EMC²

# Networking for Storage Virtualization and EMC RecoverPoint

Version 4.0

- EMC VPLEX and RecoverPoint Solutions
- Interswitch Link Hops Basics and Examples
- EMC RecoverPoint SAN Connectivity Information

Ron Dharma
Dana Lane
David Hughes

EMC²
TECHBOOKS

**Part number H8075.3**

# Contents

# Figures

# Preface

*This EMC Engineering TechBook provides a high-level overview of using a block storage virtualization solution, such as EMC VPLEX, as well as EMC RecoverPoint replication protection. It also describes how RecoverPoint can be integrated into the design of a simple or complex topology. Basic information on interswitch link hops, along with examples showing hop counts in four different fabric configurations, are included.*

*E-Lab would like to thank all the contributors to this document, including EMC engineers, EMC field personnel, and partners. Your contributions are invaluable.*

*As part of an effort to improve and enhance the performance and capabilities of its product lines, EMC periodically releases revisions of its hardware and software. Therefore, some functions described in this document may not be supported by all versions of the software or hardware currently in use. For the most up-to-date information on product features, refer to your product release notes. If a product does not function properly or does not function as described in this document, please contact your EMC representative.*

## Audience

This TechBook is intended for EMC field personnel, including technology consultants, and for the storage architect, administrator, and operator involved in acquiring, managing, operating, or designing a networked storage environment that contains EMC and host devices.

## EMC Support Matrix and E-Lab Interoperability Navigator

For the most up-to-date information, always consult the *EMC Support Matrix* (ESM), available through E-Lab Interoperability Navigator (ELN), at: http://elabnavigator.EMC.com, under the **PDFs and Guides** tab.

The *EMC Support Matrix* links within this document will take you to Powerlink where you are asked to log in to the E-Lab Interoperability Navigator. Instructions on how to best use the ELN (tutorial, queries, wizards) are provided below this **Log in** window. If you are unfamiliar with finding information on this site, please read these instructions before proceeding any further.

Under the **PDFs and Guides** tab resides a collection of printable resources for reference or download. All of the matrices, including the ESM (which does not include most software), are subsets of the E-Lab Interoperability Navigator database. Included under this tab are:

- The *EMC Support Matrix*, a complete guide to interoperable, and supportable, configurations.
- Subset matrices for specific storage families, server families, operating systems or software products.
- Host connectivity guides for complete, authoritative information on how to configure hosts effectively for various storage environments.

Under the **PDFs and Guides** tab, consult the *Internet Protocol* pdf under the "Miscellaneous" heading for EMC's policies and requirements for the *EMC Support Matrix*.

## Related documentation

Related documents include:

- The former *EMC Networked Storage Topology Guide* has been divided into several TechBooks and reference manuals. The following documents, including this one, are available through the E-Lab Interoperability Navigator, **Topology Resource Center** tab, at http://elabnavigator.EMC.com.

  These documents are also available at the following location:

http://www.emc.com/products/interoperability/topology-resource-center.htm

- *Backup and Recovery in a SAN TechBook*
- *Building Secure SANs TechBook*
- *Extended Distance Technologies TechBook*
- *Fibre Channel over Ethernet (FCoE): Data Center Bridging (DCB) Concepts and Protocols TechBook*
- *Fibre Channel over Ethernet (FCoE): Data Center Bridging (DCB) Case Studies TechBook*

- •
- *Fibre Channel SAN Topologies TechBook*
- *iSCSI SAN Topologies TechBook*
- *Networked Storage Concepts and Protocols TechBook*
- *WAN Optimization Controller Technologies TechBook*
- *EMC Connectrix SAN Products Data Reference Manual*
- *Legacy SAN Technologies Reference Manual*
- *Non-EMC SAN Products Data Reference Manual*

◆ *EMC Support Matrix*, available through E-Lab Interoperability Navigator at http://elabnavigator.EMC.com >**PDFs and Guides**

◆ RSA security solutions documentation, which can be found at http://RSA.com > **Content Library**

◆ White papers include:

- *Workload Resiliency with EMC VPLEX Best Practices Planning*

All of the following documentation and release notes can be found at http://Powerlink.EMC.com. From the toolbar, select **Support > Technical Documentation and Advisories**, then choose the appropriate Hardware/Platforms, Software, or Host Connectivity/HBAs documentation links.

Hardware documents and release notes include those on:

◆ Connectrix B series
◆ Connectrix M series
◆ Connectrix MDS (release notes only)
◆ VNX series
◆ CLARiiON
◆ Celerra
◆ Symmetrix

Software documents include those on:

◆ EMC Ionix ControlCenter
◆ RecoverPoint
◆ TimeFinder
◆ PowerPath

The following E-Lab documentation is also available:

◆ Host Connectivity Guides
◆ HBA Guides

For Cisco and Brocade documentation, refer to the vendor's website.

- http://cisco.com
- http://brocade.com

## Authors of this TechBook

This TechBook was authored by Ron Dharma, Dana Lane, and David Hughes, along with other EMC engineers, EMC field personnel, and partners.

**Ron Dharma** is a Principal Integration Engineer and team-lead for Advance Product Solution group in E-Lab. Prior to joining EMC, Ron was a SCSI software engineer, spending almost 11 years resolving integration issues in multiple SAN components. He dabbled in almost every aspect of the SAN including storage virtualization, backup and recovery, point-in-time recovery, and distance extension. Ron provided the original information in this document, and works with other contributors to update and expand the content.

**Dana Lane** is a is a Principle Systems Integration Engineer and has been with EMC for over 6 years. Dana works in E-Lab qualifying replication technologies such as RecoverPoint, Data Protection Advisor, and Axxana. Dana has over 20 years experience working in the storage environment.

**David Hughes** is a Principal Systems Integration Engineer and has been with EMC for over 15 years. For the past 5 years, David has worked in EMC E-Lab qualifying blade servers, FC/FCIP switch hardware and firmware, FC-to-iSCSI SAN Routers, EMC VNX series and CLARiiON storage systems, WAN Optimization Controllers and EMC's VPLEX. Prior to working in the E-Lab, David was an EMC Level II Technical Support subject matter expert for Brocade products, providing support across all EMC-supported FC switches and host connectivity. David also spent time working in EMC's IT and Manufacturing departments.

## Conventions used in this document

EMC uses the following conventions for special notices:

**CAUTION**

**CAUTION, used with the safety alert symbol, indicates a hazardous situation which, if not avoided, could result in minor or moderate injury.**

**IMPORTANT**

**An important notice contains information essential to software or hardware operation.**

**Note:** A note presents information that is important, but not hazard-related.

### Typographical conventions

EMC uses the following type style conventions in this document.

| | |
|---|---|
| Normal | Used in running (nonprocedural) text for:<br>• Names of interface elements (such as names of windows, dialog boxes, buttons, fields, and menus)<br>• Names of resources, attributes, pools, Boolean expressions, buttons, DQL statements, keywords, clauses, environment variables, functions, utilities<br>• URLs, pathnames, filenames, directory names, computer names, filenames, links, groups, service keys, file systems, notifications |
| **Bold** | Used in running (nonprocedural) text for:<br>• Names of commands, daemons, options, programs, processes, services, applications, utilities, kernels, notifications, system calls, man pages<br>Used in procedures for:<br>• Names of interface elements (such as names of windows, dialog boxes, buttons, fields, and menus)<br>• What user specifically selects, clicks, presses, or types |
| *Italic* | Used in all text (including procedures) for:<br>• Full titles of publications referenced in text<br>• Emphasis (for example a new term)<br>• Variables |
| `Courier` | Used for:<br>• System output, such as an error message or script<br>• URLs, complete paths, filenames, prompts, and syntax when shown outside of running text |

| | |
|---|---|
| **`Courier bold`** | Used for:<br>• Specific user input (such as commands) |
| *`Courier italic`* | Used in procedures for:<br>• Variables on command line<br>• User input variables |
| < > | 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 indicate content that you must specify (that is, x or y or z) |
| ... | Ellipses indicate nonessential information omitted from the example |

## Where to get help

EMC support, product, and licensing information can be obtained as follows.

**Product information —** For documentation, release notes, software updates, or for information about EMC products, licensing, and service, go to the EMC Powerlink website (registration required) at:

`http://Powerlink.EMC.com`

**Technical support —** For technical support, go to Powerlink and choose **Support**. On the Support page, you will see several options, including one for making a service request. Note that to open a service request, you must have a valid support agreement. Please contact your EMC sales representative for details about obtaining a valid support agreement or with questions about your account.

### We'd like to hear from you!

Your feedback on our TechBooks is important to us! We want our books to be as helpful and relevant as possible, so please feel free to send us your comments, opinions and thoughts on this or any other TechBook:

`TechBooks@emc.com`

# 1

# Virtualization

This chapter contains the following information on virtualization:

## Introduction

You can increase your storage infrastructure flexibility through storage virtualization. To enable simultaneous information access of heterogenous storage arrays from multiple storage vendors within data centers, use block storage virtualization solutions, such as EMC VPLEX™ described in this TechBook in Chapter 2, ”EMC VPLEX.”

The VPLEX solution complements EMC's virtual storage infrastructure and provides a layer supporting virtual storage between host computers running the applications of the data center and the storage arrays providing the physical storage used by these applications.

EMC block storage virtualization provides seamless data mobility and lets you manage multiple arrays from a single interface within and across a data center, providing the following benefits:

- Mobility — Achieve transparent mobility and access in and across a data center.
- Scalability — Start small and grow larger with predictable service levels.
- Performance — Improve I/O performance and reduce storage array contention with advanced data caching.
- Automation — Automate sharing, balancing, and failover of I/O across data centers.
- Resiliency — Mirror across arrays without host impact and increase high availability for critical applications.

EMC RecoverPoint provides cost-effective, continuous remote replication and continuous data protection, enabling on-demand protection and recovery to any point in time. Chapter 3, ”EMC RecoverPoint,” provides more information on RecoverPoint.

# Benefits of VPLEX and storage virtualization

EMC VPLEX™ is an inline block-based virtualization product that share a common set of features and benefits. There are numerous benefits to storage virtualization. These benefits include (but are *not* limited to) the following:

- Create an open storage environment capable of easily supporting the introduction of new technologies that are incorporated in the storage virtualizer. Customers can use heterogenous storage arrays that do not have an interoperable feature set, such as Heterogenous Array Mirroring, and accomplish the mirroring using a storage virtualizer.
- Significantly decrease the amount of planned and unplanned downtime by providing the ability for online migration and data protection across heterogeneous arrays
- Increase storage utilization and decrease the amount of "stranded" heterogeneous storage
- Reduce management complexity through single pane of glass management with simplified policies and procedures for heterogeneous storage arrays
- Improve application, storage, and network performance (in optimized configurations) through load balancing and multipathing

VPLEX provides the following key benefits:

- The combination of a virtualized data center and EMC VPLEX provides customers entirely new ways to solve IT problems and introduce new models of computing.
- VPLEX eliminates boundaries of physical storage and allows information resources to be transparently pooled and shared over distance for new levels of efficiency, control, and choice.
- VPLEX Local and VPLEX Metro products address the fundamental challenges of rapidly relocating applications and large amount of information on-demand within and across data centers, which is a key enabler of the private cloud.

Chapter 2, "EMC VPLEX," provides more details on VPLEX.

## Benefits of EMC RecoverPoint

RecoverPoint provides the following key services in a data center:

- **Intelligent fabric splitting —** Allows data to be split and replicated in a manner that is transparent to the host. By having the switch intercept host writes and duplicate, or split, those writes to the RecoverPoint appliance allows for data replication with minimal impact to the host.
- **Heterogeneous replication of data volumes —** RecoverPoint is a truly heterogeneous replication solution, enabling the replication of data between different storage array volumes. This enables the ability to migrate from a virtualized environment to a non-virtualized environment or replication between different tiers of storage, thus allowing greater flexibility in the data center.
- **Continuous data protection (CDP) —** Enables a fine granularity of recovery to a block level and refers to local replication of SAN volumes. The fine granularity provides DVR-like point-in-time snapshots of the user volume.
- **Continuous remote replication (CRR) via IP —** Allows data replication across completely separate fabrics over distances limited only by IP technology. The RecoverPoint appliances transfer data from one appliance to another via TCP/IP, greatly extending the distance between source and target volumes.
- **Continuous remote replication (CRR) via FC —** Allows data replication within a fabric over distances limited only by FC distance extension technology.
- **Compression and bandwidth reduction —** By implementing block-level compression algorithms before sending traffic over the IP network, RecoverPoint significantly reduces the network cost required for matching the RPO/RTO expectations.

For more information on RecoverPoint, refer to Chapter 3, ”EMC RecoverPoint.” Chapter 4, ”SAN Connectivity and Interswitch Link Hops,” explains how RecoverPoint can be integrated into the design of simple and complex topologies.

# 2

# EMC VPLEX

This chapter contains the following information on EMC VPLEX:

# VPLEX overview

This section contains the following information:



## Product description

The EMC VPLEX™ family is a solution for federating EMC and non-EMC storage. The VPLEX family is a hardware and software platform that resides between the servers and heterogeneous storage assets supporting a variety of arrays from various vendors. VPLEX can extend data over distance within, between, and across data centers. VPLEX simplifies storage management by allowing LUNs, provisioned from various arrays, to be managed through a centralized management interface.

**Note:** See "VPLEX product offerings" on page 23 for more information.

With a unique scale-up and scale-out architecture, VPLEX's advanced data caching and distributed cache coherency provides workload resiliency, automatic sharing, balancing and failover of storage domains, and enables both local and remote data access with predictable service levels.

**Note:** A VPLEX cabinet can accommodate up to four engines, making it easy to convert from a small configuration to a medium or large cluster configuration.

VPLEX delivers distributed, dynamic, and smart functionality into existing or new data centers to provide storage virtualization across existing boundaries.

- VPLEX is distributed, because it is a single interface for multi-vendor storage and it delivers dynamic data mobility, the ability to move applications and data in real-time, with no outage required.
- VPLEX is dynamic, because it provides data availability and flexibility as well as maintaining business through failures traditionally requiring outages or manual restore procedures.
- VPLEX is smart, because its unique AccessAnywhere™ technology can present and keep the same data consistent within and between sites and enable distributed data collaboration.

## GeoSynchrony

GeoSynchrony™ is the operating system running on VPLEX directors. GeoSynchrony is an intelligent, multitasking, locality-aware operating environment that controls the data flow for virtual storage. GeoSynchrony is:

- Optimized for mobility, availability, and collaboration
- Designed for highly available, robust operation in geographically distributed environments
- Driven by real-time I/O operations
- Intelligent about locality of access
- Provides the global directory that supports AccessAnywhere

GeoSynchrony supports your mobility, availability, and collaboration needs.

EMC VPLEX with GeoSynchrony 5.0 is a scalable, distributed, storage-federation solution that provides non-disruptive, heterogeneous data movement and volume management functionality. GeoSynchrony 5.0 runs on both the VS1 hardware and the VS2 hardware offerings. A VPLEX cluster (both VS1 and VS2) consists of a single-engine, dual-engines, or quad-engines and a management server. Each engine contains two directors. A dual-engine or quad-engine cluster also contains a pair of Fibre Channel switches for communication between directors and a pair of UPS (Uninterruptible Power Sources) for battery power backup of the Fibre Channel switches.

The management server has a public Ethernet port, which provides cluster management services when connected to the customer network.

GeoSynchrony 5.0 provides support for some features already provided by existing array and software packages that might be in use in your storage configuration.

Specifically, GeoSynchrony 5.0 now supports the following features:

- Cross connect

  You can deploy a VPLEX Metro high availability Cross Connect when two sites are within campus distance of each other (up to 1ms round trip time latency) and the sites are running VMware HA and VMware Distributed Resource Scheduler (DRS). You can then deploy A VPLEX Metro distributed volume across the two sites using a cross connect front-end configuration and install a VPLEX Witness server in a different failure domain.

- ALUA

  GeoSynchrony 5.0 supports Asymmetric Logical Unit Access (ALUA), a feature provided by some new active/passive arrays. VPLEX with GeoSynchrony 5.0 can now take advantage of arrays that support ALUA. In active/passive arrays, logical units or LUNs are normally exposed through several array ports on different paths and the characteristics of the paths might be different. ALUA calls these paths characteristics access states. ALUA provides a framework for managing these access states.

  ---

  **Note:** For more information on supported arrays, refer to the *EMC VPLEX with GeoSynchrony 5.0 Product Guide*, located on Powerlink.

  ---

The most important access states are active/optimized and active/non-optimized.

- Active optimized paths usually provide higher bandwidth than active non-optimized paths. Active/optimized paths are paths that go to the service processor of the array that owns the LUN.
- I/O that goes to the active non-optimized ports must be transferred to the service processor that owns the LUN internally. This transfer increases latency and has an impact on the array.

VPLEX is able to detect the active optimized paths and the active/non-optimized paths and performs round robin load balancing across all of the active optimized paths. Because VPLEX is aware of the active/optimized paths, it is able to provide better performance to the LUN.

With *implicit* ALUA, the array is in control of changing the access states of the paths. Therefore, if the controller that owns the LUN being accessed fails, the array changes the status of active/non-optimized ports into active/optimized ports and trespasses the LUN from the failed controller to the other controller.

With *explicit* ALUA, the host (or VPLEX) is able to change the ALUA path states. If the active/optimized path fails, VPLEX causes the active/non-optimized paths to become active/optimized paths and as a result, increase the performance. I/O can go between the controllers to access the LUN through a very fast bus. There is no need to trespass the LUN in this case.

For more information on VPLEX with GeoSynchrony, refer to the *EMC VPLEX with GeoSynchrony 5.0 Product Guide*, located on Powerlink.

## VPLEX advantages

VPLEX delivers unique and differentiated value to address three distinct needs:

- Mobility: The ability to dynamically move applications and data across different compute and storage installations, be they within the same data center, across a campus, within a geographical region and now, with VPLEX Geo, across even greater distances.
- Availability: The ability to create high-availability storage and a compute infrastructure across these same varied geographies with unmatched resiliency.
- Collaboration: The ability to provide efficient real-time data collaboration over distance for such "big data" applications as video, geographic/oceanographic research, and more.

## VPLEX management

VPLEX supports a web-based graphical user interface (GUI) and a command line interface (CLI) for managing your VPLEX implementation. For more information on using these interfaces, refer to the EMC VPLEX Management Console Help or the *EMC VPLEX CLI Guide*, available on Powerlink.

GeoSynchrony supports multiple methods of management and monitoring for the VPLEX cluster:

- Web-based GUI

  For graphical ease of management from a centralized location.
- VPLEX CLI

  For command line management of clusters.
- VPLEX Element Manager API

  Software developers and other users use the API to create scripts to run VPLEX CLI commands.
- SNMP Support for performance statistics:

  Supports retrieval of performance-related statistics as published in the VPLEX-MIB.mib.
- LDAP / AD support

  VPLEX offers Lightweight Directory Access Protocol (LDAP) or Active Directory as an authentication directory service.
- Call home

  The Call Home feature in GeoSynchrony is a leading technology that alerts EMC support personnel of warnings in VPLEX so they can arrange for proactive remote or on-site service. Certain events trigger the Call Home feature. Once a call-home event is triggered, all informational events are blocked from calling home for 8 hours.

## SAN switches

ALL EMC-recommended FC SAN switches are supported, including Brocade, Cisco, and QLogic.

## VPLEX limitations

Always refer to the *VPLEX Simple Support Matrix* for the most up-to-date support information. Refer to the *EMC VPLEX Release Notes*, available on EMC Powerlink, for the most up-to-date capacity limitations.

## VPLEX product offerings

The VPLEX family consists of the following, each briefly described in this section.

- "EMC VPLEX Local" on page 23
- "EMC VPLEX Metro" on page 24
- "VPLEX Metro configurations are formed by connecting two VPLEX clusters in separate locations at synchronous distances over EMC supported Fibre Channel, FCIP, and DWDM / WDM equipment. The distance between sites cannot exceed 5 ms RTT (round-trip time)." on page 24

Figure 1 shows the VPLEX product offerings.

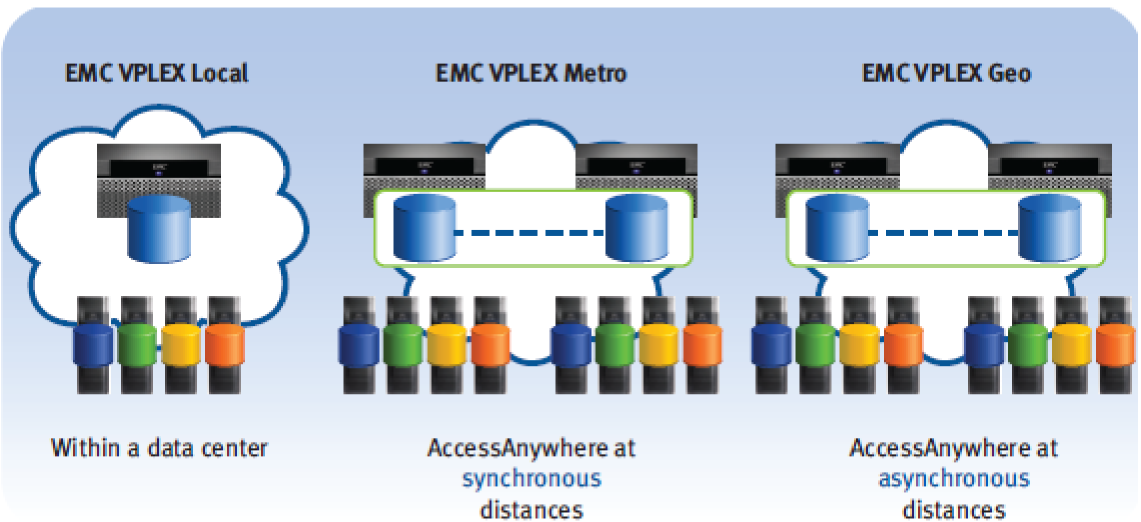


**Figure 1 VPLEX product offerings**

## EMC VPLEX Local

EMC VPLEX Local™:

- Manages data mobility and access within the data center.
- Provides simplified management and nondisruptive data mobility across heterogeneous arrays.

## EMC VPLEX Metro

EMC VPLEX Metro™:

- Delivers distributed federation capabilities and extends access between two locations at synchronous distances.
- Leverages AccessAnywhere to enable a single copy of data to be shared, accessed, and relocated over distance.

VPLEX Metro configurations are formed by connecting two VPLEX clusters in separate locations at synchronous distances over EMC supported Fibre Channel, FCIP, and DWDM / WDM equipment. The distance between sites cannot exceed 5 ms RTT (round-trip time).

### VPLEX Metro topology for DWDM and FCIP

Figure 2 shows an example of a basic VPLEX Metro topology for DWDM.

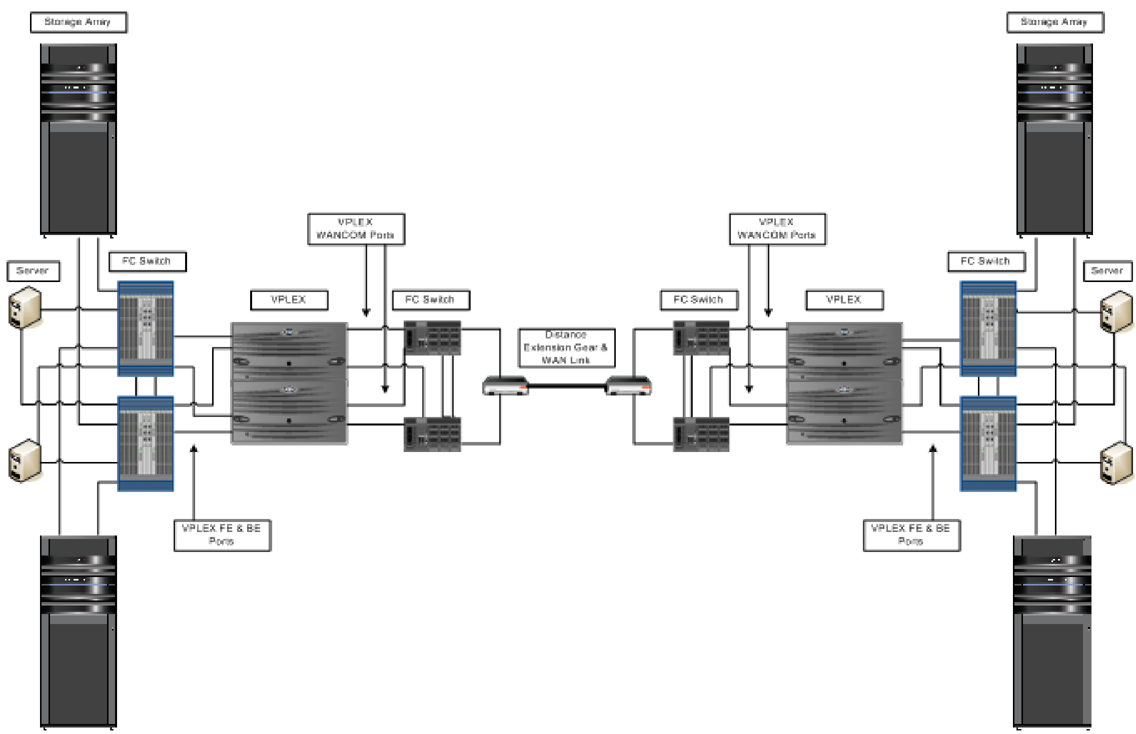


**Figure 2 Basic VPLEX Metro DWDM configuration example**

[Figure 3](#) shows an example of a basic VPLEX Metro topology for FCIP.

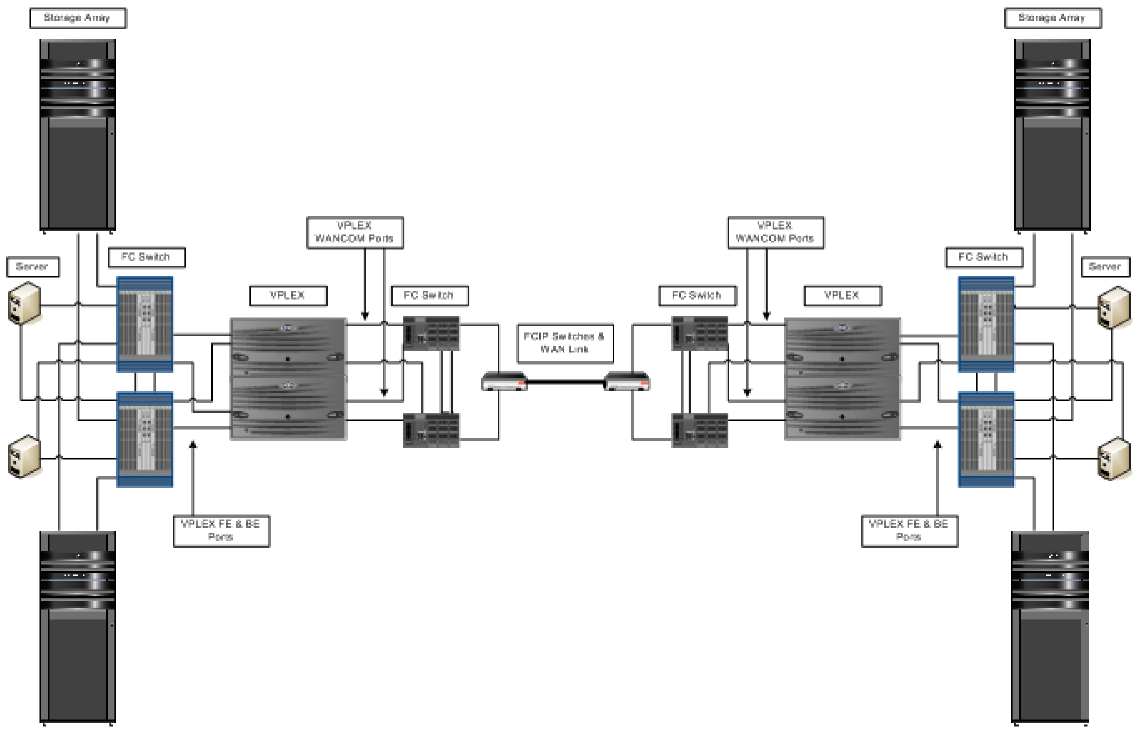


**Figure 3 Basic VPLEX Metro FCIP configuration example**

Refer to the *Extended Distance Technologies TechBook*, available through the E-Lab™ Interoperability Navigator, **Topology Resource Center** tab, at http://elabnavigator.EMC.com, for more information on DWDM and FCIP technologies.

Refer to “VPLEX documentation” on page 35 to locate VPLEX documentation for more detailed information.

## VPLEX Geo

EMC Geo™:

Provides data mobility, high availability, and collaboration between two sites within asynchronous distances.

### VPLEX Geo topology over 1/10 GbE and Distance Extension Gear

Figure 4 shows a VPLEX Geo topology over 1/10 GbE and Distance Extension Gear.

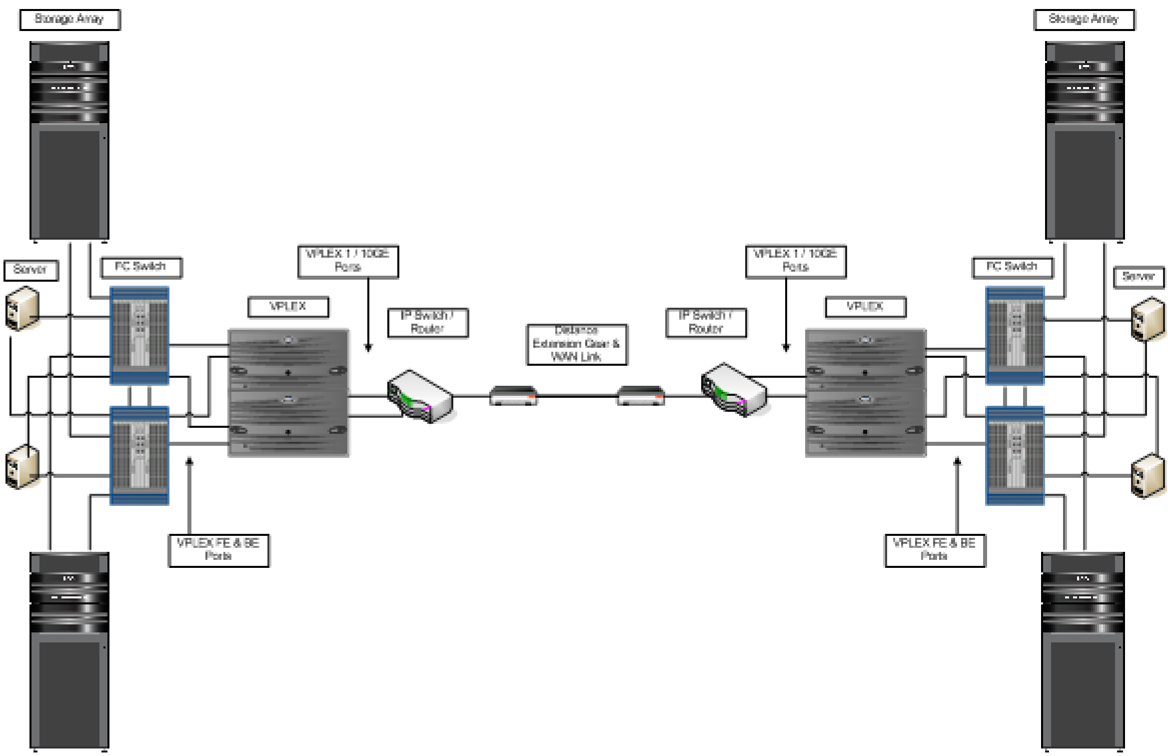


**Figure 4 VPLEX Geo topology over 1/10 GbE and Distance Extension Gear**

# VPLEX family architecture

VPLEX with GeoSynchrony 5.0 is open and heterogeneous, supporting both EMC storage and arrays from other storage vendors, such as HDS, HP, and IBM. VPLEX conforms to established world wide naming (WWN) guidelines that can be used for zoning.

VPLEX provides storage federation for operating systems and applications that support clustered file systems, including both physical and virtual server environments with VMware ESX and Microsoft Hyper-V. VPLEX supports network fabrics from Brocade and Cisco.

An example of the architecture VPLEX with GeoSynchrony 5.0 is shown Figure 5 on page 28.

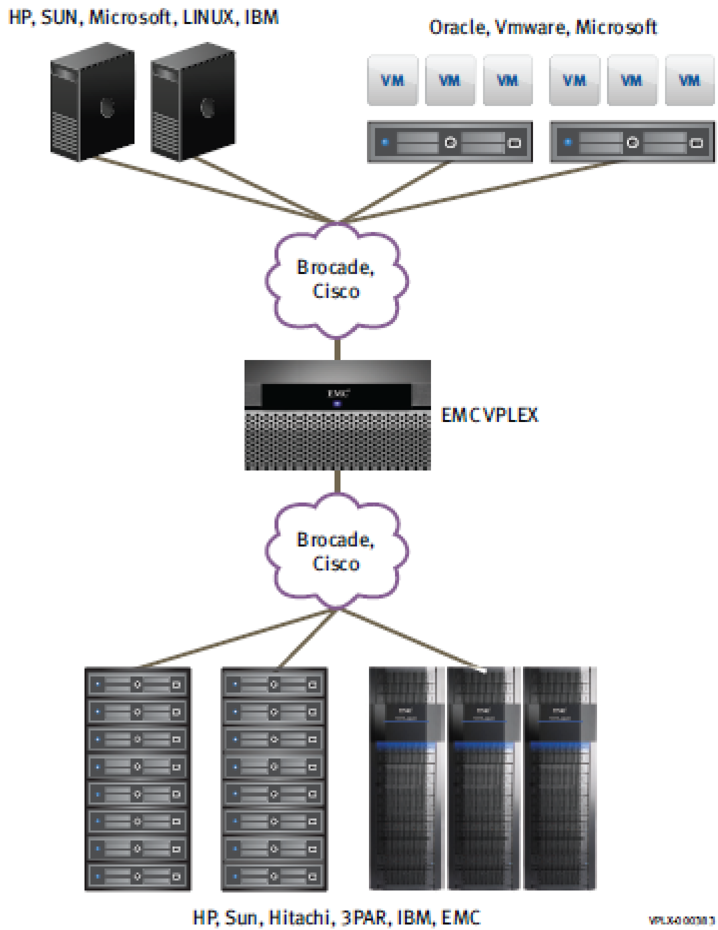


**Figure 5 Architecture highlights**

For more information and interoperability details, refer to the *EMC Simple Support Matrix, EMC VPLEX and GeoSynchrony*, available at http://elabnavigator.EMC.com under the **Simple Support Matrix** tab.

# VPLEX cluster architecture

The VPLEX family uses a unique clustering architecture to help break the boundaries of the data center and allow servers at multiple data centers to have read/write access to shared block storage devices. A VPLEX cluster can scale up through the addition of engines and scale out by connecting clusters into an EMC VPLEX Metro.

All VPLEX clusters are built from a standard engine configuration which is responsible for the federation of the I/O stream. It connects hosts and storage using the Fibre Channel protocol as the transport.

The distributed VPLEX hardware components are connected through both Ethernet or Fibre Channel cabling and respective switching hardware.

I/O modules provide front-end and back-end connectivity between SANs and to remote VPLEX clusters in VPLEX Metro or VPLEX Geo configurations.

## Management server

The management server in each VPLEX cluster provides management services that are accessible from a public IP network.

The management server coordinates event notification, data collection, VPLEX software upgrades, configuration interfaces, diagnostics, and some director-to-director communication.

Both clusters in either VPLEX Metro or VPLEX Geo configuration can be managed from a single management server.

The management server is on a dedicated, internal management IP network that provides accessibility for all major components in the cluster. The management server provides redundant internal management network IP interfaces. In addition to the internal management IP network, each management server connects to the public network, which serves as the access point.

## Fibre Channel switches

The Fibre Channel switches provide high availability and redundant connectivity between directors and engines in a dual-engine or quad-engine cluster. Each Fibre Channel switch is powered by a UPS, and has redundant I/O ports for intra-cluster communication.

The Fibre Channel switches do not connect to the front-end hosts or back-end storage.

To begin using the VPLEX cluster, you must provision and export storage so that hosts and applications can use the storage. For more details, refer to the *VPLEX Installation and Setup Guide*, located at http://Powerlink.EMC.com.

# Cache layering roles

The VPLEX cluster utilizes a write-through mode whereby all writes are written through the cache to the back-end storage. Writes are completed to the host only after they have been completed to the back-end arrays, maintaining data integrity.

All hardware resources such as CPU cycles, I/O ports, and cache memory, are pooled in a VPLEX cluster. Each cluster contributes local storage and cache resources for distributed virtual volumes within a VPLEX cluster. The Data Management (DM) component includes a per-volume caching subsystem that provides the following capabilities:

- Local Node Cache: I/O management, replacement policies, pre-fetch, and flushing (CCH) capabilities
- Distributed Cache (DMG—Directory Manager): Cache coherence, volume share group membership (distributed registration), failure recovery mechanics (fault-tolerance), RAID, and replication capabilities

## Share groups

Nodes export the same volume from a share group. This share group membership is managed through a distributed registration mechanism. Nodes within a share group collaborate to maintain cache coherence.

## Cache coherency

Cache Coherency creates a consistent global view of a volume.

Distributed Cache Coherency is maintained using a directory. There is one directory per user volume and each directory is split into chunks. These chunks exist only if they are populated. There is one directory entry per global cache page, with responsibility for:

- Tracking page owner(s) and remembering the last writer
- Locking and queuing

## Directory chunks

Directory chunks are managed by the meta-directory, which assigns and remembers chunk ownership. These chunks can migrate using Locality-Conscious Directory Migration (LCDM). This meta-directory knowledge is cached across the share group for efficiency.

## Basic read request

When a host makes a read request, VPLEX first searches its local cache. If the data is found there, it is returned to the host. If the data is not found in local cache, VPLEX searches global cache. Global cache includes all directors that are connected to one another within the VPLEX cluster. When the read is serviced from global cache, a copy is also stored in the local cache of the director from where the request originated. If a read cannot be serviced from either local or global cache, it is read directly from the back-end storage. In this case both local and global cache are updated to maintain cache coherency.

## Basic write request

When performing writes, the VPLEX system DM (Data Management) component includes a per-volume caching subsystem that utilizes a subset of the caching capabilities:

- Local Node Cache: Cache data management and back-end I/O interaction
- Distributed Cache (DMG—Directory Manager): Cache coherency, dirty data protection, and failure recovery mechanics (fault-tolerance)

A VPLEX Cluster utilizes a write-through mode whereby all writes are written through cache to the back-end storage array. Writes are completed to the host only after they have been completed to the back-end array, maintaining data integrity.

# VPLEX configuration overview

Refer to "VPLEX documentation" on page 35 to locate VPLEX documentation for more detailed information on single-, dual-, and quad-engine configurations for VPLEX with GeoSynchrony v4.x.

This section describes only the VS2 hardware for VPLEX clusters. If you are currently running on a VS1 system, see the *VPLEX with GeoSynchrony 4.2 Installation and Setup Guide* available in the Procedure Generator for a description of the VS1 hardware.

Figure 6 shows the main hardware components in a single, dual, and quad-engine VPLEX cluster.

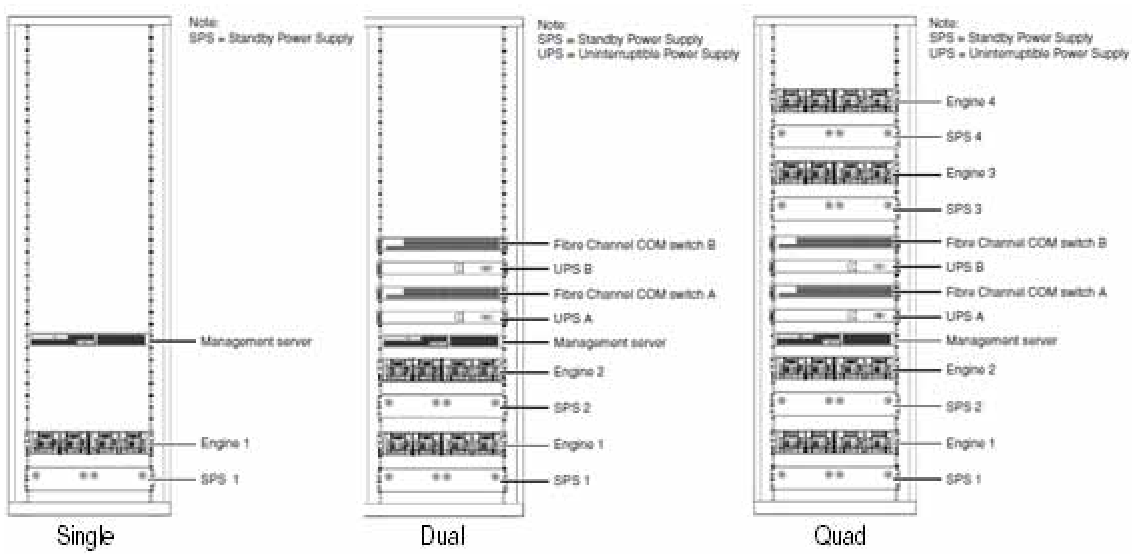


**Figure 6** **VPLEX cluster engines with GeoSynchrony v5.x**

For more information, refer to the *EMC VPLEX with GeoSynchrony 5.0 Product Guide*, located on Powerlink.

## VPLEX hardware and software components

This section describes only the VS2 hardware for VPLEX clusters. If you are currently running on a VS1 system, see the *VPLEX with GeoSynchrony 4.2 Installation and Setup Guide* available in the Procedure Generator for a description of the VS1 hardware.

The VPLEX VS2 engine for VPLEX with GeoSynchrony v5.x contains two directors, with each providing front-end and back-end I/O connections. Each of these module types are described in more detail in "The VPLEX director" section of the *EMC VPLEX with GeoSynchrony 5.0 Product Guide*, located on Powerlink.

Figure 7 shows a VPLEX engine, rear view for VPLEX with GeoSynchrony v5.x.

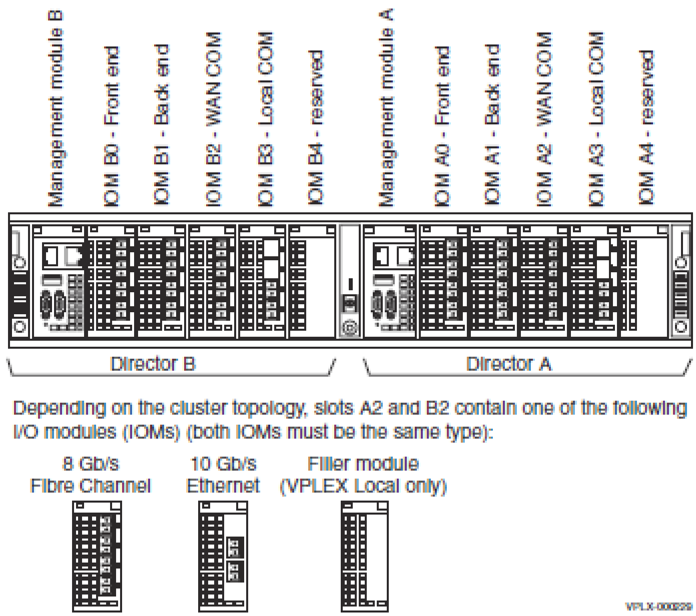


**Figure 7 VPLEX engine, rear view for VPLEX with GeoSynchrony v5.x**

For hardware and software information, refer to the VPLEX online help, available on the Management Console GUI, for 4.x and the *EMC VPLEX with GeoSynchrony 5.0 Product Guide*, located on Powerlink for more detailed information on hardware and software.

# VPLEX documentation

There are numerous VPLEX documentation, technical notes, and white papers available, some listed below, which can be found at http://Powerlink.EMC.com, including:

- *EMC VPLEX with GeoSynchrony 5.0 Product Guide*
- *EMC VPLEX with GeoSynchrony 5.0 CLI Guide*
- *EMC VPLEX with GeoSynchrony 5.0 Configuration Guide*
- *EMC VPLEX Hardware Installation Guide*
- *EMC VPLEX Release Notes*
- *Implementation and Planning Best Practices for EMC VPLEX Technical Notes*
- VPLEX online help, available on the Management Console GUI
- VPLEX Procedure Generator, available on EMC Powerlink
- *EMC Simple Support Matrix, EMC VPLEX and GeoSynchrony*, available at http://elabnavigator.EMC.com under the **Simple Support Matrix** tab.

For host-specific VPLEX information, see the appropriate Host Connectivity Guides available at http://Powerlink.EMC.com

For the most up-to-date support information, you should always refer to the *EMC Support Matrix*, **PDFs and Guides** > **VPLEX**.

# 3

# EMC RecoverPoint

This chapter contains the following information on EMC RecoverPoint:



**Note:** For detailed information on RecoverPoint with Fibre Channel over Ethernet (FCoE), refer to the “Solutions” chapter in the *Fibre Channel over Ethernet (FCoE): Data Center Bridging (DCB) Concepts and Protocols TechBook*, available through the E-Lab Interoperability Navigator, **Topology Resource Center** tab, at http://elabnavigator.EMC.com.

# RecoverPoint architecture

RecoverPoint uses a combination of hardware appliances configured out of the direct data path “out of band” and a splitter driver to replicate data from the designated source volume to the configured target volume. RecoverPoint uses intelligent Fibre Channel switches to intercept and “split” incoming writes. The intelligent Fibre Channel switch then sends a copy of the incoming write to the RecoverPoint Appliance as well as passing the original write to the original destination. The RecoverPoint appliance then handles the sending of the “split” write to the proper destination volume, either through TCP/IP or Fibre Channel.

Traffic routed by the intelligent switch does not introduce protocol overhead. All virtualized I/O is routed through ASICs on the intelligent switch or blade. The payload (SCSI data frame) is not affected by this operation, as the ASIC does not read or modify the content but rather reads only the protocol header information to make routing decisions.

Figure 8 shows an example of RecoverPoint architecture.

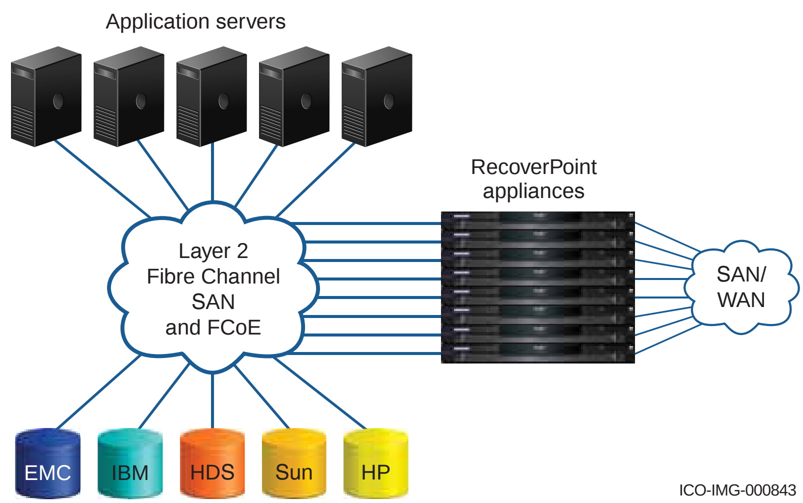


**Figure 8 RecoverPoint architecture example**

The splitter residing in the intelligent switch is responsible for duplicating the write traffic and sending the additional write to the RecoverPoint appliance, as shown in Figure 9.

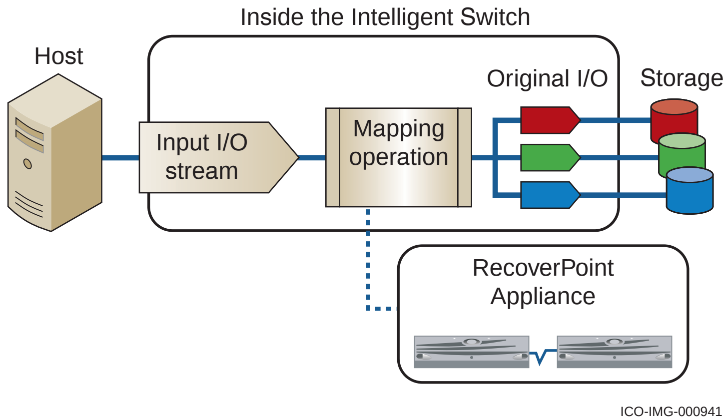


**Figure 9 Intelligent switch running RecoverPoint**

RecoverPoint offers:

- Continuous remote replication (CRR) between LUNs across two SANs
- Continuous data protection (CDP) between LUNs on the same SAN
- Concurrent local and remote (CLR) data protection between LUNs across two SANs
- Heterogeneous operating system and storage array support
- The server's I/O adapters can be FC (HBA) or FCoE (CNA)
- The storage array can be FC- or FCoE-connected, depending on the splitter strategy
- Integrated with intelligent fabric

## RecoverPoint components

There are five major components in a RecoverPoint installation:

- **Virtual Target (VT)** — This is the virtual representation of the physical storage that is used as either source or destination for replicated data.
- **Virtual Initiator (VI)** — This is the virtual representation of the physical host that is issuing I/O to the source volume.
- **Intelligent Fabric Splitter** — This is the intelligent-switch hardware that contains specialized port-level processors (ASICs) to perform virtualization operations on I/O at line speed. This functionality is available from two vendors: Brocade and Cisco.

  Brocade resides on the intelligent blade, AP4-18, which can be installed in a Brocade DCX or Brocade 48000 chassis. It can also be installed directly on the intelligent switch, AP-7600.

  Cisco resides on the intelligent blades Storage Services Module (SSM) or Multiservice Module, which can be installed in an MDS 9513, 9509, 9506, 9216i, 9216A, or 9222i. It can also be installed directly on an MDS 9222i without an intelligent blade
- **RecoverPoint Appliances (RPA)** — These appliances are Linux-based boxes that accept the “split” data and route the data to the appropriate destination volume, either using IP or Fibre Channel. The RPA also acts as the sole management interface to the RecoverPoint installation.
- **Remote Replication** — Two RecoverPoint Appliance (RPA) clusters can be connected via TCP/IP or FC in order to perform replication to a remote location. RPA clusters connected via TCP/IP for remote communication will transfer "split" data via IP to the remote cluster. The target cluster's distance from the source is only limited by the physical limitations of TCP/IP.

  RPA clusters can also be connected remotely via Fibre Channel. They can reside on the same fabric or on different fabrics, as long as the two clusters can be zoned together. The target cluster's distance from the source is again only limited by the physical limitations of FC. RPA clusters can support distance extension hardware (such as, DWDM) to extend the distance between clusters.

# Logical topology

There are *Virtual Initiators* (VI) and *Virtual Targets* (VT). A host (initiator) or storage (target) is physically connected to the fabric. The VI and VT are logical entities that are accessible from any port on the chassis or in the fabric, as shown in Figure 10.

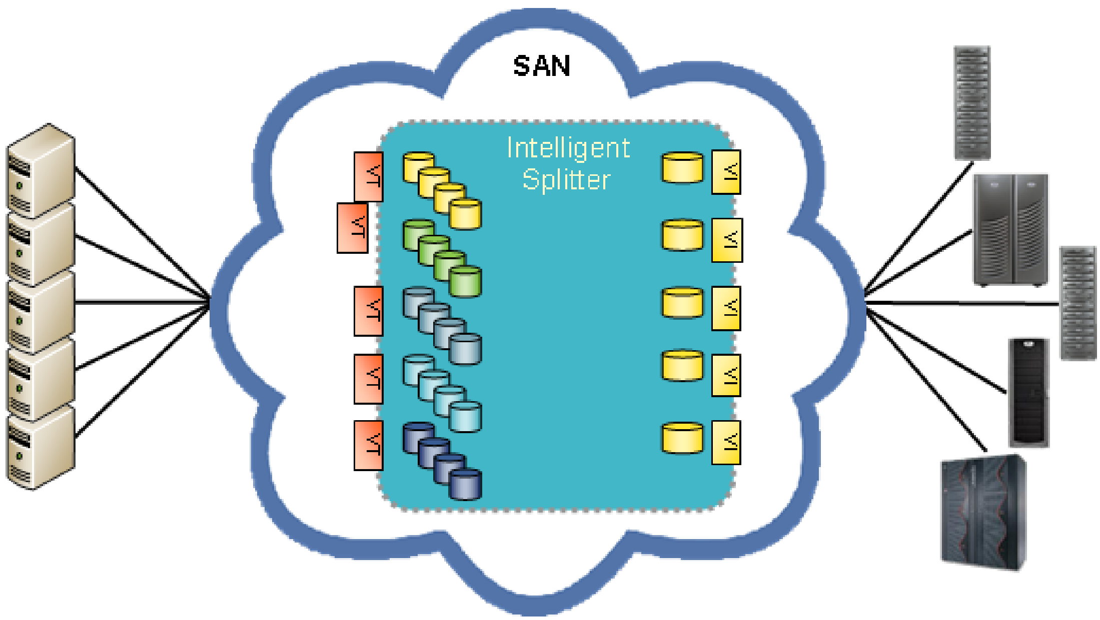


**Figure 10** **Logical topology**

The exact method of this virtualization differs between the Cisco SANTap solution and the Brocade FAP solution.

The SANTap architecture uses VSANs to physically separate the physical host from the physical storage and bridges the VSANs via virtualization internal to the intelligent switch. When a virtual target (DVT) is created, a duplicate WWN is created and placed in the "Front End" VSAN, that is the VSAN that encompasses the physical hosts. The physical hosts then communicate with the virtual target. SANTap then also creates virtual initiators (VI) and duplicates the host WWNs and places them in the "Back End" VSAN, which is the VSAN that encompasses the physical storage. The VIs then communicate to the physical storage. SANTap also creates virtual entries representing the splitter itself and communication between the RPAs and the intelligent switch happens via these virtual entries.

Figure 11 on page 42 shows example “Back End” and “Front End” VSANs. and shows the duplicate WWNs in both VSANs.

```
Example "Back End" VSAN
VSAN 150:
-------------------------------------------------------------------------
FCID        TYPE  PWWN                     (VENDOR)        FC4-TYPE:FEATURE
-------------------------------------------------------------------------
0x930009    N     50:06:04:82:d5:2e:87:52  (EMC)           scsi-fcp:both 253 ← Physical Storage
0x93000b    N     22:35:00:05:30:00:3d:e0  (Cisco)         scsi-fcp:init vir..t
0x93000c    N     22:36:00:05:30:00:3d:e0  (Cisco)         scsi-fcp:init vir..t
0x93000d    N     22:37:00:05:30:00:3d:e0  (Cisco)         scsi-fcp:init vir..t
0x93000e    N     22:38:00:05:30:00:3d:e0  (Cisco)         scsi-fcp:init vir..t
0x93000f    N     22:39:00:05:30:00:3d:e0  (Cisco)         scsi-fcp:init vir..t
0x930010    N     22:3a:00:05:30:00:3d:e0  (Cisco)         scsi-fcp:init vir..t
0x930011    N     22:3b:00:05:30:00:3d:e0  (Cisco)         scsi-fcp:init vir..t
0x930012    N     22:3c:00:05:30:00:3d:e0  (Cisco)         scsi-fcp:init vir..t
0x930013    N     22:3e:00:05:30:00:3d:e0  (Cisco)         scsi-fcp:init vir..t
0x930200    N     21:01:00:e0:8b:a6:fd:cc  (Qlogic)        scsi-fcp:init vir..t ← Virtual Initiator

Corresponding "Front End" VSAN
VSAN 151:
-------------------------------------------------------------------------
FCID        TYPE  PWWN                     (VENDOR)        FC4-TYPE:FEATURE
-------------------------------------------------------------------------
0x3a0100    N     50:06:04:82:d5:2e:87:52  (EMC)           scsi-fcp:target vi.. ← Virtual Storage (DVT)
0x3a0200    N     21:01:00:e0:8b:a6:fd:cc  (Qlogic)        scsi-fcp:init ← Physical Initiator
```

**Figure 11 Example Back End and Front End VSAN**

The Brocade FAP architecture uses Frame Redirect to channel all host I/O through the corresponding bound VI. When a host is bound to a particular physical target, a virtual target is created with a new WWN. The separation in this case is handled internally by the switch. The physical target remains zoned with the physical storage. Data would flow from the physical host to the virtual target, then from the bound VI to the physical storage.

# Local and remote replication

Replication by RecoverPoint is based on a logical entity called a *consistency group*. SAN-attached storage volumes at the primary and secondary sites, called *replication volumes*, are assigned to a consistency group to define the set of data to be replicated. An application, such as Microsoft Exchange, typically has its storage resources defined in a single consistency group so there is a mapping between an application and a consistency group.

EMC RecoverPoint replicates data over any distance:

- Within the same site or to a local bunker site some distance away

  **Note:** Refer to "CDP configurations" in the *EMC RecoverPoint Administrator's Guide,* located on Powerlink.

- To a remote site, see CRR configurations

  **Note:** Refer to "CRR configurations" in the *EMC RecoverPoint Administrator's Guide,* located on Powerlink.

- To both a local and a remote site simultaneously

  **Note:** Refer to "CLR configurations" in the *MC RecoverPoint Administrator's Guide,* located on Powerlink.

## RecoverPoint replication phases

The three major phases performed by RecoverPoint to guarantee data consistency and availability (RTO and RPO) during replication are:

- Write phase (the splitting phase)
- Transfer phase
- Distribution phase

## Intelligent switches

There are several intelligent switches that provide fabric-level I/O splitting for the RecoverPoint instance:

- Connectrix AP-7600
- Connectrix ED-48000B
- Connectrix MDS Storage Services Module (SSM)

For more information on the intelligent switches, refer to the product data information in the *EMC Connectrix Products Data TechBook*, through the E-Lab Interoperability Navigator, **Topology Resource Center** tab, at http://elabnavigator.EMC.com.

# EMC RecoverPoint with intelligent switch splitters

EMC RecoverPoint is an enterprise-scale solution designed to protect application data on heterogeneous SAN-attached servers and storage arrays. RecoverPoint runs on an out-of-path appliance and combines industry-leading continuous data protection technology with a bandwidth-efficient no data loss replication technology that protects data on local and remote sites. EMC RecoverPoint provides local and remote data protection, enabling reliable replication of data over any distance; that is, locally within the same site or remotely to another site using Fibre Channel or an IP network to send the data over a WAN. For long distances, RecoverPoint uses either Fibre Channel or IP network to send data over a WAN.

RecoverPoint uses existing Fibre Channel and Fibre Channel over Ethernet (FCoE) infrastructure to integrate seamlessly with existing host application and data storage subsystems. For detailed information on RecoverPoint with Fibre Channel over Ethernet (FCoE), refer to the “Solutions” chapter in the *Fibre Channel over Ethernet (FCoE): Data Center Bridging (DCB) Concepts and Protocols TechBook*, available through the E-Lab Interoperability Navigator, **Topology Resource Center** tab, at http://elabnavigator.EMC.com.

RecoverPoint's fabric splitter leverages intelligent SAN-switch hardware from EMC's Connectrix partners. This section focuses on topology considerations where the data splitter resides on an intelligent SAN switch.

To locate more detailed information about RecoverPoint, refer to “RecoverPoint documentation” on page 46.

## RecoverPoint documentation

There are a number of documents for RecoverPoint related information, all which can be found at http://powerlink.emc.com, including:

- Release notes

  Information in the release notes include:

  - Support parameters
  - Supported configurations
  - New features and functions
  - SAN compatibility
  - Bug fixes
  - Expected behaviors
  - Technical notes
  - Documentation issues
  - Upgrade information
- *EMC RecoverPoint Administrator's Guide*
- RecoverPoint maintenance and administration documents
- RecoverPoint installation and configuration documents
- RecoverPoint-related white papers
- *EMC Support Matrix*, located at http://elabnavigator.emc.com, PDFs and Guides >RecoverPoint

For detailed information on RecoverPoint with Fibre Channel over Ethernet (FCoE), refer to the "Solutions" chapter in the *Fibre Channel over Ethernet (FCoE): Data Center Bridging (DCB) Concepts and Protocols TechBook*, available through the E-Lab Interoperability Navigator, **Topology Resource Center** tab, at http://elabnavigator.EMC.com.

# 4

# SAN Connectivity and Interswitch Link Hops

This chapter includes SAN connectivity information and describes how RecoverPoint can be integrated into the design of a simple or complex topology.

## SAN connectivity overview

RecoverPoint use intelligent switches, which allows for seamless connectivity into a Fibre Channel switched fabric.

There are a number of different SAN topologies, classified in two categories: simple and complex topologies, which this chapter will further discuss.

RecoverPoint can be integrated into the design of either topology type.

- "Simple SAN topologies" on page 49

  **Note:** For more detailed information on simple SAN topologies, refer to the "Simple Fibre Channel SAN topologies" chapter in the *Fibre Channel SAN Technologies TechBook*, available through the E-Lab Interoperability Navigator, **Topology Resource Center** tab, at http://elabnavigator.EMC.com.

- "Complex SAN topologies" on page 54

  **Note:** For more detailed information on complex SAN topologies, refer to the "Complex Fibre Channel SAN topologies" chapter in the *Fibre Channel SAN Technologies TechBook*, available through the E-Lab Interoperability Navigator, **Topology Resource Center** tab, at http://elabnavigator.EMC.com.

# Simple SAN topologies

RecoverPoint is dependent on the intelligent switches to handle data splitting to the RecoverPoint Appliance.

The intelligent switches can act as stand-alone FC switches as well as Layer 2 FC switches. In small SANs, or with small virtualization implementations, this is one possible topology. There are a few configurations based on the choice in FC switch vendors, each discussed further in this section:


- "Connectrix B AP-7600B simple SAN topology" on page 49
- "Storage Services Module (SSM) blade in a simple SAN topology" on page 50
- "Node placement in a simple SAN topology—Connectrix B" on page 52
- "Node placement in a simple SAN topology—Connectrix MDS" on page 53


## Connectrix B AP-7600B simple SAN topology

The Connectrix B AP-7600B, when operating as a platform for the RecoverPoint instance, supports F_Port connections. That is, FC-SW nodes, such as hosts and storage, can be connected directly to any one of the 16 ports on the AP-7600B, as shown in Figure 12 on page 50. There is no need to connect through a Layer 2 switch, such as an ED-48000B or DS-4900B, for example.

**Note:** Figure 12 does not show the CPCs or the AT switches. It only displays the FC connectivity.

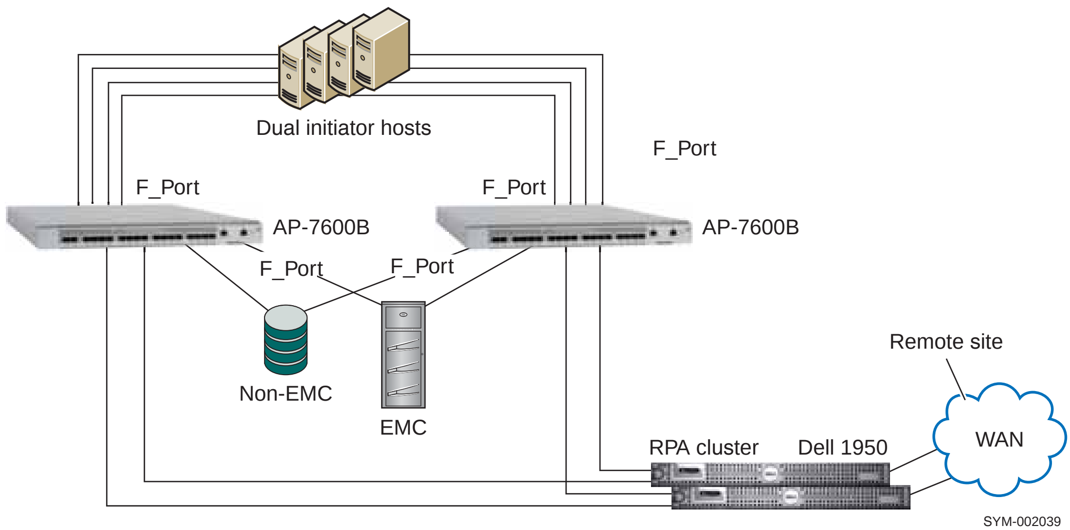


**Figure 12** **Connectrix B AP-7600B simple SAN topology**

## Storage Services Module (SSM) blade in a simple SAN topology

The Connectrix MDS Storage Services Module (SSM) can be installed in any MDS switch or director with open slots. This includes the MDS 9513, 9509, 9506, and 9216. In Figure 13 on page 51, hosts are connected to the SSM ports and storage is connected to the fixed ports on the MDS 9216i. Hosts may be also be connected to the fixed ports; however, since the SSM ports are oversubscribed, they are best suited for host connectivity.

Since the fixed ports are line rate mode (2 Gb on the MDS 9216), storage ports should be connected there to ensure maximum bandwidth.

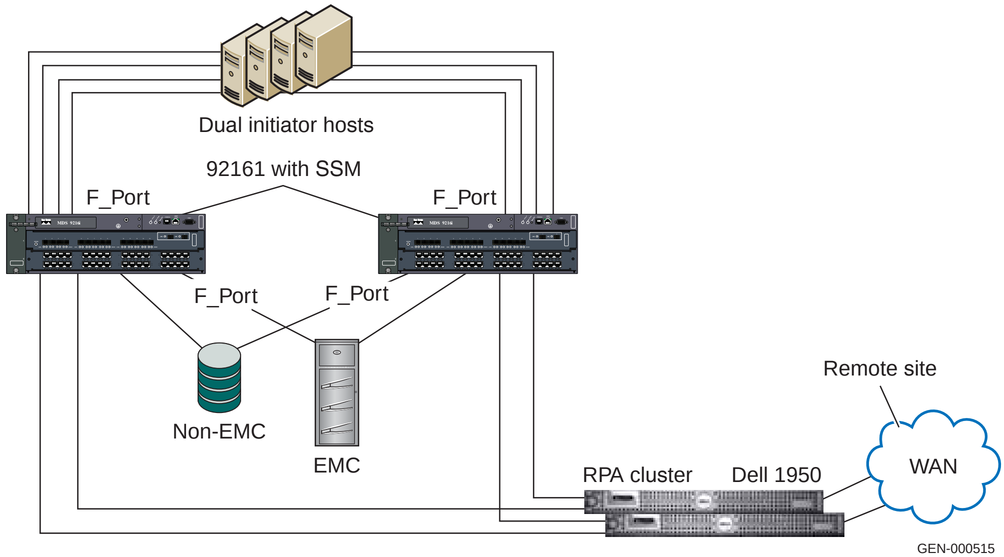


**Figure 13** **Connectrix MDS SSM blade in simple SAN topology**

Figure 14 on page 52 shows an MDS 9513 with a 24-port line card and an SSM blade. Hosts are connected to the SSM ports and storage is connected to the 24-port line card. The SSM ports are oversubscribed so they are best suited for host connectivity. The 24-port line card ports are line rate mode (4 Gb), so storage ports should be connected there to ensure maximum bandwidth.

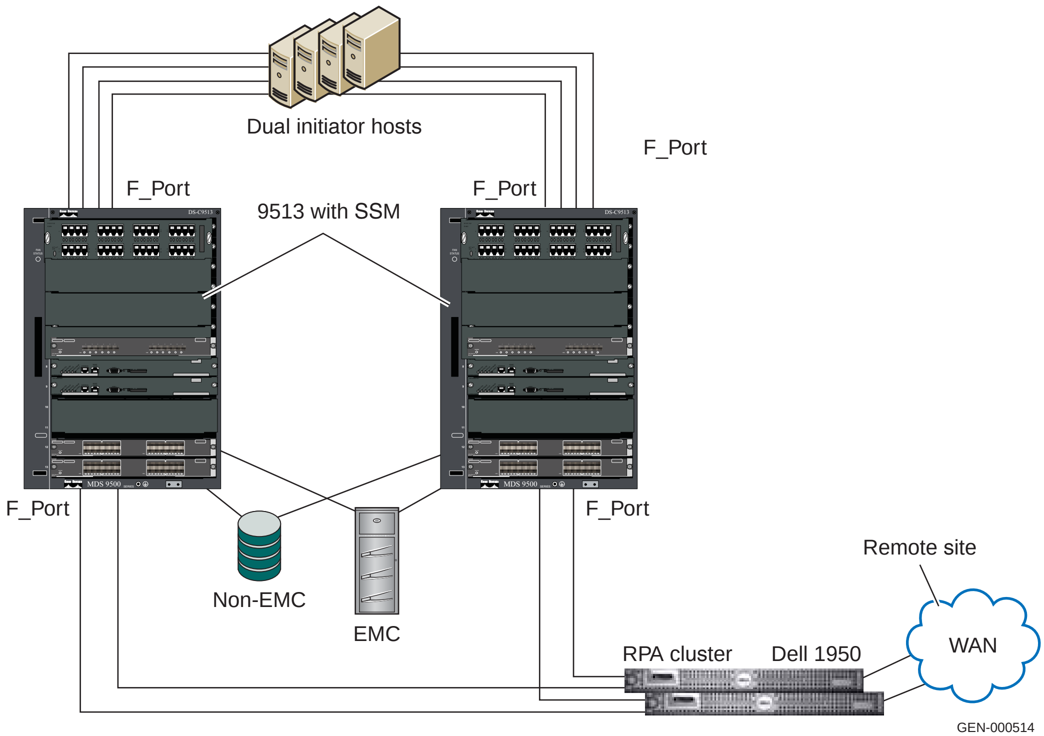


**Figure 14** **MDS 9513 with a 24-port line card and an SSM blade**

**Note:** Figure 13 on page 51 and Figure 14 do not show the CPCs or the AT switches. They only display the FC connectivity.

## Node placement in a simple SAN topology—Connectrix B

The AP-7600B, when used as a single switch, may have both hosts and storage nodes connected directly to any of the ports. It is unlikely that the AP-7600B will be used in this manner because it does not make effective use of the available ports. Since there are only 16 ports, only a mixture of 16 hosts and storage ports are possible. It is more likely that this intelligent switch will be used with more complex designs.

## Node placement in a simple SAN topology—Connectrix MDS

The Connectrix MDS single switch / director model is sometimes known as a collapsed core / edge design. There is a mixture of line rate mode (LRM) and oversubscribed mode ports (OSM). The external ports on an SSM blade are oversubscribed (same ratio as the standard 32-port line card). The Cisco switch / director can support virtualized and non-virtualized storage simultaneously. Hosts may be connected to any port in the chassis. Storage may be connected to any port in the chassis.

**Note:** Hosts accessing virtualized storage, and physical storage supplying RecoverPoint, do *not* have to be connected to the SSM.

The same best practice holds true for host and storage connectivity with or without RecoverPoint. Hosts should be connected to OSM ports whenever possible unless they have high I/O throughput requirements. Storage should be connected to LRM ports whenever possible because a storage port is a shared resource with multiple hosts requiring access.

If there are hosts using solely virtualized storage, then they require the SSM. Therefore, because a dependency exists, connect the host (whenever possible) directly to the SSM. As previously stated, host and storage nodes may be connected to any port on the switch / director.

# Complex SAN topologies

Complex SAN topologies derive their complexity from several possible conditions:

- Domain count
- Heterogeneous switch vendors in a single SAN infrastructure
- Distance extension
- Multi-protocol configurations

For the purposes of RecoverPoint, only the first two conditions will be discussed in this section.

## Domain count

SANs come in various sizes and configurations. Customers with hundreds of host and storage ports require several FC switches to be linked together through ISLs. Although the size of director class FC switches has significantly increased, core/edge topologies (as discussed in "Compound core edge switches" section in the *Fibre Channel SAN Technologies TechBook*, available through the E-Lab Interoperability Navigator, **Topology Resource Center** tab, at http://elabnavigator.EMC.com) are common because legacy switches are used in combination with the new high-density directors. These switches can be all departmental switches, all directors, or a combination of both.

A topology with numerous switches will most likely be in one of two designs, each discussed in the next sections:

- Core/edge (hosts on an edge switch and storage on a core switch). (See Figure 15 on page 56.)
- Edge/core/edge (hosts and storage on an edge switch but on different tiers). (See Figure 16 on page 57.)

## Core/edge

Figure 15 on page 56 shows a core/edge design. In this model, hosts are connected to the MDS 9506s (edge tier) and storage is connected to MDS 9513s (core tier). The core tier is the centralized location and connection point for all physical storage in this model. All I/O between the host and storage must flow over the ISLs between the edge and core. It is a one-hop logical infrastructure. (An ISL hop is the link between two switches.)

The SSM blade is located in the core. To minimize latency and increase fabric efficiency, the SSM should be co-located with the storage that it is virtualizing, just as one would do in a Layer 2 SAN.

**Note:** Physical placement of the RecoverPoint Appliances can be anywhere within the fabric and need not be connected directly to the intelligent switch.

**Note:** Direct connections to the SSM are *not* required.

- ISLs between the switches do *not* have to be connected to the SSM blade
- Hosts do *not* have to be connected to the SSM blade
- Storage should *not* be connected to the SSM blade

Since the SSM is located inside the MDS 9513 core, internal routing between blades in the chassis is *not* considered an ISL hop. There is additional latency with the virtualization ASICs; however, there is no protocol overhead associated with routing between blades in a chassis.

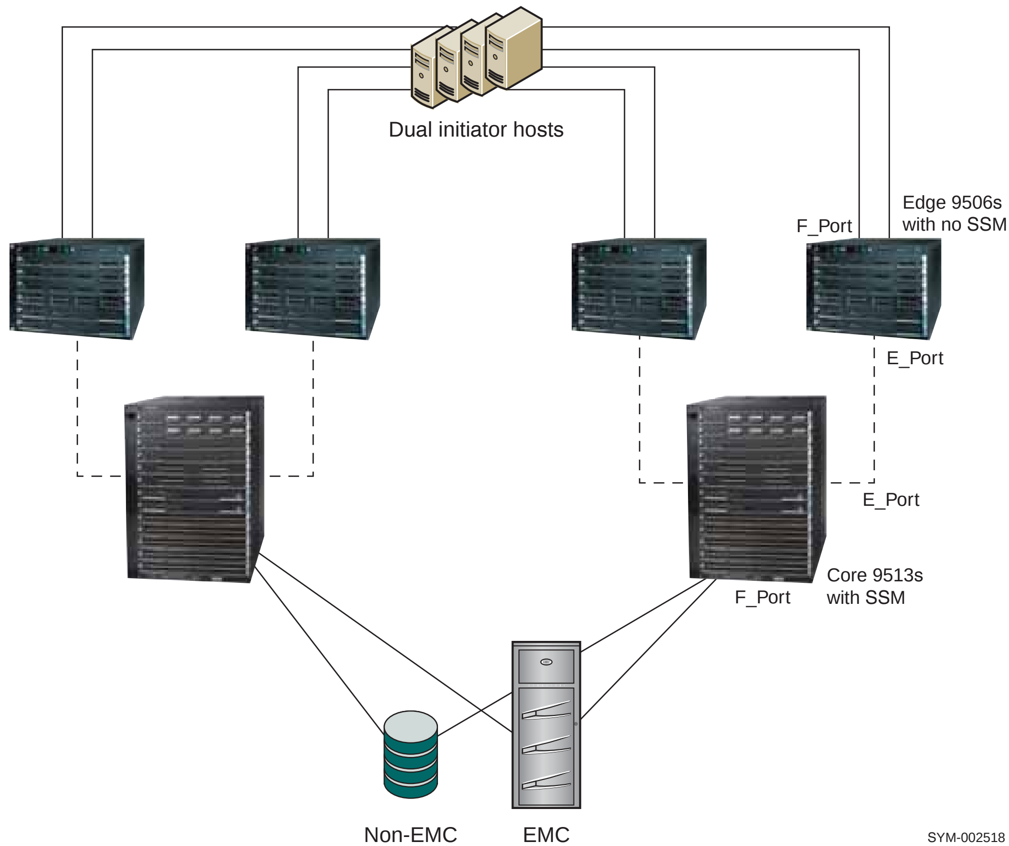


**Figure 15 Core/edge design**

Figure 16 on page 57 shows a core/edge design with a Connectrix B or RecoverPoint implementation. The AP-7600B is an external intelligent switch; therefore, it must be linked through ISLs to the core switch.

---

**Note:** Physical placement of the RecoverPoint Appliances can be anywhere within the fabric and need not be connected directly to the intelligent switch.

---

In this model, hosts are connected to the DS-32B2 (edge tier) and storage is connected to ED-48000B (core tier). The core tier is the centralized location and connection point for all physical storage in this model. All I/O between the host and storage must flow over the ISLs between the edge and core. It is a one-hop infrastructure for non-virtualized storage. However, all virtualized storage traffic must pass through at least one of the ports on the AP-7600B. Therefore, the I/O from the host must traverse an ISL between the DS-32B2 and the ED-48000B. Then, it must traverse an ISL between the ED-48000B and the AP-7600B. Finally, it must traverse an ISL back from the AP-7600B to the ED-48000B where the I/O is terminated. This is a three-hop logical topology.

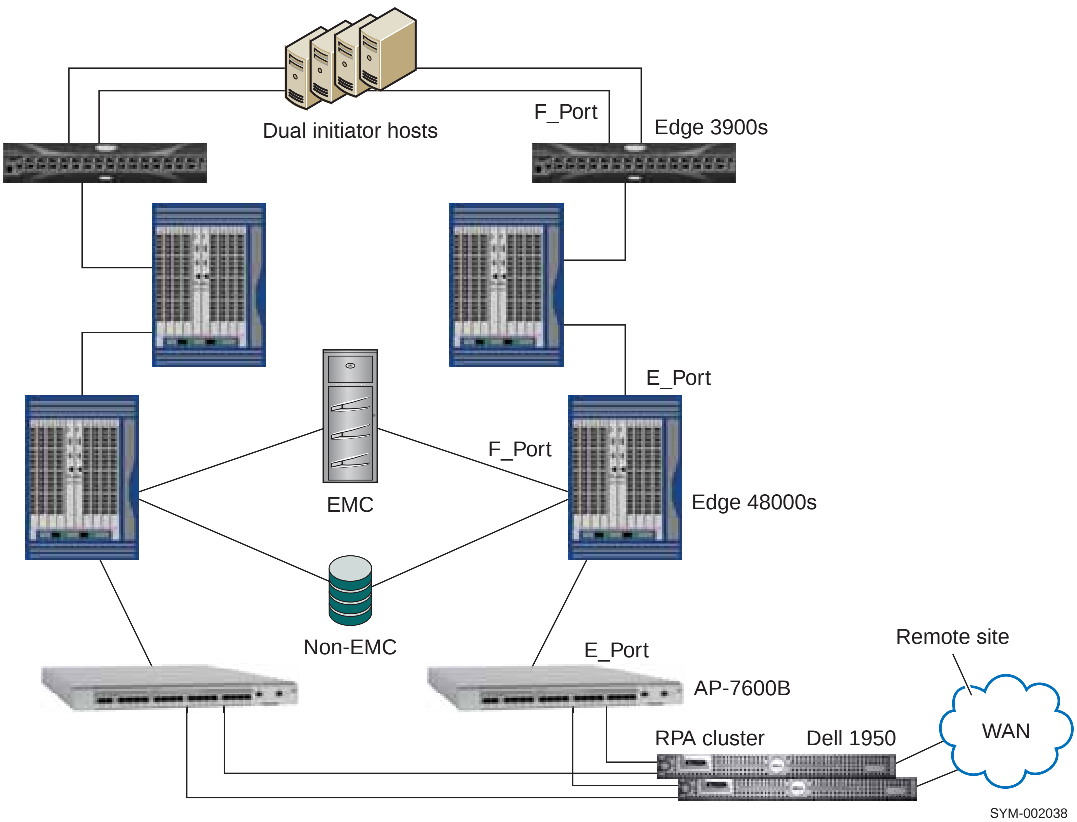


**Figure 16** **Core/edge design with a Connectrix B**

## Edge/core/edge

Figure 17 on page 59 shows an edge/core/edge design. In this model, hosts are connected to the Connectrix MDS 9506s (edge tier) and storage is connected to MDS 9506 (storage tier). The MDS 9513s act as a connectivity layer. The core tier is the centralized location and connection point for edge and storage tiers. All I/O between the host and storage must flow over the ISLs between. This is a two-hop logical topology.

Within the fabric, hosts can access virtualized and non-virtualized storage. The location of the physical storage to be virtualized will determine the location of the SSM blade. There are two possibilities:

- Storage to be virtualized is located on a single switch
- Storage to be virtualized is located on multiple switches on the storage tier

**Note:** Physical placement of the RecoverPoint Appliances can be anywhere within the fabric and need not be connected directly to the intelligent switch.

### Storage on a single switch

If the physical storage is located on one edge switch, place the SSM on the same switch. See the red box in Figure 17 on page 59.

To minimize latency and increase fabric efficiency, the SSM is co-located with the storage that it is virtualizing.

**Note:** Connections to the SSM are *not* required.

- ISLs between the switches do not have to be connected to the SSM blade
- Hosts do not have to be connected to the SSM blade
- Storage should *not* be connected to the SSM blade

Since the SSM is located inside the edge MDS 9506, internal routing between blades in the chassis is *not* considered an ISL hop. There is additional latency with the virtualization ASICs; however, there is no protocol overhead associated with routing between blades in a chassis.

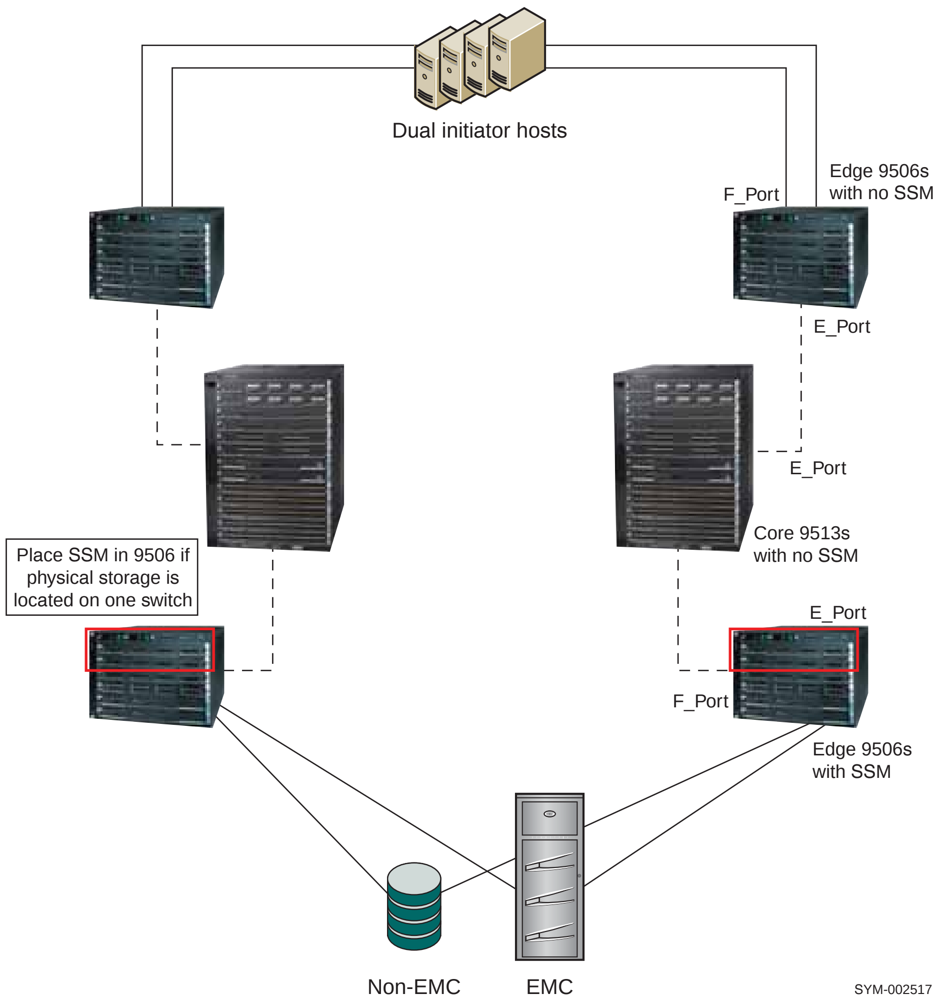


**Figure 17** **Storage on a single switch**

Figure 18 on page 60 shows a core/edge/core design with Connectrix B. The AP-7600Bs are linked via ISLs to the edge ED-48000B.

---

**Note:** Physical placement of the RecoverPoint Appliances can be anywhere within the fabric and need not be connected directly to the intelligent switch.

---

In this model, hosts are connected to the DS-32B2 (edge tier) and storage is connected to a ED-48000B, which is the other edge tier. The core ED-48000Bs are for connectivity only. All I/O between the host and storage must flow over the ISLs in the core. It is a two-hop infrastructure for non-virtualized storage. However, all virtualized storage traffic must pass through at least one of the ports on the AP-7600B. Therefore, the I/O from the host must traverse an ISL between the DS-32B2 and the ED-48000B. It must then traverse an ISL between the ED-48000B and the next ED-48000B. Then, it must traverse the ISL between the ED-48000B and the AP-7600B. Finally, it must traverse an ISL back from the AP-7600B to the ED-48000B where the I/O is terminated. This is a four-hop design that would require an RPQ. (Refer to Chapter 4, "Interswitch Link Hops," for more information on how storage and intelligent placement impacts fabric design.)

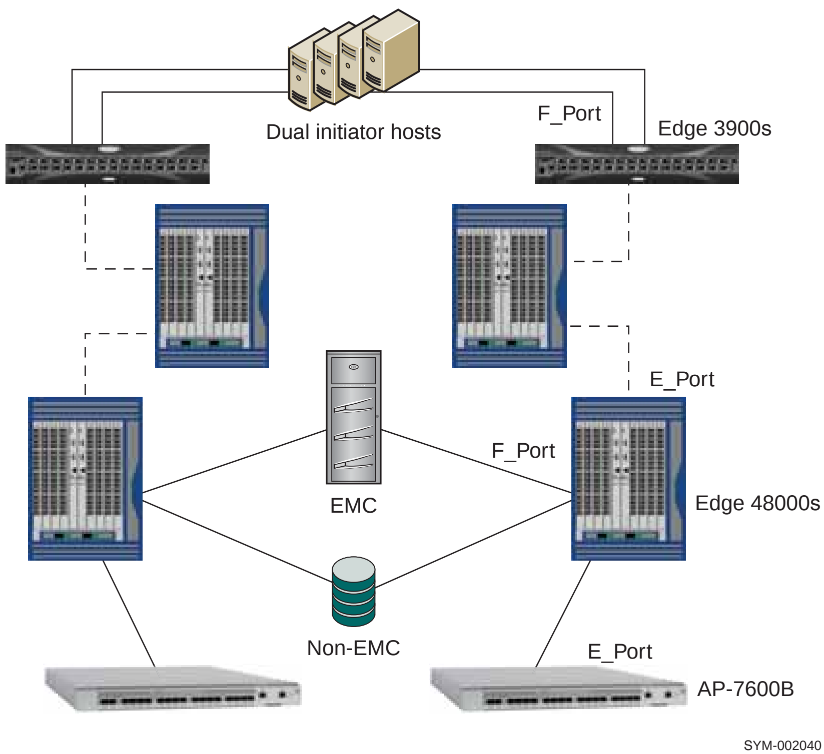


**Figure 18** **Core/edge/core design with Brocade**

### Storage on multiple switches

If the physical storage is spread amongst several edge switches, a single SSM must be located in a centralized location in the fabric. By doing this, one will achieve the highest possible efficiencies. In Figure 19 on page 62, only one side of the dual fabric configuration is shown. Because the physical storage ports are divided among multiple edge switches in the storage tier, the SSM has been placed in the connectivity layer in the core.

**Note:** Physical placement of the RecoverPoint Appliances can be anywhere within the fabric and need not be connected directly to the intelligent switch.

#### For RecoverPoint

Just as with the core/edge design (or any design), all virtualized traffic must flow through the virtualization ASICs. By locating the SSM in the connectivity layer, RecoverPoint's VIs will only need to traverse a single ISL in order to access physical storage. If the SSM was placed in one of the storage tier switches, most or some of the traffic between the SSM and storage would traverse two or more ISLs.

**Note:** Connections to the SSM are *not* required.

- ISLs between the switches do not have to be connected to the SSM blade
- Hosts do not have to be connected to the SSM blade
- Storage should *not* be connected to the SSM blade

Since the SSM is located inside the MDS 9513, internal routing between blades in the chassis is *not* considered an ISL hop. There is additional latency with the virtualization ASICs; however, there is no protocol overhead associated with routing between blades in a chassis.

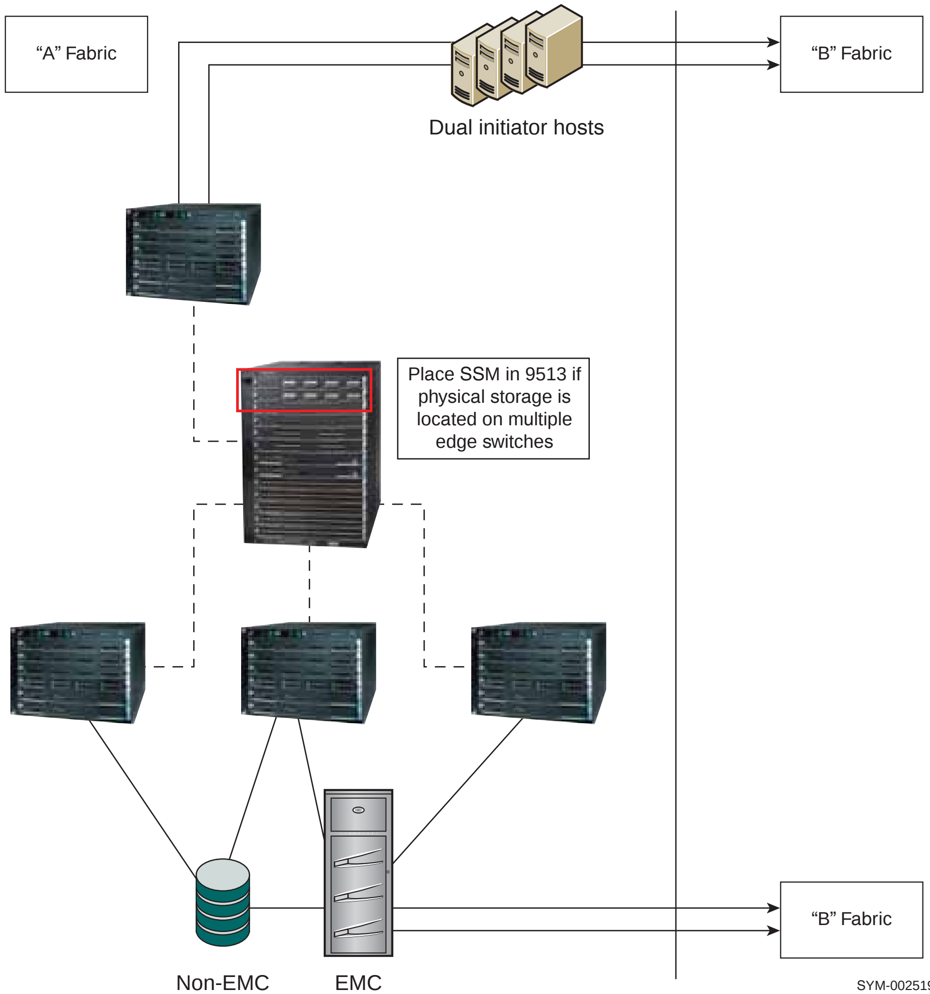


**Figure 19** **Storage on multiple switches**

Similarly in the RecoverPoint solution, the physical storage may be spread amongst several edge switches. An AP-7600B must be located in a centralized location in the fabric. This will achieve the highest possible efficiencies.

In Figure 20 on page 63, only one side of the dual fabric configuration is shown. Because the physical storage ports are divided among multiple edge switches in the storage tier, the AP-7600B has been placed in the connectivity layer in the core.

**Note:** When RecoverPoint is added to a SAN, not all physical storage has to be virtualized. Multiple storage ports may have LUNs to be virtualized. However, if more than 75% of the LUNs to be virtualized reside on a single switch (for example, Figure 17 on page 59, then locate the AP-7600B on that switch, as in Figure 17).

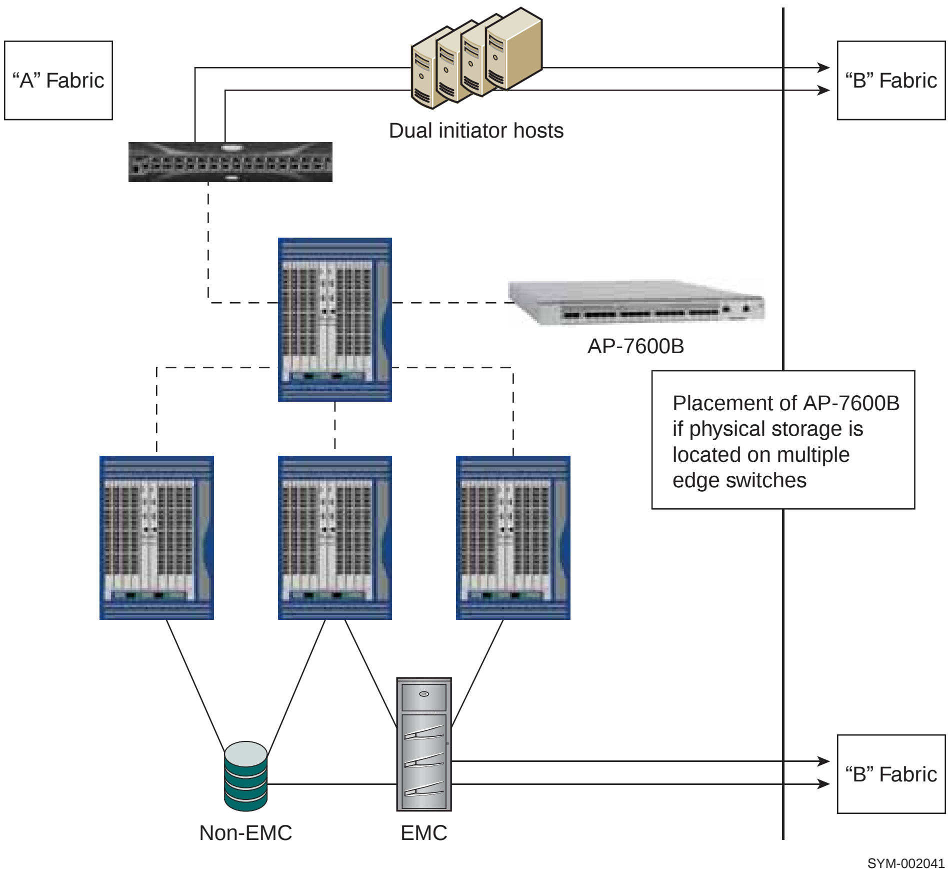


**Figure 20 Storage on multiple switches with Connectrix B**

# Interswitch link hops

This section contains the following information on interswitch link hops:



## Overview

An interswitch link hop is the physical link between two switches. The Fibre Channel switch itself is not considered a hop. The number of hops specified in the Switched Fabric Topology table of *EMC Support Matrix* states the maximum supported number of hops in a fabric under no fault conditions. For Class 3 traffic from initiator port to target port, and for Class F traffic among switches, there could be between three and five ISL hops, depending on the vendor. The physical link between host and switch as well as storage and switch is not counted as a hop. If there is a fault in the fabric, more than the *EMC Support Matrix* specified number of hops will be supported temporarily until the fault is resolved. In excess of maximum specified hops are not supported under normal operating conditions.

**Note:** In a fabric topology with **RecoverPoint**, all virtualized traffic must flow through an ASIC on the AP-7600B or the SSM blade.

In the previous chapters, recommendations were provided for placement in simple and complex fabrics. When including an intelligent switch into a fabric, traffic patterns across ISLs may change and hop counts may increase depending on the vendor and switch type used. The diagrams used in this section display hop counts in various fabric configurations.

In a Cisco MDS design, the SSM blade can be inserted in any MDS chassis with a slot. This includes the MDS 9513, 9509, 9506, and 9216. If the SSM is added to an existing fabric, the addition of the blade does not add any hops within the chassis itself. However, depending on the placement of the SSM in the environment, there may be a direct impact on the number of ISL hops.

**Note:** All virtualized traffic must pass through the SSM. All non-virtualized traffic does not have to pass through the SSM.

## Example 1 — Two-hop solution

Figure 21 on page 66 and Figure 22 on page 67, show an example of a four-switch mesh design using MDS 9506s. (Fabric 'B' is not shown for simplicity of the diagram. Fabric 'B' would be a mirror of Fabric 'A'.) Figure 21 shows a host attached to *sw1* and physical storage attached to *sw4*. The SSM module has been placed in *sw3*. The host I/O must follow a path from *sw1* to *sw3* (first hop) because of the fabric shortest path first (FSPF) algorithm inherent to FC switches. The SSM then re-maps the I/O and sends it to its physical target on *sw4* (second hop). This is a two-hop solution.

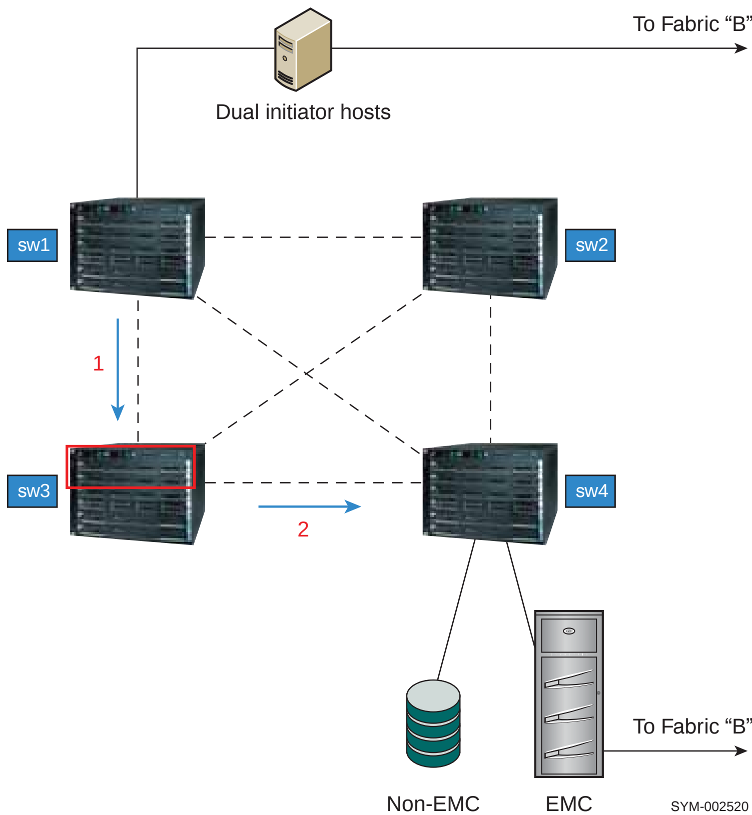


**Figure 21** **Four-switch mesh design using MDS 9506s (Example 1)**

## Example 2 — Preferred one-hop solution

Figure 22 shows the same host attached to *sw1* and the same physical storage attached to *sw4*. The SSM module has been moved to *sw4*. The host I/O must follow a path from *sw1* to *sw4*. The SSM then re-maps the I/O and sends it to its physical target on the same switch. There is no hop increase when I/O passes between blades in the MDS chassis. This is a preferred one-hop solution.

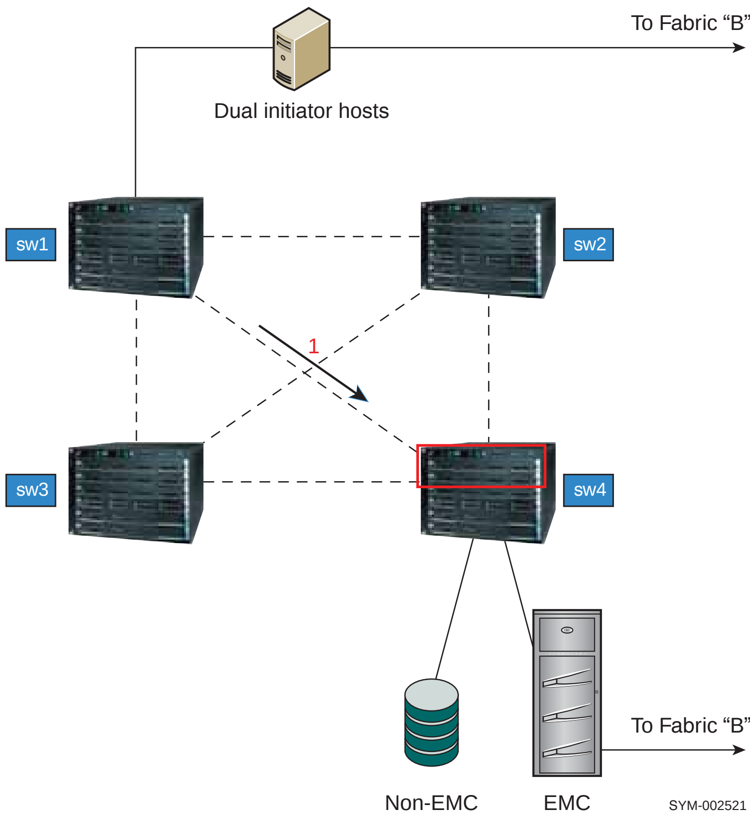


**Figure 22 Four-switch mesh design using MDS 9506s (Example 2)**

## Example 3 — Three-hop solution

Figure 23 on page 68 shows a four-switch mesh of ED-48000Bs with an AP-7600B as the intelligent switch. Because the AP-7600B is an external switch, it must be linked via an ISL to the existing fabric. Again, the host is attached to *sw1* and physical storage attached to *sw4*. The AP-7600B has been linked via an ISL to *sw4*. The host I/O must follow a path from *sw1* to *sw4* (first hop) because of the FSPF (fabric shortest path first) algorithm inherent to FC switches. The I/O is then passed from *sw4* to the AP-7600B (second hop).

The virtualization ASIC re-maps the I/O and passes it from the AP-7600B to *sw4* where the physical storage resides (third hop). This is a three-hop solution.

**Note:** The number of hops and traffic patterns would be identical in an MDS configuration if an MDS 9216 with an SSM blade which is hosting the RecoverPoint instance is attached to an existing fabric of MDS directors.

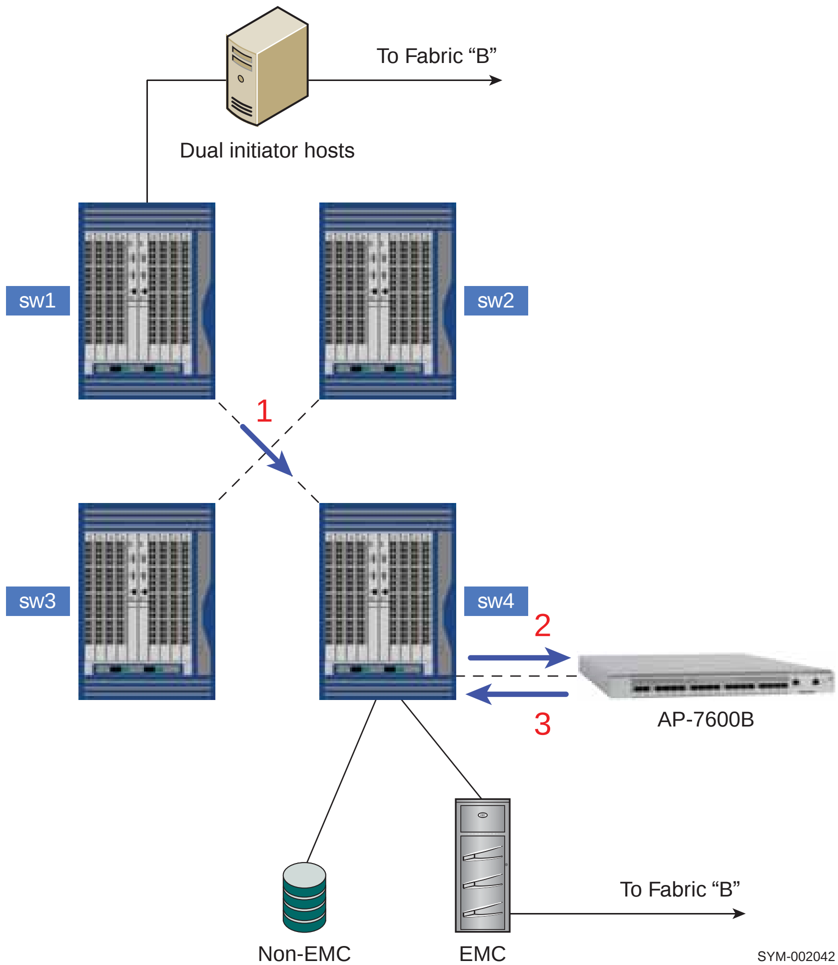


**Figure 23** **Four-switch mesh of ED-48000Bs with an AP-7600B as the intelligent switch (Example 3)**

With both mesh and complex fabric designs, it is not always possible to maintain a one-hop solution. With higher domain-count fabrics, there is a likelihood of increased hops. This is true of a virtualized and non-virtualized environment. It is also likely with the addition of intelligent switches or blades that some hosts will have more hops than others due to the placement of the physical storage in relation to the location of the intelligent switch.

## Example 4 — One-hop solution

Figure 24 on page 70 shows a four-switch mesh of MDS 9506s. Storage has been spread across *sw2*, *sw3*, and *sw4* switches. The SSM blade is located in *sw4*. (Again, Fabric 'B' is not shown for simplicity). Both *Host A* and *Host B* are accessing virtualized storage.

All I/O must pass through the SSM. I/O initially passes from *sw1* to *sw4* (first hop) because of the fabric shortest path first (FSPF) algorithm inherent to FC switches. The SSM re-maps the I/O. The final destination of the I/O *may* be the EMC storage on *sw4*.

If the I/O terminates there, it is considered a one-hop solution. However, the final destination *may* be the EMC storage on *sw2* or the non-EMC storage on *sw3*. In this case, it is a two-hop solution. *Host C* is *not* accessing virtualized storage. Its physical target is the EMC storage array on *sw2*. The I/O passes from *sw1* to *sw2* because of the FSPF algorithm inherent to FC switches. This is always a one-hop solution.

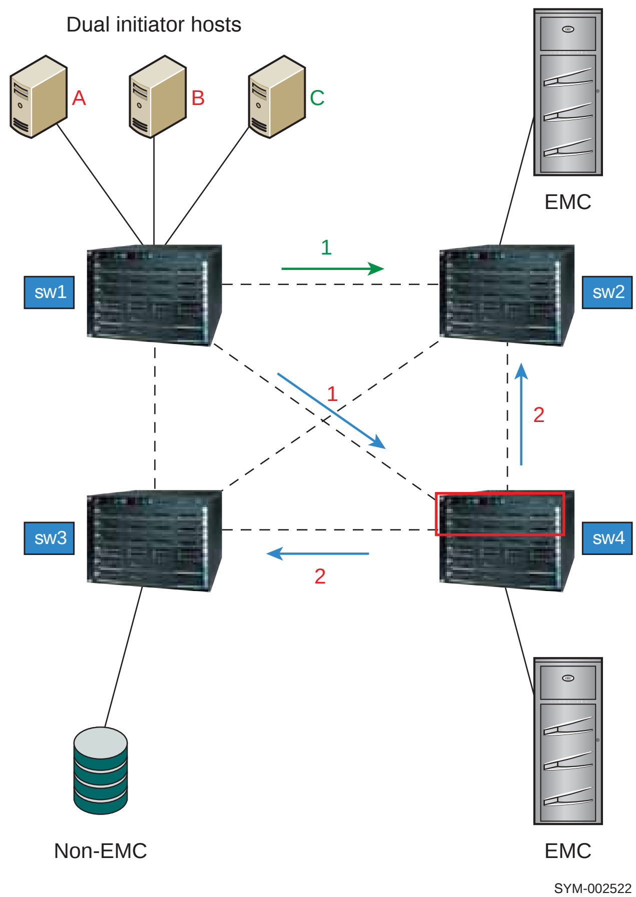


**Figure 24** **Four-switch mesh of MDS 9506s (Example 4)**

# Glossary

This glossary contains terms related to EMC products and EMC networked storage concepts.

## A

**access control** A service that allows or prohibits access to a resource. Storage management products implement access control to allow or prohibit specific users. Storage platform products implement access control, often called LUN Masking, to allow or prohibit access to volumes by Initiators (HBAs). *See also* "persistent binding" and "zoning."

**active domain ID** The domain ID actively being used by a switch. It is assigned to a switch by the principal switch.

**active zone set** The active zone set is the zone set definition currently in effect and enforced by the fabric or other entity (for example, the name server). Only one zone set at a time can be active.

**agent** An autonomous agent is a system situated within (and is part of) an environment that senses that environment, and acts on it over time in pursuit of its own agenda. Storage management software centralizes the control and monitoring of highly distributed storage infrastructure. The centralizing part of the software management system can depend on agents that are installed on the distributed parts of the infrastructure. For example, an agent (software component) can be installed on each of the hosts (servers) in an environment to allow the centralizing software to control and monitor the hosts.

**alarm** An SNMP message notifying an operator of a network problem.

**any-to-any port connectivity** A characteristic of a Fibre Channel switch that allows any port on the switch to communicate with any other port on the same switch.

**application** Application software is a defined subclass of computer software that employs the capabilities of a computer directly to a task that users want to perform. This is in contrast to system software that participates with integration of various capabilities of a computer, and typically does not directly apply these capabilities to performing tasks that benefit users. The term application refers to both the application software and its implementation which often refers to the use of an information processing system. (For example, a payroll application, an airline reservation application, or a network application.) Typically an application is installed “on top of” an operating system like Windows or Linux, and contains a user interface.

**application-specific integrated circuit (ASIC)** A circuit designed for a specific purpose, such as implementing lower-layer Fibre Channel protocols (FC-1 and FC-0). ASICs contrast with general-purpose devices such as memory chips or microprocessors, which can be used in many different applications.

**arbitration** The process of selecting one respondent from a collection of several candidates that request service concurrently.

**ASIC family** Different switch hardware platforms that utilize the same port ASIC can be grouped into collections known as an ASIC family. For example, the Fuji ASIC family which consists of the ED-64M and ED-140M run different microprocessors, but both utilize the same port ASIC to provide Fibre Channel connectivity, and are therefore in the same ASIC family. For inter operability concerns, it is useful to understand to which ASIC family a switch belongs.

**ASCII** ASCII (American Standard Code for Information Interchange), generally pronounced [aeski], is a character encoding based on the English alphabet. ASCII codes represent text in computers, communications equipment, and other devices that work with text. Most modern character encodings, which support many more characters, have a historical basis in ASCII.

**audit log** A log containing summaries of actions taken by a Connectrix Management software user that creates an audit trail of changes. Adding, modifying, or deleting user or product administration values, creates a record in the audit log that includes the date and time.

**authentication** Verification of the identity of a process or person.

## B

**backpressure** The effect on the environment leading up to the point of restriction. See "congestion."

**BB_Credit** *See* "buffer-to-buffer credit."

**beaconing** Repeated transmission of a beacon light and message until an error is corrected or bypassed. Typically used by a piece of equipment when an individual Field Replaceable Unit (FRU) needs replacement. Beaconing helps the field engineer locate the specific defective component. Some equipment management software systems such as Connectrix Manager offer beaconing capability.

**BER** *See* "bit error rate."

**bidirectional** In Fibre Channel, the capability to simultaneously communicate at maximum speeds in both directions over a link.

**bit error rate** Ratio of received bits that contain errors to total of all bits transmitted.

**blade server** A consolidation of independent servers and switch technology in the same chassis.

**blocked port** Devices communicating with a blocked port are prevented from logging in to the Fibre Channel switch containing the port or communicating with other devices attached to the switch. A blocked port continuously transmits the off-line sequence (OLS).

**bridge** A device that provides a translation service between two network segments utilizing different communication protocols. EMC supports and sells bridges that convert iSCSI storage commands from a NIC-attached server to Fibre Channel commands for a storage platform.

**broadcast** Sends a transmission to all ports in a network. Typically used in IP networks. Not typically used in Fibre Channel networks.

**broadcast frames** Data packet, also known as a broadcast packet, whose destination address specifies all computers on a network. *See also* "multicast."

**buffer** Storage area for data in transit. Buffers compensate for differences in link speeds and link congestion between devices.

**buffer-to-buffer credit** The number of receive buffers allocated by a receiving FC_Port to a transmitting FC_Port. The value is negotiated between Fibre Channel ports during link initialization. Each time a port transmits a frame it decrements this credit value. Each time a port receives an R_Rdy frame it increments this credit value. If the credit value is decremented to zero, the transmitter stops sending any new frames until the receiver has transmitted an R_Rdy frame. Buffer-to-buffer credit is particularly important in SRDF and Mirror View distance extension solutions.

# C

**Call Home** A product feature that allows the Connectrix service processor to automatically dial out to a support center and report system problems. The support center server accepts calls from the Connectrix service processor, logs reported events, and can notify one or more support center representatives. Telephone numbers and other information are configured through the Windows NT dial-up networking application. The Call Home function can be enabled and disabled through the Connectrix Product Manager.

**channel** With Open Systems, a channel is a point-to-point link that transports data from one point to another on the communication path, typically with high throughput and low latency that is generally required by storage systems. With Mainframe environments, a channel refers to the server-side of the server-storage communication path, analogous to the HBA in Open Systems.

**Class 2 Fibre Channel class of service** In Class 2 service, the fabric and destination N_Ports provide connectionless service with notification of delivery or nondelivery between the two N_Ports. Historically Class 2 service is not widely used in Fibre Channel system.

**Class 3 Fibre Channel class of service** Class 3 service provides a connectionless service without notification of delivery between N_Ports. (This is also known as datagram service.) The transmission and routing of Class 3 frames is the same as for Class 2 frames. Class 3 is the dominant class of communication used in Fibre Channel for moving data between servers and storage and may be referred to as "Ship and pray."

**Class F Fibre Channel class of service** Class F service is used for all switch-to-switch communication in a multiswitch fabric environment. It is nearly identical to class 2 from a flow control point of view.

**community** A relationship between an SNMP agent and a set of SNMP managers that defines authentication, access control, and proxy characteristics.

**community name** A name that represents an SNMP community that the agent software recognizes as a valid source for SNMP requests. An SNMP management program that sends an SNMP request to an agent program must identify the request with a community name that the agent recognizes or the agent discards the message as an authentication failure. The agent counts these failures and reports the count to the manager program upon request, or sends an authentication failure trap message to the manager program.

**community profile** Information that specifies which management objects are available to what management domain or SNMP community name.

**congestion** Occurs at the point of restriction. See "backpressure."

**connectionless** Non dedicated link. Typically used to describe a link between nodes that allows the switch to forward Class 2 or Class 3 frames as resources (ports) allow. *Contrast with* the dedicated bandwidth that is required in a Class 1 Fibre Channel Service point-to-point link.

**Connectivity Unit** A hardware component that contains hardware (and possibly software) that provides Fibre Channel connectivity across a fabric. Connectrix switches are example of Connectivity Units. This is a term popularized by the Fibre Alliance MIB, sometimes abbreviated to connunit.

**Connectrix management software** The software application that implements the management user interface for all managed Fibre Channel products, typically the Connectrix -M product line. Connectrix Management software is a client/server application with the server running on the Connectrix service processor, and clients running remotely or on the service processor.

**Connectrix service processor** An optional 1U server shipped with the Connectrix -M product line to run the Connectrix Management server software and EMC remote support application software.

**Control Unit** In mainframe environments, a Control Unit controls access to storage. It is analogous to a Target in Open Systems environments.

**core switch** Occupies central locations within the interconnections of a fabric. Generally provides the primary data paths across the fabric and the direct connections to storage devices. Connectrix directors are typically installed as core switches, but may be located anywhere in the fabric.

**credit** A numeric value that relates to the number of available BB_Credits on a Fibre Channel port. *See*“buffer-to-buffer credit”.

## D

**DASD** Direct Access Storage Device.

**default** Pertaining to an attribute, value, or option that is assumed when none is explicitly specified.

**default zone** A zone containing all attached devices that are not members of any active zone. Typically the default zone is disabled in a Connectrix M environment which prevents newly installed servers and storage from communicating until they have been provisioned.

**Dense Wavelength Division Multiplexing (DWDM)** A process that carries different data channels at different wavelengths over one pair of fiber optic links. A conventional fiber-optic system carries only one channel over a single wavelength traveling through a single fiber.

**destination ID** A field in a Fibre Channel header that specifies the destination address for a frame. The Fibre Channel header also contains a Source ID (SID). The FCID for a port contains both the SID and the DID.

**device** A piece of equipment, such as a server, switch or storage system.

**dialog box** A user interface element of a software product typically implemented as a pop-up window containing informational messages and fields for modification. Facilitates a dialog between the user and the application. Dialog box is often used interchangeably with window.

**DID** An acronym used to refer to either Domain ID or Destination ID. This ambiguity can create confusion. As a result E-Lab recommends this acronym be used to apply to Domain ID. Destination ID can be abbreviated to FCID.

**director** An enterprise-class Fibre Channel switch, such as the Connectrix ED-140M, MDS 9509, or ED-48000B. Directors deliver high availability, failure ride-through, and repair under power to insure maximum uptime for business critical applications. Major assemblies, such as power supplies, fan modules, switch controller cards, switching elements, and port modules, are all hot-swappable.

The term director may also refer to a board-level module in the Symmetrix that provides the interface between host channels (through an associated adapter module in the Symmetrix) and Symmetrix disk devices. (This description is presented here only to clarify a term used in other EMC documents.)

**DNS** *See* "domain name service name."

**domain ID** A byte-wide field in the three byte Fibre Channel address that uniquely identifies a switch in a fabric. The three fields in a FCID are domain, area, and port. A distinct Domain ID is requested from the principal switch. The principal switch allocates one Domain ID to each switch in the fabric. A user may be able to set a Preferred ID which can be requested of the Principal switch, or set an Insistent Domain ID. If two switches insist on the same DID one or both switches will segment from the fabric.

**domain name service name** Host or node name for a system that is translated to an IP address through a name server. All DNS names have a host name component and, if fully qualified, a domain component, such as *host1.abcd.com*. In this example, *host1* is the host name.

**dual-attached host** A host that has two (or more) connections to a set of devices.

## E

**E_D_TOV** A time-out period within which each data frame in a Fibre Channel sequence transmits. This avoids time-out errors at the destination Nx_Port. This function facilitates high speed recovery from dropped frames. Typically this value is 2 seconds.

**E_Port** Expansion Port, a port type in a Fibre Channel switch that attaches to another E_Port on a second Fibre Channel switch forming an Interswitch Link (ISL). This link typically conforms to the FC-SW standards developed by the T11 committee, but might not support heterogeneous inter operability.

**edge switch** Occupies the periphery of the fabric, generally providing the direct connections to host servers and management workstations. No two edge switches can be connected by interswitch links (ISLs). Connectrix departmental switches are typically installed as edge switches in a multiswitch fabric, but may be located anywhere in the fabric

**Embedded Web Server** A management interface embedded on the switch's code that offers features similar to (but not as robust as) the Connectrix Manager and Product Manager.

**error detect time out value** Defines the time the switch waits for an expected response before declaring an error condition. The error detect time out value (E_D_TOV) can be set within a range of two-tenths of a second to one second using the Connectrix switch Product Manager.

**error message** An indication that an error has been detected. *See also* "information message" *and* "warning message."

**Ethernet** A baseband LAN that allows multiple station access to the transmission medium at will without prior coordination and which avoids or resolves contention.

**event log** A record of significant events that have occurred on a Connectrix switch, such as FRU failures, degraded operation, and port problems.

**expansionport** *See* "E_Port."

**explicit fabric login** In order to join a fabric, an Nport must login to the fabric (an operation referred to as an FLOGI). Typically this is an explicit operation performed by the Nport communicating with the F_port of the switch, and is called an explicit fabric login. Some legacy Fibre Channel ports do not perform explicit login, and switch vendors perform login for ports creating an implicit login. Typically logins are explicit.

## F

**FA** Fibre Adapter, another name for a Symmetrix Fibre Channel director.

**F_Port** Fabric Port, a port type on a Fibre Channel switch. An F_Port attaches to an N_Port through a point-to-point full-duplex link connection. A G_Port automatically becomes an F_port or an E-Port depending on the port initialization process.

**fabric** One or more switching devices that interconnect Fibre Channel N_Ports, and route Fibre Channel frames based on destination IDs in the frame headers. A fabric provides discovery, path provisioning, and state change management services for a Fibre Channel environment.

**fabric element** Any active switch or director in the fabric.

**fabric login** Process used by N_Ports to establish their operating parameters including class of service, speed, and buffer-to-buffer credit value.

**fabric port** A port type (F_Port) on a Fibre Channel switch that attaches to an N_Port through a point-to-point full-duplex link connection. An N_Port is typically a host (HBA) or a storage device like Symmetrix, VNX series, or CLARiiON.

**fabric shortest path first (FSPF)** A routing algorithm implemented by Fibre Channel switches in a fabric. The algorithm seeks to minimize the number of hops traversed as a Fibre Channel frame travels from its source to its destination.

**fabric tree** A hierarchical list in Connectrix Manager of all fabrics currently known to the Connectrix service processor. The tree includes all members of the fabrics, listed by WWN or nickname.

**failover** The process of detecting a failure on an active Connectrix switch FRU and the automatic transition of functions to a backup FRU.

**fan-in/fan-out** Term used to describe the server:storage ratio, where a graphic representation of a 1:n (fan-in) or n:1 (fan-out) logical topology looks like a hand-held fan, with the wide end toward n. By convention fan-out refers to the number of server ports that share a single storage port. Fan-out consolidates a large number of server ports on a fewer number of storage ports. Fan-in refers to the number of storage ports that a single server port uses. Fan-in enlarges the storage capacity used by a server. A fan-in or fan-out rate is often referred to as just the n part of the ratio; For example, a 16:1 fan-out is also called a fan-out rate of 16, in this case 16 server ports are sharing a single storage port.

**FCP** *See* "Fibre Channel Protocol."

**FC-SW** The Fibre Channel fabric standard. The standard is developed by the T11 organization whose documentation can be found at T11.org. EMC actively participates in T11. T11 is a committee within the InterNational Committee for Information Technology (INCITS).

**fiber optics** The branch of optical technology concerned with the transmission of radiant power through fibers made of transparent materials such as glass, fused silica, and plastic.

Either a single discrete fiber or a non spatially aligned fiber bundle can be used for each information channel. Such fibers are often called optical fibers to differentiate them from fibers used in non-communication applications.

**fibre** A general term used to cover all physical media types supported by the Fibre Channel specification, such as optical fiber, twisted pair, and coaxial cable.

**Fibre Channel** The general name of an integrated set of ANSI standards that define new protocols for flexible information transfer. Logically, Fibre Channel is a high-performance serial data channel.

**Fibre Channel Protocol** A standard Fibre Channel FC-4 level protocol used to run SCSI over Fibre Channel.

**Fibre Channel switch modules** The embedded switch modules in the back plane of the blade server. See "blade server" on page 73.

**firmware** The program code (embedded software) that resides and executes on a connectivity device, such as a Connectrix switch, a Symmetrix Fibre Channel director, or a host bus adapter (HBA).

**F_Port** Fabric Port, a physical interface within the fabric. An F_Port attaches to an N_Port through a point-to-point full-duplex link connection.

**frame** A set of fields making up a unit of transmission. Each field is made of bytes. The typical Fibre Channel frame consists of fields: Start-of-frame, header, data-field, CRC, end-of-frame. The maximum frame size is 2148 bytes.

**frame header** Control information placed before the data-field when encapsulating data for network transmission. The header provides the source and destination IDs of the frame.

**FRU** Field-replaceable unit, a hardware component that can be replaced as an entire unit. The Connectrix switch Product Manager can display status for the FRUs installed in the unit.

**FSPF** Fabric Shortest Path First, an algorithm used for routing traffic. This means that, between the source and destination, only the paths that have the least amount of physical hops will be used for frame delivery.

## G

**gateway address** In TCP/IP, a device that connects two systems that use the same or different protocols.

**gigabyte (GB)** A unit of measure for storage size, loosely one billion ($10^9$) bytes. One gigabyte actually equals 1,073,741,824 bytes.

**G_Port** A port type on a Fibre Channel switch capable of acting either as an F_Port or an E_Port, depending on the port type at the other end of the link.

**GUI** Graphical user interface.

## H

**HBA** *See* “host bus adapter.”

**hexadecimal** Pertaining to a numbering system with base of 16; valid numbers use the digits 0 through 9 and characters A through F (which represent the numbers 10 through 15).

**high availability** A performance feature characterized by hardware component redundancy and hot-swappability (enabling non-disruptive maintenance). High-availability systems maximize system uptime while providing superior reliability, availability, and serviceability.

**hop** A hop refers to the number of InterSwitch Links (ISLs) a Fibre Channel frame must traverse to go from its source to its destination.

Good design practice encourages three hops or less to minimize congestion and performance management complexities.

**host bus adapter** A bus card in a host system that allows the host system to connect to the storage system. Typically the HBA communicates with the host over a PCI or PCI Express bus and has a single Fibre Channel link to the fabric. The HBA contains an embedded microprocessor with on board firmware, one or more ASICs, and a Small Form Factor Pluggable module (SFP) to connect to the Fibre Channel link.

## I

**I/O** *See* "input/output."

**in-band management** Transmission of monitoring and control functions over the Fibre Channel interface. You can also perform these functions out-of-band typically by use of the ethernet to manage Fibre Channel devices.

**information message** A message telling a user that a function is performing normally or has completed normally. User acknowledgement might or might not be required, depending on the message. *See also* "error message" *and* "warning message."

**input/output** (1) Pertaining to a device whose parts can perform an input process and an output process at the same time. (2) Pertaining to a functional unit or channel involved in an input process, output process, or both (concurrently or not), and to the data involved in such a process. (3) Pertaining to input, output, or both.

**interface** (1) A shared boundary between two functional units, defined by functional characteristics, signal characteristics, or other characteristics as appropriate. The concept includes the specification of the connection of two devices having different functions. (2) Hardware, software, or both, that links systems, programs, or devices.

**Internet Protocol** *See* "IP."

**interoperability** The ability to communicate, execute programs, or transfer data between various functional units over a network. Also refers to a Fibre Channel fabric that contains switches from more than one vendor.

**interswitch link (ISL)** Interswitch link, a physical E_Port connection between any two switches in a Fibre Channel fabric. An ISL forms a hop in a fabric.

**IP** Internet Protocol, the TCP/IP standard protocol that defines the datagram as the unit of information passed across an internet and provides the basis for connectionless, best-effort packet delivery service. IP includes the ICMP control and error message protocol as an integral part.

**IP address** A unique string of numbers that identifies a device on a network. The address consists of four groups (quadrants) of numbers delimited by periods. (This is called *dotted-decimal* notation.) All resources on the network must have an IP address. A valid IP address is in the form *nnn.nnn.nnn.nnn*, where each *nnn* is a decimal in the range 0 to 255.

**ISL** Interswitch link, a physical E_Port connection between any two switches in a Fibre Channel fabric.

## K

**kilobyte (K)** A unit of measure for storage size, loosely one thousand bytes. One kilobyte actually equals 1,024 bytes.

## L

**laser** A device that produces optical radiation using a population inversion to provide light amplification by stimulated emission of radiation and (generally) an optical resonant cavity to provide positive feedback. Laser radiation can be highly coherent temporally, spatially, or both.

**LED** Light-emitting diode.

**link** The physical connection between two devices on a switched fabric.

**link incident** A problem detected on a fiber-optic link; for example, loss of light, or invalid sequences.

**load balancing** The ability to distribute traffic over all network ports that are the same distance from the destination address by assigning different paths to different messages. Increases effective network bandwidth. EMC PowerPath software provides load-balancing services for server I/O.

**logical volume** A named unit of storage consisting of a logically contiguous set of disk sectors.

**Logical Unit Number (LUN)** A number, assigned to a storage volume, that (in combination with the storage device node's World Wide Port Name (WWPN)) represents a unique identifier for a logical volume on a storage area network.

## M

**MAC address** Media Access Control address, the hardware address of a device connected to a shared network.

**managed product** A hardware product that can be managed using the Connectrix Product Manager. For example, a Connectrix switch is a managed product.

**management session** Exists when a user logs in to the Connectrix Management software and successfully connects to the product server. The user must specify the network address of the product server at login time.

**media** The disk surface on which data is stored.

**media access control** *See* "MAC address."

**megabyte (MB)** A unit of measure for storage size, loosely one million ($10^6$) bytes. One megabyte actually equals 1,048,576 bytes.

**MIB** Management Information Base, a related set of objects (variables) containing information about a managed device and accessed through SNMP from a network management station.

**multicast** Multicast is used when multiple copies of data are to be sent to designated, multiple, destinations.

**multiswitch fabric** Fibre Channel fabric created by linking more than one switch or director together to allow communication. *See also* "ISL."

**multiswitch linking** Port-to-port connections between two switches.

## N

**name server (dNS)** A service known as the distributed Name Server provided by a Fibre Channel fabric that provides device discovery, path provisioning, and state change notification services to the N_Ports in the fabric. The service is implemented in a distributed fashion, for example, each switch in a fabric participates in providing the service. The service is addressed by the N_Ports through a Well Known Address.

**network address** A name or address that identifies a managed product, such as a Connectrix switch, or a Connectrix service processor on a TCP/IP network. The network address can be either an IP address in dotted decimal notation, or a Domain Name Service (DNS) name as administered on a customer network. All DNS names have a host name component and (if fully qualified) a domain component, such as *host1.emc.com*. In this example, *host1* is the host name and *EMC.com* is the domain component.

**nickname** A user-defined name representing a specific WWxN, typically used in a Connectrix -M management environment. The analog in the Connectrix -B and MDS environments is alias.

**node** The point at which one or more functional units connect to the network.

**N_Port** Node Port, a Fibre Channel port implemented by an end device (node) that can attach to an F_Port or directly to another N_Port through a point-to-point link connection. HBAs and storage systems implement N_Ports that connect to the fabric.

**NVRAM** Nonvolatile random access memory.

## O

**offline sequence (OLS)** The OLS Primitive Sequence is transmitted to indicate that the FC_Port transmitting the Sequence is:

a. initiating the Link Initialization Protocol

b. receiving and recognizing NOS

c. or entering the offline state

**OLS** *See* "offline sequence (OLS)".

**operating mode** Regulates what other types of switches can share a multiswitch fabric with the switch under consideration.

**operating system** Software that controls the execution of programs and that may provide such services as resource allocation, scheduling, input/output control, and data management. Although operating systems are predominantly software, partial hardware implementations are possible.

**optical cable** A fiber, multiple fibers, or a fiber bundle in a structure built to meet optical, mechanical, and environmental specifications.

**OS** *See* "operating system."

**out-of-band management** Transmission of monitoring/control functions outside of the Fibre Channel interface, typically over ethernet.

**oversubscription** The ratio of bandwidth required to bandwidth available. When all ports, associated pair-wise, in any random fashion, cannot sustain full duplex at full line-rate, the switch is oversubscribed.

## P

**parameter** A characteristic element with a variable value that is given a constant value for a specified application. Also, a user-specified value for an item in a menu; a value that the system provides when a menu is interpreted; data passed between programs or procedures.

**password** (1) A value used in authentication or a value used to establish membership in a group having specific privileges. (2) A unique string of characters known to the computer system and to a user who must specify it to gain full or limited access to a system and to the information stored within it.

**path** In a network, any route between any two nodes.

**persistent binding** Use of server-level access control configuration information to persistently bind a server device name to a specific Fibre Channel storage volume or logical unit number, through a specific HBA and storage port WWN. The address of a persistently bound device does not shift if a storage target fails to recover during a power cycle. This function is the responsibility of the HBA device driver.

**port** (1) An access point for data entry or exit. (2) A receptacle on a device to which a cable for another device is attached.

**port card** Field replaceable hardware component that provides the connection for fiber cables and performs specific device-dependent logic functions.

**port name** A symbolic name that the user defines for a particular port through the Product Manager.

**preferred domain ID** An ID configured by the fabric administrator. During the fabric build process a switch requests permission from the principal switch to use its preferred domain ID. The principal switch can deny this request by providing an alternate domain ID only if there is a conflict for the requested Domain ID. Typically a principal switch grants the non-principal switch its requested Preferred Domain ID.

**principal downstream ISL** The ISL to which each switch will forward frames originating from the principal switch.

**principal ISL** The principal ISL is the ISL that frames destined to, or coming from, the principal switch in the fabric will use. An example is an RDI frame.

**principal switch** In a multiswitch fabric, the switch that allocates domain IDs to itself and to all other switches in the fabric. There is always one principal switch in a fabric. If a switch is not connected to any other switches, it acts as its own principal switch.

**principal upstream ISL** The ISL to which each switch will forward frames destined for the principal switch. The principal switch does not have any upstream ISLs.

**product** (1) Connectivity Product, a generic name for a switch, director, or any other Fibre Channel product. (2) Managed Product, a generic hardware product that can be managed by the Product Manager (a Connectrix switch is a managed product). Note distinction from the definition for "device."

**Product Manager** A software component of Connectrix Manager software such as a Connectrix switch product manager, that implements the management user interface for a specific product. When a product instance is opened from the Connectrix Manager software products view, the corresponding product manager is invoked. The product manager is also known as an Element Manager.

**product name** A user configurable identifier assigned to a Managed Product. Typically, this name is stored on the product itself. For a Connectrix switch, the Product Name can also be accessed by an SNMP Manager as the System Name. The Product Name should align with the host name component of a Network Address.

**products view** The top-level display in the Connectrix Management software user interface that displays icons of Managed Products.

**protocol** (1) A set of semantic and syntactic rules that determines the behavior of functional units in achieving communication. (2) A specification for the format and relative timing of information exchanged between communicating parties.

## R

**R_A_TOV** *See* "resource allocation time out value."

**remote access link** The ability to communicate with a data processing facility through a remote data link.

**remote notification** The system can be programmed to notify remote sites of certain classes of events.

**remote user workstation** A workstation, such as a PC, using Connectrix Management software and Product Manager software that can access the Connectrix service processor over a LAN connection. A user at a remote workstation can perform all of the management and monitoring tasks available to a local user on the Connectrix service processor.

**resource allocation time out value** A value used to time-out operations that depend on a maximum time that an exchange can be delayed in a fabric and still be delivered. The resource allocation time-out value of (R_A_TOV) can be set within a range of two-tenths of a second to 120 seconds using the Connectrix switch product manager. The typical value is 10 seconds.

## S

**SAN** *See* "storage area network (SAN)."

**segmentation** A non-connection between two switches. Numerous reasons exist for an operational ISL to segment, including interop mode incompatibility, zoning conflicts, and domain overlaps.

**segmented E_Port** E_Port that has ceased to function as an E_Port within a multiswitch fabric due to an incompatibility between the fabrics that it joins.

**service processor** *See* "Connectrix service processor."

**session** *See* "management session."

**single attached host** A host that only has a single connection to a set of devices.

**small form factor pluggable (SFP)** An optical module implementing a shortwave or long wave optical transceiver.

**SMTP** Simple Mail Transfer Protocol, a TCP/IP protocol that allows users to create, send, and receive text messages. SMTP protocols specify how messages are passed across a link from one system to another. They do not specify how the mail application accepts, presents or stores the mail.

**SNMP** Simple Network Management Protocol, a TCP/IP protocol that generally uses the User Datagram Protocol (UDP) to exchange messages between a management information base (MIB) and a management client residing on a network.

**storage area network (SAN)** A network linking servers or workstations to disk arrays, tape backup systems, and other devices, typically over Fibre Channel and consisting of multiple fabrics.

**subnet mask** Used by a computer to determine whether another computer with which it needs to communicate is located on a local or remote network. The network mask depends upon the class of networks to which the computer is connecting. The mask indicates which digits to look at in a longer network address and allows the router to avoid handling the entire address. Subnet masking allows routers to move the packets more quickly. Typically, a subnet may represent all the machines at one geographic location, in one building, or on the same local area network.

**switch priority** Value configured into each switch in a fabric that determines its relative likelihood of becoming the fabric's principal switch.

## T

**TCP/IP** Transmission Control Protocol / Internet Protocol. TCP / IP refers to the protocols that are used on the Internet and most computer networks. TCP refers to the Transport layer that provides flow control and connection services. IP refers to the Internet Protocol level where addressing and routing are implemented.

**toggle** To change the state of a feature / function that has only two states. For example, if a feature / function is *enabled*, toggling changes the state to *disabled*.

**topology** Logical and / or physical arrangement of switches on a network.

**trap** An asynchronous (unsolicited) notification of an event originating on an SNMP-managed device and directed to a centralized SNMP Network Management Station.

## U

**unblocked port** Devices communicating with an unblocked port can log in to a Connectrix switch or a similar product and communicate with devices attached to any other unblocked port if the devices are in the same zone.

**Unicast** Unicast routing provides one or more optimal path(s) between any of two switches that make up the fabric. (This is used to send a single copy of the data to designated destinations.)

**upper layer protocol (ULP)** The protocol user of FC-4 including IPI, SCSI, IP, and SBCCS. In a device driver ULP typically refers to the operations that are managed by the class level of the driver, not the port level.

**URL** Uniform Resource Locater, the addressing system used by the World Wide Web. It describes the location of a file or server anywhere on the Internet.

## V

**virtual switch** A Fibre Channel switch function that allows users to subdivide a physical switch into multiple virtual switches. Each virtual switch consists of a subset of ports on the physical switch, and has all the properties of a Fibre Channel switch. Multiple virtual switches can be connected through ISL to form a virtual fabric or VSAN.

**virtual storage area network (VSAN)** An allocation of switch ports that can span multiple physical switches, and forms a virtual fabric. A single physical switch can sometimes host more than one VSAN.

**volume** A general term referring to an addressable logically contiguous storage space providing block I/O services.

**VSAN** Virtual Storage Area Network.

## W

**warning message** An indication that a possible error has been detected. *See also* "error message" *and* "information message."

**World Wide Name (WWN)** A unique identifier, even on global networks. The WWN is a 64-bit number (XX:XX:XX:XX:XX:XX:XX:XX). The WWN contains an OUI which uniquely determines the equipment manufacturer. OUIs are administered by the Institute of Electronic and Electrical Engineers (IEEE). The Fibre Channel environment uses two types of WWNs; a World Wide Node Name (WWNN) and a World Wide Port Name (WWPN). Typically the WWPN is used for zoning (path provisioning function).

## Z

**zone** An information object implemented by the distributed Nameserver (dNS) of a Fibre Channel switch. A zone contains a set of members which are permitted to discover and communicate with one another. The members can be identified by a WWPN or port ID. EMC recommends the use of WWPNs in zone management.

**zone set** An information object implemented by the distributed Nameserver (dNS) of a Fibre Channel switch. A Zone Set contains a set of Zones. A Zone Set is activated against a fabric, and only one Zone Set can be active in a fabric.

**zonie** A storage administrator who spends a large percentage of his workday zoning a Fibre Channel network and provisioning storage.

**zoning** Zoning allows an administrator to group several devices by function or by location. All devices connected to a connectivity product, such as a Connectrix switch, may be configured into one or more zones.