Dell EMC SRDF/Cluster Enabler Plug-in

Version 4.2.1 and later

Product Guide

REV 05

DELL EMC

Copyright © 2018-2019 Dell Inc. or its subsidiaries. All rights reserved.

Published September 2019

Dell believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS-IS”. DELL MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. USE, COPYING, AND DISTRIBUTION OF ANY DELL SOFTWARE DESCRIBED IN THIS PUBLICATION REQUIRES AN APPLICABLE SOFTWARE LICENSE.

Dell Technologies, Dell, EMC, Dell EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be the property of their respective owners. Published in the USA.

Dell EMC
Hopkinton, Massachusetts 01748-9103
1-508-1000 In North America 1-866-464-7381
www.DellEMC.com

CONTENTS

FIGURES

TABLES

PREFACE

As part of an effort to improve its product lines, Dell EMC periodically releases revisions of its software and hardware. Therefore, some functions described in this document might not be supported by all versions of the software or hardware currently in use. The product release notes provide the most up-to-date information on product features.

Contact your Dell EMC representative if a product does not function properly or does not function as described in this document.

Note: This document was accurate at publication time. New versions of this document might be released on the Dell EMC online support website. Check the Dell EMC online support website to ensure that you are using the latest version of this document.

Purpose

This guide is part of the Dell EMC Cluster Enabler for Microsoft Failover Clusters documentation set and is intended for use by system administrators during installation, system setup, and routine operations.

Audience

System administrators working with Cluster Enabler must be proficient in the use of the following.

Microsoft products

- Windows Server 2012 Standard or Datacenter editions
- Windows Server 2012 R2 Standard or Datacenter editions
- Microsoft Failover Clusters

Dell EMC products

Dell EMC storage arrays, suitable for your Cluster Enabler product version and the following applicable software:

- Solutions Enabler (SYMCLI/SYMAPI)
- Symmetrix Remote Data Facility (SRDF)
- ControlCenter Symmetrix Remote Data Facility (SRDF) Manager, if installed
- PowerPath, if installed

Related third-party documentation

The following Microsoft documentation available from microsoft.com contains information about or related to the products discussed in this guide:

- *Windows Server 2012 Clustering Whitepapers*, containing various whitepapers and datasheets about Windows Server 2012 Clustering.

DRAFT

Related documentation

The following documentation from Dell EMC contains information that may be helpful in a Cluster Enabler environment.

Dell EMC Solutions Enabler:

- *Dell EMC Solutions Enabler Array Controls and Management CLI User Guide*
- *Dell EMC Solutions Enabler SRDF Family CLI User Guide*
- *Dell EMC Solutions Enabler Installation and Configuration Guide*

Dell EMC PowerPath:

- *Dell EMC PowerPath Product Guide*

Dell EMC PowerMax:

- *Dell EMC PowerMax Family Product Guide*
- *Dell EMC PowerMaxOS 5978.221.221 Release Notes for Dell EMC PowerMax and VMAX All Flash*

VMAX All Flash:

- *Dell EMC VMAX All Flash Product Guide*
- *HYPERMAX OS 5977.1125.1125 for Dell EMC VMAX All Flash and Dell EMC VMAX3 Family Release Notes*

VMAX3 and VMAX:

- *Dell EMC VMAX3 Family Documentation Set* – Contains documentation related to the VMAX 100K, 200K, and 400K arrays.
- *Dell EMC VMAX Family Documentation Set* – Contains documentation related to the VMAX 10K, 20K, and 40K arrays.
- *Dell EMC VMAX3 Family with HYPERMAX OS Release Notes* – Details new features and any known limitations.
- *Dell EMC VMAX Family Viewer for Desktop and iPad®* – Illustrates system hardware system configurations offered for VMAX and VMAX3 arrays.
- *Dell EMC Networked Storage Topology Guide* – Provides information regarding distance restrictions for configurations of networked storage systems.
- E-Lab™ Interoperability Navigator (ELN) – Provides web-based interoperability and solution search portal. You can find the ELN at http://elabnavigator.EMC.com.

Typographical conventions used in this document

This document uses the following conventions:

Normal	Used in running (nonprocedural) text for: • Names of interface elements, such as names of windows, dialog boxes, buttons, fields, and menus • Names of resources, attributes, pools, Boolean expressions, buttons, DQL statements, keywords, clauses, environment variables, functions, and utilities
Bold	Used in running (nonprocedural) text for names of commands, daemons, options, programs, processes, services, applications, utilities, kernels, notifications, system calls, and man pages Used in procedures for: • Names of interface elements, such as names of windows, dialog boxes, buttons, fields, and menus • What the user specifically selects, clicks, presses, or types
Italic	Used in all text (including procedures) for: • Full titles of publications referenced in text • Emphasis, for example, a new term • Variables
`Courier`	Used for: • System output, such as an error message or script • URLs, complete paths, filenames, prompts, and syntax
`Courier bold`	Used for specific user input, such as commands
`Courier italic`	Used in procedures for: • Variables on the command line • User input variables
>	Separates items in a selection from a set of menus. Example: **File > Print**
< >	Angle brackets enclose parameter or variable values supplied by the user
[ ]	Square brackets enclose optional values
|	Vertical bar indicates alternate selections — the bar means “or”
{ }	Braces enclose content that the user must specify, such as x or y or z
...	Ellipses indicate nonessential information omitted from the example

Where to get help

Dell EMC support, product, and licensing information can be obtained from Dell EMC Online Support.

Note: To open a service request through Dell EMC Online Support, you need a valid support agreement. Contact your Dell EMC sales representative for details of how to obtain a valid support agreement or to answer any questions about your account.

Product information

For documentation, release notes, software updates, or for information about Dell EMC products, licensing, and service, go to Dell EMC Online Support (registration required) at:

`https://dell.com/support`

Technical support

Dell EMC offers a variety of support options.

Support by Product — Dell EMC offers consolidated, product-specific information on the Web at:

`https://support.EMC.com/products`

The Support by Product web pages offer quick links to Documentation, White Papers, Advisories (such as frequently used Knowledgebase articles), and Downloads, as well as more dynamic content, such as presentations, discussion, relevant Customer Support Forum entries, and a link to Dell EMC Live Chat.

Dell EMC Live Chat — Open a Chat or instant message session with a Dell EMC Support Engineer.

Your comments

Your comments help us to improve the accuracy, organization and overall quality of the user publications. Send your opinions of this document to:

VMAXContentFeedback@emc.com

CHAPTER 1
About Cluster Enabler

This chapter introduces the Dell EMC Cluster Enabler and shows how it provides disaster-recovery protection in geographically distributed Microsoft Failover Clusters.

Cluster Enabler overview

Cluster Enabler (CE) for Microsoft Failover Clusters is a software extension of failover clusters functionality. Cluster Enabler allows Windows Server 2012 (including R2) Standard and Datacenter editions running Microsoft Failover Clusters to operate across multiple connected storage arrays in geographically distributed clusters. Each cluster node is connected through a storage network to the storage arrays. The method of automatic failover for mirrored pairs during a node failure depends on the storage environment.

Microsoft Failover Clusters that run Cluster Enabler are referred to as *CE clusters.*

Cluster Enabler expands the range of cluster storage and management capabilities while ensuring full business-continuance protection. A Fibre Channel connection from each cluster node is made to its own storage array. Two connected storage arrays provide automatic failover of mirrored volumes during a failure of a node in a Microsoft Failover Cluster.

This connection effectively extends the distance between cluster nodes (depending on network latency) and forms a geographically distributed cluster with disaster-tolerant capabilities. The *Dell EMC Networked Storage Topology Guide* provides additional information regarding distance restrictions for your specific configuration.

Figure 1 shows an example of a Cluster Enabler configuration with two storage arrays at two sites.

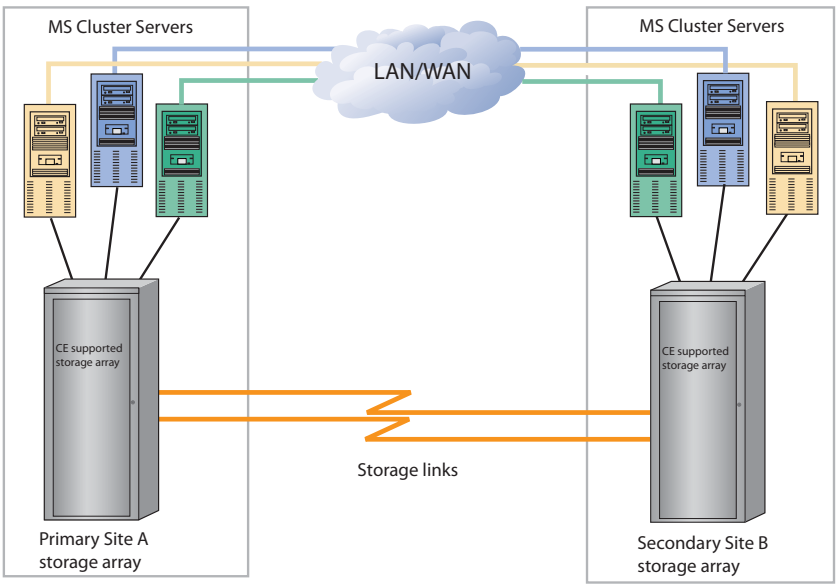

Figure 1 Overview example of a typical CE cluster configuration

Cluster Enabler plug-in architecture

Cluster Enabler for Microsoft Failover Clusters has a plug-in architecture consisting of a CE base module and separately available plug-in modules. Each CE plug-in module supports a different storage replication technology as follows:

- SRDF/Cluster Enabler for Microsoft Failover Clusters (for PowerMax, VMAX All Flash, VMAX3, and VMAX storage arrays)
- Mirrorview/Cluster Enabler for Microsoft Failover Clusters (for VNX or CX4 storage systems)
- RecoverPoint/Cluster Enabler for Microsoft Failover Clusters (for multiple RecoverPoint supported storage arrays)

The Cluster Enabler architecture supports the coexistence of multiple plug-ins, which can be installed on the same cluster node.

Note: You cannot mix replication technologies and storage configurations within the same cluster group.

Cluster Enabler components

Cluster Enabler integrates Microsoft Failover Cluster software with replication software and supported storage hardware, allowing the seamless use of disks to function as a single SCSI disk. Cluster Enabler achieves this using several components:

- **CE Manager—**An MMC-based (Microsoft Management Console) user interface that you use to configure and manage clusters for disaster recovery protection.
- **CE Resource DLL**—A Dynamic Link Library (DLL) that Microsoft Failover Clustering uses to perform group failover/failback operations for all storage group resources.
- **CE VM Resource DLL**—A DLL that Microsoft Failover Clustering uses to perform failover/failback of Hyper-V child partitions residing on Cluster Shared Volumes (CSVs).
- **CE WMI provider**—A Windows Management Instrumentation component that interfaces with the underlying storage array and performs various operations, such as failover or group creation on the storage array.
- **CE Service**—A Windows service dependent on Cluster Service, used for Quorum and CSV Device Failover, and to manage the Preferred Owners' list.
- **Quorum Filter Driver**—A component that performs arbitration or *ownership protocol* for the Microsoft Failover Clustering database quorum.

Cluster Enabler Manager interface

Cluster Enabler for Microsoft Failover Clusters provides a GUI called Cluster Enabler Manager, or CE Manager. The CE Manager provides several wizards to streamline cluster tasks and reduce the complexity of cluster management.

Using the CE Manager you can:

- Configure your Microsoft Failover Clusters for disaster-recovery protection.
- Set up and configure disk-based resources to automatically move geographically dispersed resource groups back and forth.
- Manage the storage resources available to the cluster.
- View information about the cluster and its components.

Figure 2 shows the Cluster Enabler Manager window and Table 1 summarizes the wizards.

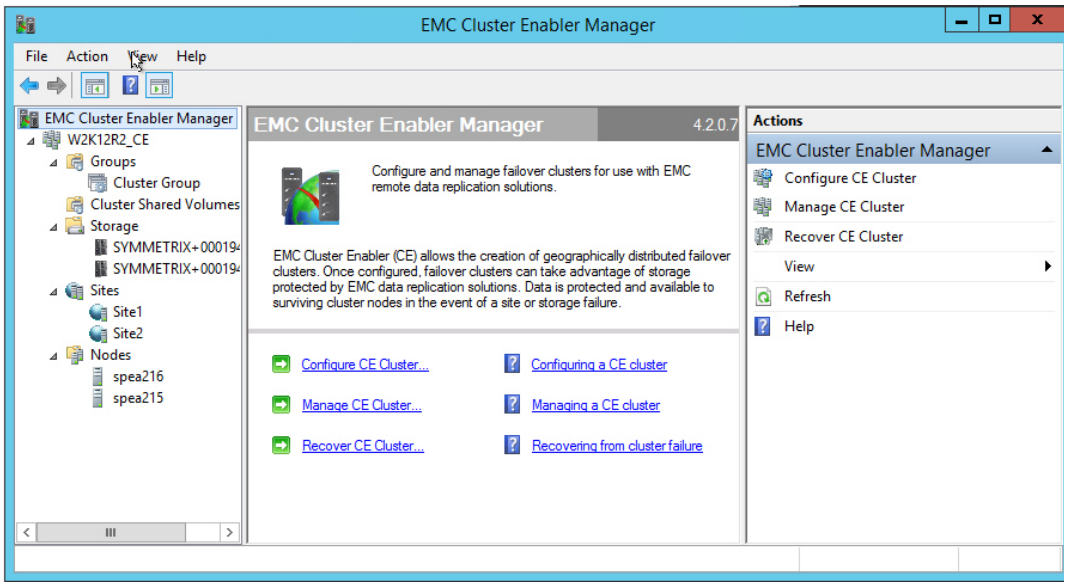

Figure 2 Cluster Enabler Manager window

Table 1 Cluster Enabler Manager Wizards

Wizard	Functionality
Configuration	Configures a CE cluster. The configuration process is the first step towards managing disaster recovery for distributed failover clusters. The Configuration Wizard leads you through the process of configuring your failover cluster for management with CE.
Create Group	Creates a CE Group, adds devices, and selects a group policy.
Modify Group	Adds or removes devices in a CE group.
Recover CE Cluster	Recovers a shared quorum cluster.
Change Quorum	Changes a cluster's quorum model type.
Update Mirror Pairs	Discovers storage, updates the storage configuration, validates the storage groups, and sets up the storage group definitions in the cluster properties database to update the mirrored pairs in a cluster.
Storage Discovery	Discovers and sets up the attached storage. Performs a storage discovery after any changes to the storage configuration.

Chapter 5, "Manage SRDF/CE," shows how to use the CE Manager and the wizards to configure and manage a CE cluster.

Cluster Enabler logging

Cluster Enabler includes a logging facility that records system activity and any errors that occur. The log files provide Dell EMC Customer Support with technical information that they can use in diagnosing and solving problems.

Characteristics of the logging facility

The logging facility has the following characteristics:

Table 2 Characteristics of the logging facility

Characteristic	Description
Log level	The log level determines how much detail the system records in the log files. Cluster Enabler provides two logging levels: • 4: all information, warning, and error messages • 5: as for level 4 with additional debug messages The initial log level, also used in normal operations, is 4.

Location of log files	As supplied, Cluster Enabler stores the log files in:
	`C:\Program Files\EMC\Cluster-Enabler\Logs`
	The name of the current log file is:
	`ce_event_trace_current.txt`

Maximum size for log files	The maximum size characteristic defines how large a log file can become before Cluster Enabler starts a new one. This size is defined in a Windows registry key and, as supplied, is 100 MB.
Retention period for log files	When the log file reaches the maximum size, Cluster Enabler closes and renames that file, and starts a new log file. The retention period determines how many log files Cluster Enabler keeps. As supplied, Cluster Enabler retains the seven most recent log files.

The initial values of the logging characteristics are suitable for many operating environments. However, you can change any characteristic to suit your site's needs. Manage CE Logging on page 90 shows how to manage the logging characteristics. That includes instructions on how you can obtain a copy of the current log file.

Appendix A lists the messages associated with the more common events that occur during Cluster Enabler operations.

Disk space requirements

The amount of disk space required depends on the logging level and the amount of cluster activity taking place. As a general guide, you might expect 50 KB each day for a normal logging level. If the logging level is set to level 5 (verbose), and cluster activity is greater than normal, you might expect 200 MB or more each day.

Microsoft Windows Server support

Cluster Enabler for Microsoft Failover Clusters runs on Microsoft Windows Server 2012 systems tailored for the x64 architecture[1].

Quorum model support

Microsoft Failover Clusters use a quorum system to determine whether a cluster is operable. There are various models for the quorum.

The following sections summarize:

- The quorum models for Microsoft Failover Clusters
- The capabilities of the Change Quorum wizard you use to change the quorum model

Quorum models

In essence, a cluster continues to operate while a majority of its resources are available. Once the cluster consists of a minority of its resources, it shuts down. There are various ways of calculating the resources in a cluster, known as a quorum models. Table 3 summarizes the models that Microsoft Failover Clustering implements.

Table 3 Quorum models for Microsoft Failover Clusters

Model name	Description
Node Majority	The cluster resources consist of the number of nodes in the cluster. This model is suited for clusters that have an odd number of nodes.
Node and Disk Majority	The cluster resources consist of the number of nodes in the cluster plus a nominated disk in the cluster, known as the disk witness. This model is suited for clusters that have an even number of nodes.
Node and File Share	The cluster resources consist of the number of nodes in the cluster plus a nominated file share, known as a file share witness. This model is suited to clusters that have an even number of nodes distributed across multiple sites.
Disk Only	The cluster resources consist of a nominated disk known as a disk witness. This is a legacy model from earlier versions of Microsoft Failover Clustering that is retained for compatibility. For new clusters, use one of the other quorum models.

The Change Quorum Wizard

Cluster Enabler includes a tool, the Change Cluster Quorum wizard, that you use to change:

- The quorum model for a cluster
- The file share witness used in a Node and File Share cluster
- The disk witness in a Disk Only cluster

Change the quorum model on page 70 shows how to use this wizard.

1. Contact Dell EMC support for information on using Windows Server 2016 or 2019.

Multiple CE cluster management

With Cluster Enabler CE Manager you can manage multiple CE clusters simultaneously, as long as all of the clusters are Windows Server 2012 clusters and are in the same domain. To manage a cluster, CE Manager runs under a domain administrator account. This account is part of local administrator group of every node of the cluster it manages.

Setting up devices on Windows Server 2012

First add all disks to Failover Cluster Management and then configure them for Cluster Enabler. By default, Failover Cluster assigns all disks to a group called *Available Storage*. Ensure that Failover Cluster can bring these disks online before using them in Cluster Enabler.

To set up devices on the Windows Server:

1. Do one of the following appropriate for your installation:
   - If there are no disks in Available Storage, ensure that all disks to be added are write-enabled on the same site (for example, site A).
   - If there are already disks in Available Storage, and you want to add more disks, ensure that those disks are write-enabled on the same site where Available Storage is online.
   - If some existing disks in Available Storage are not online, move them to the site where the Available Storage is online. If this does not solve the problem, do the following:
     1. Remove those disks from Available Storage.
     2. Move all groups and devices to the same node in Failover Cluster. Manually move the corresponding devices to ensure that devices are write-enabled on the node to which you are moving the group.
     3. Evict all remaining peer nodes.
2. Ensure that you have access to the disks where they are write-enabled. If not, restart and reformat them.
3. Right-click **Storage** in Failover Cluster Management, and select **Add a Disk**. All available disks appear in the display. Select disks to add to the cluster. All added disks appear in the group Available Storage. Verify that all disks are online in Available Storage.

The devices are now be available for use in Cluster Enabler.

Virtualization support

CE version 4.2.1 supports the following virtualization tools and features:

- Windows 2016 Hyper V
- Windows Server 2012 Hyper-V Server
- Windows Server 2012 R2 Hyper-V
- Cluster Shared Volumes
- VMware ESX Servers

Windows Server 2012 (including R2) Hyper-V server virtualization is supported for PowerMax, VMAX All Flash, VMAX3, and VMAX arrays. Once configured as a CE group, using the CE Configuration Wizard, groups with Hyper-V resources display as regular device groups.

The virtual machine (VM) and the CSV disks must first be configured in Microsoft Failover Cluster Manager.

Hyper-V

CE supports Windows Server Hyper-V server virtualization. Hyper-V is installed and managed as a role under Windows Server and requires an x64-based processor. You can use clustering to host and failover virtual machines between nodes or sites for high availability.

The clustering methods are:

Host Clustering — With host clustering, the physical host is the cluster node. If the host stops running, all of its guests (virtual machines) are restarted on another physical host. Host clustering protects against the failure of a physical host (hardware failure of a computer).

Guest Clustering — With guest clustering, a guest (that is, a virtual machine) is a cluster node, and therefore the guest runs applications that are monitored by the Cluster service. If either the guest operating system or the clustered application fails, the guest can fail over to another guest, either on the same host or on a different host. Guest clustering protects against failure of a cluster-aware application on a guest, as well as failure of an individual instance of a guest.

To get started with Hyper-V and CE version for a non-CSV disk:

1. User Server Manager to install Hyper-V.

   Instructions are available in the *Microsoft Hyper-V Getting Started Guide* on the Microsoft TechNet website.

2. Use Hyper-V Manager to create and set up a virtual machine (guest machine).

   Instructions are available in *Microsoft Hyper-V Getting Started Guide*.

3. Install an operating system on the virtual machine.
4. Install the application that you want to be highly available on the virtual machine.
5. Configure a failover cluster with the Microsoft Failover Cluster Manager for the virtual machine resources that you just created.

   The Microsoft Failover Clustering documentation contains instructions.

6. Shut down the virtual machines.

   Note: Turn off the virtual machine before adding it to the cluster.

7. Restart the virtual machines in Failover Cluster Management.
8. Open the CE Manager and configure a CE cluster using the CE Configuration Wizard.
9. On the Current Nodes wizard page, add a second node to the cluster.
10. After the node is added, follow the steps in the wizard and accept the default settings.

After the CE cluster is configured, the CE resource is part of each virtual machine service group. The physical device where the virtual machine was created is dependent on the CE resource. The CE group with the Hyper -V resource displays as a regular device group. Figure 3 shows an example of the CE Manager GUI with a Hyper-V resource.

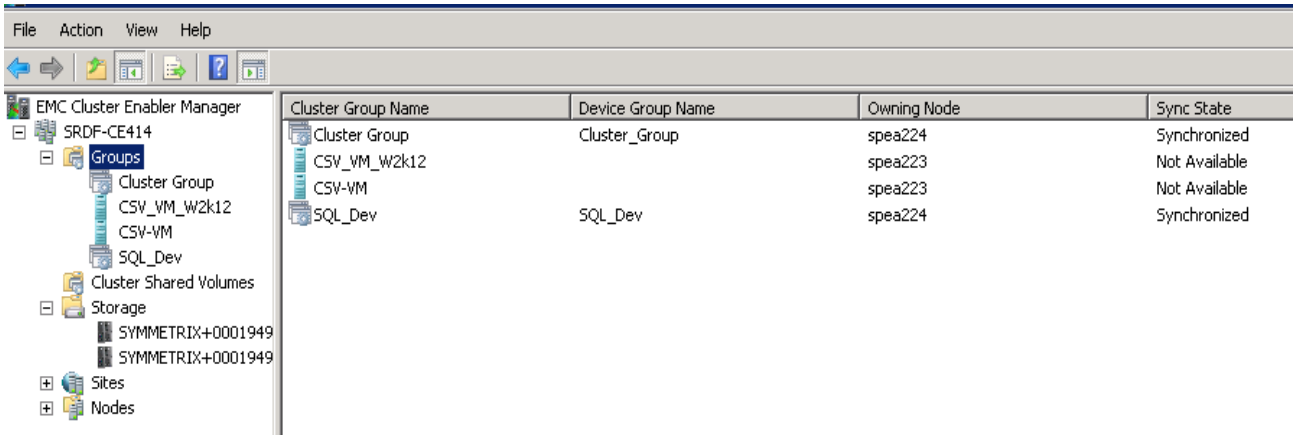

Figure 3 CE Manager with virtual machine cluster group

Cluster Shared Volumes

Cluster Enabler version 4.2.1 supports Cluster Shared Volumes (CSV). With CSV all nodes in a cluster have concurrent access to data on every CSV-enabled disk. Using Cluster Enabler, you can view the properties of or change the failover policies for a CSV disk.

For Windows Server 2012, CSV VMs can run on any node irrespective of where its CSV disk is online. This means that the VM can failover to a node where its CSV disk is marked as write-disabled.

The virtual machine and the CSV disks must first be configured in Microsoft Failover Cluster Manager. CE Manager does not allow custom resource configuration for a specific VM. Instead, run the CE Manager configuration wizard to configure all the VMs in the cluster.

Note: SRDF/CE does not support the Optimized CSV placement policies feature of Microsoft Windows Server 2012 R2. Ensure that you switch off the feature by setting the value of the CSVBalancer property to 0.

You can carry out both planned and unplanned failovers for CSV disks and virtual machines. However, in an unplanned failover of a virtual machine, only disk consistency is met.

VMware

Cluster Enabler version 4.2.1 supports the configuration of a four-node Windows Server in VMware ESX Server environments. This section shows how to configure CE in VMware environments.

CE supports two different system configurations, for either:

- A virtual machine cluster, where the virtual machines reside on two separate physical ESX servers
- A physical-virtual machine cluster, where one node is a physical host, and the other node is a virtual machine on a node in a VMware ESX cluster group

When configuring CE in VMware environments:

1. Ensure that the following applicable software and versions are installed:
   - ESX Server version 5.0 and later
   - Windows Server 2016
   - Windows Server 2012
   - Windows Server 2012 R2
   - Solutions Enabler 9.1
2. Set the timeout in the `boot.ini` file on all virtual machines to 180 seconds.

   If the `boot.ini` file currently includes only one entry, the timeout is not effective. You must populate the `boot.ini` with two separate entries. The same entry can appear twice and can be copied and pasted from the original entry. See below for an example of the `boot.ini` file.

```
[boot loader]
Timeout=180
default=multi(0)disk(0)rdisk(0)partition(2)\WINDOWS
[operating systems]
multi(0)disk(0)rdisk(0)partition(2)\WINDOWS="Microsoft Windows XP Professional"
   /noexecute=optin /fastdetect
multi(0)disk(0)rdisk(0)partition(2)\WINDOWS="Microsoft Windows XP Professional"
   /noexecute=optin /fastdetect
```

Note: No changes are necessary for physical hosts.

3. Configure a dedicated network interface card (NIC) for a heartbeat and associate the virtual NIC to a separate VLAN or network.
4. Ensure that all data devices appear to the virtual machines as raw device mapping (RDM) disks in physical compatibility mode on a separate, dedicated shared SCSI adapter.

   Note: All gatekeeper devices appear to the virtual machines as RDM disks in physical compatibility mode on a separate, dedicated SCSI adapter. The virtual SCSI adapter for the gatekeepers should not be shared with the adapter used for accessing the devices. Gatekeepers presented to the virtual machine should not be presented to any other virtual machine configured in the VMware ESX Server cluster group.

5. Follow all other VMware instructions for the configuration of Failover Clusters.

   For additional information, refer to the *Setup for Microsoft Cluster Service* technical papers available from VMware at:

   http://www.vmware.com

Additional functionality

The CE Manager has features to manage and monitor information for clusters, groups, storage devices, sites, and nodes.

Delay Failback

Delay Failback automatically modifies the Preferred Owner list for each failover cluster group so that a failover occurs to a lateral node or, if the lateral node is unavailable, to a peer node. Figure 4 shows the difference between lateral and peer nodes.

CE manages the Microsoft Failover Cluster Preferred Owner list. Whenever a group is brought online, CE examines the Preferred Owner list and determines which node is the lateral node. CE can then modify the Preferred Owner list so that the current node and its lateral partner are the first two in the list.

When a group is moved, the Preferred Owner list is modified to allow a group to fail over to a lateral node as a first option, by using failback or failover across the link. Microsoft Failover Clustering only moves a group across the link as a last resort. This prevents the failover clusters from arbitrarily performing a failback and failover across the link automatically. The Delay Failback feature delays the actual failback of a group from the primary node.

The Delay Failback feature overrides all previous configurations in all quorum-based solutions.

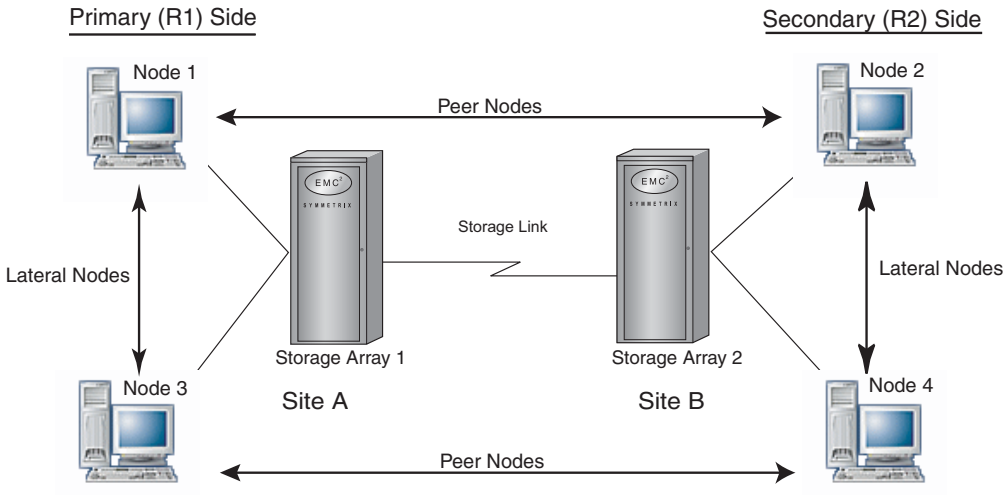

Figure 4 Lateral and peer nodes

Delay Failback runs simultaneously on all nodes. So, when a group comes online on any node, the Preferred Owner list is updated, regardless of whether it is a primary or secondary device. For example, the sequence for Group *x* on Node 1 is:

1. Delay Failback first determines if it knows the other nodes in the cluster. This information is gathered by CE during normal operations. If not, the default feature is bypassed since Delay Failback cannot differentiate between a lateral node or a peer node.
2. If Delay Failback knows the other nodes, it determines whether Group *x* is online on Node 1.

3. The Microsoft Failover Cluster Preferred Owner list is modified so that Node 1 is the first Preferred Owner, followed by the lateral node and then the peer nodes.

A cluster private property (DelayFailBackEnabled) defines whether Delay Failback is in operation. The value 0 means the feature is off, and a value of 1 means that it is on.

Mount points

Cluster Enabler provides mount points. Mount points enable you to overcome the limitation on drive letters and so a cluster can contain more than 26 volumes.

For mount points to work correctly, all related disks must belong to the same cluster group. If related disks are spread across multiple cluster groups, volumes cannot be brought online because cluster groups can be online on different nodes. To avoid this, Cluster Enabler first groups all related disks by identifying the mount points on a given disk and any disks upon which the given disk is mounted. Cluster Enabler then creates a parent/child relationship between the disks.

When you choose a disk to create a group or add a disk to an existing group, CE finds all related disks by traversing its parent/child relationships and adding every related disk to the group. It then adds appropriate dependencies between the disks so that the resources can be brought online in an orderly fashion.

Table 4 shows an example of a cluster consisting of drive letters and mount points for six volumes. The configuration shows various parent/child relationships among the disks.

For example, if you choose to add `E:\MNT1` to a group:

- `E:\MNT1` is a mount point with `E:\` as its parent.
- `E:\` is a child of `F:\`. So the group includes disk `F:\`.
- `F:\` has additional children `F:\MNT2` and `F:\MNT2\MNT3`. So the group includes these disks too.

The result of these parent/child relationships is that the group includes volumes 0BCE, 0BCF, 0BD0, 0BD1, and 0BD2. Each disk is dependent on its parent to come online. In this example, 0BCF is dependent on 0BCE, and 0BD0 is dependent on 0BCE, and so forth.

Each group is also dependent on the CE resource.

Table 4 Cluster mount point example

Drive letter and mount point	Symmetrix volume ID
`F:\`	0BCE
`F:\MNT1, E:\`	0BCF
`F:\MNT2`	0BD0
`F:\MNT2\MNT3`	0BD1
`D:\`	0BCD
`E:\MNT1`	0BD2

When you remove a device, Cluster Enabler also removes all related disks. For example, if the current mount points are `F:\`,`F:\MNT2` and `F:\MNT2\MNT3`, and if the device that corresponds to `F:\MNT2` is removed from the group, all three devices corresponding to `F:\`, `F:\MNT2`, and `F:\MNT2\MNT3` are removed.

However, if you were to first remove mount point `F:\MNT2` and then remove its corresponding device from the group, Cluster Enabler removes only the devices that correspond to `F:\MNT2` and `F:\MNT2\MNT3`. The device corresponding to `F:\` remains in the group because, after the mount point removal, it is no longer related to `F:\MNT2`.

Multiple storage array

Cluster Enabler for Microsoft Failover Clustering enables the use of multiple storage arrays by a cluster. This feature provides greater flexibility to you and your storage provisioning.

Delegating CE administration

You can manage multiple CE clusters simultaneously, as long as all of the clusters are in the same domain. To manage the cluster, CE Manager and Cluster Administrator are used with a domain account, which is part of local administrator group on every cluster node. This effectively grants full control of every cluster node to that domain account.

You can delegate the most common CE and cluster management tasks to a non-local administrator using the command-line utility `cesec.exe`.

Delegating CE administration on page 29 shows how to delegate CE administration.

System security changes

You can use `cesec.exe` command-line utility to change the following security administration privileges:

- Permit a non-local administrator to manage the cluster.
- Permit a user to make remote DCOM connections.
- Permit remote write access to the following WMI namespaces: `Root/CIMV2`, `Root/EMC`, and `Root/MSCluster`.
- Permit a user to query the Service Control Manager and to control the following CE-related services: Cluster Service (`clussvc`), CE Event Trace Service (`ce_eventtrace`), and CE Service (`cesvc`).
- Allows remote access to the CE portion of the registry (`HKLM\SOFTWARE\EMC\CE`).
- Allows the user to export CE log files by granting write access to the CE log directory (typically `C:\Program Files\EMC\Cluster-Enabler\Logs`).

Restrictions

Certain CE configuration operations are not allowed and are blocked for delegated administrators:

- CE install and uninstall
- Using the Configuration wizard to convert MS clusters to CE clusters
- Adding and deleting nodes for an existing cluster
- De-configuring a CE cluster

Viewing cluster dependency

With CE cluster dependency, you can view cluster configuration data or create a Dependency Report. The report shows all CE cluster groups and device dependencies for a cluster.

CE graphically diagrams complex storage site configurations for a CE cluster. An expanded view includes all devices involved in each site and the replication mode between sites.

You can use the dependency viewer to sort CE groups by site:

- Interconnection between devices is labeled by the mode of replication (Sync or Async).
- Remote adapter (RA) numbers are provided for each section of all configurations.
- The CSV group is listed with a CSV group name, instead of the GUID.
- The CSV Virtual Machine groups are included in the groups in which they reside.

View cluster dependency on page 73 shows how to generate a Dependency Report and includes an example.

CHAPTER 2
About SRDF/Cluster Enabler

This chapter introduces the SRDF/Cluster Enabler plug-in module. It also shows how Cluster Enabler uses the Symmetrix Remote Data Facility (SRDF) to provide disaster recovery protection in geographically distributed Microsoft Failover Clusters.

SRDF/Cluster Enabler plug-in overview

The SRDF/CE plug-in module extends the capabilities of failover clusters enabling them to operate across multiple storage arrays in geographically distributed clusters.

Each cluster node is connected through a storage network to the array (see the example in Figure 5).

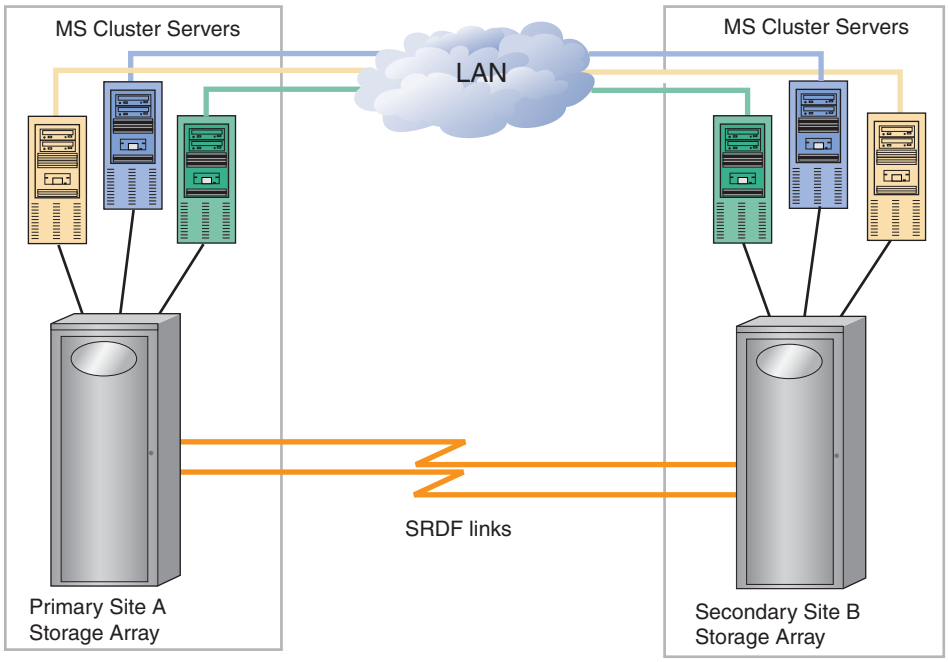

Figure 5 Example of an SRDF/CE cluster configuration

Cluster Enabler protects data from:

- Storage failures
- System failures
- Site failures

SRDF overview

SRDF is a business continuance and disaster recovery solution. SRDF uses multiple storage arrays to maintain multiple, real-time copies of logical volume data in more than one location.

SRDF duplicates data on a production site (also known as a primary or source site) to a recovery site (also known as a secondary or target site) transparently to users, applications, databases, and host processors. If the production site is not able to continue processing, data at the recovery site is current up to the last I/O transaction.

Note: The production and recovery sites are often labeled as R1 and R2, respectively.

SRDF uses include:

- Disaster Recovery
- Remote Backup
- Data Center Migration
- SDMS—Symmetrix Data Migration Service
- Data Center Decision Solutions

When production systems are unavailable, SRDF provides fast switch over to the recovery copy of the data, making data available in minutes. Business operations and related applications can resume full functionality with minimal interruption.

Protecting against data loss allows the operations and applications to resume at the secondary site. SRDF can be used:

- By itself, and data processing can be resumed by powering up a standby system and manually restarting
- In combination with other software to automatically resume operations

SRDF/CE combines Microsoft Failover Clusters and SRDF to provide a more sophisticated solution. SRDF/CE provides a wizard that you use in conjunction with the Microsoft Cluster Administrator to configure and administer the SRDF-enabled cluster.

SRDF/CE implements both SRDF/Synchronous (SRDF/S) and SRDF/Asynchronous (SRDF/A) modes of transfer.

SRDF/CE features

SRDF/CE provides the following tools and features for SRDF replication technology:

- Cluster Enabler Manager
- SRDF/Asynchronous compatibility
- SRDF/CE swap capability
- Virtual Provisioning
- Supported devices
- SRDF/CE configuration with multiple remote adapters
- Monitoring SRDF link status
- SRDF composite groups
- RDF N-X

Cluster Enabler Manager

The Cluster Enabler Manager is the facility for configuring Microsoft Failover Clusters for disaster recover. Cluster Enabler Manager also provides capabilities to both manage and monitor:

- Clusters
- Groups
- Storage devices
- Sites
- Nodes in a cluster

Chapter 5 starting on page 63 shows how to use the Cluster Enabler Manager.

SRDF/Asynchronous compatibility

SRDF/CE is compatible with SRDF/Asynchronous (SRDF/A). SRDF/A is an extended-distance asynchronous replication that uses a delta set architecture to reduce bandwidth requirements and have no impact on the performance of the host.

Asynchronous mode provides a point-in-time image on the target (R2) device that is slightly behind the source (R1) device. SRDF/A session data is transferred to the remote storage array in delta sets, eliminating the redundancy of same-track changes being transferred over the link, and so reduce the required bandwidth. SRDF/A needs only enough bandwidth to support the average production workload rather than peak workloads, provided there is sufficient cache on the storage array to support the peak workloads.

SRDF/A is intended for sites that require no host application impact while maintaining a consistent, restartable image of their data on the R2 side at all times.

SRDF/CE always enables consistency on SRDF/A groups. SRDF/A consistency ensures that applications have a consistent copy on the remote side when they failover.

SRDF/CE is compatible with PowerMaxOS, HYPERMAX OS, and Enginuity releases as outlined in the *E-Lab Interoperability Navigator*.

Note: The capacity at the target (R2) side cannot be larger than the source (R1) side. When the system fails over to the R2 side, it can never fail back since the R2 cannot resynchronize all its data back to the R1 side. In addition SRDF/A does not support a quorum group in shared quorum models. Other groups in the cluster may use synchronous or asynchronous modes, as necessary.

SRDF/CE swap capability

An R1/R2 personality swap (or R1/R2 swap) refers to swapping the RDF personality of the RDF device designations of a specified device group, so that source R1 devices become target R2 devices and target R2 devices become source R1 devices.

R1/R2 RDF swaps are available with PowerMax OS, HYPERMAX OS, or Enginuity Version 5567 or later. There are two types of R1/R2 swap:

- **FastSwap** occurs immediately after failover if the group is fully synchronized.
- **Dynamic Swap** takes longer because after failover the tracks are checked to determine whether they are synchronized, and then the swap occurs.

If you enable an R1/R2 swap for a group, SRDF/CE uses FastSwap, if it is available, during a failover. Otherwise SRDF/CE uses Dynamic Swap.

The following sections describe scenarios where an R1/R2 swap is useful.

Storage array load balancing

It is often necessary to redeploy applications and storage to a different storage array without losing disaster protection. An R1/R2 swap can enable this redeployment with minimal disruption, while offering the benefit of load balancing across two storage arrays.

For example, if you want to reconfigure an SRDF/CE environment after deciding the locations of the R1 and R2 devices, this procedure allows you to go from an active/passive configuration to active/active.

Primary data center relocation

Sometimes a primary data center needs relocation to accommodate business practices. For example, several financial institutions in New York City routinely relocate their primary data center across the Hudson River to New Jersey as part of their disaster drills. R1/R2 swaps allow these customers to run their primary applications in their New Jersey data centers. The Manhattan data centers then acts as the disaster protection site.

Post-failover temporary protection measure

You can regain a measure of protection after failing over to the remote site. If the hosts on the source side are down for maintenance, R1/R2 swap enables the relocation of production computing to the target site without giving up the security of remote data protection. When all problems are solved on the local storage array, fail over again and swap the personality of the devices back to the original configuration.

Virtual Provisioning

SRDF/CE supports Virtual Provisioning™ with SRDF/Synchronous and SRDF/Asynchronous. The *Dell EMC Solutions Enabler Array Management CLI User Guide* shows how to set up Virtual Provisioning.

Supported devices

The following types of device are compatible with SRDF/CE in point-to-point, cascaded, and concurrent configurations:

- Standard
- RAID-5
- RAID-6
- TDEVs (thin devices)
- Diskless (applicable to cascaded site A to site C configurations only)

SRDF/CE configuration with multiple remote adapters

You can configure SRDF/CE with multiple RDF links and remote adapter (RA) groups. SRDF/CE not only allows multiple RAs, but periodically tests them to ensure they are functioning. Multiple RA groups are also allowed, and these RA groups do not have to be symmetrical across all RDF links; any RA group can be allocated over a subset of the defined RDF links.

If an RDF link fails, an event log message is posted and an entry is placed in the SRDF/CE log file.

Monitoring SRDF link status

The SRDF/CE health monitoring feature for the SRDF link enables you to view link status error messages, that are recorded in the Windows event log. This feature allows you to monitor various scenarios, such as SRDF link failure.

SRDF composite groups

SRDF/CE provides SRDF composite groups (CG) that span multiple RDF groups (also called RA groups). Cascaded SRDF (see Cascaded SRDF) also uses composite groups for consistency protection during failover and failback operations.

Note: Composite groups cannot span multiple storage arrays.

Use of composite groups requires that the Solutions Enabler RDF daemon (`storrdfd`) be enabled. The *Dell EMC Solutions Enabler Installation Guide* shows how to enable the RDF daemon.

Create a CE group on page 75 shows how to create a composite CE group. Modify a CE group on page 78 shows how to add or remove devices from a composite CE group.

For additional information on composite groups, refer to the *Dell EMC Solutions Enabler SRDF Family CLI User Guide*.

RDF N-X

RDF N-X allows users to replicate data between a storage arrays running different operating environments. The replication can occur in either direction.

PowerMaxOS 5978 provides the following connectivity when creating RDF groups between PowerMaxOS 5978, HYPERMAX OS 5977 and Enginuity versions:

PowerMaxOS 5978 ↔ PowerMaxOS 5978

PowerMaxOS 5978 ↔ HYPERMAX OS 5977

PowerMax 5978 ↔ Enginuity 5876 with a SRDF N-X fix

HYPERMAX OS 5977 provides the following connectivity when creating RDF groups between HYPERMAX OS 5977 and Enginuity versions:

HYPERMAX OS 5977 ↔ HYPERMAX OS 5977

HYPERMAX OS 5977 ↔ Enginuity 5876 with fix number 67492

Concurrent SRDF

In a concurrent SRDF configuration, a single source (R1) device is remotely mirrored to two target (R2) devices simultaneously (see Figure 6). A concurrent SRDF configuration provides two identical remote copies available at any point in time. It is valuable for duplicate restarts and disaster recovery, and provides increased flexibility for data mobility and application migrations.

Concurrent SRDF technology can use two separate RA adapters in the interface link to achieve the connection between the R1 device and its two concurrent R2 mirrors. Each of the concurrent mirrors must belong to a different SRDF (RA) group.

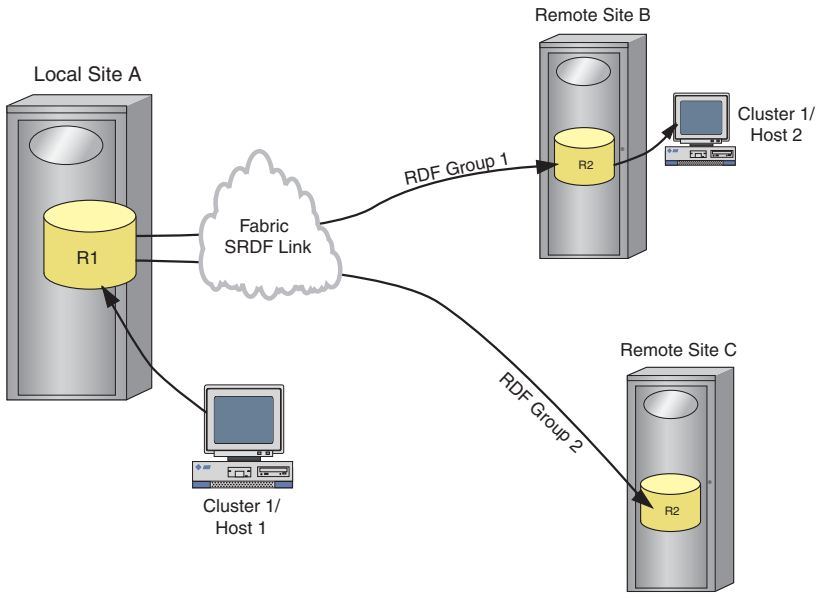

Figure 6 SRDF/CE with concurrent SRDF

The *Dell EMC Solutions Enabler SRDF Family CLI User Guide* contains details on setting up a concurrent SRDF configuration.

Restrictions and limitations

The following restrictions and limitations apply to SRDF/CE concurrent configurations:

- The CE cluster nodes are present in only two of the three sites. A configuration with CE cluster nodes present in all three sites is unsupported.
- Configurations where devices in a consistency group are mixed between concurrent, cascaded and point-to-point configurations are unsupported.
- Concurrent SRDF is not available for a configuration that includes VMAX 10K arrays or VMAXe arrays.

Supported SRDF modes for concurrent SRDF/CE configurations are:

- Synchronous for Site A to Site B (R11 → R2)
- Synchronous or asynchronous for Site A to Site C (R11 → R2)

Failover and failback behavior

This section describes the failover and failback behavior for concurrent configurations for both planned and unplanned failovers.

Planned failovers

The following occurs for planned failover scenarios in a concurrent configuration:

Failover and failback between sites A and C in a concurrent configuration (sync A → B/async A → C)

Note: CE is installed at sites A and C.

In a concurrent configuration, a planned failover between sites A and C in asynchronous mode, where SRDF does not support swap, results in sites A and B remaining as the R1 and R2. With an RDF pair state of Failed Over, where the R1 becomes write-disabled and the R2 becomes read-write enabled. Sites A and B change to an RDF pair state of invalid.

When the failback between sites A and C occurs, the configuration reverts to the original concurrent configuration, with sites A and C in asynchronous mode with an RDF pair state of consistent. Sites A and B reverts back to an RDF pair state of synchronized.

Failover and failback between sites A and B in a concurrent configuration (sync A → B/async A → C)

Note: CE is installed at sites A and B.

In a concurrent configuration, a planned failover between sites A and B in synchronous mode, a swap is performed that results in cascaded configuration (see Cascaded SRDF), where site A becomes an R21, site B becomes an R1, and site C becomes an R2.

Failover and failback between sites A and B in a concurrent configuration (sync/sync, where SRDF does not support swap)

In a concurrent configuration of sync/sync, where SRDF does not support swap, a planned failover between sites A and B in synchronous mode, results in sites A and B remaining as the R1 and R2, with an RDF pair state of Failed Over. Sites A and C change to an RDF pair state of invalid.

When the failback between sites A and B occurs, the configuration reverts to the original concurrent configuration, with sites A, B, and C in synchronous mode with an RDF pair state of synchronized.

Unplanned Failovers

The following occurs for unplanned failover scenarios in a concurrent configuration involving storage failure:

Storage failure at site A (async A → B, sync A → C)

In a concurrent configuration where the RDF mode between sites A and B is asynchronous, and synchronous between sites A and C, a storage failure at site A causes the RDF pair state between sites A and B, and A and C to become partitioned. In this failover scenario, Cluster Enabler fails over to either site B or C, depending on which node is configured as a CE site.

Storage failure at site A (sync A → B, sync A → C)

In a concurrent configuration where the RDF mode between sites A and B is synchronous, and synchronous between sites A and C, a storage failure at site A causes the RDF pair state between sites A and B, and A and C to become partitioned. In this failover scenario, Cluster Enabler fails over to either site B or C, depending on which node is configured as a CE site.

Cascaded SRDF

Cascaded SRDF is a three-way data mirroring and recovery solution that provides enhanced replication capabilities, greater interoperability, and multiple ease-of-use improvements. Cascaded SRDF allows replication between three sites without requiring the need for SRDF BCVs on the second array. A cascaded SRDF configuration does not have to use three separate sites, although that is the most common configuration for a disaster recovery solution.

The basic cascaded SRDF configuration consists of a primary site (SiteA) replicating data to a secondary site (SiteB) and replicating the same data to a tertiary site (SiteC), as shown in Figure 7. The Secondary SiteB device is labeled R21 and is:

- The R2 mirror of the Primary SiteA R1 device
- The R1 mirror of the Tertiary SiteC R2 device.

The SiteA and SiteB devices have an SRDF pair state and the SiteB and SiteC devices have an SRDF pair state.

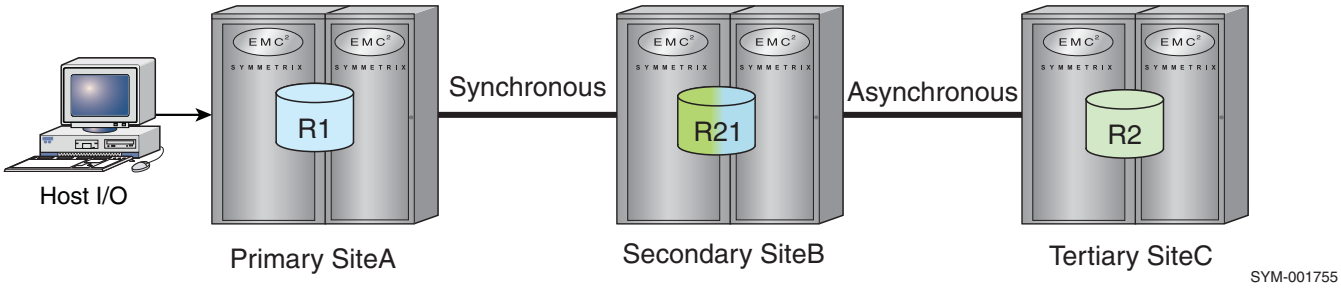

Figure 7 Sample SRDF Cascaded configuration

The *Dell EMC Solutions Enabler SRDF Family CLI User Guide* shows how to set up a cascaded SRDF configuration.

Cascaded SRDF/CE requirements

Cascaded SRDF/CE operations require the following:

- The secondary site (with the R21 devices) must be VMAX array running Enginuity 5876 and higher, a VMAX3 array running HYPERMAX OS 5977 and higher., or a PowerMax array running PowerMaxOS 5978
- R1 and R2 devices that are paired with R21 devices must be in an array that is running Enginuity 5876 and higher or HYPERMAX OS 5977 or PowerMaxOS 5978.

SRDF modes for cascaded SRDF/CE support are:

- Synchronous for SiteA to SiteB (R1 → R21)
- Asynchronous for SiteB to SiteC (R21 → R2)

Note: Currently, ACP disk mode from B to C is not available.

Restrictions and limitations

The following restrictions and limitations apply to SRDF/CE cascaded configurations:

- The CE cluster nodes are present in only two of the three sites. A configuration with CE cluster nodes present in all three sites is unsupported.
- CE can be installed only at the following sites:
  - At Sites A and B; where the replication mode is synchronous

  or
  - At sites A and C
- CE configurations where devices in a consistency group are mixed between concurrent, cascaded and point-to-point configurations are unsupported.
- A cascaded disk cannot be the quorum disk.
- Cascaded RDF devices are not discovered in CE if the R21 and R2 arrays are mapped to the same R2 host.
- Cascaded SRDF is not available for a configuration that uses VMAX 10K or VMAXe arrays.

Failover and failback behavior

This section describes the failover and failback behavior for cascaded configurations for both planned and unplanned failovers.

Planned failovers

The following occurs for planned failover scenarios in a cascaded configuration:

Failover and failback between sites A and B in a cascaded configuration

In a cascaded configuration, a planned failover between sites A and B in synchronous mode, results in a swap and hence a concurrent configuration.

When the failback between sites A and B occurs, the configuration reverts to a cascaded configuration.

Failover and failback between sites A and C in a cascaded configuration

In a cascaded configuration, a planned failover between sites A and C involves two consecutive failovers between sites A , B, and C as follows:

- A failover between sites A and B (synchronous mode)
- A failover between sites B and C (asynchronous mode)

Unplanned Failovers

The following occurs for unplanned failover scenarios in a cascaded configuration involving storage failure:

Failover from site A to site C with a storage failure at site A

In a cascaded configuration where the RDF mode between sites A and B is synchronous, and asynchronous between sites B and C, a storage failure at site A causes the RDF pair state between sites A and B to become partitioned, and consistent between sites B and C. In this failover scenario, CE fails over from B to C and the applications are transitioned (online) to site C.

When site A storage is restored, the RDF pair state between Site A and B become suspended, and consistent between B and C. To transition applications back online to site A, CE performs the following steps:

1. A failover from site A to site B
2. A failback from site C to site B
3. A failback from site B to site A

Once these steps are completed, the RDF pair state between sites A and B returns to a synchronized state, and the RDF pair state between sites B and C returns to a consistent state.

Storage failure at sites A and B

In a cascaded configuration where the RDF mode between sites A and B is synchronous, and asynchronous between sites B and C, a storage failure at both sites A and B causes sites A and B to become unreachable, with an RDF pair state of partitioned between sites B and C. If transmit idle is disabled, a failover is necessary to bring applications online at site C. If transmit idle is enabled, applications are automatically brought online at site C. When site A storage is restored, the RDF pair state between sites A and B is either suspended or split, and the RDF pair state between B and C is split.

Note: A split state scenario requires intervention of an administrator to return the SRDF pair state to Failed Over.

To transition applications back online to site A, CE performs the following steps:

1. A failover from site A to site B
2. A failback from site C to site B
3. A failback from site B to site A

Configuring cascaded SRDF with CE Manager

After all nodes have been discovered, the CE Configuration Wizard validates the presence of CE cluster nodes in two of the three storage sites. The SRDF configuration can use synchronous replication between R1 and R21 devices and asynchronous replication between R21 and R2 devices.

Pre-SRDF/CE clustering considerations

To ensure disaster recovery protection, consider the following before installing and configuring a cluster that uses SRDF/CE:

- Cabling
- Booting
- SRDF coexistence

Cabling

Avoid routing all cables through the same path, both in buildings and between sites. To provide an installation with no single point of failure, use a configuration similar to Figure 8.

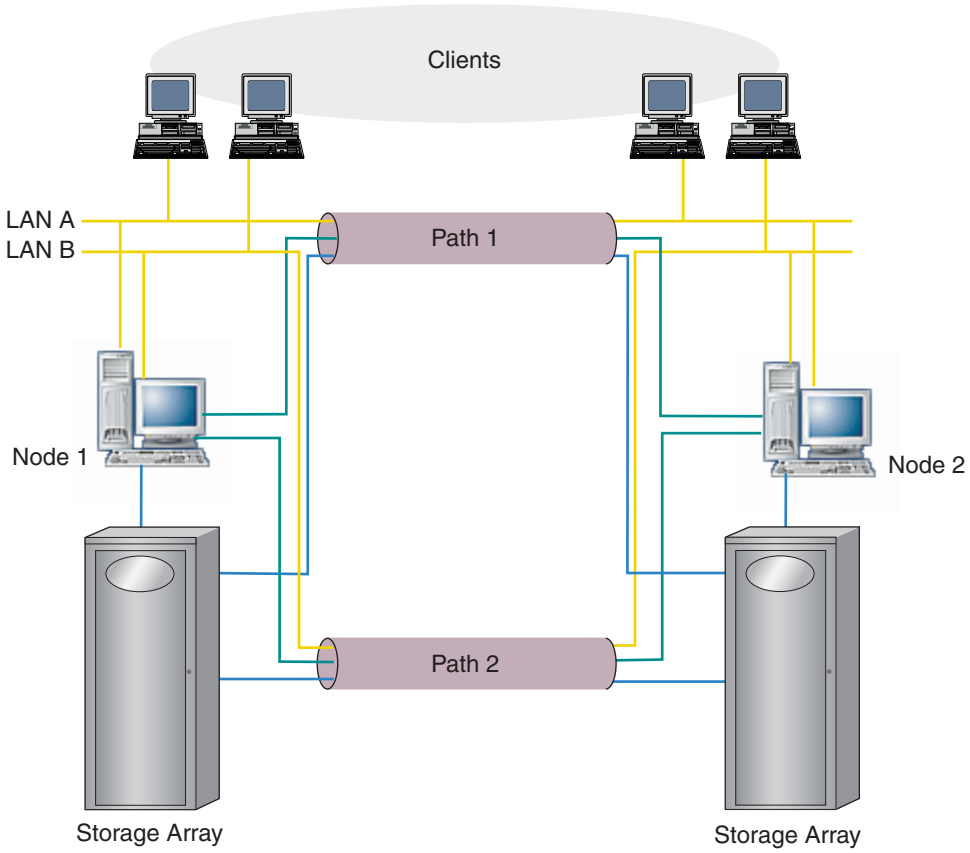

Figure 8 Suggested cabling configuration

Booting

Currently, Microsoft Failover Clusters can boot only from separate private disks (they cannot boot off the same bus). Therefore, each CE node must contain an internal disk for booting or be attached to a nonclustered disk.

SRDF coexistence

Multiple SRDF/CE clusters can share the same SRDF pair. SRDF/CE software can extend the storage array system to support up to 64 shared quorum disk clusters for each pair of arrays. There is no limit on the number of Majority Node Set clusters per pair of arrays.

SRDF/CE support matrix

Table 5 contains the SRDF/CE support matrix for Microsoft Windows Server operating systems by CE version.

Table 5 Operating system support matrix for SRDF/CE

Microsoft Windows Server OS	CE 4.2.1.26	4.2.1.17	4.2.1.12	4.2.1.4
Windows 2016	Y	Y	Y	Y
Windows 2012 R2 (standard and Datacenter)	Y	Y	Y	Y
Windows 2012 (Standard and Datacenter)	Y	Y	Y	Y
Windows 2008 R2 (Enterprise and Datacenter)	N	N	Y	Y
Windows 2008 (Enterprise and Datacenter)	N	N	Y	Y
Windows 2003 (Enterprise and Datacenter)	N	N	N	N

Table 6 shows the compatible version of Solutions Enabler and the Cluster Enabler plug-in for each CE version.

Table 6 Compatible versions of Solutions Enabler and Cluster Enabler plug-in

	4.2.1.26	4.2.1.17	4.2.1.12	4.2.1.4
Compatible version of Solutions Enabler	9.1	9.0	8.4	8.2
Compatible version of Cluster Enabler plug-in		4.2.1.10	4.2.1.6	4.2.1.4

CHAPTER 3
Clustering Concepts

This chapter describes clustering concepts for Microsoft Failover Clusters using a Cluster Enabler cluster solution and the modes of operation:

Microsoft Failover Clusters

Microsoft Failover Clusters is the clustering extension to Windows Server 2012 Enterprise and Datacenter editions. Microsoft Failover Clusters protect against failure of production server hardware or network connections. For data protection, Microsoft Failover Clusters use a protected storage subsystem. The standard failover cluster relies on RAID 1 or RAID 5 array storage to guarantee data protection.

In a typical failover cluster containing one to eight nodes, server nodes share the application workload. Typically, in a node cluster environment with *n* nodes, each node serves one-*n*th of the total number of disks and clients connected by a common SCSI bus. If one server node fails, one or several of the remaining nodes take ownership of all the disks and assume all the application workload.

Figure 9 shows an example of a two-node failover cluster and Figure 10 a four-node failover cluster.

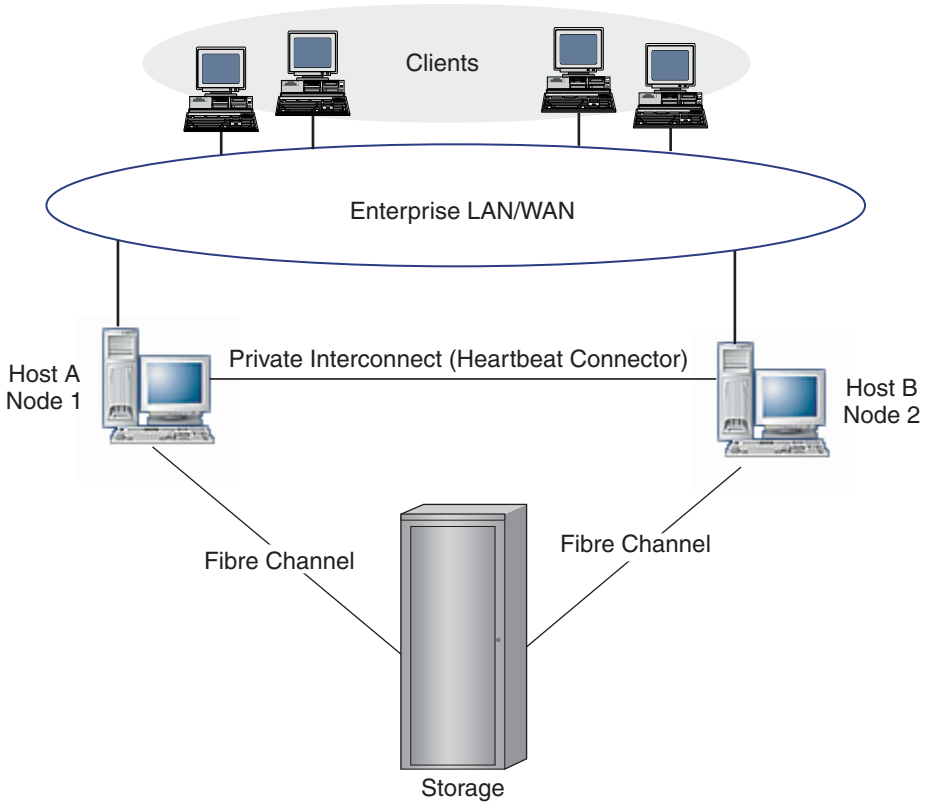

Figure 9 An example of a two-node Microsoft Failover Cluster

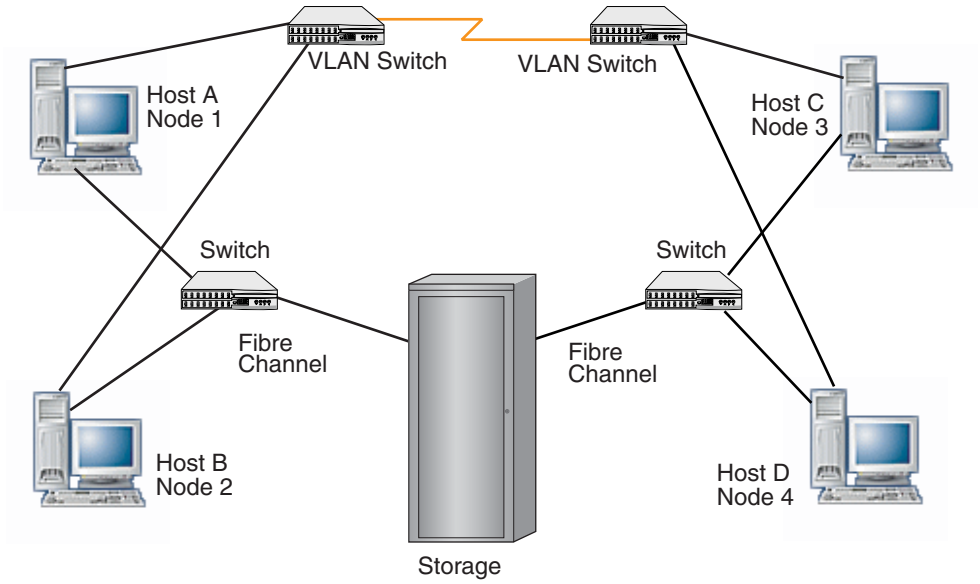

Figure 10 An example of a four-node Microsoft Failover Cluster

For information about Microsoft Failover Clusters, failover, the benefits of clustering, and its limitations, refer to the Microsoft documentation.

CE geographic cluster system

Cluster Enabler provides disaster-tolerant capabilities that enable the cluster servers to be geographically separated. Figure 11 shows a typical hardware configuration of a two-node CE cluster solution.

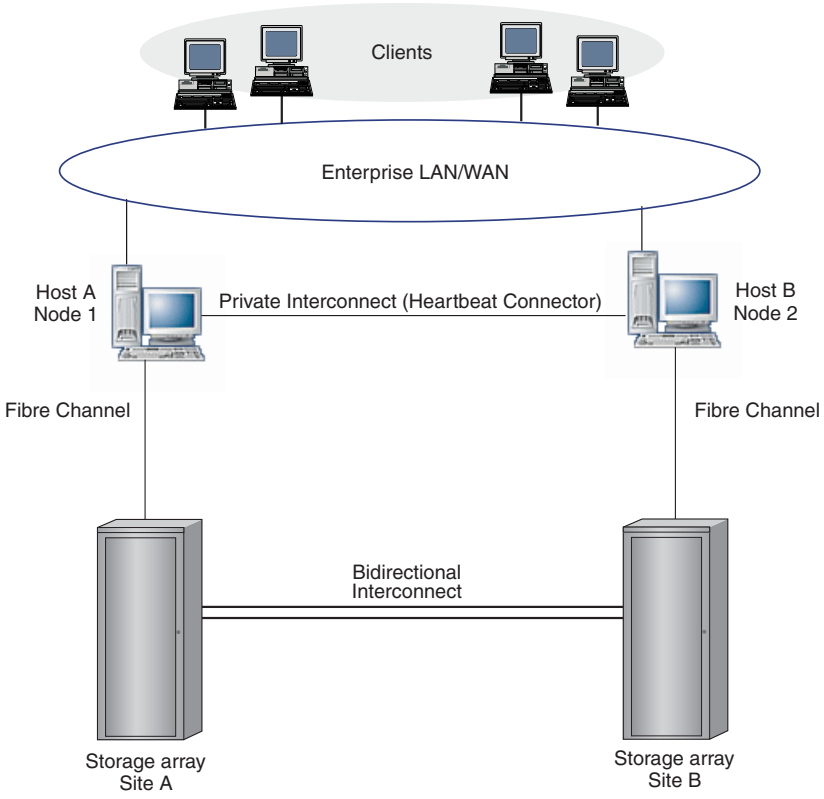

Figure 11 A geographically distributed two-node CE cluster

Network connections can provide a guaranteed maximum round-trip latency between nodes of up to 300 ms. Since many servers can connect to one storage array, it is possible to implement many clusters across this distance.

Figure 12 shows a hardware configuration of a four-node cluster solution.

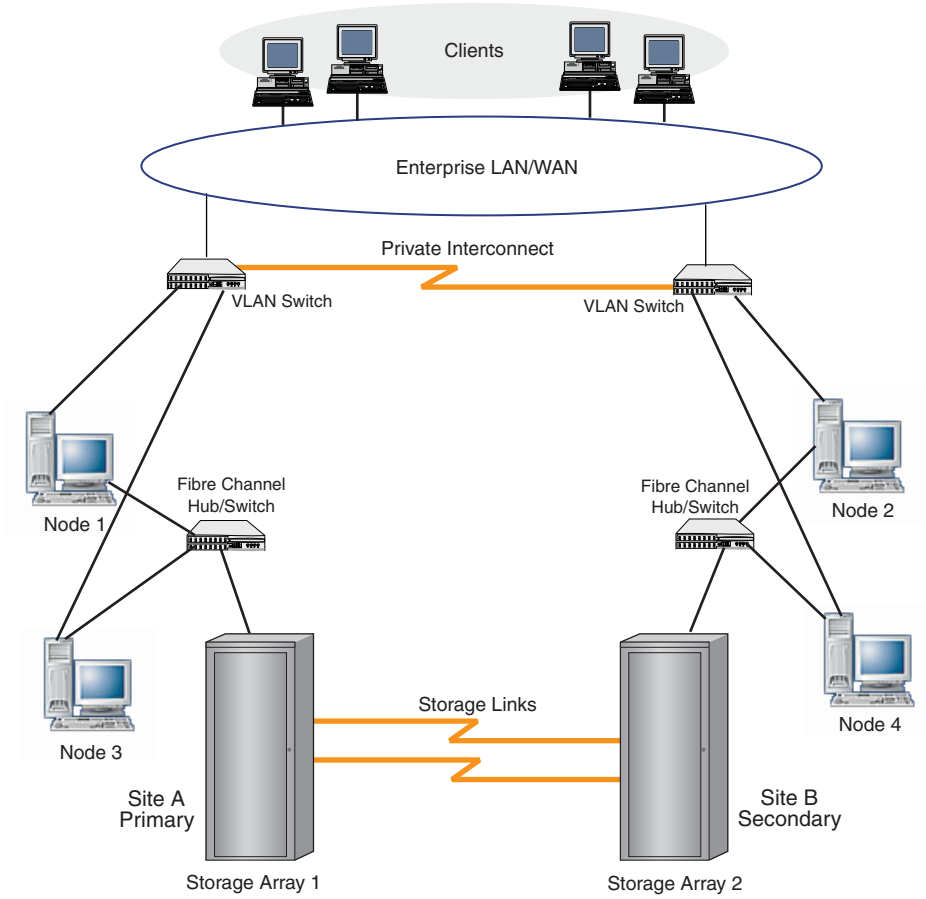

Figure 12 A geographically distributed four-node CE cluster

Cluster Enabler provides disaster-tolerant capabilities through mirroring and failover capabilities. CE allows two storage arrays to be attached using direct-connect fiber.

Cluster Enabler modes of operation

Cluster designs use different modes of operation and data-sharing mechanisms. The configuration for a CE two-node or multinode cluster in a geographically distributed cluster environment is either:

- Active/passive

or

- Active/active

Note: The terms active/active and active/passive apply to the cluster and to the applications running on the cluster. Both the cluster software and the application software must be designed for active/active operation.

Active/Passive

An active/passive cluster contains two or more nodes where all processing is done on one node during normal operation, and the work is picked up by a remaining, passive node (or nodes) only when a failure occurs on the active node. In a two-node configuration, half of the hardware is normally idle. When failover occurs, the application restarts with full performance.

Active/passive multinode clustering provides greater flexibility than the standard active/passive Microsoft failover cluster two-node cluster by providing more options in resolving failures and load distribution after server failures. For example, in a multinode cluster, the configuration may include one or more passive servers to take over the load from other servers during a site failure, or you may distribute the load among the surviving active nodes.

Active/Active

An active/active cluster contains two or more nodes where all nodes are running application software during normal operation. When a failure occurs on a node (or nodes), the work is transferred to a remaining node (or nodes) and restarted. The one or more nodes that pick up the work must then handle the processing load of both systems, and performance is usually degraded. However, all the computer hardware is used during normal operation.

Application software in a cluster environment

Software running on a cluster can be cluster aware. When software is cluster aware, it provides a restart mechanism that engages whenever the application resource is moved to another node in the cluster.

Application failover requires a restart of the application whenever failover occurs. Restart is not instantaneous. Unlike a fault-tolerant computer, a distributed cluster does not provide nonstop computing. The time that the restart takes, and the completeness of the recovery, depend on the application.

- For a transaction-oriented application (such as SQL or Exchange that contain both a database and transaction log files), the application provides a restart mechanism to recover work in progress. Usually a transaction log is used to record all work in progress. When a node fails, the information in host memory is lost, but the work can be reconstructed by applying the transaction log to the database to restart. This mechanism recovers all transactions completed before the failure. Transactions partially complete are lost and must be re-entered.
- Some applications (such as Microsoft Word or Microsoft Excel) provide a checkpoint capability. If the application experiences a failover, all work since the last disk checkpoint is lost.
- If an application has neither a database nor checkpoint capability, and does not retain state information between client requests (such as a web browser or a Microsoft Outlook client), it can fail over by reissuing the outstanding request. In this scenario, no work is lost, and no restart is needed on the server.
- If the application has neither a checkpoint nor restart capability, and it retains the state between client requests to the server, it must be rerun from the beginning when the node it is running on fails.

CHAPTER 4
Cluster Failover Behavior

This chapter describes the failover behavior of SRDF/CE in various operational modes.

Cluster failover operation

Clusters are designed to overcome failures. There are several possible types of failure and Cluster Enabler protects against more failure scenarios than local clusters can. Failure of an individual client affects only one user and is not discussed in this chapter. Figure 13 shows a geographically distributed two-node SRDF/CE cluster with two storage arrays. It also shows eight cluster elements that can fail (singly or in combination).

This section describes the following:

- SRDF/CE failover and recovery behavior
- SRDF/CE unique behavior
- Complete site failure and recovery

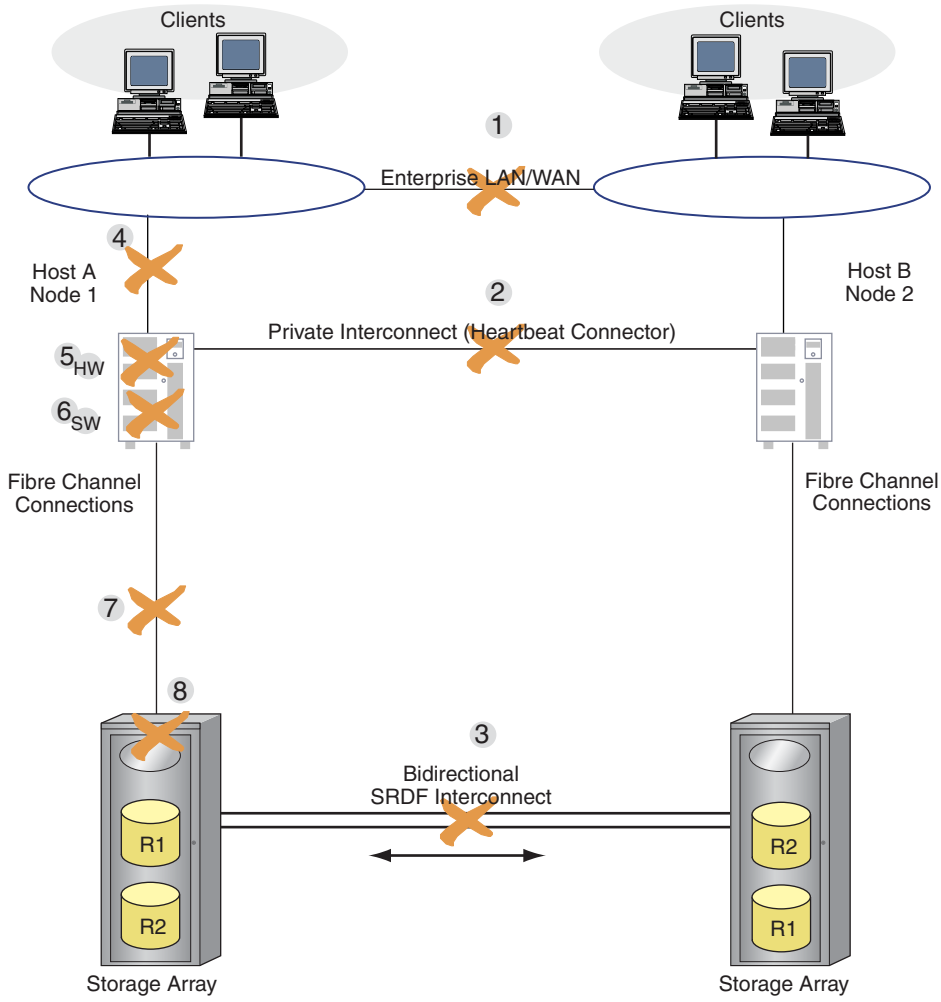

Figure 13 SRDF/Cluster Enabler failover operation

The starting condition for each of these failure scenarios is:

- Both nodes are operational.
- Node 1 owns the quorum disk.
- Both the public link (internode LAN link) and the private link (heartbeat link) are configured in Microsoft failover clusters as *enabled for all network access*.

CE provides failover and recovery operations where:

- The behavior of Cluster Enabler is the same as Microsoft failover local clusters.
- The geographic separation and disaster tolerance of Cluster Enabler causes unique behavior and provides recovery alternatives.

SRDF/CE failover and recovery behavior

The following sections introduce SRDF/Cluster Enabler failover and recovery behavior common with Microsoft Failover Clustering. The numbers in the section headings refer to the numbered failure points in Figure 13.

LAN link failure (1)

If the LAN connection between nodes fails, both servers are still available and can communicate over the heartbeat link. No failover occurs, current processing continues, and client requests from clients connected to the LAN locally continue to be serviced. Client traffic from clients connected through the LAN link fail.

Heartbeat link failure (2)

If the heartbeat link fails, Microsoft Failover Clustering routes heartbeat messages across the public LAN. Operation of the cluster continues with no failover of resources.

Host NIC failure (4)

The host is cut off from all clients. Processing continues uninterrupted on the other host. On the failed host, client input to that host fails, but current processing activities continue. Microsoft Failover Clustering detects the NIC has failed. The isolated node takes resources offline to halt processing. The other node brings the failed resources online so application failover can occur.

Server failure (5)

If the host node hardware or its operating system fails, all heartbeat messages to the remaining node cease. The remaining node then uses the quorum disk to discover the first host has failed. The remaining node then brings the resources of the failed node online and starts the applications recovery procedures.

Application software failure (6)

If an application fails, Microsoft Failover Clustering initiates a failover to the remaining node. The Cluster Enabler resource monitor makes the storage resource for the failed application available on the other node to allow application failover.

Host bus adapter failure (7)

An HBA failure is a resource failure that triggers a cluster failover operation. If both storage arrays are still running, the failover operation completes normally.

SRDF/CE unique behavior

The following sections introduce SRDF/Cluster Enabler unique behavior which is different from Microsoft Failover Clustering behavior. The numbers in the section headings refer to the numbered failure points in Figure 13.

Storage array failure (8)

When a mirrored disk fails in a storage array, it is not visible to the host because normal operations continue with the mirror, and the failed drive is hot replaced without disturbing the host. However, if an entire storage array fails, it appears to its attached server as a resource failure indistinguishable from an HBA failure. The Microsoft Failover Cluster on that server triggers a failover operation. However, because the storage array itself has failed, the remaining devices recognize that communication is lost and prevent failover from completing unless automatic failover is set as Complete site failure and recovery describes.

SRDF link failure(3)

If the link between the storage arrays fails, the Dell EMC ControlCenter® Symmetrix Manager, or the Symmetrix Management Console application notices the condition and reports an error.

The Microsoft Failover Cluster server does not notice the change (because access to existing disk resources is not disturbed). However, when SRDF/CE detects an SRDF link failure, the appropriate actions are taken (for example, synchronize the mirror group, swap the personality, and so on) when the SRDF link is restored. SRDF link failures or any failures in performing the restore action are noted in the Event Log, and in the SRDF/CE log.

Note: Upon link recovery, a synchronization operation is attempted on RDF devices that are in the suspended state. Devices in a split, mixed, or other state are not automatically synchronized. SRDF/CE detects transitions of the replication link between offline to online when the time between the two states is more than one minute. For transition times less than one minute, devices may not automatically synchronize and would require manual intervention to synchronize them.

If Microsoft Failover Clustering or a user attempts to fail over or fail back a group, and there is no link available to perform that operation, the operation fails. However, if there are multiple active lateral nodes and the groups in question are on that lateral side, lateral-to-lateral failover can occur.

You can override this behavior by enabling the Automatic Failover feature for a particular group.

Complete site failure and recovery

Local Microsoft Failover Cluster

In a local Microsoft Failover Cluster, if an entire site fails (such as from a flood, fire, or other disaster) the entire cluster fails. By contrast, with a CE cluster, each site contains only one of the two nodes in the cluster (or only one of the *n* nodes in a multinode cluster).

CE cluster

A complete site failure can be caused by either a site failure or a total communication failure (see Figure 14).

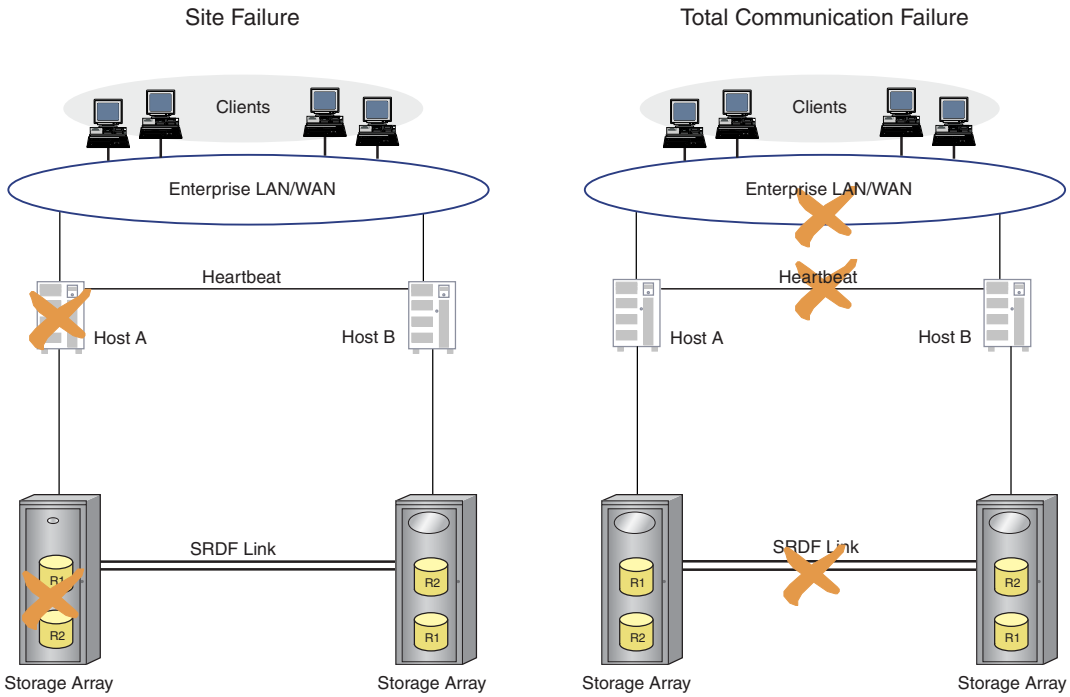

Figure 14 Types of complete site failure

Note: A surviving node cannot determine which type of failures caused the site failure.

Site (server and storage) failures (5+8)

Site failure occurs when the host and storage array both fail.

Total communication failure (1, 2 and 3)

A total communication failure can occur while the host and storage array remain operational (such as a mechanical digger rupturing the cables outside a building).

A total communication failure, while both nodes remain operational, is referred to as a split-brain condition and can lead to corruption of logical data. For example, if both sides assume the other is dead and begin processing new transactions against their copy of the data, two separate and unreconcilable copies of the data can result.

Both nodes are isolated from each other, but not from local clients. It is impossible to determine if the other node is alive. No remote client processing is possible, but running processes continue.

Response to complete site failure

In Cluster Enabler, the site failure modes determine the behavior of a cluster when a failure occurs, separating the two storage arrays and suspending remote data mirroring protection.

If a complete site failure occurs, Microsoft Failover Clustering on the surviving node detects that it is not receiving heartbeat messages. Microsoft Failover Clustering attempts to communicate with the other node using the LAN communication path.

Microsoft Failover Clustering then queries the status of the disk resource and decides whether to bring the disk resources on the local node online or to set them offline. The commands to perform this query from Microsoft Failover Clustering to Cluster Enabler are:

- **Is Alive?** — Determines whether a currently online resource is healthy and can continue to be used, or whether it and all dependent cluster resources must be taken offline.
- **Online Request** — Changes the state of an offline resource to online for a failover.

Each group's failover option setting determines how Cluster Enabler responds to queries from the Cluster Service.

Inappropriate user actions that cause groups to *bounce back* act differently. If you attempt to move the quorum group when the SRDF link is down, the Microsoft Failover Clustering destination node terminates, and the group bounces back. Active/active configurations are affected because any applications on the destination node now move. This behavior is a result of the preceding behavior.

The Cluster Enabler site failure mode settings are:

- **Restrict Group Movement** — In an SRDF link failure, this setting attempts to move disks laterally only.
- **Automatic Failover** — The Automatic Failover policy sets the group to allow automatic failover to another remote (peer) node in the event of an SRDF link failure.

Whenever a failure occurs such that mirrored data protection between sites is lost (for example, the SRDF link is down or a storage array is down), Cluster Enabler responds to the failure by not allowing any new disk groups to be brought online until communication with the other node is reestablished (unless the Automatic Failover feature is set).

NOTICE

Data Loss may occur for any group from Nodes 1 and 3 that is brought online with Automatic Failover if outstanding writes were not mirrored to the secondary site.

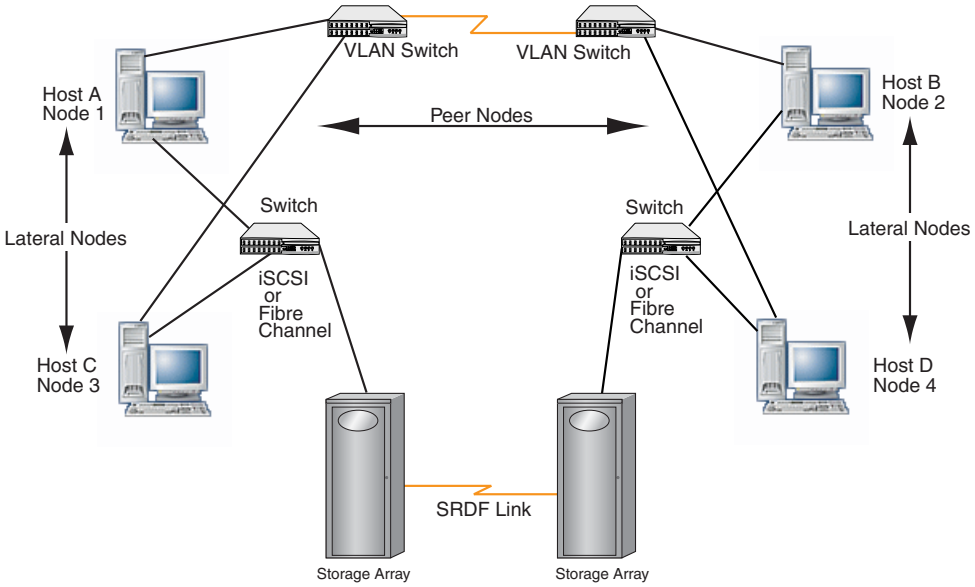

Figure 15 Lateral and peer nodes

Quorum disk-based clusters for SRDF/CE

For a cluster that uses quorum disks, the side that remains up with respect to a storage array is based on which node owns the quorum resource. In a site disaster, Failover Clusters keep all nodes up on the side owning the quorum. All resources owned by the other side are moved to the surviving side.

In the quorum disk case, SRDF/CE monitors all nodes. If tracks (data) are not owed to the surviving side, the move proceeds smoothly. If tracks are owed to the surviving side, the Automatic Failover option is required to make the move successful. If SRDF/CE detects a split-brain condition during normal group failover processing, the Automatic Failover option causes the failing site to successfully transition to the new site.

Automatic failover

In addition to the site failure mode settings, you can override the mode behavior and bring resources back online under user direction through the Automatic Failover feature. This feature lets you decide where processing is allowed to continue.

If you determine that one site is actually down, and the other site remains operational, you can use the Automatic Failover feature to:

- Override the failure mode.
- Allow disk resources to be brought online, even though SRDF is not operating and there is no mirror protection of data.

NOTICE

Use the Automatic Failover feature with great care.
Dell EMC does not recommend using the Automatic Failover feature during normal non-disaster operations because it can cause data loss.

Failure behavior when using Majority Node Set with File Share Witness

This section shows the failover behavior of the four-node cluster for Majority Node Set (MNS) with File Share Witness shown in Figure 16.

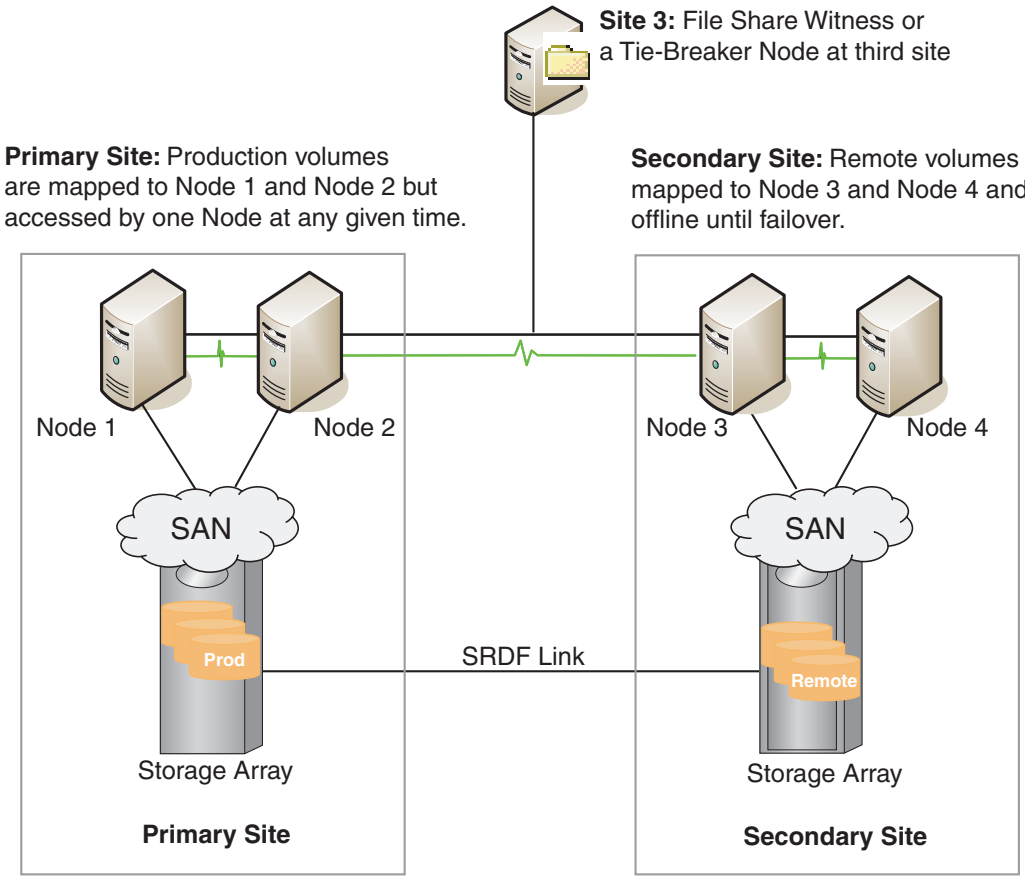

Figure 16 MNS clusters with File Share Witness

In the following examples, groups are cluster groups that contain one or more Cluster Enabler managed physical disk resources. The failover policy is Restrict Group Movement.

Storage failure at primary site

- Groups on Nodes 3 and 4 remain online but cannot failover.
- Groups on Nodes 1 and 2 move to Nodes 3 and 4 but stay offline and must be brought online manually by enabling Automatic Failover.

NOTICE

Data Loss is possible for any group from Node 1 and 2 that are brought online with Automatic Failover, if outstanding writes were not mirrored to the secondary site.

SRDF link failure

- Groups on Nodes 3 and 4 remain online but cannot failover.
- Groups on Nodes 1 and 2 remain online but cannot failover.
- To move a group to a different node, enable Automatic Failover on the destination node.

NOTICE

Data Loss is possible for any group that is moved with Automatic Failover if outstanding writes were not mirrored.

Site failure (server and storage) at primary site

- Groups on Nodes 3 and 4 remain online but cannot failover.
- Groups on Nodes 1 and 2 move to Nodes 3 and 4 but stay offline and must be brought online manually by enabling Automatic Failover.

NOTICE

Data Loss is possible for any group from Nodes 1 and 2 that are brought online with Automatic Failover if outstanding writes were not mirrored to the secondary site.

Total communication failure

- If all nodes have connectivity to the file share witness, the cluster takes two of the nodes at one site offline.
- If only one node has connectivity to the file share witness, the cluster takes the other nodes offline.
- If no nodes have connectivity to the file share witness, the entire cluster goes offline. (See Microsoft procedures for forcing an MNS cluster node online.)
- If Nodes 3 and 4 are the surviving nodes:
  - Groups on Nodes 3 and 4 remain online but cannot failover.
  - Groups on Nodes 1 and 2 move to Nodes 3 and 4 but stay offline and must be brought online manually by enabling Automatic Failover.

NOTICE

Data Loss is possible for any group from Nodes 1 and 2 that are brought online with Automatic Failover if outstanding writes were not mirrored to the secondary site.

CHAPTER 5
Manage SRDF/CE

This chapter shows how to manage an SRDF/Cluster Enabler installation.

The CE Manager

Use the Cluster Enabler (CE) Manager GUI (graphic user interface) to configure your Microsoft Failover Clusters for disaster recovery protection. The CE Manager allows you to set up and configure disk-based resources to automatically move geographically dispersed resource groups back and forth.

The CE Manager window (see Figure 17) contains a menu bar, two views, and a navigation tree. After cluster configuration, the navigation tree can be expanded to show four separate components: Groups, Storage, Sites, and Nodes.

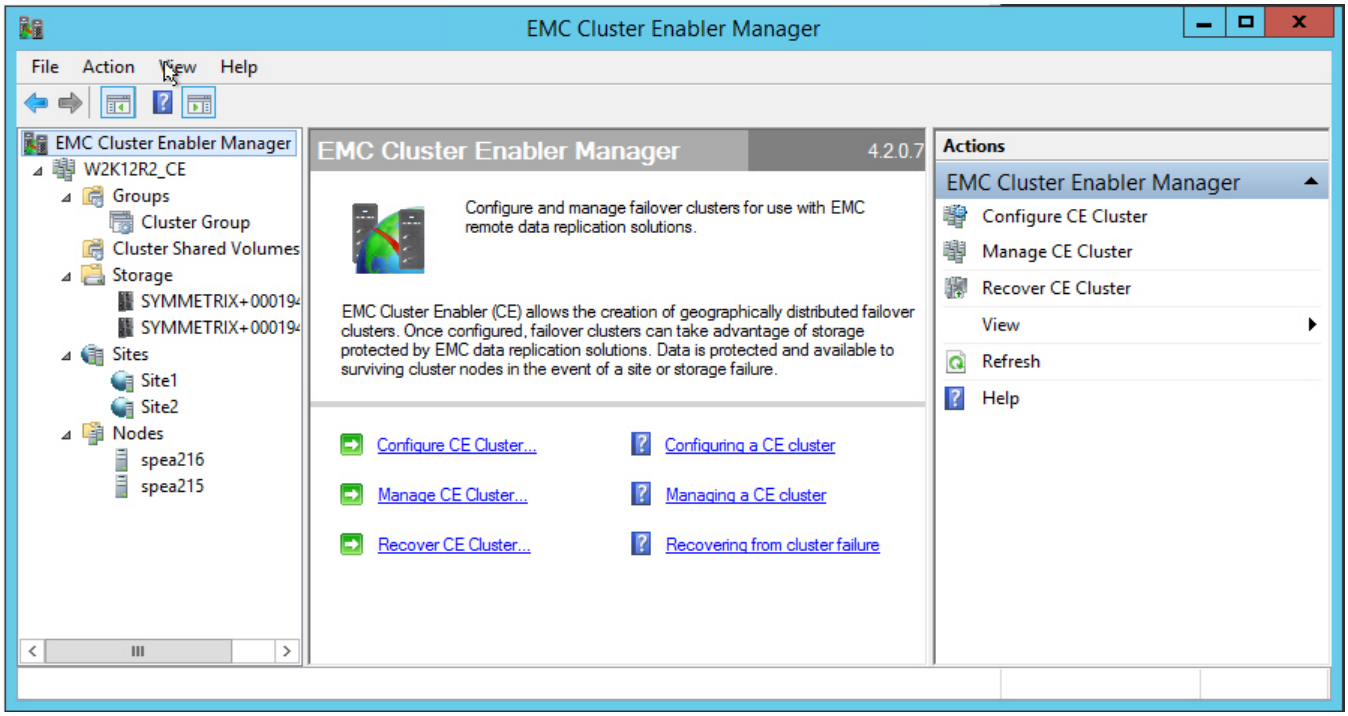

Figure 17 Cluster Enabler Manager window

Manage the configuration of a CE cluster

The configuration process is the first step towards managing disaster recovery for distributed failover clusters. The Configuration wizard configures your failover cluster for management by Cluster Enabler.

If any of the steps in the configuration process fail, the wizard displays a list of the specific errors for each node on a Summary page. Note each error and click **Finish** to exit the wizard. After fixing the reported problems, restart the configuration wizard to configure the CE cluster.

Note: Install the applicable Microsoft Failover Clustering on at least one node prior to configuring a cluster.

Configure a CE cluster

1. Select the **EMC Cluster Enabler** icon from the Navigation Tree and click the **Configure CE Cluster** link in the center pane.
2. On the Enter cluster name page (see Figure 18) enter the name of an existing cluster or cluster node in the space provided and click **Configure**. If you do not enter a name and click **Configure**, the wizard automatically detects the current clusters on the server and continues.

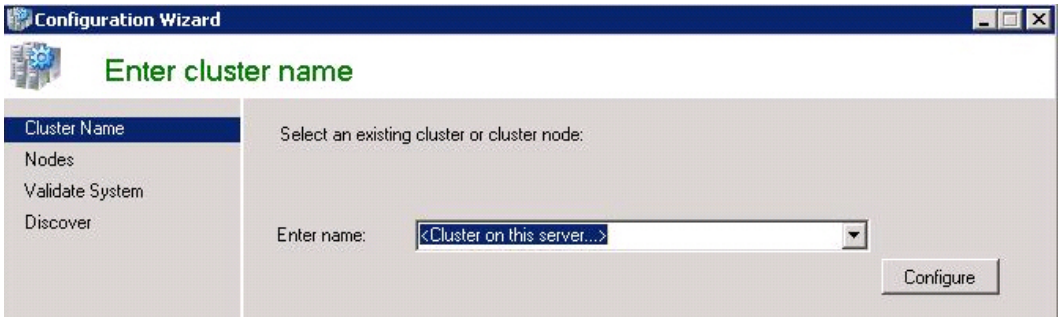

Figure 18 CE Manager Configuration Wizard

3. The Current Nodes page lists the current nodes in the cluster. To add a node, enter the node name and click **Add**.

   If you do not know the node name:

   a. Click **Browse** to browse an active directory of computers.

   b. Select a computer name from the list and click **OK**.

   c. Click **Add**.

4. Click **Next**.

   The Validating System Setup process begins.

   This process validates the system configuration by checking that the appropriate versions of Solution Enabler, Cluster Enabler, and Microsoft Failover Clustering are installed and configured.

5. When the Validation Complete notification appears, click **Next**.

   The Storage Discovery process begins.

   This process performs storage discovery for each cluster node to identify the locally-attached and remotely-attached storage.

6. When the Discover Completed notification appears, click **Next**.

   The Storage Setup process begins. This process performs storage setup for each cluster node.

7. When the Setup of Storage Configuration Completed notification appears, click **Next**.

   Note: In the case of clusters that use Node and Disk Majority or Disk Only quorum models, the storage setup process fails with a warning message when the cluster disk's topology type is not Point to Point and the replication mode is not Synchronous. If you are reconfiguring an existing CE cluster, CE manages the Cluster upon clicking the **Finish** button. Change the quorum configuration using the Change Quorum wizard and restart the Configuration wizard. If you are configuring the Cluster for the first time, change the quorum disk using the Microsoft Failover Manager and restart the CE Configuration wizard for CE to manage the cluster.

   The Validating Groups process begins. This process performs group validation for each converted failover cluster group.

8. When the Validated Groups notification appears, click **Next**.

9. When the Configuration wizard Completed Successfully notification appears, click **Finish**.

   After exiting the CE Configuration wizard, Cluster Enabler connects to the newly configured cluster. Once connected to the cluster, the configured cluster node appears in the navigation tree, located in the left pane.

10. Double-click the cluster icon to expand the cluster and view folders for Groups, Storage, Sites, and Nodes (see Figure 19). You are now ready to manage the cluster.

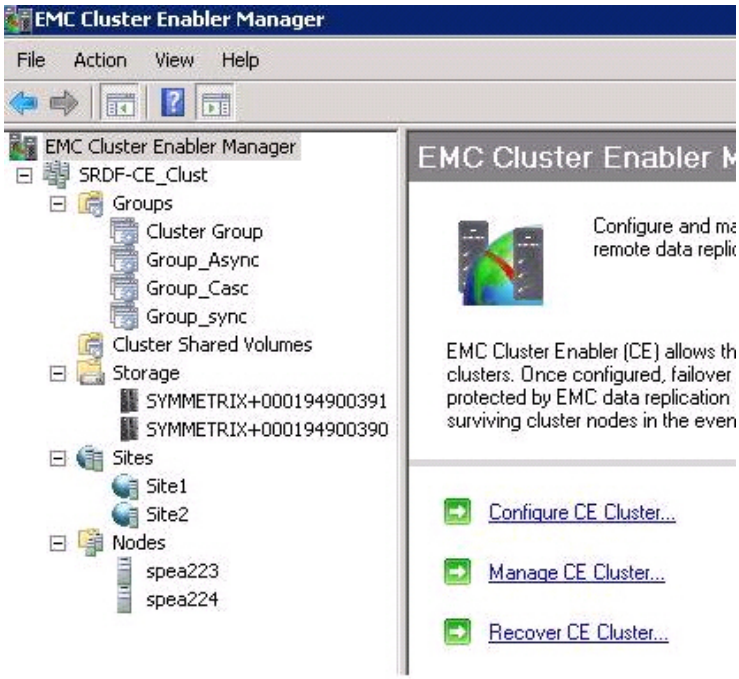

Figure 19 CE Manager expanded navigation tree

Add nodes to a CE cluster

You can also use the CE Configuration wizard to add nodes to a cluster. Always add new nodes using CE Manager and not Microsoft Failover Cluster.

Manage a CE cluster

Once you have configured the CE cluster use CE Manager to manage the cluster:

1. Select the **EMC Cluster Enabler** icon from the Navigation Console Tree and click the **Manage CE Cluster** link in the center pane.
2. Enter the cluster name in the selection box. If you do not enter a name, the default is to automatically connect to a cluster accessible on the server. Click **OK**.

   Once connected to the cluster, the configured cluster node appears in the navigation tree.
3. Double-click the cluster icon to expand the cluster and view the Groups, Storage, Sites and Nodes folders.

You can:

- Discover storage
- Update mirrored pairs
- Change the quorum model
- Manage CSV disks
- View cluster dependency

Note: If you manage a CE cluster from a management station that is not part of the cluster, ensure that the cluster nodes and on the management station run the same version of CE.

Discover storage

Use the Storage Discover wizard to discover and setup the attached storage. Carry out storage discovery after making any changes to the storage configuration.

1. Select the **Cluster** icon from the Navigation Console Tree and select **Action > Storage Discover**.

   The Storage Discovery page appears and the wizard searches for new storage devices.
2. When the Discover Completed notification appears, click **Next**.

   Note: If any storage discovery process fails, the wizard lists the discovery errors for each node on the Summary page. Note each error and click **Finish** to exit the wizard. Resolve the errors and restart the wizard.

   The Storage Setup page appears.
3. When the Set up of Storage Configuration Completed notification appears, click **Next**.

   The Summary page appears.
4. When the Discovered all Nodes notification appears, click **Finish**.

   Cluster Enabler refreshes the CE cluster to reflect any storage changes.

Update Mirrored Pairs

Use the Update Mirror Pairs wizard to update the mirrored pairs in a cluster. The wizard helps you to:

- Discover storage
- Update the storage configuration
- Validate the storage groups
- Set up the storage group definitions in the cluster properties database to update the mirrored pairs in a cluster

To update mirrored pairs:

1. Failover CE groups to an R1 node before modifying their R1-R2 relationship.
2. Shutdown the R2 (passive) Node.
3. From R1 node, modify the R1-R2 pairing with a new R2 device.
4. Establish the pairing and wait for the RDF state to be synchronized.
5. Bring up the R2 (passive) node.
6. Manage the cluster from CE manager on the R1 node.
7. Select the **Cluster** icon in the navigation tree and select **Action > More Actions... > Update Mirror Pairs**.

   The first page of the wizard appears and storage discovery process begins.
8. When the Discover Complete notification appears, click **Next**.

   The Storage setup process sets up the storage configuration.
9. When the Setup of Storage Configuration Completed notification appears, click **Next**.

   The Validating Groups process runs to validate each group in the cluster.
10. When the Validated Groups notification appears, click **Next**.

    The Updating Storage Mirror Relationships process runs to update the mirrored pairs in the groups.
11. When the Update Mirror Pairs for groups notification appears, click **Next**.
12. When Update Mirror Pairs Completed Successfully appears, click **Finish**.

Once the Update Mirror Pairs wizard completes successfully, the internal configuration database updates to reflect the new R1/R2 relationship. Once updated, groups can be failed over between the arrays with the new R2 pairs.

Change the quorum model

Use the Change Quorum Model wizard to change:

- The quorum model of the cluster
- The quorum disk
- The file share used in Node and File Share Majority clusters
- The cluster number

Once your Microsoft cluster is configured as a CE cluster, always use the Change Quorum wizard for all quorum model changes.

Figure 20 shows the first page of the Change Quorum wizard.

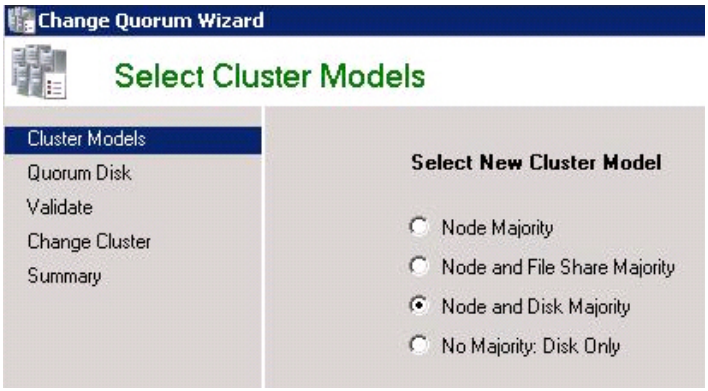

Figure 20 Quorum models

To change the quorum model:

1. Select the cluster icon in the navigation tree and then select **Action > More Actions... > Change Quorum Model**.

   The Select Cluster Models page appears (see Figure 20).

2. Select the quorum model that the cluster is to use.
3. The next step depends on the quorum model you chose:

Table 7 Defining the disk or file share for the quorum

Quorum model	Action
Node and File Share Majority	Select the file share that you want to use and click **Next**.
Node and Disk Majority	Select the quorum disk that you want to use and Click **Next**. Ensure that the disk you select has the following characteristics: • Point-to-Point topology • Synchronous replication mode
Disk Majority	

   The Select Cluster Number page appears.

4. Select a number from the **Select a Cluster Number** list box and click **Next**.

   The list box contains only unused cluster numbers. To view a list of all numbers, including those in use, click **Show All Cluster Numbers**. However, you must select a number that is not in use.

   The Validate Cluster Model process runs.

5. When the Validation of Cluster Model Successfully notification appears, click **Next**.

   The Change Cluster Model process runs to apply the new settings.

6. When Change Cluster Model Successfully appears, click **Next**.
7. When the Changed Cluster Model Successfully notification appears, click **Finish**.

Manage CSV disks

Convert CSV disks for CE

Before you can manage CSV disks with CE Manager, convert the CSV disks using the CE Configuration wizard. Use the CE Configuration wizard (see "Configure a CE cluster" on page 65) to configure CSV as you would a CE cluster. Move all VMs to the CSV primary site before configuration or they fail over automatically.

Note: If I/O is attempted on a cluster node containing an R2 CSV disk, the node (and only that node) transitions to redirected access. The node returns to direct access only when the mirror is promoted or swapped to a primary mirror.

Note: When CSV are configured for CE, there are no disk resources listed under the new virtual machine. Disk resources are listed under Cluster Shared Volumes.

Management operations

Once CSV disks are converted, use the CE Manager to manage them. The CSV Folder view displays the set of VMs residing on each CSV disk. The CSV disk and the VM details appear in a tree view (see Figure 21). The parent node contains all of the CSV-related data (such as CSV Path, Owning Node, Device group name, Sync State). VM details are grouped under the appropriate CSV Parent Node, on which the VM resides.

This representation allows you to see the set of VMs hosted on each CSV disk, and whether, the CSV disk is configured using CE or not. You can expand the tree view by selecting **Expand All** or collapse it by selecting **Collapse All**, which is useful if there many CSV disks to manage.

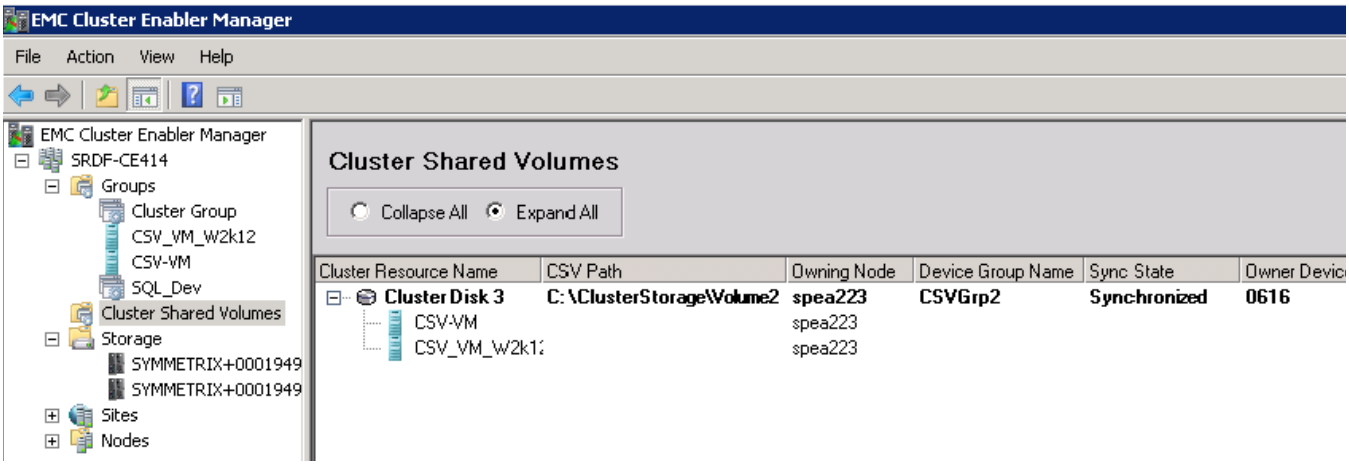

Figure 21 Cluster Shared Volumes tree view

You can:

- Change the failover policy for a CSV disk
- Remove a CSV from CE Manager control

Change the failover policy

To change the failover policy:

1. Right click a CSV in the navigation tree and select the **Properties**.
2. On the Policies tab, select the required failover behavior:

Table 8 CSV failover policies

Failover policy	Description
Restrict Group Movement	The CSV disk cannot fail over to a peer node. In a replication link failure, this setting only attempts to move disk laterally. If the replications link is up, this setting has no impact.
Automatic Failover	The CSV disk can automatically failover to any remote site node in the event of a replication link failure.

3. Click **OK**.

Note: CSV can use SRDF/Synchronous mode only. So, the SRDF/Asynchronous option is unavailable in the **Advanced** tab setting of the selected CSV disk's properties. The failover behavior for a CSV disk is currently limited to the Restrict Group Movement policy only and is the default selection.

Remove a CSV from the control of CE Manager

You can remove a CSV from the control of the CE Manager. In turn this removes failover support to remote nodes, and you cannot use CE Manager to control the virtual machines that are dependent on the CSV.

To remove a CSV from the control of CE Manager:

1. Right-click on a cluster disk and select **Deconfigure CSV From CE**.

   CE Manager asks you to confirm the action.

2. Click **Yes** to remove the disk or **No** to retain it.

View cluster dependency

1. Open the CE Manager, select the cluster in the Navigation Tree, and do one of the following:

   - Select **Action** > **View Dependency**.
   - Right click on the selected cluster and select **View Dependency** from the pop-up menu.

   There may be a pause before the diagram appears while CE gathers site information.

2. Select each cluster group, and double-click the disk objects to expand the view for each site. Devices use a color code by site that is defined on the right-side of the display.

   You can use the **Expand All** or **Collapse All** buttons to expand or collapse group details.

3. Click **Sort Group by Site** to change the Site view.

4. To preview a printable diagram, click **Print Preview**.

   To print the diagram click **Print**.

Figure 22 shows a sample Dependency Report.

The CE Dependency Viewer provides both the CSV disk group name and the path for a configured CSV disk. If the disk is configured, the group name displays as:

```
CSVGrp1 [C:\ClusterStorage\Volume1].
```

Otherwise, only the CSV Path appears as:

```
[C:\ClusterStorage\Volume1].
```

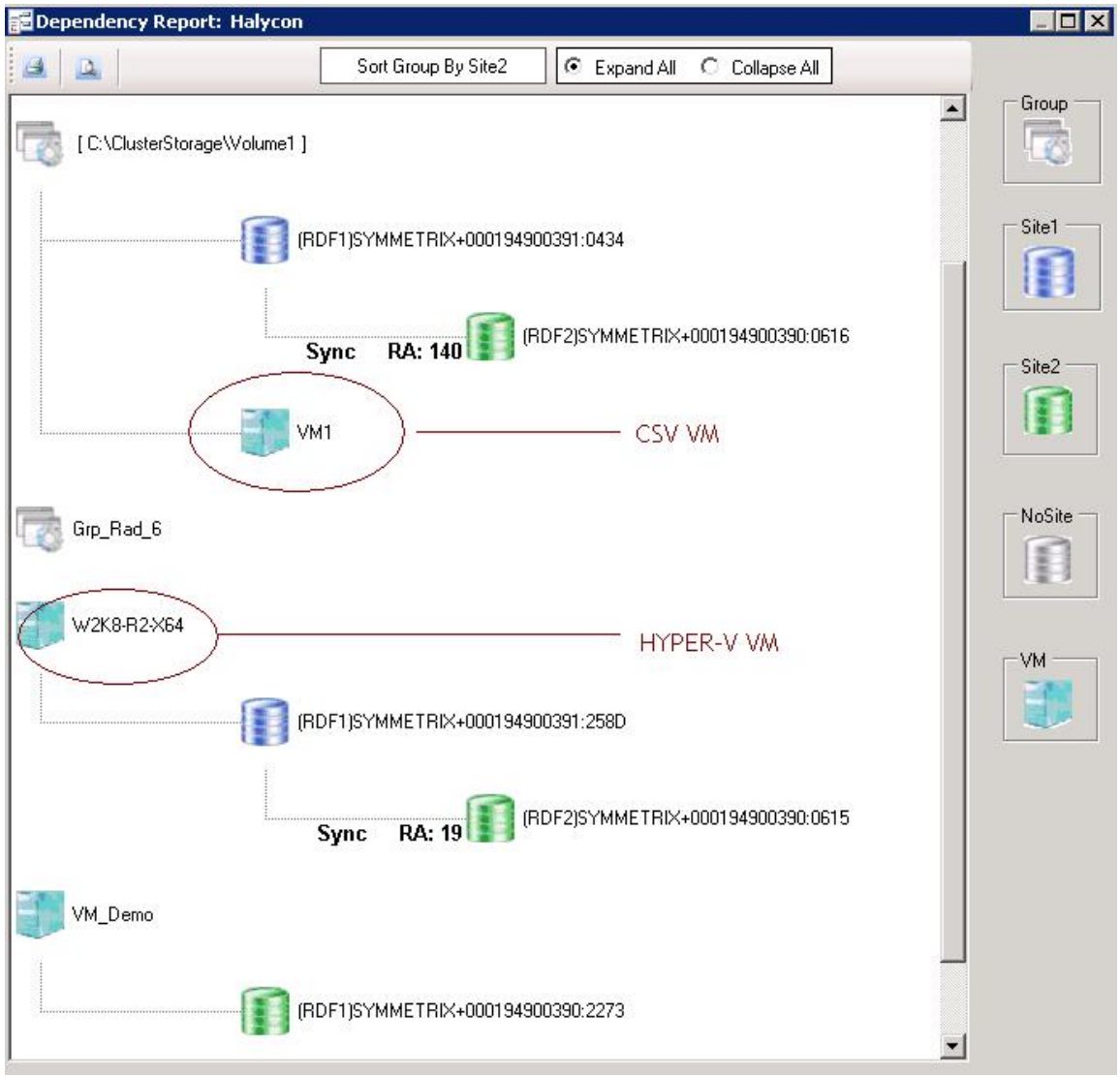

Figure 22 Sample Dependency Report

The following notations apply to Figure 22:

CSV VM — VM resides on the CSV disk.
HYPER-V VM — VM resides on regular RDF disk.

Manage a CE group

For CE cluster groups you can:

- Create a CE group
- Modify a CE group
- Dismantle a CE group and convert it back to a regular cluster group
- Delete a CE group

Create a CE group

1. Click the **Groups** icon from the Navigation Tree and select **Action > Create Group** from the menu bar.

   Cluster Enabler reads the storage configuration and displays the first page of the Create Group wizard appears.

   Note: A mirrored pair needs to be present on the array before you can create a group. Run the Storage Discover wizard to detect a newly created mirrored pair by right- clicking on the cluster name or clicking the **Discover** button in the Select Devices page of the Create Group wizard.

   ---

2. Type a unique name for the group in **Group Name** and click **Create**.
3. On the next page (see Figure 23) select devices for inclusion in the new group. Select the check box for each device and click **Next**.

   The Select Group Policy page appears.

   Notes:

   - By default, all available configured VMAX storage appears in collapsed view. Click **Expand All** to expand the tree view.
   - Use the checkboxes to select the types of device to display: **Async**, **Cascaded**, and **Concurrent**. For example, selecting the **Async** checkbox displays all SRDF asynchronous capable devices within in the same RA group, mapped to the nodes.
   - An error message appears if the selected type of devices are used up or not available.
   - Selecting devices from a single RA group creates a device group. Selecting devices from multiple VMAX RA groups creates a composite group.
   - You cannot create a composite group on VMAX 10K or VMAXe arrays.
   - Any devices other than the replication mode synchronous or asynchronous are not listed as available for device selection.

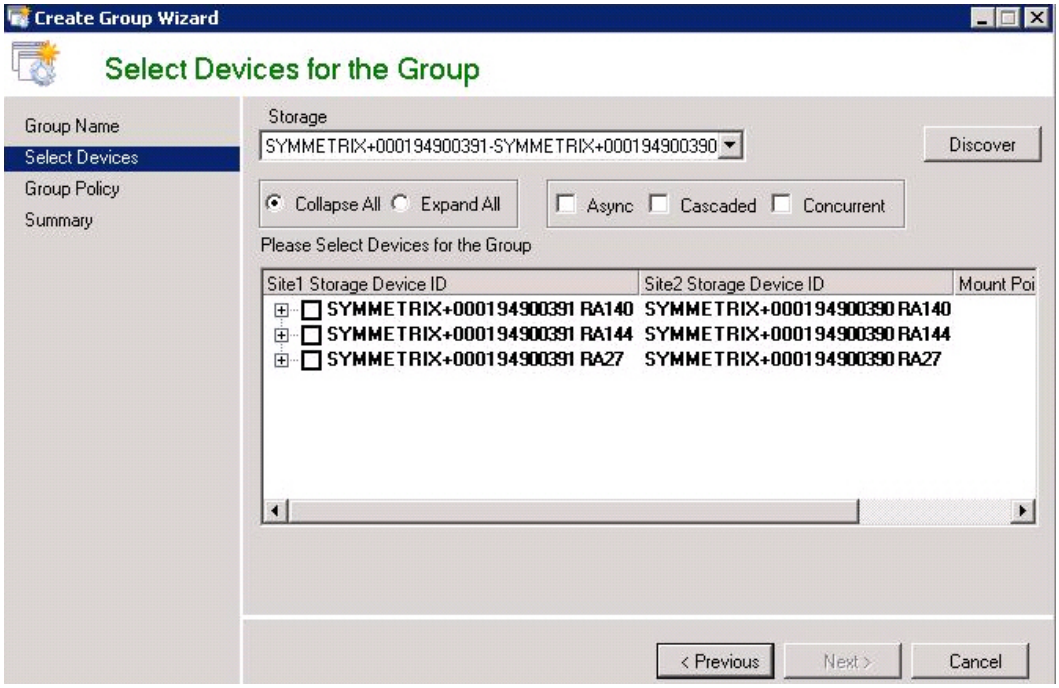

Figure 23 Create a CE Group: Select Devices for the Group

4. Select a policy for the group from the list box (see Figure 24) and click **Next**.

   The available policies are:

Table 9 Cluster group failover policies

Failover policy	Description
Restrict Group Movement	The group cannot fail over to a peer node. In a replication link failure, this setting only attempts to move disk laterally. If the replications link is up, this setting has no impact.
Automatic Failover	The group can automatically failover to any remote site node in the event of a replication link failure.

5. When the Group Created Successfully notification appears, click **Finish**.

Cluster Enabler refreshes the CE Cluster. Once complete, the group you created appears under Groups. If you do not see the newly created group, select **Action > Refresh**.

Note: You can use the **Asynchronous** check box to convert a synchronous capable device to an asynchronous capable device (or the reverse). If you do this and the appropriate SRDF/Asynchronous or SRDF/Synchronous license does not exist, group creation fails with one of the following messages:

The following Array Licenses are required on Symmetrix: Sync
The following Array Licenses are required on Symmetrix: Async
The following Solutions Enabler Licenses are required on node:
< nodename>

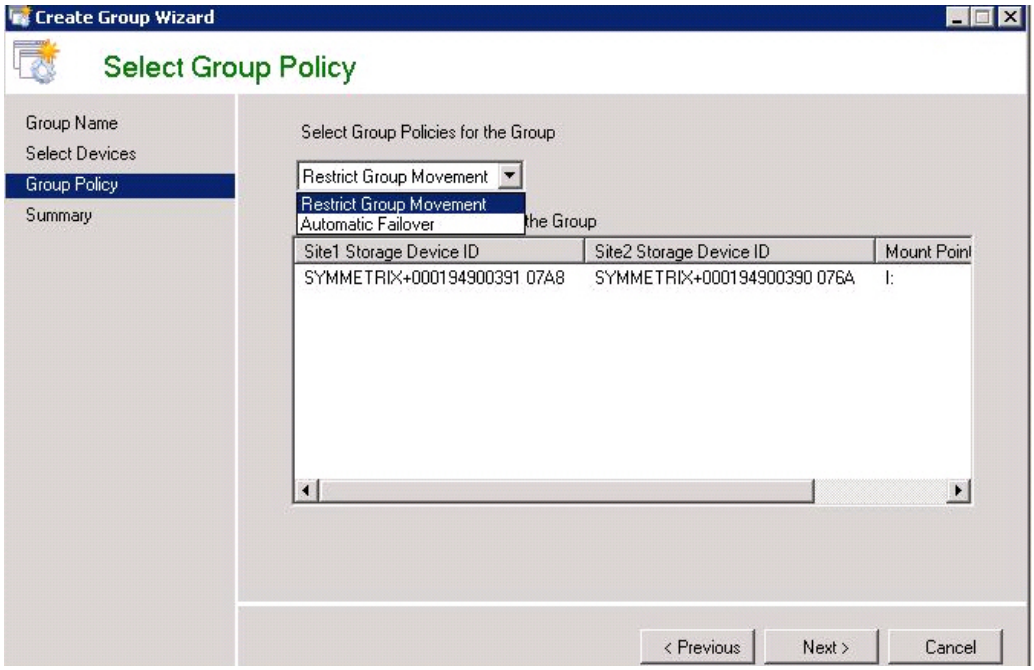

Figure 24 Create a CE Group: Select Group Policy

Modify a CE group

1. Click the **Groups** icon in the navigation tree and select **Action > Modify Group**.

   The Storage Synchronization process runs and then the first page of the Modify Group wizard appears.

   Note: A mirrored pair needs to be present on the array before you can modify a group. Run the Storage Discover wizard to detect a newly created mirrored pair by right-clicking on the **cluster name** or clicking the **Discover** button in the Select Devices page of the Modify Group wizard.

   ---

2. From the Select Devices page (see Figure 25), select **Add Devices** or **Delete Devices** in the **Select Action** list box.

   A list of available devices that you can add or remove appears. RA group pairs and the devices contained within the RA group appear in a tree view. Initially, the RA Groups are collapsed. Select **Expand All** to expand the tree view to see individual devices within each group.

3. Select the check boxes for the devices you want to add or delete and click **Next**.

   Selecting an RA group, selects all devices in that group.

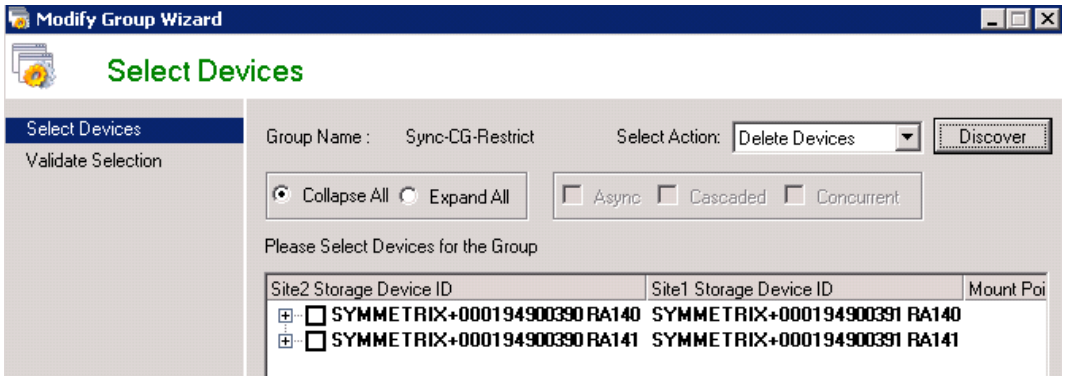

Figure 25 Modify a CE Group: Select Devices

4. When the Validate Selection page appears, click **Next**.
5. When the Group Modified Successfully notification appears, click **Finish**.

Cluster Enabler refreshes the CE cluster. Upon completion of the refresh, you see the updated group information for the devices that you added or deleted. If you do not see that information, select **Action > Refresh**.

Note: You can use the **Asynchronous** check box to convert a synchronous capable device to an asynchronous capable device (or the reverse). If you do this and the appropriate SRDF/Asynchronous or SRDF/Synchronous license does not exist, group creation fails with one of the following messages:

The following Array Licenses are required on Symmetrix: Sync
The following Array Licenses are required on Symmetrix: Async
The following Solutions Enabler Licenses are required on node:
< nodename>

Notes:

- Selecting **Add Devices** displays ungrouped devices from all RA group pairs.
- Selecting **Delete Devices** for a composite group displays all devices under each RA group pair in this group.
- If deleting devices in a composite group, and you delete all devices from within one RA group, the modified group automatically converts to a device group and RDF consistency is disabled.
- If the group is a device group, and you add devices from another RA group, the group automatically converts to a composite group and RDF consistency is enabled. A warning message displays in this case.
- Any devices other than the replication mode synchronous or asynchronous are not listed for device selection.
- You cannot create a composite group on the VMAX 10K or VMAXe arrays.

Configure a CE Group

1. Select a group listed under the Groups icon and select **Action > Configure CE Group** (see Figure 26).

   The Configure CE Group action is available for non-converted Microsoft Failover Cluster groups and unavailable for all existing CE Groups.

   A dialog box appears asking you to confirm the action.

2. Click **Yes** to configure the Microsoft Failover Cluster group to a CE group.

If the group is configured as a CE Group, CE failover support to the remote nodes is added. The CE resource is added, but the Microsoft Failover Cluster physical disk resources are not changed. Converting a Microsoft Failover cluster group to a CE group enables the Delete, Modify Group, and Deconfigure Group CE operations.

When configuring a VM Group through CE, if the underlying CSV disk group is not configured, it is configured first before the VM Group is configured. You can also configure the CSV disk group by itself through CE.

Note: The Configure CE Group Option fails with the following error message if the Microsoft Failover Cluster Group does not have a disk in the group:

```
The Group <group_name> does not have the Clustered Disk
```

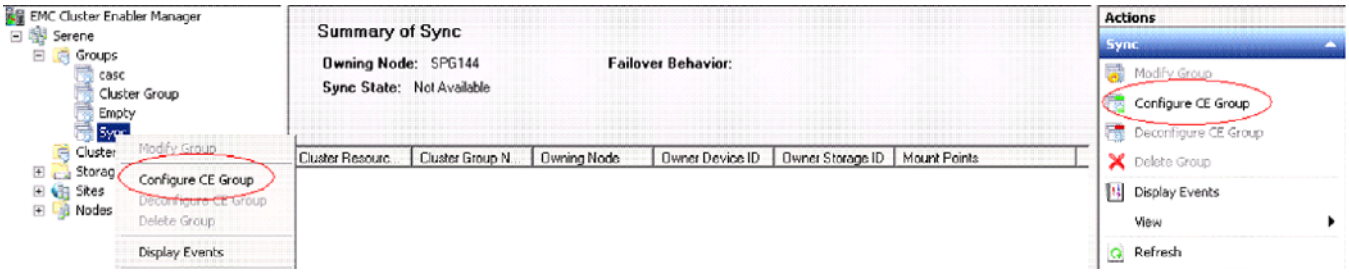

Figure 26 Configure CE Group option

Deconfigure a CE group

1. Select a group listed under the **Groups** icon located in the navigation tree and select **Action > Deconfigure CE Group**.

   A dialog appears asking you to confirm the action.

2. Click **Yes** to deconfigure the group.

You can deconfigure a VM group only if you deconfigure its underlying CSV disk group from Cluster Enabler. Deconfiguring the CSV disk group deconfigures all dependent VMs.

Note: If the group is deconfigured, CE failover support to the remote nodes is no longer operational. To make group failover operational again, reconfigure the cluster group using the CE Configuration Wizard in the CE Manager.

Delete a CE group

1. Select a group listed under the **Groups** icon in the navigation tree and select **Action > Delete Group**.

   A dialog appears asking you to confirm the action.

2. Click **Yes** to delete the group.

Note: Deleting a CE group deconfigures the group and removes it from the cluster.

Manage storage

Cluster Enabler provides the following means of managing storage:

- View device information
- Add device to and remove them from a group

View device information

Use the **Storage** icon in the navigation tree to view device information for the storage arrays. Click the icon to see a list of arrays. Then click on the name of an array to view summary information about it (see Figure 27).

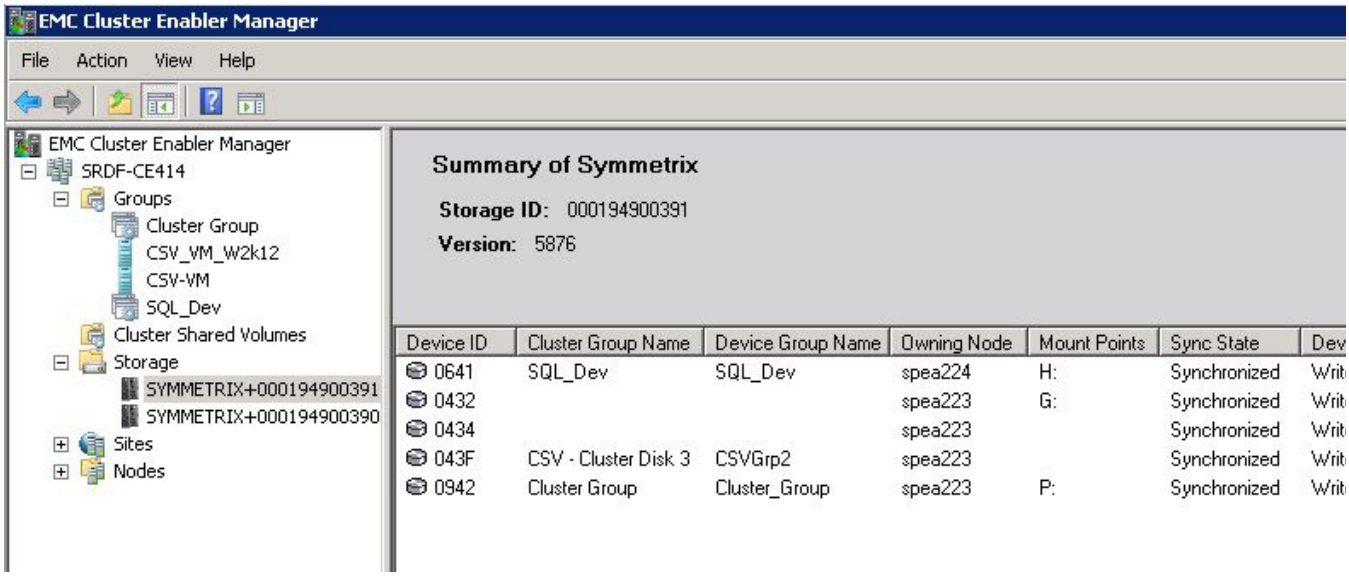

Figure 27 Example of Symmetrix storage array view

The display shows the name (Storage ID) of the array and the version of PowerMaxOS, HYPERMAX OS, or Enginuity. Beneath that is a table of information about the devices in the array. Table 10 defines the information shown for each device.

Table 10 Storage information

Column	Description
Device ID	All SRDF R1/R2 device IDs that are mapped to any cluster member node.
Cluster Group Name	The name of the CE Group that device belongs to.
Device Group Name	The name of the SYMAPI device group or composite group that the device belongs to. This name is derived from name of the Cluster Group.
Owning Node	If a device belongs to a cluster group, the owning node information is obtained directly from Microsoft Failover Cluster. Otherwise, the owning node is the name of a node where the device is write-enabled.
Mount Points	The mount point of the physical drive on the owning node.
Sync State	The RDF state for the group. The *Dell EMC Solutions Enabler SRDF Family CLI User Guide* provides a listing of all possible RDF states.

Table 10 Storage information (continued)

Column	Description
Device Status	The SRDF R1/R2 device status. The possible device status states are Ready, Not Ready, and Write-Disabled.
Capacity MB	The device capacity in megabytes.
Swap Capable	Whether the device is swap capable.
Async Capable	Whether the device is asynchronous capable.
WWN	The device's World Wide Name (WWN).
Logical Device	The logical device name (if applicable).
RDF Type	The RDF device type: R1 or R2.
RA Group	The name of the RA group that the device belongs to.
R1 Invalid Tracks	The number of invalid R1 tracks (if any).
R2 Invalid Tracks	The number of invalid R2 tracks (if any).
RDF Async Lag Time	The lag time between the target (R2) device and the source (R1) device in an SRDF/Asynchronous environment.
Invista WWN	The Invista device's World Wide Name (WWN).

Add and remove devices

While the array device information is available, you can add a device to or remove one from a group:

1. Select the device in the list and click the **Add to Group** or **Remove from Group** icon in the Action panel (see Figure 28).

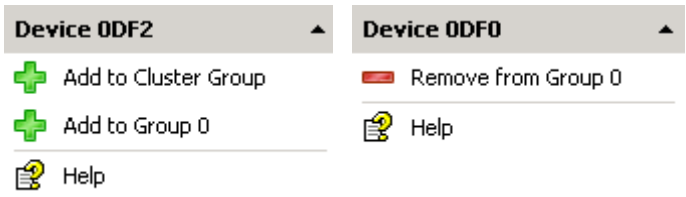

Figure 28 CE Manager storage actions

   Cluster Enabler opens the Modify Group wizard at the validation step.

2. Click **Next** to add or remove your selection.
3. When the Group Modified Successfully notification appears, click **Finish**.

View information

Cluster Enabler provides summary information about CE Groups, Storage, Sites and Nodes. This information appears when you select an icon in the navigation tree.

Group information

1. Click the **Groups** icon in the navigation tree.

   Summary information on all the groups appears in the center pane (see Figure 29).

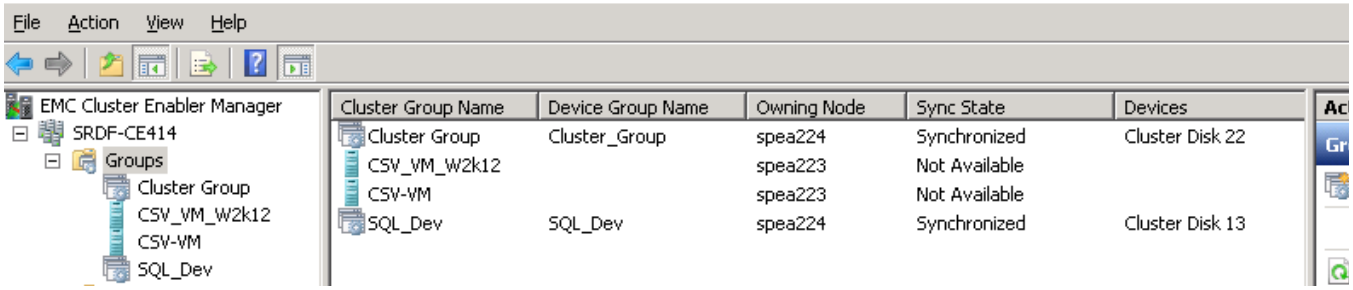

Figure 29 The Groups component

2. Double click a group name to display the group details (see Figure 30) .

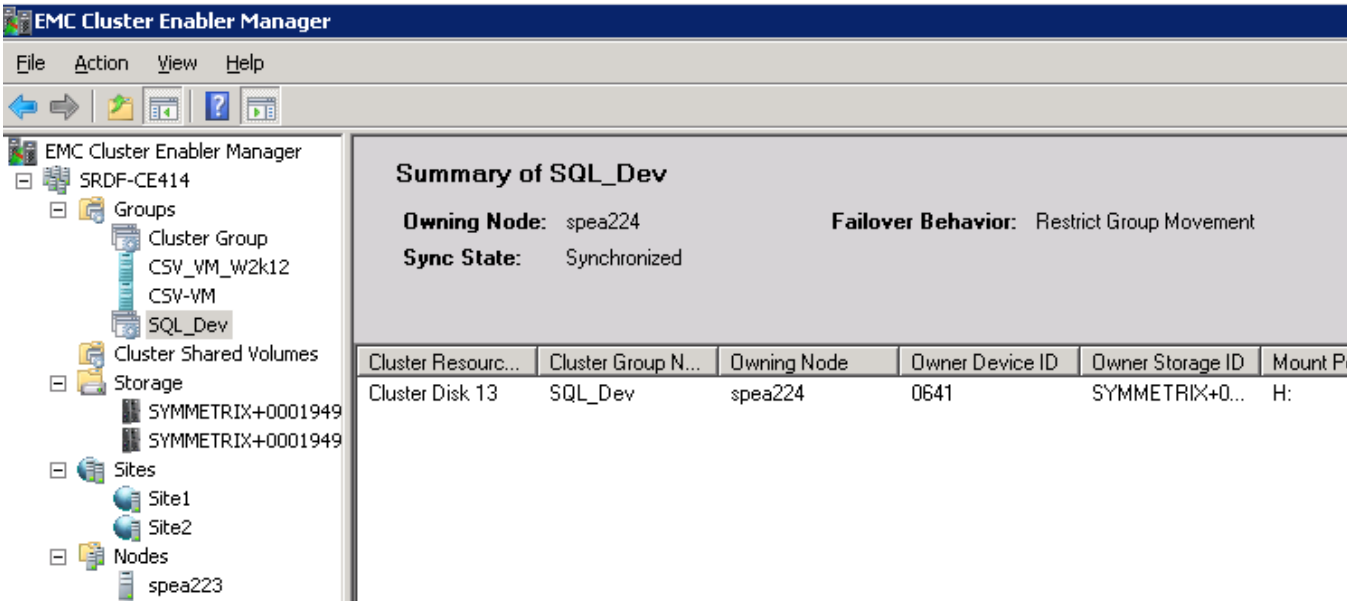

Figure 30 Group information

Table 11 defines the information shown in the summary and detailed displays.

Table 11 Group information

Column	Description
Cluster Group Name	The CE Group name to which the device belongs.
Device Group Name/ Consistency Group Name	The SYMAPI device group or composite group name to which the Symmetrix device belongs; derived from Cluster Group name.
Owning Node	The name of the failover cluster node that owns the particular group. This information is obtained directly from MS Failover Cluster. Only groups that are part of the cluster display.
Sync State	The RDF state for the group. The *Dell EMC Solutions Enabler SRDF Family CLI User Guide* lists the RDF states,
Devices	Listed by disk resource name in the cluster.
Cluster Resource Name	Listed by physical disk resource name.
Owner Device ID	The storage array's device ID mapped to the owning node (such as, ODEC, ODED).
Owner Storage ID	The storage array ID (such as, Symmetrix+00187900830).
Mount Points	The mount point of the physical drive on the owning node.

Click the **Display Events** icon in the Action pane to display event information in the lower tier of the center pane. Table 12 defines the information shown in the display.

Table 12 Group event information

Column	Description
Date/Time	The date and time that the recorded event occurred.
Computer Name	The computer name on which the event occurred.
Group Name	The group name to which the event occurred.
Message	A detailed message of the event type.

If the group is a virtual machine, the icon for the group changes and CSV-related information about the VM appears (see Figure 31). Table 13 defines the information shown in the display.

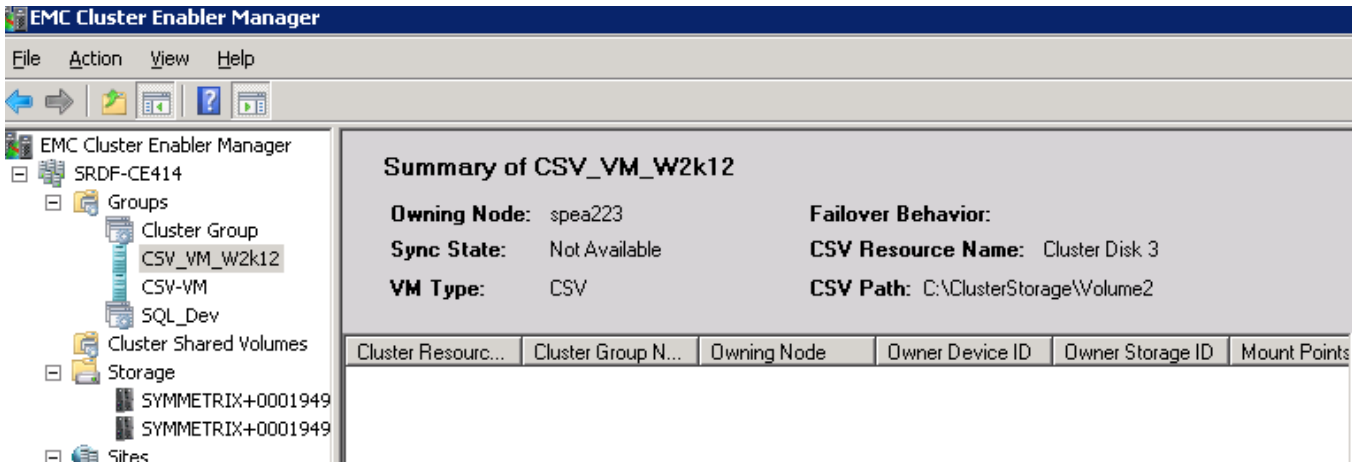

Figure 31 VM group information

Table 13 CSV Group information

Column	Description
Cluster Recource Name	Listed by physical disk resource name.
CSV Path	The Windows path where the CSV disk is mounted. Usually shown as `C:\ClusterStorage\Volume`.
Owning Node	The name of the failover cluster node that owns the particular group. This information is obtained directly from MS Failover Cluster. Only groups that are part of the cluster display.
Device Group	The name of the SYMAPI device group that the device belongs to; derived from Cluster Group name.
Sync State	The RDF state for the group. The *Dell EMC Solutions Enabler SRDF Family CLI User Guide* lists the RDF states.

Node information

1. Click the **Nodes** icon in the navigation tree.

   Summary information on all nodes appears in the center pane (see Figure 32).

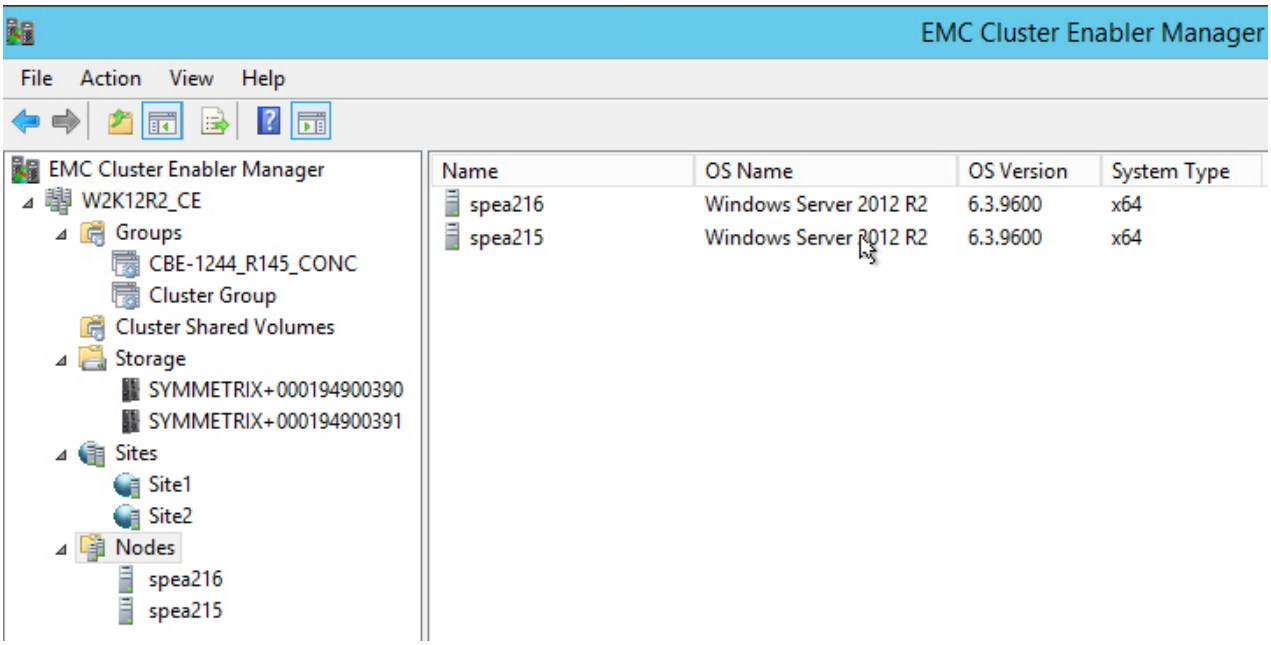

Figure 32 The Nodes component

2. Double click a node name to display the node details (see Figure 33).

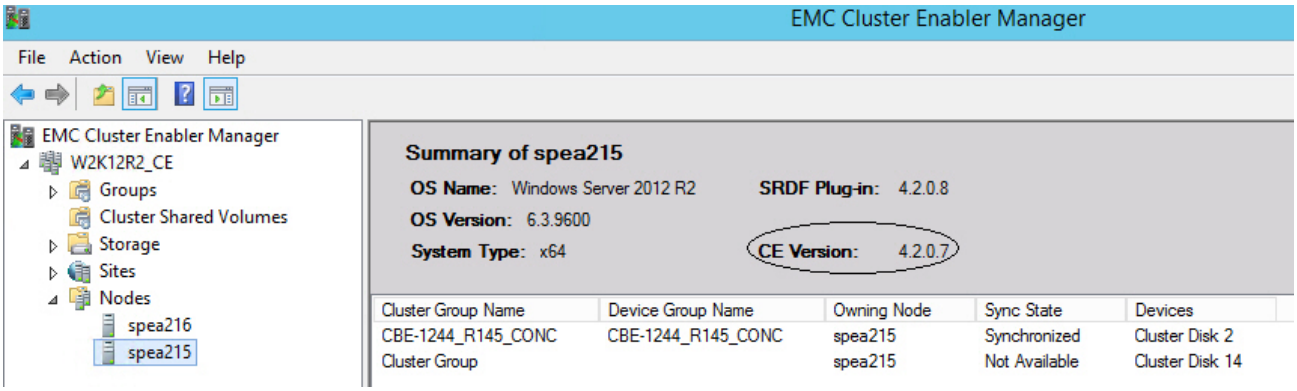

Figure 33 Node information

Table 14 lists the information in the summary and detailed displays.

Table 14 Node information

Column	Description
Name	The node name.
OS Name	The Windows operating system (such as, 2012 SP2).
OS Version	The Windows operating system version (such as, 6.3.9600).
System Type	The Windows system type (such as, x64).
CE Plug-in	The CE Plug-in version (such as 4.2.1.12).
CE Version	The CE Base version installed on the selected node.
Cluster Group Name	The name of the CE Group to which the device belongs.
Device Group Name	The name of the SYMAPI device group or composite group that the device belongs to; derived from Cluster Group name.
Owning Node	The name of the failover cluster node that owns the particular group. This information is obtained directly from MS Failover Cluster. Only groups that are part of the cluster can appear in this column.
Sync State	The RDF state for the group. The *Dell EMC Solutions Enabler SRDF Family CLI User Guide* provides a listing of all possible RDF states,
Devices	Listed by cluster resource name.

Click the **Display Events** icon in the Action pane to display event information in the lower tier of the center pane. Table 15 defines the information shown for each event.

Table 15 Node event information

Column	Description
Date/Time	The date and time that the recorded event occurred.
Computer Name	The computer name on which the event occurred.
Group Name	The group name to in which the event occurred.
Message	A detailed message of the event type.

Site information

1. Click the **Sites** icon in the navigation tree.

   Summary information on all sites appears in the center pane (see Figure 34).

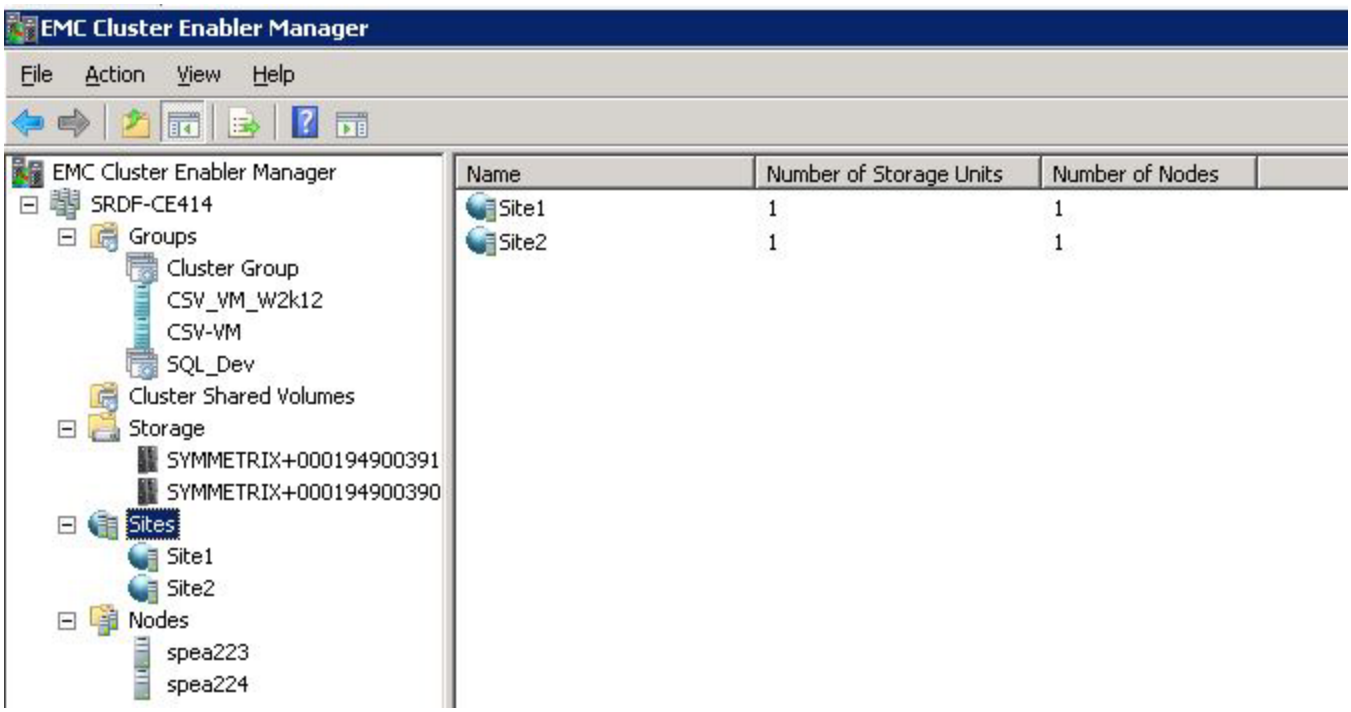

Figure 34 The Sites component

2. Double click a site name to display the site details (see Figure 35).

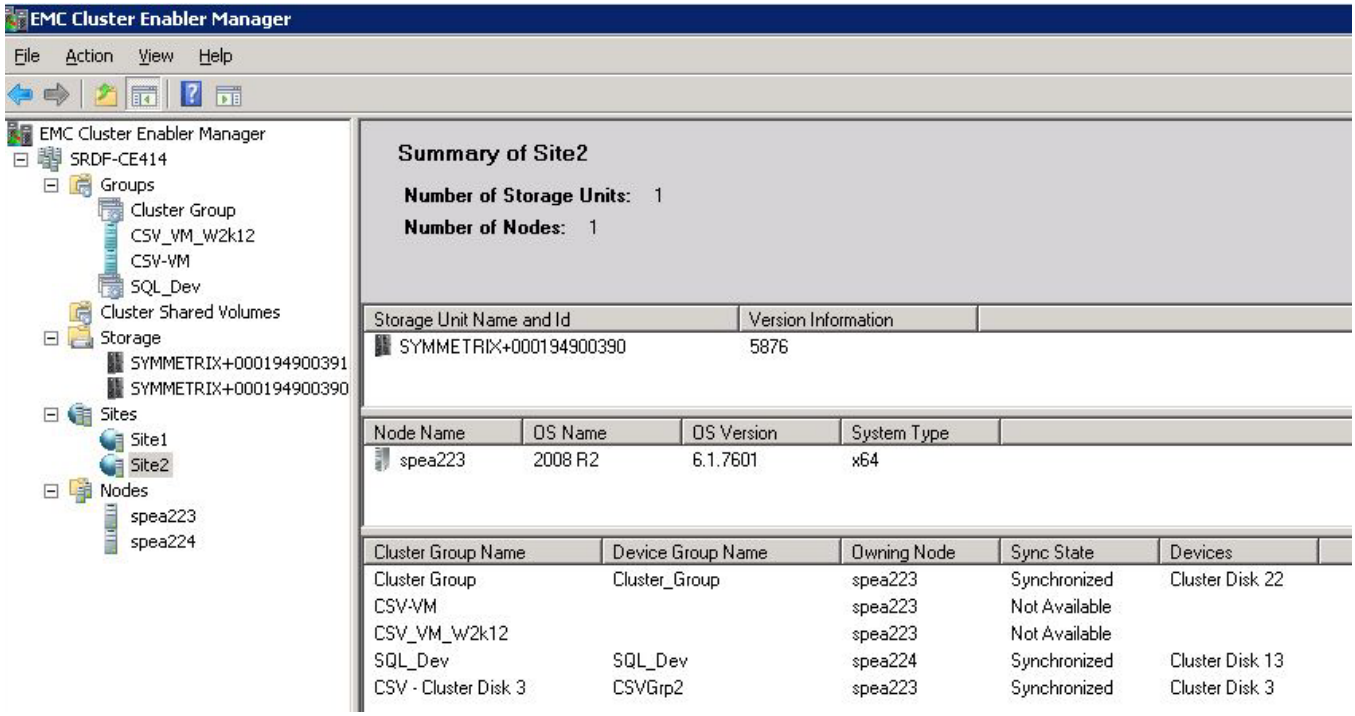

Figure 35 Site information

Table 16 defines the information in the summary and detailed displays.

Table 16 Site component information

Column	Description
Name	The site name.
Number of Storage Units	The number of storage units for this site.
Number of Nodes	The number of nodes for this site.
Storage Unit Name & ID	The storage array ID (such as, Symmetrix+00187900830).
Version Information	The version of PowerMaxOS, HYPERMAX OS, or Enginuity that the array is running.
Node Name	The node name.
OS Name	The Windows operating system (such as, 2012 SP2).
OS Version	The Windows operating system version (such as, 6.2.3790).
System Type	The Windows system type (such as, X86).
Cluster Group Name	The name of the CE Group that the device belongs to.
Device Group Name	The name of the SYMAPI device group or composite group that the device belongs to; derived from Cluster Group name.
Owning Node	The name of the failover cluster node that owns the particular group. This information is obtained directly from Microsoft Failover Clustering. Only groups that are part of Microsoft Failover Clustering display.
Sync State	The RDF state for the group. The *Dell EMC Solutions Enabler SRDF Family CLI User Guide* lists the RDF states.
Devices	The devices by cluster resource name.

Note: To change the name of a site right-click its name and select **Rename** from the context menu. The site name becomes editable. Alternatively, right-click the site name and select **Properties** from the context menu. You can then change the site name in the **Properties** tab.

Manage CE Logging

You can modify the following characteristics of the logging facility:

- Log level
- Location of the log file
- Retention period for log files
- Maximum size of a log file

In addition, you can extract the current log to a dump file.

Set the log level

1. Open a command prompt and stop the ce_eventrace service:

   ```
   net stop ce_eventrace
   ```

2. Start the registry editor:

   ```
   regedit
   ```

3. Edit the value of the following key:

   ```
   HKEY_LOCAL_MACHINE\SOFTWARE\EMC\CE\Configuration\
   EventTraceLevel
   ```

   The values that the key can have are:

   - `4`: the log file contains error, warning, and informational messages.
   - `5`: the log file contains error, warning, informational, and verbose messages.

     This level greatly increases the amount of data that is sent to the log file.

4. Exit from the registry editor and restart the ce_eventrace service:

   ```
   net start cd_eventrace
   ```

Set the location of the log file

1. Ensure that the directory you want to use for log files exists and has write permission.
2. Open a command prompt and stop the ce_eventrace service:

   ```
   net stop ce_eventrace
   ```

3. Start the registry editor:

   ```
   regedit
   ```

4. Set the value of the following key to the path of the directory to hold log files:

   ```
   HKEY_LOCAL_MACHINE\SOFTWARE\EMC\CE\Configuration\EventTraceDirectory
   ```

   Terminate the path with a trailing backslash (\). In addition, ensure that the directory exists before you make this change.

5. Exit from the registry editor and restart the ce_eventrace service by typing:

   ```
   net start ce_eventtrace
   ```

Set the retention period for log files

1. Open a command prompt and start the registry editor:

   ```
   regedit
   ```

2. Set the value of the following key to the number of log files that you want to keep at any one time:

   ```
   HKEY_LOCAL_MACHINE\SOFTWARE\EMC\CE\Configuration\ EventTraceLogRetention
   ```

Set the maximum size of a log file

1. Open a command prompt and start the registry editor:

   ```
   regedit
   ```

2. Set the value of the following key to the maximum size of a log file (in MB):

   ```
   HKEY_LOCAL_MACHINE\SOFTWARE\EMC\CE\Configuration\EventTraceFileMaxSize
   ```

Extract the current log to a dump file

Open a command prompt and type:

```
CE_EventTraceDump.exe
```

This command creates a dump file. The format of the file name is:

ce_event_trace_*yymmddhhmmss*.txt

Here, *yy* is the last two digits of the current year, *mm* is the number of the month, *dd* is the day of the month and *hhmmss* is the time of day.

To specify the name of the dump file, use the -o option:

CE_EventTraceDump.exe -o *filename*

Replace *filename* with the path for the output file.

For example, to store the output from the command in `C:\MyLogFiles\ce_event_trace.txt` use:

```
CE_EventTraceDump.exe -o C:\MyLogFiles\ce_event_trace.txt
```

Control Delay Failback

You can view the current setting of Delay Failback and you can change that value.

View

1. Open a command prompt and type:

   ```
   cluster /priv
   ```

2. Locate the entry **DelayFailbackEnabled**.

   A value of 1 means that the feature is on and a value of 0 means the feature is off.

   If the property does not appear in the output, Delay Failback is on.

Change

1. Open a command prompt.
2. Type:

   ```
   cluster /priv DelayFailbackEnabled=n:DWORD
   ```

   Replace *n* with `1` to turn Delay Failback on or `0` to turn it off.

Restore and recovery operations

This section sets out some of the restore and recovery operations for different types of failures.

- Restore a failed SRDF site
- Recover a SRDF backup site in case of primary site failures
- Recover from an SRDF link failure
- Restrict group movement and recovery
- Recover from a corrupt quorum log
- Replace a storage array
- The Recover CE Cluster Wizard

Note: Some of these procedures use *CEInstallDir* to refer to the Cluster Enabler installation directory (by default, `C:\Program Files\EMC\ClusterEnabler`). When entering this path always use a capital letter for the drive letter.

Restore a failed SRDF site

To restore your storage system after a site failure occurs with all links lost:

1. Reconstruct the cluster.
2. Restore groups that failed over.

Reconstruct the cluster

1. Restore SRDF and IP links.
2. Restart all nodes.
3. Open CE Manager and connect to the cluster.
4. Run the Storage Discover Wizard.

   Any groups that are failed over to a secondary site are in a split state. Groups that are not failed over are in suspended state. You can safely bring the groups that did not failover to a secondary site online at this point.

Restore groups that failed over

The following instructions show how to restore groups that failed over to a secondary site.

IMPORTANT

Choosing the wrong option for restore could cause data loss. Contact EMC support if you have any question about the commands to use.

When the secondary site has good data that you want to copy to the primary site:

1. Open the command line prompt.
2. Navigate to the Cluster Enabler installation directory.

3. Set the `SYMCLI_DB_FILE` environment variable to *`CEInstallDir`*`\SRDFCESymapi.db`.
4. Restore every failed-over group:

   ```
   symrdf -g <failed over group name> restore -incr
   ```

5. Monitor the group state by typing the following command:

   ```
   symrdf -g <groupname> query
   ```

When the primary site has good data that you want to copy to secondary site:

1. Open the command line prompt.
2. Navigate to the Cluster Enabler installation directory.
3. Set the `SYMCLI_DB_FILE` environment variable to *`CEInstallDir`*`\SRDFCESymapi.db`.
4. Re-synchronization the primary and secondary sites for each failed over group:

   ```
   symrdf -g <failed over group name> establish  -incr
   ```

5. Monitor the group state by typing the following command:

   ```
   symrdf -g <groupname> query
   ```

Recover a SRDF backup site in case of primary site failures

The following procedure shows how to recover a backup site when the primary site fails. Cluster Enabler lets you set the failover option on a group basis.

For Majority Node Sites clusters

You can restart the backup site using the `/forcequorum` option as described in Microsoft cluster documentation.

For Shared Quorum models

If the Cluster Group (the group in which the quorum disk is a member) does not have the failover option set to Automatic Failover, the group does not failover to any secondary node and therefore the cluster cannot be started. On one of the secondary nodes, use the Recover CE Wizard to start the cluster in Safe Mode. This starts the cluster service on this node with just the Cluster Group.

Once you have cluster service running on the secondary site:

1. Open CE Manager and connect to the cluster.
2. For each group that you want to failover, change the failover policy to Automatic Failover.
3. Use the Microsoft Cluster Administrator/Failover Cluster Manager to bring all of these groups online.

At this point the cluster is running with required services at the backup site.

Recover from an SRDF link failure

The following procedures show how to recover from an SRDF link failure.

IMPORTANT

Choosing the wrong option for restore could cause data loss. Contact Dell EMC support if you have any question about the commands to use.

Groups that failed over on the RDF link

For groups that failed-over on the RDF link, when the link was in a failed state:

1. Choose the remote mirror that has valid user/application data.
2. Move the group to a node that has a valid mirror mapped.
3. Restore the SRDF link.
4. Open CE Manager and run the Storage Discover wizard.
5. Open the command line prompt.
6. Set the `SYMCLI_DB_FILE` environment variable to *`CEInstallDir`*`\SRDFCESymapi.db`
7. If the R1 has valid data, type:

   ```
   symrdf -g <groupname> establish -incr
   ```

   If the R2 has valid data, type:

   ```
   symrdf -g <groupname> restore -incr
   ```

Groups that remained online

For groups that remained online on the same side as before the link failure:

1. Restore the SRDF link.
2. Open CE Manager and run the Storage Discover wizard.

Restrict group movement and recovery

A resource can take a long time to come online when all of the following conditions exist:

- A CE group has the "Restrict Group Movement" policy set.
- The RDF link is down.
- The user manually tries to move the group to a node that is connected to a different storage array.

For example, if the user tries to move group G1 from the R1 side to the R2 side when the RDF link is down, the Microsoft cluster's preferred owner logic attempts to bring the group online on the R2 side as expected.

But since the restrict group movement policy is set for the CE group, Microsoft cluster fails the resource on the R2 side nodes. This is correct behavior and is expected, but it may take a long time for the resource to fail on all the R2 nodes before coming back

online on one of the nodes on the R1 side. This is because by default the cluster tries to bring the group online 3 times on each node. The more nodes there are in the cluster, the longer it takes for the Microsoft cluster preferred owner logic to complete. To minimize this undesirable effect you can change the property of the resources to "Do not Restart". This minimizes the number of retries and reduce the time required to bring the group online.

Recover from a corrupt quorum log

An article on recovering from a corrupt quorum log is available in Microsoft Knowledge Base Article 172951 at the following web address:

support.microsoft.com/kb/172951

Replace a storage array

Use the following procedure to replace a storage array. This procedure assumes that:

- All RDF groups are Dynamic
- All failover cluster groups are configured for Swapping RDF personalities (SwapEnabled) during failover

To replace a storage array:

1. Change the Microsoft Failover Cluster service start up to Manual on all cluster nodes.
2. Failover all groups to the array that you are NOT replacing. Now the groups are online on R1 side of the RDF device.
3. Shutdown all nodes that are attached to the array that is being replaced.
4. Replace the R2 array and establish new R1/R2 relations.
5. Bring the SRDF link up and synchronize R2 with R1 data.
6. Wait for the synchronization to complete.
7. Adjust device masks on all nodes connected to new array, so that the devices are correctly mapped to these hosts.
8. Reboot the nodes attached to new array.
9. Open CE Manager on one of the nodes connected to R1 side.
10. Choose **Actions > More Actions... > Update Mirror Pairs** and step through the wizard processes.
11. Once Update Mirror Pair wizard completes successfully, CE updates its internal configuration database to reflect new R1/R2 relations. At this point you should be able to failover groups between these arrays.
12. Set the Cluster Service Startup type to Automatic on all cluster nodes.

The Recover CE Cluster Wizard

Cluster Enabler provides a wizard to help you recover a failed shared quorum cluster by bringing the cluster online on a single node. The shared quorum cluster fails to come online in a site failover scenario where the failover option for a quorum group is set to Restrict Group Movement. The Recover CE Cluster Wizard changes the failover policy on quorum group to Automatic Failover and then brings the cluster online on the node. You can then use the Create Group Wizard to change other groups' failover policies and bring them online appropriately.

Note: The Recover CE Cluster Wizard is useful for shared quorum clusters. To force a Majority Node Set (MNS) cluster node to form a cluster use the /forcequorum option as set out in the Microsoft Clusters documentation.

Follow these steps to automatically recover and restore a shared quorum cluster using the Recover CE Cluster Wizard:

1. Click the **EMC Cluster Enabler Manager** icon in the navigation tree and select **Action > Recover CE Cluster**. The first page of the Recover CE Cluster Wizard appears.

   Note: When running the Recover CE Wizard to recover a CE cluster, run the wizard on a failed node only when the entire cluster is down.

   ---

2. The Enter Node Name page appears (see Figure 36). Type the **Cluster Name** and **Node Name**, click **Validate**. The Recover CE Wizard should only be run on a single node.

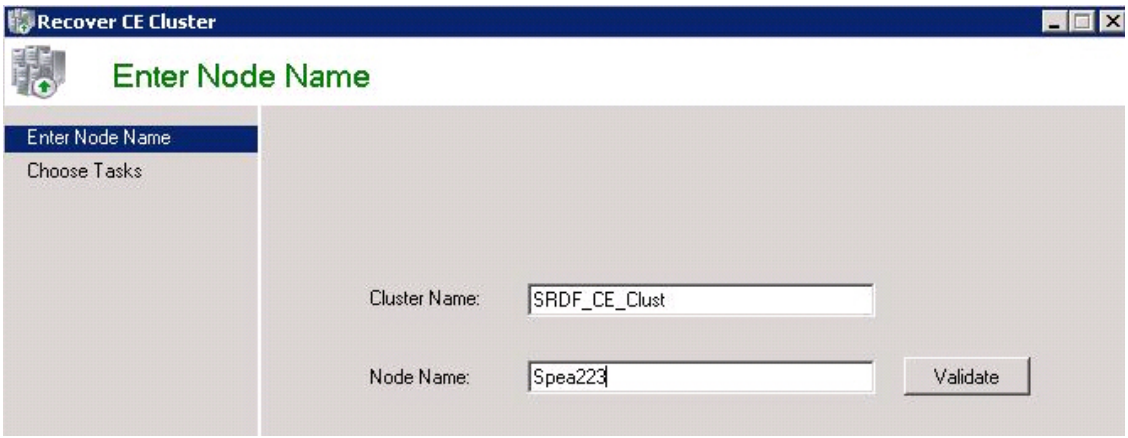

Figure 36 Recover CE Cluster Enter Node Name

3. The Choose Tasks page appears (see Figure 37). Do one of the following:
   - Select **Resolve Cluster Number** start to resolve a cluster number for a Shared Quorum model cluster and recover the cluster.
   - Select **Start Cluster in Safe Mode** to restart a cluster in safe mode and bring the cluster online using previous CE cluster settings.
4. Click **Next**.

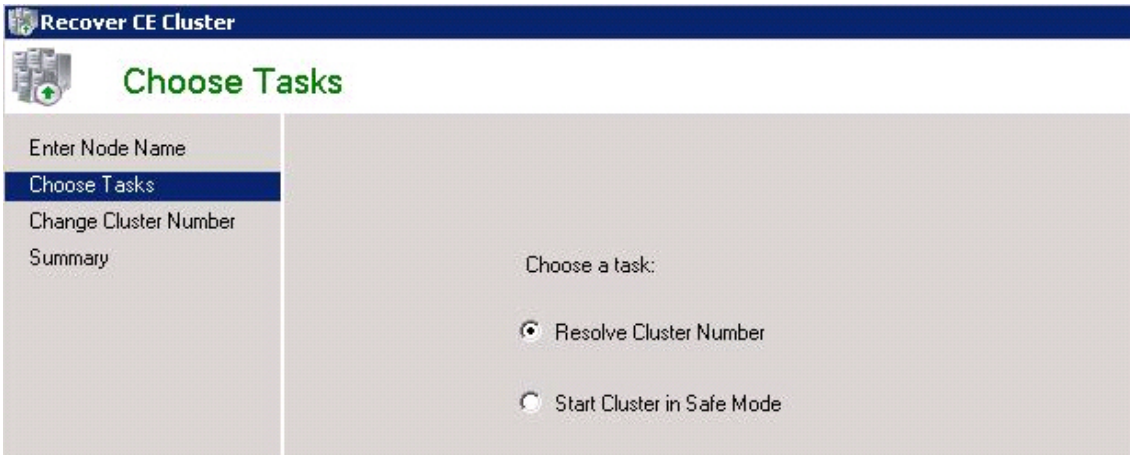

Figure 37 Recover CE Cluster Choose Tasks

5. If you selected **Resolve Cluster Number** in step 3 and are recovering a shared quorum model cluster, the following screen appears.

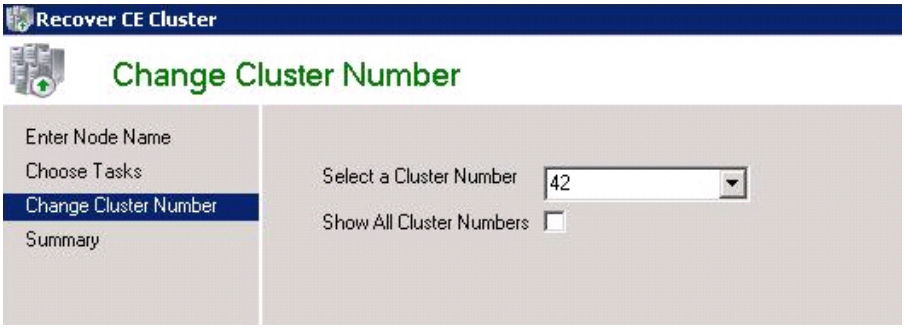

Figure 38 Recover CE Cluster Change Cluster Number

   The wizard generates a list all of the available cluster numbers.

6. From the **Select a Cluster Number** list box, select the Cluster Number that you want to use for the cluster and click **Next**.

   Select **Show All Cluster Numbers** to view all of the cluster numbers both used and unused for the system. Do not select a number that is already in use.

7. When the Started Cluster Service Successfully notification appears, click **Finish**. The Summary page appears and Cluster Enabler restarts the cluster service for the CE cluster.

Configure a custom resource

This section shows how to create and modify a custom resource using the CE Manager. A custom resource could be a Veritas volume or other third-party resource. Once a CE Group is created and the custom resource volumes are added, the storage resource is comprised of storage array disks.

Before you can configure a custom resource using the CE Manager, set up the custom resource using the vendor's resource software (for example, Veritas Volume Manager). Then manually add the custom resource to Microsoft Failover Clusters. For example, a custom resource is of the Resource Type Generic Application. Figure 39 shows a custom resource named test in Group 4 as displayed from the Microsoft Cluster Administrator application.

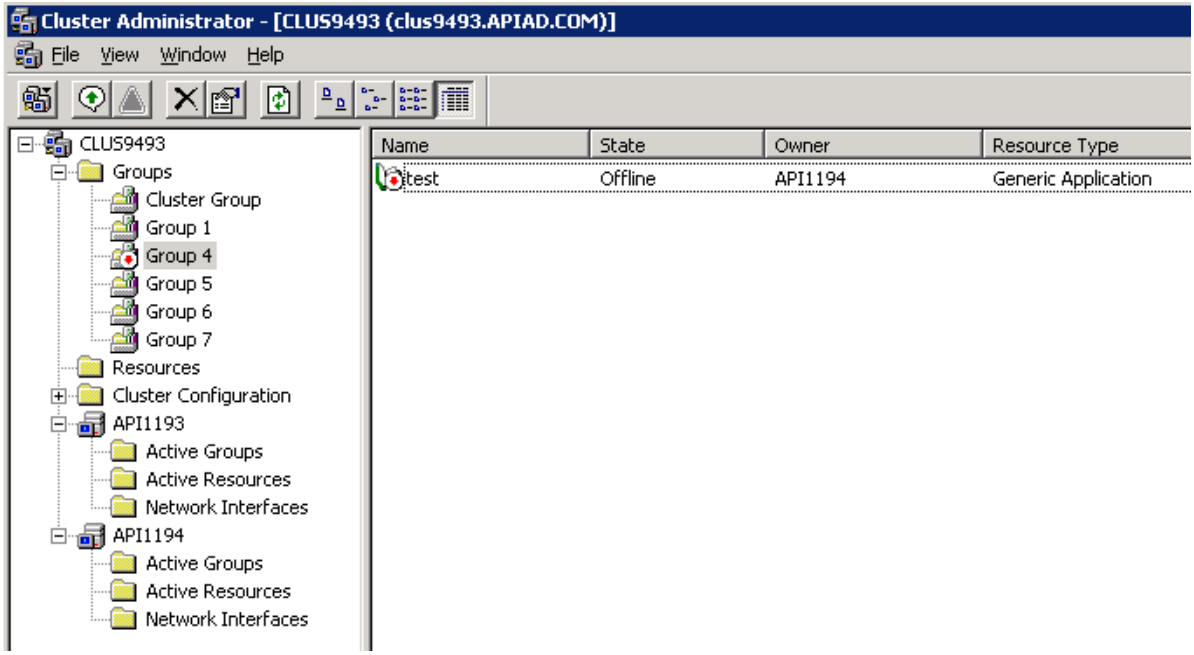

Figure 39 Microsoft Cluster Administrator, Generic Application Resource Type

For Cluster Enabler to recognize a third-party resource, it must be added to cluster properties. Figure 40 displays the CustomResourceTypes as listed in the Cluster Enabler cluster properties, which is viewable from the command line.

```
C:\Program Files\EMC\Cluster-Enabler>cluster /priv
Listing private properties for '':

T  Cluster              Name                           Value
-- -------------------- ------------------------------ ----------------------
M                       CustomResourceTypes            <?xml version="1.0"?>
M                       CustomResourceTypes            <CustomResTypeList>
M                       CustomResourceTypes              <CustomResType>Volume Manager Disk Group</C
ustomResType>
M                       CustomResourceTypes            </CustomResTypeList>
M                       SiteList                       <?xml version="1.0"?>
<SiteList>
  <Site>
    <SiteName>Site1</SiteName>
    <Node>API1194</Node>
    <Storage>SYMMETRIX+000187900830</Storage>
  </Site>
  <Site>
    <SiteName>Site2</SiteName>
    <Node>API1193</Node>
    <Storage>SYMMETRIX+000187900848</Storage>
  </Site>
</SiteList>
D                       ClusterNumber                  38 (0x26)
D                       ISCECluster                    1 (0x1)
```

Figure 40 Cluster properties

If you would like to use another third-party resource (for example, Generic Application), type the following command:

```
cluster /priv CustomResourceTypes="<?xmlversion=1.0?>,"<CustomResTypeList>,
   "<CustomResType>Volume Manager Disk Group</CustomResType>,
   "<CustomResType>Generic Application</CustomResType>,
"</CustomResTypeList>:MULTISTR
```

Figure 41 shows the changed cluster properties with Generic Application added to CustomResourceTypes.

.

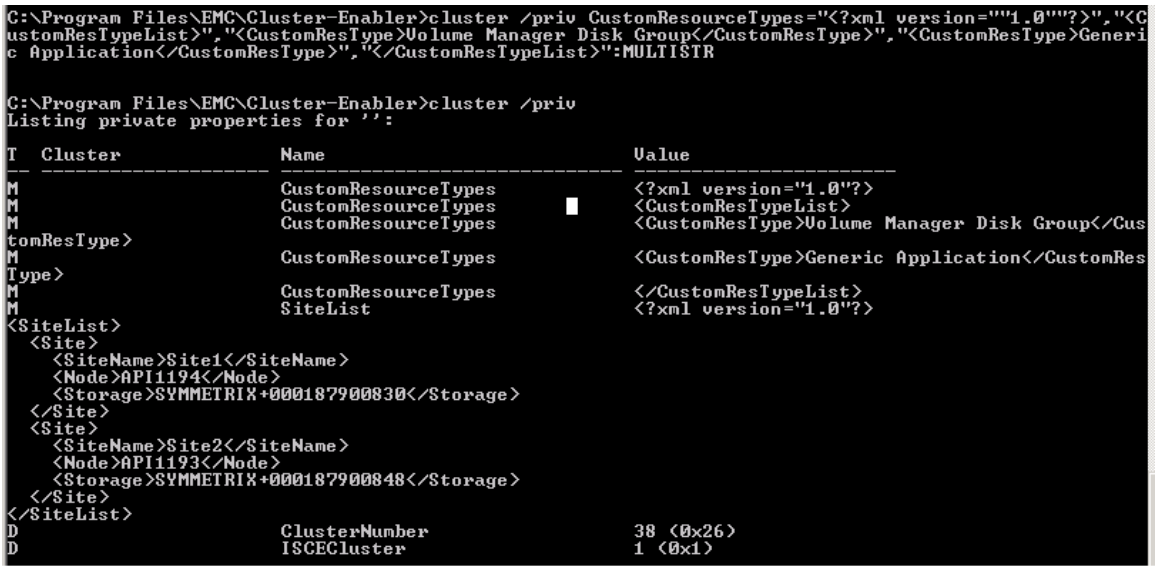

Figure 41 Cluster properties with Generic Application

After you have configured the custom resource for Microsoft failover clusters, you can use the CE manager Create Group Wizard to create a custom resource CE Group.

Create a custom resource CE Group

1. Open the CE Manager, click the **Groups** icon in the Navigation tree and select **Action > Create Group**. CE loads configuration data and then displays the first page of the Create Group Wizard.

   The **Enter Group Name** dialog appears.

   Note: A mirrored pair needs to be present on the array before attempting to create a group. Run the Storage Discover Wizard to detect a newly created mirrored pair by right-clicking on the cluster name or clicking the **Discover** button in the Select Devices page of the Create Group Wizard.

   ---

2. Enter the exact same **Group Name** as displayed in the Microsoft Cluster Administrator and click **Create**.

   The next page enables you to select devices for inclusion in the new group. The wizard recognizes that this is a custom resource group and displays a warning to that effect.

Note: Cluster Enabler creates a SYMAPI device group for the storage array. It also creates the corresponding CE resource which it makes the custom resource dependent on. Physical disk resources are not created in the failover cluster by Cluster Enabler.

3. Select the appropriate devices from the list using the check boxes. Then click **Next**.

   The Select Group Policy page appears (see Figure 42).

   Notes:

   - Click **Expand All** to expand the tree to show all devices.
   - Use the checkboxes to select the types of device to display: **Async**, **Cascaded**, and **Concurrent**. For example, selecting the **Async** checkbox displays all SRDF asynchronous capable devices within in the same RA group.
   - An error message appears if selected type of devices is used up or not available.
   - Selecting devices from a single RA group, creates a device group. Selecting devices from multiple RA groups, creates a composite group.

4. Select for the group from the list box and click **Next**.

   The available policies are:

Table 17 Custom resource group failover policies

Failover policy	Description
Restrict Group Movement	The group cannot fail over to a peer node. In a replication link failure, this setting only attempts to move disk laterally. If the replications link is up, this setting has no impact.
Automatic Failover	The group can automatically failover to any remote site node in the event of a replication link failure.

5. When the Group Created Successfully notification appears, click **Finish**.

   Cluster Enabler refreshes the CE cluster. Upon completion of the refresh, the group that you created appears under Groups. If you do not see the new group, select **Action > Refresh**.

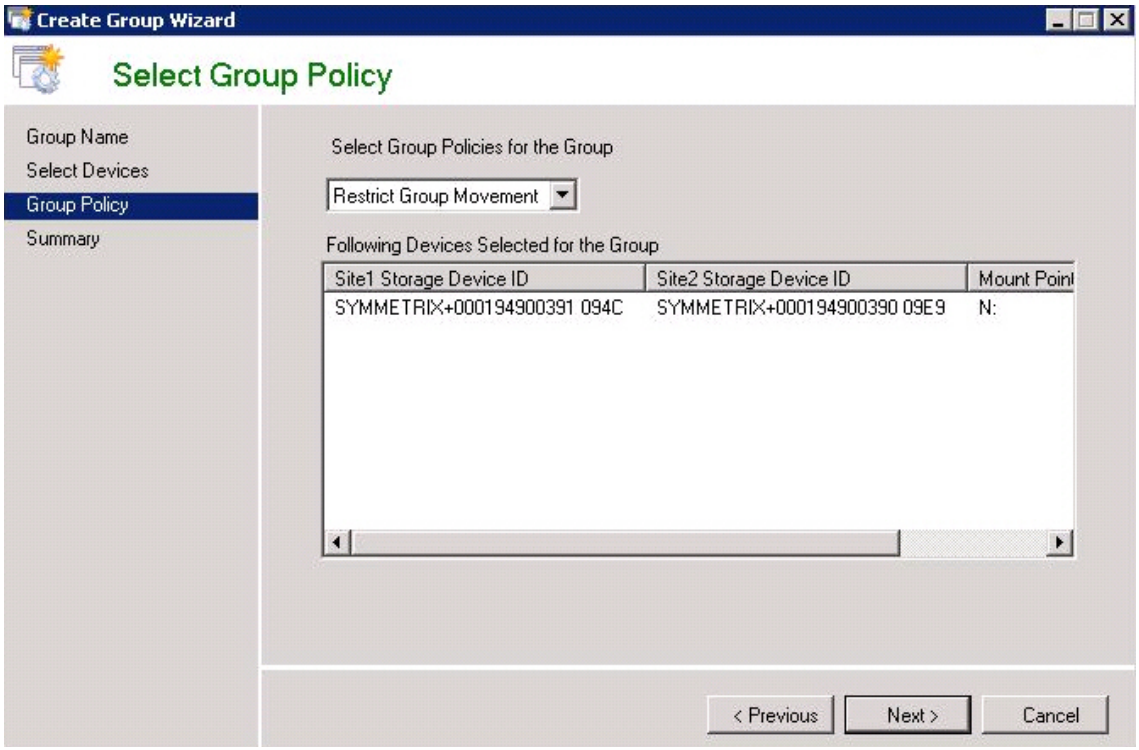

Figure 42 Select Group Policy, custom resource

6. Open the Microsoft Cluster Administrator application and select the custom resource. A resource of resource type "EMC Cluster Enabler" is visible in the custom resource. Figure 43 shows an example for a resource named EMC_Group 4.

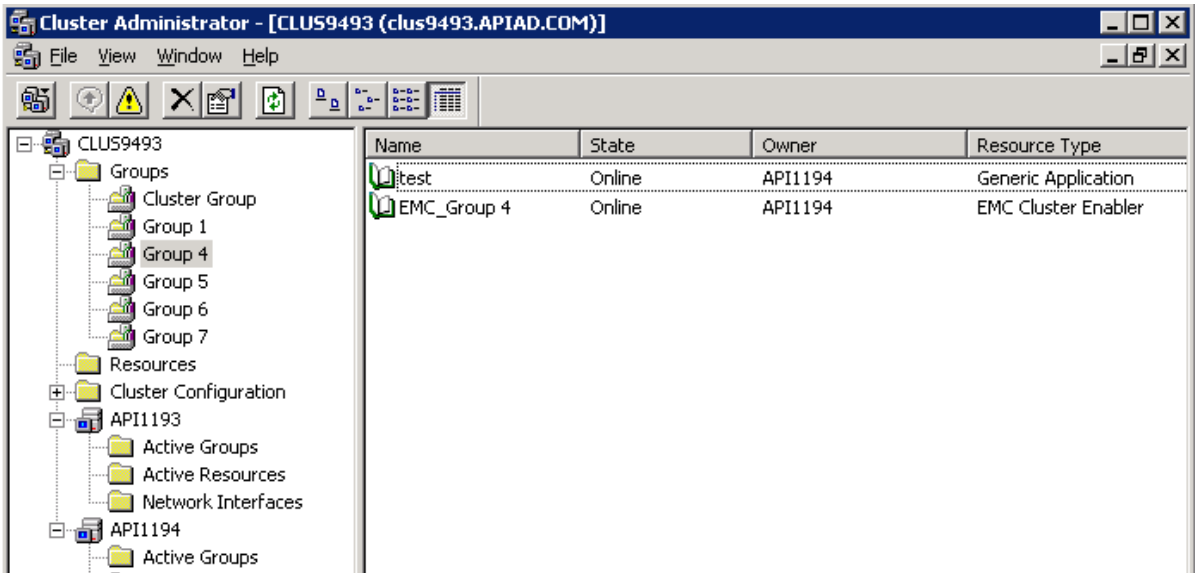

Figure 43 Microsoft Cluster Administrator, EMC_Group 4

Edit a custom resource CE Group

If the composition of an underlying custom resource changes, make the same changes to the CE Group custom resource by adding or deleting devices.

1. In CE Manager, click the **Group** icon in the navigation tree, select the group to modify and select **Action > Modify Group**.

   CE reads the storage configuration and then displays the first page of the Modify Group Wizard.

   Note: A mirrored pair needs to be present on the array before you can modify a group. Use the Storage Discover Wizard to detect a newly created mirrored pair by right-clicking on the cluster name or clicking the **Discover** button in the Select Devices page of the Modify group wizard.

   ---

2. From the Select Devices page, select **Add Devices** or **Delete Devices** in the **Select Action** list box.

   A list of available devices that you can add or remove appears. RA group pairs and the devices contained within the RA group are in a tree view. Initially, the RA Groups are shown in collapsed view. Select **Expand All** to expand the tree view to see individual devices within each group.

3. Select the check boxes for the devices you want to add or delete and click **Next**.

   Selecting the RA group, selects all devices in that group.

4. When the Validate Selection page appears (see Figure 44) click **Next**.

   The wizard recognizes that this is a custom resource group and displays a warning to that effect.

   Note: Only the storage group and the corresponding CE resource are modified. No physical disk resources are added to the failover cluster.

   ---

5. When the Group Modified Successfully notification appears, click **Finish**.

Cluster Enabler refreshes the CE cluster (see Figure 45). Upon completion of the refresh, the updated group information reflects the devices added or deleted. If you do not see the updated group information, select **Action > Refresh**.

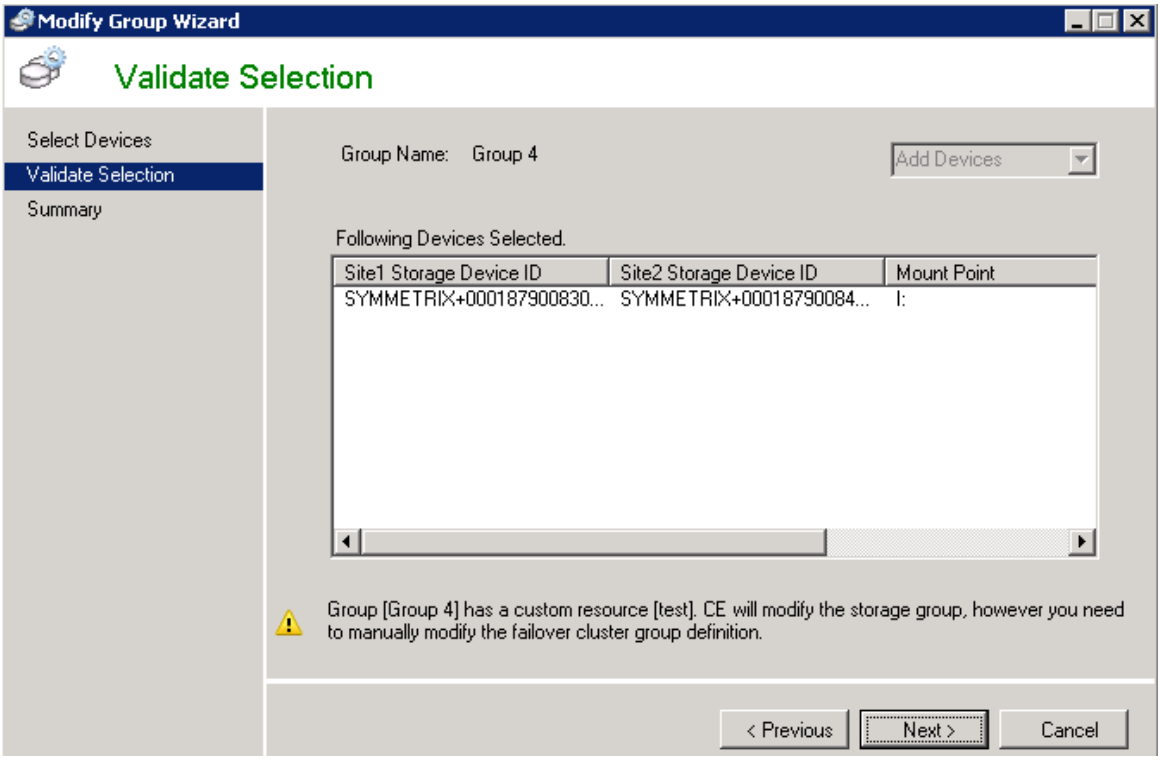

Figure 44 Validate selection, custom resource

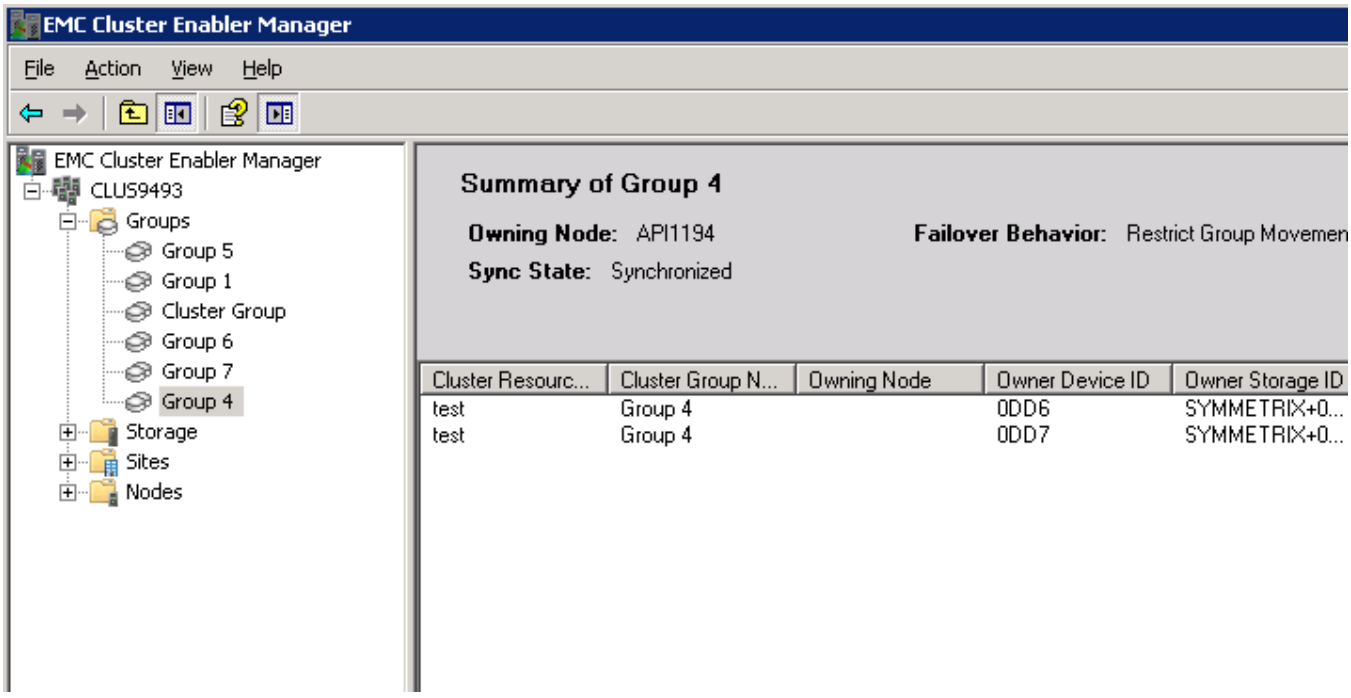

Figure 45 Summary of Group 4, custom resource

Delegate CE administration

Use the `cesec.exe` utility, located in the Cluster Enabler installation directory (typically, `C:\Program Files\EMC\Cluster-Enabler`) to delegate CE and cluster management tasks to a non-local administrator.

You can:

- Add and remove users and groups who have delegated administration
- List security settings for Cluster Enabler

Add and remove delegated users and groups

1. Open a command prompt.
2. Type:

   `cesc.exe -ce` *`action principal`*

   Replace:

   action with `set` or `remove`, as appropriate.

   principal with the identity of the user or group. Use these formats:

   *domain**user*
   user@*domain*.com
   *domain**group*
   group@*domain*.com

   Here, *domain* is the name of a domain, *user* is the name of a user, and *group* is the name of a user group.
3. Repeat steps 1 and 2 on all other nodes in the cluster.

List security settings

1. Open a command prompt.
2. Type:

   ```
   cesc.exe -oc list
   ```

APPENDIX A
Windows log messages

The Windows event log displays descriptive messages for some of the more common events encountered when using Cluster Enabler as listed in Table 18. There are three event types:

- Information
- Warning
- Error

Table 18 Windows event log messages (page 1 of 2)

Event ID	Event type	Description	Action
1	Informational	Generic ID used to report informational messages.	Action varies based on description text.
2	Warning	Generic ID used to report warning messages.	Action varies based on description text.
3	Error	Generic ID used to report error messages.	Action varies based on description text.
4	Informational	Informational message generated when a group comes online successfully.	No action necessary.
5	Error	Error message generated when a group fails to come online.	The description text indicates the name of the group that failed to come online. Look at previous event log messages and application logs to find the root cause of the failure.
6	Error	An unexpected application error occurred.	1. Turn on verbose logging (logging level 5) and repeat the action. 2. If failure occurs again, save the Windows event log and the CE application log, and contact Dell EMC support.
7	Error	The link between the storage arrays is down for storage group (*GroupName*).	Use storage array CLI interfaces to determine the root cause of the problem.
8	Informational	The link between the storage arrays is replicating data to the remote storage array.	No action necessary.
9	Error	Communication or data access to the WMI (Windows Management Instrumentation component) service failed.	1. Read the event log messages and application logs to find the root cause of the problem. 2. If failure occurs again, save the Windows event log and the CE application log, and contact Dell EMC support.
10	Error	A failure occurred while reading or writing storage group information.	1. Turn on verbose logging (logging level 5) and repeat the action. 2. If failure occurs again, save the Windows event log and the CE application log, and contact Dell EMC support.

Table 18 Windows event log messages (page 2 of 2)

Event ID	Event type	Description	Action
11	Error	A failure occurred while reading or writing storage group information to the cluster registry.	1. Turn on verbose logging (logging level 5) and repeat the action. 2. If failure occurs again, save the Windows event log and the CE application log, and contact Dell EMC support.
12	Error	A failure occurred while deleting a mirror group.	Read the event log messages and application logs to find the root cause of the problem.
13	Error	A failure occurred while creating a mirror group.	Read the event log messages and application logs to find the root cause of the problem.

GLOSSARY

This glossary contains terms related to the Cluster Enabler software.

A

agent An installed program designed to control a particular resource type. Each type of resource supported in a cluster is associated with an agent.

availability The ability to continue to provide a service even during hardware or software failure.

C

cache Random access electronic storage used to retain frequently used data between the CPU and either a hard disk or slower RAM. It speeds up general data flow because a cache can be accessed quickly.

CDP See "continuous data protection (CDP)".

client A computer using services or resources provided by a remote machine, called a server. Often, communications software has a separate version for the client, or guest, and the server, or host.

Clients create a TCP/IP session with a service in the cluster using a known IP address. This address appears to the cluster software as a resource in the same group as the application providing the service. In a failure, the Cluster Service moves the entire group to another system.

client failover The response of a client machine after resource failure on the server for the client caused a resource failover. A client detects a failure in the session and reconnects in exactly the same manner as the original connection. The IP address is now available on another machine, and the connection is quickly reestablished. In this simple case, all information related to the original session not committed to disk is lost. This provides higher availability, but no fault tolerance for the service. Applications can use transactions to guarantee the client request is committed to the server database to gain fault-tolerant semantics.

CLR See "continuous local and remote replication (CLR)".

cluster A group of two or more independent computers addressed and used as a single system.

cluster-aware software Software that provides a restart mechanism invoked whenever the application resource is moved to another node in the cluster.

cluster service The collection of software on each node that manages all cluster-specific activity.

Cluster Shared Volumes Cluster Shared Volumes (CSV) is a Microsoft Failover Clustering feature that allows all nodes in a cluster concurrent access to data on every CSV-enabled shared disk.

consistency group — For RecoverPoint, a consistency group is a data set consisting of the production source and its replicas. A consistency group comprises the production source volumes and either a local replica, remote replica, or both. Each consistency group contains as many replication sets as there are volumes in the production storage to replicate.

continuous asynchronous — A RecoverPoint replication mode where each write transaction is acknowledged locally at the source side and then sent to the target side. The primary advantage of continuous-asynchronous replication is its ability to provide synchronous-like replication without degrading the performance of host applications.

continuous data protection (CDP) — A RecoverPoint configuration that uses a methodology that continuously captures or tracks data modifications and stores changes independent of the primary data, enabling recovery points from any point in the past. CDP provides fine granularities of restorations to infinitely variable recovery points.

continuous local and remote replication (CLR) — A RecoverPoint configuration that includes both a CDP and a CRR copy, providing concurrent local and remote data protection. In RecoverPoint, the CDP copy is normally used for operational recovery, while the CRR copy is normally used for disaster recovery.

continuous remote replication (CRR) — A RecoverPoint configuration where data is transferred between two sites over Fibre Channel or a WAN. In this configuration, the RPAs, storage and splitters exist at both the local and the remote site.

continuous synchronous — A RecoverPoint replication mode. In continuous synchronous replication, the host application that initiates the write waits for an acknowledgment from the replica before continuing. Replication in synchronous mode produces a replica that is 100% up to date with the production source.

create mirror — To establish a remote mirror, that is, use the remote mirror software to create data structures on one or more LUNs on specific storage systems, such that one is the primary image and the other is a secondary image.

CRR — See “continuous remote replication (CRR)”.

D

data center migrations — A function that reduces application outage to minutes instead of hours.

dependency — The requirement of one resource needing another resource to function properly. The Cluster Enabler resource becomes a dependency for physical disk resources in the cluster. Therefore, any operations performed on the disk resource cannot be completed until the Cluster Enabler resource has been invoked.

device — A uniquely addressable part of the storage array consisting of a set of access arms, the associated disk surfaces, and the electronic circuitry required to locate, read, and write data. Also called a LUN (logical unit number).

device group — A grouping of several devices established to provide configuration, status, and performance data on the collective devices within the group.

device number — The value that logically identifies a disk device in a string. See also “LUN”.

disaster recovery — A function that recovers data at the disaster recovery site in minutes rather than days.

discover — A discover action performed in the Cluster Enabler Configuration Wizard scans the storage array connected to the current node and gathers device information.

F

failback — The action of moving a resource back to the cluster member designated to be the resource's Preferred Owner. By default, resources are owned by their Preferred Owner, so a failback would only occur if the resource moved from its Preferred Owner. This is likely the result of a failover.

failover — The process of taking one or more resources offline on one cluster member and bringing them online on another cluster member.

fault-tolerant — Continuous operation in case of failure. A fault-tolerant system can be created using two or more computers that duplicate all processing, or having one system stand by if the other fails. It can also be built with redundant processors, control units, and peripherals. Fault-tolerant operation requires backup power in a power failure. It may also imply duplication of systems in disparate locations in the event of natural catastrophe or vandalism.

FDDI — An acronym for Fiber Distributed Data Interface.

Fibre Channel — A high-speed serial interface capable of data transfer rates of up to 400 MB/s.

forced failover — A CE feature allowing you to automatically keep a cluster up on a particular array or arrays in a total site disaster.

forced quorum — Software functionality allowing the cluster to be forced up in the event that total communication is lost between nodes and Microsoft Failover Cluster. Microsoft Failover Cluster wants to shut down the cluster to avoid a split-brain condition.

See “split-brain condition”.

G

graphical user interface (GUI) — A method that allows users to interact with the computer and its special applications based on graphics instead of text. GUIs use icons, pictures, and menus and use a mouse as well as a keyboard to accept input.

group — A collection of resources to be managed as a single unit. Usually, a group contains all elements needed to run a specific application and for client systems to connect to the service provided by the application. Groups allow an administrator to combine resources into larger logical units and manage them as a unit. Operations performed on a group affect all resources contained within that group.

H

HBA See "host bus adapter (HBA)".

heartbeat A polling communication mechanism used by the cluster processes to determine whether the other members of the cluster are alive and working or have failed. If the heartbeat is not functioning, a failover is initiated, and another node in the cluster takes over the services.

high availability The characteristic of a computer system/computing environment that allows it to continue to provide applications and access to data if a single component or resource fails. Service is interrupted for only a brief time, and may or may not be apparent to the end users.

host bus adapter (HBA) A device circuit board that provides an interface between the SCSI bus and the computer I/O bus (for example, PCI, EISA, microchannel).

I

I/O Input/output.

identifier (ID) A sequence of bits or characters that identifies a program, device, controller, or system.

L

lateral node Nodes connected to the same Symmetrix array.

LUN A logical unit number (LUN) is a unique identifier used on a SCSI bus that enables it to differentiate between up to eight separate storage devices (each of which is a logical unit). See also, "device number".

M

Microsoft Management Console (MMC) A Microsoft user interface (UI) framework for use in administrating different components of the Microsoft Windows operating platform. This framework is used to host-specific UI/control extensions called *snap-ins*. Use snap-ins to administer both local and remote computers. Third-party snap-ins can be written for use with MMC.

mirrored pair A device comprising two hypervolumes with all data recorded twice—once on each disk drive.

MMC See "Microsoft Management Console (MMC)".

N

network interface card (NIC) A device that provides network communication capabilities to and from a computer system.

Node Majority A quorum-capable resource based on replicating data to local disks associated with a majority of cluster nodes. MNS enables you to create a server cluster without shared disk for the quorum resource. Cluster Enabler allows you to configure an MNS cluster on Windows Server 2012 Enterprise and Datacenter Editions.

nodes — Members of a cluster. Also referred to as systems. A node contains a CPU, disk, and network resource.

O

offline — The state of a resource or group that classifies it as unavailable. When used in context with a cluster member, offline implies the cluster member may not be booted, or the cluster service on the node in question may not be functioning properly.

online — The state of a resource or group that classifies it as available. When used in context with a cluster member, online implies the other cluster members are receiving heartbeats from the cluster member in question. See also "resource".

Q

quorum disk — An ordinary disk volume used as a special communication mechanism between server systems. In a Microsoft failover cluster, a small amount of cluster system data (a few megabytes) is stored on this volume. The SCSI-3 `Reserve` and `Reset` commands are used to move quorum-disk ownership back and forth between nodes. If the heartbeat mechanism fails, the quorum disk is used for each node to verify whether the other node is still functioning. Because not all disk products implement these multihost SCSI-3 commands, not all disk products will work in a failover cluster environment. Thus, Microsoft is very rigorous in providing the Cluster/RAID category of tests to qualify disks (refer to Microsoft's Hardware Compatibility List) capable of running with Microsoft failover cluster software).

R

RAID — Redundant array of independent disks. Data is stored on multiple magnetic or optical disk drives to increase output performance and storage capacities and to provide varying degrees of redundancy and fault tolerance. Instead of storing valuable data on a single hard disk that could fail at any time, RAID ensures a backup copy of all information always exists by spreading data among multiple hard disks.

Replication set — A RecoverPoint term. A storage volume in the production source that is replicated must have a corresponding volume at each copy. A replication set is production volume and its associated volume at the local copy, the remote replica, or both.

resource — An object managed by the Cluster Service that sees all resources as identical opaque objects. Resources may include physical hardware devices, such as disk drives and network cards, or logical items, such as disk partitions, TCP/IP addresses, entire applications, and databases. A resource is said to be online on a node when it is providing its service on that specific node.

resource failback — The movement of resources back to their preferred location in the cluster. This is usually done under manual user control to avoid a situation where a resource is failed back, and then immediately fails over again because of an unresolved node problem. Microsoft Failover Cluster also allows automatic failback and provides a timing window to try to avoid repeated failovers.

resource failover — The process where control of a resource moves to another node of a cluster. Failover can be initiated automatically or manually. When initiated automatically, the cluster management software detects a failure of server node hardware or an application. When manually initiated, the cluster administrator uses the Cluster Administrator software application.

resource group — A collection of resources to be managed as a single unit. Usually a group contains all elements needed to run a specific application, and for client systems to connect to the service provided by the application. Groups allow an administrator to combine resources into larger logical units and manage them together. Operations performed on a group affect all resources contained within that group.

S

scalability — The ability to add new components to a storage system as system load increases.

SCSI — Small Computer System Interface. SCSI is a high-speed parallel interface used to connect microcomputers to SCSI peripheral devices, such as disks, printers, and other computers and local area networks.

snap-in — See “Microsoft Management Console (MMC)”.

snapshot — A RecoverPoint term. A snapshot is the difference between one consistent image of stored data and the next. Snapshots are taken seconds apart. The application writes to storage; at the same time, the splitter provides a second copy of the writes to the RecoverPoint appliance.

snapshot replication mode — A RecoverPoint replication mode that only transfers data that has changed between one consistent image of the storage subsystem and the next. By definition, snapshot replication produces a replica that is not up to date.

split-brain condition — A total communication failure while both nodes remain operational. A split-brain condition is a potential cause of logical data corruption. For example, if both sides assume the other is dead and begin processing new transactions against their copy of the data, two separate and unreconcilable copies of the data can be created.

stretch cluster — A Microsoft cluster that is geographically distributed across multiple physical locations.

V

virtual servers — See “nodes”.

W

workload migrations — Similar to data center migrations; especially useful for minimizing outages during preventative maintenance of hardware or software.