EMC® Symmetrix® Remote Data Facility (SRDF®)

Connectivity Guide

P/N 300-003-885
REV A10

Contents

Figures

Tables

	Title	Page

EMC Symmetrix Remote Data Facility (SRDF) Connectivity Guide

Preface

As part of an effort to improve and enhance the performance and capabilities of its product line, EMC from time to time releases revisions of its hardware and software. Therefore, some functions described in this document may not be supported by all revisions 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.

EMC Symmetrix Remote Data Facility (SRDF) is the industry-leading business-continuity, information-protection, and content-replication software-controlled solution. SRDF can extend the physical fiber connections between Symmetrix systems, using telecommunication networking technology.

Note: For simplicity, this document uses the term *SRDF* to represent all the EMC SRDF family products.

Audience   This guide is intended for use by EMC customers, partners, technical architects, field support, and customer service engineers involved with the design and deployment of SRDF networks. The audience is expected to be familiar with SRDF. More basic information on SRDF can be found in the EMC Symmetrix Remote Data Facility (SRDF) product guides, available at https://support.emc.com.

Organization   This document is intended to provide the reader with a connectivity-focused guide for deploying SRDF. To accomplish this goal, this guide is structured around various building blocks needed for successful SRDF network implementations and includes the following information:

◆ General information on SRDF technology

◆ Symmetrix hardware and Enginuity information relevant to SRDF

◆ Various network protocols available for SRDF deployments

◆ Various network topologies and transports available for SRDF deployments

Various levels of technical information will be provided as well as best practices where applicable.

This guide may reference non-EMC products and technologies. It is not within the scope of this guide to go into any detail on third-party products. For more information on third-party products, refer to third-party documentation and/or the *EMC Networked Storage Topology Guide* available through E-Lab Interoperability Navigator at http://elabnavigator.EMC.com.

Note: The implementations described in this document for medium-distance or long-distance solutions can be used locally as well. Proper planning will include all useful alternatives and can often benefit from existing equipment or other local considerations. Additionally, some limitations quoted in this document are subject to relatively frequent change. Contact your EMC Customer Support representative for the latest distance information before planning an implementation.

Related documentation   For documentation, refer to the Symmetrix Remote Data Facility (SRDF) product page, or other relevant product pages, at https://support.emc.com. This site also contains Documentation Portfolios for VMAX 40K, VMAX 20K/VMAX, and VMAX 10K series that include the hardware platform product guide, physical planning guide, SRDF product guide, and front-end port layout diagrams, as appropriate.

Related documents and tools include:

◆ EMC Symmetrix Remote Data Facility (SRDF) product guides

◆ *Symmetrix Remote Data Facility (SRDF) Two-Site Interfamily Connectivity* tool

- ◆ Symmetrix product guides
- ◆ Symmetrix physical planning guides

The following are available through the E-Lab Interoperability Navigator at http://elabnavigator.EMC.com.

- ◆ EMC TechBooks, **Topology Resource Center** tab
- ◆ *EMC Support Matrix* (ESM), **PDFs and Guides** tab

  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 product.
  - • 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*.

For detailed information regarding solutions that involve other qualified vendors' equipment, refer to the appropriate product specification manuals available from each vendor, or ask your site representative to contact the appropriate EMC partner vendor or EMC Technical Support/Engineering.

Conventions used in this guide

EMC uses the following conventions for notes, cautions, warnings, and danger notices.

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

IMPORTANT

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

⚠ **CAUTION**

A caution contains information essential to avoid data loss or damage to the system or equipment. The caution may apply to hardware or software.

Typographical conventions

EMC uses the following type style conventions in this document.

Bold	Use for names of interface elements, such as names of windows, dialog boxes, buttons, fields, tab names, key names, and menu paths (what the user specifically selects or clicks)
Italic	Use for full titles of publications referenced in text
`Monospace`	Use for: • System output, such as an error message or script • System code • Pathnames, filenames, prompts, and syntax • Commands and options
`Monospace italic`	Use for variables.
`Monospace bold`	Use for user input.
[ ]	Square brackets enclose optional values
|	Vertical bar indicates alternate selections — the bar means "or"
{ }	Braces enclose content that the user must specify, such as x or y or z
...	Ellipses indicate nonessential information omitted from the example

Where to get help   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 Online Support website at https://support.emc.com.

Technical support

For technical support, go to the EMC Online Support website at https://support.emc.com. You must have a valid support agreement. Please contact your EMC sales representative for details about obtaining a valid support agreement or to answer any questions about your account.

Your comments   Your suggestions will help us continue to improve the accuracy, organization, and overall quality of our user publications. Please send your opinion of this document to:

techpub_comments@EMC.com

EMC Symmetrix Remote Data Facility (SRDF) Connectivity Guide

SRDF Overview and Architecture

This chapter contains the following information:

SRDF overview

The EMC® Symmetrix® Remote Data Facility (SRDF®) family of remote replication software offers various levels of Symmetrix-based business continuance and disaster recovery solutions. The SRDF products offer the capability to maintain multiple copies of data, independent of the host and operating system. SRDF configurations require at least two Symmetrix systems, also known as the *primary* and the *secondary* system. Both sites can be located in the same room, in different buildings within the same campus, or hundreds to thousands of kilometers apart.

Supported features can differ depending on the hardware platforms. For the most up-to-date information, refer to the appropriate product guides, or the following SRDF product guides, located at https://support.emc.com.

Note: For simplicity, this document uses the term SRDF to represent all EMC SRDF-related products, including: SRDF/A, SRDF/AR (multi-hop and single-hop), SRDF/DM, and SRDF/CG. Most of the document focuses on SRDF/S, but generally it applies to any of the variations.

By maintaining copies of data in different physical locations, SRDF enables you to perform the following operations with minimal impact on normal business processing:

- Disaster restart
- Disaster restart testing
- Recovery from planned outages
- Remote backup
- Data center migration
- Data replication and mobility

SRDF concepts

The EMC Symmetrix Remote Data Facility (SRDF) product guides, located at https://support.emc.com, Symmetrix Remote Data Facility (SRDF) product page, Documentation, provides detailed information on the following SRDF concepts:

- SRDF devices
- SRDF device state
- SRDF groups
- SRDF links

Basic SRDF configuration

A basic SRDF configuration consisting of a primary site and a secondary site is illustrated in . At the primary site, a local host connects to Symmetrix A. The device containing the production data to be remotely mirrored is called the primary (or R1) device. At the secondary site, a second (optional) host connects to Symmetrix B with the secondary (R2) device containing the remotely mirrored data. The Symmetrix systems communicate through SRDF links.

SRDF data (host writes or tracks to be copied) is maintained in cache slots of global memory in the Symmetrix, containing the R1 devices. Since the release of EMC Enginuity™ Operating Environment for Symmetrix, version 5670, the actual SRDF work queue (I/O jobs) is distributed to individual SRDF directors' local cache (queues) by the front-end host adapters (HA) for SRDF/S and SRDF/A, and by the disk adapters (DA) for adaptive copy and SRDF/AR and SRDF/DM. The distribution of the work queues to the SRDF directors eliminates multiple memory accesses and lock/unlock cache slot contention. Each SRDF link will therefore operate independently at its own speed and latency. If the SRDF links are on a balanced network (speed and latency), then the SRDF workload will be well-balanced between directors.

Note: SRDF replicates only open systems data from a Symmetrix DMX, VMAX 40K, or VMAX 20K/VMAX array to a Symmetrix VMAX 10K or VMAXe array. Mainframe or IBM i data that resides on the Symmetrix DMX, VMAX 40K, or VMAX 20K/VMAX array cannot be replicated to the Symmetrix VMAX 10K or VMAXe array by using SRDF solutions.

Figure 1 illustrates two basic SRDF configurations.
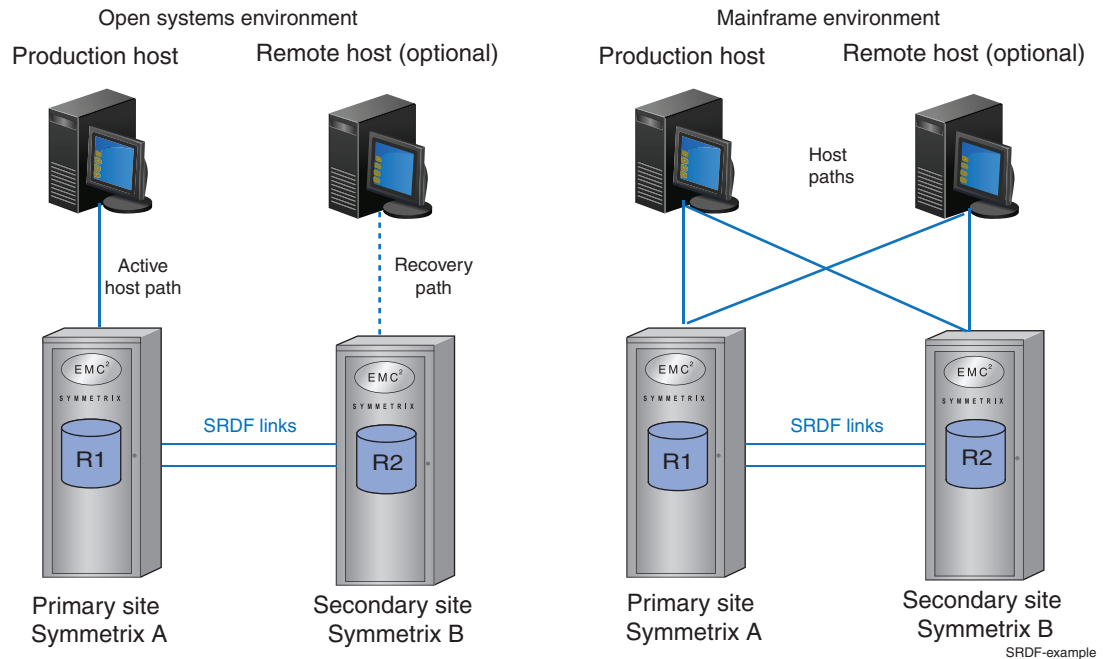
Figure 1   **Basic SRDF configurations in open systems and mainframe environments**

Note: The VMAX 10K Series does not support mainframe.

For information on typical SRDF configurations, refer to the appropriate EMC Symmetrix Remote Data Facility (SRDF) product guides located at https://support.emc.com, Symmetrix Remote Data Facility (SRDF) product page, Documentation.

The SRDF Two-site Interfamily Connectivity application on EMC Online Support (https://support.emc.com) provides details about SRDF two-site interfamily connectivity options. On the home page of EMC Online Support, search with the keywords **SRDF Two-site Interfamily Connectivity**.

You can also configure three-site SRDF solutions with different Symmetrix hardware models running different Enginuity versions. Refer to the appropriate product guides for more information.

SRDF base family products

The SRDF family consists of three base solutions, each discussed further in this section:

◆ SRDF/Synchronous (SRDF/S), page 20

◆ SRDF/Asynchronous (SRDF/A), page 23

◆ SRDF/Data Mobility (SRDF/DM), page 24

Also included in this section is more details on "SRDF family options" on page 26.

More details can be found in the appropriate SRDF product guides, located at https://support.emc.com.

Figure 2 provides a brief description and SRDF site configuration of SRDF base solutions.

Description	SRDF site configuration
SRDF/S • Synchronous mode • No data exposure • Limited distance (up to 125 miles (200 km))	Primary — R1 — Limited distance Synchronous — R2 — Secondary (SRDF-OptionsS)
SRDF/A • Asynchronous mode • RPO seconds before the point of failure • Unlimited distance	Primary — R1 — Unlimited distance Asynchronous — R2 — Secondary (SRDF-OptionsA)
SRDF/DM • Adaptive copy mode • Designed for migration and remote replication only • Unlimited distance	Primary — R1 — Unlimited distance Adaptive copy — R2 — Secondary (SRDF-OptionsDM)

Figure 2   SRDF base product offerings

Note: Best practices for these modes of operation can be found in "Best practices, recommendations, and considerations" on page 106.

SRDF/S

SRDF *synchronous* (SRDF/S) is a disaster-restart solution that operates in synchronous mode and maintains a (synchronous) mirrored copy of production data (R1 devices) in a physically separated Symmetrix array (R2 devices) within an SRDF configuration. SRDF/S is a building block of several multisite disaster-restart options such as SRDF/Star and SRDF/AR.. Figure 3 on page 20 shows this SRDF mode of operation.

① I/O write received from host / server into primary cache
② I/O is transmitted to secondary cache
③ Receipt acknowledgment is provided by secondary back to primary cache
④ Ending status is presented to host / server
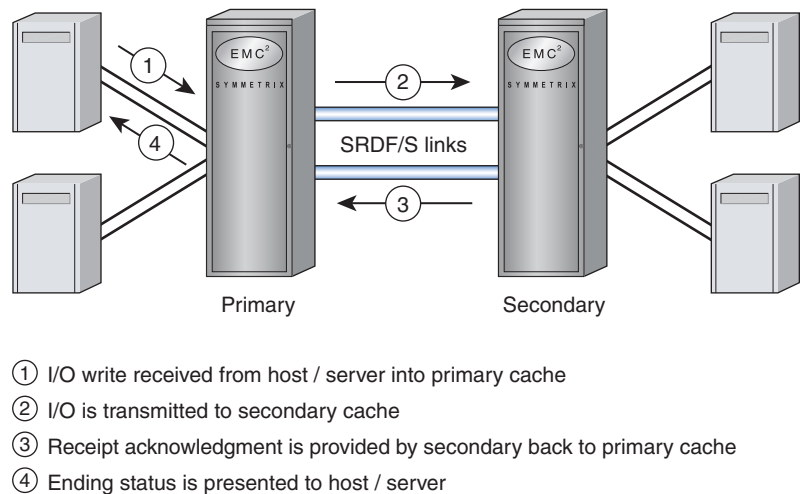
SYM-002508

Figure 3   SRDF/S overview

SRDF/S offers the following major features and benefits:

- ◆ High data availability
- ◆ High performance
- ◆ Flexible configurations
- ◆ Host and applications software transparency
- ◆ Automatic recovery from a component or link failure
- ◆ Significantly reduced recovery time after a disaster
- ◆ Reduced backup and recovery costs
- ◆ Reduced disaster recovery complexity, planning, and testing

The SRDF/S operation is transparent to the host operating system and host applications. It does not require additional host software for

duplicating data on the participating Symmetrix units. SRDF/S offers greater flexibility through additional modes of operation, specifically:

◆   Semi-synchronous mode (*no longer supported with SRDF/S*)
◆   Adaptive copy write-pending mode
◆   Adaptive copy disk mode

◆   Asynchronous mode

Synchronous and semi-synchronous modes of operation

Each SRDF base solution operates in a specific mode of operation. SRDF modes of operation are either *primary* or *secondary*. Once set, the primary mode is the default mode of operation for a given SRDF device, range of SRDF volumes, or an SRDF group. SRDF/S is the primary mode of operation.

Synchronous   Available with the SRDF/S product offering, synchronous mode maintains a real-time mirror image of data between the primary and secondary volumes. Data must be successfully stored in both the local and remote Symmetrix systems before an acknowledgement is sent to the primary site host.

> **Note:** In the mainframe PAV/MA or with FBA Multiwrite environment, SRDF synchronous mode supports multiple concurrent write I/Os at the volume level.

Primary modes for SRDF R1 devices are *synchronous* (journal 0) and *semi-synchronous* (journal 1). Synchronous SRDF *writes* from the host to the *primary* do not finish until they are acknowledged at the primary that the complete write is resident in the secondary Symmetrix system's cache. This means that the logical volume is busy at the primary host (including read-following-write operations) throughout the SRDF operation.

Semi-synchronous   Semi-synchronous mode is supported with Enginuity version prior to 5773. Used mainly for the extended distance solution, semi-synchronous mode, shown in Figure 4 on page 22, allows the primary and secondary volumes to be out of synchronization by one write I/O operation. Data must be successfully stored in the Symmetrix system containing the primary volume before an acknowledgement is sent to the local host.
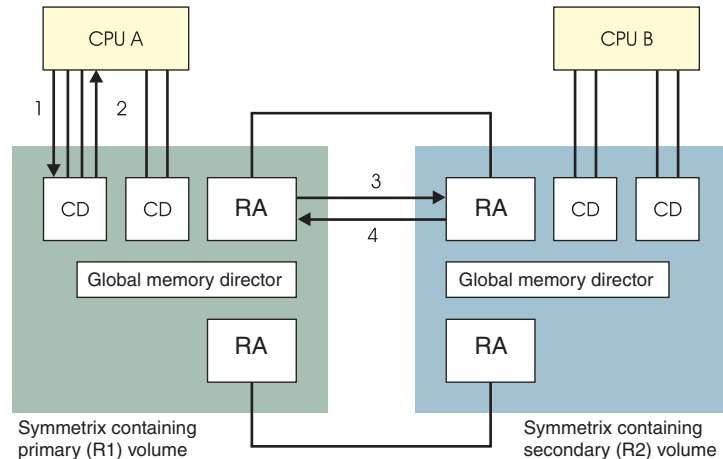
Figure 4 **Semi-synchronous mode**

Semi-synchronous mode will not allow the next write operation to a primary device until a positive acknowledgement is received from the remote Symmetrix system that the first write operation was received in the remote Symmetrix global memory. However, any number of read operations can be performed to the primary device while awaiting acknowledgement of the first write operation.

In semi-synchronous mode, the first write to a primary volume is acknowledged locally, allowing subsequent reads to be accepted. However, a second write to the same volume is deferred until the first write is acknowledged as completed to the secondary Symmetrix system's cache.

Semi-synchronous mode does not maintain dependent write consistency across multiple volumes, which can lead to inconsistencies of the remote copy. It does help relieve read queuing behind an SRDF write operation. Most heavy write volumes, such as database logs and sequential write updates, do not benefit from semi-synchronous mode. Best practice is to place "hot" SRDF volumes that are incurring performance problems into their own SRDF group and run that SRDF group in SRDF/A mode.

Note: Semi-synchronous mode is *not* supported with FICON mainframe host adapters and *not* supported with any host adapter beginning with Symmetrix DMX-3. In the mainframe PAV/MA environment, semi-synchronous mode is not recommended because it provides no performance benefit. SRDF users seeking the increased parallelism offered by PAVs should run SRDF in synchronous mode.

SRDF/A

SRDF/A is a disaster-restart solution that mirrors data from the R1 devices while maintaining a dependent-write consistent copy of the data on the R2 devices at all times. The dependent-write consistent copy of the data at the secondary site is typically only seconds behind the primary site. SRDF/A session data is transferred to the secondary array in cycles. This mechanism eliminates the redundancy caused by multiple same-track changes written within the same cycle being transferred over the SRDF links and potentially reduces network bandwidth requirements. SRDF/A provides a long-distance replication solution. This level of protection is intended for users who require a fast host response time while maintaining a dependent-write consistent, restartable image of data at the secondary site.

To guarantee the high levels of data consistency, disaster-restart protection, and low I/O response time of SRDF/A, the following advanced SRDF/A features are available:

- SRDF/A Multi Session Consistency (SRDF/A MSC)

- Tunable cache utilization

- Reserve Capacity (Transmit Idle and Delta Set Extension)

- Transfer log

- Write folding

- Write pacing

- Tolerance mode

For more information on these SRDF/A features, refer to the appropriate EMC Symmetrix Remote Data Facility (SRDF) product guides, located at https://support.emc.com, Symmetrix Remote Data Facility (SRDF) product page, Documentation.

SRDF/DM

SRDF/DM is a two-site SRDF data migration and replication solution that operates in only adaptive copy modes. SRDF/DM enables fast data transfer from R1 to R2 devices over extended distances.

SRDF adaptive copy

SRDF *adaptive copy* modes provide long-distance bulk data transfers for data center relocations and content replication. Figure 5 shows SRDF adaptive copy mode of operation.
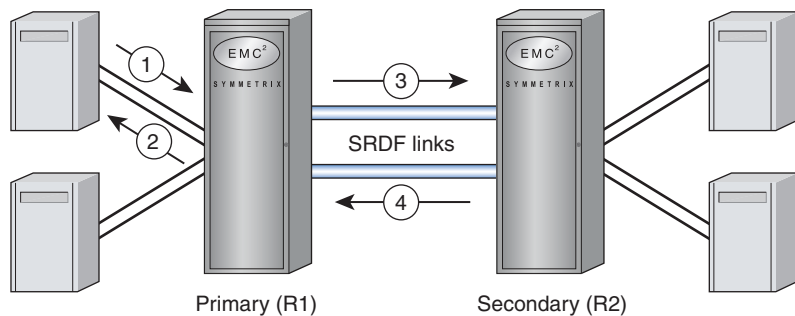
1. I/O write received from host / server into primary cache
2. Ending status is presented to host / server
   If Write Pending (WP) mode: Data stays in cache until DA finds WP track and schedules destage to local disk and RDF copy task to R2
   If Disk mode: Track marked Invalid R2 until DA schedules RDF copy task to R2. Data in cache slot is freed.
3. I/O is transmitted across SRDF link to secondary cache
4. Reciept acknowledgment from secondary RA will remove WP R2 or Invalid R2 state for primary cache slot

Figure 5   SRDF adaptive copy overview

For information on SRDF adaptive copy, refer to the appropriate EMC Symmetrix Remote Data Facility (SRDF) product guides, located at https://support.emc.com, Symmetrix Remote Data Facility (SRDF) product page, Documentation.

SRDF adaptive copy disk

Figure 6 on page 25 depicts the process timeline for SRDF adaptive copy disk mode. The host director (HA) writes data to the local (R1) cache, sets R2 Invalid Track flag, and completes the host write. After destaging of the data, the DA scans for invalid tracks and sends a copy request to the RA to copy the track to R2. If in adaptive copy W/P mode, HA sets R2 W/P, instead of Invalid Track flag, and the data is held in local global memory (cache) until the DA recognizes the R2 W/P flag, sends a copy request to one of the RAs, and the W/P slot(s) (which may not be a full track, as in disk mode) are copied to target R2 cache.
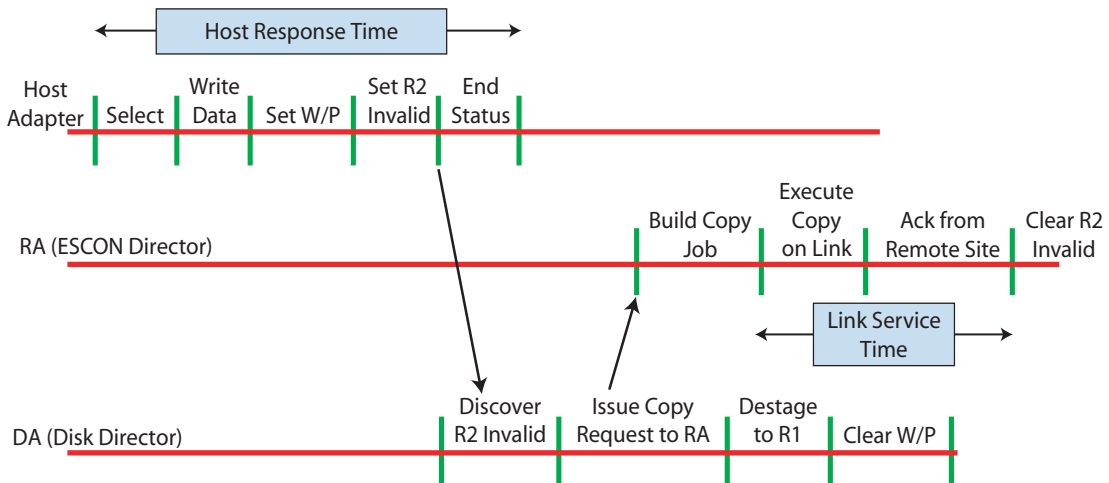
➣ Data is held in local cache until destaged to source R1 volume.

➣ R2 invalid tracks are retained until track sent across RDF link to target cache.

➣ Only full tracks (32 KB for open systems, 64 KB for DMX-3 or later, 56 KB for mainframe) are sent across the link.

➣ For adaptive copy - write pending mode, R2 write pending bit is set and cleared after write to R2.

Figure 6   Adaptive copy disk mode timeline

SRDF family options

There are a number of additional options and features that can be added to the base solutions to solve specific service level requirements. These options include:

◆ SRDF/Automated Replication (SRDF/AR) solutions for meeting very specific, remote replication service-level requirements.

◆ SRDF/Star for advanced multisite failover with continuous protection.

◆ SRDF/Consistency Groups (SRDF/CG) for data consistency.

◆ SRDF/Cluster Enabler (SRDF/CE) for integration with host-based clustering products such as Microsoft Failover Clusters.

For information on other SRDF options, concepts, and operations, refer to the appropriate EMC Symmetrix Remote Data Facility (SRDF) product guides, located at https://support.emc.com, Symmetrix Remote Data Facility (SRDF) product page, Documentation.

Note: Always refer to these product guides for the most current support information.

Figure 7 on page 27 provides a description and SRDF site configuration for these four options.

Description	SRDF site configuration
SRDF/AR • Combines SRDF/DM and SRDF/S • Unlimited distance • Disaster restart solution • Requires EMC TimeFinder®	
SRDF/ CG • SRDF/A and/or SRDF/S mode • Used when co-dependent data resides on multiple physical locations • Ensures a complete, restartable point-in-time remote copy	
SRDF/Star • SRDF/CG with SRDF/S and SRDF/A • Three-site disaster recovery solution providing a zero data loss recovery • Cascaded SRDF/Star (top) • SRDF/EDP with DL R21 to reduce the cost of drives (top) • Concurrent SRDF/Star (bottom)	
SRDF/CE • Complete solution for restarting operations in cluster environments such as MSCS with Microsoft Failover Clusters • Can be implemented with SRDF/S and SRDF/A	

Figure 7   SRDF options

Enginuity/Symmetrix hardware changes

Table 1 summarizes changes for EMC Symmetrix DMX® and EMC Symmetrix VMAX® systems since EMC Enginuity version 5671 that affect SRDF.

Table 1   Enginuity changes affecting SRDF (page 1 of 2)

Feature	Hardware/ Feature type	DMX-1/2 5671	DMX-3/4 5773	VMAX 5874	VMAX 5875	VMAX 40K, VMAX 20K/ VMAX 5876	VMAX 10K 5876
Device support for SRDF	SRDF	64K	64K	64K	64K	64K	8K
Device limit per RA	SRDF	32K	64K	64K	64K	64K	8K
SRDF groups per system	SRDF	64	128	250	250	250	32
SRDF groups per port	SRDF	8	32	64	64	64	32
Fibre Channel port	Hardware	1/2/4 Gb	1/2/4 Gb	2/4/8 Gb	2/4/8 Gb	2/4/8 Gb	2/4/8 Gb
GbE port (Short wave optical port only)	Hardware	1 Gb	1 Gb	1 Gb	1 /10 Gb	1 /10 Gb	1 /10 Gb
Maximum number of SRDF Ports per system	Hardware	8	8	32	32	32	8
Minimum SRDF/A Cycle Times	SRDF	5 secs and 15 secs with MSC	1 sec and 3 secs with MSC	1 sec and 3 secs with MSC	1 sec and 3 secs with MSC	1 sec and 3 secs with MSC	1 sec and 3 secs with MSC
Multi-track write support - *streaming of large sequential writes improves batch and DB reorg performance*	SRDF	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled
SRDF Delta Set Extension	SRDF	Supported	Supported	Supported	Supported	Supported	Supported
Mainframe Multiple Writes per volume with PAV	SRDF	Supported	Supported	Supported	Supported	Supported	Not Supported
SRDF static flow control:	SRDF						
Gigabit Ethernet		Enabled	Enabled	Enabled	Disabled	Disabled	Disabled
Fibre Channel		Supported	Supported	Supported	Disabled	Disabled	Disabled
SRDF dynamic flow control:	SRDF						
Gigabit Ethernet		N/A	N/A	N/A	Enabled	Enabled	Enabled
Fibre Channel		N/A	N/A	N/A	Enabled	Enabled	Enabled
SRDF-A dynamic parameters: Minimum Cycle Time; Maximum Host Throttle Time; Cache utilization	SRDF/A	Supported	Supported	Supported	Supported	Supported	Supported

Table 1   Enginuity changes affecting SRDF (page 2 of 2)

Feature	Hardware/ Feature type	DMX-1/2 5671	DMX-3/4 5773	VMAX 5874	VMAX 5875	VMAX 40K, VMAX 20K/ VMAX 5876	VMAX 10K 5876
Transmit Idle	SRDF	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled
Fibre Channel Single Round Trip (SiRT)	SRDF/A	Supported	Enabled	Enabled	Enabled	Enabled	Enabled
GbE compression [c]	Hardware	Supported	Supported	Supported	Supported	Supported	Not Supported
	Software	N/A	N/A	Supported	Supported	Supported	Supported
Fibre Channel compression [c]	Software	N/A	N/A	Supported	Supported	Supported	Supported
	Hardware	N/A	N/A	N/A	N/A	Supported on VMAX 40K only	Not Supported
64K track size for OS: • (2) 32K writes from DMX-3 or later models to DMX-2 or earlier models • Mainframe still uses 56K (no impact)	SRDF	Not supported	Enabled	Enabled	Enabled	Enabled	Enabled
IPv6 feature on GbE [b]	SRDF	N/A	Supported	Supported	Supported	Supported	Supported
IPsec encryption on 1 GbE ports [b], [c]	SRDF	N/A	Supported	Supported	Supported	Supported	Supported
	Hardware	N/A	Supported	Supported	Supported	Supported	Supported
Open Systems Multiple Writes per volume (<8)	SRDF	N/A	Supported	Supported	Supported	Supported	Supported
Mixed SRDF modes on SRDF ports for 5876	SRDF	N/A	N/A	N/A	N/A	Supported	Supported
Dynamic copy performance for adaptive copy	SRDF	N/A	N/A	N/A	N/A	Enabled	Enabled

a. For 64K device support, you need to have a minimum of two SRDF groups.

b. Requires new MPCD card for IPsec/IPV6 support on DMX: 293-801-982 = 4 GE; 293-801-984 = 2 FC + 2 GE; 293-801-986 = (3) FICON + (1) GbE, in place of 293-801-972 to -976 card for IPv4.

c. Hardware compression is an optional feature that requires different hardware with compression/encryption support. Contact your EMC Service Representative for details.

2

Symmetrix SRDF Hardware

This chapter contains the following information on Symmetrix SRDF hardware:

Symmetrix VMAX Series systems

EMC Symmetrix VMAX systems are enterprise-class storage platforms intended for open systems and mainframe computing.

Note: The VMAX 10K Series does not support mainframe.

This section contains the following information:

◆ "Symmetrix VMAX Series engines" on page 32

◆ "Host connectivity and front-end ports" on page 36

Refer to the appropriate Symmetrix product guides, located at https://support.emc.com, for detailed information.

Symmetrix VMAX Series engines

This section provides a brief description of the VMAX Series engines:

◆ "VMAX 40K" on page 32

◆ "VMAX 20K/VMAX" on page 34

◆ "VMAX 10K Series" on page 35

If a system has a single engine, physical memory mirrors are internal to the engine.

If a system has multiple engines, physical memory mirrors are provided between engines.

Dual write technology is maintained by the system. In the event of a director or memory failure, data is obtained from the redundant copy.

VMAX 40K The Symmetrix VMAX 40K system is the highest-performance, highest-capacity storage system designed for the hybrid cloud and other demanding data center environments. The Symmetrix VMAX 40K is a distributed multi-engine storage system that can non-disruptively scale from one to eight VMAX 40K Engines. Systems are configured around a central system bay and up to 10 adjacent storage bays integrating standard 3.5" drives or high-density 2.5" drives for space or power constrained data centers. Each drive array enclosure (DAE) can support up from 25 2.5" drives or 15 3.5" drives.

The VMAX 40K also offers the ability to mix standard and high density drives within a system (Enginuity 5876 Q4 2012 SR required). For example, mixed can be in 8 adjacent storage bays (seven high density storage bays and one standard storage bay).

Symmetrix VMAX 40K dispersed system bays allow a single VMAX 40K to be separated across two locations by up to 82 feet (25 meters) enabling deployment in dense data center environments where floor loading and other physical limitations and obstructions would otherwise limit configurations.

VMAX 40K systems have high-performance engines, including six-core, x86 family processors that run a maximum of 3.066 GHz.

Figure 9 shows a Symmetrix VMAX 40K engine.
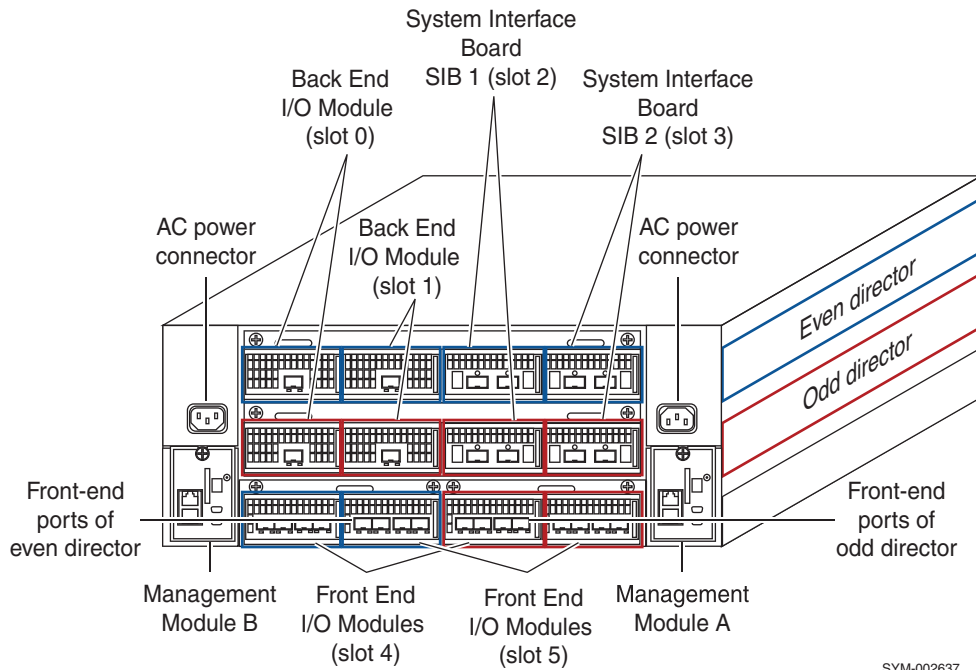
Figure 8   Symmetrix VMAX 40K engine

For detailed information on the VMAX 40K, refer to the *EMC Symmetrix VMAX 40K Series Product Guide*. All planning tasks, including space, power and cooling requirements, are described in the *EMC Symmetrix VMAX 40K Series Planning Guide*. These documents are located at https://support.emc.com.

VMAX 20K/VMAX   The Symmetrix VMAX 20K/VMAX system is the high-end, scalable storage array with a system bay and separate rollup storage bays.

The system scales from a single Symmetrix VMAX 20K/VMAX Engine system with one storage bay to a large eight engine system with a maximum of ten storage bays. Online system upgrades are achieved by adding single or multiple engines or additional storage bays. Each engine contains two Symmetrix VMAX 20K/VMAX directors with extensive CPU processing power, physical memory, front-end ports, and back-end ports. Drive capacity is increased by installing 4 Gb/s disk array enclosures (DAEs) to the storage bay.

VMAX 20K/VMAX systems have high-performance engines, including quad-core, x86 family processors that run a maximum of 2.33 GHz.

Figure 9 shows a Symmetrix VMAX 20K/VMAX engine.
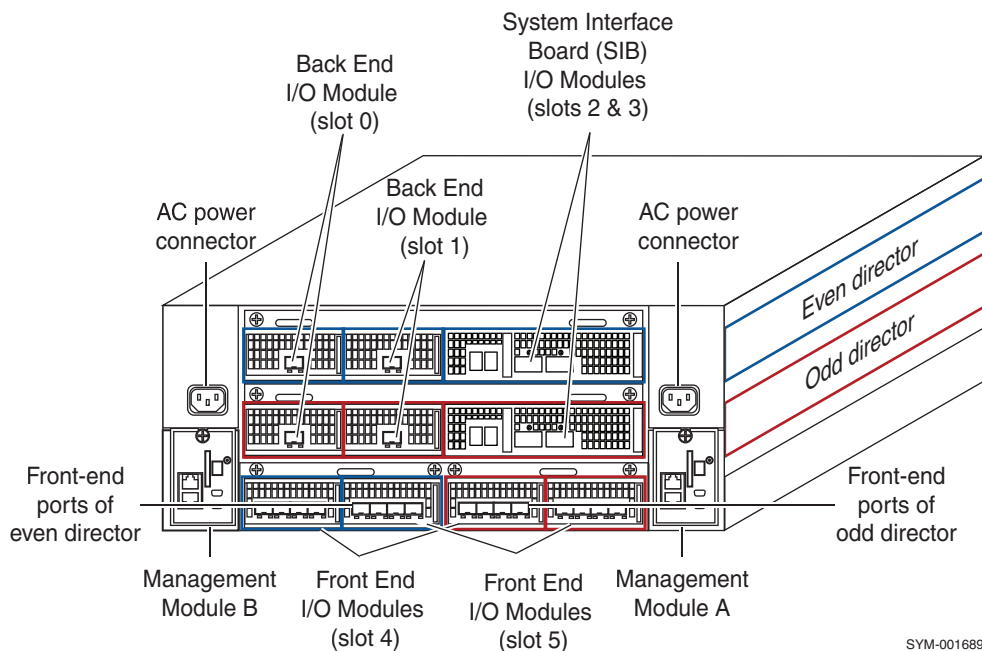
Figure 9   Symmetrix VMAX 20K/VMAX engine

For detailed information on the VMAX 20K, refer to the *EMC Symmetrix VMAX 20K Series Product Guide*. All planning tasks, including space, power and cooling requirements, are described in the *EMC Symmetrix VMAX 20K Series Planning Guide*.These documents are located at https://support.emc.com.

VMAX 10K Series   The EMC Symmetrix VMAX 10K Series includes systems with Serial Number xxx987xxxx, systems with Serial Number xxx959xxxx, and VMAXe. In this document, they are referred to as VMAX 10K Series. However, each model has its own product guides, which can be found at https://support.emc.com.

VMAX 10K systems is a multi-controller, scale-out architecture with consolidation and efficiency for the enterprise with one to four system bays (one to four VMAX 10K Engines) and up to three storage bays. The system scales from a single engine (in one system bay) with no storage bays to a four engine (in four system bays) with three additional storage bays. The Symmetrix VMAX 10K can be dispersed by up to 10 meters (32 feet) to solve floor-loading problems or to work around obstacles in the data center.

VMAX 10K systems have high-performance engines, including quad-core, x86 family processors that run at a maximum of 2.4 GHz.

The VMAX 10K Series engine is shown in Figure 10.
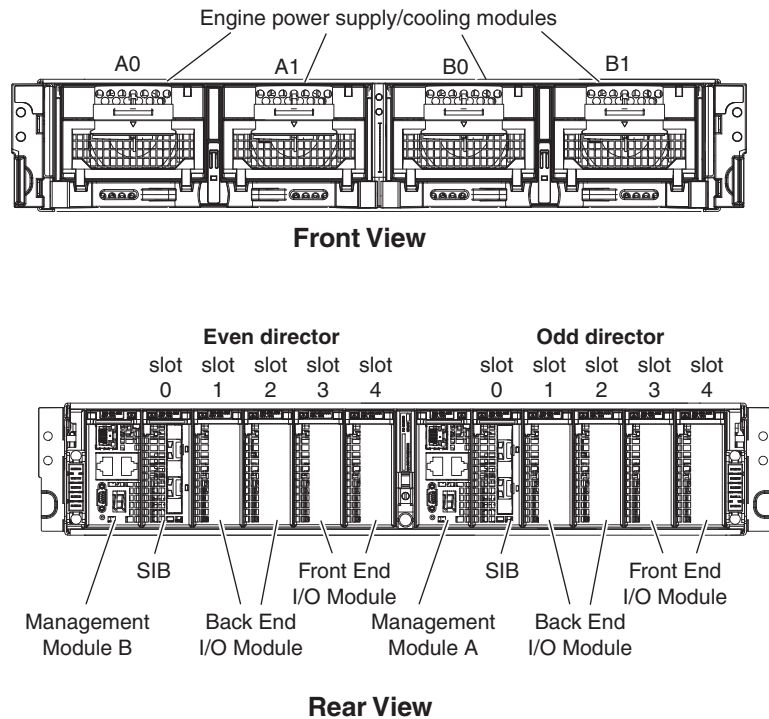
Figure 10 **VMAX 10K Series engine**

For detailed information on the VMAX 10K, refer to the appropriate EMC Symmetrix VMAX 10K Series product guides. All planning tasks, including space, power and cooling requirements, are described in the *EMC Symmetrix VMAX 10K Series Planning Guide*. These documents are located at https://support.emc.com.

Host connectivity and front-end ports

This section contains basic host connectivity and front-end port information for:

VMAX 40K and VMAX 20K/VMAX   The Symmetrix VMAX 40K supports dual connections to the host or server by allowing multiple connections from a host (direct attached) or a Fibre Channel switch (SAN) to a Symmetrix VMAX 40K.

Multiple connections are distributed across front-end I/O modules and engines to ensure uninterrupted access in the event of an I/O failure. A minimum of two connections per server or SAN is required to provide full redundancy.

Symmetrix VMAX 40K systems provide I/O modules that support:

◆ Fibre Channel/FCoE/iSCSI for host, SRDF, Recover Point, and Federated Tiered Storage connectivity in open systems environments

◆ FICON for host connections in mainframe environments, and FC or GbE when SRDF is used.

Front-end I/O modules are ordered in pairs and installed in engine slots 4 and 5. To provide redundancy, front-end I/O modules connect to alternate directors. Each engine supports four I/O modules (two pairs).

Front-end I/O module configurations can be found in the appropriate VMAX 40K or VMAX 20K/VMAX Series product guides, located at https://support.emc.com in the Symmetrix VMAX 40K Series Documentation Portfolio or Symmetrix VMAX 20K/VMAX Series Documentation Portfolio.

Front-end port diagrams for VMAX 40K can be found at https://support.emc.com in the Symmetrix VMAX 40K Series Documentation Portfolio.

Front-end port diagrams for VMAX20K/VMAX can be found at https://support.emc.com in the Symmetrix VMAX 20K/VMAX Series Documentation Portfolio.

VMAX 10K Series   The Symmetrix VMAX 10K supports dual connections to the host or server by allowing multiple connections from a host (direct attached) or a Fibre Channel switch (SAN) to a Symmetrix VMAX 10K. Multiple connections are distributed across Front End I/O Modules and engines to ensure uninterrupted access in the event of an I/O failure.

Front-end I/O module configuration examples can be found in the appropriate VMAX 10K Series product guides, located at https://support.emc.com.

Front-end port diagrams can be found at https://support.emc.com in the appropriate Symmetrix VMAX 10K Series Documentation Portfolio.

Symmetrix DMX Series systems and earlier

This section contains the following information on the Symmetrix DMX Series, and earlier, systems:

SRDF director hardware

SRDF director hardware consists of a director and adapter board set. This hardware provides the communications physical layer for SRDF data and information exchanges between Symmetrix systems. SRDF director hardware includes any of four types of director/adapter board sets, depending on the protocol:

◆ ESCON remote adapter (RA), page 40.

> **Note:** *ESCON is not supported with Enginuity version 5874.*

◆ Fibre Channel remote adapter (RF), page 41.

◆ GbE remote adapter (RE), page 43.

◆ Multiprotocol Channel Director (MPCD), supported with Symmetrix DMX series, and configurable for SRDF over Gigabit Ethernet (GbE), page 44. With DMX-3, Fibre Channel was added. The MPCD is also configurable for iSCSI, FICON, or combinations of these protocols.

The SRDF network options listed above can be used for any host environment. SRDF uses a storage protocol based on the Gigabit Ethernet, Fibre Channel FC-4, or ESCON specifications to remotely mirror data between Symmetrix systems. The host attachment, I/O protocol, and disk data structures that each host requires are independent of the SRDF operation between Symmetrix systems.

Note: Some restrictions apply in mixed environments with iSeries and other host types. For more information, consult EMC Customer Support.

⚠️ **CAUTION**

EMC recommends that the Symmetrix system has at least two director board sets for SRDF protection; one processor on each board for SRDF to provide redundant links in case a communications link fails or in the unlikely event a director fails. Having two director boards for SRDF avoids a potential single point of failure if one processor on the director is used for SRDF and the other processor(s) is/are used to connect to a host.

SRDF director functions

The SRDF director/adapter board sets, further described in "Symmetrix DMX Series systems and earlier" on page 38, provide the link connections. Along with the fiber-optic protocol support, they are integral in controlling communications between two Symmetrix systems in an SRDF configuration.

GbE remote directors (RE) or Fibre Channel remote directors (RF) maintain a peer-to-peer relationship at the transport layer of communications.

The ESCON remote director (RA) board set that normally sends data across an SRDF link is known as an *RA-1*. An RA-1 functions like a host channel interface. The ESCON RA board set that normally receives data sent across an SRDF link is known as an *RA-2*. An RA-2 functions like a storage director interface.

An RA-1 and its corresponding RA-2 are known as an *RA pair*.

◆ With Symmetrix 3xxx or 5xxx models, there can be multiple RA pairs in an SRDF configuration, up to a maximum of 16 pairs.

◆ With Symmetrix 8xxx models, an optional four-processor ESCON board can be used for SRDF.

◆ With Symmetrix DMX-4 models, an optional eight-port four-processor ESCON board can be used for SRDF.

◆ ESCON is *not* supported with Symmetrix VMAX and Symmetrix VMAX SE systems.

Note: Each Symmetrix product guide, located at https://support.emc.com, Symmetrix Remote Data Facility (SRDF) product page, Documentation, provides detailed descriptions of the SRDF-supported hardware functionality.

ESCON director (RA)

Note: ESCON is not supported with Enginuity version 5874 and later.

A Symmetrix 8000 RA contains four individual processors (CPUs) (Figure 11 on page 41). A Symmetrix 3x30/5x30 RA contains two processors. Each processor can provide one SRDF port connection. On one board, a processor can be configured with host interface emulation while the other processor(s) is configured with SRDF emulation. Alternatively, all processors on the board can be configured for SRDF emulation. All ESCON ports are 1310 nm multimode.

Each processor provides two host connections if not configured for SRDF. *If one port is configured for SRDF, the other port on that processor cannot be used.*

The Symmetrix DMX ESCON director contains four processors (CPUs) and eight interfaces to host mainframe systems or four interfaces for SRDF connections. Each processor can provide one SRDF port connection. On one board, a processor can be configured with host interface emulation while the other processor(s) is configured with SRDF emulation. Alternatively, all processors on the board can be configured for SRDF emulation.
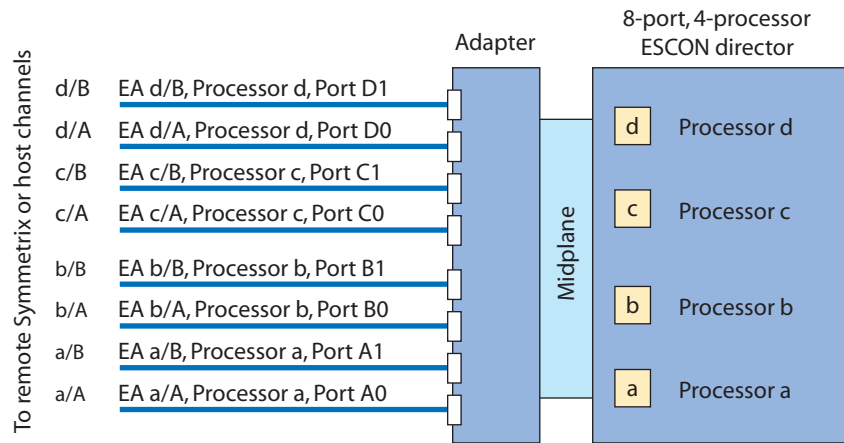
GEN-000204

Figure 11   DMX 8-port ESCON director designations

Fibre Channel director (RF)

Enginuity version 5x66 and later supports SRDF Fibre Channel emulation.

Note: All Fibre Channel director ports have seven (7) Buffer-to-Buffer Credits (BB_Credits).

Symmetrix RF directors include:

◆ **Four-Port/Two-Processor RF** — Symmetrix 8000 series systems support a four-port RF that supports SRDF operations on a single port on each processor.

One processor on a board can be configured with host interface emulation while the other is configured with SRDF emulation. Alternatively, both processors can be configured for SRDF operation. Each processor provides two host connections if not configured for SRDF.

Note: If one port is configured for SRDF, the other port on that processor cannot be used.

◆ **Eight-Port/Four-Processor RF** — Symmetrix DMX-series Fibre Channel directors support up to four SRDF/Fibre Channel links (one per processor or slice) at 4 Gb/s (see Figure 12 on page 42.) The Fibre Channel ports will auto-negotiate to match 1 Gb/s ports on older Symmetrix models or Fibre Channel switches. Each processor provides two host connections if not configured for SRDF.

Note: If one port is configured for SRDF, the other port on that processor cannot be used. SRDF must be configured using Port 0 on each processor slice, starting with D0 downwards.

GEN-000205
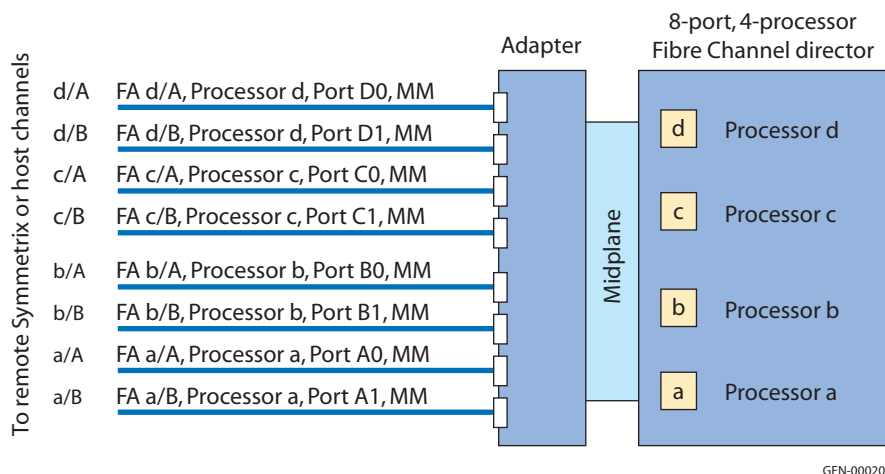
Figure 12   DMX 8-port Fibre Channel front-end host director designations

Notes   Example of director type designation:

FA =	Fibre Channel director
FA d/A =	Defines Processor d, Port A
Port A =	Equivalent to Processor d, Port DO
MM =	Multi-mode
SM =	Single mode

Gigabit Ethernet director (RE) on Symmetrix 8000

The Symmetrix 8000 series (Figure 13) supports a two-processor, two-port Gigabit Ethernet (GbE) card, with the following features:

◆ Native SRDF IP support

◆ Model DP3-GBENET — Two GbE ports per card

◆ Multiple TCP connections per GbE port

◆ All 850 nm multimode LC ports

◆ 256 KB buffers for extended distance (TCP windowing)

Both ports on the GbE card can be used *only* as SRDF GbE ports. No other emulations are supported.

Channel director operator panel designations

Physical channel designations

Adapter

2-port, 2-processor GigE remote channel director

To remote Symmetrix

b/A EF b/A, Processor b, Port A | A | b Processor b

a/A EF a/A, Processor a, Port A | A | a Processor a

System buses:
Odd-numbered directors connect to the top-high and bottom-low buses.

Even-numbered directors connect to the top-low and bottom-high buses.

GEN-000207

Figure 13   Symmetrix 8000 2-port GbE remote director channel designations

Multiprotocol Channel Director on the Symmetrix DMX series

The Symmetrix DMX series supports the four-processor/two-port and the four-processor/four-port Multiprotocol Channel Director (MPCD) configured for GbE (MPCD-GbE), iSCSI, or FICON. This section discusses Multiprotocol Channel Directors on:

◆ "DMX-1 and DMX-2" on page 44
◆ "DMX-3" on page 45
◆ "DMX-4" on page 46

DMX-1 and DMX-2
On a four-port director, for DMX-1 and DMX-2, the ports can be intermixed with SRDF (GbE) and host (iSCSI or FICON) emulations, as shown in Figure 14.
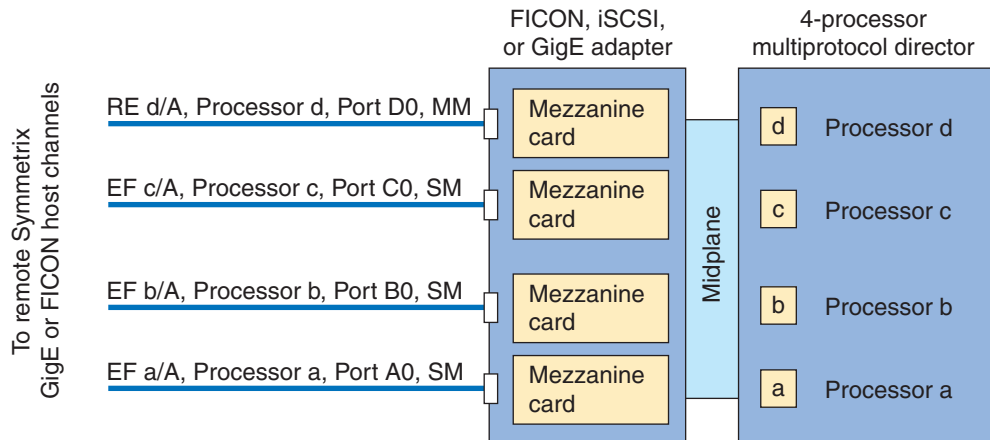
GEN-000206

Figure 14   DMX-1/2 Multiprotocol Director with FICON, iSCSI, and GbE remote director designations (one MM and three SM ports)

Notes
Example of director type designation:

Port A =	Equivalent to Processor d, Port DO
MM =	Multimode
SM =	Single mode

This adapter can be configured in the combinations of FICON, iSCSI, and GbE port assignments, as listed in Table 2:

Table 2   Combination port assignments

Processor	FICON, iSCSI, and GbE port assignments	
d	GbE MM	iSCSI MM
c	FICON SM	FICON SM
b	FICON SM	FICON SM
a	FICON SM	FICON SM

For more information, refer to the "Specifications" appendix and the "Planning and Installation" appendix in the DMX800, DMX1000, DMX2000, or DMX3000 product guides,located at https://support.emc.com, Symmetrix product page, Documentation.

DMX-3   The Multiprotocol Director for DMX-3 added support for Fibre Channel (FA) and SRDF over Fibre Channel (RF), as shown in Figure 15 on page 45. The FA ports can be configured two ports per processor, while SRDF always requires a dedicated processor. RF links will use Port 0, while Port 1 of the processor (or slice) will be disabled.

When SRDF over Fibre Channel (RF) ports are configured along with host FA ports, the SRDF RF ports will always start with a D0 port and end with an A0 port.
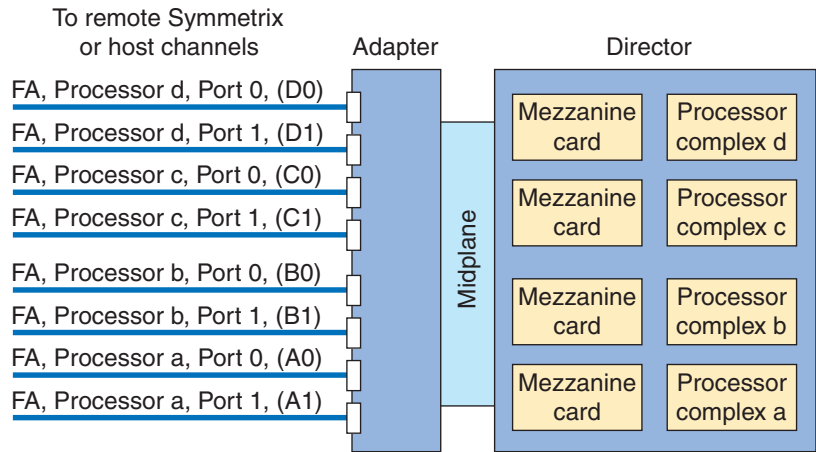
Figure 15   DMX-3 Multiprotocol Director with Fibre Channel, GbE, and iSCSI, and FICON port designations

Notes

Example of director type designations:

FA =	Fibre Channel director
EF =	FICON Channel director
SE =	iSCSI Channel director
RF=	SRDF over Fibre Channel director
RE =	GbE Remote director
MM =	Multimode
SM =	Single mode

For more information, refer to the *EMC Symmetrix DMX-3 Product Guide* and the *EMC Symmetrix DMX-3 Physical Planning Guide*, located at https://support.emc.com, Symmetrix product page, Documentation

DMX-4 Figure 16 shows the port designations for the DMX-4 Multiprotocol Director.
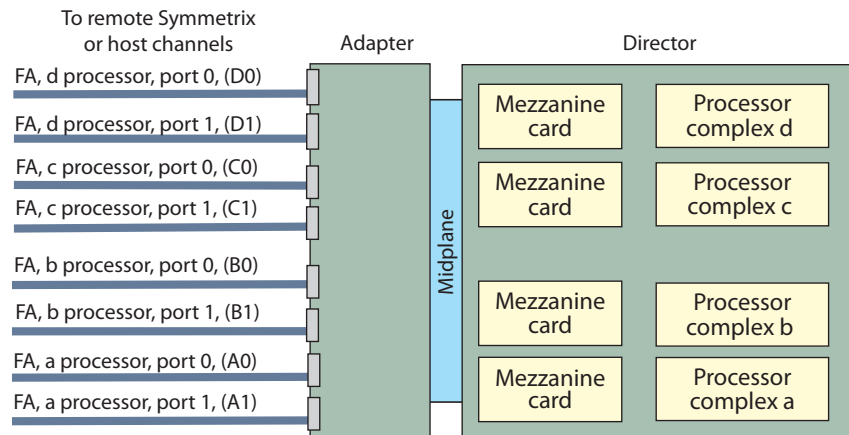
Figure 16 **DMX-4 Multiprotocol Director with Fibre Channel, GbE, and iSCSI, and FICON port designations**

Notes

Example of director type designations:

FA =	Fibre Channel director
EF =	FICON Channel director
SE =	iSCSI Channel director
RE =	GbE Remote director
MM =	Multimode
SM =	Single mode
SE 03d =	iSCSI director 03, processor d, which is located in slot 2 of the card cage

For more information, refer to the *EMC Symmetrix DMX-4 Product Guide* and the *EMC Symmetrix DMX-4 Physical Planning Guide*, located at https://support.emc.com, Symmetrix product page, Documentation.

3

SRDF Network Protocols

SRDF uses Gigabit Ethernet, Fibre Channel, or ESCON network protocols to move data across the SRDF links. This chapter discusses each of these protocols:

Gigabit Ethernet

This section contains information on implementing SRDF over Gigabit Ethernet and IP.

Generic considerations

Latency   Latency is generally referred to in milliseconds (ms) and as the combined round trip time (RTT) between primary and secondary Symmetrix systems. Latency should be as low as possible at the given distances separating the sites. The round trip latency of the SRDF network is the time (ms) it takes for a full-size Ethernet packet (1500 bytes) to travel from the GbE port in one Symmetrix, over an IP network, to the Gigabit port in another Symmetrix, and back again. One millisecond is equal to 0.001 seconds; therefore, 1000 ms is equal to one second. Therefore, it takes roughly 1 ms for light to travel 200 km or 125 circuit miles one way.

Flow control   Flow controls are necessary because senders and receivers are often unmatched in capacity and processing power. A receiver might not be able to process packets at the same speed as the sender, causing the receiver to send a pause message to its link partner to temporarily reduce the amount of data it's transmitting. With Gigabit Ethernet, if buffers fill, packets are dropped. The goal of flow-control mechanisms is to prevent dropped packets that then must be retransmitted.

Note: In order for flow control to work properly with Gigabit Ethernet switch ports, auto negotiation must be **on**.

Starting with Enginuity version 5670, the GbE director has a built-in flow control mechanism to control the amount of data placed in the SRDF queue to be transmitted to the secondary side. The SRDF flow control mechanism keeps track of how many I/Os are being sent across the link, and dynamically adjusts the number of I/Os waiting in the job queue. This prevents I/Os from timing out while in the queue or when sent across the link.

There are two additional ways to control traffic rates to prevent network overflow issues.

- ◆ TCP send/receive queue size
- ◆ Link speed limit settings to match available bandwidth

When packets are lost in transit over the network, TCP has Slow Start and Congestion Avoidance algorithms to reduce the number of packets being sent. When packet loss is encountered, Slow Start will start by sending a single packet and waiting for an ACK (acknowledgement), then send two packets and wait for an ACK, then four packets, and so on. This exponential increase will continue until it reaches a window size of half the size of the one in which the packet loss had occurred. After this halfway value is reached, Congestion Avoidance takes over and the increase in the window size should be at most one packet each round trip time.

This mechanism was designed for handling small numbers of packets in a network with packet loss. When the network becomes congested, packets are dropped. Traditional TCP has a tendency to overreact to packet drops by halving the TCP window size. The resulting reduction in throughput that can occur in traditional TCP implementations is unacceptable to many storage applications.

Starting with Enginuity version 5671, DMX GbE supports the option to disable TCP Slow Start.

Note: This option is only available through EMC Customer Service.

For dedicated bandwidth and loss-less network configurations, it is recommended to disable TCP Slow Start. By disabling TCP Slow Start, the DMX GbE director will not back off to 5% I/O rate during retransmissions. Instead, it will try to maintain I/O rate just below the Speed Limit setting. It is also recommended to set Speed Limit at 90% of available bandwidth per GbE port.

Starting with Enginuity version 5772, TCP Slow Start is disabled by default.

Network quality   A well-designed network should provide latencies that are as low as possible, with the number of hops kept to a minimum given the geographical distance separating the sites.

For SRDF/S, the acceptable round trip time (RTT) is dependent on application tolerance to latency. A best practice, however, is for latency between SRDF ports to not exceed 5ms RTT. Packet loss should never exceed 0.01%.

IMPORTANT

GbE ports and IP networks may add significant overhead (latency) for SRDF/S so the additional O/H should be included in modeling SRDF/S host response time impact.

SRDF/A and SRDF adaptive copy are extensively tested against what is referred to as the *distance envelope*. This involves a series of stringent tests to ensure that SRDF continues to function with latencies to a maximum of 200 ms RTT and a packet loss up to 1%. It is strongly recommended, however, that packet loss not exceed 0.1% for these SRDF modes.

Networks that experience packet loss greater than the recommended figures mentioned above are considered poor-quality links and will severely impact throughput rates. SRDF can be impacted as follows:

◆ SRDF traffic slows down due to packet retransmissions and retries

◆ With SRDF/S, host response times will significantly increase resulting in application timeouts

◆ SRDF links might start continuously bouncing

◆ Symmetrix remote adapter resources are wasted on unhealthy links

Packets should not be delivered out of order. This will result in SRDF link bounces or will create a major degradation in performance. The rapid variation in the latency on the network (known as jitter) should also be as low as possible.

Note: Latency jitter (the variance in latency) should not exceed either 25% of the average network latency or 25ms, which ever is the lower number.

For poor-quality networks and/or high latency (~100 ms) and low bandwidth, it is highly recommended to evaluate the use of WAN optimization devices qualified by EMC E-Lab™ to improve GbE performance.

Note: WAN optimization appliances deployed within IP networks are qualified by EMC. Consult the *EMC Support Matrix* for WAN optimization appliances that can be used with SRDF. For more information on WAN optimization, refer to "WAN optimization" on page 102.

Figure 17 shows an example of the impact of packet loss over varying
degrees of latency. In this example, slow start is enabled.
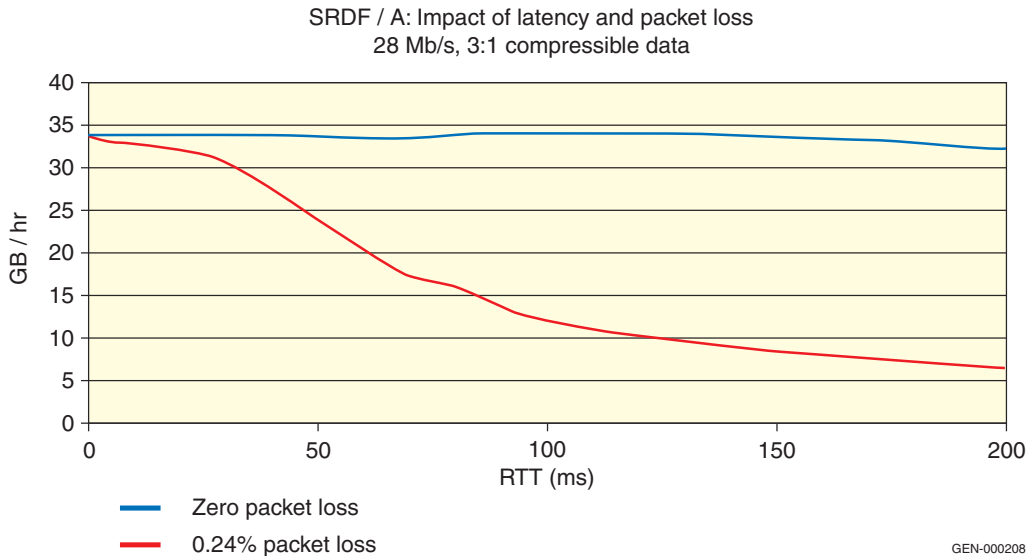
Figure 17   SRDF/A latency and packet loss

MTU size   A standard Ethernet packet is 1500 bytes. Larger packets (jumbo
frames) are any packet larger than 1500 bytes up to 9000 bytes.
Maximum transmission unit (MTU) is the largest TCP packet size
that can be sent across the network without the packet being
fragmented. Packet fragmentation can have a negative impact. For
instance, packets may not be reassembled or packets may be
dropped. Fragmentation can cause such a high processor load on the
device performing the fragmentation (for example, routers,
encryptors, etc.) that it can severally impact (limit) throughput
and/or cause the dropping of packets. To prevent packet
fragmentation, verify that every device in the network path has the
same MTU value set and that the MTU size is not exceeded.

Compression   Data compression is the process of encoding information using fewer
bits than an unencoded representation would use through the use of
specific encoding schemes. This is accomplished through algorithms
looking for data having statistical redundancy. Compression
programs simply eliminate the redundancy. Instead of repeating data
over and over again, a compression program will store the data once
and then refer back to it whenever it appears in the original data
stream.

Dedicated external compression devices are available today and are capable of achieving ratios much greater than devices that are not dedicated to compression. Keep in mind the following characteristics:

◆ Software compression is typically a lower cost option suited more for lower bandwidth environments.

◆ Hardware compression is typically a higher cost option, but more suitable for higher bandwidth environments.

◆ Efficiency of the compression is dependent on the algorithms used and the compressibility of the data (random versus sequential).

Encryption   Encryption devices can be used to ensure the secrecy of data transmitted between two points. Encryption refers to algorithmic schemes that encode plain text into nonreadable form or cyphertext. The receiver of the encrypted text uses a *key* to decrypt the message, returning it to its original plain text form. A key is a randomly or user-generated set of numbers and/or characters used to encrypt and/or decrypt information. There are many types of encryption and not all are reliable.

Note: Third-party encryption devices are not qualified by EMC. These devices can be used with SRDF. Recommended encryption devices can be found in the *EMC Select Catalog* available at https://support.emc.com.

Switched SRDF over GbE   Figure 18 on page 55 is an example of switched SRDF over GbE configuration. The GbE director port establishes logical paths through GbE switches and routers, just like Switched Fibre Channel via Fibre Channel switches. The GbE logical paths are called TCP connections and also support point-to-multipoint configurations. Unlike SRDF over FC, GbE does not require external buffering over distance since each logical path (TCP connection) has its own buffering, as shown in Figure 18.
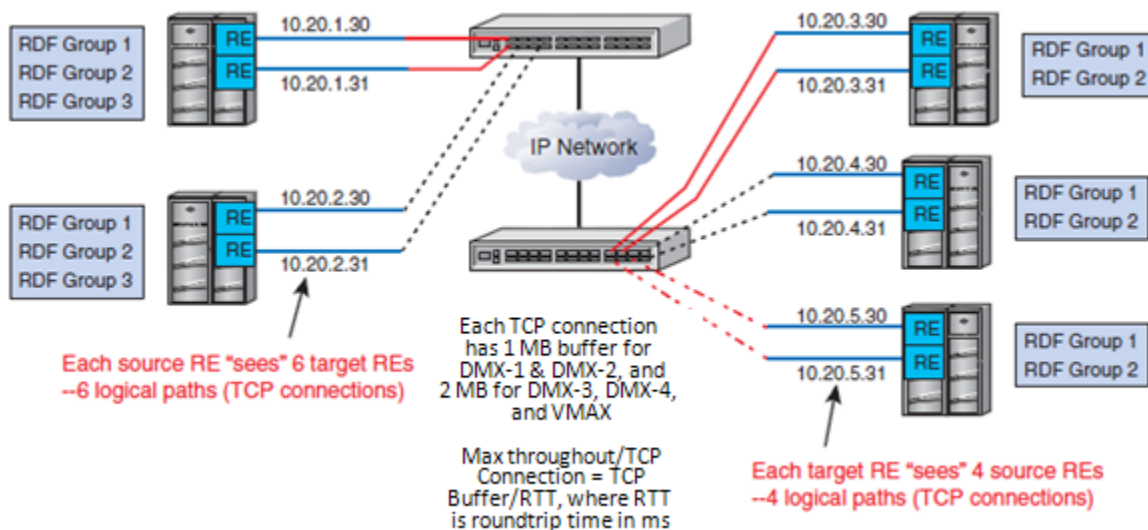
Figure 18   SRDF / GbE in switched network

MPCD and GbE director considerations

For GbE best practices, refer to "GbE" on page 113.

Speed limit GbE ports have the ability to speed limit (throttle) the amount of traffic that it sends out of each port. The speed limit value can be changed dynamically (not persistent unless set at group level) or it can be set in the Symmetrix Configuration file (persistent). Speed limit values can be set in whole number values from 10-1000 Mb. It is best practice recommendation to always set the speed limit when full bandwidth (1 Gb) per port is not available. 10 GbE ports can be speed limited from 10 - 10,000 MBs.

Note: Speed limits are set by EMC Customer Service.

Speed limit can be configured on a per port basis or across a group of ports (recommended) using a new feature in 5772 code called "auto-adjust" speed limit. With auto-adjusting speed limits, you will combine a set of SRDF over GbE directors into a single speed limit group with an aggregate bandwidth for the entire speed limit group. Each SRDF port will automatically set its speed limit based on its share of the aggregate bandwidth, which is the total bandwidth of the

speed limit group divided by the number of active members in the speed limit group. SRDF directors are added to the speed limit group one at a time after the speed limit group is created with its initial SRDF director. A member of a speed limit group can become inactive, for example, if its link fails for a period of time.

If auto-adjust speed limit is not used, it is recommended to set the speed limit/throttle rate on each individual GbE ports. To determine the optimal speed limit for an individual GbE port, first look at the total amount of network bandwidth available for SRDF and then set the speed limit on each port, ensuring that the sum of all the ports and their speed limit values do not exceed the available network bandwidth.

MTU GbE ports support jumbo frames, which must be set in the Symmetrix Configuration file (bin file). However, the entire network, end-to-end, must be configured to support jumbo frames (up to 9000 bytes) or the packets may be fragmented or dropped. The MTU can be lowered to less than 1500 bytes. In cases where VPN encryption or other tunneling occurs within the network, it may be necessary to compensate for the additional overhead caused by these encapsulation protocols. In cases such as these, additional bytes may be added to the packet. The GbE ports will try to negotiate the maximum MTU (as set in the configuration file) across the network.

Note: For Symmetrix 10 GbE ports, it is recommended to set MTU = 9000 bytes, if supported by the network.

Note: Since some devices do not negotiate the MTU with the GbE ports, problems can arise where you may need to hard-set the desired MTU in the Symmetrix Configuration file.

TCP connection scaling In order to get the maximum benefit from scaling, it is important that all GbE ports on the primary and secondary Symmetrix can see each other across the network environment. When Symmetrix GbE director ports are connected to a switched or routed network, a TCP connection is established between all ports. For example, a 4x4 configuration would have 16 TCP connections per side. Performance across multiple TCP sessions scales very well.

Note: VMAX can scale performance nearly linearly (DMX ~ 75%) over distance by adding TCP sessions, using the following rule:

Maximum throughput per RE = n TCP sessions x 2 MB/RTT(ms). Obviously, this is dependent on maximum throughput of RE using available bandwidth and compression rates.

In the example shown in Figure 19, adding TCP sessions improves performance in a network with packet loss.
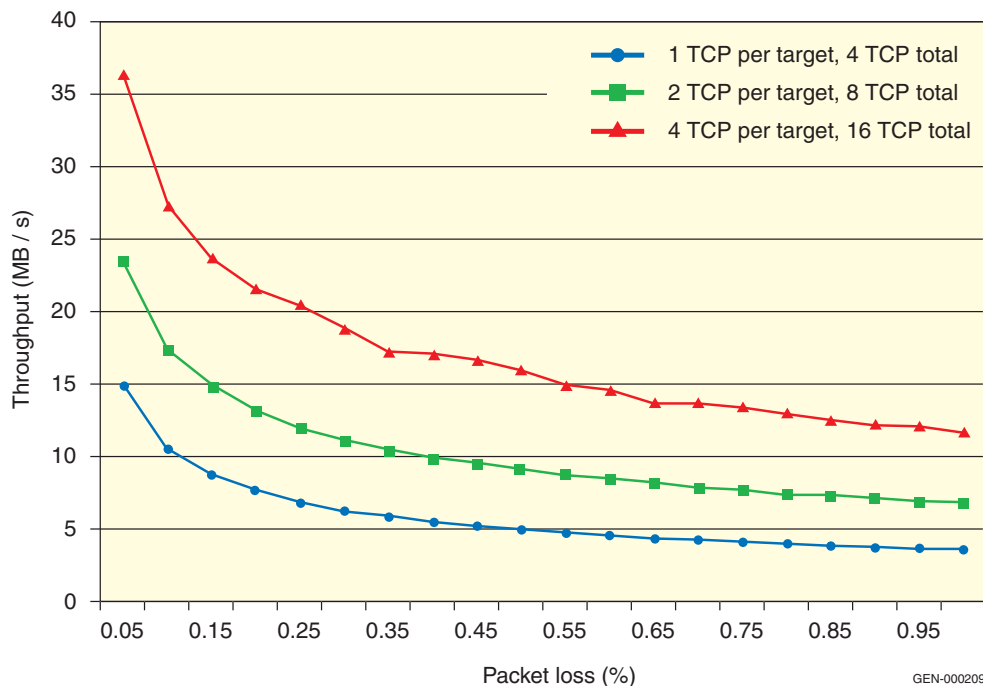
Figure 19   Throughput increase with additional TCP connections example

Network quality   The network quality of an SRDF link is now tested with Enginuity version 5875 ucode and later. VMAX systems on both sides of the link must be running 5875 or later to take advantage of this feature. This feature takes the following actions:

◆   Qualify a questionable SRDF connection before allowing traffic to be sent

◆   Gradually slows down intervals between unsuccessful SRDF link re-establishment attempts

◆ Puts bouncing SRDF links out of service

A qualification test for any SRDF connection can be manually triggered. This feature is supported for both GbE and Fibre Channel SRDF links.

Compression

Compression was first available with the introduction of the Symmetrix DMX. The GbE ports on the MPCD and the GbE director interface boards have a hardware compression chip. Depending on the compressibility of the data, this can lower network bandwidth requirements. The SRDF data is compressed prior to encapsulation into IP packets.

Hardware compression for SRDF traffic over GbE is available through a compression chip:

◆ On the MPCD and GbE director boards (Symmetrix DMX systems)

◆ On the compression-capable Front End I/O Module Carriers (Symmetrix VMAX systems)

Enginuity version 5875 or earlier require that hardware compression for GbE SRDF be enabled in the Symmetrix configuration file.

Beginning with Enginuity version 5875, users can control hardware compression over GbE SRDF links through an additional SRDF group-level attribute by using EMC host-based SRDF software. Hardware compression is enabled only if both the configuration and the SRDF group level setting are set.

Beginning with Enginuity version 5874, software compression can be used along with hardware compression. This may result in better compression, as shown in Figure 20, but needs to be monitored for possible negative effects.

Note: Software and hardware compression stats can only be monitored via Inlines. They are not exported to host-based software.

Figure 20   Software and hardware compression (Enginuity v5874 and later)

Software compression can be enabled at the SRDF group level and SRDF mode of operation. Both sides of the link must support software compression.

Note: Higher host response times can result due to the software compression. If host response time is not a concern, bandwidth savings can be realized through the additional software compression.

Security   IPsec (Internet Protocol Security) and IPv6 are available on the GbE/IPsec Multiprotocol Channel Director beginning with Enginuity version 5773 microcode. The embedded line-grade encryption co-processor for this director provides strong security without degrading performance. Pre-shared encryption keys simplify management while maintaining security for data in transit.

These directors can co-exist in the same system with any existing DMX-4 channel directors. With the new director, IPsec can be enabled on a port-by-port basis. The IP stack version can be similarly configured per port.

The number of IPsec sessions are limited to 100 per port. IPsec enabled and IPsec disabled channel directors can co-exist in the same Symmetrix.

The GbE/IPsec channel director contains an embedded line-grade encryption co-processor to enable IP Security (IPsec) for SDRF-over-IP connections. IPsec and IPv6 capabilities will be configured via Symmetrix Management Console (SMC) or by EMC via the Service Processor. Status and configuration can also be monitored and viewed with SYMCLI commands.

The benefits of this director are as follows:

◆ Reduce the security risks associated with transmitting data over IP networks.

◆ Symmetrix-native encryption hardware eliminates the need for encryption appliances, reducing cost and complexity.

◆ Allows customers to effectively manage Symmetrix communication and encryption with one interface – Symmetrix Management Console (SMC).

◆ Addresses government and corporate mandates for IPv6 capable products.

◆ Support for the industry-leading AES 256, AES 128, and 3DES encryption algorithms.

For additional details on IPsec configuration and encryption, refer to the *Symmetrix 5773 GbE and Multiprotocol IPv4/6 Channel Directors Technical Note* located at https://support.emc.com, Symmetrix product page, Documentation.

Fibre Channel

This section contains information on deploying SRDF over Fibre Channel.

Generic considerations

This section provides information about considerations, common problem areas, limitations, etc., when provisioning SRDF configurations over a Fibre Channel switched fabric topology and direct connect scenarios. The information provided in this section focuses on practical checkpoints that may be helpful when setting up these environments. There are numerous areas that can affect SRDF connectivity and performance (anywhere from the physical to upper application layers).

Latency/round trip conditions

Latency is generally referred to in milliseconds (ms) as the combined round trip time (RTT) between primary and secondary Symmetrix. A FC-Frame by itself takes approximately 1 millisecond to traverse a one-way distance of 200 km from primary-transmitter to secondary-receiver (E_Port to E_Port or F_Port to F_Port).

For example, if two locations are 200 km apart, since standard Fibre Channel protocol requires two round trips for a write I/), then 4 milliseconds of latency (2 X RTT) will be added to the write operation. As more network components are attached to the configuration for pure Fibre Channel environments, latency will naturally increase. This latency can be caused by network components such as HBAs, switches, fiber optics, distance extension devices, as well as factors such as cable purity. The SRDF application layer will contribute additional delays on top of the network components.

Table 3 shows the time and latency for one-way distances.

Table 3   **Time/Latency to distance**

Time/Latency	Distance in kilometers (One way)	Distance in miles (One way)
5 microseconds (us)	1 kilometer (km)	0.62 miles
1 millisecond (ms)	200 kilometers (km)	125 miles

Using Symmetrix tools (inlines, symcli, Symmetrix Management Console, etc.) or external performance monitoring tools, the round trip time can be determined for troubleshooting any SRDF latency issues within the Fibre Channel network.

BB_Credits
Fibre Channel uses the BB_Credits (Buffer-to-Buffer Credits) mechanism for hardware-based flow control. This means that a port has the ability to pace the frame flow into its processing buffers. This mechanism eliminates the need of switching hardware to discard frames due to high congestion. EMC testing has shown this mechanism to be extremely effective in its speed and robustness.

BB_Credit management occurs between any two Fibre Channel ports that are connected. For example:

◆ One N_Port and one F_Port

◆ Two E_Ports

◆ Two N_Ports in a point-to-point topology

The standard provides a frame-acknowledgement mechanism in which an R_RDY (Receiver Ready) primitive is sent from the receiving port to the transmitting port for every available buffer on the receiving side. The transmitting port maintains a count of free receiver buffers, and will continue to send frames if the count is greater than zero.

The algorithm is as follows:

1. The transmitter's count initializes to the BB_Credit value established when the ports exchange parameters at login.

2. The transmitting port decrements the count per transmitted frame.

3. The transmitting port increments the count per R_RDY it receives from the receiving port.

Figure 21 provides a view of the BB_Credit mechanism.
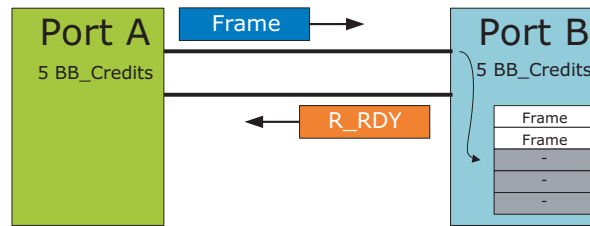
Figure 21   **BB_Credit mechanism**

As viewed from Port A's perspective, when a link is established with Port B, BB_Credit information is exchanged. In this case, Port B provided a BB_Credit count of 5 to Port A. For Port A, this means it can transmit up to five Fibre Channel frames without receiving an R_RDY.

BB_Credit applied to distance

Applying the BB_Credit theory to practical Fibre Channel distance calculations involves another step in which we establish the connection between BB_Credit counts and distance. As a rule of thumb, a typical Fibre Channel frame size of 2 KB occupies approximately 4 km on a long fiber.

BB_Credit calculation for Fibre Channel

Assuming the following is true:

◆ Light propagation in glass is 5 microsec/km, or 5^9 sec/m.
◆ Frame size is 2148 bytes/frame.
◆ Fibre Channel bit rate depends on the Fibre Channel speed.

Speed	Bit rate
2 Gb/s (200 MB/s)	2.125 Gb/s
4 Gb/s (400 MB/s)	4.25 Gb/s
8 Gb/s	8.5 Gb/s
10 Gb/s (1200 MB/s)	10.51875 Gb/s
16 Gb/s	14.025 Gb/s

Note: For more information, refer to the *EMC Networked Storage Topology Guide*, available through E-Lab Interoperability Navigator at: http://elabnavigator.EMC.com.

◆ To calculate for BB_Credits required for 100% utilized link, use the following formulas.

Speed	Formula
2 Gb/s	BB_Credit = ROUNDUP [2 * one-way distance in km/4] * 2
4 Gb/s	BB_Credit = ROUNDUP [2 * one-way distance in km/4] * 4
8 Gb/s	BB_Credit = ROUNDUP [2 * one-way distance in km/4] * 8
10 Gb/s	BB_Credit=ROUNDUP [2 * one-way distance in km/4] * 12
16 Gb/s	BB_Credit = ROUNDUP [2 * one-way distance in km/4] * 16

The factor of **2** in the formulas accounts for the time it takes the light to travel the entire round trip distance: frame from transmitter to receiver and R_RDY back to transmitter.

Table 4 shows the distances for a fully-utilized link based on the number of BB_Credits.

Table 4   BB_Credits compared to distance for a fully-utilized link

BB_Credits for:					Distance for a fully utilized link
2 Gb/s Fibre Channel	4 Gb/s Fibre Channel	8 Gb/s Fibre Channel	10 Gb/s Fibre Channel	16 Gb/s Fibre Channel	
4	8	16	20	32	4 km
16	32	64	80	128	16 km
32	64	128	160	256	32 km
120	240	480	600	960	120 km

Maximum distances shown in Table 4 assume 100% utilization of the ISL. If the ISL is not fully utilized, greater distances can be achieved since more BB_Credits become available. For example, for a 2 Gb/s switch port with 120 BB_Credits and with an ISL that is only 50% utilized, the maximum distance is 240 km.

The BB_Credit mechanism allows the transmitting node (Port A) to keep sending Fibre Channel frames before receiving an R_RDY from Port B (as illustrated in Figure 22). This creates a streaming effect, and will not throttle down frame transmissions from Port A.

Note: Link distances depend not only on available BB_Credits, but also on the type of fiber used. See Appendix A, "Cable Specifications," to determine maximum direct connect distances supported by FC cable types.
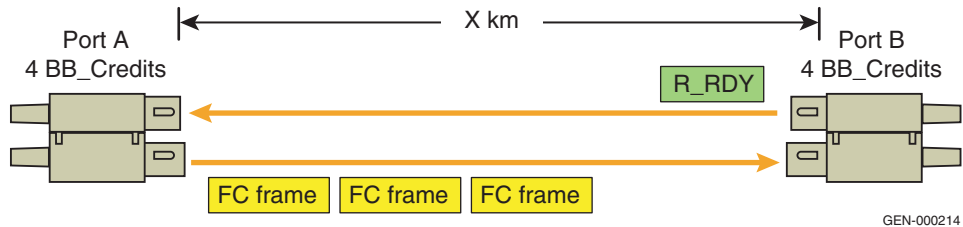
Figure 22   BB_Credit allowing multiple frame transmission before R_RDY

The amount of buffers directly correlates to the achievable distance and throughput. Buffering in a distance extension environment can be handled at any one of the following points:

◆ Fibre Channel switch

◆ Symmetrix DMX FA/RF port (when no Fibre Channel switch is present) has 7 BB_Credits

◆ Symmetrix VMAX FA/RF port (when no Fibre Channel switch is present) has 20 BB_Credits

◆ Distance Extension device

In the example shown in Figure 23 on page 66, there could be up to five frames on the link at one time: one (1) being processed and five (5) credits being returned. Therefore, at 2 Gb/s, approximately 10 credits are required to maintain full bandwidth.

A good rule to follow is:

At 2 Gb/s with 2 KB payload, approximately 1 credit per km is needed.
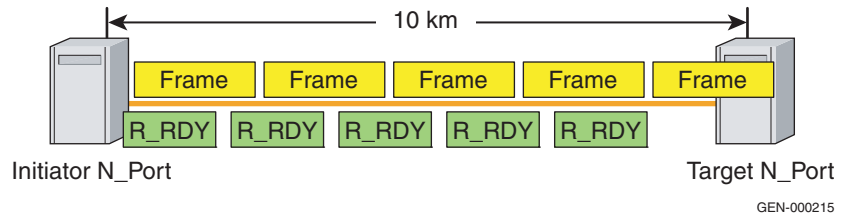
Figure 23 **2 Gb/s on 10 km link example**

In Figure 24, BB_Credits and flow control are managed by Symmetrix FA ports. In this scenario, the distance extension devices are not offering extended buffers and are "transparent."
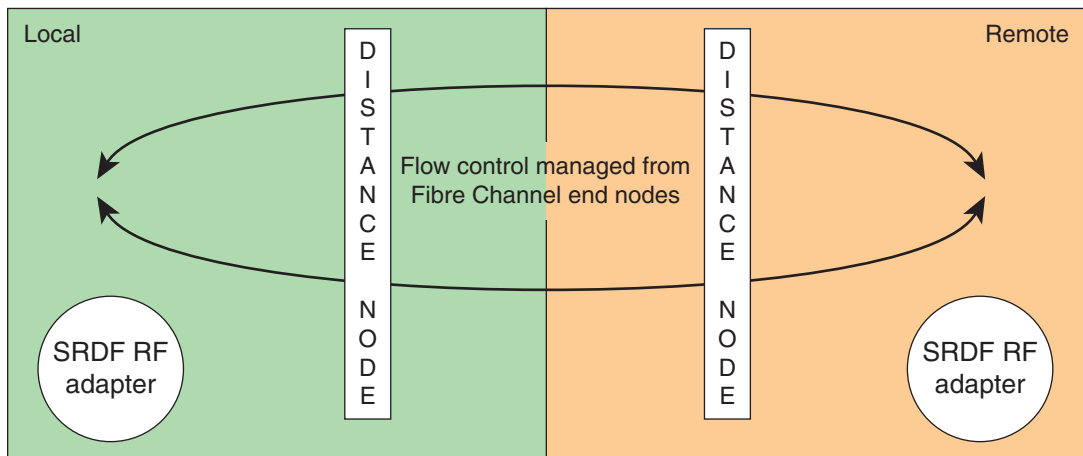
Figure 24 **Flow control managed by Symmetrix RF adapters (without buffering from distance extension devices)**

In Figure 25, BB_Credits are provided by the Fibre Channel switches. The distance extension device is transparent and does not participate in BB_Credit flow control.

Link speed, latency, and the amount of available credits will determine the performance characteristics of these configurations.

Note: Refer to the *EMC Networked Storage Topology Guide* available at http://elabnavigator.EMC.com, or the vendor switch documentation for details around BB_Credit support and flow control implementation.
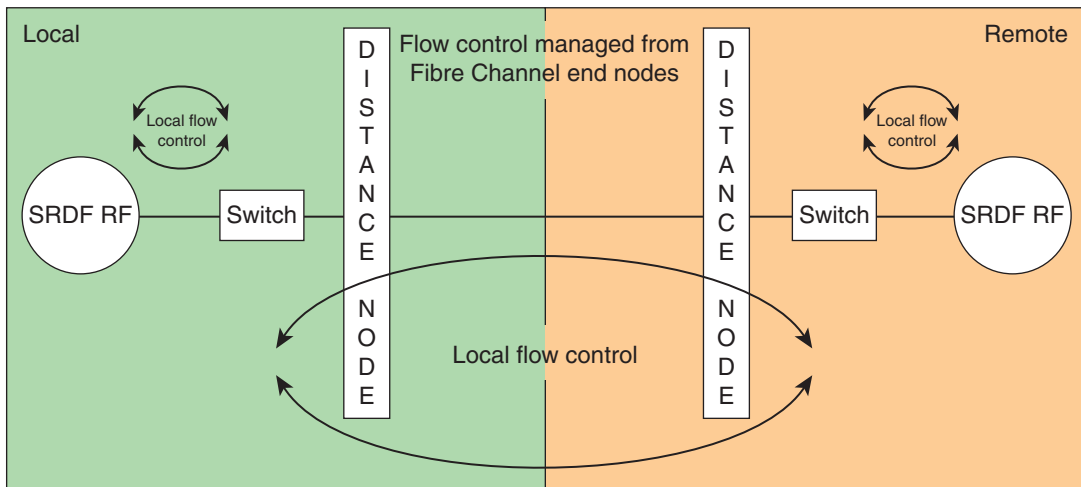
Figure 25   **Flow control managed by Fibre Channel switch (without buffering from distance extension devices)**

Determining sufficient amount of BB_Credits is crucial when provisioning Fibre Channel environments prior to utilization. Miscalculating the amount of credits may lead to performance degradation due to credit starvation.

Note: EMC recommends adding 20% margin to calculated BB_Credit values to account for spikes in traffic.

Credit starvation occurs when the number of available credits reaches zero preventing all forms of Fibre Channel transmissions from occurring. Once this condition is reached a timeout value will be triggered causing the link to reinitialize. To avoid this condition, sufficient BB_Credits must be available to meet the latency and performance requirements for the particular SRDF deployment.

The standard Fibre Channel flow control and BB_Credit mechanism is adequate for most short-haul deployments. With longer distance deployments however, the Fibre Channel flow control model is not as effective. Additional buffering and WAN-optimized flow control are often needed.

Figure 26 shows a configuration where the distance extension devices are providing additional buffering and flow control mechanisms for the purpose of increasing distances between locations. To accomplish this, the Fibre Channel end nodes are provided with immediate R_RDY responses with every "sent" FC-frame. This occurs within the local flow control segments. The distance extension nodes, in turn, implement their own buffering and WAN-optimized flow control.
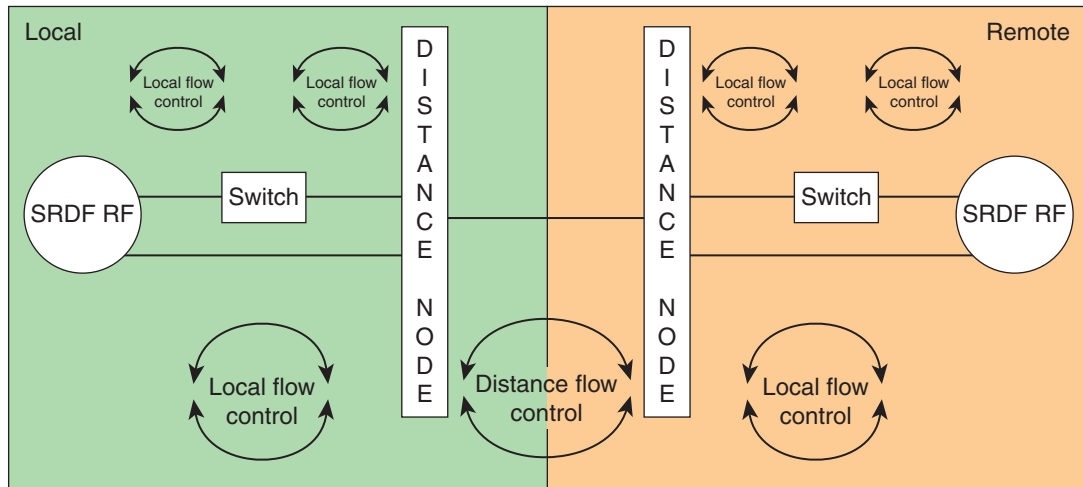
Figure 26   Flow control (with buffering from distance extension devices)

Refer to the distance extension vendor documentation for detailed information on each vendor's buffering and flow control implementations.

Fibre Channel link initialization (FC timeouts)

For link initialization of a Fibre Channel port, Fibre Channel specifications state that the maximum tolerable response time for a response is 100 milliseconds round trip time. This timeframe coincides with the limited timeframe of the Receiver-Transmitter Timeout Value (R_T_TOV), which is how long an FC-port listens for a link response to a link service before an error is detected.

For all Fibre Channel-only ISL configurations, R_T_TOV is a mandatory limitation.

For transparent distance extension environments the Fibre Channel end nodes (SRDF RF adapter or Fibre Channel switch) will be responsible for link initialization. The distance separating local and remote Fibre Channel switches must not exceed the 100 milliseconds

R_T_TOV time-out value or the Fibre Channel link between the Fibre Channel switches will be disabled.

Although the specification around R_T_TOV is stringent within the Fibre Channel specification, this does not limit SRDF to latencies up to 100 milliseconds round trip. For greater latencies to be possible, there are proprietary implementations via distance extension devices that enable link initialization to be handled locally between the local Fibre Channel port and the attached distance extension client port. The distance extension client port will imitate the responses of the remote Fibre Channel port within the 100 milliseconds window to the local Fibre Channel port. (This type of configuration was shown in Figure 26 on page 68.)

As discussed previously, the distance extension nodes can implement their own WAN-optimized flow control and buffering. These distance extension environments are considered non-transparent and require customized configuration of the distance extension devices and any Fibre Channel switches handling local initialization between the SRDF RF adapter and/or Fibre Channel switch.

SRDF with SiRT SiRT (single round trip) for Fibre Channel SRDF directors (RFs) was introduced in Enginuity version 5772 for SRDF/S mode only. It is dynamically enabled for SRDF/S links > 12 Km for block sizes up to 32K on DMX and 64K on VMAX. SiRT is compatible with fast write/write acceleration switches and extenders, as it will measure link latency and disable automatically if connected to these devices. As a best practice, it is recommended that either the EMC SiRT feature or the third-party fast write feature should be used. Both should not be enabled simultaneously.

The Fibre Channel SiRT feature for the Fibre Channel director can be set to **Off** or **Automatic**. When set to Automatic, this feature will only accelerate write I/Os using criteria based on latency and I/O size.

Note: SiRT defaults to Automatic (Enabled) in VMAX for SRDF/S. It automatically disables for Adaptive Copy and SRDF/A.

Figure 27 shows the normal write process without the SiRT feature.
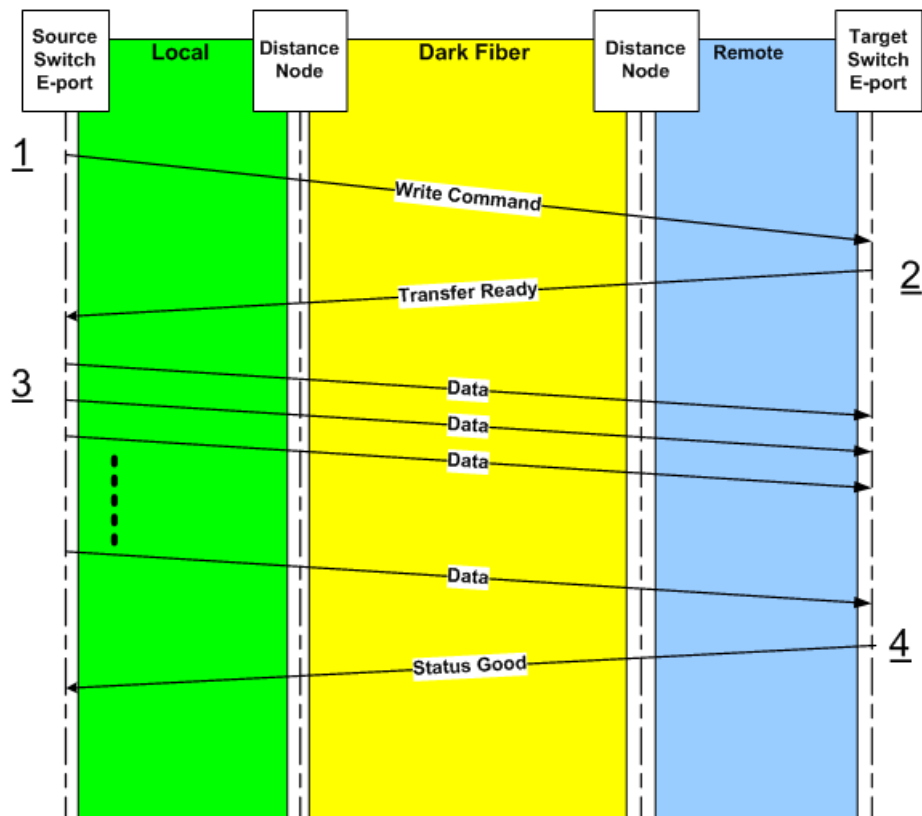
Figure 27   Normal write command process

The intended purpose of this feature is to maintain SRDF/S synchronicity while improving performance by localizing the transfer-ready response to the local RF port, thereby reducing an unnecessary acknowledgement response (trip) over the dark fiber distance (step 2 in Figure 27). Immediate benefits are apparent upon activation in transparent SRDF synchronous distance extension environments. If applicable, multiple SRDF synchronous links can maximize their I/O performance over the network (transparent WDM environment).

In the example shown in Figure 28, RF1 (R1 F_Port) and RF2 (R2 F_Port) are managing the SiRT flow control.
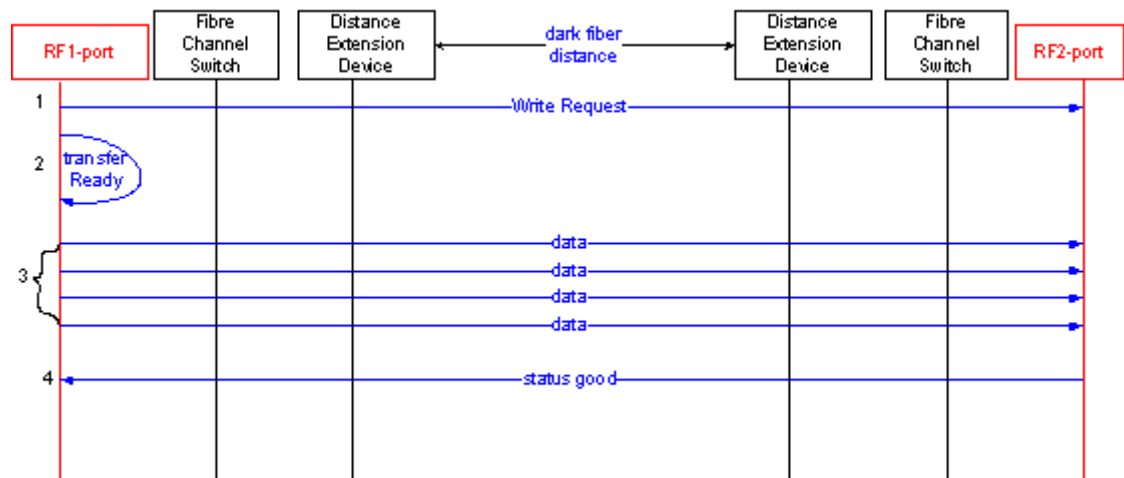
Figure 28   **SRDF SiRT**

Legend:

Red	RF-ports with SiRT activated.
Blue	A step-by-step of a single write command with SiRT enabled.

Fast write/ write acceleration   EMC Connectrix and other third-party products offer single round trip for Fibre Channel capabilities (fast write/write acceleration) that can also increase SRDF throughput for direct-attach or Fibre Channel switched fabric configurations over extended distances. It is transparent to SRDF FC links and is used for all SRDF modes to decrease response time (SRDF/S) or improve performance over long distance links (mostly for adaptive copy and SRDF/AR, but also for some SRDF/A configurations).

Figure 29 shows a write command with fast write features.
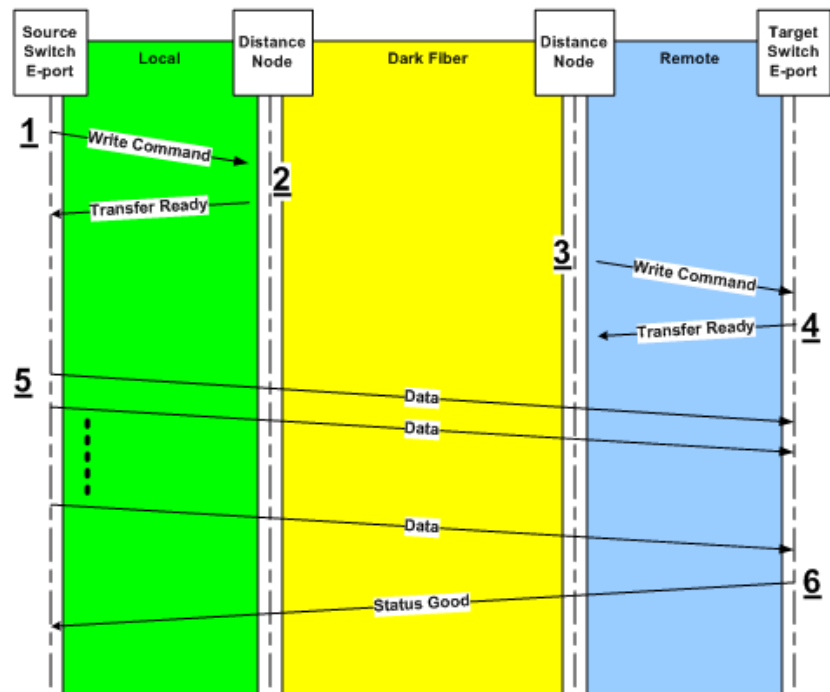
Figure 29   **Write command with SiRT**

For Connectrix or third-party products, refer to the *EMC Support Matrix* and the *EMC Networked Storage Topology Guide* available at http://elabnavigator.EMC.com to verify which of these products are supported for SRDF configurations.

⚠ **IMPORTANT**

Not all products offering this feature are supported with SRDF due to unique write commands utilized by SRDF.

SiRT with distance vendor write acceleration
With this in mind SiRT usage, in combination with the distance extension device-offered write acceleration mode, must be addressed. Essentially for environments where the distance extension device is

already servicing write commands on an E_Port level, it is recommended to disable SiRT. Refer to Figure 30.
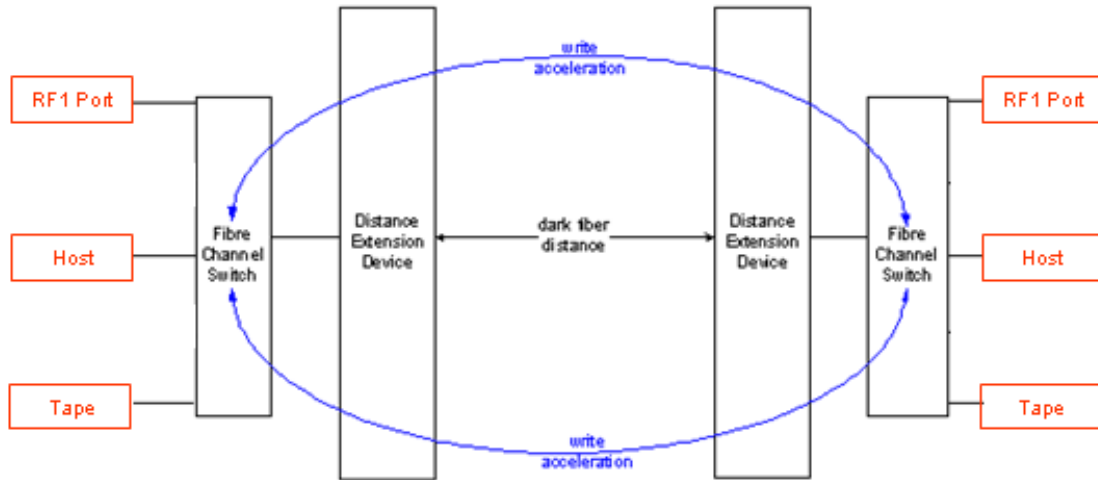
Figure 30   **All F_Ports will benefit**

Legend:

Red	RF ports benefiting from distance extension device, write acceleration.
Blue	Scope.

In Figure 30, by enabling the write acceleration feature on the distance extension device, potentially all F_Ports (RF ports, FA ports, tape, etc.) issuing writes traversing across the E_Port attached to the distance extension client port can also take advantage of the throughput benefits from the activated write acceleration feature.

Compression   Data compression is the process of encoding information using fewer bits than an unencoded representation would use through the use of specific encoding schemes. This is accomplished through algorithms that look for data having statistical redundancy. Compression programs simply eliminate the redundancy. Instead of repeating data over and over again, a compression program will store the data once and then refer back to it whenever it appears in the original data stream.

Note: Hardware and software compression is available for GbE SRDF only. Fibre Channel SRDF can use software compression only. Software compression must be enabled since it is disabled by default.

Compression for Fibre Channel exists and is supported by EMC with various distance extension and switch products. Keep in mind the following characteristics:

◆ Software compression is typically a lower cost option suited more for lower bandwidth environments (for example, OC-3).

◆ Efficiency of the compression is dependent on the algorithms used and the compressibility of the data (random versus sequential).

◆ Compression is not recommended where the available bandwidth meets or exceeds the throughput requirements.

Encryption   Encryption devices can be used to ensure the secrecy of data transmitted between two points. Encryption refers to algorithmic schemes that encode plain text into nonreadable form or cyphertext. Currently, there are no supported solutions that encrypt native Fibre Channel data inflight. Encryption must therefore occur on the WAN using either IP or SONET/SDH encryption devices.

Forward error control   Forward error correction has been widely deployed at the physical-link layer to check for and correct bit errors on WAN links. Even when a network's physical link is free from bit errors, packets may still get dropped in transit on WAN links because of congestion.

Packet-level forward error control (FEC) is an error control technique available with some distance extension devices where additional packets are inserted within the data stream that allow for reconstruction of data (if corrupted) without the need for retransmission. Retransmissions, especially with synchronous SRDF, can be extremely costly and should be avoided whenever possible.

Fabric isolation   Fabric isolation is a method to bound fabric definition. Boundaries can be defined functionally or geographically. The goal is to define the fabrics in such a way that events (such as network problems, administrative configurations changes, or hardware issues) occurring in one fabric will not impact other fabrics. The challenge with isolated fabrics is how to share resources amongst the fabrics.

In distance extension networks, isolating traffic locally while minimizing traffic across the WAN leads to SRDF deployments which are higher performing and less problematic.

Fabric isolation can be achieved in one of the following ways, each discussed in this section:

Dedicated fabrics

Isolation between fabrics can be achieved by using dedicated switches to form the link across the WAN. In a typical scenario each site (local and remote) will have its operational fabric where hosts access storage. A third fabric is created for traffic across the WAN. The WAN fabric will not be connected to the local and remote operational fabrics.

Disadvantages of this method include the need for additional hardware and the inability to share resources between the local and remote sites.

VSAN

Virtual SAN (VSAN) is used to allocate ports across different physical switches to create a virtual fabric. It acts as a self-contained fabric despite sharing hardware resources on a switch or multiple switches. Ports within a switch can be partitioned into multiple VSANs. In addition, multiple switches can join available ports to form a single VSAN. VSAN contains dedicated fabric services designed for enhanced scalability, resilience, and independence. This is especially useful in segregating service operations and failover events between high availability resources.

Each VSAN contains its own complement of hardware-enforced zones, dedicated fabric services, and management capabilities, just as if the VSAN were configured as a separate physical fabric. Therefore, VSANs are designed to allow more efficient SAN utilization and flexibility.

VSAN provides the same advantages that a dedicated fabric provides, however, it eliminates the need for additional hardware. Resource sharing between VSANs can be accomplished by using the Inter-VSAN Routing (IVR) feature.

Figure 31 shows an example of VSAN implementation. VSAN:A contains remote traffic. VSAN:B and VSAN:C contains local SAN traffic isolated to individual sites.
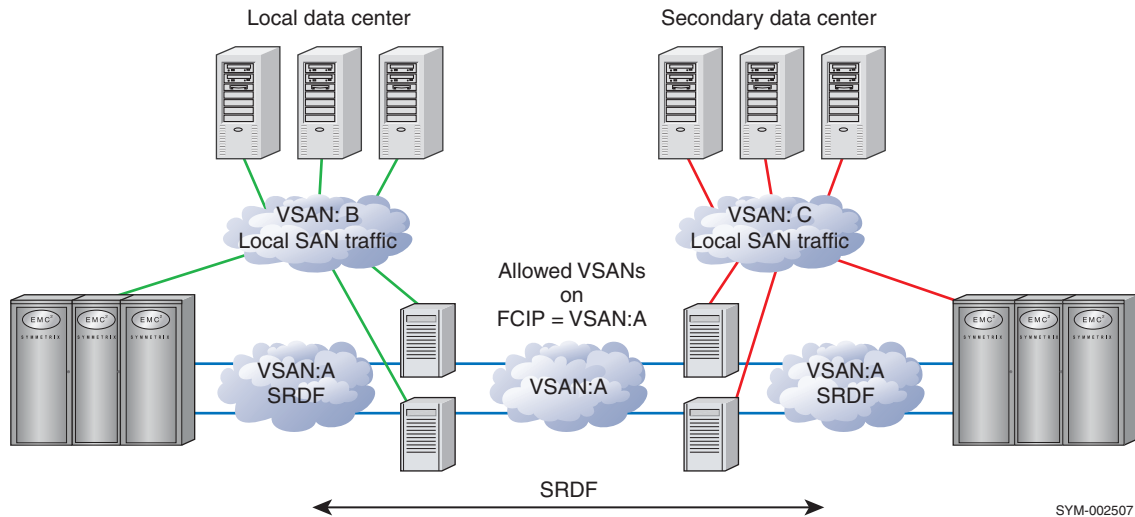
Figure 31   VSAN example

Note: Refer to the *EMC Networked Storage Topology Guide*, available at http://elabnavigator.EMC.com for more information on VSAN solutions.

FC routing

FC routing is a method for delivering efficient resource consolidation and utilization across separate fabrics while maintaining the management, isolation, security, and stability of dedicated fabrics. FC routing connects two or more physically separate fabrics to enable shared access to storage resources from any fabric.

Only the shared resources will be presented (routed) on each side of the WAN. All events occurring in one fabric will impact only the routed devices. This capability allows customers to scale and share resources across their fabrics while retaining the administration and fault isolation benefits of dedicated, separately managed, fabrics.

FC routing can be also be used inside the data center with similar benefits. For more information on specific FC routing solutions, refer to the *EMC Networked Storage Topology Guide*, available at http://elabnavigator.EMC.com.

Fibre Channel director considerations

This section provides information to consider.

SRDF static flow control

In addition to the Fibre Channel mechanisms and restrictions, SRDF static (legacy) flow control can be governed by the *SRDF Flow Control Algorithm* which allocates the number of I/Os that can be sent from the SRDF layer. This feature, introduced for SRDF over Fibre Channel in Enginuity version 5671, should be enabled for all extended distance asynchronous links (SRDF/A, SRDF/AR, and Adaptive Copy).It should *not* be used for SRDF/S and is disabled by default.

The SRDF algorithm first tracks, analyzes, and monitors for certain I/O conditions. It then determines the optimal number of I/Os (jobs) to send out on the SRDF links, based on certain RTT windows: 0-40ms; 40-120ms; and over 120ms. The algorithm maintains a manageable number of I/Os (job queues) while making efforts to avoid time-out scenarios. The Fibre Channel flow control is managed from the primary (sending) side of the SRDF Fibre Channel (RF) links, based on "activity" of each of the secondary ports, and monitoring for primary-side "resource shortage" reported by the Fibre Channel link hardware.

Network quality

The network quality of an SRDF link is now tested with Enginuity version 5875 ucode and later. VMAX systems on both sides of the link must be running 5875 or later to take advantage of this feature. This feature takes the following actions:

- Qualify a questionable SRDF connection before allowing traffic to be sent

- Gradually slows down intervals between unsuccessful SRDF link re-establishment attempts

- Puts bouncing SRDF links out of service

A qualification test for any SRDF connection can be manually triggered. This feature is supported for both GbE and Fibre Channel SRDF links.

Compression

Beginning with Enginuity version 5874, software compression is available with the Fibre Channel port. Software compression can be enabled at the SRDF group level and SRDF mode of operation. Both sides of the link must support software compression. Hardware compression is also supported for Fibre Channel on VMAX 40K only and can be enabled at the SRDF group level.

⚠ **IMPORTANT**

Software and hardware compression should be disabled when Write Acceleration is enabled on an external Fibre Channel switch.

Note: Higher host response times can result due to the software compression. If host response time is not a concern, bandwidth savings can be realized.

FCIP and iFCP

FCIP FCIP is a tunneling protocol that transports Fibre Channel ISL traffic. FCIP uses IP as the transport protocol. An FCIP link tunnels all SRDF traffic between locations and may have one or more TCP connections between a pair of IP nodes for the tunnel end-points. From the Fibre Channel fabric view, an FCIP link is an ISL transporting all Fibre Channel control and data frames between switches, with the IP network and protocols transparent. One can configure one or more ISLs between Fibre Channel switches using FCIP links.

Proper network design is required given the speed and bandwidth differences between Fibre Channel and a typical IP network. This design should take into consideration the management of congestion and overload conditions. Fibre Channel fabrics connected via FCIP links are not isolated unless isolation is achieved using one of the fabric isolation methods mention earlier.

iFCP iFCP is a gateway-to-gateway protocol to provide communication over IP between two Fibre Channel end-devices residing in two different fabrics. An iFCP session is created between a pair of iFCP supporting devices for each pair of Fibre Channel devices. An iFCP gateway transports Fibre Channel device-to-device frames over TCP. An iFCP session uses a TCP connection for transport and IPsec for security, and manages Fibre Channel frame transport, data integrity, address translation, and session management for a pair of Fibre Channel devices. Since an iFCP gateway handles the communications between a pair of Fibre Channel devices, it only transports device-to-device frames over the session, and hence, the Fibre Channel fabrics across the session are fully isolated and independent.

While FCIP by itself does not provide fabric isolation, iFCP is a Fibre Channel routing protocol and benefits from all the advantages of fabric isolation and FC routing.

Note: Consult the *Distance Extension Solution* table of the *EMC Support Matrix* and the *EMC Networked Storage Topology Guide*, available at http://elabnavigator.EMC.com, for more details on supported FCIP and iFCP devices.

FC-SONET

This section provides information on FC-SONET and the Generic Framing Procedure (GFP).

Generic Framing Procedure

The Generic Framing Procedure (GFP) was originally developed for the purpose of encapsulating IP packets or Ethernet MAC frames into SONET. GFP is an efficient protocol that simplifies the mapping of user data into SONET frames.

GFP's importance to SRDF is that it provides simple encapsulation for native storage protocols while promising interoperability between any vendor's GFP-compatible equipment.

GFP-T is a specific implementation of the protocol that "transparently" encapsulates any 8B/10B encoded frame into SONET packets for transport across the WAN. GFP-T can be used for transport of a wide range of protocols including ESCON, FICON, Fibre Channel, Gigabit Ethernet, HDTV and InfiniBand.

Using GFP-T, 8B/10B characters are decoded and re-encoded into a 64B/65B word. Eight of these words are then grouped together into an octet along with control and error flags. Eight octets are then assembled into a "superblock," which is scrambled and subject to cyclic redundancy checks. The resulting packet is fully SONET compatible.

Note: Refer to the *Distance Extension Solution* table of the *EMC Support Matrix* and the *EMC Networked Storage Topology Guide*, available at http://elabnavigator.EMC.com for more details on supported FC-SONET devices.

ISL load balancing of SRDF traffic

The FC-SW standard does not provide for load balancing when there are multiple paths at the same cost, leaving it up to the switch vendor to establish algorithms to balance the load across multiple ISLs.

The value ISL load balancing brings to SRDF traffic is the ability to optimize storage network performance in congested fabrics. Properly sized, Service Level Agreements should be met without SRDF, Subscriber (server to storage), or other Business Continuance traffic monopolizing the bandwidth of the ISL trunk.

In most storage networks, there will be peaks of congestion, with some paths running at their limit while others are barely used. SRDF and other traffic patterns should be analyzed to better determine if the intelligent balancing of data flows would be beneficial in reducing bottlenecks at these peaks. You can analyze traffic patterns and examine bandwidth reporting via EMC Work Load Analyzer, Ionix™ ControlCenter®, Connectrix Manager, or other such utilities. One indicator to implement ISL load balancing would be observing repeated occurrences of an 80+% utilization of an ISL with little utilization of an adjacent ISL of equal cost.

In some cases, parallel ISL links may suffice for sustained performance and availability (Figure 32). This is particularly true when ISL traffic is limited to SRDF with little to no subscriber (e.g., server to storage) traffic.

SYM-002505
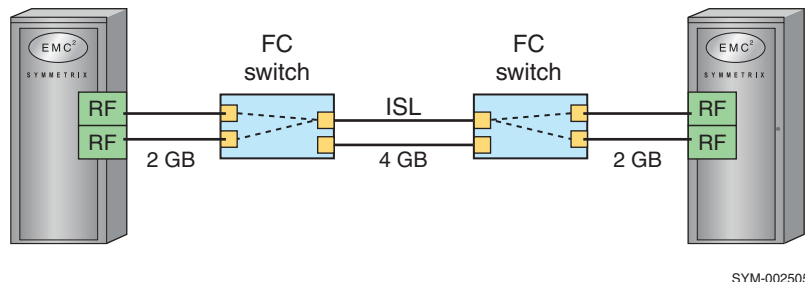
Figure 32   Parallel ISL links, employing ISL load sharing example

In the example shown in Figure 32, the two 4 Gb ISLs, if properly configured with the necessary BB_Credits, are capable of handling the peak bandwidth requirements of the 2 Gb RF ports. This configuration does not employ ISL load balancing. It employs ISL load sharing. In high traffic parallel ISL configurations, switches utilize the FC-SW-2 standard, Fabric Shortest Path First (FSPF), in a round-robin method to share load amongst ISLs of equal cost. This is a static connection defined when the fabric is built. Adding more parallel ISL connections may reduce congestion, but this eventually leads to a more costly, inflexible method of load sharing. ISL load balancing provides a more efficient solution for an oversubscribed ISL.

Load balancing should also be implemented if any ISL is observed to be over-utilized while an adjacent ISL is minimally utilized. Implementing load balancing under either of these conditions will allow for a more intelligently balanced data path.

Figure 33 on page 82 provides an example. In this case, the effective bandwidth is dynamically shared amongst both SRDF and other traffic. Ideally, the number of hops between RF and server N_Ports should be kept to one, and the fabric should be logically segregated between subscriber traffic (e.g., server to storage) and Business Continuance traffic (e.g., SRDF and Backup) whenever practical.

Key:

= Load Balancing

= Subscriber traffic
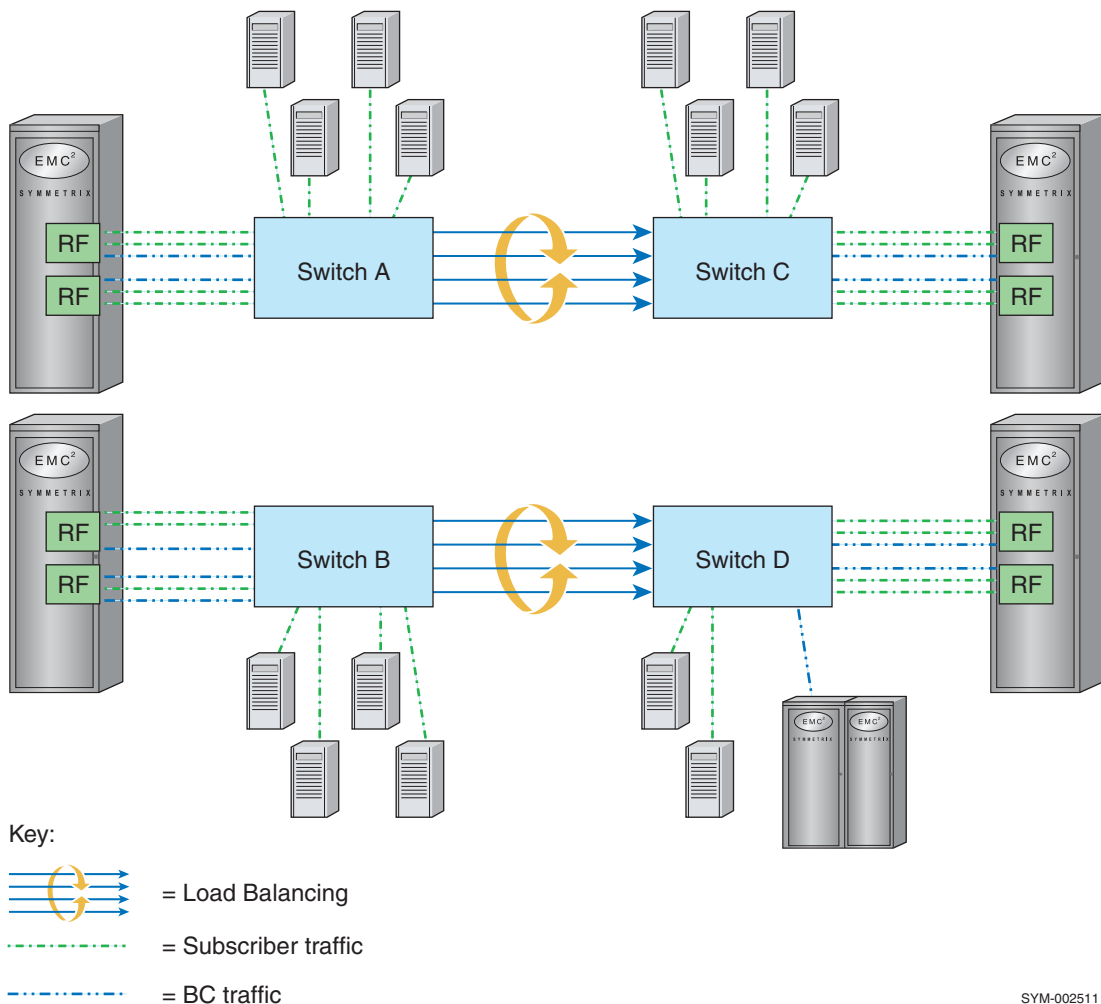
= BC traffic

SYM-002511

Figure 33   **ISL traffic example**

Two types of load balancing are common in SANs:

◆ Frame-based

◆ Exchange based

Frame-based load balancing is not recommended because it is sensitive to cable lengths, which in turn can cause out-of-order frame delivery.

Since it is always a best practice to ensure enough ISLs to keep link utilization below 70%, it is recommended to always use exchange-based ISL load balancing for SRDF over distance to properly load balance and ensure no out-of-order frame delivery. Exchange-based load balancing is based on the SID, DID, and Fibre Channel originator exchange ID (OXID). Traffic is load balanced at the Fibre Channel exchange level where every exchange can take a different path through the fabric regardless of differences in cable length or latency. While SRDF can handle out-of-order frames and CRC errors via abort and retry algorithms, there can be substantial performance penalty in doing so. Monitoring Symmetrix error logs for high occurrences of BF2D, xx1E, and AB3E events is one method of determining if out-of-order and discarded frames are impacting performance. These events indicate that there are issues external to the Symmetrix.

Performance and distance considerations

Until recently, a delineating factor in EMC environments was the distance limitation of 200 km. Today, however, that has changed. For WDM, SONET, and SDH networks, the defining factor for distance supported in all cases depends on the following:

◆ Adherence to the R_T_TOV timer (defined as the latency between any two Fibre Channel N_Ports; cannot exceed 100 ms round trip)

Note: Third party distance extension devices may provide the ability for local Fibre Channel link initialization whereby the 100 ms (R_T_TOV) timer limitation may no longer be a limitation. For more details, refer to "Fibre Channel link initialization (FC timeouts)" on page 68.

◆ The specific tolerance of the customer application to increased latency

◆ Any additional buffering provided by SAN or distance equipment

◆ The distance the network equipment provider is capable of supporting

ESCON

SRDF over ESCON uses a Symmetrix RA director, which is basically an EA director configured to run SRDF emulation code.

RA directors employ an ESCON protocol schema, with one RA (RA1) in each connection assuming a role similar to a mainframe channel (host), and the other (RA2) assuming the role of a mainframe control unit (device). For this reason, significant protocol overhead is incurred in cases where a primary (R1) volume is configured to an RA2 director (bidirectional).

Connections in this environment may be through direct multimode cable connections between RA directors in Symmetrix systems up to 3 km apart. (Refer to Figure 34 on page 85.) For distance limitations regarding cable types for each cable segment, refer to Table 5 on page 120.

Each ESCON connection requires one pair of fiber-optic strands (*xmit* and *rcv*). RA directors can be used for SRDF through an ESCON director; however, the ESCON director ports must be configured to be static and dedicated. IBM ESCON Manager software may be used to control the ESCON director ports, but will present invalid status upon query operations to SRDF director ports.

Because ESCON directors provide their own transmit signals at each port, the effective distance of the overall 3 km link can be doubled. By using ESCON extender/repeater units or the ESCON director XDF feature, these distances can be extended up to 66 km for campus-mode SRDF. Distance extension devices, such as DWDM and ESCON over IP devices, can also be used to extend SRDF over ESCON links to hundreds, and even thousands, of kms.

Refer to the *EMC Support Matrix* and the *EMC Networked Storage Topology Guide*, available at http://elabnavigator.EMC.com, for more details on supported devices.

Note: For SRDF distances beyond 15 km, or connections via channel extenders, FarPoint™ mode should be used to eliminate distance limitations and minimize performance penalties.

Unidirectional Configuration

Campus/extended distance solution
- Writes flow from R1 on an RA-1 (*channel* protocol) to R2 on an RA-2 (*controller* protocol).

Dual-directional Configuration

Extended distance solution
- Two unidirectional SRDF RA groups (primary and secondary in each Symmetrix system).

Bidirectional Configuration

Campus solution
- Each Symmetrix system has both primary and secondary volumes.
- RA links transmit data in both directions.
- FarPoint is not supported.
- Performance degradation occurs in RA-2 to RA-1 direction.

SYM-002504

Figure 34   SRDF ESCON topologies (Point-to-point)

FarPoint

FarPoint is an SRDF feature used *only* with ESCON extended distance solutions (and certain ESCON campus solutions) to optimize the performance of the SRDF links. This feature works by allowing each RA to transmit multiple I/Os, in series, over each SRDF link in pipeline fashion.

Standard SRDF without FarPoint

The standard ESCON SRDF protocol (without SRDF FarPoint) allows only one write I/O to occupy a communication link at a given time. The next write I/O occurs only after the remote Symmetrix system has acknowledged receipt of the previous I/O from the local Symmetrix system and an ending status is presented to the local Symmetrix system as shown in the top half of Figure 35 on page 87.

SRDF with FarPoint

With SRDF FarPoint, a single remote adapter (R1) can serially transmit more than one write I/O over the SRDF links as shown in the bottom half of Figure 35. This method (referred to as *pipelining*) enables the SRDF communication links to be more fully utilized, depending on the distribution of application I/O write activity across multiple logical volumes. The number of write I/Os that can be placed on the link at one time is determined by the capacity of the type of SRDF link being used. With Enginuity version 5x67 or later, multiple copy tasks can be placed on the SRDF link for a single logical volume (up to eight per volume).

Symmetrix   SRDF without FarPoint   Symmetrix

Symmetrix   SRDF with FarPoint   Symmetrix
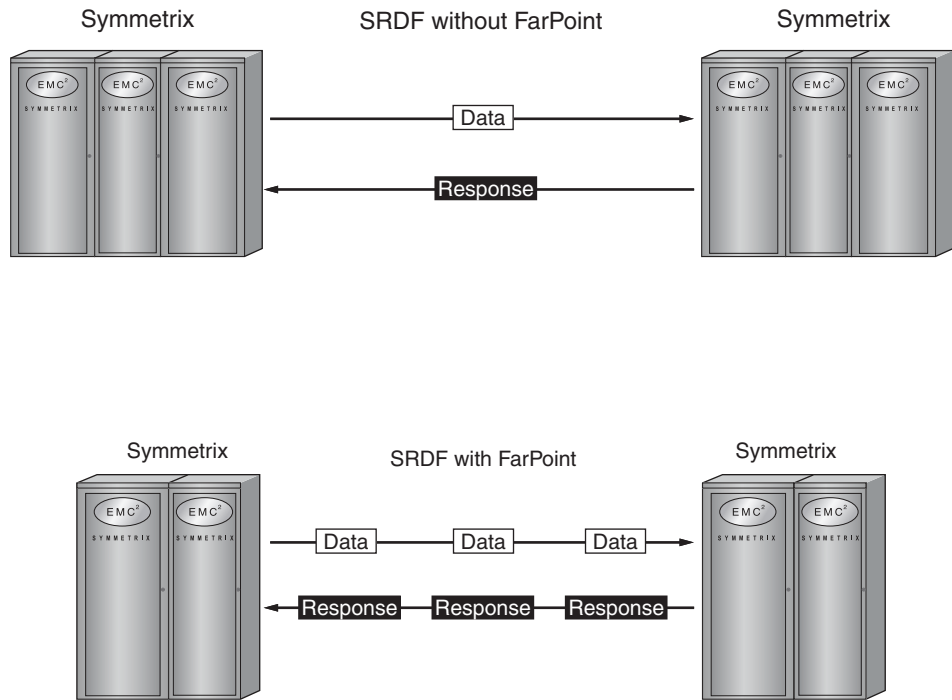
SYM-002515

Figure 35   SRDF with and without FarPoint

Preserving synchronization

From the point of view of the host, FarPoint does not change the SRDF protocol. In synchronous mode, the Symmetrix system returns a completion status to the host only after the write operation is completed to cache on the remote (R2-side) Symmetrix.

Without FarPoint, the Symmetrix system waits for one write operation to complete before sending the next write to the same device (volume). With FarPoint, the Symmetrix system, while waiting for the status of the first write operation, uses the free link bandwidth to send another write operation from a different primary volume.

Interaction with the host remains unchanged, so the synchronous condition is fully preserved, and the data on the remote SRDF device is 100 percent consistent from the host's point of view.

Servicing remote reads

Remote read operations (mainly used for recovery purposes), as well as other special I/Os, are serviced using the standard SRDF protocol (without FarPoint). In this situation, the write pipeline is cleared

before any read operations are performed. In other words, all FarPoint write I/Os in process on all the FarPoint-enabled RA links are completed, all links drop out of FarPoint to campus mode, and then the read I/O is completed.

Since FarPoint is uni-directional (not peer-to-peer protocol like Fibre Channel), any *write* from the RA2 back to the RA1 (a restore operation) is actually a requested *read* for the RA1 director on the R1 side of the link.

This entails 2 ½ round trips for each individual read I/O and since only one I/O is on the link at a time, the performance over distance degrades rapidly.

You should always swap R1/R2 and RA1/RA2 personalities to change the direction of FarPoint I/O before doing a restore/ ComeHome operation.

Performance considerations

An SRDF configuration using FarPoint can improve performance substantially at the Symmetrix system level across a group of SRDF devices depending on the length of the SRDF links. Even on a short link, the FarPoint solution may provide some benefits due to the pipelining and one round trip per write.

However, write operations to a single device are still serialized by the host, so in the case of a synchronous device, a new write operation does not start until the previous write operation to that device receives a status from the remote system.

This means that the maximum number of write operations per device is still low, and using FarPoint in this situation does not improve system performance.

Note: The introduction of PAV/MA and striping does increase I/O concurrency at the volume level, potentially improving throughput in FarPoint configurations by having more active SRDF volumes.

Network Topologies and Implementations

This chapter discusses various physical and transport options available for deploying SRDF networks using various metropolitan and wide-area topologies, and some best practices for deploying SRDF in these environment.

Transport options

This section discusses the following transport options:

◆ "Dark fiber," next

◆ "IP" on page 95

Dark fiber

Dark fiber is very expensive and in some cases, depending on the distance separating the sites, may be impractical for use as a WAN. Dark fiber is better suited for use in a metropolitan area network (MAN).

In cases where the distance is not an issue and large amounts of bandwidth are needed, dark fiber may be the best choice. Synchronous SRDF applications requiring high throughput and low latency are excellent candidates for DWDM/CWDM over dark fiber. For instance, SRDF/S configurations within Metropolitan Area Networks or MANs) are commonly deployed using Fibre Channel directors or switches with DWDM over dark fiber fiber in a FC-switched Fabric deployment.

Generally, different Fibre Channel directors/switches will support varying ranges of bb-credit amounts. On a FC-port, BB-credit amounts can range from 1 - 6752 BB_Credits (depending upon specific Fibre Channel switch types, models, and/or Fibre Channel switch line cards or blades used), which is more than sufficient for SRDF/S distances. Refer the EMC Switch Topology guides and/or switch vendor documents for maximum FC-port BB-Credit amounts and BB-Credit calculations/allocations on a particular Fibre Channel switch, blade, online-card over MAN distances.

Following are the various options available for dark fiber deployments, each discussed further in this section:

◆ "Long-wave optics," next

◆ "CWDM" on page 91

◆ "DWDM" on page 91

◆ "SONET/SDH" on page 92

◆ "ATM" on page 94

◆ "T1/T3, E1/E3" on page 94

Long-wave optics   Long-wave optics operating at the 1310nm wavelength are available for use with Symmetrix and Fibre Channel switches. As of the time of this writing, EMC offers various long-wave optics for Fibre Channel switches supporting distances of 10, 20, and 35 km. For Symmetrix RF cards, EMC offers a single long-wave optic capable of supporting 10 km.

When any of these optics are directly connected to dark fiber, only a single I/O stream can be driven across the fiber. If the customer's throughput requirements are satisfied through the use of these long-wave optics, this can be a viable solution.

CWDM   Coarse Wave Division Multiplexing (CWDM) combines up to 16 wavelengths onto a single fiber. CWDM technology uses an ITU standard 20 nm spacing between the wavelengths, from 1310 nm to 1610 nm. With CWDM technology, since the wavelengths are relatively far apart (compared to DWDM), the lasers and transponders are generally inexpensive compared to DWDM.

Best practice is to use CWDM when:

◆   Customer's throughput requirements can be satisfied by the number of wavelengths available through CWDM.

◆   Cost of DWDM is prohibitive.

◆   Distance between locations is less than 100 km.

◆   Manageability requirements are met versus generally more robust management capabilities available with DWDM.

DWDM   Dense Wave Division Multiplexing (DWDM) combines up to 160 wavelengths onto a single fiber. DWDM technology uses an ITU standard 0.8 nm spacing between the wavelengths with a reference frequency fixed at 1552.52 nm. With DWDM technology, the lasers and transponders are fairly expensive when compared to CWDM. Since wavelength spacing is so small, added components and controls are needed to maintain the precision of lasers.

Best practice is to use DWDM when:

◆   Customer has high throughput requirements that cannot be satisfied by CWDM.

◆   Distance between locations is greater than that supported by CWDM. DWDM systems can support distances separated by hundred's of kms through the use of high powered lasers and amplifiers.

◆ Manageability, in general, is more robust with DWDM systems. Therefore, if the customer requires high levels of manageability, DWDM systems should be considered.

SONET/SDH   SONET (Synchronous Optical Network) and SDH (Synchronous Digital Hierarchy) are a set of related standards for synchronous data transmission over fiber optic networks. SONET/SDH is the core backbone of all major ISPs, providing low latency, high availability, zero loss, and high bandwidth, all of which make it ideally suited for replicating large amounts of data over a WAN. SONET/SDH is also found in MANs and can be a suitable transport for lower bandwidth storage applications or for those customers without access to dark fiber.

SONET is the United States version of the standard published by the American National Standards Institute (ANSI). SDH is the international version of the standard published by the International Telecommunications Union (ITU).

SONET/SDH is not a communications protocol in and of itself, but rather a generic and all-purpose transport container for moving both voice and data. SONET/SDH can be used directly to support either ATM or Packet over SONET (PoS) networking.

The basic SONET/SDH signal operates at 51.840 Mb/s and is designated STS-1 (Synchronous Transport Signal one). The STS-1 frame is the basic unit of transmission in SONET/SDH. Higher speed circuits are formed by aggregating multiple STS-1 circuits. When sizing a network that will be using a SONET/SDH transport mechanism, make sure you take into account the overhead associated with SONET/SDH (see).

Next generation SONET/SDH systems provide further enhancements through Generic Framing Procedure (GFP) and Virtual Concatenation (VCAT):

◆ GFP allows flexible multiplexing so that data from multiple clients sessions can be sent over the same link. GFP supports the transport of both packet-oriented (e.g., Ethernet, IP, etc.) and character-oriented (e.g., Fibre Channel) data. GFP is a framing procedure for mapping variable-length payloads into SONET/SDH synchronous-payload envelopes.

The value of GFP is it provides simple encapsulation for native data protocols (such as Ethernet or Fibre Channel) with extremely low overhead requirements (see Figure 36 on page 93). GFP mapping to SONET also promises interoperability by allowing any vendor's GFP-compatible equipment to pass and offload GFP payloads.
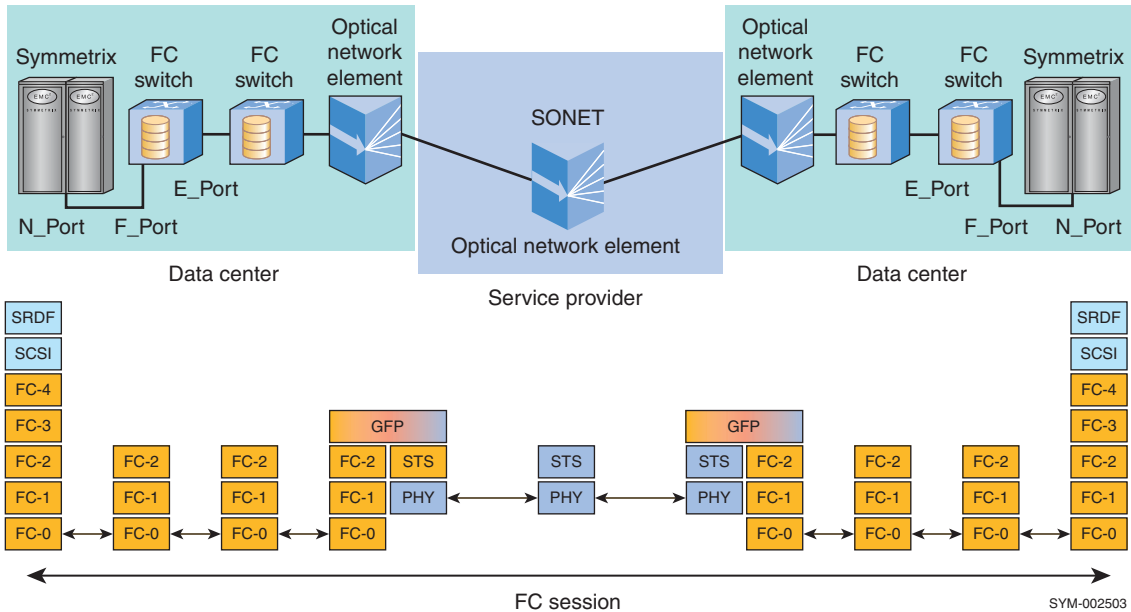
Figure 36   Fibre Channel storage extension over SONET/SDH

Note that the actual GFP mapping does not necessarily occur at the SONET/SDH device. If the Fibre Channel or Ethernet signal originates at a WDM device, these devices often contain cards that can wrap the native signal within GFP before converting the signal to a WDM wavelength. In this case, a SONET/SDH payload will already exist at handoff from WDM to the SONET/SDH network.

◆ Virtual Concatenation (VCAT) enables multiple low-capacity SONET/SDH circuits to be grouped together by inverse multiplexing into a higher capacity link (Figure 37 on page 94). For example, pre-VCAT, if a customer had an OC-3 and needed more bandwidth, an OC-12 had to be purchased. The OC-12 was

required even if the customer required just a bit more bandwidth than available with an OC-3. STS-3c-2v (equivalent to OC-6), or STS-3c-3v (equivalent to OC-9) are options available with VCAT.

VCAT groups a number (x) of virtual containers (STS-1/3c), using the combined payload (STS-n-Xv) to match the required bandwidth.
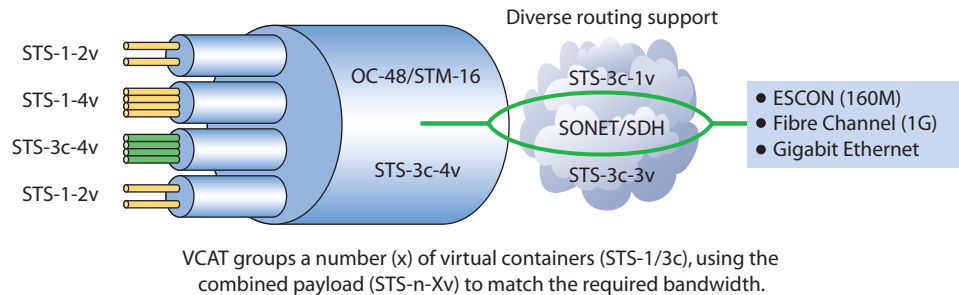
GEN-000140

Figure 37   VCAT

ATM   ATM (Asynchronous Transfer Mode) is a network protocol that transmits data in 53-byte cells. An ATM cell has a 5-byte header and 48-byte data payload. ATM is used as a protocol over SONET. Since ATM was designed, networks have become much faster, removing the need for small cells to reduce *jitter* (rapid variation in the latency on the network). Some believe this removes the need for ATM in the network backbone.

The hardware for implementing the service adaptation for IP packets is expensive at very high speeds. Specifically, the cost of segmentation and reassembly (SAR) hardware at OC-3 and above speeds makes ATM less competitive for IP than Packet over SONET.

T1/T3, E1/E3   The T designator is used to define a digitally multiplexed carrier line.

- ◆ A T1 is made up of 24 64 K-bit channels multiplexed together with an additional 8 K bits for framing to form a 1.544 Mb circuit (commonly referred to as a T1 line).

- ◆ A T3 line is basically a T1 on a larger scale. A T3 is made up of 672 64K-bit channels to make a 44.73 Mb circuit.

- ◆ E1/E3 circuits (E1 = 2.048 Mb; E3 = 34.396 Mb) are used worldwide with the exception of North America and Japan. E and T carriers, although very similar, are incompatible.

⚠️ **CAUTION**

T1s and E1s are not recommended for use with SRDF, except for ESCON FarPoint. It is possible to multiplex them together to get the minimum 10 Mb of bandwidth required for SRDF over Fibre Channel and GbE, but it is critical they be configured properly and are properly load-balanced so the packets do not arrive at the secondary side out of order.

IP

This section contains network considerations including, bandwidth, quality, and latency.

Network considerations

Many factors need to be considered when selecting the appropriate WAN connection. These include bandwidth, network quality, and latency, each discussed next.

Bandwidth

An important part of designing an SRDF storage network solution is allowing for the necessary amount of bandwidth to handle present requirements, as well as being able to scale to future needs.

SONET/SDH has the ability to dynamically scale bandwidth by using Link Capacity Adjustment Scheme (LCAS) to increase or decrease the bandwidth of virtual concatenated containers (VCAT).

For other types of bandwidth (such as T3 circuits), where more bandwidth is required, it is possible to have a burstable or bandwidth-on-demand circuit. Burstable circuits are more cost-effective in cases where the amount of provisioned bandwidth is only exceeded for a short period of time each day, week, or month.

Network quality

Latency should be as low as possible, given the distance separating the sites, with a minimum number of hops. SRDF has been tested against what is referred to as the *Distance Envelope*. This involves a series of tests to make sure that SRDF will continue to function as long as the latency does not exceed 200ms RTT and packet loss does not exceed 1%.

A quality link should have packet loss no greater than 0.01%. SRDF can tolerate packet loss up to 1%. If the packet loss is greater than 0.01% it is not considered a quality link and will typically result in

throughput that is severally impacted. Packets should not be delivered out of order since this can cause the SRDF link to bounce or result in a major drop in performance. The rapid variation in the latency on the network (*jitter*) should also be very low. It is recommended that all of these aspects be covered in an SLA (Service Level Agreement). An SLA is a contract from the ISP (Internet Service Provider) guaranteeing the quality of the link.

Latency

Latency is the time it takes (in milliseconds) for a full Ethernet packet (1500 bytes) to traverse the network, typically quoted in round trip time (RTT). The amount of latency is the time it takes for light to travel over the fiber miles (1 msec per 125 (200 km) one-way) as well as any additional delay added due to networking devices located between the primary and the secondary Symmetrix.

Latency is often used interchangeably with distance. While they are related, they do not refer to the same thing. As alluded to above, there is a certain amount of latency resulting from the time it takes for light to travel from one location to another. However, additional latency can result from other factors such as network components and fiber quality.

Metro and WAN implementations

This section discusses the following topologies that can be used to deploy SRDF:

Metro (local) implementation

Figure 38 shows an example where CWDM edge devices at the customer sites are using GFP to map multiple transparent Fibre Channel and GbE services into OC-48 (STS-48C) SONET payloads which are in turn carried as OC-48 lambdas (wavelengths) to the metro edge DWDM office. These signals are then transported as an OC-48 lambdas over DWDM to telco carrier point of presence (POP). From there, the OC-48 signals are demultiplexed from DWDM and sent to a SONET multiplexer and transported over extended distance via an OC-192 SONET network. At the final metro edge site, the reverse process occurs, taking the GFP mapped OC-48 signal back to the corresponding transparent service (FC, GbE, etc.).
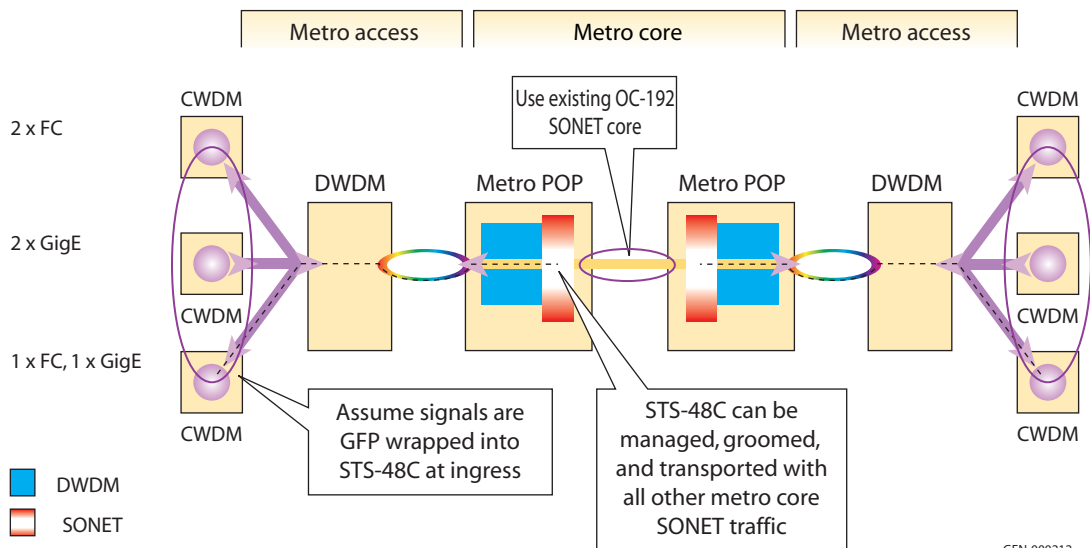
Figure 38   SRDF Metro configuration example

WAN extended distance implementation

Customers using SRDF in extended distance configurations have many connectivity options, including native GbE, Fibre Channel over IP and Fibre Channel over SONET. The most important consideration is the reliability and robustness of the WAN network. It is certainly the most costly component, but also the key to a successful and well-performing SRDF application. Bandwidth sizing, based on SRDF mode and workload, depends heavily on network quality and also on connectivity features such as compression, load-balancing, and traffic shaping/prioritization. Figure 39 on page 99 is an example of an extended distance SRDF configuration that utilizes a SONET network for Symmetrix GbE director ports. The same WAN configuration could also support FCIP and iFCP extension over dedicated SONET network, rather than a routed IP network.
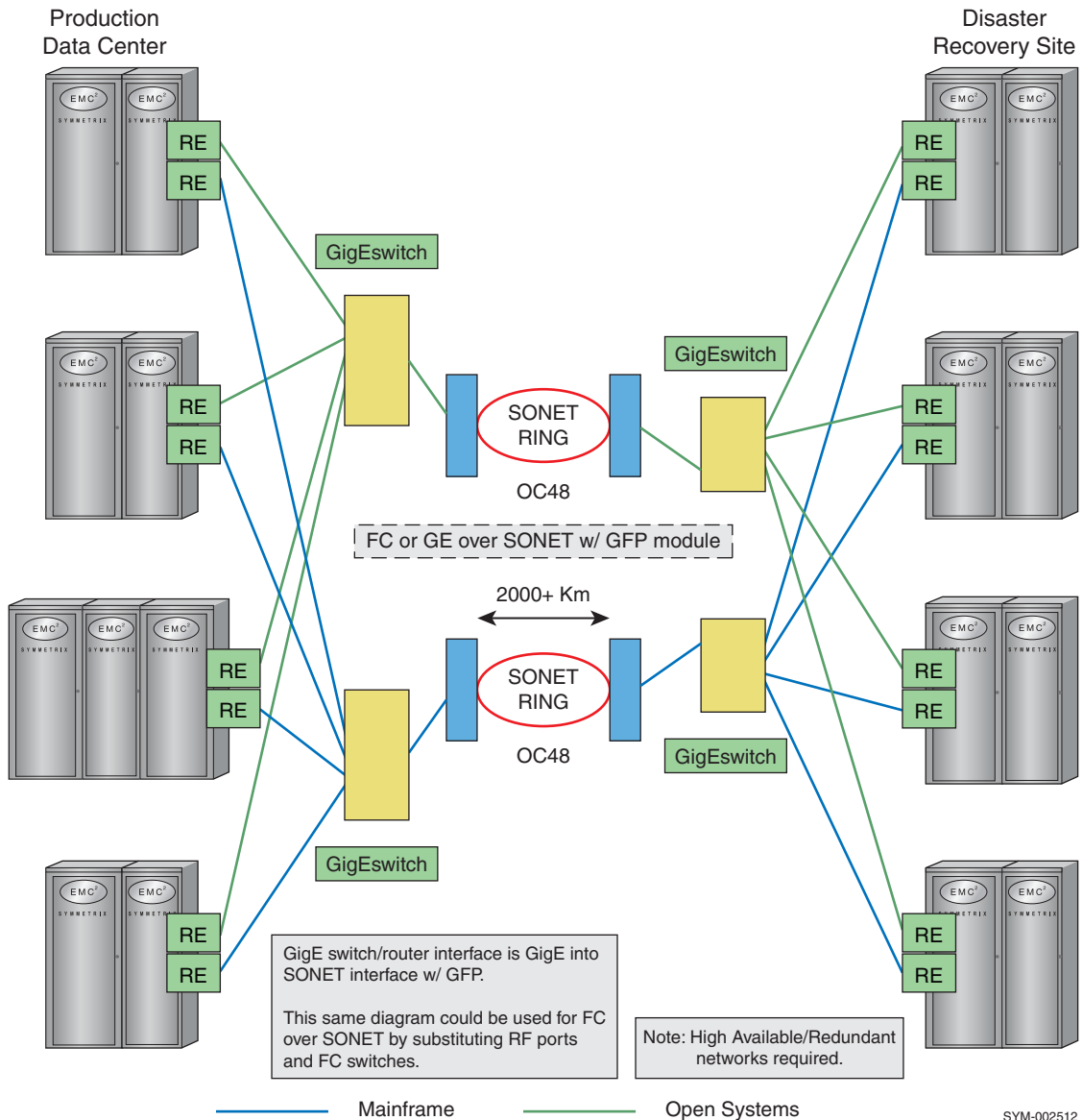
Figure 39   SRDF network configuration example

Combined Metro/WAN implementation

Many EMC customers require the highest levels of disaster recovery. A common architecture used to attain this level of disaster recovery is through the deployment of Concurrent SRDF via SRDF/S within a metropolitan area and SRDF/A over the WAN.

Figure 40 shows a sample deployment where SRDF/S over DWDM is used between the primary and Metro Secondary data centers. SRDF/A, in turn, is used over SONET to locations over metro and WAN networks. This architecture provides the highest level of recoverability in the event of either local or metropolitan-wide disasters.
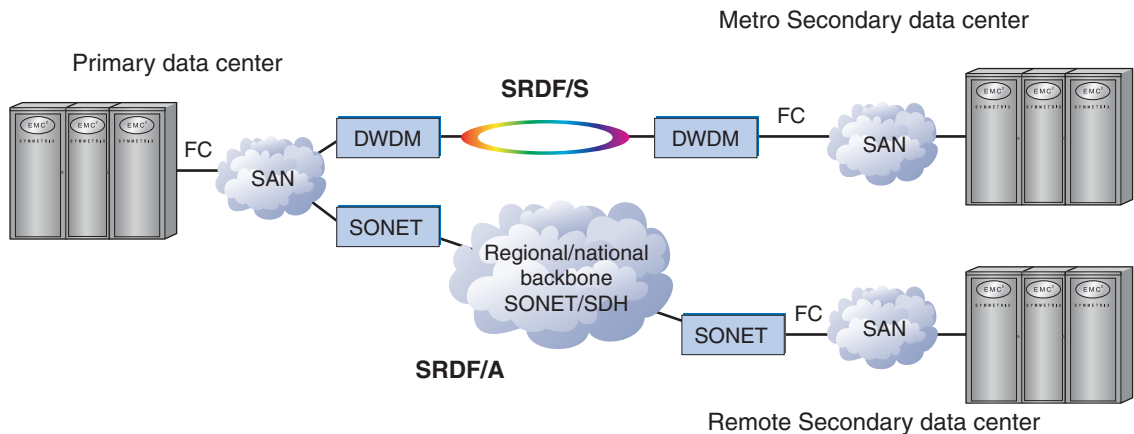
Figure 40   Concurrent SRDF/A example

Another example of a combined metro/WAN deployment is through the use of SRDF/AR multihop. An SRDF multihop topology allows you to string three sites together where (as shown in Figure 41 on page 101) a third SRDF site (Site C) is providing business continuance backup to the Metro Secondary SRDF site (Site B). In this multihop scheme, Site C is two hops (SRDF links) away, remotely backing up both the production site (Site A) and the Metro Secondary site (Site B) in the first hop.

Site A
Production
Host

SRDF/S
(Hop 1)

Site B
Metro Secondary
(Bunker)

SRDF
Adaptive Copy
(Hop 2)

Site C
Remote
Secondary

FC   SAN   DWDM   DWDM   FC   SAN   R2 BCV/R1   SAN   SONET   Regional/national backbone SONET/SDH   SONET   SAN   FC   R2

R1

Symmetrix   Symmetrix   Symmetrix

SYM-002510
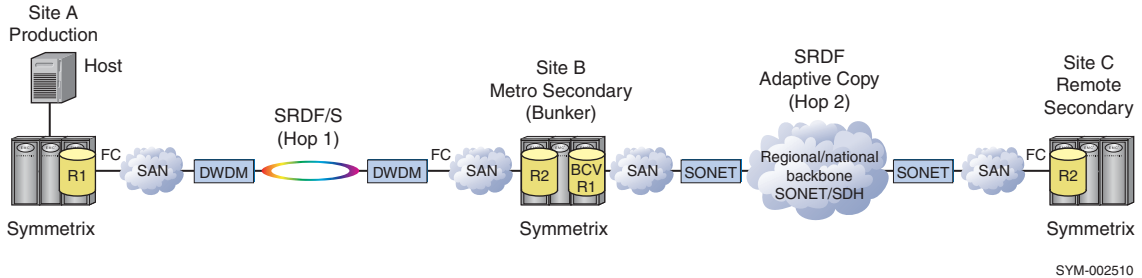
Figure 41   SRDF/AR Multihop example

With multihop operations, you can fully automate backup copying by using the SRDF Automated Replication (SRDF/AR) facility.

In the example shown in Figure 40 on page 100 and Figure 41, above, the Fibre Channel and SAN components could be interchanged with IP components for an equivalent IP configuration.

Note: SRDF flow control should be enabled for increased stability of the SRDF link. Further tuning of SRDF flow control can be made to improve performance. For more information, contact your EMC Customer Service representative.

WAN optimization

The typical configurations used for IP WAN with SRDF can be done using Fibre Channel or MPCD (multi-protocol channel director) ports with Gb/E capability out of the Symmetrix to a corresponding port on a given device. Vendors today supply WAN and data optimization devices with the capability to intercept either FCIP- or Gb/E- based traffic and perform optimization on the data flows.

The two types of architectures used are known as *in-line /bridge* mode or *out-of-path /routed* mode for connectivity.

An in-line (also known as *bump in the wire*) configuration denotes the insertion of the WAN optimizer in the path between the inbound and outbound data flow from the LAN and the WAN circuit, as shown in Figure 42.
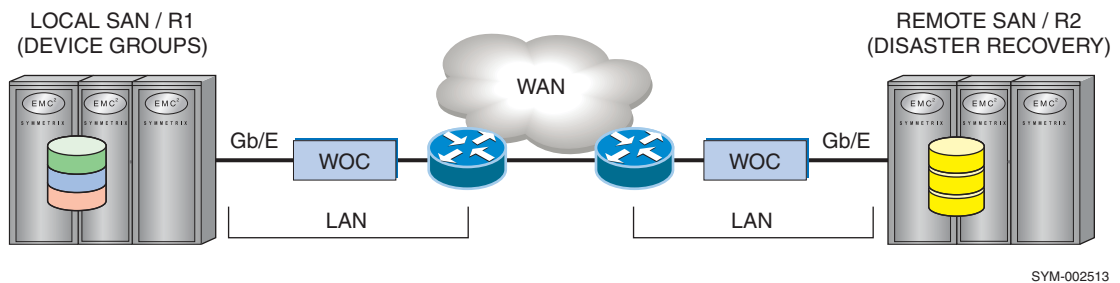

SYM-002513

Figure 42   **In-line configuration**

Out-of-path configurations involves the use of router redirection mechanisms such as PBR, WCCP, VRRP, or other proxy wherein the LAN and WAN traffic is redirected to the WAN and data optimization device for optimization, as shown in .

LOCAL SAN / R1
(DEVICE GROUPS)

WAN

REMOTE SAN / R2
(DISASTER RECOVERY)

Gb/E

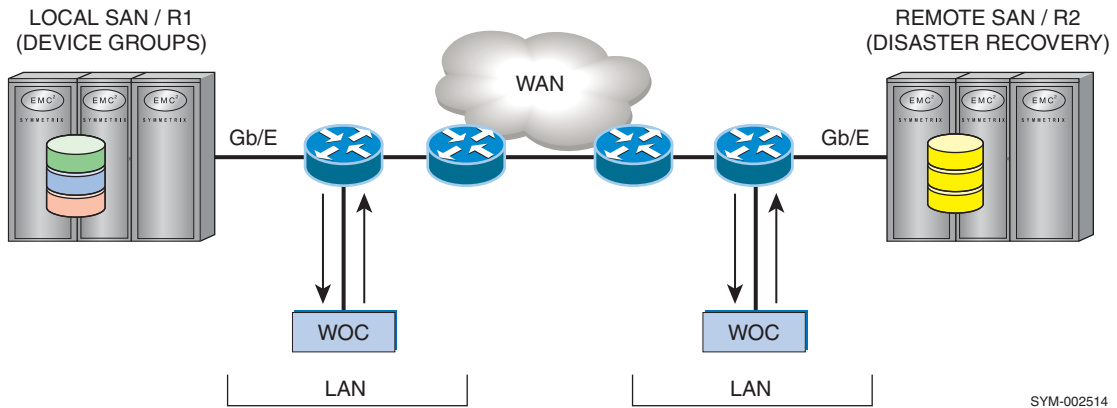Gb/E

WOC

WOC

LAN

LAN

SYM-002514

Figure 43   Out-of-path configuration

◆   PBR (Policy Based Routing) can be configured (depending on
    router/switch vendor) by single or several network centric
    characteristics such as the source and destination, packet size,
    protocol, etc.

◆   WCCPv2 (Web Cache Communication Protocol), like PBR,
    redirects traffic-based network-centric characteristics as it passes
    through the router, but also has load balancing capabilities.

◆   VRRP (Virtual Router Redundancy Protocol) is the original
    redundancy protocol that allows routing devices on the network
    to determine available paths. WAN optimizers can take
    advantage of this protocol for high availability.

Overall, WAN optimizers can be aligned with customer networking
best practices, and it should be clear to the customer that applications
using these devices can and should be prioritized based on their
WAN bandwidth/throughput requirements.

IP links can have high latency and/or packet loss. WAN optimization
appliances can be used to maximize the amount of data that can be
transmitted over a link. In some cases, these appliances may be a
necessity, depending on performance requirements. WAN
optimization can occur at layer 3/4 (network layer/transport), layer 4
(payload), and layer 4/7 (the session/presentation/ application
layer).

TCP was developed as a local area network (LAN) protocol.
However, with the advancement of the Internet it was expanded to be
used over the WAN. Over time TCP has been enhanced, but even

WAN optimization   **103**

with these enhancements TCP is still *not* well suited for WAN use for many applications.

The primary factors that directly impact TCP's ability to be optimized over the WAN are latency, packet loss, and the amount of bandwidth to be used. It is these factors on which the layer 3/4 optimization products focus. Many of these optimization products will re-encapsulate the packets into UDP or some other protocol. Some products create tunnels of one sort or another to perform their peer-to-peer connection between appliances for the optimized data.

Layer 3 optimization refers to the payload (data) within the packet and focuses on the reduction of actual payload as it passes over the network through the use of data compression, or data deduplication engines (DDEs). Compression is performed through the use of data compression algorithms (refer to "Compression" on page 53). A second type of data reduction is data deduplication using large data pattern tables and associated pointers (fingerprints). Large amounts of memory and/or hard-drive storage are used to store these pattern tables and pointers. These appliances build tables with pointers associated to the data that is to be reduced. Identical tables are built in the optimization appliances on both sides of the WAN with only smaller pointers passed over the network (versus resending data). While typical LZ compression ratio is about 2:1, DDE ratios can range greatly, depending on many factors. In general, 5:1 (and sometimes much higher ratios) can be achieved.

Layer 4/7 optimization is what is called the "application" layer of optimization. This area of optimization can take many approaches that can vary widely, but are generally done through the use of application-aware optimization engines. The actions taken by these engines can result in benefits, including reductions in the number of transactions that occur over the network or more efficient use of bandwidth. It is also at this layer the TCP optimization occurs.

Note: WAN optimization appliances deployed within IP networks are qualified by EMC. Consult the *WAN Optimization TechBook*, available through E-Lab Interoperability Navigator at http://elabnavigator.EMC.com, **Topology Resource Center** tab.

⚠️ **IMPORTANT**

Compression should be disabled on the GbE ports on the MPCD and the GbE director when a WAN optimization device employing data optimization is used. If compression is enabled on the GbE ports on the MPCD and the GbE director, data optimization benefits will be severely impacted resulting in increased WAN bandwidth needs.

Best practices, recommendations, and considerations

This section provides best practices, recommendations, and considerations for:

SRDF over a Metro or WAN

This section provides some best practices, recommendations, and configurations for involving SRDF over a Metro or WAN:

- SRDF zoning:
  - If the administrator so chooses, SRDF Single Zoning, where multiple sources and targets are contained in a single zone, can be used.
  - An SRDF zone should strictly contain RF ports.
  - If multiple zones are used, zoning must be designed in such a way that functionality, such as Dynamic SRDF, meets customer requirements. For example, if Dynamic SRDF is or will be in use, zoning requirements may change. In the case of Dynamic SRDF, any RFs that may have Dynamic SRDF connectivity established to one another must be in the same zone.
- A maximum of 10 RFs per switch is recommended. For example, say you have a two site configuration with four DMXs in each site. Each DMX contributes four SRDF ports. No single switch should have more than ten RFs connected to it. In this example, a minimum of two switches would need to be deployed at each site. The main reason for restricting a switch to ten RFs is due to NameServer traffic. NameServer traffic is an important consideration and needs to be kept to a minimum in order to

minimize link recovery times when RSCNs occur. By distributing across multiple switches, processing of NameServer traffic is also able to scale.

◆ Since many SRDF/S configurations use Fibre Channel and DWDM technology, ensure that the customer is using quality fiber and connectors and that the DWDM and Fibre Channel switch optics are properly configured and well within optical link budgets. Also ensure latency is minimized by choosing fibre runs carefully and by minimizing network equipment (switches, extenders, routers, etc) overhead and delay. Also ensure that selected Fibre Channel switches provide sufficient BB_Credits for planned distance (refer to "BB_Credit applied to distance" on page 63 for more information), including planning for 1K Fibre Channel frame size on DMX-1 and DMX-2. Figure 44 shows an example of Metro SRDF/S DWDM options for 2 G Fibre Channel.
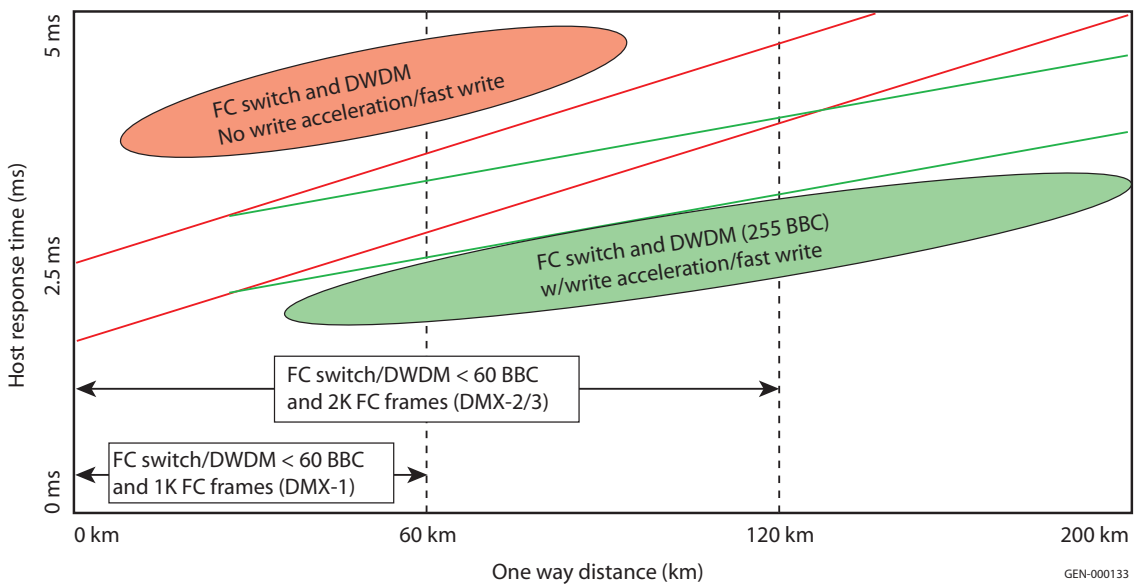

Figure 44   Metro SRDF/S DWDM connectivity options

◆ For point-to-point, low throughput SRDF deployments, best practice is to deploy a minimum of two links. For higher throughput SRDF/S deployments, best practice is to deploy at least four Fibre Channel (RF) links per Symmetrix since DMX scales well beyond 8 links. Ensure enough links to keep link utilization around 50% to minimize host response time for SRDF/S.

◆ For Metro SRDF/S implementations using DWDM or extended distance WAN connectivity, it is not recommended to extend Fibre Channel fabrics containing local host/storage traffic between data centers along with SRDF/S. Best practice is to isolate local host/storage fabrics within data centers so that link outages cannot cause disruption to the host connections to the SAN. Extension of ISL links between data centers over DWDM or WAN links can cause host timeouts (up to 25 seconds) if the SANs are not isolated, and the fabric goes through a rebuild process due to ISL failures. Recommendation is to use Fibre Channel features, such VSANs or FC routing, to prevent merging of the local and remote fabrics.

◆ Each Fibre Channel (RF) or GbE link should have a minimum of 10 Mb/s of dedicated bandwidth. On GbE interfaces, always set the speed limit in the Symmetrix Configuration file for each port, if bandwidth is less than 1 Gb/s per GbE port. Speed Limit should be set to (90% * available bandwidth / # GbE ports). When using Fibre Channel over extended distances, consider using SiRT on the SRDF/S links, or fast write/write acceleration in FC switches to eliminate one of two round trips per FC write I/O. Ensure that the group Speed Limit (if available) is enabled across all REs.

◆ As best practice, do not share RA ports between SRDF/S and SRDF/A groups. When doing data migrations (adaptive copy), use separate RAs, when possible, to minimize performance impacts.

◆ Packet loss should be zero, but should never exceed 0.1%. Latency over distance is a given. Packet loss is not. Design the network to be resilient by having diverse network paths with the ability to self-heal and transparently failover. Ensure the network is able to scale to allow for future growth.

SRDF/AR and adaptive copy

This section provides recommendations and considerations for SRDF/AR and adaptive copy:

◆ SRDF operations for SRDF/AR and adaptive copy Disk mode transmit full tracks across the links as SRDF copy tasks. DMX-1 and DMX-2 use 32K (OS) and 56K (MF) tracks, while DMX-3, DMX-4, and VMAX use 64K and 56K tracks, respectively. The exception for DMX-3, DMX-4, and VMAX is when the R2-side Symmetrix is an older model. In this case, the R1-side DMX-3, DMX-4, and VMAX will send 64K OS tracks as two 32K tracks to the older Symmetrix. This does *not* affect mainframe I/O because of the standard 56K track size.

◆ The SRDF copy tasks (invalid tracks owed to the secondary Symmetrix) are controlled by the DAs. The DAs are responsible for fetching the invalid tracks and placing them on the SRDF copy queues. The priority of the background copy tasks versus the higher priority destaging of writes in cache to physical disk is controlled by the SRDF QoS setting and the DA copy task (also referred to as copy job) parameter.

◆ This DA copy track parameter is the number of tracks that the DAs will place on the SRDF copy queue at a time. In Enginuity version 5x71, the default value was doubled from 40 (28 hex) tracks to 80 (50 hex) tracks per SRDF group. For VMAX, this setting is 256 (100 hex) tracks per SRDF group. Since this is the maximum number of tracks that can be on the SRDF queue per SRDF group, performance will be throttled over high latency links. For long-distance replication, the parameter may be increased by EMC Customer Service to improve performance (max value is 255). For example, two SRDF links at 40 ms RTT: Assuming Enginuity version 5x71, 80 tracks X 32K/track / 40 ms = 64 MB/s peak (2 RE (GbE) or 2 RF w/ fast write support). This throughput would drop about 20-30% (~50 MB/s) if two round trips over Fibre Channel are used.

◆ Spreading the volumes across multiple SRDF groups for long-haul data migrations will also provide a multiplier for performance, but be aware that adjusting the DA copy task setting may impact DA destaging of host writes from cache and host response time. For Enginuity version 5670 on DMX-1 and DMX-2, if production is active during data migrations, do not increase copy task setting beyond 70 (46 hex) and use SRDF QoS

to limit front-end impact during the replication process. Keep in mind that adaptive copy mode will run much more efficiently if many volumes are active (use host striping and metal, etc.) and spread across the physical disks.

SRDF/A

This section contains some considerations and best practices for SRDF/A.

Some considerations for SRDF/A include:

SRDF/A, unlike adaptive copy, is not throttled by DA activity since the SRDF updates are retained directly in cache for transfer to the secondary Symmetrix. SRDF/A will burst at the SRDF link limits, whether its measured in maximum I/Os per second per link for smaller writes, or MB/s for large writes, since the I/Os are only throttled by the following two parameters:

◆ Number of I/Os allowed on the SRDF link per volume: 8 for ESCON; 64 for Fibre Channel or GbE

◆ Number of SRDF I/Os allowed on the SRDF link: 400

Typically, the two round trips (FC) versus one (GbE or Write Acceleration) does not impact SRDF/A performance since SRDF/A will send continuous stream of writes until the cycle is empty. If the links are kept full, there is no waiting for round trip delay and SRDF/A performance is not impacted by two round trips. Fibre Channel fast writes will improve performance for SRDF/A when there are "gaps" on the links due to retransmissions after link errors, or small writes that throttle SRDF throughput due to hitting I/Os per second limits on the Fibre Channel or GbE links. The number of active volumes in a cycle can also affect performance on links where the links may be hitting the 64 writes/volume limit. Both Fibre Channel and GbE links load balance on disparate links (different speeds/latencies).

Some best practices for SRDF/A include:

◆ Do not over-subscribe the link! SRDF/A does not put data on the link in a steady stream. It does it in bursts at the SRDF link limit. If you have several SRDF groups, and if these SRDF groups have different peak I/O periods, one might think that it would be possible to set the throttle rate for an individual SRDF group to a level that is less than the overall bandwidth. However, the overall combined throttle rate must be calculated with regard to all SRDF

groups in relation to the available bandwidth. Because of how SRDF/A works, in that it is always sending changes and will utilize all available bandwidth (i.e., equal to the throttle rate), careful planning must be done to prevent this sort of overrun of the network. An example of an SRDF/A cycle is shown in Figure 45.
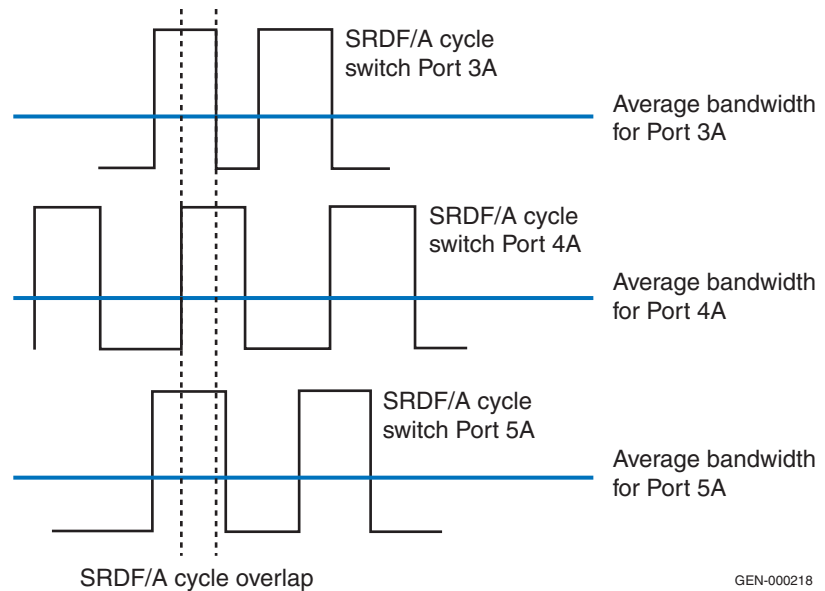

SRDF/A cycle switch Port 3A

Average bandwidth for Port 3A

SRDF/A cycle switch Port 4A

Average bandwidth for Port 4A

SRDF/A cycle switch Port 5A

Average bandwidth for Port 5A

SRDF/A cycle overlap   GEN-000218

Figure 45   SRDF/A cycle

◆ Do not mix SRDF/S and SRDF/A groups on the same RAs. SRDF/A performance on shared RAs may be impacted during heavy SRDF/S write activity because of the SRDF/S queuing algorithm.

◆ Do not mix data migrations or resyncs and active SRDF/A sessions on same RA ports, if possible, due to performance impact.

◆ For Enterprise SRDF/A solutions where the network is designed to guarantee bandwidth that far exceeds the workload (and is also fully redundant), you need not worry about efficient compression to maximize link utilization, nor locality of reference concerns, because bandwidth is not a bottleneck. In these scenarios, consider decreasing the SRDF/A cycle time to improve average cycle variation (minimize secondary delay and RPO) and improve recovery time after workload spikes.

The problem with high bandwidth and 30-second cycles is that the SRDF transfers complete in a fraction of the 30-second cycle, then must wait for the 30-second timer to expire rather than trying to switch cycles as soon as the data transfer completes to the R2s. A shorter cycle will also flatten out the peaks more efficiently by decreasing the amount of cycle elongation during workload spikes.

In single session SRDF/A, you can reduce the cycle time. Refer to the appropriate MSC documentation for more information.

SRDF/Star

The following is a best practice and recommendation for SRDF/Star:

The three-site WAN network for Star must be designed so that all WAN links are active during Concurrent SRDF operation between production data center and the bunker and DR sites (A to B and A through B to C, as shown in Figure 46 on page 113. This provides better performance and ensures that the WAN links are functional when a failover to site B or C occurs, rather than having a "standby" link between Bunker (site B) and DR site (C) that is only used when a disaster is declared.

The example configuration shown in Figure 46 uses Fibre Channel switches to provide exchange-based routing across all ISLs and FSPF to round robin between all available WAN paths. Because the Fibre Channel switches have ISLs that can access all WAN links (directly from production to DR (A-C) and via DWDM to Bunker site, then onto second WAN link to DR site (A-B-C)), equal cost FSPF for the two WAN links guarantees that both WANs will be active from production site for asynchronous replication (SRDF/A) to DR site.
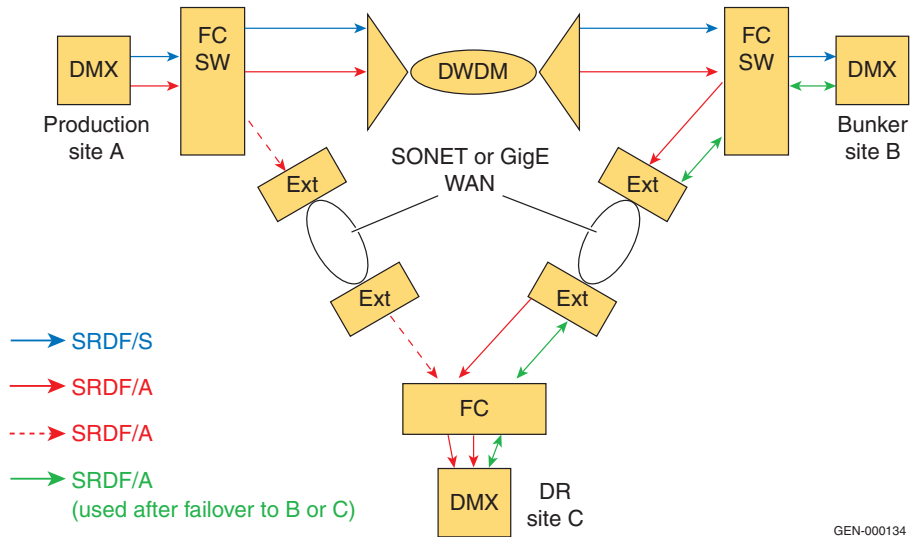
Figure 46 **SRDF/Star example**

For more information on SRDF/Star, it benefits, limitations, and how it works, refer to the *EMC Symmetrix Remote Data Facility (SRDF) Product Guide*, located at https://support.emc.com, Symmetrix Remote Data Facility (SRDF) product page, Documentation.

GbE

This section provides best practices, recommendations, and considerations for GbE:

◆ Always set the speed limit for each and every GbE port. The combined throughput of all GbE ports should not exceed 95% of the available dedicated bandwidth for SRDF.

◆ Set the maximum individual GbE port throttle setting to be no more than 950 Mbs and to never utilize the default zero (0) setting. This should be done whether or not WAN optimizers are being used or if there is available bandwidth above the combined GbE port throughput level.

◆ A TCP session (one per logical path) on DMX has default window size of 1 MB (2 MB for DMX-3, DMX-4, or VMAX), which determines maximum throughput per session, based on actual window size/RTT (round trip time in ms). The actual window

Best practices, recommendations, and considerations

size depends on available bandwidth and network delay (Bandwidth * RTT). In a switched configuration, where all primary GbE ports "see" all secondary GbE ports, then the number of TCP sessions is the number of primary ports multiplied by the number of secondary ports. GbE, like switched FC, supports point-to-multipoint configurations.

◆ To prevent link problems, it is recommended to use minimum of 10 Mb/s per GbE port (RPQ for > 6 Mb/s). If the link is not on a quality (.01% or less packet loss) network, RTT should be limited to under 60 ms to prevent link recovery and performance problems. For networks that have packet loss problems that are over 40 ms RTT, WAN optimization products should be considered to help improve performance through optimized retransmissions and recovery. For Enginuity version 5873 and later, it is a best practice is to utilize the auto-adjust speed limit feature across a group of GbE ports. Any number of SRDF over GbE ports can be grouped together using an aggregate bandwidth for the entire group. The speed limit feature will automatically adjust the speed limits for the individual ports to maintain that bandwidth into the network should any of the group's ports become inactive (for any reason).

Note: Ensure that the customers' IP networks are using end-to-end routing (not packet level) to prevent out-of-order packets. GbE will re-order packets, but it will affect performance.

◆ Consider the use of a WAN optimization appliance if any one or a combination of bandwidth, latency, and/or packet loss are limiting factors. When using these appliances, please note the following:

• If the WAN optimizer does not support auto disabling of compression for the GbE ports, compression should be disabled (in the Symmetrix configuration) for each GbE port on the MPCD and the GbE director port, when a WAN optimization device employing data optimization is used. If compression is enabled on the GbE ports on the MPCD and the GbE director, data optimization benefits will be severely impacted resulting in increased WAN bandwidth needs.

• When using GbE, SRDF flow control is on by default for increased stability of the SRDF link. Further tuning of SRDF flow control can be made to improve performance. For more

information, contact your EMC Customer Service representative. For more information on flow control, refer to "Flow control" on page 50.

- Refer to the *EMC WAN Optimization Controller Technologies TechBook*, available on the E-Lab Navigator, at http://elabnavigator.EMC.com, **Topology Resource Center** tab, regarding configuration details and settings for specific WAN optimization appliances when used with SRDF. SRDF link stability and/or performance can be severely compromised if the correct settings on the WAN optimization appliances are not made. There are many factors that impact how much optimization will occur, therefore results will vary.

◆ DMX to Symmetrix 8000 GbE configurations must be carefully planned. In these configurations, compression is auto-disabled so bandwidth is not efficiently used. GbE flow control is not available with Symmetrix 8000, so low-bandwidth, high-latency links will not be able to effectively utilize the bandwidth. Consider using external compression products. Additional bandwidth provisioning may also be required.

◆ Spanning Tree Protocol (STP) is commonly used in many large networks where three or more switches are attached together within the network. One downside of STP is that when a change occurs within the switched fabric, an updating of the mapping tables has to take place. During this update process, all traffic is stopped within the VLANs that were affected by the change until the update has finished. This update typically takes anywhere from 30 to 60 seconds. SRDF links can drop due to the length of the outage unless Link Limbo timer and Transmit Idles features are invoked for the SRDF links.

On Cisco switches that transport SRDF traffic and where the VLAN that they utilize could be affected by a STP reconvergence, it is recommended that a feature called *portfast*, *fastlink*, or *fastbackbone* be enabled. This feature allows for an instant failover without having to wait for the reconvergence time. For non-Cisco switch equipment, it should be determined if such a feature is supported and if it possible for this feature to be enabled. If this feature is not enabled, it is recommended that if a known forced reconvergence is to occur (for example, maintenance within a network) that appropriate measures be taken to lessen the effect of the timeout on SRDF traffic. For instance, if a maintenance

event is scheduled and SRDF/A is active, it would be recommended to change to SRDF adaptive copy mode during this period.

Fibre Channel

This section provides best practices, recommendations, and considerations for Fibre Channel.

◆ Consider the use of a WAN optimization appliance if any one or a combination of bandwidth, latency, and/or packet loss are limiting factors. When using these appliances, please note the following:

• Compression should be disabled (in the Symmetrix configuration) on any device (such as FCIP channel extenders) prior to ingress of a WAN optimization device employing data deduplication. If compression is enabled, data deduplication benefits will be severely impacted resulting in increased WAN bandwidth needs.

• SRDF flow control should be enabled for increased stability of the SRDF links. Further tuning of SRDF flow control can be made to improve performance. For more information, please contact your EMC Customer Service representative. Refer to Table 1 on page 28 for information on flow control support. For more information on flow control, refer to "Flow control" on page 50.

• Refer to the *EMC Networked Storage Topology Guide* regarding configuration details and settings for specific WAN optimization appliances and FCIP devices when used with SRDF. SRDF link stability and/or performance can be severely compromised if the correct settings on the WAN optimization appliances and FCIP devices are not made.

• In general, it has been observed that optimization ratios are higher with GbE ports on the MPCD and the GbE director vs. FCIP. There are many factors that impact how much optimization will occur, therefore results will vary.

• In general, it has been observed that optimization ratios are higher with SRDF Adaptive Copy than SRDF/A. There are many factors that impact how much optimization will occur, therefore results will vary.

ESCON/FarPoint

This section provides best practices, recommendations, and considerations for ESCON/FarPoint:

◆ ESCON FarPoint requires at least 1.5 Mb/s of bandwidth per RA and supports T1/E1, T3/E3 point-to-point WAN circuits, as well as ATM/OC3, SONET, and IP networks, using extenders. Throughput per RA link is limited to 10-12 MB/s, but for low-bandwidth (T3), low-throughput SRDF applications, FarPoint is still a viable configuration. As the legacy SRDF protocol, it is limited by scalability (8 ESCON links < performance of single Fibre Channel link) and uni-directional limitations.

◆ FarPoint protocol eliminates standard ESCON flow control/hand-shaking to maximize performance over distance. There are only two parameters (settable in Symmetrix Configuration file or by EMC Customer Service) to manage the performance and streaming of data on a FarPoint RA link.

  • The first parameter is used for troubleshooting, controls the number of idles between 1 K ESCON frames, and therefore directly affects the peak throughput (10-12 MB/s) of the links.

  • The second, and more important, parameter is often used by EMC Customer Service personnel to "tune" the FarPoint links is the FarPoint Buffer setting. This parameter (Symmetrix Configuration file default is 200,000 bytes per RA) is the maximum amount of data that an RA can pipeline on the link (truncated to multiples of full write I/Os, since partial writes cannot be transmitted on an SRDF link). This value divided by round trip time (RTT) determines maximum throughput (MB/s) per RA.

  The value is frequently decreased for low bandwidth links (speed matching) to prevent over-running network extender buffering. For long distance data migrations, the FarPoint Buffer size can be increased by EMC Customer Service to improve FarPoint performance. The RA pipeline depth (average I/O on the link) and round trip times are displayed in the 'A2, F3' RA statistics as queue depth and Echo time.

A

Cable Specifications

This appendix contains information on network cable specifications and supported distances.

Cable specifications and supported distances

Table 5 provides network cable specifications and supported distances.

A *cable segment* is a physical cable connecting one node to another.

Table 5   Cable specifications

Primary SRDF links	Network supported		Nominal distance
SRDF/ESCON	Direct fiber, multimode:	62.5/125 μm cable	3 km/cable segment
		50/125 μm cable	2 km/cable segment
	• Fiber repeaters/converters (e.g., McDATA 9191 to McDATA 9191 using 9 μm cable) • Three converters/repeaters maximum allowed		30 km/cable segment
	• Other fiber repeaters/converters (using 9 μm cable) Three converters/repeaters maximum allowed		20 km/cable segment [a]
	WDM (MAN)		Varies [b,d]
	LAN / WAN — T1/E1, T3/E3, ATM, IP		Varies [d]

a. Typical best-case distance specifications are quoted; refer to specific vendor's equipment for current information.

b. Limited by vendor's DWDM laser technology and the type of cable for the Symmetrix system-to-DWDM cable segment.

c. Optics availability varies with Fibre Channel switch model. Refer to the *EMC Support Matrix* and the *EMC Networked Storage Topology Guide*, available through E-Lab Interoperability Navigator at: http://elabnavigator.EMC.com for the latest supported distances.

d. Use of SRDF synchronous mode is limited by the application's ability to tolerate delay introduced by distance. Buffering techniques, such as SRDF/A, SRDF adaptive copy mode, and TimeFinder® BCVs, are available to address distance requirements.

Table 6 shows linked distances for both Fibre Channel and GbE links.

Table 6   **Link distance with speed and fiber type (meters)**

Speed	OM1 Link Distance 62.4um core 200 MHZ km [a]	OM2 Link Distance 50um core 500 MHZ km [a]	OM3 Link Distance 50um core 2000 MHZ km [a]	OM4 Link Distance 50um core 4700 MHZ km [a]	OS1 Link Distance 9um core ~ Infinite MHZ km [a]
1 GFC	300	500	860	[a]	10,000
2 GFC	150	300	500	[a]	10,000
4 GFC	50	150	380	400	10,000
8 GFC	21	50	150	190	10,000
10 GFC	33	82	300	[a]	10,000
16 GFC	15	35	100	125	10,000

a.   The link distance on OM4 fiber has not been defined for these speeds.

CAUTION

Some vendors place very powerful lasers in their systems to improve signal strength. Other vendors have receivers that can be damaged or even destroyed by this intense light. Based on the connection distance involved, light attenuation may be needed between equipment. Check input and output ranges carefully before operational testing.

Bandwidth by WAN Link

This appendix provides bandwidth ratings for the various WAN links.

Bandwidth by WAN link

Table 7 provides bandwidth ratings for the various WAN links. Performance varies depending on many factors, including I/O size, configuration, and media.

Note: The availability of links and the "last mile" cost can have a significant impact on total bandwidth costs and can vary substantially by region.

Table 7   Bandwidth by WAN link

Electrical level	SONET[a] level	SDH[a] level	Payload rate (Mb/s)	Overhead rate (Mb/s)	Rated bandwidth (MB/s)	Comments
1 GbE	N/A	N/A	1025		125	
10 GbE			10250		1250	
DS1/T1	N/A	N/A			0.193	Can inverse multiple links to act like one
DS3/T3	N/A	N/A			5.5	
E1	N/A	N/A			0.256	International link
E3	N/A	N/A			4.3	International link
STS-1	OC-1	N/A	50.112	1.728	6.4	
STS-3	OC-3	STM-1	150.336	5.184	19.4	
STS-12	OC-12	STM-4	601.344	20.736	777.7	
STS-48	OC-48	STM-16	2405.376	82.944	311	
STS-192	OC-192	STM-64	9621.504	331.776	1244	
STS-768	OC-768	STM-256	38486.016	1327.104	4976	

a. SONET and SDH (the European equivalent of SONET) are optical equivalents of electrical STS signals.

Glossary

This glossary contains terms related to disk storage subsystems. Many of these terms are used in this manual.

A

ATM
Asynchronous Transfer Mode, a high-speed, connection-oriented switching technology that can transmit voice, video, and data traffic simultaneously through fixed-length packets called *cells*. ATM is used as a protocol for LAN and WAN traffic, as well as a transport protocol for IP packets. ATM is generally associated with SONET, which provides the optical link layer.

adaptive copy disk mode (AD)
Symmetrix SRDF Adaptive copy disk mode is designed for bulk data transfer. Host write tasks accumulate on the primary volume rather than in global memory. A background process destages the write tasks to the corresponding secondary volume. When a skew value is reached, the primary volume reverts to its primary mode of operation, either synchronous or semi-synchronous, whichever is currently specified.

adaptive copy mode
Symmetrix SRDF Adaptive copy modes facilitate data sharing and migration. These modes allow the primary and secondary volumes to be more than one I/O out of synchronization. The maximum number of I/Os that can be out of synchronization is known as the maximum skew value. The default value is equal to the entire logical volume. The maximum skew value for a volume can be set using the SRDF monitoring and control software.

adaptive copy write pending mode (AW)	An SRDF mode of operation that stores new data for a remotely mirrored pair in the cache of the local Symmetrix storage subsystem until it can be successfully written to both the primary and secondary volumes.
Asynchronous SRDF	SRDF/Asynchronous (SRDF/A) mode for SRDF secondary devices provides a dependent-write consistent copy on the R2 device, which is only slightly behind the currency of the primary R1 device. SRDF/A session data is transferred to the remote Symmetrix system in cycles or *delta sets*, which minimizes the redundancy of same-track changes being transferred over the SRDF links.
automated replication	SRDF/Automated Replication (SRDF/AR) performs automated consistent replication of data from standard devices and from BCV R1 devices over SRDF links to the remote SRDF pair.

B

back-up	The process of creating a copy of data to ensure against accidental loss.
BCV device	A Business Continuance Volume (BCV) is Symmetrix storage device that holds a mirror copy of the primary data on the Standard device. A BCV Device contains a copy of data from a standard Symmetrix device that is online for regular I/O operation from its host(s). The Standard device and its assigned and mirroring BCV device are known as a BCV pair.
BCV pair	The Standard device and its assigned and mirroring BCV device are known as a BCV pair.
bidirectional SRDF link	If an SRDF group contains both primary and secondary volumes, write operations move data in both directions over the SRDF links for that group. This is called an SRDF bidirectional configuration.
bit	The smallest unit of computer memory. A bit can hold a value of zero or one.
Buffer-to-Buffer Credits (BB_Credits)	The number of internal 2 KB buffers on a communications chip in a Fibre Channel device. More credits for longer distance enables more efficient use of the available link capacity.
business continuance	Solutions that maintain mirror images of data at the device level in Symmetrix systems located locally and in physically separate sites.

Both TimeFinder and SRDF operations provide a recovery solution for component site failures between locally and remotely located storage devices.

business continuance volumes (BCVs) A standard Symmetrix device with special attributes that allow it to independently support applications and processes, such as backup operations, restore operations, Decision Support operations, and application testing. BCV devices are available through the *EMC TimeFinder* software.

byte Any 8-bit unit of data storage. For SRDF applications with FC, GbE, and ESCON, a byte is 10 bits.

C

cache Random access electronic storage used to retain frequently used data from disk for faster access by the host.

cache slot Unit of cache equivalent to one track.

Cascaded SRDF A three-site SRDF replication configuration that provides a data recovery solution. This uses a device (R21) that is both and R2 to an R1 mirror and an R1 to an R2 mirror.

change tracker A tool available to EMC field engineers to aid in the estimation of bandwidth required in an SRDF configuration. Change Tracker allows a user to monitor the update activity on Symmetrix devices, and is especially useful for sizing SRDF in a multihop configuration.

Change Tracker consists of two programs: a collector program that gathers the update activity statistics, and a reporter program that analyzes the data and formats it into useful reports.

channel director The component in the Symmetrix subsystem that interfaces between the host and data storage. It transfers data between the host and global memory.

clone TimeFinder/Clone operations allow you to create point-in-time copies of a source (prime or backup) device on multiple target devices.

compatible peer Symmetrix Enginuity version 5568 (and later) enables the Symmetrix system to support native IBM Peer-to-Peer Remote Copy (PPRC)

commands through a Symmetrix feature called EMC Compatible Peer.

concurrent SRDF   Supports the ability for a single primary volume to be remotely mirrored to two secondary volumes concurrently. This feature is called Concurrent SRDF and is supported in ESCON, Fibre Channel, and Gigabit Ethernet SRDF configurations. In Concurrent SRDF, R11 volumes must be assigned to two *different* SRDF groups. This is because each remote mirror can only be assigned to one SRDF group, even if the two secondary volumes reside in the same Symmetrix system.

consistency group   A consistency group (CG), known as SRDF/CG) is a group comprised of primary SRDF devices, which have been enabled for remote point-in-time consistency. These SRDF consistency groups operate in unison to preserve the integrity and dependent-write consistency of a database distributed across multiple Symmetrix systems.

Count-Key-Data (CKD)   A data recording format employing self-defining record formats, used in mainframe hosts, in which each record is represented by a count area that identifies the record and specifies its format, an optional key area that may be used to identify the data area contents, and a data area that contains the user data for the record. CKD can also refer to a set of channel commands that are accepted by a device that employs the CKD recording format.

cycle time   Time elapsed between SRDF/A cycle switch operations.

D

DA   Disk adapter. Generally the term *DA* applies to the disk director on all Symmetrix models except the DMX-800. The disk director and its disk adapter provide the interface between cache and the disk devices.

In the DMX-800, DA functionality is provided by a multifunction front-end/back-end (FEBE) board. For the purpose of this document, the term *DA* applies to the cache-to-disk interface in any Symmetrix model.

dark fiber/lit fiber   Fiber-optic cable that is inactive (dark) until the customer lights it up with new traffic. When leasing dark fiber, the customer has full use of the bandwidth on that fiber. In a lit (currently used) fiber service, a

wavelength or specific protocol such as GbE has been provided to the customer's building or data center. It is already lit up with other traffic since it is a shared fiber. Note that a band in a DWDM fiber could be considered dark if it is not yet used.

data availability Access to user data at all times by user applications.

Delta Sets SRDF/A cycles.

DSE SRDF/A Delta Set Extension (DSE) (Enginuity version 5772 and later) enables asynchronous replication operations to remain active in the event system cache resources are becoming in danger of reaching the system write pending or SRDF/A max cache limit. This functionality is achieved by offloading some or all of the active Delta Set data into preconfigured storage pools. DSE can be enabled on both primary and secondary sites (*EMC recommended*).

destage The asynchronous write of new or updated data from global memory to disk drive.

device A uniquely addressable part of the Symmetrix subsystem that consists of a set of access arms, the associated disk surfaces, and the electronic circuitry required to locate, read, and write data.

device address The hexadecimal value that uniquely defines a physical I/O device on a channel path in an z/OS environment.

device number The value that logically identifies a Symmetrix logical volume.

disk director The component in the Symmetrix subsystem that interfaces between cache and the disk devices.

director Host adapters in the Symmetrix system provide (through adapters) the interfaces between the host and data storage. SRDF directors are host adapters (GbE, Fibre Channel, or ESCON) that are configured for SRDF.

disk director The component in the Symmetrix subsystem that interfaces between global memory and the disk drives.

dual-directional SRDF link For an ESCON-based extended-distance SRDF configuration (for example, using E1, E3, T1, T3, and ATM links and/or IP network connections) that require data to move in two directions, a dual-directional link configuration may be required.

DWDM Dense Wavelength Division Multiplexing, a process in which multiple different or multiple individual channels of data are carried at different wavelengths over one pair of fiber links. DWDM assigns one wavelength of light (lambda) to each data stream applied. These wavelengths are then multiplexed through a single fiber pair. The process is protocol independent. The channel module locks in on the bit rate assigned. Distance is limited by the power of the laser, typically 80 km or less. Ratios of 32:1 protected channels are available for enterprises, and 128:1 for carrier-class systems.

E

E1 Also known as CEPT1, the 2.048 Mb/s rate used by European CEPT carriers to transmit 30 64 Kb/s digital channels for voice or data calls, plus a 64 Kb/s signaling channel and a 64 Kb/s channel for framing and maintenance. E1 has similar latency characteristics to T1. Refer to T1.

E3 Also known as CEPT3, the 34.368 Mb/s rate used by European CEPT carriers to transmit 16 CEPT1s plus overhead. E3 has similar latency characteristics to T3. Refer to T3.

ECA Enginuity Consistency Assist (ECA) provides consistent split operations across multiple, heterogeneous hosts. Used in consistency groups, TimeFinder, and SRDF/AR split operations for consistent and restartable copy of a database management system.

Enginuity The operating environment for Symmetrix 8000 series and newer systems that provides high-speed, highly available storage management functionality. The term is equivalent to *microcode* in older Symmetrix systems.

ESCON An IBM mainframe serial-I/O channel path that allows attachment of one or more control units to an ESA/390 channel subsystem by using optical fiber links. Best case throughput is 17 MB/s. This forms the basis of the original transport layer for SRDF between Symmetrix systems and still works very well today.

ESCON director Device that provides a dynamic switching function and extended link path lengths (with XDF capability) when attaching a mainframe ESCON channel or SRDF director (RA1, RA2).

ESP Enterprise Storage Platform. Symmetrix ESP is a functional enhancement that allows simultaneous storage and access of

mainframe data and open systems data on the same Symmetrix system.

F

FA   Symmetrix Fibre Channel director. When this director is configured for SRDF, it becomes an RF.

fabric   Fibre Channel employs a fabric to connect devices. A fabric can be as simple as a single cable connecting two devices. The term is often used to describe a more complex network utilizing hubs, switches, and gateways.

FarPoint   FarPoint is an SRDF feature used only with ESCON extended distance solutions (and certain ESCON campus solutions) to optimize the performance of the SRDF links. This feature works by allowing each SRDF director to transmit multiple I/Os, in series, over each SRDF link.

fast write   In Symmetrix, a write operation at cache speed that does not require immediate transfer of data to disk. The data is written directly to cache and is available for later destaging.

FBA   Fixed Block Architecture. Disk drive data storage format using fixed-size data blocks.

Fibre Channel   Open systems high-performance serial-I/O channel path, with data rates up to 200 MB/s. Fibre Channel transports SCSI commands and data in serial frames, which can be routed in a switched environment.

Fibre Channel combines the performance, hardware-based functionality, and reliability of I/O channels with the distance and routing benefits of serialized network protocols. This forms the basis of the second-generation transport layer for SRDF.

FICON   An I/O interface based on the Fibre Channel architecture. In this new interface, the ESCON protocols have been mapped to the FC-4 layer, that is, the Upper Level Protocol layer, of the Fibre Channel Protocol. It is used in the S/390 and z/Series environments.

Frame   Data packet format in an ESCON, Fibre Channel, or GbE environment.

G

gatekeeper One of the disk storage devices, otherwise part of the storage capability, becomes designated as the Gatekeeper to handle all feature command I/O between host and the Symmetrix system. Typically, multiple gatekeepers are assigned to a Symmetrix system.

Gb/s Gigabits per second.

gigabyte (GB) 10^9 bytes. The GB_2 value is based on the convention:
1 GB = 1024 * 1024 * 1024 bytes.
The GB_{10} value is based on the convention:
1 GB = 1000 * 1000 * 1000 bytes.
Although these values are expressed differently, their values are equivalent.

GbE Gigabit Ethernet, an Ethernet framing format operating at 1024 Mb/s, primarily over fiber. It is a full-duplex implementation using a variable-length payload of 64-byte to 1514-byte packets. GbE is used primarily as a LAN backbone.

global memory Random-access electronic system storage used to retain frequently used data for faster access.

group The physical disks related to each other for common parity protection when implementing the RAID 5 option.

H

host adapter The component in the Symmetrix subsystem that interfaces between the host and data storage. It transfers data between the host and cache.

hypervolume extension The ability to define more than one logical volume on a single physical disk drive making use of its full formatted capacity. These logical volumes are user-selectable in size. The minimum volume size is one cylinder and the maximum size depends on the disk drive capacity and the emulation mode selected.

I

I/O device An addressable input/output unit such as a disk drive.

IP Internet Protocol, designed to provide addressing and routing at a level capable of being routed over multiple transports, enabling the creation of internets. Transports are rapidly consolidating to Ethernet for the LAN, and SONET and dark fiber for the WAN.

IPL Initial Program Loading.

IPv6 IPv6 is the latest version of the Internet Protocol (IP), sometimes called IPng. IPv6 is designed to solve the challenges of the current IP version, IPv4, by providing support for more addresses, better security, reduced administrative workload, and enhanced support for mobile devices.

IPsec IPsec (Internet Protocol Security) is a framework for security of network communication. IPsec allows two computers to trust each other's identity and share a private key for each communications session. Based on this session key, the computers establish an encrypted communications channel and can verify that data they receive is what was originally sent.

internet Spelled with a lowercase "I," a collection of networks interconnected by a set of routers that allow them to function as a single, large virtual network.

Spelled with an uppercase "I," the largest internet in the world, consisting of large national backbone nets and a variety of regional and local campus networks all over the world. The Internet uses the Internet Protocol suite.

ISL Interswitch link. The link between two Fibre Channel switches in a fabric.

K

kilobyte (KB) The KB_2 value is based on the convention:
1 KB = 1024 bytes.
The KB_{10} value is based on the convention:
1 KB = 1000 bytes.
Although these values are expressed differently, their values are equivalent.

L

link path A single ESCON Light Emitting Diode (LED) fiber optic connection between the two Symmetrix storage subsystems. A minimum of two to a maximum of eight links can exist between the two units.

local volume A Symmetrix logical volume that is not participating in SRDF operations. All CPUs attached to the Symmetrix may access it for read/write operations. It is available for local mirroring or dynamic sparing operations to the Symmetrix storage subsystem in which it resides only.

logical volume A user-addressable unit of storage. In the Symmetrix subsystem, the user can define multiple logical volumes on a single physical disk device.

LPAR Logical Partition or Logical Partitioning Mode. Mainframe system function that allows different operating systems to run concurrently in separate logical partitions.

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

M

Mb/s Megabits per second.

media The disk surface on which data is stored.

megabyte (MB) The MB_2 value is based on the convention:
1 MB = 1024 * 1024 bytes.
The MB_{10} value is based on the convention:
1 MB = 1000 * 1000 bytes.
Although these values are expressed differently, their values are equivalent.

metadevice and metamember Metadevices allow for the designation of individual devices to be concatenated together, creating larger storage devices. When the Metahead (first device in the member chain) is addressed with a command, the Enginuity process determines which metadevice is the target for the command execution. Metadevice addressing, by striping, divides each Metamember device into a series of stripes, then addresses a stripe from each device before finally advancing to

the next stripe on the first device. The striping of data across multiple drives effectively creates definable cylinder stripes.

mirroring   The Symmetrix option that maintains two identical copies of a designated volume on separate disks. Each volume automatically updates during a write operation. If one disk drive fails, the Symmetrix system automatically uses the other disk drive.

mirrored pair   A logical volume with all data recorded twice, once on each of two different physical devices.

MPCD   Multiprotocol Channel Director, a two-port or four-port Symmetrix director that can be configured for GbE, iSCSI, or FICON. Refer to "Multiprotocol Channel Director on the Symmetrix DMX series" on page 44.

MTU   Maximum Transmission Unit.

multimode fiber   Cable with a glass diameter of either 62.5 μm or 50 μm. Multimode (MM) is less expensive than single-mode (SM) cable, but has less range because the larger diameter allows more reflections off the sides of the glass, producing a fuzzy optical signal over long distances. Single-mode cable is much smaller in diameter and there is less reflection, enabling an acceptable signal to travel farther. Multimode fiber works well up to 500 meters.

multi-hop SRDF   An SRDF multi-hop topology allows you to daisy-chain three Symmetrix sites where a third site is providing business continuance backup to the remote RDF site. In a multi-hop scheme, the last site in the chain is two hops (SRDF links) away from the local or primary site.

multisession consistency (MSC) mode   Beginning with Enginuity version 5x71 for mainframe and open systems, SRDF/A is supported in configurations where there are multiple primary Symmetrix systems and/or multiple primary Symmetrix SRDF groups connected to multiple secondary Symmetrix systems and/or multiple secondary Symmetrix SRDF groups. This is referred to as SRDF/A Multi-Session Consistency or SRDF/A MSC. SRDF/A MSC configurations can also support mixed open systems and mainframe data controlled within the same SRDF/A MSC session.

MVS   Multiple Virtual Storage (mainframe systems only).

N

network attached storage
A NAS device is attached to a TCP/IP-based network (LAN or WAN), and accessed using CIFS and NFS-specialized I/O protocols for file access and file sharing.

network topology
A physical arrangement of nodes and interconnecting communication links in networks based on application requirements and geographical distribution of users.

O

offline
Pertaining to equipment or devices not under control of the processor.

offline SRDF link
An SRDF link is offline if one or more of the following occurs: The remote adapter is offline - link disabled. The remote adapter is online but the link is offline (damaged or disconnected cable or other damaged hardware) - link disabled.

online
Pertaining to equipment or devices under control of the processor.

online SRDF link
An SRDF link is online when the following occurs: The remote adapter is operational and enabled on both sides of the SRDF configuration. The Symmetrix systems are configured properly on both sides of the SRDF configuration. The external link infrastructure components are operational.

open system
A system whose characteristics comply with standards made available throughout the industry, and therefore can be connected to other systems that comply with the same standards.

P

peer-to-peer remote copy (PPRC)
A hardware-based remote copy option that provides a synchronous volume copy across storage subsystems for disaster recovery, device migration, and workload migration.

permanent link loss
If SRDF/A experiences a permanent link loss, it drops all of the devices on the link to not ready state. This results in all data in the active and inactive primary Symmetrix cycles (capture and transmit delta sets) being changed from write pending for the remote mirror to owed to the remote mirror. In addition, any new write I/Os on the primary Symmetrix system result in tracks being marked owed to the

remote mirror. All of these tracks are owed to the secondary Symmetrix once the links are restored.

PoS   Packets over SONET, a protocol that places the IP layer directly above the SONET layer, and while offering quality of service, eliminates the overhead of using the ATM protocol as a transport between the SONET and IP packet.

primary volume (R1)   Also known as *source* volume. A Symmetrix logical volume that is participating in SRDF operations. It resides in the "local" Symmetrix storage subsystem. All CPUs attached to the Symmetrix may access a primary volume for read/write operations. All writes to this volume are mirrored to a "remote" Symmetrix storage subsystem. A primary volume is not available for local mirroring operations.

PVC   Permanent Virtual Circuit, (also called virtual channel connection in ATM), established by administrative means, much like a leased or dedicated real circuit.

Q

QoS   Quality of Service (QoS) is a Symmetrix management area that provides the Administrator performance tuning tools to optimize the storage system's dynamic I/O handling paths. QoS features allow you to adjust data transfer rates and caching parameters to meet your current storage environment.

R

RA   ESCON remote adapter. Refer to .

RAID 5   RAID 5 data protection implements data striping and rotating parity across all hypervolumes of a RAID 5 device. RAID 5 (3+1) configurations consists of four Symmetrix devices with parity and data striped across each device. RAID 5 (7+1) configurations consists of eight Symmetrix devices with data and parity striped across each device.

RAID 1/0   A protection method that combines RAID 1 and RAID 0; used in open systems environments.

RAID 10   A protection method that combines RAID 1 and RAID 0; used in mainframe environments.

RAID 6	A protection method that supports the ability to rebuild data in the event that two drives within the RAID group fail. RAID 6(6+2) configurations consists of eight Symmetrix devices with parity and data striped across each device. RAID 5 (14+2) configurations consists of sixteen Symmetrix devices with data and parity striped across each device.
RE	GbE remote adapter. Refer to "Gigabit Ethernet director (RE) on Symmetrix 8000" on page 43.
read hit	The name of the action taken when data requested by the read operation is in global memory.
read miss	The name of the action taken when data requested by the read operation is not in global memory.
reconnect	The function that occurs when a target selects an initiator to continue an operation after a disconnect.
reconnection	A reconnection exists from the assertion of the BSY signal in a RESELECTION phase until the next BUS FREE phase occurs. A reconnection can only occur between a target and an initiator.
Reserve Capacity	SRDF/A Reserve Capacity enhances SRDF/A's ability to maintain an operational state when encountering network resource constraints that would have previously suspended SRDF/A operations. With SRDF/A Reserve Capacity functions enabled, additional resource allocation can be applied to address temporary workload peaks, periods of network congestion, or even transient network outages.
resynchronization	A track image copy from the primary volume to the secondary volume of only the tracks which have changed. Resynchronization is the process of copying data from primary to secondary or secondary to primary sites since the data was last synchronized.
RF	Fibre Channel remote adapter. Refer to "Fibre Channel director (RF)" on page 41.
rolling disaster	A rolling disaster is a series of events that lead up to a complete disaster. For example, the loss of a communication link occurs prior to a site failure. Most disasters are rolling disasters; their duration may be only milliseconds or up to hours.

S

SCSI   Small Computer System Interface.

secondary volume (R2)   Also known as *target* volume. A Symmetrix logical volume that is participating in SRDF operations. It resides in the "remote" Symmetrix storage subsystem. It is paired with a primary volume in the local Symmetrix storage subsystem and receives all write data from its mirrored pair. This volume is not accessed by user applications during normal I/O operations. A secondary volume is not available for local mirroring or dynamic sparing operations.

semisynchronous mode   An SRDF mode of operation that provides an asynchronous mode of operation. Applications are notified that an I/O (or I/O chain) is complete once the data is in the cache of the RA1 Symmetrix storage subsystem. Any new data is then written to cache in the RA2 Symmetrix storage subsystem. The RA2 Symmetrix storage subsystem acknowledges receipt of the data once it is secure in its cache. If source tracks are pending transfer to a secondary volume, and a second write is attempted to the primary, Symmetrix will disconnect (non-immediate retry request), and wait for the pending track to transfer to the RA2 Symmetrix storage subsystem.

Used mainly for an extended distance SRDF solution, semi-synchronous mode allows the primary and secondary volumes to be out of synchronization by one write I/O operation. Data must be successfully stored in the Symmetrix system containing the primary volume before an acknowledgement is sent to the local host. This mode is not supported for FICON or at Enginuity version 5772 and later.

single-mode fiber   Cable with a glass diameter of 9 µm. The smaller glass diameter (compared to MM) results in much less light reflection off the sides of the glass and, therefore, fewer wave fronts in the signal. This increases the effective distance an acceptable optical signal can travel.

Single-mode fiber (SM) connections between Symmetrix systems can be up to 10 km if single-mode ports on the Symmetrix system are used. Long runs between Fibre Channel switches are typically done with single-mode fiber.

SMC   Symmetrix Management Console.

SONET   Synchronous Optical Network Transport, a physical layer technology designed to provide a universal transmission and multiplexing

scheme, with transmission rates in the multiple-Gb/s range, and a sophisticated operations and management system. SONET is the transport layer between the network protocol and the physical fiber that provides the framing and add/drop multiplexing function. Throughput ranges from OC-1 at 51.84 Mb/s to OC-256 at 13 Gb/s. OC-192 is currently being deployed for long-haul traffic within carrier networks, connecting data centers and local offices.

source volume (R1)   Also known as *primary* volume. A Symmetrix logical volume that is participating in SRDF operations. It resides in the "local" Symmetrix storage subsystem. All CPUs attached to the Symmetrix may access a primary volume for read/write operations. All writes to this volume are mirrored to a "remote" Symmetrix storage subsystem. A primary volume is not available for local mirroring operations.

SRDF   Symmetrix Remote Data Facility. SRDF consists of the microcode and hardware required to support Symmetrix remote mirroring.

SRDF group   SRDF groups define relationships between Symmetrix systems. An SRDF group is a set of SRDF director port connections configured to communicate with another set of SRDF director ports in another Symmetrix system. Logical volumes (devices) are assigned to SRDF groups.

SRDF link   One end-to-end SRDF connection between a given pair of Symmetrix systems.

SRDF/Asynchronous (SRDF/A)   A mode of remote replication that allows customers to asynchronously replicate data while maintaining a dependent write consistent copy of the data on the secondary (R2) device sat all times. The dependent write consistent point-in time copy of the data at the remote site is typically only seconds behind the primary (R1) site. SRDF/A session data is transferred to the secondary Symmetrix system in cycles (or delta sets), eliminating the redundancy of multiple same-track changes being transferred over the SRDF links, potentially reducing the required bandwidth.

SRDF/Automated Replication (SRDF/AR)   An automation solution that uses both SRDF and TimeFinder to provide a periodic asynchronous replication of a restartable data image. A single-hop SRDF/AR configuration is used to permit controlled data loss (depending on the cycle time). For protection over greater distances, a multi-hop SRDF/AR configuration can provide long distance disaster restart with zero data loss at a middle or "bunker" site.

SRDF/Consistency Groups (SRDF/CG) An SRDF product offering designed to ensure the dependent-write consistency of data remotely mirrored by the SRDF operations in the event of a rolling disaster.

SRDF/Data Mobility (SRDF/DM) An SRDF product offering that permits operation in SRDF adaptive copy mode only and is designed for data replication and/or migration between two or more Symmetrix systems. SRDF/DM transfers data from primary volumes to secondary volumes, permitting information to be shared, content to be distributed, and access to be local to additional processing environments. Adaptive copy mode enables applications using that volume to avoid propagation delays while data is transferred to the remote site.

SRDF/DM supports all Symmetrix systems and all Enginuity levels that support SRDF, and can be used for local or remote transfers. *See* adaptive copy mode.

SRDF/Star SRDF/Star is a data protection and failure recovery solution that covers three geographically dispersed data centers in a triangular topology. The primary site, by way of SRDF links, uses concurrent or Cascading SRDF capabilities to mirror the same production data synchronously to one secondary site and asynchronously to another secondary site for a total business continuance solution.

SRDF/Synchronous (SRDF/S) A business continuance solution that maintains a real-time (synchronous) copy of data at the logical volume level.

stage The process of writing data from a disk drive to global memory.

standard device A standard device is a Symmetrix volume that *could be* participating in SRDF operations, but *is being* used for TimeFinder/Clone/Snap operations.

striping The process of segmenting logically sequential data and writing the segments to multiple physical drives.

synchronous mode Available with the SRDF/S product offering, synchronous mode maintains a real-time mirror image of data between the primary and secondary volumes. Data must be successfully stored in both the local and remote Symmetrix systems before an acknowledgement is sent to the primary site host.

T

T1 Point-to-point digital communications circuit (DS1) provided by North American carriers for voice and data transmission (1.544 Mb/s). T1 has similar latency characteristics to E1. See also "E1".

T3 Point-to-point digital communications circuit (DS3) provided by carriers for voice and data transmission (44.5 Mb/s in U.S. or 34 Mb/s in Europe-E3); can be divided into 24 separate 64 Kb channels. T3 has similar latency characteristics to E3. See also "E3".

target volume (R2) Also known as *secondary* volume. A Symmetrix logical volume that is participating in SRDF operations. It resides in the "remote" Symmetrix storage subsystem. It is paired with a primary volume in the local Symmetrix storage subsystem and receives all write data from its mirrored pair. This volume is not accessed by user applications during normal I/O operations. A secondary volume is not available for local mirroring or dynamic sparing operations.

TCP Transmission Control Protocol.

temporary link loss If SRDF/A suffers a temporary loss (<10 seconds by default) on all of the SRDF links, the SRDF/A state remains active and data continues to accumulate in global memory. This may result in an elongated cycle, but the secondary Symmetrix dependent-write consistency is not compromised and the primary and secondary Symmetrix device relationships are not suspended.

terabyte (TB) The TB_2 value is based on the convention:
1 TB = 1024 * 1024 * 1024 * 1024 bytes.
The TB_{10} value is based on the convention:
1 TB = 1000 * 1000 * 1000 * 1000 bytes.
Although these values are expressed differently, their values are equivalent.

thin device A host-accessible device that has no storage directly associated with it.

thin pool A collection of data devices that provide the storage capacity for thin devices.

TimeFinder EMC family of software providing replication solutions, leveraging Symmetrix architecture. The EMC TimeFinder family of local replication allows users to non-disruptively create and manage

point-in-time and dependent write consistent copies of data. This allows operational processes, such as backup, reporting, and application testing to be performed independent of the primary application to maximize service levels without impacting performance or availability.

Transmit Idle A Reserve Capacity enhancement to EMC's SRDF/A feature that provides the capability of dynamically and transparently extending the Capture, Transmit, and Receive phases of the SRDF/A cycle while masking the effects of an "all SRDF links lost" event. Without the SRDF/A Transmit Idle enhancement, an "all SRDF links lost" event would normally result in the abnormal termination of SRDF/A. The SRDF/A Transmit Idle enhancement has been specifically designed to prevent this event from occurring.

U

unidirectional SRDF link If all primary (R1) volumes reside in one Symmetrix system and all secondary (R2) volumes reside in another Symmetrix system, write operations move in one direction, from primary to secondary. In a unidirectional configuration, data moves in the same direction between all devices in the SRDF group.

V

volume A general term referring to a storage device. In the Symmetrix subsystem, a volume corresponds to a part of a physical disk drive that emulates a single disk drive to the host.

W

Write Folding Write Folding in SRDF/A environments improves the efficiency of the SRDF network links. Even if there are multiple data updates, (i.e. repeated writes) in the same cycle, the systems send the data across the SRDF links only once.

write hit There is room in global memory for the data presented by the write operation.

write miss There is no room in global memory for the data presented by the write operation.

Z

z/OS   A widely-used operating system for the IBM zSeries mainframe computers that uses 64-bit real storage.

EMC Symmetrix Remote Data Facility (SRDF) Connectivity Guide