partner if the partner fails

Managing a cluster in normal mode

Normal mode management tasks

You manage a cluster in normal mode by performing the following actions:

- ◆ Monitoring cluster status
- ◆ Viewing information about the cluster
 - ❖ Displaying the partner’s name
 - ❖ Displaying disk information
- ◆ Enabling and disabling takeover
- ◆ Enabling and disabling immediate takeover of a panicked partner
- ◆ Halting a node without takeover
- ◆ Performing a takeover

Monitoring cluster status

To determine whether the cluster feature is enabled and whether the other node in the cluster is up, complete the following step.

Step	Action
1	Enter the following command: <pre>cf status</pre> Example: <pre>filer1> cf status</pre> Cluster enabled, filer2 is up. This verifies the link between the clustered nodes and tells you that both filer1 and filer2 are functioning and available for takeover. Result: If Data ONTAP returns “Cluster disabled,” you disabled the cluster feature. You can reenable clustering using the <code>cf enable</code> command. Note _____ Data ONTAP can disable cluster failover if a software or hardware problem exists that prevents a successful takeover. In this case, the message returned from the <code>cf status</code> command describes the reason why clustering is disabled. _____

Description of cluster status messages

The following table shows some of the messages that the `cf status` command can display.

Message	Meaning
Cluster enabled, <i>partner_name</i> is up.	The cluster is operating normally.
<i>partner_name_1</i> has taken over <i>partner_name_2</i> .	One node in the cluster took over the other node in the cluster.
Interconnect not present.	The system does not recognize the existence of a cluster interconnect adapter.
Interconnect is down.	The cluster interconnect adapter cannot access the partner. This might be due to cabling problems or the partner might be down.
Interconnect is up.	The cluster interconnect adapter is active and can transmit data to the partner.
<i>partner_name_1</i> has detected a mailbox disk error, take over of <i>partner_name_2</i> disabled.	One node in the cluster cannot access multiple mailbox disks. Check access to both the local and partner root volumes and mirrors, if they exist. Also check for disk or FC-AL problems or offline storage adapters.
<i>partner_name_2</i> may be down and has disabled take over by <i>partner_name_1</i> .	One node in the cluster might be down.
Version mismatch	The partner node has an incompatible version of Data ONTAP.
<i>partner_name_1</i> is attempting takeover of <i>partner_name_2</i> . takeover is in module <i>n</i> of <i>N</i> modules.	A takeover is being attempted (includes information about how far the takeover has progressed).

Message	Meaning
<i>partner_name_1</i> has taken over <i>partner_name_2</i> , giveback in progress. giveback is in module <i>n</i> of <i>N</i> modules.	A giveback is being attempted (includes information about how far the giveback has progressed).
<i>partner_name_1</i> has taken over <i>partner_name_2</i> , <i>partner_name_2</i> is ready for giveback.	The takeover node received information that the failed node is ready for giveback.
<i>partner_name_1</i> has taken over <i>partner_name_2</i> , <i>partner_name_2</i> is ready for giveback. Automatic giveback is disabled due to exceeding retry count.	The takeover node received information that the failed node is ready for giveback, but giveback cannot take place because the number of retries exceeded the limit.

Displaying the partner's name

To determine the name of the other node in the cluster, complete the following step.

Step	Action
1	Enter the following command: <pre>cf partner</pre> Note _____ If the node does not yet know the name of its partner, because the cluster is new, this command returns “partner”. _____

Displaying disk information on a cluster

In a cluster, the `sysconfig` and `aggr status` command outputs are special in that they display information about both nodes.

For each node, the `sysconfig` command output displays disks on both FC-AL loop A and FC-AL loop B:

- ◆ The information about disks on FC-AL loop A is the same as for storage systems not in a cluster.

- ◆ The information about disks on FC-AL loop B is for hardware only; the `sysconfig` command displays information about the adapters supporting the disks. The command does not show whether a disk on FC-AL loop B is a file system disk, spare disk, or parity disk.

To display disk information for both nodes in a cluster, complete the following step.

Step	Action
1	Enter one of the following commands: <pre>sysconfig -r</pre> or <pre>aggr status -r</pre>

Enabling and disabling takeover

You might want to disable takeover if you are doing maintenance that typically causes a takeover. You can reenable takeover after you finish maintenance.

To enable or disable takeover on a cluster, complete the following step.

Step	Action
1	Enter the following command: <pre>cf enable disable</pre> Use <code>cf enable</code> to enable takeover or <code>cf disable</code> to disable takeover. Note _____ You can enable or disable takeover from either node in the cluster. _____

Enabling and disabling immediate takeover of a panicked partner

You can enable a clustered node to take over immediately when its partner panics. This shortens the time between the initial failure and when service is fully restored, because a cluster can fail over in less time than it takes for a storage system to recover from a panic. However, the subsequent giveback causes another brief outage, so if service disruption is an issue (for example, if you have CIFS clients), you might not want to incur the initial takeover.

Note_____

If the panicked node needs to do a core dump and a spare disk is available, the core dump is directed to the spare disk and the partner takes over immediately (within 60 seconds) without waiting until the core dump is completed.

If the panicked node needs to do a core dump and no spare disk is available, the node does not perform a core dump; instead, it prints a message that a spare disk is needed. When the node reboots, it logs that no core dump took place.

To enable or disable immediate takeover of a panicked partner, complete the following steps.

Step	Action
1	Ensure that you enabled cluster takeover with the <code>cf enable</code> command.
2	Enter the following command: <code>options cf.takeover.on_panic [on off]</code> Use <code>on</code> to enable immediate takeover of a failed partner or <code>off</code> to disable immediate takeover. If this option is disabled, normal takeover procedures apply. Note_____ If you enter this command on one node in a cluster, the value applies to both nodes in the cluster. The setting of this option is persistent across reboots.

Halting a node without takeover

You can halt the node and prevent its partner from taking over. For example, you might need to perform maintenance on both the storage system and its disks and want to avoid an attempt by the partner node to write to those disks.

To halt a node in a cluster without causing a takeover, complete the following step.

Step	Action
1	Enter the following command: <code>halt -f</code>

Commands for performing a takeover

You can initiate a takeover on a node in a cluster to perform maintenance on that node while still serving the data on its disks to users. The following table lists and describes the commands you can use when initiating a takeover.

Command	Description
<code>cf takeover</code>	Initiates a takeover of the partner of the local node. Takeover is aborted if a core dump is in progress on the partner (if the <code>cf.takeover.on_panic</code> option is set to <code>off</code>). The takeover starts either after the partner halts successfully or after a timeout.
<code>cf takeover -f</code>	Initiates an immediate takeover of the partner of the local node regardless of whether the other node is dumping its core. The partner node is not allowed to halt gracefully.

Command	Description
<code>cf forcetakeover</code>	Tells the cluster monitor to ignore some configuration problems that would otherwise prevent a takeover, such as unsynchronized NVRAM due to a faulty cluster interconnect connection. It then initiates a takeover of the partner of the local node. Attention — Use <code>cf forcetakeover</code> only in certain controlled circumstances, such as when you cannot get <code>cf takeover</code> to succeed. You can lose up to 10 seconds worth of data that was committed to NVRAM but not to disk when you use this command.
<code>cf forcetakeover -d</code>	Initiates a takeover of the local partner even in the absence of a quorum of partner mailbox disks. The <code>cf forcetakeover -d</code> command is valid only if the <code>cluster_remote</code> license is enabled. Attention — Use the <code>-d</code> option only after you verify that the partner is down. Note — The <code>-d</code> option is used in conjunction with RAID mirroring to recover from disasters in which one partner in a cluster is not available. For more information, see the <i>Data Protection Online Backup and Recovery Guide</i>.
<code>cf forcetakeover -n</code>	Initiates a takeover for a nondisruptive upgrade. For more information, see the <i>Upgrade Guide</i>.

Configuring when takeover occurs

Reasons for takeover

Takeovers can happen for several different reasons. Some system errors must cause a takeover; for example, when a clustered storage system loses power, it automatically fails over to the other cluster node.

However, for some system errors, a takeover is optional, depending on how you configure your cluster. The following table outlines which system errors can cause a takeover to occur, and whether you can configure the cluster for that error.

System error	Option used to configure	Default value	Notes
A clustered system undergoes a system failure and cannot reboot.	<code>cf.takeover.on_failure</code> set to On	On	You should leave this option enabled unless instructed otherwise by technical support.
A clustered system undergoes a software or system failure leading to a panic.	<code>cf.takeover.on_panic</code> set to On	Off, unless: ◆ FCP or iSCSI is licensed ◆ the system is a FAS270	For more information, see “Enabling and disabling immediate takeover of a panicked partner” on page 122.
There is a mismatch between the disks that one clustered system can see and the disks that the other system can see.	<code>cf.takeover.on_disk_shelf_miscompare</code> set to On	Off	For more information, see “How disk shelf comparison takeover works” on page 127.

System error	Option used to configure	Default value	Notes
All the network interface cards (NICs) or VIFS enabled for negotiated failover on a node failed.	<code>cf.takeover.on_network_interface_failure</code> set to <code>On</code> , <code>cf.takeover.on_network_interface_failure.policy</code> set to <code>all_nics</code>	By default, takeover on network failure is disabled.	To enable a network interface for negotiated failover, you use the <code>ifconfig if_name nfo</code> command. For more information, see the <i>Network Management Guide</i> .
One or more of the NICs or VIFS enabled for negotiated failover failed.	<code>cf.takeover.on_network_interface_failure</code> set to <code>On</code> , <code>cf.takeover.on_network_interface_failure.policy</code> set to <code>any_nic</code>	By default, takeover on network failure is disabled.	To enable a network interface or VIF for negotiated failover, you use the <code>ifconfig if_name nfo</code> command. For more information, see the <i>Network Management Guide</i> .
A clustered system fails within 60 seconds of booting up.	<code>cf.takeover.on_short_uptime</code> set to <code>On</code>	<code>On</code>	Changing the value of this option on one node automatically updates the option on the partner node.
A clustered system cannot send heartbeat messages to its partner.	n/a		You cannot prevent this condition from causing a takeover.
You halt one of the clustered systems without using the <code>-f</code> flag.	n/a		You cannot prevent this condition from causing a takeover.
You initiate a takeover manually using the <code>cf takeover</code> command.	n/a		You cannot prevent this condition from causing a takeover.

Specifying the time period before takeover

You can specify how long (in seconds) a partner in a cluster can be unresponsive before the other partner takes over.

To specify the time period before takeover, complete the following step.

Step	Action
1	Enter the following command: <code>options cf.takeover.detection.seconds number_of_seconds</code> The valid values for <i>number_of_seconds</i> are 10 through 180; the default is 15.

Both partners do not need to have the same value for this option. Thus, you can have one partner that takes over more quickly than the other.

Note_____

If your cluster is failing over only because one of the nodes is too busy to respond to its partner, increase the value of the `cf.takeover.detection.seconds` option on the partner.

Attention_____

If the specified time is less than 15 seconds, unnecessary takeovers can occur. Use caution when assigning a takeover time of less than 15 seconds.

How disk shelf comparison takeover works

When communication between clustered nodes is first established through the cluster interconnect adapters, the nodes exchange a list of disk shelves that are visible on the A and B loops of each node. If, later, a system sees that the B loop shelf count on its partner is greater than its local A loop shelf count, the system concludes that it is impaired and prompts its partner to initiate a takeover.

Note_____

Disk shelf comparison does not function for clusters using software disk ownership, or for Fabric Attached MetroClusters.

Configuring VIFs or interfaces for automatic takeover

After you configure your interfaces or VIFs to allow takeovers and givebacks to be completed successfully, you can also optionally configure them to trigger automatic takeover if any or all of them experience a non-transient failure.

To configure a VIF or interface for automatic takeover, complete the following steps.

Step	Action
1	For every VIF or interface on which you want to enable automatic takeover, enter the following command: <pre>ifconfig interface_name nfo</pre> For more information about this command, see the <code>na_ifconfig</code> man page or the <i>Network Management Guide</i>. For more information about configuring automatic takeover, see “Configuring when takeover occurs” on page 125.
2	Update the <code>/etc/rc</code> file so that your changes persist across reboots.
3	The default policy is that takeover only occurs if <i>all</i> the NICs or VIFs on a node that are configured for automatic takeover fail. If you want takeover to occur if <i>any</i> NIC or VIF configured for automatic takeover fails, enter the following command: <pre>options cf.takeover.on_network_interface_failure.policy any_nic</pre>
4	Enter the following command to enable the feature: <pre>options cf.takeover.on_network_interface_failure enable</pre>

Managing a cluster in takeover mode

Takeover mode prompt

In takeover mode, the prompt looks like this:

```
filer1(takeover) >
```

Determining why takeover occurred

To determine why a takeover occurred, complete the following step.

Step	Action
1	Enter the following command: <pre>cf status</pre> Results: This command can display the following information: ◆ Whether the cluster is enabled or disabled ◆ Whether a takeover is imminent due to a negotiated failover ◆ Whether a takeover occurred, and the reason for the takeover

Statistics in takeover mode

In takeover mode, statistics for some commands differ from the statistics in normal mode in the following ways:

- ◆ Each display reflects the sum of operations that take place on the takeover node plus the operations on the failed node. The display does not differentiate between the operations on the takeover node and the operations on the failed node.
- ◆ The statistics displayed by each of these commands are cumulative.
- ◆ After giving back the failed partner’s resources, the takeover node does not subtract the statistics it performed for the failed node in takeover mode.
- ◆ The giveback does not reset (zero out) the statistics.

To get accurate statistics from a command after a giveback, you can reset the statistics as described in the man page for the command you are using.

Note

You can have different settings on each node for SNMP options, but any statistics gathered while a node has been taken over do not distinguish between nodes.

Managing an emulated node

What an emulated node is

An emulated node is a software copy of the failed node that is hosted by the takeover node. You access the emulated node in partner mode by using the `partner` command, as described in the following sections.

Management exceptions for emulated nodes

You manage an emulated node as you do any other storage system, including managing disks, with the following exceptions, which are described in greater detail later in this section:

- ◆ An emulated node can access only its own disks.
- ◆ Some commands are unavailable.
- ◆ Some displays differ from normal displays.

Accessing the emulated node from the takeover node

You access the emulated node from the takeover node in takeover mode with the `partner` command. You can issue the `partner` command in two forms:

- ◆ The `partner` command without an argument toggles between partner mode, in which you manage the emulated node, and takeover mode, in which you manage the takeover node.
- ◆ The `partner` command with a Data ONTAP command as an argument executes the command on the emulated node in partner mode and then returns to takeover mode.

Using the partner command without arguments: To toggle between partner mode and takeover mode, complete the following step.

Step	Action
1	From the takeover prompt, enter the following command: partner Results: The prompt changes to the partner mode prompt, which has the following form: <i>emulated_node/takeover_node></i> Example: The following example shows the change from takeover mode to partner mode and back: <pre>filer1(takeover)> partner Login from console: filer2 Thu Aug 20 16:44:39 GMT [filer1: rc]: Login from console: filer2 filer2/filer1> partner Logoff from console: filer2 filer1(takeover)> Thu Aug 20 16:44:54 GMT [filer1: rc]: Logoff from console: filer2 filer1(takeover)></pre>

Using the partner command with arguments: To use the partner command with a Data ONTAP command as an argument, complete the following step.

Step	Action
1	From the takeover prompt, enter the following command: partner <i>command</i> <i>command</i> is the command you want to initiate on the emulated node. Example: <code>filer1(takeover)> partner cf status</code> filer2 has been taken over by filer1. filer1(takeover)>

Accessing the emulated node remotely

You can also access the emulated node remotely.

Using Telnet: With Telnet, you can directly access only the takeover node.

To access the emulated node by using a Telnet session, complete the following steps.

Step	Action
1	Enter the following command: <code>telnet takeover_node</code>
2	Log in to the takeover node.
3	Enter the following command: <code>partner</code>

Using a Remote Shell connection: To execute a command on the emulated node using `rsh`, complete the following step.

Step	Action
1	Enter the following command: <code>rsh failed_node command</code> <i>failed_node</i> is the name of the failed node. <i>command</i> is the Data ONTAP command you want to run. Example: In the following example, <code>filer2</code> is the failed node. <code>rsh filer2 df</code>

Emulated node command exceptions

Almost all the commands that are available to a takeover node are available on the emulated node. Some commands, however, are either unavailable or behave differently in emulated mode. The commands in the following subsections are the exceptions.

Unavailable commands: The following commands are not available on an emulated node:

- ◆ `cf disable`
- ◆ `cf enable`
- ◆ `cf forcegiveback`
- ◆ `cf forcetakeover`
- ◆ `cf giveback`
- ◆ `cf takeover`
- ◆ `date`
- ◆ `halt`
- ◆ `ifconfig partner`
- ◆ `ifconfig -partner`
- ◆ `ifconfig mtsize`
- ◆ `license cluster`
- ◆ `rdate`
- ◆ `reboot`
- ◆ `timezone`

Commands with different behaviors: The following table lists the commands that behave differently on an emulated node than they do on a takeover node, and describes the behavior.

Command	Difference
<code>ifconfig</code> <code>[interface]</code>	Displays the following: ◆ Emulated interface mappings based on the failed node <code>/etc/rc</code> file rather than the takeover node interface mappings ◆ Emulated interface names rather than the interface names of the takeover node ◆ Only interfaces that have been configured, rather than all interfaces, configured or not, as displayed on the takeover node
<code>mt</code>	Uses the tape devices on the takeover node because the failed node has no access to its tape devices.
<code>netstat -i</code>	Appends a plus sign (+) to shared interfaces. A shared interface is one that has two IP addresses assigned to it: an IP address for the node in which it physically resides and an IP address for its partner node in the cluster.

Command	Difference
<code>sysconfig</code>	When it displays hardware information, the <code>sysconfig</code> command displays information only about the hardware that is attached to the takeover node. It does not display information about the hardware that is attached only to the failed node. For example, the disk adapter information that the <code>partner sysconfig -r</code> command displays is about the disk adapters on the takeover node.
<code>uptime</code>	Displays how long the failed node has been down and the host name of the takeover node.
<code>aggr status</code>	When it displays hardware information, the <code>aggr status</code> command displays information only about the hardware that is attached to the takeover node. It does not display information about the hardware that is attached only to the failed node. For example, the disk adapter information that the <code>partner aggr status -r</code> command displays is about the disk adapters on the takeover node.

Backing up and restoring from the emulated node

Emulated node tape drives are inaccessible

You can perform dumps and restores on the emulated node. However, neither the takeover node nor the emulated node can access the tape drives or other peripheral devices on the failed node.

Prerequisites to performing scheduled dumps on an emulated node

Any dump commands directed to the failed node's tape drives are executed on the takeover node's tape drives. Therefore, any dump commands that you execute using a scheduler, such as the `cron` command, succeed only under the following conditions:

- ◆ The device names are the same on both nodes in the cluster.
- ◆ The dump commands for the takeover node and the emulated node are not scheduled to occur during the same time period; the takeover node and the emulated node cannot access the tape drives simultaneously.

Performing dumps and restores for a failed node

Because the peripheral devices for a failed node are inaccessible, you perform dumps and restores for a failed node by using the emulated node (available using the `partner` command on the takeover node), making sure that you use a peripheral device attached to the takeover node.

For more information about performing dumps and restores, see the *Data Protection Tape Backup and Recovery Guide*.

Example using the restore command in partner mode:

```
filer1 (takeover)> partner
filer1/filer2> restore [options [arguments]]
```

Example using the restore command as an argument to the partner command:

```
filer1 (takeover)> partner restore [options [arguments]]
```

Managing giveback

Giveback management tasks

Managing giveback involved the following tasks:

- ◆ [“Configuring giveback”](#) on page 137
- ◆ [“Performing a giveback”](#) on page 139
- ◆ [“Enabling automatic giveback”](#) on page 142

Managing giveback

Configuring giveback

Giveback configuration tasks

- Configuring giveback involves the following tasks:
- ◆ Shortening giveback time
 - ◆ Setting giveback delay time for CIFS clients
 - ◆ Setting giveback to terminate long-running processes

About shortening giveback time

You can shorten the client service outage during giveback by enabling an option (`cf.giveback.check.partner`) that checks whether the failed node is already recovered and waiting for giveback—that is, waiting for disk reservation release—when you enter the `cf giveback` command at the console of the takeover node. If the failed node is waiting for giveback, giveback proceeds as normal; if not, giveback is canceled.

Note—
This option works only with platforms booted from CompactFlash. To determine whether your storage system boots from CompactFlash, see your Hardware and Service guide.

Enabling shorter giveback time

To enable shorter giveback time, complete the following step.

Step	Action
1	Enter the following command: <pre>options cf.giveback.check.partner {on off}</pre> Note— This option is Off by default, except for platforms booted from CompactFlash, where it is On by default.

Setting giveback delay time for CIFS clients

This option specifies the number of minutes to delay an automatic giveback before terminating CIFS clients that have open files. During the delay, the system periodically sends notices to the affected clients. If you specify 0, CIFS clients are terminated immediately.

This option is used only if automatic giveback is On.

To set takeover delay time for CIFS clients, complete the following step.

Step	Action
1	Enter the following command: <code>options cf.giveback.auto.cifs.terminate.minutes <i>minutes</i></code> Valid values for <i>minutes</i> are 0 through 999. The default is 5 minutes.

About setting giveback to terminate long-running processes

This option, when on, specifies that automatic giveback should immediately terminate long-running operations (dump/restore, vol verify, and so on) when initiating an automatic giveback. When this option is off, the automatic giveback is deferred until the long-running operations are complete. This option is used only if automatic giveback is On.

Setting giveback to terminate long-running processes

To set automatic giveback to terminate long-running processes, complete the following step.

Step	Action
1	Enter the following command: <code>options cf.giveback.auto.terminate.bigjobs {on off}</code> The <code>on</code> argument enables this option. The <code>off</code> argument disables this option. This option is On by default.

Managing giveback

Performing a giveback

Types of giveback operations

- You can perform the following types of giveback operations from the takeover node:
- ◆ Normal giveback
 - ◆ Giveback and terminate processes
 - ◆ Forced giveback

Initiating normal giveback

To initiate normal giveback, complete the following step.

Step	Action
1	Enter the following command on the command line of the takeover node: <pre>cf giveback</pre> Note _____ If the giveback fails, there might be a process running that prevents giveback. You can wait and repeat the command, or you can follow the steps outlined in “Terminating processes to initiate giveback” on page 139. After a giveback, the takeover node’s ability to take over its partner automatically is not reenabled until the partner reboots successfully. If the partner fails to reboot, you can enter the <code>cf takeover</code> command to initiate a takeover of the partner manually. _____

Terminating processes to initiate giveback

Some storage system processes prevent a giveback from being completed. If this happens, you can either wait for the process to be complete, or you can terminate the process early to initiate a giveback operation.

Processes that can prevent giveback: The `cf giveback` command does not succeed if any of the following events are taking place on the emulated node at the same time:

- ◆ Dump and restore processes
- ◆ Outstanding CIFS sessions
- ◆ RAID disk additions
- ◆ Volume creation (traditional volume or FlexVol™ volume)
- ◆ Aggregate creation
- ◆ Disk ownership assignment
- ◆ Disks being added to a volume (`vol add`)
- ◆ Snapshot™ copy creation, deletion, or renaming
- ◆ Quota initialization
- ◆ Advanced mode repair operations, such as `wafliron`
- ◆ Storage system panics
- ◆ Backup dump and restore operations
- ◆ SnapMirror® transfers (if the partner is a SnapMirror destination)
- ◆ SnapVault™ restorations
- ◆ Disk sanitization operations

Terminating the process and initiating a giveback: To terminate a process early and initiate giveback to the failed node, complete the following step.

Step	Action
1	On the takeover node, enter the following command: <pre>cf giveback -f</pre> Note _____ If you invoke <code>cf giveback -f</code> and it fails, the takeover node panics. After the takeover node recovers, you can attempt the <code>giveback -f</code> command again. If you cannot successfully perform a giveback using the <code>-f</code> option, you must force giveback using the <code>forcegiveback</code> command. See “Forcing giveback” on page 141.

Forcing giveback

The takeover node might detect an error condition on the failed node that typically prevents a complete giveback, such as data not being flushed from NVRAM to the failed node’s disks.

To force the takeover node to give back the resources of the failed node even if the takeover node detects an error that typically prevents a complete giveback, complete the following step.

Step	Action
1	On the takeover node, enter the following command: <code>cf forcegiveback</code> Attention _____ Use <code>cf forcegiveback</code> only when you cannot get <code>cf giveback</code> to succeed. When you use this command, you risk losing any data committed to NVRAM but not to disk. _____

What happens when a giveback is interrupted

If the takeover node experiences a failure or a power outage during the giveback process, the giveback process stops and the takeover node returns to takeover mode when the failure is repaired or the power is restored.

Enabling automatic giveback

When automatic giveback is initiated

Data ONTAP initiates an automatic giveback when one node of a cluster is in takeover mode and the failed node is repaired and reboots. The repaired node boots into Data ONTAP, and the node in takeover mode detects this and initiates a giveback.

Note_____This option works only with platforms booted from CompactFlash. To determine whether your storage system boots from CompactFlash, see your Hardware and Service guide.

Enabling or disabling automatic giveback

To enable or disable automatic giveback, complete the following step.

Step	Action
1	Enter the following command: <pre>options cf.giveback.auto.enable {off on}</pre> The <code>on</code> value enables automatic giveback. The <code>off</code> value disables automatic giveback. This option is Off by default.

Configuring network interfaces for clusters

Interface configuration tasks

Configuring network interfaces for clusters involves the following tasks:

- ◆ “[Understanding interface configurations for takeover](#)” on page 144
- ◆ “[Configuring shared interfaces](#)” on page 148
- ◆ “[Configuring dedicated and standby interfaces](#)” on page 151

Note

For information about configuring a cluster to use FCP, see “How Data ONTAP supports FCP with clustered systems” in the *Block Access Management Guide*.

Understanding interface configurations for takeover

Types of takeover configurations

The following table lists the three types of interface configurations that you can enable in a cluster.

Interface type	Description
Shared	This type of interface supports both the local and partner nodes. It contains both the local node and partner node IP addresses. During takeover, it supports the identity of both nodes.
Dedicated	This type of interface only supports the node in which it is installed. It contains the local node IP address only and does not participate in network communication beyond local node support during takeover. It is paired with a standby interface.
Standby	This type of interface is on the local node, but only contains the IP address of the partner node. It is paired with a dedicated interface.

Note

Most cluster interfaces are configured as shared interfaces because they do not require an extra NIC. However, some Asynchronous Transfer Mode (ATM) configurations require interfaces to be configured as dedicated and standby.

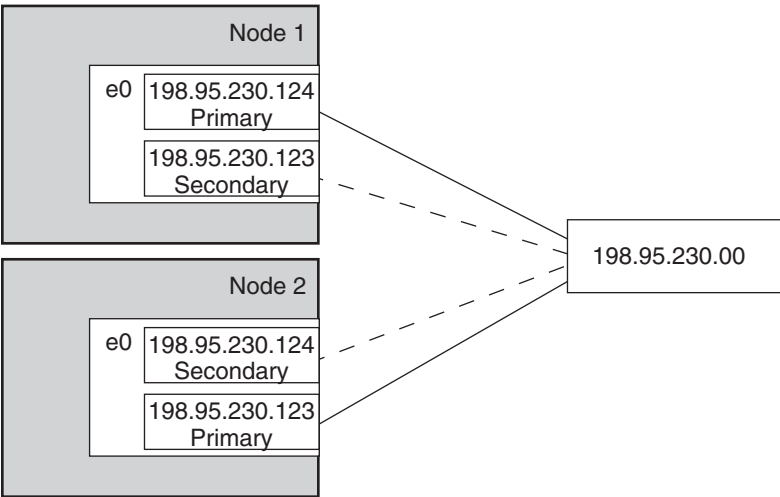
Interface roles in normal and takeover modes

The following table shows the role of each interface type in normal and takeover mode.

Interface type	Normal mode	Takeover mode
Shared	Supports the identity of the local node	Supports the identity of both the local node and the failed node
Dedicated	Supports the identity of the local node	Supports the identity of the local node
Standby	Idle	Supports the identity of the failed node

Takeover configuration with shared interfaces

In the following configuration illustration, you use two NICs to provide the two interfaces.

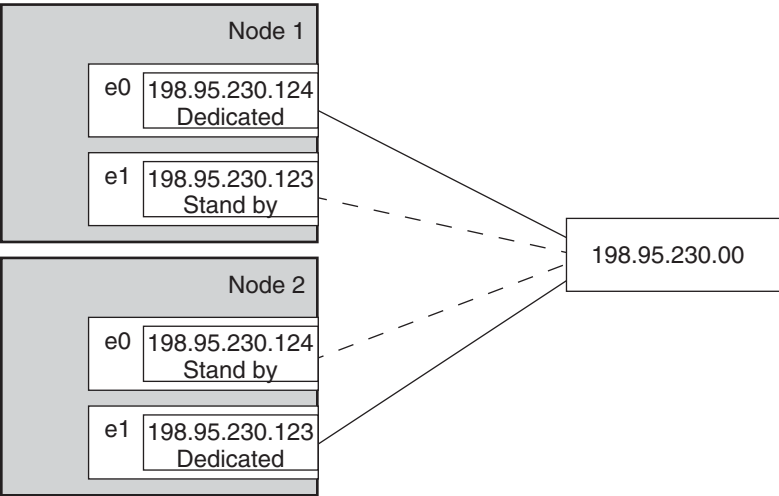


If Node 1 fails, interface e0 on Node 1 stops functioning, but the secondary address on e0 on Node 2 handles the Node 1 network connection with the 230 network.

If Node 2 fails, e0 on Node 2 stops functioning, but e0 on Node 1 substitutes for the failed interface and handles the Node 2 network connection with the 230 network.

Takeover configuration with dedicated and standby interfaces

In the following configuration illustration, you use two NICs for each interface, one on each storage system. One NIC acts as a dedicated interface and the other acts as a standby interface.



If Node 1 fails, interface e0 on Node 1 stops functioning, but e0 on Node 2 substitutes for the failed interface and handles the Node 1 network connection with the 230 network.

If Node 2 fails, e1 on Node 2 stops functioning, but e1 on Node 1 substitutes for the failed interface and handles the Node 2 network connection with the 230 network.

Interface types and configurations

The following table lists the configurations supported by each type of interface in a cluster.

Interface	Shared	Dedicated	Standby	Partner parameter
Ethernet	X	X	X	IP address or interface name
Gigabit Ethernet	X	X	X	IP address or interface name
FDDI	X	X	X	IP address or interface name

Interface	Shared	Dedicated	Standby	Partner parameter
Virtual interface	X	X	X	Virtual interface name
VLAN interface	X	X	X	IP address or interface name
ATM FORE/IP		X	X	IP address or FORE/IP interface name
ATM adapter for LANE	X ¹	X	X	ATM adapter interface name

¹ATM adapters in a cluster can be shared for LANE provided that the two adapters do not have Emulated LANs in common.

Configuring shared interfaces

Prerequisites for configuring shared interfaces

To configure a pair of shared interfaces so that one can replace the other if one storage system fails and is taken over, you must have available the following items:

- ◆ Two interfaces, one installed on each node in the cluster
- ◆ Two IP addresses, the address of one interface on each node in the cluster

Emulated LAN considerations

For LANE, when an adapter is shared, an Emulated LAN cannot be configured more than once on the adapter after takeover. This means that if both nodes in a cluster pair are on the same Emulated LAN, each node needs a minimum of two ATM cards to support full takeover.

Configuring shared interfaces

To configure a pair of shared interfaces for takeover, complete the following steps.

Step	Action
1	Enter the following command on the command line of both nodes in a cluster <i>and</i> in the <code>/etc/rc</code> file of both nodes, so that the command is permanent: <pre>ifconfig interface local_address partner partner_address {other_options}</pre> <i>interface</i> is the name of the local interface. <i>local_address</i> is the IP address of the local interface. <i>partner_address</i> is the IP address or interface name that the local interface services when the local node takes over the other node in the cluster. <i>other_options</i> denotes whatever other options are needed to correctly configure the interface in your network environment. For more information about the <code>ifconfig</code> command, see the <i>Network Management Guide</i> and the <code>na_ifconfig</code> man page. Example : <pre>ifconfig e0 199.9.204.255 partner 199.9.204.254</pre> Note_____ When you configure virtual interfaces for takeover, you must specify the interface name and not the IP address. _____
2	If you want to configure your interfaces for automatic takeover in case of NIC failure, follow the steps outlined in “Configuring VIFs or interfaces for automatic takeover” on page 128.

Emulated LAN example

A cluster pair might be configured with Emulated LANs as follows.

Storage system and adapter	Interfaces and Emulated LANs	Storage system and adapter	Interfaces and Emulated LANs
appliance1-a4	el0 default el1 nineKMTU	appliance1-a5	el2 eighteenKMTU
appliance2-a4	el0 default el1 nineKMTU	appliance2-a5	el2 eighteenKMTU

To achieve full takeover, you would use these commands on both nodes:

```
ifconfig a4 partner a5
ifconfig a5 partner a4
```

Note_____

If a4 and partner a5 are configured with common Emulated LANs, a4 does not take over those Emulated LANs.

Virtual interface example

The following is an example of a virtual interface configuration command:

```
ifconfig vif0 partner vif0
```

VLAN interface example

The following is an example of a VLAN interface configuration command:

```
ifconfig e9-100 partner e7-100
```

Configuring network interfaces for clusters

Configuring dedicated and standby interfaces

Prerequisites for configuring interfaces

- To configure a pair of interfaces so that the standby interface can replace a dedicated interface on the partner node in a cluster, you must have the following items available:
- ◆ Four interfaces, two on each node in the cluster
 - ◆ Two IP addresses or interface names, the address of one interface or interface name for each node in the cluster

Rules for specifying interfaces

- Keep these rules in mind when specifying interfaces:
- ◆ You can use interface names for specifying all interfaces.
 - ◆ You cannot use IP addresses when specifying ATM interfaces for LANE and virtual interfaces.

Configuring dedicated and standby interfaces

To configure dedicated and standby interfaces for a cluster made up of nodeA and nodeB, complete the following steps.

Step	Action
1	On nodeA, enter the following command on the command line and also enter it in the /etc/rc file, so that the command is permanent: <pre>ifconfig interfaceA1 addressA1 {other_options}</pre> <i>interfaceA1</i> is the name of the dedicated local interface for nodeA. <i>addressA1</i> is the IP address of the dedicated local interface for nodeA. <i>other_options</i> denotes whatever other options are needed to correctly configure the interface in your network environment. For more information about the <code>ifconfig</code> command, see the <i>Network Management Guide</i> and the <code>na_ifconfig</code> man page. Result: The dedicated local interface for nodeA is configured.

Step	Action
2	Also on nodeA, enter the following command on the command line and in the /etc/rc file: <pre>ifconfig interfaceA2 partner addressB1</pre> <i>interfaceA2</i> is the name of the standby interface for nodeA. <i>addressB1</i> is the IP address or interface name of the dedicated interface of nodeB. Note When you configure virtual interfaces for takeover, you must specify the interface name and not the IP address. Result: The standby interface for nodeA is configured to take over the dedicated interface of nodeB on takeover.
3	On nodeB, enter the following command on the command line and in the /etc/rc file: <pre>ifconfig interfaceB1 addressB1 {other_options}</pre> <i>interfaceB1</i> is the name of the dedicated local interface for nodeB. <i>addressB1</i> is the IP address of the dedicated local interface for nodeB. <i>other_options</i> denotes whatever other options are needed to correctly configure the interface in your network environment. For more information about the <code>ifconfig</code> command, see the <i>Network Management Guide</i> and the <code>na_ifconfig</code> man page. Result: The dedicated local interface for nodeB is configured.

Step	Action
4	Also on nodeB, enter the following command on the command line and in the /etc/rc file: <pre>ifconfig interfaceB2 partner addressA1</pre> <i>interfaceB2</i> is the name of the standby interface on nodeB. <i>addressA1</i> is the IP address or interface name of the dedicated interface for nodeA. Note When you configure virtual interfaces for takeover, you must specify the interface name and not the IP address. Result: The standby interface on nodeB is configured to take over the dedicated interface of nodeA on takeover.
5	If you want to configure your interfaces for automatic takeover in case of NIC failure, follow the steps outlined in “Configuring VIFs or interfaces for automatic takeover” on page 128.

Checking cluster configuration

About the Cluster Configuration Checker utility

NetApp provides a Cluster Configuration Checker utility that you can download and run to protect against the most common configuration causes of failover problems between clustered nodes, including

- ◆ Mismatched licenses
- ◆ Mismatched option settings
- ◆ Incorrectly configured network interfaces
- ◆ Different versions of Data ONTAP on the local and partner nodes
- ◆ Differences in the cfmodes settings between the two nodes

Where you can run the utility

You can run the utility, `cf-config-check.cgi`, as a command from a UNIX shell, or you can install the Common Gateway Interface (CGI) script on a UNIX Web server and invoke it from a Web browser.

rsh requirements

To run the Cluster Configuration Checker utility, `rsh` access to both clustered nodes is required.

Downloading and running the checker

To download and run the Cluster Configuration Checker, log in to the NOW Web site (<http://now.netapp.com/>) and go to Software Downloads > Tools & Utilities. Click “Cluster Configuration Checker (cf-config-check.cgi).”

Follow the directions for downloading and running the utility.

Correcting takeover or giveback failures

About this section The procedure for correcting a takeover or giveback failure on an F800 series cluster is different from that of a FAS900 and FAS3000 series cluster. See one of the following sections according to your system hardware:

- ◆ [“Troubleshooting F800 series takeover or giveback failures”](#)
- ◆ [“Troubleshooting FAS900 and FAS3000 series takeover or giveback failures”](#) on page 156

**Troubleshooting
F800 series
takeover or
giveback failures**

If takeover or giveback fails on an F800 series cluster, complete the following steps.

Step	Action
1	Check the Channel B cabling on the disk shelves, and tighten any loose cables.
2	Check the cabling to your cluster adapters, making sure that the cabling is correct, that the cables are properly seated in the adapters, and that the cables are secured to the adapters.
3	Check communication between the nodes by entering the following command: <code>cf status</code> Result: If there is a cabling error or the interconnect link is down, the system displays an error message indicating the error.
4	If you have not already done so, run the Cluster Configuration Checker utility. For details, see “Checking cluster configuration” on page 154.
5	Check network connectivity. For details, see the <i>Data ONTAP Network Administration Guide</i> .
6	Reboot the cluster and rerun the takeover and giveback tests.
7	If you still do not have clustering enabled, contact technical support.

Troubleshooting FAS900 and FAS3000 series takeover or giveback failures

If takeover or giveback fails on a FAS900 series or FAS3000 series cluster, complete the following steps.

Note

You must reboot the cluster after completing each test.

To correct a takeover or giveback failure, complete the following steps.

Step	Action	
1	Check communication between the local and partner nodes by entering the following command and observing the messages: <pre>cf status</pre> Result: If there is a problem with cabling or the adapter, the appropriate message is displayed. Recable or replace the cluster adapter, as needed.	
2	If the error message indicates...	Then...
	A cluster adapter error	Check the cluster adapter cabling. Make sure that the cabling is correct and properly seated at both ends of the cable.
	That the NVRAM5 adapter is in the wrong slot number	Check the NVRAM5 slot number. Move it to the correct slot if it is in the wrong one.
	A Channel B cabling error	Check the cabling of the Channel B disk shelf loops and reseal and tighten any loose cables.
	A networking error	Check for network connectivity. See the <i>Network Administration Guide</i> for more information.
3	If you have not already done so, run the Cluster Configuration Checker script, as described in “ Checking cluster configuration ” on page 154. Correct any errors or differences displayed in the output.	
4	Reboot the cluster and rerun the takeover and giveback tests.	
5	If you still do not have clustering enabled, contact technical support.	

Removing clustering

Tasks for removing clustering

- Removing clustering is a two-step process:
- ◆ Disable clustering software, such as the cluster license
 - ◆ Remove the clustering hardware
 - ❖ [“Removing F800 series clustering hardware”](#) on page 158
 - ❖ [“Removing FAS900 and FAS3000 series clustering hardware”](#) on page 160

Disabling clustering software

This procedure applies to F800, FAS900, and FAS3000 series systems. To disable the clustering software, complete the following steps.

Step	Action
1	Enter the following command on either node console: <code>cf disable</code>
2	Disable the cluster license by entering the following command: <code>license delete cluster</code>
3	Open the /etc/rc file with a text editor and remove references to the partner node in the ifconfig entries, as shown in the following example: Original entry: <code>ifconfig e0 199.9.204.254 partner 199.9.204.255</code> Edited entry: <code>ifconfig e0 199.9.204.254</code>
4	Repeat Step 1 through Step 3 on the partner node.
5	Proceed to the appropriate following section based on your hardware.

Removing F800 series clustering hardware

Tasks for removing the hardware

Removing F800 series clustering hardware is a two-part process:

- ◆ Recabling your disk shelves
- ◆ Removing the hardware

Recabling Channel B for redundant disk loops

To recable your Channel B loops and convert your cluster to stand-alone storage systems, complete the following steps.

Note _____
This procedure assumes that you have DS14 disk shelves and dual Fibre Channel adapters.

Step	Action
1	Halt both nodes by entering the following command at each console: <code>halt</code>
2	Ground yourself, then power off the nodes and disk shelves.
3	Unplug the local node from the partner node’s Channel B disk shelf and do the same for the partner node.
4	Connect the local node to the open Channel B loop in its local disk shelves.
5	Repeat Step 4 for the partner node.
6	Go to “ Removing the hardware ” on page 159.

Removing the hardware

To remove the clustering hardware and extra FC-AL adapters, complete the following steps.

Step	Action
1	Ground yourself, using the antistatic wrist strap.
2	Remove the interconnect cables from the cluster adapter.
3	Open the back of the local node, as described in the <i>F800 Series Hardware Guide</i> .
4	If applicable, unlock and remove the GBIC from the adapter (typically used with VI cluster adapters).
5	Unscrew the adapter from the chassis.
6	Remove the adapter.
7	Replace the slot cover and secure it with the screw from the adapter.
8	Remove any extra FC-AL or dual Fibre Channel adapters.
9	Close and secure the back of the storage system, as described in the <i>F800 Series Hardware Guide</i> .
10	Power on first the disk shelves and then the storage system; press the Delete key during boot to get to the OK prompt.
11	Unset the partner system ID by entering the following command at the OK prompt: <code>unsetenv partner-sysid</code>
12	Boot the storage system by entering the following command at the OK prompt: <code>boot</code>
13	Check the status of clustering on the storage system by entering the following command on the console: <code>cf status</code> Result: Cluster monitoring is disabled, as indicated by the following message: Cluster monitor not initialized
14	Repeat Step 1 through Step 13 for the partner node.

Removing FAS900 and FAS3000 series clustering hardware

Removing the hardware

To remove FAS900 or FAS3000 series clustering hardware and recable the node’s disk shelves, complete the following steps.

Note See the System Configuration Guide for the most up-to-date slot assignments.

Step	Action	
1	Halt both nodes by entering the following command on each console: <code>halt</code>	
2	Turn off the power to each node, then turn off the power to the disk shelves and unplug them from the power source.	
3	Ground yourself, then remove the cluster interconnect cables from both nodes. See the <i>Hardware and Service Guide</i> for additional information.	
4	If you are converting a cluster to a stand-alone storage system for a...	Then...
	FAS920, 940, or 960 with IB adapter	Open both nodes, remove the cluster adapters, and then close the nodes.
	FAS900 series standard cluster with NVRAM5 adapter	Remove the NVRAM5 adapter from slot 11 and install it in slot 10.
	FAS980 Fabric Attached MetroCluster with NVRAM5 adapter	Completely remove the VI cluster adapter from slot 6.
	FAS3000 series standard cluster with NVRAM5 adapter	Move the NVRAM5 adapter to slot 1.

Step	Action	
5	Unplug the local node from the partner node's Channel B disk shelf and do the same for the partner node.	
6	If you are converting to a...	Then...
	System with nonmirrored disks	a. Disconnect all cabling from the Channel B loop on the local node. b. Repeat for the partner node.
	System with mirrored disks or a redundant Channel B loop	a. Connect the local node to the open Channel B loop in its local disk shelves, as described in the appropriate disk shelf guide. b. Repeat for the partner node.
7	Power on the disk shelves, then the individual nodes, monitoring the system console for error messages during the boot process.	
8	Run all system diagnostics at the boot prompt by entering the following command on the system console: <code>boot diags</code>	
9	Unset the partner system ID by entering the following command at the prompt: <code>unsetenv partner-sysid</code>	
10	Boot the node by entering the following command: <code>boot</code>	
11	Check clustering status by entering the following command: <code>cf status</code> Result: If clustering is disabled, you see the following output: Cluster monitor not initialized	
12	Repeat Step 1 through Step 11 for the partner node.	

About this chapter

This chapter describes how to use the optional MetroCluster feature of Data ONTAP to recover data in situations such as prolonged power outages or natural disasters.

Note

This chapter describes recovering data on storage systems. For information about protecting data on other networked equipment—application servers, for example—see the documentation for that equipment.

Topics in this chapter

This chapter discusses the following topics:

- ◆ “[Recognizing a disaster](#)” on page 164
- ◆ “[Recovering from a disaster](#)” on page 166

Recognizing a disaster

Failures that require disaster recovery

The disaster recovery procedure is an extreme measure that you should use only if the failure disrupts all communication from one MetroCluster site to the other for a prolonged period of time.

The following are examples of disasters that could cause such a failure:

- ◆ Fire
- ◆ Earthquake
- ◆ Prolonged power outages at a site
- ◆ Prolonged loss of connectivity from clients to the storage systems at a site

Ways to determine whether a disaster occurred

It is critical that you follow a predefined procedure to confirm that a disaster occurred. The procedure should include determining the status of the disaster site by

- ◆ Using external interfaces to the storage system, such as the following:
 - ❖ Ping
 - ❖ Remote shell
 - ❖ FilerView® administration tool
- ◆ Using network management tools to verify connectivity to the disaster site
- ◆ Physically inspecting the disaster site, if possible

You should declare a disaster only after verifying that service cannot be restored.

Failures that do not require disaster recovery

If you can reestablish the cluster relationship after fixing the problem, you should not perform the disaster recovery procedure.

Do not perform the disaster recovery procedure for the following failures:

- ◆ A failure of the cluster interconnect between the two sites. This can be caused by the following:
 - ❖ Failure of the interconnect cable
 - ❖ Failure of one of the VI cluster adapters
 - ❖ If using switches, a failure of the SFP connecting a cluster node to the switch

With this type of failure, both nodes of the cluster remain running. Automatic takeover is disabled because Data ONTAP cannot synchronize the NVRAM logs. After you fix the problem and reestablish the connection, the cluster nodes resynchronize their NVRAM logs and the cluster returns to normal operation.

- ◆ The storage from one site (site A) is not accessible to the cluster node at the other site (site B). This can be caused by the following:
 - ❖ Failure of any of the cables connecting the storage at one site to the cluster node at the other site or switch
 - ❖ If using switches, failure of any of the SFPs connecting the storage to the switch or the cluster node to the switch
 - ❖ Failure of the Fibre Channel adapter on the cluster node
 - ❖ Failure of a storage disk shelf (LRC, ESH, ESH2, or AT-FCX module; power; access to disk shelves; and so on)

With this type of failure, you see a “mailbox disk invalid” message on the console of the storage system that cannot see the storage. After you fix the problem and reestablish the connection, the cluster returns to normal operation.

- ◆ If you are using switches, the Inter-Switch Link (ISL) between each pair of switches fails.

With this type of failure, both nodes of the cluster remain running. You see a “mailbox disk invalid” message because a storage system at one site cannot see the storage at the other site. You also see a message because the two nodes cannot communicate with each other. After you fix the problem and reestablish the connection, the cluster nodes resynchronize their NVRAM logs and the cluster returns to normal operation.

Recovering from a disaster

What to do

After determining that the failure to the MetroCluster configuration is a disaster, you should do the following:

- ◆ Ensure that the surviving node is isolated from its partner. See [“Restricting access to the disaster site cluster node”](#) on page 166.
- ◆ Force the surviving node to take over the functions of its partner. See [“Forcing a node into takeover mode”](#) on page 168.
- ◆ Recover access to the failed partner’s data by completing one of the following tasks:
 - ❖ If you are using file-access protocols, remount the failed partner’s volumes. See [“Remounting volumes of the failed node”](#) on page 168.
 - ❖ If you are using iSCSI, bring the failed partner’s LUNs online.
- ◆ Fix the cluster at the disaster site. See [“Fixing failures caused by the disaster”](#) on page 170.
- ◆ Re-create the MetroCluster configuration. See [“Reestablishing the MetroCluster configuration”](#) on page 170.

Restricting access to the disaster site cluster node

You must restrict access to the disaster site cluster node to prevent the cluster node from resuming service. If you do not restrict access, you risk the possibility of data corruption.

You can restrict access to the disaster site cluster node in the following ways:

- ◆ Turn off power to the disaster site cluster node.
- ◆ Use manual fencing.

Turning off power to the disaster site cluster node: To turn off power to the disaster site cluster node, complete the following step.

Step	Action
1	Switch off the power at the back of the storage system.

This is the preferred method for restricting access to the disaster site cluster node. You can perform this task at the disaster site or remotely, if you have that capability.

Using manual fencing: You can use manual fencing as an alternative to turning off power to the disaster site cluster node. The manual fencing method restricts access using software and physical means.

To manually fence off the disaster site cluster node, complete the following steps.

Step	Action	
1	Disconnect the cluster interconnect and Fibre Channel adapter cables of the cluster node at the surviving site.	
2	If you are using...	Then fencing is achieved by...
	Application failover	Using any application-specified method that either prevents the application from restarting at the disaster site or prevents the application clients from accessing the application servers at the disaster site. Methods can include turning off the application server, removing an application server from the network, or any other method that prevents the application server from running applications.
	IP failover	Using network management procedures to ensure that the storage systems at the disaster site are isolated from the external public network.

Forcing a node into takeover mode

To force the surviving node into takeover mode, complete the following step.

Step	Action
1	Enter the following command on the surviving cluster node: <pre>cf forcetakeover -d</pre> Result: Data ONTAP causes the following to occur: ◆ The surviving cluster node takes over the functions of the failed partner. ◆ The mirrored relationships between the two plexes of mirrored aggregates are broken, thereby creating two unmirrored aggregates. This is called splitting the mirrored aggregates. The overall result of using the <code>cf forcetakeover -d</code> command is that a cluster node at the surviving site is running in takeover mode with all the data in unmirrored aggregates.

Remounting volumes of the failed node

You must remount the volumes of the failed node because the volumes are accessed through the surviving node. For more information about mounting volumes, see the *File Access and Protocols Management Guide*.

To remount the volumes, complete the following steps.

Step	Action
1	On an NFS client at the surviving site, create a directory to act as a mount point. Example: <code>mkdir /n/toaster/home</code>
2	Mount the volume. Example: <code>mount toaster:/vol/vol0/home /n/toaster/home</code>

Recovering LUNs of the failed node

If you have a MetroCluster configuration with iSCSI-attached hosts, you must actively track the state of LUNs (track whether they are online or offline) on the node at each site. If there is a failure to a MetroCluster configuration that qualifies as a disaster and the node at one site is inaccessible, all LUNs in the

aggregates that were mirrored at the surviving site are offline. There is no way to distinguish the LUNs that were offline before the disaster from the LUNs that were online before the disaster unless you have been tracking their status.

When you recover access to the failed node's LUNs, it is important to bring back online only the LUNs that were online before the disaster. To avoid igroup mapping conflicts, do not bring a LUN online if it was offline before the disaster. For example, suppose you have two LUNs with IDs of 0 mapped to the same igroup, but one of these LUNs was offline before the disaster. If you bring the previously offline LUN online first, you cannot bring the second LUN online because you cannot have two LUNs with the same ID mapped to the same host.

To recover access to the failed node's LUNs, complete the following steps.

Step	Action
1	Identify the LUNs that were online before the disaster occurred.
2	Make sure that the LUNs are mapped to an igroup that contains the hosts attached to the surviving cluster node. For more information about mapping LUNs to igroups, see your <i>Block Access Management Guide</i>.
3	On the surviving cluster node, enter the following command: <pre>lun online lun-path ...</pre> <i>lun-path</i> is the path to the LUN you want to bring online. You can specify more than one path to bring multiple LUNs online. Example 1: <code>lun online /vol/vol1/lun0</code> Example 2: <code>lun online /vol/vol1/lun0 /vol/vol1/lun1</code> Note _____ After you bring LUNs back online, you might have to perform some application or host-side recovery procedures. For more information, see the documentation for your application and for your host operating system.

Fixing failures caused by the disaster

You need to fix the failures caused by the disaster, if possible. For example, if a prolonged power outage to one of the MetroCluster sites caused the failure, restoring the power fixes the failure.

To fix failures, complete the following steps.

Step	Action
1	Ensure that the storage system and disk shelves have power.
2	Ensure that all Fibre Channel adapters, VI cluster adapters, cables, and switch ports are connected and working.

You cannot fix failures if the disaster causes a site to be destroyed. For example, a fire or an earthquake could destroy one of the MetroCluster sites. In this case, you fix the failure by creating a new MetroCluster configured partner at a different site.

After the node at the surviving site can see the disk shelves at the disaster site, Data ONTAP renames the volumes in the mirrored aggregates that were split by adding a number in parenthesis to the volume name. For example, if the volume name was vol1 before the disaster and the split, the renamed volume name could be vol1(1).

Reestablishing the MetroCluster configuration

Depending on the state of a mirrored aggregate before you forced the surviving cluster to take over its partner, you use one of two procedures to reestablish the MetroCluster configuration:

- ◆ If the mirrored aggregate was in a normal state before the forced takeover, you can rejoin the two aggregates to reestablish the MetroCluster configuration. This is the most typical case.
- ◆ If the mirrored aggregate was in an initial resynchronization state (level-0) before the forced takeover, you cannot rejoin the two aggregates. You must re-create the synchronous mirror to reestablish the MetroCluster configuration.

Rejoining the two aggregates: To return the cluster to normal operation by rejoining the two aggregates, complete the following steps.

Step	Action
1	Validate that you can access the remote storage by entering the following command: <pre>aggr status -r</pre>
2	Turn on power to the cluster node at the disaster site. Result: After the cluster node at the disaster site boots, it displays the following message: <pre>Waiting for Giveback...</pre>
3	Determine which aggregates are at the surviving site and which aggregates are at the disaster site by entering the following command: <pre>aggr status</pre> Aggregates at the disaster site show plexes that are in a failed state with an out-of-date status. Aggregates at the surviving site show plexes as online.
4	If aggregates at the disaster site are online, take them offline by entering the following command for each online aggregate: <pre>aggr offline disaster_aggr</pre> <i>disaster_aggr</i> is the name of the aggregate at the disaster site. Note_____ An error message appears if the aggregate is already offline. _____
5	Re-create the mirrored aggregates by entering the following command for each aggregate that was split: <pre>aggr mirror aggr_name -v disaster_aggr</pre> <i>aggr_name</i> is the aggregate on the surviving site's cluster node. <i>disaster_aggr</i> is the aggregate on the disaster site's cluster node. Result: The <i>aggr_name</i> aggregate rejoins the <i>disaster_aggr</i> aggregate to reestablish the MetroCluster configuration.

Step	Action
6	Enter the following command at the partner cluster node: cf giveback Result: The cluster node at the disaster site reboots.

Example of rejoining aggregates: The following example shows the commands and status output when you rejoin aggregates to reestablish the MetroCluster configuration.

First, the aggregate status of the disaster site's storage after reestablishing access to the cluster partner at the surviving site is shown.

```
filer1> aggr status -r
Aggregate mir (online, normal) (zoned checksums)
  Plex /mir/plex5 (online, normal, active)
    RAID group /filer1/plex5/rg0 (normal)

RAID Disk Device HA  SHELF BAY CHAN  Used (MB/blks) Phys (MB/blks)
-----
parity  8a.2  8a   0    2  FC:B  34500/70656000 35003/71687368
data    8a.8  8a   1    0  FC:B  34500/70656000 35003/71687368

Aggregate mir(1) (failed, out-of-date) (zoned checksums)
  Plex /mir(1)/plex1 (offline, normal, out-of-date)
    RAID group /mir(1)/plex1/rg0 (normal)

RAID Disk Device HA  SHELF BAY CHAN  Used (MB/blks) Phys (MB/blks)
-----
parity  6a.0  6a   0    0  FC:B  34500/70656000 35003/71687368
data    6a.1  6a   0    1  FC:B  34500/70656000 35003/71687368

Plex /mir(1)/plex5 (offline, failed, out-of-date)
```

Next, the mirror is reestablished using the `aggr mirror -v` command.

Note

The node at the surviving site is called `filer1`; the node at the disaster site is called `filer2`.

```
filer1> aggr mirror mir -v mir(1)
This will destroy the contents of mir(1).  Are you sure? y
```

```
Mon Nov 18 15:36:59 GMT [filer1:
raid.mirror.resync.snapertok:info]: mir: created mirror
resynchronization snapshot mirror_resync.1118153658(filer2)
Mon Nov 18 15:36:59 GMT [filer1: raid.rg.resync.start:notice]:
/mir/plex6/rg0: start resynchronization (level 1)
Mon Nov 18 15:36:59 GMT [filer1: raid.mirror.resync.start:notice]:
/mir: start resynchronize to target /mir/plex6
```

After the aggregates rejoin, the synchronous mirrors of the MetroCluster configuration are reestablished.

```
filer1> aggr status -r mir
Aggregate mir (online, mirrored) (zoned checksums)
  Plex /mir/plex5 (online, normal, active)
    RAID group /mir/plex5/rg0 (normal)
```

RAID Disk Device	HA	SHELF	BAY	CHAN	Used (MB/blks)	Phys (MB/blks)
-----	-----	-----	-----	-----	-----	-----
parity 8a.2	8a	0	2	FC:B	34500/70656000	35003/71687368
data 8a.8	8a	1	0	FC:B	34500/70656000	35003/71687368

```
  Plex /mir/plex6 (online, normal, active)
    RAID group /mir/plex6/rg0 (normal)
```

RAID Disk Device	HA	SHELF	BAY	CHAN	Used (MB/blks)	Phys (MB/blks)
-----	-----	-----	-----	-----	-----	-----
parity 6a.0	6a	0	0	FC:B	34500/70656000	35003/71687368
data 6a.1	6a	0	1	FC:B	34500/70656000	35003/71687368:

Re-creating mirrored aggregates: To return the cluster to normal operation by re-creating the MetroCluster mirror, complete the following steps.

Step	Action
1	Validate that you can access the remote storage by entering the following command: <code>aggr status -r</code> Note_____ A (level-0 resync in progress) message indicates that a plex cannot be rejoined.

Step	Action
2	Turn on the power to the cluster node at the disaster site. Result: After the cluster node at the disaster site boots up, it displays the following: Waiting for Giveback...
3	If the aggregates at the disaster site are online, take them offline by entering the following command for each aggregate that was split: <pre>aggr offline disaster_aggr</pre> <i>disaster_aggr</i> is the name of the aggregate at the disaster site. Note_____ An error message appears if the aggregate is already offline. _____
4	Destroy every target plex that is in a level-0 resync state by entering the following command: <pre>aggr destroy plex_name</pre> For more information about the SyncMirror feature, see the <i>Data ONTAP Data Protection Online Backup and Recovery Guide</i>.
5	Re-create the mirrored aggregates by entering the following command for each aggregate that was split: <pre>aggr mirror plex_name</pre>
6	Enter the following command at the partner cluster node: <pre>cf giveback</pre> Result: The cluster node at the disaster site reboots.

Example of re-creating a mirrored aggregate: The following example shows the commands and status output when re-creating aggregates to reestablish the MetroCluster configuration.

First, the aggregate status of the disaster site's storage after reestablishing access to the cluster partner at the surviving site is shown.

```
filer1>aggr status -r
Aggregate mir1 (online, normal) (zoned checksums)
  Plex /mir1/plex0 (online, normal, active)
```

```
RAID group /mir1/plex0/rg0 (normal)
```

RAID Disk	Device	HA	SHELF	BAY	CHAN	Used (MB/blks)	Phys (MB/blks)
parity	8a.3	8a	0	3	FC:B	34500/70656000	35003/71687368
data	8a.4	8a	0	4	FC:B	34500/70656000	35003/71687368
data	8a.6	8a	0	6	FC:B	34500/70656000	35003/71687368
data	8a.5	8a	0	5	FC:B	34500/70656000	35003/71687368

```
Aggregate mir1(1) (failed, partial) (zoned checksums)
Plex /mir1(1)/plex0 (offline, failed, inactive)
```

```
Plex /mir1(1)/plex6 (online, normal, resyncing)
RAID group /mir1(1)/plex6/rg0 (level-0 resync in progress)
```

RAID Disk	Device	HA	SHELF	BAY	CHAN	Used (MB/blks)	Phys (MB/blks)
parity	6a.6	6a	0	6	FC:B	34500/70656000	35003/71687368
data	6a.2	6a	0	2	FC:B	34500/70656000	35003/71687368
data	6a.3	6a	0	3	FC:B	34500/70656000	35003/71687368
data	6a.5	6a	0	5	FC:B	34500/70656000	35003/71687368

The mir1(1)/plex6 plex shows that a level-0 resynchronization was in progress; therefore, an attempt to rejoin the plexes fails, as shown in the following output:

```
filer1> aggr mirror mir1 -v mir1(1)
aggr mirror: Illegal mirror state for aggregate 'mir1(1)'
```

Because the mir1(1)/plex6 plex had a level-0 resynchronization in progress, the mir1(1) aggregate must be destroyed and the mir aggregate remirrored to reestablish a synchronous mirror, as shown in the following output:

```
filer1> aggr mirror mir1 -v mir1(1)
aggr mirror: Illegal mirror state for aggregate 'mir1(1)'
filer1> aggr destroy mir1(1)
Are you sure you want to destroy this aggregate? y
Aggregate 'mir1(1)' destroyed.
filer1> aggr mirror mir1
Creation of a mirror plex with 4 disks has been initiated. The
disks need to be zeroed before addition to the aggregate. The
process has been initiated and you will be notified via the system
log as disks are added.
```


About this chapter

This chapter describes how to add storage to the cluster and how to upgrade or replace storage hardware.

Topics in this chapter

This chapter discusses the following topics:

- ◆ [“Adding disk shelves to a cluster”](#) on page 178
- ◆ [“Installing a disk shelf in an existing mixed-shelf loop”](#) on page 182
- ◆ [“Hot-adding a DS14, DS14mk2 FC, or DS14mk2 AT to an existing loop”](#) on page 187
- ◆ [“Upgrading or replacing LRC, ESH, ESH2, or AT-FCX modules”](#) on page 193
- ◆ [“Hot-swap disk loop error messages on Fabric Attached MetroClusters”](#) on page 206

Adding disk shelves to a cluster

About adding disk shelves to a cluster

For FAS3000 and FAS900 series systems: You can add a DS14/DS14mk2 FC disk shelf to a cluster that is made up of DS14/DS14mk2 FC disk shelves without powering down your cluster. This enables you to add (but not swap) disk shelves without a service interruption. However, you cannot add more disk shelves to your cluster than your cluster can support. See <http://now.netapp.com> for the amount of storage your cluster can support.

Attention

Hot-adding a disk shelf is different from hot-swapping a disk shelf. Hot-swapping a disk shelf, which means removing an existing disk shelf and installing a new one in its place, is not supported. Your system supports hot-adding of disk shelves only.

For F800 series systems: One FC-AL interface can connect up to eight daisy-chained disk shelves (four for DS14 disk shelves) per FC-AL adapter. You must add an additional FC-AL adapter to each of your storage systems for every set of eight disk shelves you add, up to the maximum of FC-AL adapters supported by your storage system. See the *F800 Series Field Service Guide* for FC-AL installation information.

Note

See the *System Configuration Guide* at <http://now.netapp.com> for the maximum number of FC-AL adapters your storage system supports, and for additional restrictions.

Restrictions for adding disk shelves to any cluster

The following table lists some of the restrictions for adding disk shelves to your cluster. See the *DiskShelf14 and DiskShelf14mk2 FC Hardware Guide* for a detailed discussion of restrictions.

If your cluster has...	And you are...	Then...
FC7 or FC8 disk shelves	Adding FC9 disk shelves in the same loop	Disk shelf 0 must be an FC7 disk shelf.
FC7, FC8, or FC9 disk shelves	Adding DS14/DS14mk2 FC disk shelves with LRCs	A DS14/DS14mk2 FC must be the last disk shelf in the loop.
DS14/DS14mk2 FC disk shelves with ESH or ESH2 modules	Hot-adding a disk shelf	◆ Disks should be pre-zeroed and used as spares. ◆ You can only hot-add a DS14/DS14mk2 FC disk shelf to an existing DS14/DS14mk2 FC disk shelf. ◆ A DS14/DS14mk2 FC must be the last disk shelf in the loop. ◆ DS14/DS14mk2 FC disk shelves with ESH or ESH2 modules can only be added to disk shelf loops with ESH or ESH2 modules. ◆ If the new disk shelf is a DS14mk2 FC, the loop speed switch must be set for the appropriate installation.
DS14mk2 AT disk shelves with AT-FCX modules	Hot-adding a disk shelf	◆ Disks should be pre-zeroed and used as spares. ◆ You can only hot-add a DS14mk2 AT disk shelf to an existing DS14mk2 AT disk shelf. ◆ DS14mk2 AT disk shelves with AT-FCX modules can only be added to disk shelf loops with AT-FCX modules.

Adding a disk shelf to a cluster

To add a disk shelf to a cluster, complete the following steps.

Step	Action	
1	Determine whether disk shelf counting is enabled by entering the following command: <code>options cf.takeover.on_disk_shelf_miscompare</code>	
2	If the disk shelf counting option is set to On, turn it off by entering the following command: <code>options cf.takeover.on_disk_shelf_miscompare off</code>	
3	Install the new disk shelf in your cabinet or equipment rack, as described in the <i>DiskShelf14 and DiskShelf14mk2 FC Hardware Guide</i> .	
4	Verify that the new disk shelf ID is unique by entering the following command: <code>fcstat device_map</code> Result: Shelf IDs for all Fibre Channel devices are listed.	
5	If...	Then, for loops with...
	Disk shelves in the loop to which you are adding the new disk shelf have LRCs	◆ DS14 only: Go to Step 6. ◆ A mix of DS14 and DS14mk2 FC: Set the loop speed to 1 Gb. Attention — An incorrectly set loop speed results in an open loop condition.
5	Disk shelves in the loop to which you are adding the new disk shelf have either ESH or ESH2 modules exclusively, or AT-FCX modules exclusively	◆ DS14 only or a mix of DS14 and DS14mk2 FC: Set the loop speed to 1 Gb. ◆ DS14mk2 FC only that are running at 1 Gb: Set the loop speed to 1 Gb. ◆ DS14mk2 FC only that are running at 2 Gb: Set the loop speed to 2 Gb. ◆ AT-FCX modules, the loop speed is set automatically Attention — An incorrectly set loop speed causes the storage system to panic. Disk shelves containing AT-FCX modules cannot be on the same loop as ESH/ESH2 modules.

Step	Action
6	Plug in the new disk shelf and turn it on. Then wait 60 seconds for the disk shelf to fully power up and all electronics to come online.
7	Attach the cable to the Input port of the new disk shelf.
8	Attach the other end of the cable to the Output port of the last existing disk shelf in the loop.
9	Set both terminate switches (except for ESH2 and AT-FCX, which don't have them) on the new disk shelf to On.
10	Set both terminate switches (except for ESH2 and AT-FCX) on the previous disk shelf to Off.
11	If disk shelf counting was on, reenable it by entering the following command: <code>options cf.takeover.on_disk_shelf_miscompare on</code>

Installing a disk shelf in an existing mixed-shelf loop

Assumptions about the existing mixed-shelf loop

This section is based on the assumption that the last disk shelf in the existing mixed-shelf loop connected to your storage system is not a DS14 or DS14mk2 FC disk shelf.

A DS14mk2 AT disk shelf is only compatible with similar shelves and cannot be added to a mixed-shelf loop.

Note

If the last disk shelf in your storage system is a DS14, DS14mk2 FC, or DS14mk2 AT, you can add the disk shelf to the loop without shutting down the system. Go to [“Hot-adding a DS14, DS14mk2 FC, or DS14mk2 AT to an existing loop”](#) on page 187.

Attention

Do not attempt to add a DS14/DS14mk2 FC disk shelf that has ESH or ESH2 modules to a mixed-shelf loop.

Installing the disk shelf

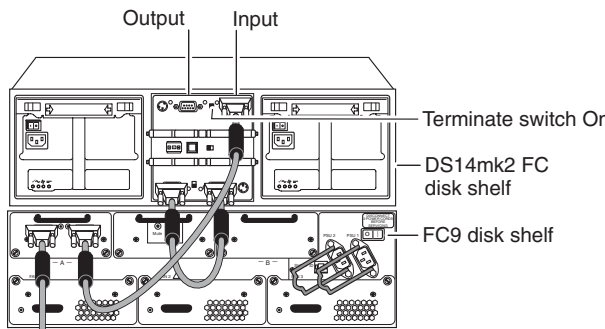
To install a DS14/DS14mk2 FC disk shelf into an existing mixed-shelf loop on a storage system, complete the following steps.

Attention

If you are installing a DS14 or DS14mk2 FC disk shelf into a loop that contains FC7, FC8, or FC9 disk shelves, the DS14 or DS14mk2 FC must be the last disk shelf in the loop and must have the highest disk shelf ID in the loop to avoid address conflicts.

Step	Action
1	Put on the antistatic wrist strap and grounding leash.

Step	Action	
2	Shut down the storage system by entering the following command at the console: <pre>halt</pre> Attention _____ Always use the <code>halt</code> command to perform a clean shutdown.	
3	Verify that the LCD display at the front of your storage system displays the following message: <pre>Halted</pre>	
4	Turn off the power to your storage system and the disk shelves.	
5	Install the disk shelf, attach grounding cables, and connect the disk shelf to a power source. See the <i>DiskShelf14 and DiskShelf14mk2 FC Hardware Guide</i> for more information.	
6	If the last disk shelf in the loop is...	Then...
	An FC7 or FC8 disk shelf	Remove the terminator from the Output/Terminator connector and go to Step 7 .
	An FC9 disk shelf	Go to Step 7 .
7	Connect one end of the provided cable to the Output/Terminator connector of the LRC A on the last disk shelf in the loop.	

Step	Action
8	Connect the other end of the cable to the LRC A Input on the disk shelf that you are adding. The following illustration shows a DS14mk2 FC disk shelf installed above an FC9 disk shelf.  The diagram illustrates the physical connection between two disk shelves. The top shelf is labeled 'DS14mk2 FC disk shelf' and has an 'Output' port on the left and an 'Input' port on the right. A cable is connected from the 'Input' port of the DS14mk2 shelf to the 'LRC A' input of the 'FC9 disk shelf' below it. The FC9 shelf has multiple ports, including 'LRC A' and 'LRC B'. A 'Terminate switch On' label points to a switch on the DS14mk2 shelf.

Step	Action	
9	If the DS14/ DS14mk2 FC...	Then...
	Is the only disk shelf being added to the loop	a. Connect one end of the provided cable to the LRC B Input on the disk shelf. b. Connect the other end of the cable to the LRC B Output on the disk shelf. c. Turn the terminate switches on LRCs A and B to the On position. LRC A On terminate position Note _____ Because LRC B on the disk shelf is inverted, the On termination position on LRC B is up instead of down. _____
	Is the first of additional DS14/DS14mk2 FC disk shelves being added to the loop	a. Turn the terminate switch on the first disk shelf in the mixed-shelf loop to the Off position for LRC A (up) and LRC B (down). b. Connect one end of the provided cable to the LRC B Output on the first disk shelf in the mixed-shelf loop. c. Connect the other end of the cable to the LRC B Input on the next disk shelf that you added to the mixed-shelf loop. d. Verify that the terminate switch on all but the last disk shelf is in the Off position for LRC A (up) and LRC B (down). The terminate switches on the last disk shelf should be in the On position.

Step	Action	
10	Give the disk shelf a unique disk shelf ID. See the <i>DiskShelf14</i> and <i>DiskShelf14mk2 FC Hardware Guide</i> for additional information. Note _____ A valid shelf ID is 1 through 7. Only use disk shelf ID 7 if it is the last ID available on this loop. A shelf ID that is not from 1 through 7 causes the drive addresses to default to those of a shelf with the ID switch set to 7, even though the Shelf ID indicator in the front operation panel displays a dash (-).	
11	If your disk shelf is a...	Then...
	DS14mk2 FC with an LRC module	Set the loop speed switch to 1 Gb. Attention _____ An incorrectly set loop speed results in an open loop condition.
	DS14	Go to Step 12 .
12	Turn on the power first to the disk shelves and then to the storage system.	
13	Reboot the storage system.	

Hot-adding a DS14, DS14mk2 FC, or DS14mk2 AT to an existing loop

About this section This section provides information about how to hot-add a DS14, DS14mk2 FC, or DS14mk2 AT disk shelf to an existing loop. It also tabulates the error messages that appear on your storage system console if the attempt at hot-adding was unsuccessful.

Note The hot-added disk shelf must have new drives that will be used as spares only.

Requirements for hot-adding To hot-add a DS14, DS14mk2 FC, or DS14mk2 AT to an existing loop, the following requirements must be met.

If the disk shelf being hot-added...	Then...
Has LRC modules	◆ The disk shelves in the existing loop must have a DS14, DS14mk2 FC, or DS14mk2 AT disk shelf with LRCs as the last shelf. ◆ You must follow the procedure “Hot-adding a disk shelf to an existing loop” on page 188 exactly, if you are hot-adding to a mixed-shelf loop. Attention Failure to follow the procedure exactly causes the last FC9 in the loop to automatically terminate the loop. The storage system loses contact with the existing last DS14, DS14mk2 FC, or DS14mk2 AT in the loop and crashes.
Has ESH or ESH2 modules	The disk shelves in the existing loop must have only DS14, DS14mk2 FC, or DS14mk2 AT disk shelves with ESH or ESH2 modules in the loop.
Has AT-FCX modules	The disk shelves in the existing loop must have only DS14mk2 AT disk shelves with AT-FCX modules in the loop.

Hot-adding a disk shelf to an existing loop

To hot-add a disk shelf to an existing loop, complete the following steps.

Step	Action
1	Put on the antistatic wrist strap and grounding leash.
2	Install the disk shelf. <i>See the DiskShelf14, DiskShelf14mk2 FC, or DiskShelf14mk2 AT Hardware Guide for more information.</i>
3	Ground the disk shelf using the provided grounding cable. <i>See the DiskShelf14, DiskShelf14mk2 FC, or DiskShelf14mk2 AT Hardware Guide for more information.</i>
4	Connect the disk shelf to a power source. <i>See the DiskShelf14, DiskShelf14mk2 FC, or DiskShelf14mk2 AT Hardware Guide for more information.</i>
5	Give the disk shelf a unique disk shelf ID. <i>See the DiskShelf14, DiskShelf14mk2 FC, or DiskShelf14mk2 AT Hardware Guide for more information.</i> Note _____ Only use disk shelf ID 1 through 6. Disk shelf ID 7 is reserved for other purposes. _____
6	Verify that the disk shelf ID is not being used in the loop by entering the following command at the console: <code>fcstat device_map adaptername</code> Attention _____ An invalid disk shelf ID causes the storage system to panic. _____

Step	Action	
7	If disk shelves in the loop have...	Then, for loops with...
	LRC modules	◆ DS14 only: Go to Step 8. ◆ A mix of DS14 and DS14mk2 FC: Set the loop speed to 1 Gb. Attention _____ An incorrectly set loop speed results in an open loop condition.
	ESH, ESH2, or AT-FCX modules	◆ DS14 only or a mix of DS14 and DS14mk2 FC: Set the loop speed to 1 Gb. ◆ DS14mk2 FC only that are running at 1 Gb: Set the loop speed to 1 Gb. ◆ DS14mk2 FC only that are running at 2 Gb: Set the loop speed to 2 Gb. Attention _____ An incorrectly set loop speed causes the storage system to panic. _____ ◆ DS14mk2 AT: Loop speed is fixed and automatically set at 2 Gb.

Step	Action	
8	If the disk shelf you added has...	Then...
	ESH2 modules	Turn on the power to the disk shelf and wait 60 seconds for the electronics to finish initializing.
	LRC or ESH modules	Turn on the power to the disk shelf and wait 60 seconds for the electronics to finish initializing. Attention ————— You must power up the new disk shelf before you connect it to the last disk shelf in the loop because there must be approximately a 60-second time difference between Step 8 (powering on the disk shelf) and Step 14 (turning off the terminate switch) of this procedure.
	AT-FCX modules	Turn on the power to the disk shelf and wait 60 seconds for the electronics to finish initializing. Note ————— The AT-FCX module is self terminating as long as no cable is plugged into the Out port of the last shelf.
9	Connect one end of the provided cable to the LRC, ESH, ESH2, or AT-FCX Channel A Input on the new disk shelf.	
10	Connect the other end of the cable to the LRC, ESH, ESH2, or AT-FCX Channel A Output on the last disk shelf in the existing loop.	
11	Connect one end of the provided cable to the LRC, ESH, ESH2, or AT-FCX Channel B Input on the new disk shelf.	
12	Connect the other end of the cable to the LRC, ESH, ESH2, or AT-FCX Channel B Output on the last disk shelf in the existing loop.	

Step	Action	
13	Verify that all the cables are securely fastened. Attention ————— Poorly secured cables cause the storage system to panic over an open loop.	
14	If the disk shelf you added has...	Then...
	ESH2 or AT-FCX modules	Result: In 60 seconds, the storage system recognizes the hot-added disk shelf.
	LRC or ESH modules	a. Turn the terminate switch of the newly added disk shelf to the On position for LRC or ESH Channel A (down) and LRC or ESH Channel B (up). Attention ————— Failure to place the terminate switches in the On position causes the storage system to panic over an open loop. b. Turn the terminate switch of the now second-to-last disk shelf to the Off position for LRC or ESH Channel A (up) and LRC or ESH Channel B (down). Result: In 60 seconds, the storage system recognizes the hot-added disk shelf.

Error messages

The following error messages appear on your storage system console if your attempt at hot-adding the DS14/DS14mk2 FC is unsuccessful.

Error message	Explanation
Open loop panic	One of three reasons cause this error message to appear: ◆ The terminate switch on LRC or ESH Channel A and LRC or ESH Channel B on the new disk shelf are not in the On position. ◆ The shelf-to-shelf cable between the now second-to-last disk shelf and the newly added disk shelf is defective or is not securely fastened. ◆ The speed of the newly added DS14mk2 FC disk shelf is incorrectly set.
Soft address panic	One of two reasons cause this error message to appear: ◆ There is an invalid disk shelf ID. ◆ The power was turned on before the disk shelf ID was changed and the disk shelf was not power-cycled after the disk shelf ID was changed.
Speed mismatch termination	The ESH modules on the disk shelf detected a speed mismatch between the preceding disk shelf and this disk shelf and is reporting them as automatic termination errors.

Upgrading or replacing LRC, ESH, ESH2, or AT-FCX modules

About the modules The LRC, ESH, ESH2, or AT-FCX module in a DS14, DS14mk2 FC, or DS14mk2 AT includes a SCSI-3 Enclosure Services Processor. It maintains the integrity of the loop when disks are swapped and provides signal retiming for enhanced loop stability. There are two modules in the middle of the rear of the disk shelf, one for Channel A and one for Channel B.

Note _____
The Input and Output ports on module B on the DS14/DS14mk2 FC are the reverse of module A.

Connectors in an LRC, ESH, ESH2, or AT-FCX: The module has the following connectors.

Module connector	Function
Input	Provides the interface between the disk shelf and the storage system.
Output	Provides the interface between two disk shelves to create a loop of daisy-chained disk shelves.

Restrictions for changing module types If you plan to change the type of any module in your cluster, make sure you understand the following restrictions about module types:

- ◆ You cannot mix LRC and ESH modules in the same loop. Doing so will result in loop failure.
- ◆ To replace an ESH with an ESH2, you must
 - ❖ be running Data ONTAP 6.4.4 or later
 - ❖ replace all ESH modules in the same shelf with ESH2 modules

Best practices for changing modules The following list describes important guidelines and requirements you should understand and practice when changing modules in your cluster.

- ◆ Whenever you remove a module from a cluster, you need to know whether the path you will disrupt is redundant. If it is, you can remove the module without interfering with the storage system’s ability to serve data. However,

if that module provides the only path to any disk in your cluster, you must take action to ensure that you do not incur system downtime.

For more information, see “[Understanding single and dual pathing in clusters](#)” on page 196.

- ◆ When you replace a module, make sure that the replacement module’s termination switch is in the same position as the module it is replacing.

Note

ESH2 modules are self-terminating; this guideline does not apply to ESH2 modules.

- ◆ If you replace a module with a different type of module, make sure that you also change the cables if necessary.
For more information about supported cable types, see the Hardware Guide for your disk shelf.
- ◆ Always wait 30 seconds after inserting any module before reattaching any cables in that loop.

Testing the modules

To test your LRC, ESH, ESH2, or AT-FCX modules after you replace or upgrade them, complete the following steps on both nodes.

Step	Action
1	Verify that all disk shelves with LRC, ESH, ESH2, or AT-FCX modules are functioning properly by entering the following command: <code>environ shelf</code> Results: The system reports the modules as good.
2	Verify that there are no missing disks by entering the following command: <code>aggr status -r</code> Result: Local disks displayed on the local node should be displayed as partner disks on the partner node, and vice-versa.
3	Verify that you can create and retrieve files on both nodes for each licensed protocol.

For detailed information

For details about upgrading or replacing LRC, ESH, ESH2, or AT-FCX modules, see the following topics:

- ◆ [“Understanding single and dual pathing in clusters”](#) on page 196
- ◆ [“Upgrading to ESH or ESH2 modules”](#) on page 199
- ◆ [“Hot-swapping an LRC, ESH, ESH2, or AT-FCX module”](#) on page 203

Understanding single and dual pathing in clusters

What a dual-path configuration is

Some clusters have two paths to each of their disks; this is called a dual-path configuration. Fabric Attached MetroClusters and clusters using the SharedStorage functionality are dual-path configurations.

Why understanding your paths is important

If you want to remove a module from your cluster, you need to know whether the path you will disrupt is redundant. If it is, you can remove the module without interfering with the storage system’s ability to serve data. On the other hand, if that module provides the only path to any of the disks in your cluster, you must take action to ensure that you do not incur system downtime.

How to determine path status for your cluster

You can determine whether any module in your system provides the only path to any disk by using the `storage show disk -p` command at your system console. This command displays the following information for every disk in the cluster:

- ◆ Primary port
- ◆ Secondary port
- ◆ Shelf
- ◆ Bay

Dual-path example output: The following example shows what the `storage show disk -p` command output might look like for a dual-path cluster:

PRIMARY	PORT	SECONDARY	PORT	SHELF	BAY
-----	----	-----	----	-----	-----
0c.112	A	0b.112	B	7	0
0b.113	B	0c.113	A	7	1
0b.114	B	0c.114	A	7	2
0c.115	A	0b.115	B	7	3
0c.116	A	0b.116	B	7	4
0c.117	A	0b.117	B	7	5
0b.118	B	0c.118	A	7	6
0b.119	B	0c.119	A	7	7
0b.120	B	0c.120	A	7	8
0c.121	A	0b.121	B	7	9
0c.122	A	0b.122	B	7	10
0b.123	B	0c.123	A	7	11

Notice that every disk (e.g. 0c.112/0b.112) has two ports active. This means that you do not need to fail over one system before removing modules from the system.

Attention

Make sure that *every* disk has two paths. Even in a cluster configured for dual paths, a hardware or configuration problem can cause one or more disks to have only one path. If any disk in your cluster has only one path, you must treat that loop as if it were in a single-path cluster when removing modules.

Example output for a cluster that does not have dual paths: The following example shows what the `storage show disk -p` command output might look like for a cluster that does not use dual paths:

```
filer1> storage show disk -p
PRIMARY PORT SECONDARY PORT SHELF BAY
-----
```

5b.16	B			1	0
5b.17	B			1	1
5b.18	B			1	2
5b.19	B			1	3
5b.20	B			1	4
5b.21	B			1	5
5b.22	B			1	6
5b.23	B			1	7
5b.24	B			1	8
5b.25	B			1	9
5b.26	B			1	10
5b.27	B			1	11
5b.28	B			1	12
5b.29	B			1	13
5b.32	B			2	0
5b.33	B			2	1
5b.34	B			2	2
5b.35	B			2	3
5b.36	B			2	4
5b.37	B			2	5
5b.38	B			2	6
5b.39	B			2	7
5b.40	B			2	8
5b.41	B			2	9
5b.42	B			2	10
5b.43	B			2	11
5b.44	B			2	12
5b.45	B			2	13

For this cluster, there is only one path to each disk. This means that you cannot remove a module from the cluster, thereby disabling that path, without first performing a takeover.

Upgrading to ESH or ESH2 modules

Reasons for upgrading

If your cluster has disk shelves that have LRC modules and you add a DS14/DS14mk2 FC disk shelf that has ESH or ESH2 modules, you need to replace the LRC modules with ESH or ESH2 modules in the existing DS14 disk shelves.

What disk shelf upgrades and downgrades are supported

The following table describes which types of hot-upgrade and hot-downgrade each type of diskshelf supports.

Attention — You must hot-upgrade or hot-downgrade all disk shelves in the loop.

Disk shelf loop speed	Hot-upgrade	Hot-downgrade
DS14 (only available at 1-Gb loop speed)	LRC to ESH or ESH2 hot-upgrade is supported.	ESH or ESH2 to LRC hot-downgrade is supported.
DS14mk2 FC set at 1-Gb loop speed	LRC to ESH or ESH2 hot-upgrade is supported.	ESH or ESH2 to LRC hot-downgrade is supported.
DS14mk2 FC set at 2-Gb loop speed	◆ Not applicable ◆ ESH to ESH2 is supported. See“Hot-swapping a module” on page 203.	◆ ESH or ESH2 to LRC hot-downgrade is not supported. ◆ ESH2 to ESH hot-downgrade is supported.

Upgrading an LRC module to an ESH or ESH2 module

To upgrade your DS14/DS14mk2 FC disk shelves to use ESH or ESH2 modules, complete the following steps. Keep in mind that at any given time a storage system that is up and running needs to have at least one path to any shelf.

Note

Do not mix LRC and ESH modules in the same loop; loop failure results.

Replacing an ESH with an ESH2 requires that you upgrade to Data ONTAP 6.4.4 or later, and replace both ESH modules in the disk shelf with ESH2 modules.

For more information about upgrading disk shelves, see the *DiskShelf14*, *DiskShelf14mk2 FC*, or *DiskShelf14mk2 AT Hardware Guide*.

Step	Action
1	Determine which loop you will upgrade first, and determine whether any disks in the cluster are single-pathed through that loop. For more information, see “ How to determine path status for your cluster ” on page 196.
2	If any disks use this loop for their only path to a controller, complete the following steps: a. On one node, which we will call Node A, enter the following command: <code>cf takeover</code> b. Wait for takeover to be complete and make sure that the partner node, or Node B, reboots and is waiting for giveback. Result: You can now replace all modules on all loops attached to Node B.
3	Note the terminate switch position of the LRCs you are replacing in this loop. If the new modules have terminate switches, they should be set to the same setting as the LRCs they are replacing. Note _____ The ESH2 is self-terminating and does not have a terminate switch. _____
4	While grounded, unplug the cabling from the dormant loop and note the cable locations.
5	Remove the LRC modules from the disk shelves that do not have a cable attached.

Step	Action
6	Insert all of the new modules. Attention _____ After inserting all modules, wait 30 seconds before proceeding to the next step. _____
7	Recable the disk shelves to their original locations.
8	Check the operation of the new modules by entering the following command from the console of the node that is still running: <pre>environ shelf</pre> Result: The node reports the status of the modified shelves as good.
9	If there is more than one loop attached to the node that has been taken over, repeat Step 3 through Step 8 for all other loops on that node.
10	If you performed a cluster takeover previously, complete the following steps: - At the console of the takeover node, return control of Node B's disk shelves by entering the following command: <pre>cf giveback</pre> - Wait for the giveback to complete before proceeding to the next step.
11	Check the operation of the new modules by entering the following command on the console of both nodes: <pre>environ shelf</pre> Result: Each node reports the status of the modified shelves as good.
12	Check the next loop you will upgrade to determine whether any disks in the cluster are single-pathed through that loop. For more information, see "How to determine path status for your cluster" on page 196

Step	Action
13	If any disks use this loop for their only path to a controller, complete the following steps: a. On Node B, enter the following command: <code>cf takeover</code> b. Wait for takeover to be complete and make sure that the partner node, or Node A, reboots and is waiting for giveback. Result: You can now replace all modules on all loops attached to Node A.
14	Repeat Step 3 through Step 11 , with Node B as the takeover node.
15	Test the upgraded modules, as described in “ Testing the modules ” on page 194.
16	Test the configuration, as described in “ Stage 6: Checking and testing cluster configuration ” on page 112.

Hot-swapping an LRC, ESH, ESH2, or AT-FCX module

When to use this procedure

You use this procedure when you need to replace a faulty module.

Note—————
When you hot-swap a LRC, ESH, ESH2, or AT-FCX module, you must ensure that you never disable the only path to a disk, which will result in a system outage.

For more information, see “[Understanding single and dual pathing in clusters](#)” on page 196.

Hot-swapping a module

To hot-swap a single module, complete the following steps.

Step	Action
1	Verify that your storage system meets the minimum software requirements to support the disk shelf modules that you are hot-swapping. See the <i>DiskShelf14</i> , <i>DiskShelf14mk2 FC</i> , or <i>DiskShelf14mk2 AT Hardware Guide</i> for more information.
2	Determine which loop contains the module you are removing, and determine whether any disks in the cluster are single-pathed through that loop. For more information, see “ How to determine path status for your cluster ” on page 196.
3	If you performed a cluster takeover previously, complete the following steps: a. Return control of Node A’s disk shelves by entering the following command at Node B’s console: cf giveback b. Wait for the giveback to complete before proceeding to the next step.

Step	Action
4	If any disks use this loop for their only path to a controller, complete the following steps: - Follow the cables from the module you want to replace back to one of the cluster nodes. We will call this node Node A. - At the <i>Node B</i> console, enter the following command: cf takeover - Wait for takeover to be complete and make sure that the partner node, or Node A, reboots and is waiting for giveback. Result: The modules in the loop that are attached to Node A can now be replaced.
5	Put on the antistatic wrist strap and grounding leash.
6	Disconnect the module that you are removing from the Fibre Channel cabling.
7	Using the thumb and index finger of both hands, press the levers on the CAM mechanism on the module to release it and pull it out of the disk shelf.
8	Slide the replacement module into the slot at the rear of the disk shelf and push the levers of the CAM mechanism into place. Attention _____ Do not use excessive force when sliding the module into the disk shelf; you might damage the connector. Wait 30 seconds after inserting the module before proceeding to the next step. _____
9	Reconnect the Fibre Channel cabling.
10	Check the operation of the new modules by entering the following command from the console of both nodes: environ shelf

Step	Action
11	Test the replacement module, as described in “ Testing the modules ” on page 194.
12	Test the configuration, as described in “ Stage 6: Checking and testing cluster configuration ” on page 112.

Hot-swap disk loop error messages on Fabric Attached MetroClusters

About the error messages

A Fabric Attached MetroCluster generates error messages when a disk drive is removed from or added to a disk loop. These messages are informational and require no action from you.

Messages and descriptions

The following table lists and describes informational error messages you might encounter when removing or inserting a disk into a loop.

Error message	Description
[node1: ispf_timeout_1:warning]: Met1_sw1:11.34 (NxN2N2Nbae) (NxN359da8N,Nx28:NN6Ne3dN:N1NN,N/N,4951/N/N ,29927/N): command timeout, quiescing drive to allow outstanding I/O to complete.	When the disk drive is pulled from the bay, data flow to all drives in the loop is interrupted. The error message indicates that the data flow was interrupted and the system is trying to reestablish contact with that loop.
[node1: ispf_timeout_1:error]: node2:11.34 (NxN2N2Nbae): global device timer timeout, initiating device recovery.	The system is attempting to recover from the data flow interruption.
[node1: scsi.cmd.abortedByHost:error]: Device Met1_sw1:11.34: Command aborted by host adapter: HA status Nx4: cdb Nx28:NN6Ne3dN:N1NN.	The current command to the drive loop is aborted. The system is continuing to recover the drive loop.
[node1: scsi.cmd.retrySuccess:info]: Device Met1_sw1:11.39: request successful after retry #1: cdb Nx28:N162416N:N1NN.	Contact was reestablished with the disk loop.
[node1: ses_admin:info]: initiating rescan on channel Met1_sw1:11.	The system is scanning the disk loops to determine what was added or removed.
[node1: raid.disk.inserted:info]: Disk Met1_sw1:11.72 (S/N 3HZNL8HLNNNN7323DX8A) has been inserted into the system.	The system found a new disk drive added to the loop.
[node1: dyn_dev_qual_admin:info]: Firmware is up- to-date on all disk drives.	The system checked the firmware revisions on all disks and found that it is up to date.

About this chapter

This chapter describes the process of changing hardware components using the nondisruptive cluster upgrade method.

For more information about nondisruptive Data ONTAP upgrades of clustered systems, see the *Upgrade Guide*.

Note

Nondisruptive hardware upgrades to FCP attached storage systems are not supported.

See the hardware documentation for your storage systems before upgrading hardware components.

When hardware upgrades can be performed

You can perform hardware maintenance and upgrades using the nondisruptive clustered upgrade method. You can perform nondisruptive hardware upgrades independently or in conjunction with a Data ONTAP upgrade of clustered systems.

If hardware and software are upgraded together, the new software must be downloaded first. If you are upgrading within a release family, you can upgrade the hardware and Data ONTAP together. If you are upgrading between release families, you need to upgrade Data ONTAP first, which might cause a small disruption, but you can use the nondisruptive upgrade procedure to cover the longer period required for hardware upgrades.

In a nondisruptive upgrade, each node is successively “failed” and its partner is put in takeover state at one point in the procedure. While the “failed” node is in the `waiting for giveback` state, it can be interrupted and brought to the `firmware ok` prompt or powered off.

Replacing a component nondisruptively

You can use the nondisruptive cluster upgrade method to perform the following hardware upgrade procedures:

- ◆ Replacing the head (with the same type as before, and with the same adapters)
- ◆ Replacing the motherboard

- ◆ Replacing or adding an adapter
You can replace NVRAM, disk, or NIC components, either with the same component or with an improved component (for example, you can upgrade from 2-port to 4-port Gigabit Ethernet (GbE), or 1-port to 2-port Fibre Channel).
- ◆ Replacing the cluster interconnect card
- ◆ Installing onboard firmware on various platforms (for example, the Baseboard Management Controller (BMC) on FAS900 series storage systems)

You can also run diagnostics on various components of the “failed” machine (for example, motherboard or NVRAM).

Note

Running diagnostics on parts of the machine that are still visible to or in use by the partner—notably disks and the cluster interconnect card—can present problems. See the *Diagnostics Guide*.

Replacing a component nondisruptively involves the following tasks:

- ◆ Removing the component
- ◆ Installing the new component

Removing the component: To remove a hardware component from a node, complete the following steps.

Step	Action
1	Put on an antistatic wrist strap.
2	Take over the node that has the component you want to replace by entering the following command from its partner’s console: <code>cf takeover</code> Result: The node takes over the partner node. You see a message similar to the following: <code>takeover completed</code>
3	Turn off the power to the target node, then unplug it from the power source.

Step	Action
4	Remove any applicable cabling to the component, then remove the component, as described in the <i>Hardware and Service Guide</i> or the hardware documentation for that component.

Installing the new component: To install a component in a node, complete the following steps.

Step	Action
1	While grounded, install the new component, as described in the <i>Hardware and Service Guide</i> or the hardware documentation for that component.
2	Close the node and reconnect any applicable cables to the component.
3	Plug the node back into the power source, and then apply power.
4	After the node boots and is at the ‘Waiting for Giveback’ prompt, run giveback by entering the following command on the local node’s console: <pre>cf giveback</pre> Result: The partner node reboots and functions normally. A successful giveback ends with the following message on the local node: <pre>giveback completed</pre>

About this appendix This appendix lists the regulatory notices you need to be aware of when installing and operating your NetApp equipment.

Regulatory notices

FCC notices (U.S. only)

Network Appliance devices are designed for a CFR 47 (Code Federal Regulations) Part 15 Class A environment.

The FCC and Network Appliance guarantee the user's rights to operate this equipment only if the user complies with the following rules and regulations:

- ◆ Install and operate this equipment in accordance with the specifications and instructions in this guide.
- ◆ Modify this equipment only in the ways specified by Network Appliance.
- ◆ Use shielded cables with metallic RFI/EMI connector hoods to maintain compliance with applicable emissions standards.
- ◆ If the system has nine or more fibre channel disk shelves, install the system in two or three NetApp System Cabinets to maintain performance within Part 15 of CFR 47 regulations.

Compliance with Part 15 of CFR 47

This equipment has been tested and found compliant with Part 15 of the CFR 47 rules for Class A digital devices. These rules are designed to provide reasonable protection from interference to electronics equipment operated in a commercial environment.

Operation of this device is subject to the following two conditions:

- ◆ This device cannot cause harmful interference.
- ◆ This device must accept any interference received, including interference that may cause undesired operation.

Compliance with ICES-003

This Class A digital apparatus complies with Canadian ICES-003.

Cet appareil numérique de la classe A conforme à la norme NMB-003 du Canada.

Compliance with EN regulations

Marking by the symbol **CE** indicates compliance of this Network Appliance device to the EMC Directive and the Low Voltage Directive of the European Union. Such marking is indicative that this Network Appliance device meets the technical standards listed in “Declaration of Conformity,” later in this appendix.

Attention

This is a Class A product. In a domestic environment this product may cause radio interference, in which case the user may be required to take adequate measures.

Bureau of Standards, Metrology, and Inspections notice (BSMI, Taiwan only)

警告使用者：

這是甲類的資訊產品，在居住的環境中使用時，可能會造成射頻干擾，在這種情況下，使用者會被要求採取某些適當的對策。

Translation of the BSMI notice:

Warning: This is a Class A product. In a domestic environment this product may cause radio interference, in which case the user may be required to take adequate measures.

Voluntary Control Council for Interference by Information Technology Equipment (VCCI, Japan)

この装置は、情報処理装置等電波障害自主規制協議会（VCCI）の基準に基づくクラスA情報技術装置です。この装置を家庭環境で使用すると電波妨害を引き起こすことがあります。この場合には使用者が適切な対策を講ずるよう要求されることがあります。

Translation of the VCCI-A notice:

This is a Class A product based on the standard of the Voluntary Control Council for Interference by Information Technology Equipment (VCCI). If this equipment is used in a domestic environment, radio disturbance may arise. If such trouble occurs, the user may be required to take corrective actions.

Declaration of Conformity

**Network Appliance, Inc.,
495 East Java Drive
Sunnyvale, California, 94089, U.S.A.**

declare under our sole responsibility that the products

Type of equipment	Description	Model number	Year of introduction
Network Appliance	NetApp FAS3000 series storage systems	FAS3020, FAS3050	2005
	NetApp FAS900 series storage systems	FAS920, FAS940, FAS960, FAS980	2003, 2004
	NetApp F810 storage system	F810	2001
	NetApp F820 storage system	F820	2000
	NetApp F825 storage system	F825	2002
	NetApp F840 storage system	F840	2000
	NetApp F880 storage system	F880	2001

to which this declaration relates conform to the following standards:

EN 60950:2002, Information Technology Equipment (Safety)

EN 55022:1998, Emissions Requirements for Information Technology Equipment

EN 50024:1998, Immunity Requirements for Information Technology Equipment

EN 60825-1: 1994+All, Safety of Laser/LED Equipment

EN 61000-3-2:2000 Limits for Harmonic Current Emissions

EN 61000-3-3:1995/A1:2001 Limitation of Voltage Fluctuations and Flicker in Low-Voltage Supply Systems

following the provisions of the directives listed below:

73/23/EEC, Low Voltage Directive (Product Safety)

89/336/EEC, Electromagnetic Compatibility Directive

Part number: 210-00583

Feature update history

The following table lists and describes the history of changes made to this manual. When a change is implemented, it applies to the release in which it was implemented and all subsequent releases, unless otherwise specified.

Feature updates	Feature first implemented in	Feature release date
◆ Updates for FAS920 ◆ Update for NVRAM5 ◆ Illustration updates	Data ONTAP 6.5.1	May 2004
◆ Updates for NVRAM5 support in FAS900 series clusters, except for MetroCluster ◆ Failover event cause-and-effect table ◆ Declaration of Conformity update ◆ Addition of cluster failover and single-point-of-failure analysis	Data ONTAP 7.0	November 2004
◆ FAS3000 series clustering information ◆ Corrections were made to the <i>Upgrading an LRC to ESH/ESH2/AT-FCX</i> procedure	Data ONTAP 7.0.1	April 2005
◆ Incorporation of the Cluster Administration chapter from the <i>Data ONTAP System Administration Guide</i> and the Disaster Protection Using MetroCluster appendix from the <i>Data ONTAP Data Protection Online Backup and Recovery Guide</i>.	Data ONTAP 7.1	June 2005
◆ Updated MetroCluster information for FAS3000 series	Data ONTAP 7.1	October 2005
◆ Updated module replacement information ◆ Fixed problem in Brocade switch configuration information	Data ONTAP 7.1	December 2005

Index

A

adapters

- cabling F800 series cluster 65
- cabling FAS3000 series cluster 44
- cabling FAS900 series cluster 56
- dual-port FC-AL 45
- installing 209
- NVRAM5 3, 44, 56
- removing 208

adding a disk shelf

- to mirrored cluster 179
- to standard cluster 180

aggregates

- re-creating mirrored after disaster 173
- rejoining after disaster 171

AT-FCX modules, about 193

audience xiii

B

bring up 105

- configuring interfaces for 110
- first time 96
- manually setting options for 108

C

cabling

- Channel A loop in mirrored FAS3000 cluster 38
- Channel A loop in mirrored FAS900 cluster 50
- Channel A loop in standard F800 cluster 59
- Channel A loop in standard FAS3000 cluster 36
- Channel A loop in standard FAS900 cluster 47
- Channel B loop in mirrored FAS3000 cluster 42
- Channel B loop in mirrored FAS900 cluster 54
- Channel B loop in standard F800 cluster 62
- Channel B loop in standard FAS3000 cluster 40
- Channel B loop in standard FAS900 cluster 52

dual-port FC-AL 45

- FC-ALs in FAS900 series mirrored cluster 45
- location of FAS900 dual-port FC-ALs 45
- preparing equipment racks for 34
- preparing system cabinets for 32
- required 30, 69

cf giveback command 139

cf-config-check.cgi utility 154

checking configuration through a utility 154

CIFS sessions terminated on takeover 117

cluster adapter

- cabling for F800 series 65
- cabling for FAS3000 series 44
- cabling for FAS900 series 56

cluster configuration checker 154

Cluster failover (CFO) 17

cluster types

- comparison 2
- Fabric Attached MetroCluster 12
- installed in equipment racks 28
- installed in system cabinets 28
- mirrored 7
- nonswitched MetroCluster 9
- standard 3

clusters

- benefits of 3
- characteristics of 5
- configuration information for 120
- configuring interfaces for 110
- definition of 1
- disabling 157
- displaying disk information 121
- enabling and disabling takeover (cf) 121
- first-time bring up of 96
- hardware components described 4
- managing in normal mode 118

command exceptions for emulated nodes 132

commands

- cf (enables and disables takeover) 121
- cf forcegiveback 141
- cf forcetakeover (forces a takeover) 124
- cf forcetakeover (forces takeover) 168

- cf giveback 139
- cf giveback (enables giveback) 114
- cf partner (displays partner's name) 120
- cf status (determines status) 129
- cf status (monitors status) 118
- cf takeover (enables takeover) 112
- cf takeover (initiates takeover) 123
- cfclear clears configuration) 75
- cfgdisable (disables configuration) 75
- cfgsave (saves configuration) 75
- configshow (shows switch settings) 76
- configure (switch parameters) 75
- halt (halts system without takeover) 123
- license add (license cluster) 103
- partner (accesses emulated node) 130
- storage show disk -p 196
- switchdisable (disables switch) 75
- switchenable (enables switch) 76
- switchname (sets switch name) 76
- sysconfig/aggr status (displays disk information) 121
- takeover (description of all takeover commands) 123
- version (displays switch firmware) 74
- communications regulations 211
- configuration
 - displaying information for cluster 120
 - reestablishing after disaster 170
 - testing 112
- configuring
 - dedicated and standby interfaces 151
 - interfaces 111
 - shared interfaces 149
- Controller failover 17
- conventions
 - command xv
 - formatting xvi
 - keyboard xv

D

Data ONTAP

- in a Fabric Attached MetroCluster 14
- in a mirrored cluster 11

- in a nonswitched MetroCluster 11
- in a standard cluster 3
- dedicated and standby interfaces, diagram 146
- dedicated interfaces
 - configuring using ifconfig 151
 - configuring using setup 99
 - described 144
- dedicated interfaces, configuring 151
- definitions
 - appliance xiv
 - disk shelf xiv
 - F800 xiii
 - Fabric Attached MetroCluster xiv
 - FAS3000 xiii
 - FAS900 xiii
 - FC HBA xiv
 - FC HBA for Disk xiv
 - mirrored cluster xiv
 - nonswitched MetroCluster xiv
 - switch fabric 12
- diagnostics, running 101
- disabling clustering 157
- disaster recovery
 - forcing takeover 168
 - manual fencing 167
 - restricting access to failed node 166
 - steps 166
 - using MetroCluster 163
 - when not to perform 164
- disasters
 - determining whether occurred 164
 - recognizing 164
- disk information, displaying 121
- disk shelf
 - about adding 178
 - adding to cluster 180
 - adding to mirrored clusters 179
 - adding to standard cluster 180
 - installing in existing mixed-shelf loop 182
 - restrictions for adding 179
 - upgrades supported 199
 - upgrading modules for 199
- documentation, required 29, 68
- dual paths, about 196

E

- eliminating single-point-of-failure 18
- emulated LANs, considerations in clusters 148
- emulated node
 - accessing from takeover node 130
 - accessing from takeover node (partner) 130
 - accessing remotely 132
 - command exceptions 132
 - description of 130
 - dumps and restores 135
- equipment racks
 - installing in 28
 - preparation of 34
- ESH modules, about 193
- ESH2 modules, about 193

F

- Fabric Attached MetroClusters
 - advantages of 13
 - behavior of Data ONTAP with 14
 - cabling the local node and disk shelves in 78
 - cabling the remote node and disk shelves in 85
 - configuring the switches in 74
 - planning worksheet example 73
 - restrictions 14
- Fabric Attached Metroclusters
 - limitations 16
- failover
 - cause-and-effect table 20
 - determining status (cf status) 129
 - determining status in a cluster 129
- FC-AL adapters, installing 178
- fencing, manual 167
- forcing a takeover 124
- forcing giveback 141

G

- GBIC, removing 159
- giveback
 - description of 117
 - executing 139
 - forcing 141
 - processes that prevent giveback 139

- testing 112

H

- halting system without takeover (halt) 123
- hardware
 - cluster components 4
 - installing an adapter 209
 - removing an adapter 208
 - removing F800 series clustering 158
 - removing FAS900 and FAS3000 series clustering 160
 - upgrading nondisruptively 207

I

- installing FC-AL adapters 178
- installing, adapters 209
- interface configurations
 - dedicated 144
 - shared 144
 - standby 144
- interface types and configurations 146

L

- licenses
 - disabling cluster 157
 - enabling cluster 103
 - required 103
- LRC modules, about 193
- lun commands, lun online 169
- LUNs, bringing online 169

M

- mailbox disks 5
- manual fencing 167
- MetroClusters
 - converting to a Fabric Attached 93
 - defined 163
 - Fabric Attached
 - See* Fabric Attached MetroClusters
 - non-switched
 - about 9
 - advantages of 10

- Data ONTAP and 11
 - illustration of 9
 - setup requirements for 11
 - reestablishing configuration after disaster 170
- mirrored clusters
 - about 7
 - adding a disk shelf to 179
 - advantages of 7
 - behavior of Data ONTAP with 11
 - cabling Channel A loop for FAS3000 series in 38
 - cabling Channel A loop for FAS900 series in 50
 - cabling Channel B loop for FAS3000 series in 42
 - cabling Channel B loop for FAS900 series in 54
 - FC-ALs in FAS900 series 45
 - restrictions 8, 11
 - setup requirements for 8
- modules
 - about 193
 - best practices for changing 193
 - hot-adding to existing loop 187
 - hot-swapping 203
 - restrictions for changing types 193
 - testing 194
 - upgrading 199
- mounting volumes 168

N

- negotiated failover *See* failover
- network interfaces
 - cluster configuration for takeover 146, 151
 - emulated LAN considerations in clusters 148
 - types and configurations in a cluster 146
- nondisruptive cluster upgrades, hardware 207
- non-switched MetroClusters *See* MetroClusters

O

- options
 - recommended to be identical between nodes 107
 - required to be identical between nodes 106

setting 105

P

- parameters
 - required to be identical between nodes 108
 - setting 105
- partner command 130
- partner name, displaying (cf partner) 120
- preparing equipment racks 34

R

- racks, installing in equipment 28
- reestablishing cluster configuration 170
- removing hardware 158, 160
- requirements
 - cable 30, 69
 - documentation 29, 68
 - equipment 30, 69
 - tool 30, 69
- restrictions
 - Fabric Attached MetroCluster 14
 - in mirrored clusters 8, 11
- rsh, using to access after takeover 117

S

- safety information xi
- setting options and parameters 105
- setup, running on clusters 97
- shared interfaces
 - configuring using ifconfig 149
 - configuring using setup 98
 - described 144
 - diagram 145
- single-point-of-failure
 - eliminating 18
- SNMP protocol, in clustered takeover mode 129
- special messages xvi
- standard cluster
 - adding a disk shelf to 180
 - behavior of Data ONTAP with 3
 - cabling Channel A loop for F800 series in 59
 - cabling Channel A loop for FAS3000 series in 36

- cabling Channel A loop for FAS900 series in 47
 - cabling Channel B loop for F800 series in 62
 - cabling Channel B loop for FAS3000 series in 40
 - cabling Channel B loop for FAS900 series in 52
 - contents of 3
- standby interfaces
 - configuring using ifconfig 151
 - configuring using setup 99
 - described 144
- status messages, descriptions of 119
- status, monitoring cluster 118
- switch fabric 12
- switches
 - banks 12
 - cabling the local node and disk shelves with 78
 - cabling the remote node and disk shelves with 85
 - checking firmware revision for 74
 - configuring for MetroCluster 74
 - disabling 75
 - pools 12
 - quadrants 12
 - setting the parameters 75
- system cabinets
 - installing in 28
 - preparing for cabling 32

T

- takeover
 - CIFS sessions and 117

- configuring with dedicated and hot standby interfaces 146
 - forcing for disaster recovery 168
 - mode with SNMP 129
 - rsh access after 117
 - statistics 129
 - telnet access after 117
 - testing 112
 - what happens after 117
 - what happens during 117
 - when occurs 116
 - with SNMP 129
- telnet, using to access after takeover 117
- tools, required 30, 69
- troubleshooting
 - giveback failure on F800 series 155
 - giveback failure on FAS3000 series 156
 - giveback failure on FAS900 series 156
 - takeover failure on F800 series 155
 - takeover failure on FAS3000 series 156
 - takeover failure on FAS900 series 156

U

- upgrading
 - disk shelf modules 199
 - hardware nondisruptively 207

V

- VI cluster adapter, removing 159
- VIFs, using to reduce SPOF 97
- volumes, mounting 168

