V-Series Systems
MetroCluster Guide

NetApp, Inc.
495 East Java Drive
Sunnyvale, CA 94089 U.S.A.
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 210-04515_A0
July 2009

Copyright and trademark information

Copyright information

Trademark information

Table of Contents

Preface

About this guide

This guide provides information for V-Series systems in MetroCluster environments for all versions of Data ONTAP that support MetroClusters.

Note

Data ONTAP software runs on multiple hardware platforms. This documentation might describe features that are not supported on your platform.

Audience

This guide is for system administrators who are familiar with operating systems such as UNIX® and Windows® and who will be installing V-Series systems. This guide does not discuss basic system or network administration topics, such as IP addressing, routing, and network topology; it emphasizes the characteristics of the V-Series system.

Relationship of this guide to other guides

This guide is intended to be used in conjunction with other information in the V-Series and Data ONTAP libraries. The following table describes the relationships between this guide and other documentation.

Guide name	Information includes...
Installation Requirements and Reference Guide	◆ General guidelines for creating and making array LUNs available to V-Series systems. ◆ Quick start installation instructions for connecting devices together and for installing Data ONTAP on a V-Series system that uses only third-party storage. ◆ Reference information. ◆ Detailed background information including layout in aggregates and checksums.
Implementation Guides	◆ Vendor-specific details about how to set up a storage array to work with V-Series systems. ◆ More detailed configuration examples than are provided in the *Installation Requirements and Reference Guide*.

Guide name	Information includes...
Implementation Guide for Native Disk Shelves	Information about setting up the storage on the native disk shelves connected to the V-Series system.
V-Series *Setup, Installation, and Management Guide* (Data ONTAP 7.3.x releases) or *Data ONTAP software setup guides* (for releases later than 7.3.x)	Detailed steps for setting up the V-Series system, including information about installing Data ONTAP software for installations using only third-party storage. These guides are most helpful to installers who are new to Data ONTAP setup and installation.
Data ONTAP guides	Detailed information about all Data ONTAP features used by all systems running Data ONTAP, for example, storage features and data protection features.

See the V-Series *Support Matrix* for details about Data ONTAP releases that support V-Series, supported switches, supported firmware, capacity, and maximum array LUN count.

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

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

Understanding MetroClusters 1

About this chapter

This chapter defines a MetroCluster and describes the two kinds of MetroCluster configurations: stretch and fabric-attached.

Topics in this chapter

This chapter contains the following topics:

Generic storage allocation terms used in this document

array LUN: This guide uses the term *array LUN* (logical unit number) to describe an area on the storage array that is available for a V-Series system or a non V-Series host to read data from or write data to. You might be accustomed to hearing a different term to describe this area; the term varies among vendors and sometimes among platforms for the same vendor. See the V-Series *Implementation Guide* for your storage array type for the specific term used for your platforms.

Note

The V-Series system can virtualize the storage provided by third-party storage arrays and serve it up as LUNs to applications and customers outside the V-Series system (for example, through iSCSI). Clients are unaware of where such a LUN is stored. Data ONTAP software product guides refer to this type of storage as a LUN. You can think about this type of storage as a front-end LUN or Data ONTAP-served LUN.

HA pair: Two storage systems (nodes) whose controllers are connected to each other either directly or through switches. In some versions of Data ONTAP, this configuration is referred to as an *active/active configuration*.

disks, disk shelves: The terms *disk* and *disk shelf* in this document refer to native storage connected to the V-Series system. These terms do not refer to disks or disk shelves on a third-party storage array.

native disks, native disk shelves: Disks and disk shelves that are sold as local storage for systems that run Data ONTAP software.

What a MetroCluster is

Definition of MetroCluster

A V-Series system MetroCluster configuration is a hardware and software solution that provides storage redundancy and disaster recovery over long distances. Like mirrored HA pairs, MetroCluster configurations contain two complete copies of the specified data volumes or file systems. These copies, called *plexes*, are continuously and synchronously updated every time Data ONTAP writes data to the storage arrays. In a MetroCluster configuration, each plex must be on a separate set of LUNs and each LUN set must be on a different storage array. One plex is the local plex and the second (mirrored) plex is on the remote storage array.

There are two types of MetroCluster configurations:

- Stretch MetroCluster
- Fabric-attached MetroCluster

What a stretch MetroCluster is

A stretch MetroCluster configuration is a direct connection of two or more mirrored HA pairs over limited distances (campus-sized areas.) This configuration provides storage redundancy and disaster recovery across distances up to 500 meters. The V-Series systems connect to the storage arrays over redundant paths in a switched fabric configuration.

What a fabric-attached MetroCluster is

A fabric-attached MetroCluster is a switch connected configuration of two or more mirrored HA pairs that provides storage redundancy and disaster recovery across long distances. A fabric-attached MetroCluster configuration is the same as a stretch MetroCluster configuration except for the following features:

- It uses the switched fabric to connect the local V-Series system with the remote V-Series system.
- It provides storage redundancy and disaster recovery across longer distances than a stretch MetroCluster. For V-Series systems, the distance is up to 30 km. (FAS systems support a longer distance.)

Attention

For V-Series systems, fabric-attached MetroCluster configurations that exceed 30 km require a Product Variation Request (PVR). See the V-Series *Support Matrix* on the NOW™ Web site (http://now.netapp.com) for information about submitting a PVR.

Data protection features unique to a MetroCluster

A MetroCluster provides the same single point of failure protections that are available on mirrored HA pairs, plus the following additional data protection that is not available with mirrored HA pairs:

- Redundancy over longer distances than mirrored HA pairs.
- Automatic access to data on local array LUNs and volumes by clients if a V-Series system interconnect or switch interconnect fails.

  To ensure data integrity, mirrors are disabled and data is automatically resynchronized after the interconnect is reestablished.
- Ability to manually force a failover when an entire node is destroyed or unavailable (for example, in the case of a fire or long-term power loss).

  If a MetroCluster site disaster occurs and all connections between the sites are lost, you can enter a manual command to enable the surviving node to take over the function of its partner. Mirrors are disabled and data is accessed on the surviving partner. To reestablish the MetroCluster, you must fix the failure. For example, restore power or, if the partner is destroyed, create a new MetroCluster partner at a different location.

Dual-controller systems with a MetroCluster

Dual-controller (31xx) systems are available in single or dual-controller platforms. If the systems are populated with a single controller, the single controller on each system can support a single MetroCluster.

Stretch MetroCluster configuration on 31xx systems

A stretch MetroCluster can be configured between a pair of 31xx systems in which each system has a single controller (rather than two).

To implement the stretch MetroCluster, an FC-VI adapter must be installed in each controller to provide the cluster interconnect between the systems. When the FC-VI adapter is installed in the system, the internal InfiniBand interconnect is automatically disabled. This is different than other stretch MetroClusters, which use NVRAM adapters to provide the interconnect.

Fabric-attached MetroCluster configuration on 31xx systems

A fabric-attached MetroCluster can be configured between a pair of 31xx systems in which each system has a single controller (rather than two).

When the system detects the presence of an FC-VI adapter, which connects the controller to the switch fabric, the internal InfiniBand connection is automatically deactivated.

Requirements for Installing a MetroCluster 2

About this chapter

This chapter provides requirements for installing a MetroCluster configuration with V-Series systems.

Note

See the V-Series *Support Matrix* for information about MetroCluster support in different Data ONTAP releases.

Topics in this chapter

This chapter contains the following topics:

Hardware requirements for MetroCluster

V-Series system requirements

The requirements for V-Series systems in a MetroCluster are as follows:

- The V-Series systems must be the same platform type at both sites.
- The platform type must support a MetroCluster configurations.
- You cannot exceed the maximum storage capacity for each V-Series system.

  See the V-Series *Support Matrix* for the most current information about storage capacity for each platform type.
- V-Series systems on which native disk shelves are installed cannot be deployed in a MetroCluster configuration.

Note

MetroCluster configurations that do not conform to the requirements on the V-Series *Support Matrix* might be qualified by NetApp professional services. Contact your NetApp sales representative for additional information.

Stretch MetroCluster interconnect requirements

For stretch MetroCluster configurations (covering distances up to 500 meters), the interconnect requirements are as follows:

- Two NVRAM5 or NVRAM6 cluster interconnect cards—provided with each V-Series system
- Four copper-to-fiber adapters—two on each NVRAM, required for cluster remote and SyncMirror operations
- Two 4xIB MPO optical cluster interconnect cables—one on each NVRAM card
- VI-MC card for dual-controller configurations—see the *System Configuration Guide* for information about supported cards.
- Four long distance Small Form Plugable adapters (SFPs)—one per switch

Fabric-attached MetroCluster interconnect requirements

For fabric-attached MetroCluster configurations (covering distances of up to 30 kilometers for V-Series systems), the interconnect requirements are as follows:

- Two VI-MC cards—one each on V-Series system
- Four long distance Small Form Plugable adapters (SFPs)—one per switch

See the *System Configuration Guide* for information about supported VI-MC cards.

Switch requirements

The requirements for switches used in a MetroCluster are as follows:

- You can use mixed switch types, but the switches must be the same for each fabric.
- Switches must use firmware that is supported for V-Series system MetroCluster configurations.
- To ensure redundancy, connect each V-Series system to storage using redundant components in a switched fabric configuration.

Inter-Switch Link (ISL) requirements

Only one ISL is supported per fabric. See the V-Series *Support Matrix* for details about ISL configuration.

Storage array requirements

The requirements for storage arrays used in a MetroCluster are as follows:

- The storage arrays must be symmetric. This means that the array LUNs on both sides of the mirror must be the same size and from the same family of storage arrays. Disk types (FC or SATA) must also be the same.

  For example, if you have a Hitachi 9585 with FC drives at one site, you can have a Hitachi 9570 with FC drives at the other side because they are in the same family.

  See the V-Series *Implementation Guide* for your storage array for more information about which arrays are in the same family.
- Storage arrays that provide tiered configurations (Hitachi) must use the same tiered topology on each side of the MetroCluster.
- You must mirror the root volume to enable successful takeover.

SyncMirror software requirements for MetroCluster

Requirement for SyncMirror

V-Series systems in MetroCluster configurations require SyncMirror® software. MetroCluster configurations use SyncMirror to build a system that can continue to serve data after the complete loss of one of the nodes and the storage at that site. SyncMirror creates two physically separated copies of an aggregate. The two copies, known as *plexes*, are simultaneously updated. Therefore, the copies are always identical and data consistency is retained.

You can have both mirrored and unmirrored volumes in a MetroCluster configuration. However, the MetroCluster configuration can preserve data only if volumes are mirrored. Unmirrored volumes are lost if the storage on which they reside is destroyed.

Note

You must connect unmirrored storage to both nodes, just as for mirrored storage. You cannot have storage that is connected to only one node in a MetroCluster configuration.

The configuration requirements for Data ONTAP mirrored volumes are as follows:

- The root volume must be mirrored with one plex at each site.
- Data volumes should be mirrored with one plex at each site.

See the details about SyncMirror in the Data ONTAP *Data Protection and Online Backup Guide*.

License requirements for SyncMirror

The following software licenses must be enabled on the V-Series system at the local and remote sites:

- cluster: Provides automatic hardware failover capability between sites.
- syncmirror_local: Provides an up-to-date copy of data at the remote and local sites. Data is ready for access after failover without intervention from the administrator.
- cluster_remote: Provides features that enable the administrator to declare site disaster and initiate a site failover via a single command.

Note

MetroCluster configurations might require an additional switch license to support ISLs. Refer to the switch documentation for complete information about switch and license requirements.

Requirements for SyncMirror use with third-party storage

To set up SyncMirror with third-party storage, you must fulfill standard requirements for any SyncMirror deployment plus a number of requirements that are unique to setting up SyncMirror with third-party storage. For information about the requirements for using SyncMirror for mirroring third-party storage, see the Data ONTAP *Data Protection and Online Backup Guide*.

Importance of planning for your SyncMirror deployment

You should plan for your SyncMirror deployment before you assign array LUNs to your V-Series systems. You need to make sure that you size the array LUNs for your mirrored aggregates correctly and, when you assign an array LUN to a V-Series system, that you correctly identify whether the LUN is in local storage or remote storage. You do this by specifying the local pool (pool 0) or remote pool (pool1), as appropriate, when you assign an array LUN to a V-Series system.

See the Data ONTAP *Data Protection and Online Backup Guide* for examples of how to assign array LUNs to the appropriate pool.

Attention

If the SyncMirror license is not installed when you assign an array LUN to your V-Series system, Data ONTAP specifies the local pool (pool0) for the LUN. After you install the SyncMirror license, you would have to unassign the LUNs that are in the remote location that you are using for mirroring. You must then reassign each of those LUNs to the V-Series system, this time specifying a pool parameter of pool1. If the LUNs in the remote location were added to an aggregate before the SyncMirror license was installed, you would have to destroy the aggregates that the LUNs were added to, unassign the LUNs, reassign them to the system in pool1, and recreate the aggregates.

Fabric-attached MetroCluster FC-VI traffic in-order delivery requirement

Requirement for in-order delivery of FC-VI traffic

Fabric-attached configurations require in-order delivery of FC-VI (Fibre Channel Virtual Interface) traffic. When there are multiple ISLs in a fabric-attached MetroCluster, and in some circumstances when there is a single ISL, you need to configure your switches to ensure in-order-delivery of FC-VI traffic.

Requirement for configurations with only one ISL

Some switches (such as the Brocade 48000) have multiple paths through the switch itself, which could result in out-of-order delivery even when there is a single ISL per fabric. If your configuration includes a switch with multiple paths through the switch, be sure to configure the switch and fabric to ensure in-order-delivery.

Requirements for configurations with more than one ISL

In fabric-attached MetroCluster configurations with more than one ISL within a fabric, you must configure the fabric that transports the FC-VI traffic to guarantee in-order-delivery of all frames and sequences. This means that you must configure each switch to guarantee in-order-delivery during each of the following operational states:

- Normal operations, during which there is more than one path between a source port and a destination port
- Normal operations, during which a switch could probe the FC-VI port
- Fabric reconfiguration events

Requirement for switches that can probe the FC-VI port

Some switches in a fabric-attached MetroCluster configuration are configured to probe devices that do not register with the fabric name server. In these fabrics, the N-port discovery process can cause the FC-VI port to enter a “failed” state. To prevent this loss of access, you must disable device probing on these fabrics.

Zoning requirements for FC-VI ports

You need to zone the FC-VI ports on the FC adapters end-to-end across the fabric. You must zone the FC-VI ports as follows:

- Zone FC-VI port “a” on the V-Series system on one fabric to the FC-VI port “a” on the V-Series system on the other fabric

- Zone FC-VI port “b” on the V-Series system on one fabric to the FC-VI port “b” on the V-Series system on the other fabric

Ensure that you test your zoning configuration of the FC-V1 ports to ensure that you do not include “a” and “b” ports in the same zone. See “Testing zoning of FC-VI ports” on page 49.

When you set up the zones for the FC-VI ports you need to provide the WWPN for the FC-VI ports. You can obtain the WWPN for the FC-VI ports from the following sources:

- The switch
- `sysconfig -M` output

  Look for entries such as the following:

```
!Qlogic 2352 FCVI Cluster Interconnect Adapter<adapter_WWPN>
```

Where to find information about switch configuration

For information about how to configure your switch for in-order delivery of FC-VI traffic, see the documentation for your McDATA, Brocade, or Cisco switch.

Recommended MetroCluster Configurations 3

About this chapter

This chapter shows the best practice recommendations for stretch and fabric-attached MetroCluster configurations with V-Series systems.

Attention

Fabric-attached MetroCluster configurations that do not conform to the best practice configuration require a PVR.

Topics in this chapter

This chapter contains the following topics:

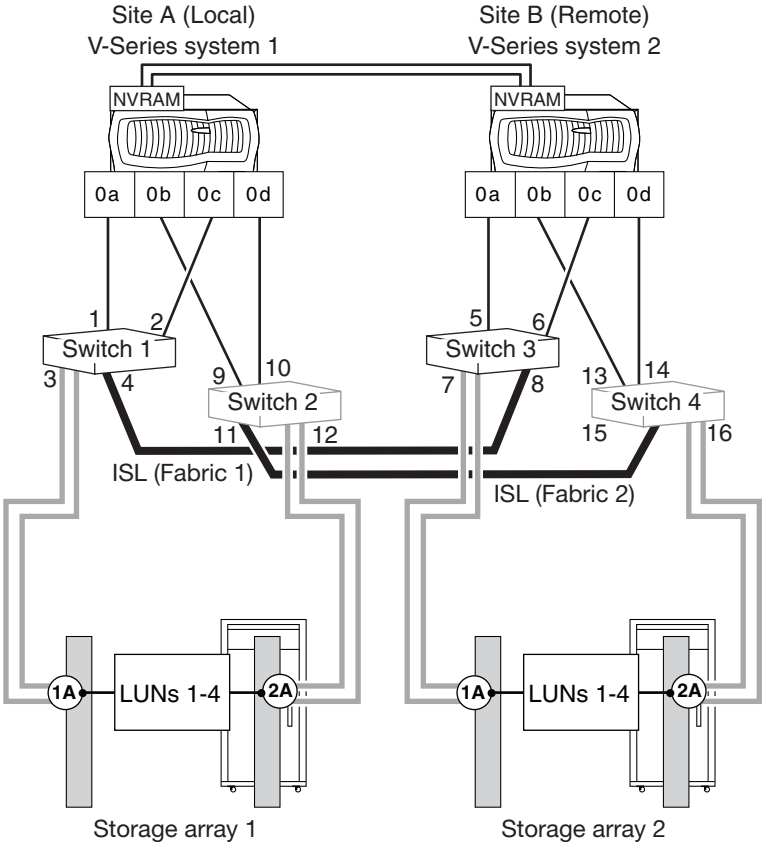
Recommended stretch MetroCluster configuration

When a PVR is required for a stretch MetroCluster

MetroCluster configurations that do not conform to this best practice configuration for a stretch MetroCluster require a PVR. See the V-Series *Support Matrix* for details about configuration requirements and supported systems.

Best practice stretch MetroCluster configuration

The following illustration shows the components and best practice configuration of a stretch MetroCluster configuration.



Connections

This section describes the connections for the stretch MetroCluster configuration in the previous illustration.

V-Series system interconnect: V-Series systems are connected directly.

The following shows NVRAM connections.

V-Series System 1	V-Series System 2
NVRAM port L02 Ph2	NVRAM port L02 Ph2
NVRAM port L01 Ph1	NVRAM port L01 Ph1

The following shows connections for dual-controller systems.

V-Series System 1	V-Series System 2
VI -MC port A	VI-MC port A
VI-MC port B	VI-MC port B

Inter-Switch Link connections (ISL): Stretch MetroCluster configurations use a switched fabric to connect the local half of the configuration to the remote half of the configuration. In the previous illustration, Switches 1 and 3 are connected to each other (Fabric 1). Switches 2 and 4 are also connected (Fabric 2).

- The first fabric in the MetroCluster begins from Switch 1 on Site A (local) and is completed by connecting the ISL cable to the first switch on Site B (remote)—Switch 3.
- The second fabric is created using Switch 2 on Site A (local), connected through a second ISL cable, to the second switch on Site B (remote) — Switch 4

The following table describes the ISLs in this configuration.

ISL Connection	Switch	Fabric
Port 4 on switch 1	Port 8 on switch 3	Fabric 1
Port 11 on switch 2	Port 15 on switch 4	Fabric 2

V-Series system-switch connections: The best practice connections in the previous illustration eliminate a single point of failure in the following ways:

- FC initiator ports on the same Fibre Channel controller chip (for example, port 0a and 0b) connect to alternate fabrics.

  Note

  See the V-Series *Installation Requirements and Reference Guide* for specific information about port pairs.

- Multiple paths and zones ensure that FC initiator ports on the same controller chips access the array LUN from different V-Series systems and switch fabrics.

The following table describes the connections from the V-Series system to the switch.

V-Series System Port	Switch
V-Series system 1: FC port 0a	Switch 1: Port 1
V-Series system 1: FC port 0b	Switch 2: Port 9
V-Series system 1: FC port 0c	Switch 1: Port 2
V-Series system 1: FC port 0d	Switch 2: Port 10
V-Series system 2: FC port 0a	Switch 3: Port 5
V-Series system 2: FC port 0b	Switch 4: Port 13
V-Series system 2: FC port 0c	Switch 3: Port 6
V-Series system 2: FC port 0d	Switch 4: Port 14

Storage array-switch connections: The following describes the best practice connections from the storage array to the switch.

- Ports 1A and 2A on each storage array connect to alternate fabrics.
- V-Series systems are configured to access any array LUN on two storage array paths (1A and 2A).

Storage Array Port	Switch	Fabric
Array 1: Port 1A	Switch 1: Port 3	1
Array 1: Port 2A	Switch 2: Port 12	2

Storage Array Port	Switch	Fabric
Array 2: Port 1A	Switch 3: Port 7	1
Array 2: Port 2A	Switch 4: Port 16	2

Recommended fabric-attached MetroCluster configuration

When a PVR is required for a fabric-attached MetroCluster

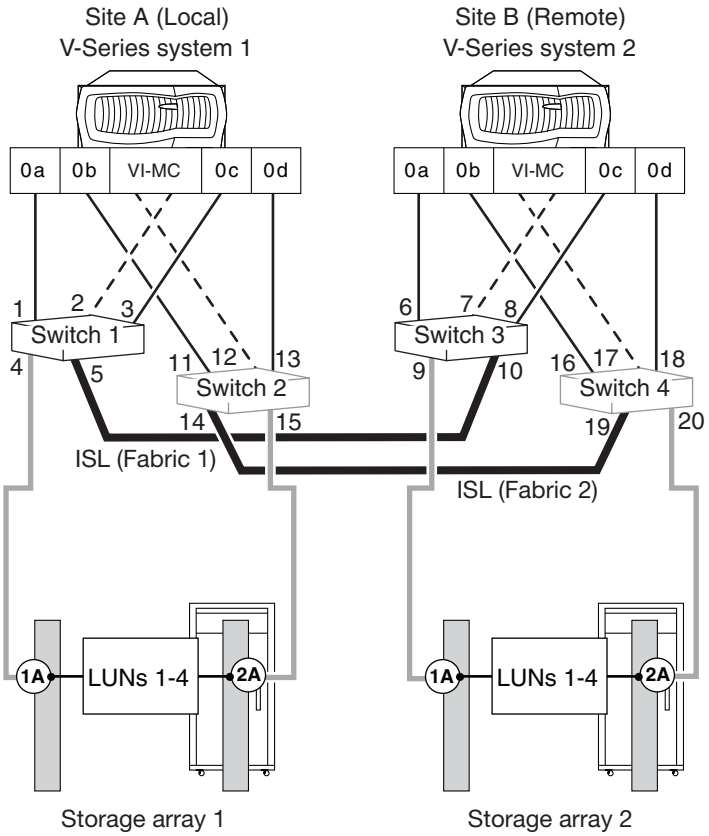
Fabric-attached MetroCluster configurations that do not conform to the best practice configuration require a PVR. The following restrictions apply:

- A fabric-attached MetroCluster implemented on a SAN fabric that uses more than one ISL between any pair of switches in the data path for the FC-VI adapter requires a PVR.
- Configurations that span distances greater than 30 km require a PVR.

See the V-Series *Support Matrix* for details about configuration requirements, ISL requirements, and supported systems.

Best practice fabric-attached MetroCluster configuration

The following illustration shows the components and best practice configuration of a fabric-attached MetroCluster configuration. The fabric-attached V-Series system MetroCluster shown provides the same single point of failure protections that are available for all mirrored HA pairs.



Connections

This section describes the connections for the fabric-attached MetroCluster configuration in the previous illustration.

V-Series system interconnect: V-Series systems are interconnected by connecting ports A and B on the VI-MC card to alternate switch fabrics.

V-Series system	Switch	Fabric
V-Series system 1: Port A	Switch 2, Port 12	2
V-Series system 1: Port B	Switch 1, Port 2	1

V-Series system	Switch	Fabric
V-Series system 2: Port A	Switch 4, port 17	2
V-Series system 2: Port B	Switch 3, port 7	1

Inter-Switch Link connections (ISL): Fabric-attached MetroCluster configurations use a switched fabric to connect the local half of the configuration to the remote half of the configuration. In the previous illustration, Switches 1 and 3 are connected to each other (Fabric 1). Switches 2 and 4 are also connected (Fabric 2).

- The first fabric in the MetroCluster begins from Switch 1 on Site A (local) and is completed by connecting the ISL cable to the first switch on Site B (remote)—Switch 3.
- The second fabric is created using Switch 2 on Site A (local), connected through a second ISL cable, to the second switch on Site B (remote)—Switch 4

The following table describes the ISLs in this configuration.

ISL connection	Switch	Fabric
Port 5 on switch 1	Port 10 on switch 3	Fabric 1
Port 14 on switch 2	Port 19 on switch 4	Fabric 2

V-Series system-switch connections: The best practice connections in the previous illustration eliminate a single point of failure in the following ways:

- FC initiator ports on the same Fibre Channel controller chip (for example, port 0a and 0b) connect to alternate fabrics.

  Note

  See the V-Series *Installation Requirements and Reference Guide* for specific information about port pairs.

- Multiple paths and zones ensure that FC initiator ports on the same controller chips access the array LUN from different V-Series systems and switch fabrics.

The following table describes the connections from the V-Series system to the switch.

V-Series system port	Switch
V-Series system 1: FC port 0a	Switch 1: Port 1
V-Series system 1: FC port 0b	Switch 2: Port 11
V-Series system 1: FC port 0c	Switch 1: Port 3
V-Series system 1: FC port 0d	Switch 2: Port 13
V-Series system 2: FC port 0a	Switch 3: Port 6
V-Series system 2: FC port 0b	Switch 4: Port 16
V-Series system 2: FC port 0c	Switch 3: Port 8
V-Series system 2: FC port 0d	Switch 4: Port 18

Storage array-switch connections: The following describes the best practice connections from the storage array to the switch.

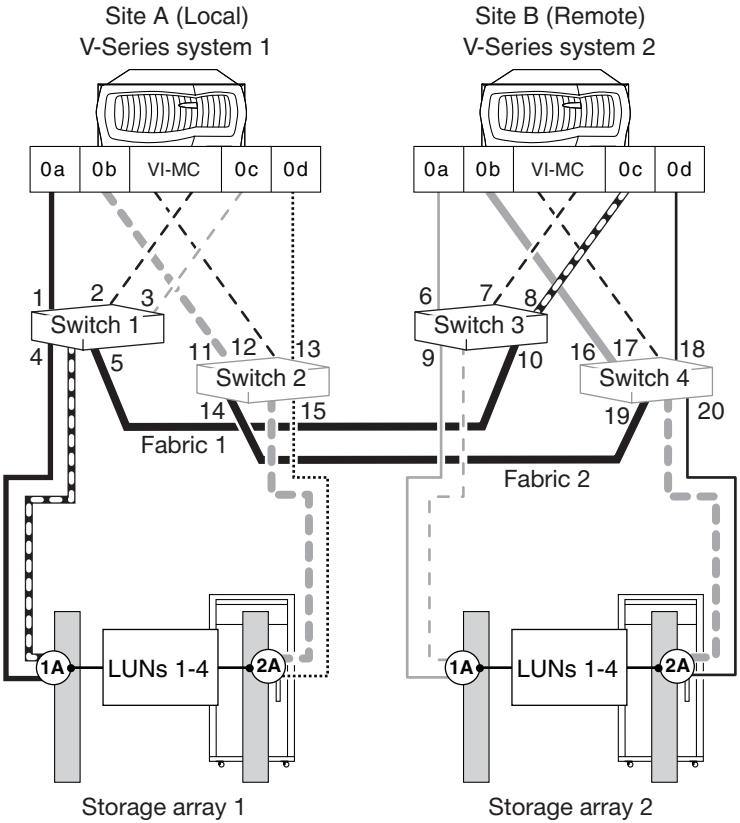
- Ports 1A and 2A on each storage array connect to alternate fabrics.
- V-Series systems are configured to access any array LUN on two storage array paths (1A and 2A).

Storage Array Port	Switch	Fabric
Array 1: Port 1A	Switch 1: Port 4	1
Array 1: Port 2A	Switch 2: Port 15	2
Array 2: Port 1A	Switch 3: Port 9	1
Array 2: Port 2A	Switch 4: Port 20	2

MetroCluster zoning example

Sample zoning

The following example shows single initiator zoning for a MetroCluster configuration.



In the example, four array LUNs are allocated on each storage array for the MetroCluster. LUNs of equal size are provisioned on the storage arrays at both sites, which is a SyncMirror requirement. Each FC initiator port on each V-Series system has a path to every V-Series LUN on the storage arrays. The ports on the storage array are redundant, and are configured as follows:

- On Storage Array 1, Ports 1A and 2A are a port pair. Because they are alternate paths, both ports can access LUNs 1 through 4.

- On Storage Array 2, Ports 1A and 2A are a port pair. Because they are alternate paths, both ports can access LUNs 1 through 4.

Switches are zoned so that there are only two paths to each array LUN, one unique path from each V-Series FC initiator port through each switch. If there are multiple connections between a V-Series system and the switch, the best practice recommendation is to put each connection into a separate zone.

Zones: The following table describes the V-Series systems zones.

Zone	Connection
Zone 1	V-Series system 1: Port 0a Storage array 1: Port 1A
Zone 2	V-Series system 2: Port 0a Storage array 2: Port 1A
Zone 3	V-Series system 1: Port 0b Storage array 2: Port 2A
Zone 4	V-Series system 2: Port 0b Storage array 1: Port 2 A
Zone 5	V-Series system 1: Port 0c Storage array 2: Port 1A
Zone 6	V-Series system 2: Port 0c Storage array 1: Port 1A
Zone 7	V-Series system 1: Port 0d Storage array 1: Port 2A
Zone 8	V-Series system 2: Port 0d Storage array 2: Port 2A

Installing a MetroCluster 4

About this chapter

This chapter contains instructions for cabling a stretch or fabric-attached MetroCluster. See Chapter 5, "Data ONTAP Configuration and Verification," on page 47 for an overview of what you need to configure in Data ONTAP to complete set up of the V-Series systems in your deployment and verify that your MetroCluster is working.

Attention

MetroCluster configurations that do not conform to the best practice recommendations in Chapter 3, "Recommended MetroCluster Configurations," on page 13 require a PVR. For details about configuration requirements and supported systems, see the V-Series *Support Matrix*.

Topics in this document

This document contains the following topics:

MetroCluster installation overview

The following table provides an overview of the tasks for planning and connecting your MetroCluster installation.

Stage	Task	For information, see...
1	Ensure that your MetroCluster environment meets the minimum V-Series requirements, for example, supported MetroCluster V-Series system models, supported switches, supported switch firmware, appropriate MetroCluster licenses.	◆ Chapter 2, "Requirements for Installing a MetroCluster," on page 5 ◆ The Data ONTAP *Data Protection and Online Backup Guide* ◆ The V-Series *Support Matrix*
2	Plan the size of, number of, location of, and pool assignment for LUNs on the storage array. Plan for LUN security,	◆ V-Series *Installation Requirements and Reference Guide* ◆ The Data ONTAP *Data Protection and Online Backup Guide*
3	Plan the port-to-port connectivity scheme between the V-Series system, the switch, and the storage array. **Note** The number of LUNs supported on a storage array port varies, depending on the V-Series system model and storage array that you deploy.	◆ Chapter 4, "Installing a MetroCluster," on page 25 ◆ V-Series *Installation Requirements and Reference Guide* (for information about V-Series and storage array port usage)
4	Create the LUNs on the storage arrays.	◆ V-Series *Installation Requirements and Reference Guide* ◆ V-Series *Implementation Guide* for your storage array ◆ Your storage array documentation
5	Configure LUN security	◆ V-Series *Installation Requirements and Reference Guide* ◆ Your storage array documentation

Stage	Task	For information, see...
6	Install NVRAM or VI-MC adapters on the V-Series systems. **Note** MetroClusters that are configured on new V-Series systems are shipped with the correct NVRAM or VI-MC card. If you are configuring a MetroCluster on existing V-Series systems, you might need to install a new card or move an existing card to another slot.	◆ The *System Configuration Guide* (for information about supported VI-MC cards).
7	Cable the local V-Series system to the fabric.	"Connecting the local V-Series systems" on page 30.
8	Cable the remote V-Series system to the fabric.	"Connecting the remote V-Series systems" on page 35.
9	Connect the switched fabric.	"Connecting the switch fabric" on page 39.
10	Cable the storage array to the fabric.	"Connecting the fabric and storage array" on page 42.
11	Configure zoning.	"Configuring zoning" on page 45.

Best practice cabling guidelines

Guidelines for partner-to-partner connections

Guidelines for partner-to-partner connections are as follows:

- The cabling for MetroCluster partners is different on each node. You should distinguish these differences by using different names.

  This guide refers to the local partner as V-Series system 1, Site A, and the remote partner as V-Series system 2, Site B. Before cabling your MetroCluster, record which V-Series system is Site A and which is Site B.
- Each port on the V-Series system should be connected to the same fabric.

  For example, if Port A of VI-MC card on the local node is connected to Switch 2 and Port A of the VI-MC card on the remote node is connected to Switch 4, then Switch 2 and Switch 4 must be connected by the ISL, thereby connecting the nodes to the same fabric. See "Connecting the switch fabric" on page 39 and "Recommended fabric-attached MetroCluster configuration" on page 18 for examples of these connections.

Guidelines for V-Series system-switch connections

Guidelines for connecting a V-Series system to a switch are as follows:

- Fibre Channel ports on the same channel controller chip should never be connected to the same switch. You must connect one port to one switch and the other port to the other switch. For example, if onboard port 0a is connected to Switch 3, you should not connect onboard port 0b to Switch 3; you must connect port 0b to Switch 4.
- If you are using a redundant FC initiator port pair, connecting both ports of the pair to the same switch port number can simplify cabling and management of the MetroCluster configuration. For example, if port 0a is connected to Port 1, Switch 1, you should connect port 0b to Port 1, Switch 2.
- All switches within a fabric must be the same switch model and have the same number of licensed ports.

Guidelines for Inter-Switch connections

You can connect the Inter-Switch Link (ISL) to any available switch port.

Guidelines for V-Series system-storage array connections

Guidelines for connecting a V-Series system to the storage array are as follows:

- To eliminate configurations that contain a single point of failure, the ports on the storage array that you select to access a given LUN must connect to different V-Series systems, switches, and FC ports.
- Data ONTAP requires access to a specific array LUN on only two of the storage array ports. Check the port configuration of each storage array in the MetroCluster configuration to ensure conformance with this requirement.

Connecting the local V-Series systems

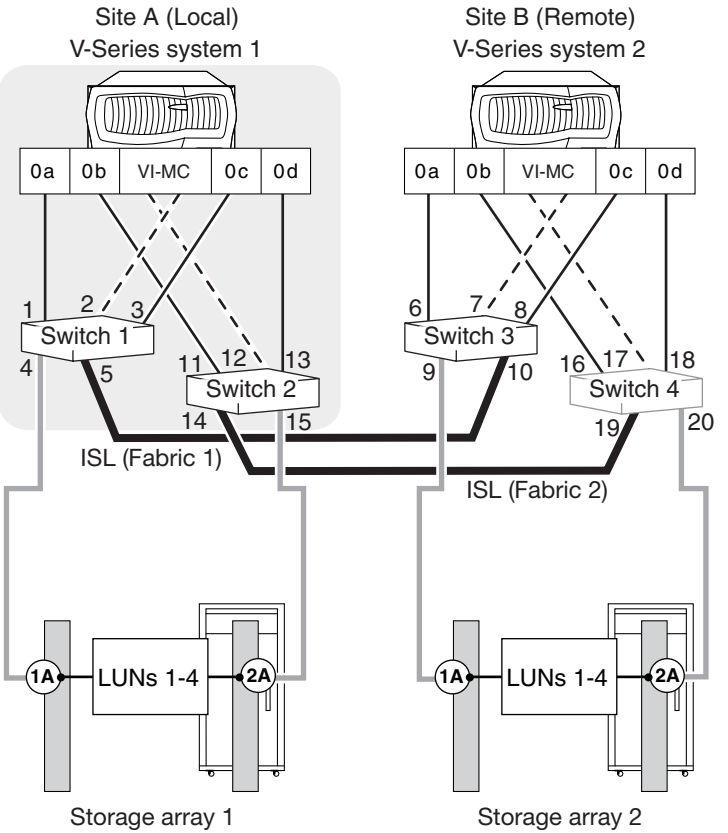
Before you begin

Before you begin this procedure, make sure you have done the following:

- Completed your MetroCluster configuration plan
- Configured LUNs on the storage array

Example local site

The following example shows the local V-Series system connections for this procedure.



Connecting the local V-Series system

Complete the following steps to connect a local V-Series system to the fabric.

Step	Action	Diagrams and references
1	Connect the local V-Series system to the remote V-Series system. Locate the NVRAM or VI-MC interconnect module and take one of the following actions ◆ For fabric-attached configurations (VI-MC adapter), do the following: ❖ Cable VI-MC Port A to one fabric (Port A to Switch 2 in the previous example). ❖ Cable VI-MC Port B to the alternate fabric (Port B to Switch 1 in the previous example). ◆ For stretch configurations with an NVRAM adapter, do the following: ❖ Install a copper-to-fiber converter in L01 Ph1, and cable L01 Ph1 on the local node to L01 Ph1 on the remote node. ❖ Install a copper-to-fiber converter in L02 Ph2, and cable L02 Ph2 on the local node to L02 Ph2 on the remote node. ◆ For dual-controller configurations with VI-MC cards, do the following: ❖ Connect the VI-MC Port A on system 1 to the VI-MC Port A on system 2. ❖ Connect the VI-MC Port B on system 1 to the VI-MC Port B on system 2.	◆ For a fabric-attached MetroCluster diagram, see “Example local site” on page 30. ◆ For a stretch MetroCluster diagram, see “Recommended stretch MetroCluster configuration” on page 14.

Step	Action	Diagrams and references
2	Connect the V-Series system FC initiator ports to the switched fabric, as follows: **a.** Identify one of the two FC initiator port pairs on the V-Series system. Cable one port of the pair to one fabric. Cable the other port of the pair to the alternate fabric. (In the previous example, these are ports 0a and 0b.) **Note** FC initiator port pairs are Fibre Channel initiator ports on the same Fibre Channel controller chip. See the V-Series *Installation Requirements and Reference Guide* for specific information about port pairs. **b.** Identify the other onboard FC initiator port pair on the V-Series system. Cable one port of the pair to one fabric. Cable another port of the pair to the alternate fabric. (In the previous example, these are ports 0c and od).	◆ For FC initiator port pairs, see the V-Series system *Installation Requirements and Reference Guide.* ◆ For a diagram of the local MetroCluster site, see “Example local site” on page 30.
3	(Optional) Connect the V-Series system to a tape backup device through a separate FC initiator port or SCSI tape adapter.	
4	Connect a console cable to the console port on the V-Series system. Use the RJ-45 to DB-9 adapter that is included with your system. Connect the console cable to the adapter.	
5	Install the cable management tray. Pinch the arms of the tray and fit the holes in the arms through the motherboard tray pins. Then push the cables into the cable holders, thread the adapter cables through the top rows of the cable holders, and thread the port cables through the lower cable holders.	See http://now.netapp.com/NOW/knowledge/docs/hardware/hardware_index.shtml.

Step	Action	Diagrams and references
6	Connect the V-Series system to the Ethernet network. Plug the network cable into the networking port. If you are connecting more than one network cable to the network, connect to the ports sequentially. Use the cable management tray to direct all the cabling from your system.	
7	(Optional) Connect the remote LAN module (RLM) from the back of the V-Series system to the network using an Ethernet cable. **Attention** The network switch port for the RLM connection must negotiate down to 10/100 or autonegotiate.	
8	If applicable, turn on any tape backup devices.	
9	For each power supply on the V-Series system, take the following steps: **a.** Ensure that the power switch is in the Off (0) position. **b.** Connect the socket end of the power cord to the power plug on the power supply. **c.** Secure the power cord with the retaining adjustable clip on the power supply. **d.** Plug the other end of the power cord into a grounded electrical outlet. **Attention** To obtain power supply redundancy, you must connect the second power supply to a separate AC circuit.	

Step	Action	Diagrams and references
10	Start a communications program. You must use some form of communications program to perform initial network setup and V-Series configuration. You can start a communications program through RLM or through the console after connecting to the serial port.	

Connecting the remote V-Series systems

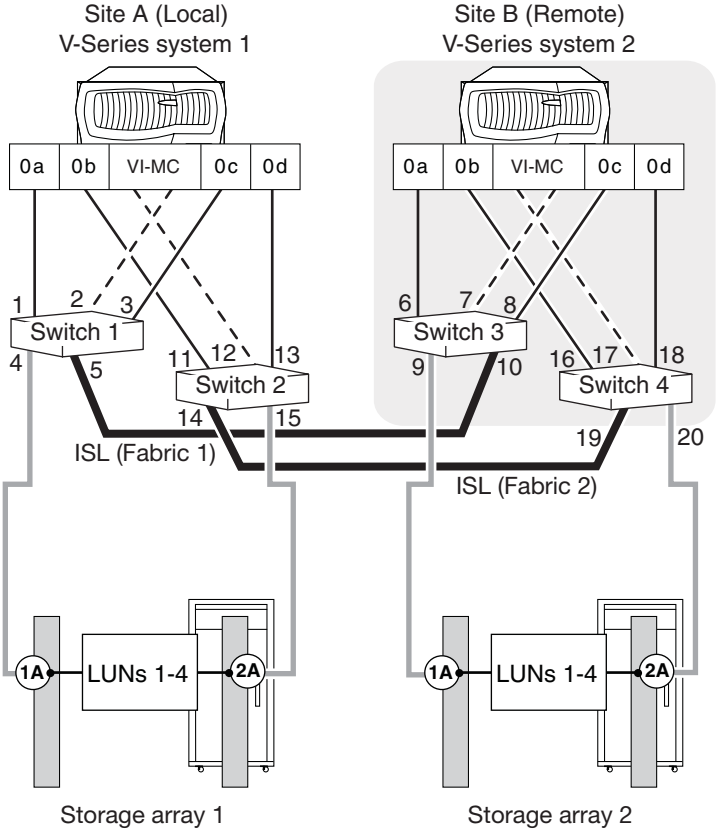
Before you begin

Before you begin this procedure you should have done the following:

- Completed your MetroCluster configuration plan
- Configured LUNs on the storage array
- Connected the local V-Series system to the fabric

Sample remote site

The following illustration shows the remote V-Series system connections for this procedure.



Connecting the remote V-Series system

Complete the following steps to connect a remote V-Series system to the fabric.

Step	Action	Diagrams and references
1	Connect the local V-Series system to the remote V-Series system. ◆ For fabric-attached configurations (VI-MC adapter), do the following: ❖ Cable VI-MC Port A to one fabric (Port A to Switch 4 in the previous example). Cable VI-MC Port B to the alternate fabric (Port B to Switch 3 in the previous example).	◆ For a fabric-attached MetroCluster diagram, see "Example local site" on page 30. ◆ For a stretch MetroCluster diagram, see "Recommended stretch MetroCluster configuration" on page 14.
2	Connect the V-Series system FC initiator ports to the switched fabric, as follows: **a.** Identify one of the two FC initiator port pairs on the V-Series system. Cable one port of the pair to one fabric. Cable another port of the pair to the alternate (In the previous example, these are ports 0a and 0b.) **Note** FC initiator port pairs are Fibre Channel initiator ports on the same Fibre Channel controller chip. See the V-Series *Installation Requirements and Reference Guide* for specific information about port pairs. **b.** Identify the other onboard FC initiator port pair on the V-Series system. Cable one port of the pair to one fabric. Cable the other port of the pair to the alternative fabric. (In the previous example, these are ports 0c and 0d.)	◆ For a diagram of the local MetroCluster site, see "Example local site" on page 30.
3	(Optional) Connect the V-Series system to a tape backup device through a separate FC initiator port or SCSI tape adapter.	

Step	Action	Diagrams and references
4	Connect a console cable to the console port on the V-Series system. Use the RJ-45 to DB-9 adapter that is included with your system. Connect the console cable to the adapter.	
5	Install the cable management tray. Pinch the arms of the tray and fit the holes in the arms through the motherboard tray pins. Then push the cables into the cable holders, thread the adapter cables through the top rows of the cable holders, and thread the port cables through the lower cable holders.	See http://now.netapp.com/NOW/knowledge/docs/hardware/hardware_index.shtml.
6	Connect the V-Series system to the Ethernet network. Plug the network cable into the networking port. If you are connecting more than one network cable to the network, connect to the ports sequentially. Use the cable management tray to direct all the cabling from your system.	
7	(Optional) Connect the remote LAN module (RLM) from the back of the V-Series system to the network using an Ethernet cable. **Attention** The network switch port for the RLM connection must negotiate down to 10/100 or autonegotiate.	
8	If applicable, turn on any tape backup devices.	

Step	Action	Diagrams and references
9	For each power supply on the V-Series system, complete the following: **a.** Ensure that the power switch is in the Off (0) position. **b.** Connect the socket end of the power cord to the power plug on the power supply. **c.** Secure the power cord with the retaining adjustable clip on the power supply. **d.** Plug the other end of the power cord into a grounded electrical outlet. **Attention** To obtain power supply redundancy, you must connect the second power supply to a separate AC circuit.	
10	Start a communications program. You must use some form of communications program to perform initial network setup and V-Series configuration. You can start a communications program through RLM or through the console after connecting to the serial port.	

Connecting the switch fabric

Before you begin

Before you begin this procedure you should have done the following:

- Completed your MetroCluster configuration plan
- Configured LUNs on the storage array
- Connected the local V-Series system to the fabric
- Connected the remote V-Series system to the fabric

Attention

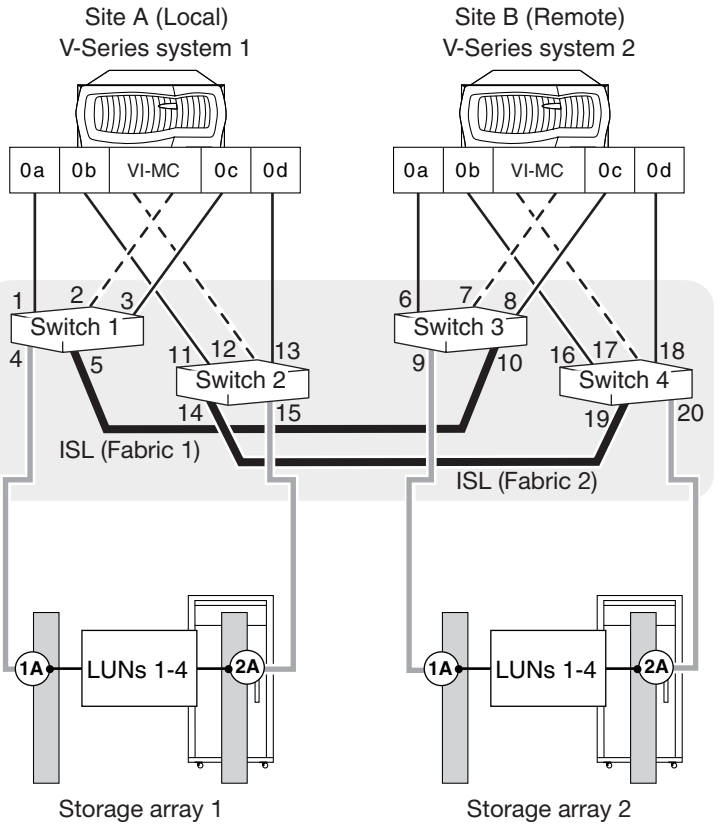
MetroCluster configurations that do not conform to the best practice configurations in this guide require a PVR. See Chapter 3, "Recommended MetroCluster Configurations," on page 13. For details about the requirements, and for complete information about the latest configuration requirements and supported systems, see the V-Series *Support Matrix*.

Sample switch fabric

The following illustration shows the switch fabric for this procedure.

Note

If you are configuring a new switch fabric, you must set the domain ID to a unique value for each fabric. Refer to switch documentation for additional details.


Site A (Local)
V-Series system 1
Site B (Remote)
V-Series system 2
0a
0b
VI-MC
0c
0d
0a
0b
VI-MC
0c
0d
1
2
3
Switch 1
4
5
11
12
13
Switch 2
14
15
6
7
8
Switch 3
9
10
16
17
18
Switch 4
19
20
ISL (Fabric 1)
ISL (Fabric 2)
1A
LUNs 1-4
2A
1A
LUNs 1-4
2A
Storage array 1
Storage array 2

Connecting the switch fabric

Complete the following steps to connect the switched fabric and apply power.

Step	Action	Diagrams and references
1	Connect the switched fabric. ◆ Connect an ISL cable to a switch on one fabric to another switch on the same fabric. In the previous example, fabric 1, Switch 1, Port 5 connects to Fabric 1, Switch 3, Port 10. ◆ Connect an ISL cable on a switch on the alternate fabric to another switch on the alternate fabric. In the previous example, Fabric 2, Switch 2, Port 14 connects to Fabric 2, Switch 4, Port 19. **Note** You must install a long distance SFP adapter in each port that you use to connect an ISL cable. You may also need to install an additional switch license to provide ISL support.	For a switch diagram, see "Sample switch fabric" on page 39.
2	Apply power to the switched fabric. Make sure that all switch IDs are set; then turn on each switch 10 minutes apart from one another.	

Connecting the fabric and storage array

Before you begin

Before you begin this procedure you should have done the following:

- Completed your MetroCluster configuration plan
- Configured LUNs on the storage array
- Connected the local V-Series system to the fabric
- Connected the remote V-Series system to the fabric
- Connected the switch fabric and powered up the switches

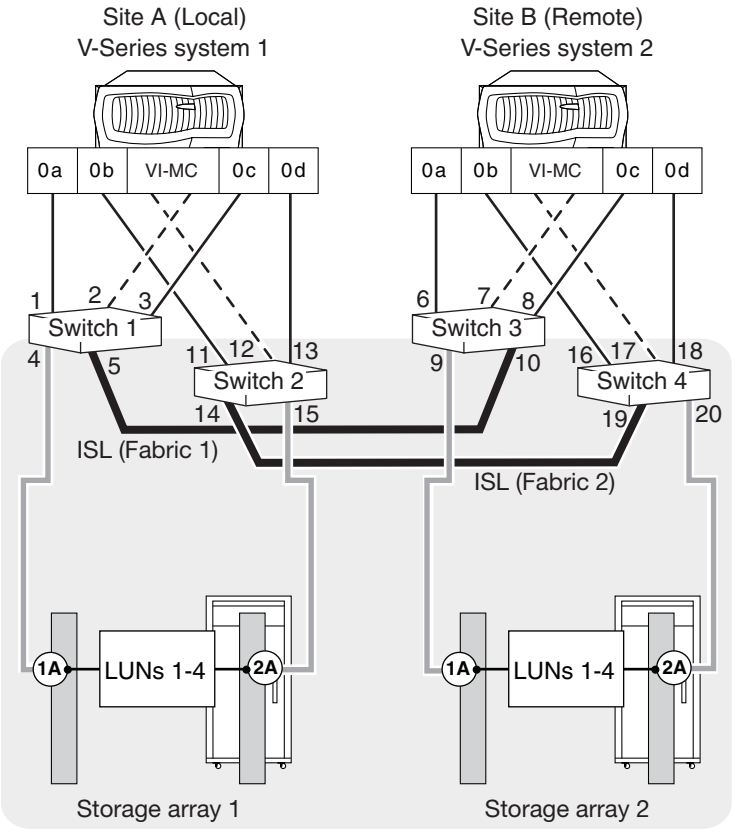
Configuration requirements

Storage array requirements vary, depending largely on the storage array model that you deploy. Before you connect the storage array to the fabric, check the V-Series *Implementation Guide* for your storage array to verify that the ports on your storage array do not access more than the number of array LUNs and host groups supported on that model.

All storage arrays, regardless of model, must be configured to allow Data ONTAP access to a specific LUN on two (primary and secondary) storage array ports.

Sample storage array connections

The following illustration shows connections between the switched fabric and the storage arrays.


Site A (Local)
V-Series system 1
Site B (Remote)
V-Series system 2
0a
0b
VI-MC
0c
0d
0a
0b
VI-MC
0c
0d
1
2
3
Switch 1
4
5
11
12
13
Switch 2
14
15
6
7
8
Switch 3
9
10
16
17
18
Switch 4
19
20
ISL (Fabric 1)
ISL (Fabric 2)
1A
LUNs 1-4
2A
1A
LUNs 1-4
2A
Storage array 1
Storage array 2

Connecting the storage arrays

Complete the following steps to connect the storage arrays to the switched fabric.

Step	Action	Diagram
1	Connect the ports on the storage arrays at Site A: ◆ Connect controller 1A to any port on one fabric. In the previous example, this is Switch 1, Port 4, Fabric 1. ◆ Connect controller 2A to any port on the alternate fabric. In the previous example, this is Switch 2, Port 15, Fabric 2. ◆ Connect additional controller ports and fabric, as required by your MetroCluster configuration.	For a storage array connection diagram, see "Sample storage array connections" on page 42.
2	Connect the ports on the storage arrays at Site B: ◆ Connect controller 1A to any port on one fabric. In the previous example, this is Switch 3, Port 9, Fabric 1. ◆ Connect controller 2A to any port on the alternate fabric. In the previous example, this is Switch 4, Port 20, Fabric 2. ◆ Connect additional controller ports and fabric, as required by your MetroCluster configuration.	

Configuring zoning

Before you begin

Before you configure zoning you should have done the following:

- Completed your MetroCluster configuration plan.
- Configured LUNs on the storage array
- Connected the local V-Series system to the fabric
- Connected the remote V-Series system to the fabric
- Connected the switch fabric
- Connected the fabric and storage array

Zoning requirements

Use switch zoning to define paths between connected nodes based on the node's unique WWN. It is recommended that you use single-initiator zoning, which limits each zone to a single V-Series system FC initiator port. Single-initiator zoning also improves discovery and boot time because the V-Series system FC initiators do not attempt to discover each other. See "MetroCluster zoning example" on page 22 for an example.

For more information about switches and zoning with V-Series systems, see the following:

- V-Series *Installation Requirements and Reference Guide*
- V-Series *Support Matrix*

Note

The configuration and firmware requirements for switches in a MetroCluster environment might differ from those in other V-Series system configurations. Always refer to the V-Series *Support Matrix* for the latest information about switch requirements.

Refer to your vender's switch documentation for information about how to configure zones on your switch.

Zoning requirements for fabric-attached MetroClusters

See "Fabric-attached MetroCluster FC-VI traffic in-order delivery requirement" on page 10 to determine whether there are any zoning requirements for your MetroCluster related to FC-VI traffic in-order delivery.

Data ONTAP Configuration and Verification 5

About this chapter

This chapter describes how to configure and test a MetroCluster configuration. Read this chapter after you have connected and powered up all the components in your MetroCluster environment.

Topics in this chapter

This chapter contains the following topics:

Data ONTAP configuration overview

The following table provides a high-level overview of what you need to do after connecting the devices in your MetroCluster.

Stage	Task	For information, see...
1	Set up the V-Series systems. This includes the following: **a.** Assigning array LUNs to the V-Series systems **b.** Creating the root volume **c.** Providing system-specific information about your V-Series system in the setup program (for example, host name and IP address)	V-Series *Setup, Installation, and Management Guide* (Data ONTAP 7.3.x releases) or the appropriate *Data ONTAP software setup guide* (for releases later than 7.3.x)
2	Install licenses, including the v-series license, licenses for SyncMirror, and Data ONTAP software.	
3	Create one or more mirrored aggregates.	Data ONTAP *Data Protection Guide*
4	Test the MetroCluster configuration.	"Testing the MetroCluster configuration" on page 49
5	Verify network and protocol setup.	Data ONTAP *File Access and Protocol Management Guide*
6	Configure other Data ONTAP features, for example, data protection features.	Data ONTAP guides

Testing the MetroCluster configuration

Testing zoning of FC-VI ports

If you had to configure the switches in your fabric-attached MetroCluster to ensure in-order-delivery of FC-VI traffic, you should test the zoning you configured for the FC-VI ports to ensure that you have not crossed the FC-VI “a” ports and “b” ports in your zones.

See “Fabric-attached MetroCluster FC-VI traffic in-order delivery requirement” on page 10 for information about requirements related to FC-VI traffic in-order delivery and zoning.

Site verification validation

Follow these steps at each MetroCluster site to test the paths, FC ports, and switch configuration in the MetroCluster environment.

Step	Action
1	Enter the following command to display path information from each V-Series system to the array LUNs: `storage show disk -p` You should see two paths to each array LUN.

Step	Action

2	To test FC ports, enter the following commands for each FC port in the MetroCluster configuration:
	a. Display the state of the HBA port.
	`fcadmin link_state`
	b. Take an FC port offline to simulate a port failure or a cable pull.
	`fcadmin offline portname`
	For example:
	`fcadmin offline 0a`
	c. Display disk path information. (The display should show only one path)
	`storage show disk -p`
	d. Bring the FC port online.
	`fcadmin online portname`
	For example:
	`fcadmin online 0a`
	e. Verify that both paths are online. Enter:
	`storage show disk -p`

Step	Action

3	Simulate a switch failure or a storage array controller failure for each switch fabric.
	
	a. Offline all FC ports on both V-Series systems that are attached to one fabric.
	
	`fcadmin offline portname`
	
	For example:
	
	`fcadmin offline 0a`
	
	`fcadmin offline 0b`
	
	`fcadmin offline 0c`
	
	`fcadmin offline 0d`
	
	b. Verify that all HBAs are disabled.
	
	`fcadmin linkstate`
	
	c. Do a takeover from site A.
	
	`cf takeover`
	
	d. When site B is in takeover mode, do a giveback from Site A.
	
	`cf giveback`

Disaster recovery verification

Follow these steps on each site in the MetroCluster configuration to simulate disaster failure and recovery.

Attention

Simulating a site failure and recovery involves disabling and degrading mirrors. You should never simulate disaster site failure and recovery in production environments without prior planning and downtime.

Step	Action
1	Disable the cluster interconnect between the V-Series systems.

Step	Action
2	Power down one of the V-Series systems. (This system will simulate the site of the disaster).

3	On the surviving system complete the following steps:
	a. Enter the command to activate a forced, manual takeover:
	`cf forcetakeover -d`
	b. Validate that forced takeover has occurred.
	`cf status`
	c. Validate the status of the aggregates, and of the volumes that they include.
	`aggr status`
	`vol status`
	The display should show an online status for aggregates and volumes, but mirrors, which are disabled, should display as "degraded."

4	Reconnect the cluster interconnect.

Step	Action
5	On the surviving system, enter the following commands: **a.** Rejoin the aggregates by entering the following command: **`aggr mirror surviving_aggregate -v victim_aggregate`** **Note** Data ONTAP uses parentheses to indicate the degraded aggregate. For example, `aggr0_b` indicates the un-degraded aggregate, and `aggr0_b(1)` is degraded. **b.** Validate that forced takeover has occurred by entering the following command: **`cf status`** **c.** Validate aggregate and volume status by entering the following command: **`aggr status`** **`vol status`** The display should show an online and mirrored status for aggregates and volumes.
6	Power up the V-Series system that simulated the disaster.
7	On the surviving system, enter the following command to re-activate the site that simulated the disaster. (This is the site you powered down in step 2.) **`cf giveback`**
8	To validate that the HA pair, aggregates, and mirrors are online and operational, enter the following at both sites: ◆ **`cf status`** ◆ **`aggr status`** ◆ **`vol status`** The display should indicate that the HA pair is enabled, and that aggregates and volumes are online and mirrored.

Disaster Recovery Using MetroCluster 6

About this chapter

This chapter describes how to use MetroCluster to recover data in situations such as prolonged power outages or natural disasters.

Note

This chapter describes recovering data on storage arrays that provide storage for V-Series systems. For information about protecting data on other network equipment—application servers, for example—see the documentation for that equipment.

Topics in this chapter

This chapter discusses the following topics:

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 V-Series 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 V-Series systems, 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 MetroCluster connection 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 node to the switch

  With this type of failure, both nodes remain running. Automatic takeover is disabled because Data ONTAP cannot synchronize the NVRAM logs. After

you fix the problem and reestablish the connection, the nodes resynchronize their NVRAM logs and the MetroCluster returns to normal operation.

- The storage from one site (site A) is not accessible to the 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 node at the other site or switch
  - If using switches, failure of any of the SFPs connecting the storage to the switch or the node to the switch
  - Failure of the Fibre Channel adapter on the node

  With this type of failure, you see a "mailbox disk invalid" message on the console of the V-Series systems that cannot see the storage. After you fix the problem and reestablish the connection, the MetroCluster 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 remain running. You see a "mailbox disk invalid" message because a V-Series 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 nodes resynchronize their NVRAM logs and the MetroCluster 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 node" on page 58.
- Force the surviving node to take over the functions of its partner. See "Forcing a node into takeover mode" on page 60.
- 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 60.
  - If you are using iSCSI, bring the failed partner's front-end LUNs (Data ONTAP-served LUNs) online.
- Fix any problems at the disaster site. See "Fixing failures caused by the disaster" on page 62.
- Re-create the MetroCluster configuration. See "Reestablishing the MetroCluster configuration" on page 62.

Restricting access to the disaster site node

You must restrict access to the disaster site node to prevent the 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 node in the following ways:

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

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

Step	Action
1	Switch off the power at the back of the V-Series system.

This is the preferred method for restricting access to the disaster site 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 node. The manual fencing method restricts access using software and physical means.

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

Step	Action	
1	Disconnect the cluster interconnect and Fibre Channel adapter cables of the 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 V-Series 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 node: `cf forcetakeover -d` **Result:** Data ONTAP causes the following to occur: ◆ The surviving 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 `cf forcetakeover -d` command is that a 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 Data ONTAP *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:** `mkdir /n/toaster/home`
2	Mount the volume. **Example:** `mount toaster:/vol/vol0/home /n/toaster/home`

Recovering front-end LUNs of the failed node

If you have a MetroCluster configuration with iSCSI-attached hosts, you must actively track the state of front-end 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 front-end LUNs in the aggregates that were mirrored at the surviving site are

offline. There is no way to distinguish the front-end LUNs that were offline before the disaster from the front-end LUNs that were online before the disaster unless you have been tracking their status.

When you recover access to the failed node's front-end LUNs, it is important to bring back online only the front-end LUNs that were online before the disaster. To avoid igroup mapping conflicts, do not bring a front-end LUN online if it was offline before the disaster. For example, suppose you have two front-end 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 front-end 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 front-end LUNs, complete the following steps.

Step	Action
1	Identify the front-end LUNs that were online before the disaster occurred.
2	Make sure that the front-end LUNs are mapped to an igroup that contains the hosts attached to the surviving node. For more information about mapping LUNs to igroups, see the Data ONTAP *Block Access Management Guide*.
3	On the surviving node, enter the following command: **`lun online lun-path ...`** *lun-path* is the path to the front-end LUN you want to bring online. You can specify more than one path to bring multiple front-end LUNs online. **Example 1:** `lun online /vol/vol1/lun0` **Example 2:** `lun online /vol/vol1/lun0 /vol/vol1/lun1` **Note** After you bring front-end 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 V-Series system has 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 storage arrays 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 node 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 MetroCluster 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: **`aggr status -r`**
2	Turn on power to the node at the disaster site. **Result:** After the node at the disaster site boots, it displays the following message: `Waiting for Giveback...`
3	Determine which aggregates are at the surviving site and which aggregates are at the disaster site by entering the following command: **`aggr status`** 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: **`aggr offline`** ***`disaster_aggr`*** *disaster_aggr* 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: **`aggr mirror`** ***`aggr_name`*** **`-v`** ***`disaster_aggr`*** *aggr_name* is the aggregate on the surviving site's node. *disaster_aggr* is the aggregate on the disaster site's node. **Result:** The *aggr_name* aggregate rejoins the *disaster_aggr* aggregate to reestablish the MetroCluster configuration.

Step	Action
6	Enter the following command at the partner node: **cf giveback** **Result:** The 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 partner node at the surviving site is shown.

```
vseries1> aggr status -r
Aggregate aggr0 (online, raid0, mirrored) (block checksums)
Plex /aggr0/plex0 (online, normal, active, pool0)
RAID group /aggr0/plex0/rg0 (normal)

RAID Disk Device HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)
  Phys (MB/blks)
--------------------------------------------------------
data vns_b200e_9:2.126L2 0d -  - 0  LUN   N/A 48796/99934208
49289/100945152

Plex /aggr0/plex2 (online, normal, active, pool1)
RAID group /aggr0/plex2/rg0 (normal)

RAID Disk Device HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)
Phys (MB/blks)
--------------------------------------------------------
data vns_b200e_8:5.126L2   0b  - - 1    LUN   N/A  48796/99934208
49297/100961280

Pool1 spare disks (empty)
Pool0 spare disks (empty)

Partner disks

RAID Disk Device HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)
Phys (MB/blks)
--------------------------------------------------------
partner vns_b200e_9:2.126L1  0d  - - 1  LUN   N/A  0/0
55552/113770496
```

```
partner vns_b200e_8:5.126L1  0b  - - 0  LUN   N/A  0/0
55459/113581440
```

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

Note

The node at the surviving site is called vseries1; the node at the disaster site is called vseries2.

```
vseries1> 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 [vseries1:
raid.mirror.resync.snapcrtok:info]: mir: created mirror
resynchronization snapshot mirror_resync.1118153658(vseries2)
Mon Nov 18 15:36:59 GMT [vseries1: raid.rg.resync.start:notice]:
/mir/plex6/rg0: start resynchronization (level 1)
Mon Nov 18 15:36:59 GMT [vseries1:
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.

```
vseries1>aggr status -r
Aggregate aggr0 (online, raid0, mirrored) (block checksums)
Plex /aggr0/plex0 (online, normal, active, pool0)
RAID group /aggr0/plex0/rg0 (normal)

RAID Disk Device HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)
  Phys (MB/blks)
---------------------------------------------------------
data vns_b200e_9:2.126L2 0d -  - 0  LUN   N/A 48796/99934208
49289/100945152

Plex /aggr0/plex2 (online, normal, active, pool1)
RAID group /aggr0/plex2/rg0 (normal)

RAID Disk Device HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)
Phys (MB/blks)
---------------------------------------------------------
data vns_b200e_8:5.126L2   0b  - - 1    LUN   N/A  48796/99934208
49297/100961280

Pool1 spare disks (empty)
Pool0 spare disks (empty)

Partner disks
```

```
RAID Disk Device HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)
Phys (MB/blks)
------------------------------------------------------------
partner vns_b200e_9:2.126L1  0d  - - 1  LUN   N/A  0/0
55552/113770496

partner vns_b200e_8:5.126L1  0b  - - 0  LUN   N/A  0/0
55459/113581440
```

Re-creating mirrored aggregates: To return the MetroCluster 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: **`aggr status -r`** **Note** A `(level-0 resync in progress)` message indicates that a plex cannot be rejoined.
2	Turn on the power to the node at the disaster site. **Result:** After the 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: **`aggr offline`** ***`disaster_aggr`*** *disaster_aggr* is the name of the aggregate at the disaster site. **Note** An error message appears if the aggregate is already offline.

Step	Action
4	Destroy every target plex that is in a level-0 resync state by entering the following command: **`aggr destroy plex_name`** For more information about the SyncMirror feature, see the Data ONTAP *Data Protection Online Backup and Recovery Guide*.
5	Re-create the mirrored aggregates by entering the following command for each aggregate that was split: **`aggr mirror plex_name`**
6	Enter the following command at the partner node: **`cf giveback`** **Result:** The 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 partner at the surviving site is shown.

```
vseries1> aggr status -r
Aggregate aggr0 (online, raid0, mirrored) (block checksums)
Plex /aggr0/plex0 (online, normal, active, pool0)
RAID group /aggr0/plex0/rg0 (normal)

RAID Disk Device HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)
  Phys (MB/blks)
--------------------------------------------------------------
data vns_b200e_9:2.126L2 0d -  - 0  LUN   N/A 48796/99934208
49289/100945152

Plex /aggr0/plex2 (online, normal, active, pool1)
RAID group /aggr0/plex2/rg0 (normal)

RAID Disk Device HA  SHELF BAY CHAN Pool Type  RPM  Used (MB/blks)
Phys (MB/blks)
--------------------------------------------------------------
data vns_b200e_8:5.126L2   0b  - - 1    LUN   N/A  48796/99934208
49297/100961280

Pool1 spare disks (empty)
```

```
Pool0 spare disks (empty)

Partner disks

RAID Disk Device HA SHELF BAY CHAN Pool Type RPM Used (MB/blks)
Phys (MB/blks)
------------------------------------------------------------
partner vns_b200e_9:2.126L1  0d  - - 1  LUN   N/A  0/0
55552/113770496

partner vns_b200e_8:5.126L1  0b  - - 0  LUN   N/A  0/0
55459/113581440
```

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:

```
vseries1> 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:

```
vseries1> aggr mirror mir1 -v mir1(1)
aggr mirror: Illegal mirror state for aggregate 'mir1(1)'
vseries1> aggr destroy mir1(1)
Are you sure you want to destroy this aggregate? y
Aggregate 'mir1(1)' destroyed.
vseries1> 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.
```

Index

H

I

L

M

N

P

R

S

U

V

W

Z