



EMC Symmetrix Remote Data Facility (SRDF)

PRODUCT GUIDE

**P/N 300-001-165
REV A03**

EMC Corporation

Corporate Headquarters:
Hopkinton, MA 01748 -9103
1-508 -435 -1000
www.emc.com

Copyright © 2001, 2002, 2003, 2004, 2005 EMC Corporation. All rights reserved.

Printed January, 2005

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

THE INFORMATION IN THIS PUBLICATION IS PROVIDED “AS IS.” EMC CORPORATION MAKES NO REPRESENTATIONS OR WARRANTIES OF ANY KIND WITH RESPECT TO THE INFORMATION IN THIS PUBLICATION, AND SPECIFICALLY DISCLAIMS IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

Trademark Information

EMC², EMC, Symmetrix, Celerra, CLARiiON, CLARAlert, Documentum, HighRoad, Legato, Navisphere, PowerPath, ResourcePak, SnapView/IP, SRDF, TimeFinder, VisualSAN, and where information lives are registered trademarks and EMC Automated Networked Storage, EMC ControlCenter, EMC Developers Program, EMC OnCourse, EMC Proven, EMC Snap, Access Logix, AutoAdvice, Automated Resource Manager, AutoSwap, AVALONidm, C-Clip, Celerra Replicator, Centera, CentraStar, CLARevent, Connectrix, CopyCross, CopyPoint, DatabaseXtender, Direct Matrix, Direct Matrix Architecture, EDM, E-Lab, Enginuity, FarPoint, FLARE, GeoSpan, InfoMover, MirrorView, NetWin, OnAlert, OpenScale, Powerlink, PowerVolume, RepliCare, SafeLine, SAN Architect, SAN Copy, SAN Manager, SDMS, SnapSure, SnapView, StorageScope, SupportMate, SymmAPI, SymmEnabler, Symmetrix DMX, Universal Data Tone, and VisualSRM are trademarks of EMC Corporation. All other trademarks used herein are the property of their respective owners.

Preface	xi
----------------------	----

Chapter 1 Introducing SRDF

SRDF Product Overview	1-2
SRDF/Synchronous (SRDF/S)	1-3
SRDF/Asynchronous	1-4
SRDF/Data Mobility (SRDF/DM)	1-5
SRDF/Automated Replication (SRDF/AR)	1-5
SRDF/Star	1-5
SRDF/Consistency Groups (SRDF/CG)	1-6
SRDF/Cluster Enabler for MSCS or VCS	1-7
SRDF/S – SRDF/A Mode Change	1-7
SRDF Functional Overview	1-8
Basic SRDF Configuration	1-8
SRDF Interfamily Connectivity	1-9
SRDF Supported Data Protection Options	1-11
Monitoring and Controlling SRDF	1-11
SRDF Director Hardware	1-12
SRDF Director Functions	1-12
ESCON Director (RA)	1-13
Fibre Channel Director (RF)	1-13
Gigabit Ethernet Director (RE) on Symmetrix 8000	1-14
Multiprotocol Channel Director on the Symmetrix DMX Series	1-15
SRDF Configurations	1-16
SRDF Configuration Using Switched Fibre Channel	1-19
SRDF Fully Switched Fabric Connectivity	1-19
Switched and Concurrent SRDF	1-21

SRDF with Native GigE.....	1-22
SRDF and FarPoint.....	1-23

Chapter 2 **SRDF Technical Concepts**

SRDF Volume Types	2-2
Primary (Source) Volumes.....	2-2
Secondary (Target) Volumes	2-3
Local Volumes	2-3
Dynamic SRDF Devices.....	2-3
EMC Compatible Peer	2-5
SRDF Groups.....	2-6
Dynamic SRDF Groups	2-6
SRDF Volumes: Special Considerations	2-7
Dynamic Sparing with Remotely Mirrored Pairs (SRDF) ..	2-7
Open Systems Metavolumes.....	2-7
Symmetrix RAID 10 (Mirrored Striped Mainframe Volumes with DMSP).....	2-8
PPRC Command Support	2-9
SRDF Link Configurations	2-10
SRDF Unidirectional Link Configuration	2-10
SRDF Bidirectional Link Configuration	2-10
SRDF Dual-Directional Link Configuration	2-10
SRDF Link and Volume States	2-11
SRDF Link States	2-11
Configuration Settings Affecting Device Ready and Link States.....	2-11
Logical Volume States	2-12
SRDF Volumes States (Symmetrix View)	2-13
Symmetrix Volume States (Host View)	2-14
Host Accessibility	2-14
Primary Modes of Operation	2-16
Synchronous Mode.....	2-16
Semi-Synchronous Mode.....	2-17
Secondary Modes of Operation	2-19
Adaptive Copy Modes.....	2-19
Additional SRDF Modes and Attributes	2-21
Domino Modes	2-21
Invalid Tracks Attribute.....	2-22
SRDF System-Level Attributes	2-22
Concurrent SRDF	2-23
SRDF/Consistency Groups (SRDF/CG).....	2-25
How a Consistency Group Works.....	2-25
Continuous Processing	2-27

z/OS Technical Considerations	2-28
-------------------------------------	------

Chapter 3 **SRDF Operations**

Write Operations	3-2
Write Operations in a Unidirectional or Dual-Directional Configuration	3-2
Write Operations in an ESCON Bidirectional Configuration	3-2
Read Operations.....	3-3
Primary Volume Read Operations.....	3-3
Secondary Volume Read Operations	3-3
Recovery Operations	3-5
Failover to the Secondary Symmetrix System	3-5
Failback to the Primary Symmetrix System.....	3-5
Recovery for a Large Number of Invalid Tracks	3-6
Business Continuance Using SRDF	3-7
Business Continuance Using SRDF and TimeFinder	3-9
Using TimeFinder/Mirror BCVs with Primary Devices....	3-9
Using a BCV as a Primary (Source) Device.....	3-9
Using TimeFinder BCVs with Secondary Devices	3-10
Using a BCV as a Secondary (Target) Device.....	3-10
SRDF Remote Command Support.....	3-11
SRDF Multi-Hop	3-11
R1/R2 Swap	3-13
R1/R2 Swap Procedure History	3-13
Dynamic R1/R2 Swap.....	3-14
Migrating Data From R1 to a Larger R2 Device	3-15

Chapter 4 **SRDF/Asynchronous Operations**

SRDF/Asynchronous Overview	4-2
SRDF/A Benefits.....	4-2
Requirements and Limitations.....	4-3
Ensuring Data Consistency: The Dependent Write Principle...	4-4
How Does SRDF/A Work?	4-6
Symmetrix Ordered Write Processing.....	4-6
Dependent Write Consistency	4-7
SRDF/A Data Flow Description.....	4-7
Ensuring Data Consistency: The Cycle Switch Process.....	4-10
SRDF Options and Features.....	4-12
Multiple SRDF/A Groups per Symmetrix.....	4-12
Concurrent SRDF/A and SRDF/S from a Single Source	

Volume	4-12
SRDF/Star	4-13
SRDF/A to SRDF/S Mode Change Feature	4-13
Dynamic SRDF with SRDF/A	4-13
SRDF/A States	4-14
SRDF/A Mode Transitions and Recovery	4-15
NR State (System Startup)	4-15
Inactive State	4-15
Switching to SRDF/Asynchronous Mode	4-15
SRDF/A to Adaptive Copy	4-16
Active State: System Working in Asynchronous Mode, and R2 is Consistent	4-17
Dropping Out of Active State	4-17
Deactivating SRDF/A	4-17
SRDF/A Recovery Scenarios	4-18
Temporary Link Loss	4-18
Permanent Link Loss	4-18
Primary Symmetrix Global Memory Full Condition	4-19
Failback from R2	4-19
SRDF/A Multi-Session Consistency	4-20
SRDF/A Consistency Groups' Multi-Session Consistency: Operational Overview	4-21
Entering SRDF/CG for SRDF/A Multi-Session Consistency	4-21
Performing a Consistent Cycle Switch	4-23
SRDF/A Multi-Session Recovery	4-25
R1 Side Recovery	4-25
R2 Side Recovery	4-25
SRDF/A Multi-Session Failure Scenarios	4-28
SRDF/A MSC Host Failure	4-28
Tolerance Mode	4-28
 Chapter 5	
SRDF/Star Operations	
SRDF/Star Overview	5-2
SRDF/Star Benefits	5-4
Known Requirements and Limitations at this Release	5-4
How SRDF/Star Works	5-5
SRDF/Star Automation	5-7
 Index	i-1

Figures

1-1	Basic SRDF Configuration	1-9
1-2	Two Production Sites and One Recovery Site	1-16
1-3	Data Vaulting Solution	1-17
1-4	Sites Containing Both Primary and Secondary Devices	1-18
1-5	Switched SRDF with Multiple Primary and Secondary Devices	1-19
1-6	Switched SRDF over Fibre Channel with Enginuity 5x67	1-20
1-7	Switched and Concurrent SRDF Configuration Example	1-21
1-8	Switched GigE Network Configuration Example	1-23
1-9	SRDF with and without FarPoint	1-24
2-1	Dynamic SRDF	2-4
2-2	PPRC and GDPS Support	2-9
2-3	SRDF Logical Volume State	2-12
2-4	Synchronous Mode	2-17
2-5	Semi-Synchronous Mode	2-18
2-6	Concurrent SRDF Configuration	2-23
2-7	Primary and Secondary Relationships	2-25
2-8	Failed Link between Primary 2 and Target 1	2-26
2-9	Primaries 1, 2, and 3 in a Consistency Group	2-27
2-10	Failed Link between Primary 2 and Target 1	2-27
3-1	SRDF Business Continuation	3-7
3-2	Primary-to-Secondary Resynchronization	3-8
3-3	Secondary-to-Primary Resynchronization	3-8
3-4	SRDF Single-Hop Configuration (BCV Functioning as a Primary SRDF Device)	3-10
3-5	SRDF Multi-Hop Configuration	3-12
3-6	R1/R2 Swap Concept	3-13
4-1	SRDF/Asynchronous	4-6
4-2	SRDF/A Locality of Reference Example	4-8
4-3	SRDF/A Allowed State Transitions	4-14
4-4	SRDF/A Multigroup Fan-In and Fan-Out Configurations	4-20
4-5	Multi-Session Consistency	4-22

4-6 Multi-Session R2 Side Recovery4-26

5-1 Concurrent SRDF Configured for SRDF/Star Support5-3

5-2 SRDF/Star Configuration Reference5-7

Tables

1-1	Symmetrix Model Numbers	1-3
1-2	SRDF Interfamily Connectivity Support	1-10
1-3	Data Protection Options for SRDF Disk Devices	1-11
2-1	Primary (R1) Volume Accessibility	2-15
2-2	Secondary (R2) Volume Accessibility	2-15

As part of its effort to continuously improve and enhance the performance and capabilities of the EMC product line, EMC periodically releases new versions of both the EMC Enginuity Operating Environment and Symmetrix Remote Data Facility (SRDF). Therefore, some functions described in this guide may not be supported by all versions of Enginuity currently in use. For the most up-to-date information on product features, refer to your product release notes. If an SRDF feature does not function properly or does not function as described in this guide, please contact the EMC Customer Support Center for assistance.

Audience

This guide provides an overview of SRDF and SRDF family of product offerings for users requiring a basic knowledge of SRDF concepts and operation. Detailed SRDF installation and configuration information is available in the documents listed in *Conventions Used in This Guide* on page xii.

Organization

This guide contains the following information:

Chapter 1, *Introducing SRDF*, introduces SRDF and explains concepts such as SRDF volume types, hardware, and sample configurations.

Chapter 2, *SRDF Technical Concepts*, describes SRDF links, modes of operation, system-level attributes, and consistency groups.

Chapter 3, *SRDF Operations*, explains basic SRDF operations: write, read, recovery, and business continuance.

Chapter 4, *SRDF/Asynchronous Operations*, provides a technical overview of SRDF/Asynchronous (SRDF/A).

Chapter 5, *SRDF/Star Operations*, provides a technical overview of SRDF/Star.

Related Documentation

For additional information on SRDF and all Symmetrix systems-related publications, contact your EMC Sales Representative or refer to the EMC Powerlink website at:

<http://powerlink.emc.com>

For information on configuring SRDF, refer to:

- ◆ *EMC Support Matrix* at <http://www.emc.com/horizontal/interoperability/matrices/EMCSupportMatrix.pdf>
- ◆ *EMC Networked Storage Topology Guide* at <http://powerlink.emc.com>

Follow these links: **Support, Document Library, Solutions Guides**

- ◆ *SRDF Product Guide* at <http://powerlink.emc.com>

Follow these links: **Support, Document Library, Software, SRDF**

- ◆ *Symmetrix (Model or Series) Product Guide* at <http://powerlink.emc.com>

Follow these links: **Support, Document Library, Symmetrix**

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 and caution notices.

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



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.

AVANT GARDE	Keystrokes
Palatino, bold	◆ Dialog box, button, icon, and menu items in text ◆ Selections you can make from the user interface, including buttons, icons, options, and field names
<i>Palatino, italic</i>	◆ New terms or unique word usage in text ◆ Command line arguments when used in text ◆ Book titles
<i>Courier, italic</i>	Arguments used in examples of command line syntax.
Courier	System prompts and displays and specific filenames or complete paths. For example: working root directory [/user/emc]: c:\Program Files\EMC\Symapi\db
Courier, bold	◆ User entry. For example: sympoll -p ◆ Options in command line syntax

Where to Get Help Obtain technical support by calling your local sales office.

For service, call:

United States: (800) 782-4362 (SVC-4EMC)

Canada: (800) 543-4782 (543-4SVC)

Worldwide: (508) 497-7901

and ask for Customer Support.

**Sales and Customer
Support Contacts**

For the list of EMC sales locations, please access the EMC home page at:

<http://www.emc.com/contact/>

For additional information on the EMC products and services available to customers and partners, refer to the EMC Powerlink website at:

<http://powerlink.emc.com>

Your Comments

Your suggestions help us improve the accuracy, organization, and overall quality of the user publications. Please send a message to **techpub_comments@emc.com** with your opinions of this guide.

Introducing SRDF

This chapter introduces SRDF and explains basic concepts including SRDF volume types, director hardware, and example configurations.

This chapter contains the following topics:

- ◆ SRDF Product Overview 1-2
- ◆ SRDF Functional Overview 1-8
- ◆ SRDF Director Hardware 1-12
- ◆ SRDF Configurations 1-16

SRDF Product Overview

The Symmetrix Remote Data Facility (SRDF®) family of replication software offers various levels of Symmetrix based business continuance and disaster recovery solutions. The SRDF products offer the capability to maintain multiple, host-independent, mirrored copies of data. The Symmetrix® systems can be in the same room, in different buildings within the same campus, or hundreds to thousands of kilometers apart.

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

The SRDF family consists of three base solutions:

- ◆ SRDF/Synchronous (SRDF/S, formerly SRDF) – High-performance, host-independent, real-time synchronous remote replication from one Symmetrix to one or more Symmetrix systems.
- ◆ SRDF/Asynchronous (SRDF/A) – High-performance extended distance asynchronous replication using a Delta Set architecture for reduced bandwidth requirements and no host performance impact.
- ◆ SRDF/Data Mobility (SRDF/DM) – Rapidly transfer data from source volumes to remote volumes anywhere in the world, permitting information to be shared, content to be distributed, and information to be consolidated for centralized parallel processing activities.

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 multi-site failover with continuous protection

- ◆ SRDF/Consistency Groups (SRDF/CG) for data consistency
- ◆ SRDF/Cluster Enabler for integration with host-based clustering products such as MSCS and VCS

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/CG, and SRDF/DM. Most of the document focuses on SRDF/S, but generally it applies to any of the variations.

The SRDF family of products can also be used with the EMC® TimeFinder® family of products, which includes TimeFinder/Mirror, TimeFinder/Clone, and TimeFinder/Snap. For simplicity, this document uses the term TimeFinder to represent all EMC TimeFinder related products.

This product guide refers to the Symmetrix models shown in Table 1-1.

Table 1-1 Symmetrix Model Numbers

Cabinet Description	Symmetrix 3xxx/5xxx Series	Symmetrix 8000 Series	Symmetrix DMX Series
1/2 Bay	3330/5330 and 3630/5630	8230	DMX800 (rackmount)
1 Bay	3430/5430, 3830/3832 ^a /5830	8430, 8530	DMX1000
2 Bay			DMX2000
3 Bay	3700/5700 and 3930/5930	8730, 8830	DMX3000

a. The Symmetrix 3832 is a 3830 model with a split backplane.

SRDF/Synchronous (SRDF/S)

Symmetrix Remote Data Facility/Synchronous (SRDF/S) is a business continuance solution that maintains a real-time (synchronous) copy of data at the logical volume level in Symmetrix 3xxx, 5xxx, 8xxx, or Symmetrix DMX™ systems in the same or separate locations.

Symmetrix 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
- ◆ Increased integrity of recovery procedures

- ◆ Reduced backup and recovery costs
- ◆ Reduced disaster recovery complexity, planning, testing, and so forth

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
- ◆ Adaptive copy write-pending mode
- ◆ Adaptive copy disk mode

For more information on these SRDF/S modes of operation, refer to *Primary Modes of Operation* on page 2-16 and *Secondary Modes of Operation* on page 2-19.

SRDF/Asynchronous

Beginning with Enginuity™ Version 5670, the Symmetrix DMX system supports the asynchronous replication product named SRDF/Asynchronous (SRDF/A). SRDF/A is another 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 at all times. The point-in-time copy of the data at the remote side is only slightly behind the primary (R1) side. SRDF/A session data is transferred to the secondary Symmetrix system in predefined timed cycles or *delta sets*, eliminating the redundancy of multiple same track changes being transferred over the link, potentially reducing the required bandwidth.

SRDF/A provides a long-distance replication solution with minimal impact on performance. This level of protection is intended for customers who require minimal host application impact while maintaining a consistent, restartable image of their data on the R2 side. In the event of a disaster at the R1 site or if SRDF links are lost during data transfer, a partial delta set of data can be discarded, preserving consistency on the R2 with a data loss of no more than two SRDF/A cycles.

For more detailed information on SRDF/A, refer to Chapter 4, *SRDF/Asynchronous Operations*.

SRDF/Data Mobility (SRDF/DM)

SRDF/Data Mobility (SRDF/DM) is 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 quickly transfers data from source volumes to remote 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.

For more information on adaptive copy mode operation, refer to *Adaptive Copy Modes* on page 2-19.

SRDF/Automated Replication (SRDF/AR)

EMC SRDF/ Automated Replication (SRDF/AR) is an automation solution that uses both SRDF and TimeFinder to provide a periodic asynchronous replication of a restartable data image. You can use a single-hop SRDF/AR configuration that permits controlled data loss (depending on the cycle time). However, if greater protection is required, a multi-hop SRDF/AR configuration can provide long distance disaster restart with zero data loss.

For more information on single-hop and multi-hop configurations, refer to *Business Continance Using SRDF and TimeFinder* on page 3-9.

Compared to traditional disaster recovery solutions with their long recovery time and high data loss, disaster restart solutions using SRDF/AR provide remote restart with a short restart time and low data loss.

SRDF/AR offers customers data protection with dependent write consistency across a distance. This protection is accomplished by using geographically separated replicas with hardware and software products from EMC Corporation. The SRDF/AR process can be implemented with TimeFinder/Mirror in a mainframe z/OS environment, or through UNIX and Windows environments using Solutions Enabler's `symreplicate` command line interface.

SRDF/Star

Available at Enginuity level 5671, SRDF/Star provides advanced multi-site business continuity protection available for mainframe environments. It enables concurrent SRDF/S and SRDF/A

operations from the same source volumes with the ability to incrementally establish an SRDF/A session between the two remote sites in the event of a primary site outage—a capability only available through SRDF/Star software.

This capability takes the promise of concurrent synchronous and asynchronous operations (from the same source device) to its logical conclusion. SRDF/Star allows you to quickly re-establish protection between the two remote sites in the event of a primary site failure, and then just as quickly restore the primary site when conditions permit.

With SRDF/Star, enterprises can quickly resynchronize the SRDF/S and SRDF/A copies by replicating only the differences between the sessions—allowing for much faster resumption of protected services after a source site failure.

SRDF/Consistency Groups (SRDF/CG)

EMC SRDF/Consistency Groups (SRDF/CG) is an SRDF product offering designed to ensure the consistency of the data remotely mirrored by the SRDF operations in the event of a *rolling disaster*. Most applications, and in particular database management systems (DBMSs), have *dependent write* logic imbedded in them to ensure data integrity if a failure occurs in:

- ◆ The host processor
- ◆ The software
- ◆ The storage subsystem

A dependent write is a write that will not be issued by an application until a prior, related write I/O operation has completed. An example of dependent write is a database update.

For example, when updating a database, a DBMS takes the following steps:

1. The DBMS writes to the disk containing the transaction log.
2. The DBMS writes the data to the actual database dataset.
3. The DBMS writes again to the log volume to indicate that the database update was made.

In a remote disk copy environment, data consistency cannot be ensured if one of these writes was remotely mirrored, but its predecessor was not remotely mirrored. This could occur, for example, in a rolling disaster where a communication loss occurs that affects only a subset of the disk controllers that are performing the remote copy function.

SRDF/CG prevents a rolling disaster from affecting the integrity of the data at the remote site. When SRDF/CG detects any write I/O to a volume that cannot communicate with its remote mirror, SRDF/CG suspends the remote mirroring for all volumes defined to the consistency group before completing the intercepted I/O and returning control to the application. In this way, SRDF/CG prevents dependent I/O from reaching its remote mirror in the case where a predecessor I/O only gets as far as the local mirror.

For more information on consistency group operations, refer to *SRDF/Consistency Groups (SRDF/CG)* on page 2-25.

SRDF/Cluster Enabler for MSCS or VCS

SRDF/Cluster Enabler for MSCS or VCS provides high availability and automated failover through storage-based replication and server clustering through SRDF/S and MSCS or VCS.

SRDF/S – SRDF/A Mode Change

SRDF/S – SRDF/A Mode Change is a feature available at Engenuity level 5671 that enables the user to switch between asynchronous mode and synchronous mode while maintaining dependent write consistency on the remote side during the transition period. This only applies to a single SRDF/A group. An example of its use is of day processing where transactions are so many that the application might not tolerate the latency involved with handling the transaction load while in synchronous mode. The user might then revert to SRDF/A and in so doing keep point-in-time dependent write consistency on the remote side without affecting the end user response times. Once the load situation is passed the user can once again transition back to SRDF/S mode.

For more information on SRDF/S – SRDF/A Mode Change, refer to the EMC *Solutions Enabler Symmetrix SRDF CLI Product Guide* and the *Symmetrix SRDF Host Component for z/OS Product Guide*

SRDF Functional Overview

Symmetrix systems use local mirroring, or RAID 1, as one method of protecting data by maintaining data on both a production volume and a mirror volume within the same storage unit. SRDF uses a method of data protection known as remote mirroring. Remote mirroring is similar to local mirroring, except that the production volume resides in one storage unit while its remote mirrors (up to two remote mirrors are supported with Concurrent SRDF) reside in a different storage unit.

When your main storage systems are down for a planned or unplanned event, SRDF enables fast switchover from the primary (source) data to the secondary (target) copy. The local SRDF device, known as the *primary* (source, R1) device, is configured in a pairing relationship with a *secondary* (target, R2) device, forming an SRDF pair. While the R2 device is mirrored with the R1 device, the R2 device is either write disabled or *not ready* to its host. (*Not ready* means disabled for both reads and writes.) After the R2 device becomes synchronized with its R1 device, you can split the R2 device from the R1 device at any time, making the R2 device fully accessible again to its host. After the split, the secondary (target, R2) device contains valid data and is available for performing business continuance tasks through its original device address or restoring (copying) data to the primary (source, R1) device if there is a loss of data on that device.

For more information on SRDF operations, refer to Chapter 3.

Basic SRDF Configuration

A basic SRDF configuration consisting of a production site and a recovery site is illustrated in Figure 1-1. At the production site, a local host connects to Symmetrix A. The device containing the production data to be remotely mirrored is called the primary (source) or R1 volume. At the recovery site, a second (optional) host connects to Symmetrix B with the secondary (target) or R2 volume containing the remotely mirrored data. The Symmetrix systems communicate through SRDF links.

Additional SRDF configurations are explained in *SRDF Configurations* on page 1-16.

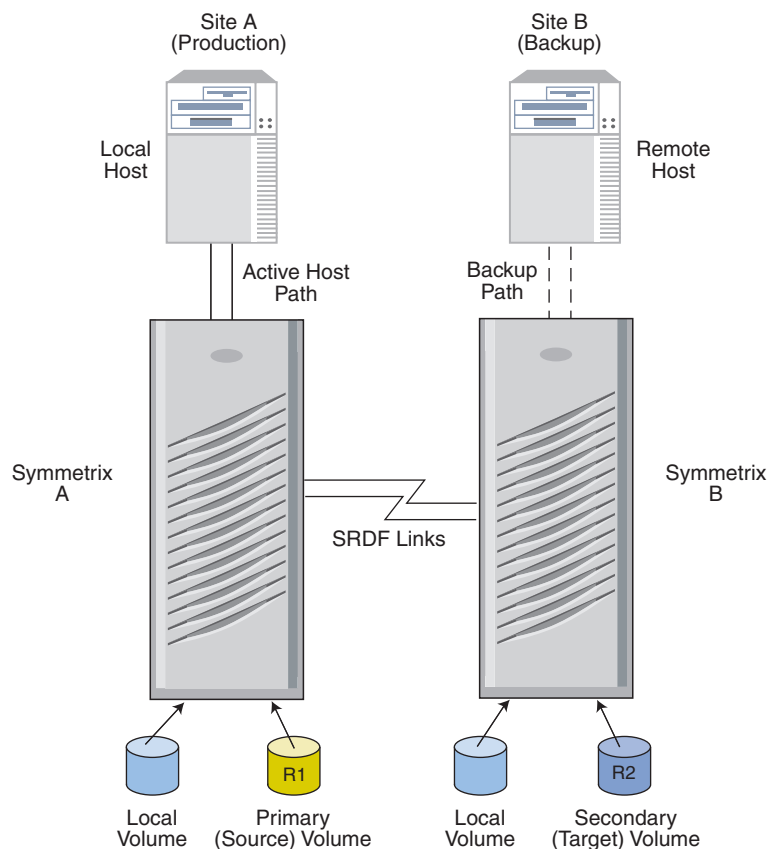


Figure 1-1 Basic SRDF Configuration

SRDF Interfamily Connectivity

SRDF supports connectivity between Symmetrix models running various levels of Enginuity code. Enginuity code 5x67 is the first Symmetrix family of code that supports connectivity with subsequent families of code (*interfamily*) to allow for nondisruptive code upgrades in SRDF environments. This feature's usefulness is realized with all Enginuity releases from level 5x68 and beyond.

Interfamily SRDF enables you to upgrade a network of Symmetrix systems over a period of weeks or months while maintaining SRDF protection. This connectivity support is described in Table 1-2 on page 1-10.

Table 1-2 SRDF Interfamily Connectivity Support

Symmetrix Models	Symmetrix 3x30/5x30 Running 5267	Symmetrix 8000 Running 5567	Symmetrix 8000 Running 5568	Symmetrix DMX Running 5669	Symmetrix DMX Running 5670	Symmetrix DMX Running 5671
Symmetrix 3x30/5x30 Running 5267	Yes	Yes	Yes	No	Yes ^{a b}	Yes ^{a b}
Symmetrix 8000 Running 5567	Yes	Yes	Yes	No	Yes ^{a b}	Yes
Symmetrix 8000 Running 5568	Yes	Yes	Yes	Yes	Yes	Yes
Symmetrix DMX Running 5669	No	No	Yes	Yes	Yes	No
Symmetrix DMX Running 5670	Yes ^{a b}	Yes ^{a b}	Yes	Yes	Yes ^c	Yes ^d
Symmetrix DMX Running 5671	Yes ^{a b}	Yes ^{a b}	Yes	No	No	Yes

- a. Used for migration to higher levels of Enginuity code only.
- b. Only ESCON and Fibre Channel connections (no FarPoint) are supported.
- c. SRDF/A is only supported with Enginuity level 5670 and higher.
- d. SRDF/Star Mainframe is only supported with Enginuity level 5671 and higher.

SRDF Supported Data Protection Options

The Symmetrix data protection options described in Table 1-3 can be purchased separately and implemented into the Symmetrix SRDF operation.

Table 1-3 Data Protection Options for SRDF Disk Devices ^a

Data Protection Option	Description
Mirroring (RAID Level 1)	Provides the highest level of performance and availability for all mission-critical and business-critical applications by maintaining a duplicate copy of a volume on two disk devices.
Parity RAID	Provides: - High performance dependent upon volume layout and external striping - High availability - data from lost volume is regenerated from remaining members A Parity RAID (3+1) group consists of three data volumes to one parity volume. A Parity RAID (7+1) group consists of seven data volumes to one parity volume.
RAID 5	Provides: - High performance with automatic striping across hypervolumes - High availability - lost hypervolumes regenerated from remaining members RAID 5 is configured in (3+1) and (7+1) groups. RAID 5 technology stripes data and distributes parity blocks across all the disk drives in the disk array.
Permanent Member Sparing	Replaces a faulty drive automatically from a list of available spares residing in the Symmetrix system without CE involvement on site.
Dynamic Sparing	Increases data availability by copying the data on a failing volume to a spare volume until the original device is replaced. Dynamic Sparing is used as additional protection for mirrored Parity RAID and SRDF volumes.
Symmetrix RAID 10	RAID 10 (one-zero) is a mirroring feature with striping used for mainframe environments.

a. When configuring multiple data protection options, consult your EMC Sales Representative for configuration rules.

Monitoring and Controlling SRDF

An EMC representative installs and initially configures SRDF at your site using the Symmetrix service processor.

After SDRF is up and running, you can monitor and control its operation by purchasing the appropriate EMC ControlCenter® software, Solutions Enabler for UNIX and Windows environments, or the SRDF Host Component for z/OS. For more information on SRDF related software, refer to the EMC Powerlink website at: <http://powerlink.emc.com> or contact your EMC Sales Representative.

SRDF Director Hardware

SRDF director hardware consists of a director and adapter board set. This hardware provides the communications transport 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:

- ◆ Multiprotocol Channel Director (MPCD), supported with Symmetrix DMX series, and configurable for SRDF over Gigabit Ethernet (GigE). (The MPCD is also configurable for iSCSi, FICON, or combinations.)
- ◆ GigE remote adapter (RE), supported with Symmetrix 8000 series.
- ◆ Fibre Channel remote adapter (RF).
- ◆ ESCON remote adapter (RA).

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.

Some restrictions apply in mixed environments with AS/400 and other host types. For more information, consult EMC Customer Support.



CAUTION

A single board configured in a Symmetrix system is a potential single point of failure and is not recommended in protection environments. Such a failure could occur if only one processor on the director is used for SRDF and the other processor(s) is/are used to connect to a host. The Symmetrix system should contain at least two director boards for an SRDF configuration; use 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.

SRDF Director Functions

The SRDF director/adapter board sets described above provide the link connections, fiber-optic protocol support, and communications control between two Symmetrix systems in an SRDF configuration.

GigE 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 DMX1000, DMX2000, and DMX3000 models, an optional four-processor ESCON board can be used for SRDF.

ESCON Director (RA)

A Symmetrix 8000 RA contains four individual processors (CPUs), and 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.*

At Enginuity Version 5670, the Symmetrix DMX ESCON director contains four processors (CPUs) and eight interfaces to host mainframe systems or 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.

Fibre Channel Director (RF)

Enginuity Version 5x66 and later support SRDF Fibre Channel emulation. Symmetrix RF directors include:

- ◆ **Two-Port/Two-Processor RF**—(Enginuity versions 5x66 and 5x67) Each processor on the two-port RF provides one host connection. One processor on a board can be configured with host

interface emulation while the other processor is configured with SRDF emulation. Alternatively, both processors can be configured for SRDF operation. Fibre Channel ports are either 850 nm multimode or 1310 nm single-mode.

- ◆ **Four-Port/Two-Processor RF**—Symmetrix 8000 series systems (Engenuity Version 5568) 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.

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) at 2 Gb/s. 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.

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

Gigabit Ethernet Director (RE) on Symmetrix 8000

The Symmetrix 8000 series supports a two-processor, two-port Gigabit Ethernet (GigE) card, with these features:

- ◆ Native SRDF IP support
- ◆ Model DP3-GBENET — Two GigE ports per card
- ◆ All 850 nm multimode ports
- ◆ Multiple TCP connections per GigE port
- ◆ 256 KB buffers for extended distance (TCP windowing)

The GigE director requires Symmetrix Engenuity level 5568.54.20 or higher.

Both ports on the GigE card can be used only as SRDF GigE ports; no other emulations are supported.

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 GigE (MPCD-GigE), iSCSI, or FICON.

The MPCD requires Symmetrix Engenuity level 5670 or higher.

On a two-port director, one port can be configured with host interface emulation (iSCSI or FICON) while the other port is configured for SRDF emulation. Alternatively, both ports can be configured for SRDF GigE operation.

On a four-port director, the ports can also be intermixed with SRDF (GigE) and host (iSCSI or FICON) emulations.

MPCD features:

- ◆ Native SRDF IP support
- ◆ Two-port and four-port models: combination FICON/GigE or iSCSI/GigE

All GigE ports are 850 nm multimode.

- ◆ Multiple TCP connections per GigE port
- ◆ On-board hardware compression:
 - Can be configured per TCP connection
 - Will auto-disable for Symmetrix 8000 series GigE connections
 - Compression is done on the SRDF level, not the TCP level

SRDF Configurations

This section provides examples of typical SRDF configurations.

In Figure 1-2, two production sites (A and C) send data across SRDF links to one recovery site (B).

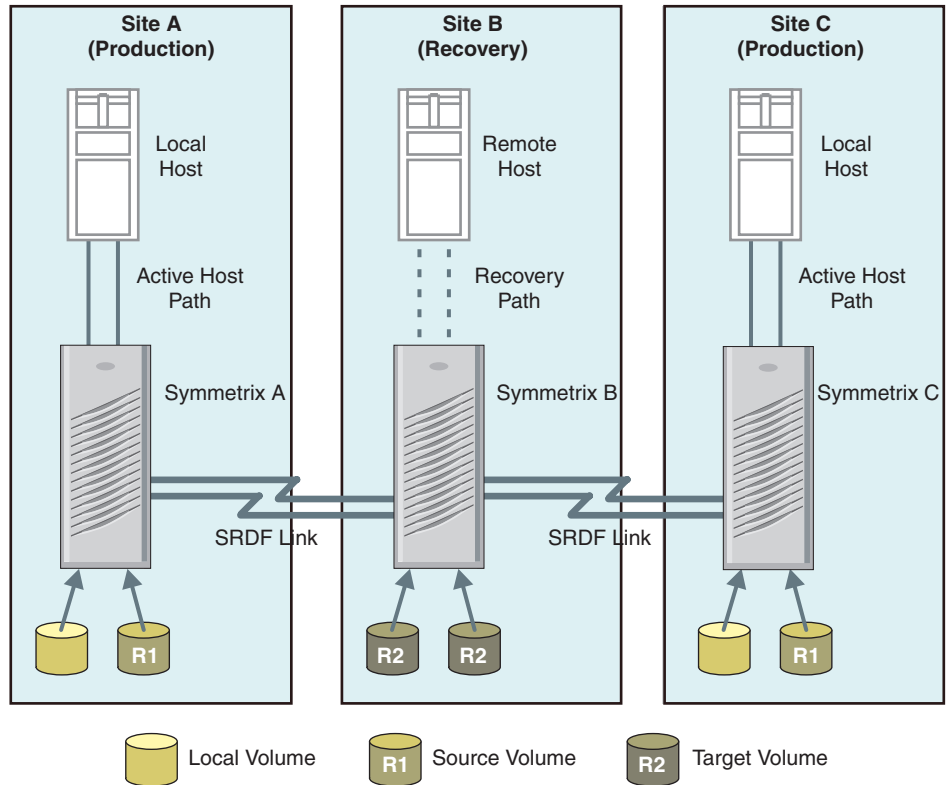


Figure 1-2 Two Production Sites and One Recovery Site

In Figure 1-3, one recovery site (G) provides a data vaulting solution for six production sites (A through F).

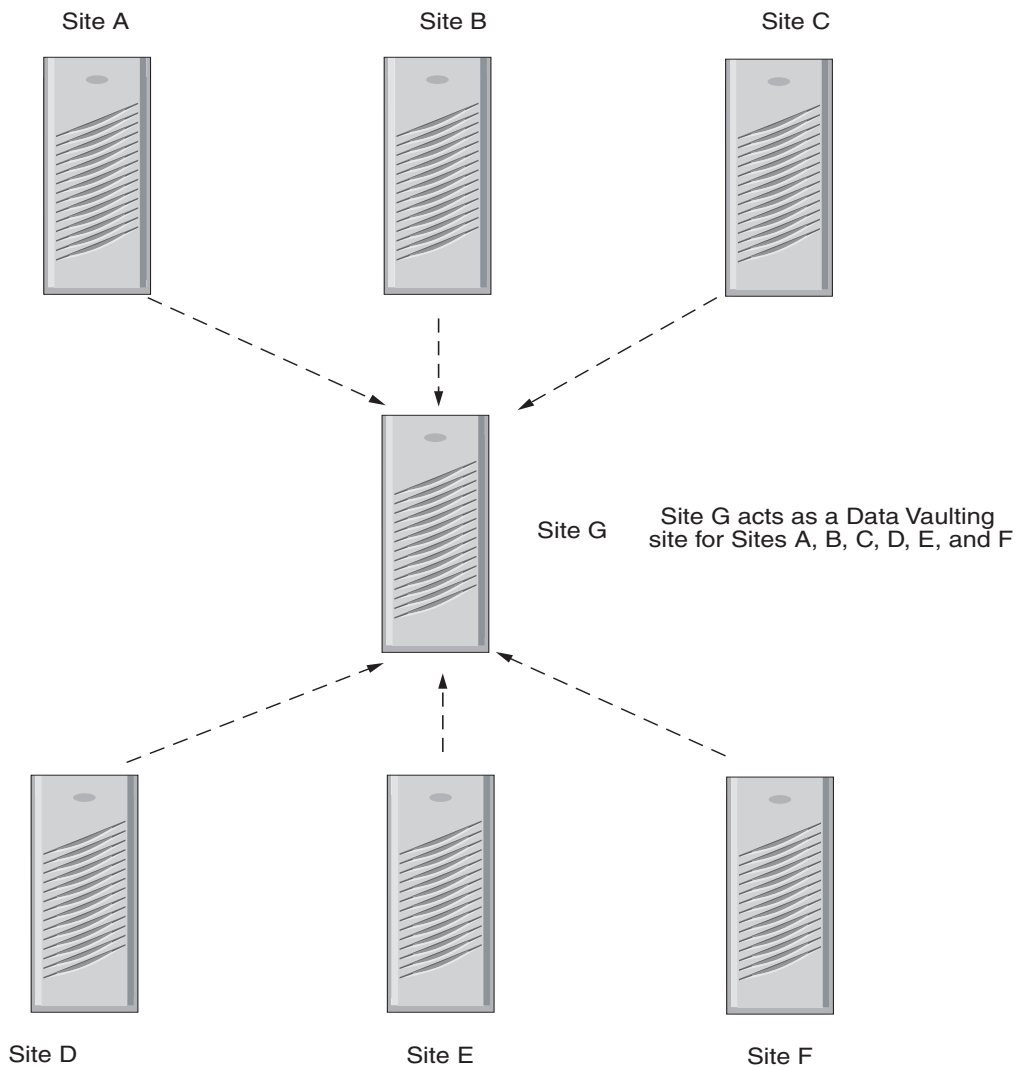


Figure 1-3 Data Vaulting Solution

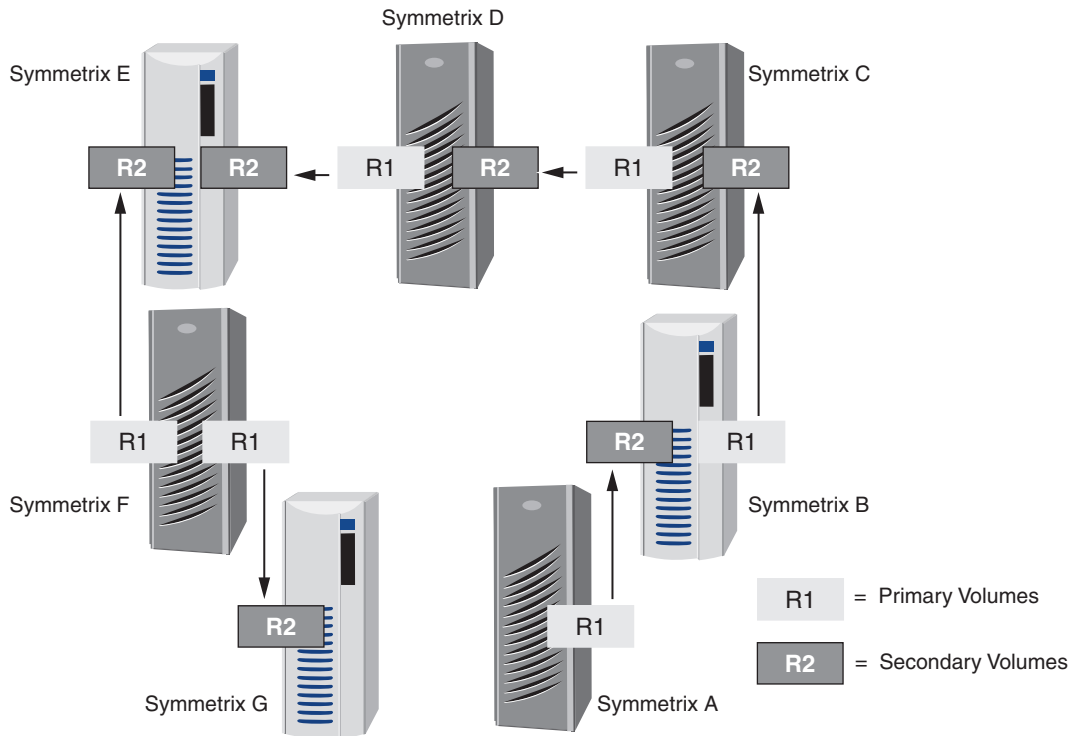


Figure 1-4 Sites Containing Both Primary and Secondary Devices

SRDF Configuration Using Switched Fibre Channel

A more flexible *switched SRDF* configuration using Fibre Channel switches connected through E_Ports is shown in Figure 1-5. Note that in this configuration, multiple primary (source) R1 devices are remotely connected using a Storage Area Network (SAN) to multiple secondary (target) R2 devices.

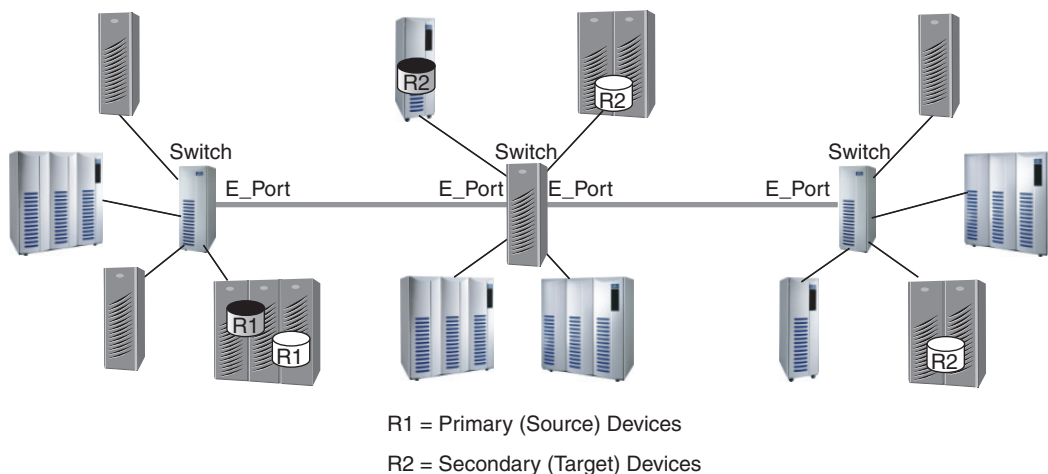


Figure 1-5 Switched SRDF with Multiple Primary and Secondary Devices

SRDF Fully Switched Fabric Connectivity

Enginuity code 5x67 and later supports fully switched open SRDF (fibre) connections. Switched SRDF allows all RFs in the Symmetrix systems connected to the fabric to communicate with each other, providing greater connectivity and configuration flexibility. The switched SRDF involves non-blocking switching devices that interconnect two or more nodes. Symmetrix systems in a switched SRDF topology can have each port pair running bi-directionally. Switched SRDF also enables fewer RFs to be present on the fabric, depending on system performance and redundancy requirements.

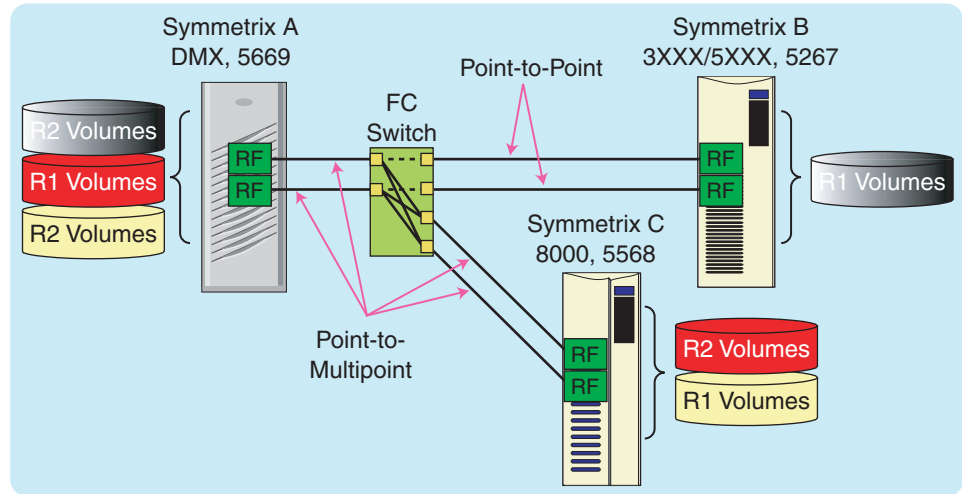


Figure 1-6 Switched SRDF over Fibre Channel with Engenuity 5x67

There are four logical paths from Symmetrix A to Symmetrix C (two logical paths on each RF director to each RF director in Symmetrix C). When Symmetrix A is communicating with Symmetrix B, there is a single logical path from each RF director to one other RF director. This is a point-to-point connection, the same as was provided with earlier Engenuity levels. Also note that all three Symmetrix systems have primary (source, or R1) devices. The ability to support point-to-multipoint connectivity is not related to primary (source) or secondary (target) device relationships.

Switched and Concurrent SRDF

As illustrated in Figure 1-7, Symmetrix A is communicating with Symmetrix B and Symmetrix C simultaneously through two RF directors, as well as using Concurrent SRDF to mirror its primary devices (R1s) to a second set of secondary devices (R2s) in Symmetrix D. Each RF director in Symmetrix A is configured to support two SRDF groups. When Symmetrix A is communicating with Symmetrix C, either of its RF directors can communicate with either RF director in Symmetrix C.

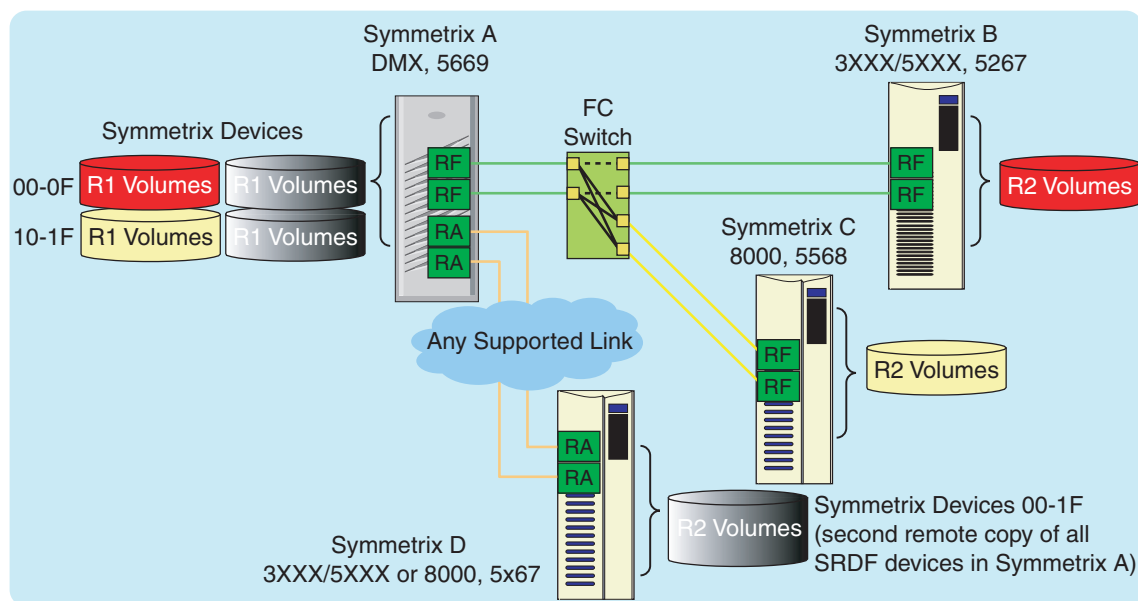


Figure 1-7 Switched and Concurrent SRDF Configuration Example

There are four logical paths from Symmetrix A to Symmetrix C (two logical paths on each RF director to each RF director in Symmetrix C). When Symmetrix A is communicating with Symmetrix B, there is a single logical path from each RF director to one other RF director. This is a point-to-point connection, the same as was provided with earlier Engenuity levels.

The ability to support point-to-multipoint connectivity is not related to primary (source) or secondary (target) device relationships. A separate SRDF group in Symmetrix A is configured with FarPoint™ on ESCON RAs for Concurrent SRDF to Symmetrix D, possibly using adaptive copy – disk mode and/or FarPoint.

For Concurrent SRDF, the primary SRDF mode is a synchronous or semi-synchronous mode of Symmetrix A's primary devices (R1s); however, the secondary mode of adaptive copy can be enabled for either one or both secondary R2s. In the example shown in Figure 1-7, the RA links would be extended via channel extenders over a WAN using adaptive copy mode and/or FarPoint for asynchronous replication to Symmetrix D.

As of Symmetrix Enginuity level 5671, SRDF/A is supported with Concurrent SRDF.

With Enginuity Version 5567 and the Symmetrix 8000 systems, the same local primary volume can be remotely mirrored in two different locations. This functionality is called *Concurrent SRDF*. The Symmetrix 3XXX/5XXX systems can be configured with a secondary volume that is one of two remote copies from a Symmetrix 8000 primary volume; however, the Symmetrix 3XXX/5XXX systems cannot be configured to provide two remote copies of the same primary volume.

For more information on Concurrent SRDF, refer to *Concurrent SRDF* on page 2-23.

SRDF with Native GigE

Native IP support for any SRDF based product on Symmetrix systems is based on GigE technology that enables direct Symmetrix system-to-IP-network attachment. This increases the options for Symmetrix system to Symmetrix system connectivity, and allows a Symmetrix system to connect to existing Ethernet infrastructure and directly access high-speed data transmission conduits via IP.

Symmetrix DMX series systems provide native IP support through the Multiprotocol Channel Director (MPCD). Symmetrix 8000 series systems provide native IP support through the GigE remote director. The MPCD and GigE remote directors provide comparable functionality, with the exception of data compression, which is a feature of the MPCD only. Unless otherwise specified, the following information applies to both Symmetrix 8000 series and Symmetrix DMX GigE remote director configurations.

Figure 1-8 illustrates GigE connectivity into an IP network infrastructure. Each Symmetrix GigE source port will establish a logical TCP connection to each GigE target port. Hence, this configuration will have four logical TCP connections between the Symmetrix pair.

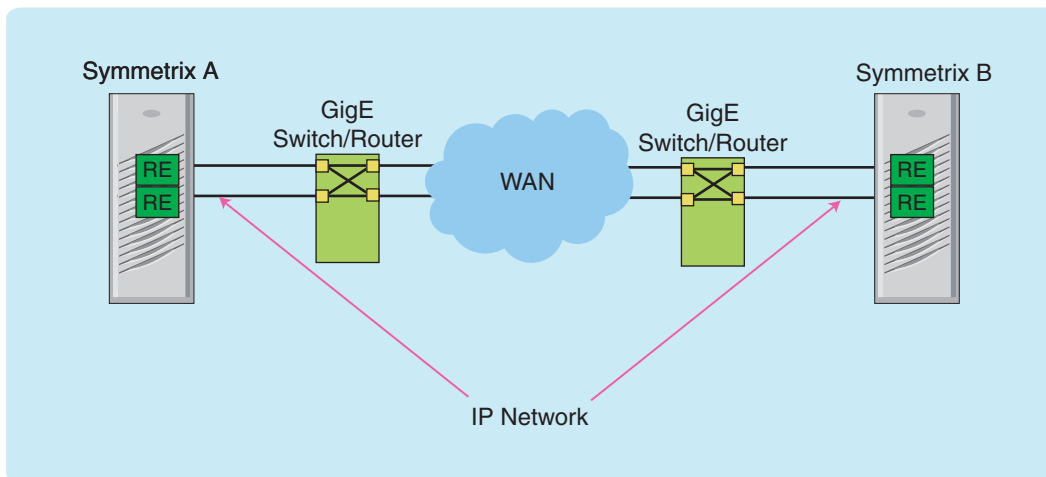


Figure 1-8 Switched GigE Network Configuration Example

SRDF and 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.

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 1-9 on page 1-24.

SRDF with FarPoint

With SRDF FarPoint, a single remote adapter can serially transmit more than one write I/O over the SRDF link. This method (referred to as *pipelining*) enables the SRDF communication link 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 Engenuity level

5x67 or later, multiple copy tasks can be placed on the SRDF link for a single logical volume.

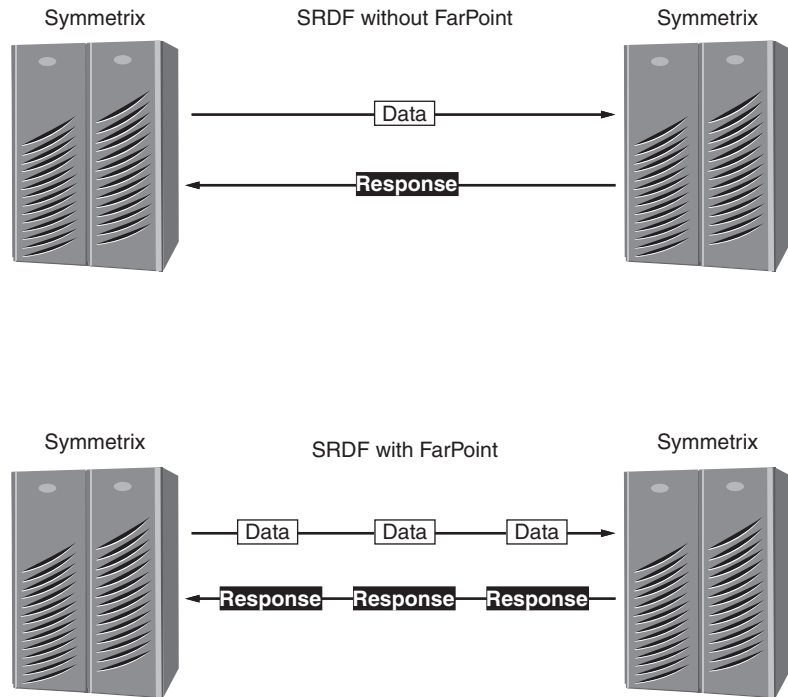


Figure 1-9 SRDF with and without FarPoint

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 performed on the remote machine. Without FarPoint, the Symmetrix system waits for one write operation to complete before sending the next one. 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.

**Performance Impact
of FarPoint**

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 link. Even on a short link, the FarPoint solution may provide some benefits.

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.

The introduction of PAV/MA does increase I/O concurrency at the volume level, potentially improving throughput in FarPoint configurations.

This chapter describes SRDF links, modes of operation, system-level attributes, and consistency groups.

This chapter contains the following topics:

- ◆ SRDF Volume Types2-2
- ◆ SRDF Groups2-6
- ◆ SRDF Volumes: Special Considerations2-7
- ◆ SRDF Link Configurations.....2-10
- ◆ SRDF Link and Volume States2-11
- ◆ Primary Modes of Operation2-16
- ◆ Secondary Modes of Operation2-19
- ◆ Additional SRDF Modes and Attributes2-21
- ◆ Concurrent SRDF2-23
- ◆ SRDF/Consistency Groups (SRDF/CG)2-25

SRDF Volume Types

SRDF refers to Symmetrix volumes as:

- ◆ Primary (source) volumes (R1)
- ◆ Secondary (target) volumes (R2)
- ◆ Local volumes
- ◆ Dynamic SRDF devices
- ◆ PPRC mode devices

In the context of this discussion, *volumes* are synonymous with *devices*.

You can assign additional primary (source) and secondary (target) volume pairs without halting SRDF operations. This eliminates the need to perform an initial Enginuity microcode program load (IMPL) when you add new data volumes.

Primary (Source) Volumes

Primary (source) volumes contain production data that is mirrored in a different Symmetrix system. Primary volumes are also referred to as *R1* volumes. Updates to a primary volume are automatically mirrored to a secondary (target) volume in the other Symmetrix system.

Primary volumes can be locally protected by:

- ◆ A dynamic spare (refer to *Dynamic Sparing with Remotely Mirrored Pairs (SRDF)* on page 2-7 for more information)
- ◆ RAID 1, conventional mirroring, (the primary volume is then referred to as a *mirrored pair*)
- ◆ RAID 10, (the primary volume is a striped mirrored device)
- ◆ RAID-S protection (the primary volume is a RAID-S data volume)
- ◆ Symmetrix Parity RAID (3+1) and Parity RAID (7+1) protection (the primary volume is a Parity RAID data volume)
- ◆ RAID 5 (3+1) and RAID 5 (7+1) protection (the primary volume is a RAID 5 data volume)

Additionally, a primary volume can be paired with a Business Continuance Volume (BCV) to provide an additional working copy of the data at the same location.

Secondary (Target) Volumes

Secondary (target) volumes contain a mirrored copy of data from a primary volume. Secondary volumes are also referred to as R2 volumes. As with primary volumes, secondary volumes can be locally protected by:

- ◆ A dynamic spare
- ◆ RAID 1, conventional mirroring, (the secondary volume is then referred to as a mirrored pair)
- ◆ RAID 10 (the secondary volume is a striped mirrored device)
- ◆ RAID-S protection (the secondary volume is actually a RAID-S data volume)
- ◆ Symmetrix Parity RAID (3+1) and Parity RAID (7+1) protection (the secondary volume is a Parity RAID data volume)
- ◆ RAID 5 (3+1) and RAID 5 (7+1) protection (the secondary volume is a Parity RAID data volume)

A secondary volume can also be paired with a BCV to provide an additional working copy of the data at the same location.

Local Volumes

Local volumes can reside on an SRDF enabled Symmetrix system, but they do not participate in SRDF activity. Local volumes are typically protected either by RAID 1, RAID 10, RAID-S, Parity RAID (3+1), Parity RAID (7+1), RAID 5 (3+1), RAID 5 (7+1), or a dynamic spare.

Dynamic SRDF Devices

Since Enginuity Version 5567, devices can be configured to be Dynamic SRDF capable devices. The only Dynamic SRDF functionality delivered in Enginuity Version 5567 is R1/R2 personality swap. Further, this capability is only allowed in non-FarPoint configurations. The Dynamic SRDF functionality, delivered in Enginuity Version 5568, enables you to create, delete, and swap SRDF pairs, using EMC host-based SRDF control software, while the Symmetrix system is in operation. This functionality allows R1 and R2 device personalities to be swapped within an SRDF pair. Using Dynamic SRDF technology, you can establish SRDF device pairs from non-SRDF devices, and then synchronize and manage them in the same way as configured SRDF pairs. Figure 2-1 shows Dynamic SRDF connections in a switched fabric environment.

At Enginuity level 5671, SRDF/A devices can also be configured as Dynamic SRDF capable devices.

Dynamic SRDF is supported over the following topologies:

- ◆ ESCON point-to-point connection (RA)
- ◆ Fibre Channel point-to-point connection (RF)
- ◆ Switched Fibre Channel fabric connection (RF)
- ◆ GigE connection (RE)

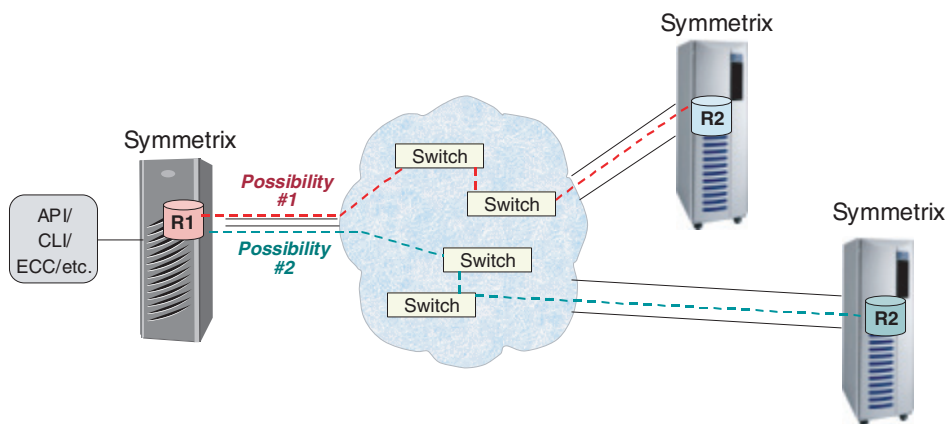


Figure 2-1 Dynamic SRDF

Dynamic SRDF provides the following three capabilities:

- ◆ Personality swap between primary and secondary volumes
- ◆ Terminate and reestablish a relationship with a new secondary volume
- ◆ Create a new primary/secondary pair relationship from non SRDF devices

The above functionality is not available for devices that are protected by Parity RAID.

This feature, when combined with dynamic SRDF groups, allows the user complete control over their SRDF configuration. Using EMC host-based software, the user can create or remove RA directors, SRDF groups, and SRDF devices. (Refer to *Dynamic SRDF Groups* on page 2-6.)

EMC Compatible Peer

Enginuity code 5568 (or higher) enables the Symmetrix system to support native IBM Peer-to-Peer Remote Copy (PPRC) commands through an Symmetrix feature called EMC Compatible Peer. Compatible Peer mode requires Dynamic SRDF capable devices and Enginuity level 5568 or higher. A dynamic SRDF device becomes a PPRC mode device upon receipt of a PPRC CESTPAIR command. From the time that the device enters PPRC mode, no host-based SRDF control software can operate on the device; it is truly a PPRC device and can only be controlled by PPRC commands received from the host. Control of these devices via the Symmetrix service processor or other *out-of-band* software is not supported.

Compatible Peer implements PPRC version 1, architectural level 2, beginning with Enginuity level 5568. Enginuity level 5671 adds support for PPRC version 1, architectural levels 3 and 4 (Hyper-swap support).

Advanced SRDF features such as Concurrent SRDF or SRDF/Asynchronous are not supported on PPRC devices. Parity RAID data protection is not supported on PPRC devices. All other Symmetrix data protection mechanisms are supported.

For more detailed PPRC implementations, refer to *GDPS: Installation and Testing Guide*.

SRDF Groups

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.

An SRDF group that is configured through Symmetrix configuration is called a static SRDF group. With Enginuity code up to and including 5x68, each Symmetrix system supported up to 16 static SRDF groups.

Enginuity code 5669 increases the maximum number of definable SRDF groups from 16 to 64. This enhancement allows customers more flexibility when working with the following SRDF features:

- ◆ Peer-to-Peer Remote Copy (PPRC) support
- ◆ Switched SRDF
- ◆ Concurrent SRDF
- ◆ Dynamic SRDF

Dynamic SRDF Groups

At Enginuity code 5669, a user can dynamically create empty SRDF groups and dynamically associate the groups with Fibre Channel or GigE SRDF directors. Removing dynamic SRDF groups is also possible. Both of these operations are accomplished using EMC host-based SRDF control software. Dynamic SRDF groups created through this mechanism are persistent through Symmetrix power on or IMPL.

PPRC groups require static SRDF groups. Dynamic groups are not supported for PPRC.

This feature, when combined with dynamic SRDF devices, allows the user complete control over their SRDF configuration. Using EMC host-based software, the user can create or remove RA directors, SRDF groups, and SRDF devices.

SRDF Volumes: Special Considerations

The following Symmetrix features require special considerations when used in an SRDF environment:

- ◆ Dynamic Sparing with Remotely Mirrored Pairs (SRDF)
- ◆ Open systems metavolumes
- ◆ Mainframe RAID 10 volumes
- ◆ PPRC command support

Dynamic Sparing with Remotely Mirrored Pairs (SRDF)

Typically, when a disk drive fails, it does not fail suddenly, it fails over a period of time during which it gives clues that it is failing. The Symmetrix system monitors the operation of all its disk drives and can detect when a disk is beginning to fail.

A Symmetrix system can be configured with physical disk devices known as *dynamic spares*, or *hot spares*. As the name suggests, the purpose of a dynamic spare is to take the place of a failed or failing disk device. If a disk drive begins to fail, the Symmetrix system automatically invokes the dynamic spare.

When the Dynamic Sparing option is invoked for a remotely mirrored SRDF pair, the Symmetrix system automatically activates an available spare in the Symmetrix system containing the failing device and copies data from the failing device to the spare. The system continues processing I/O requests with the spare functioning as one volume of a mirrored pair while the failing device and its remote mirror all operate without interruption. If the Symmetrix system cannot copy all data from the failing device to the spare, it will retrieve the unavailable data from the *good* member of the remote pair.

Open Systems Metavolumes

For open systems hosts attached to Symmetrix systems you can create a logical device that spans multiple Symmetrix logical volumes. This logical device is known as a *metavolume*.

A metavolume is a logical volume set created from individual Symmetrix logical volumes that can comprise one-to-multiple physical disks. Metavolumes are functionally the same as logical volume sets implemented with host volume manager software. Some metavolumes are created to define volumes larger than the current Symmetrix maximum hypervolume size of approximately 32 GB.

Physically, a metavolume is two or more Symmetrix hypervolumes presented to a host as a single addressable device. The metavolume consists of a *head device*, some number of *member devices* (optional), and a *tail device*.

Symmetrix metavolumes can contain up to 255 logical volumes and be up to 7.65 TB in size. Metavolumes can be composed of nonsequential and nonadjacent volumes. In an SRDF configuration, if you create metavolumes for the primary volumes, you must also create them for the secondary volumes. The primary and secondary meta members must be equal in number.

When configuring a metavolume, each metavolume device is counted as a single logical volume. Using metavolumes will reduce the maximum number of host-visible devices. That is, each member of the metavolume must be counted toward the maximum number of host-supported Symmetrix logical volumes. Refer to the specific Symmetrix model product guide for information on supported Symmetrix logical volume limits.

Symmetrix RAID 10 (Mirrored Striped Mainframe Volumes with DMSP)

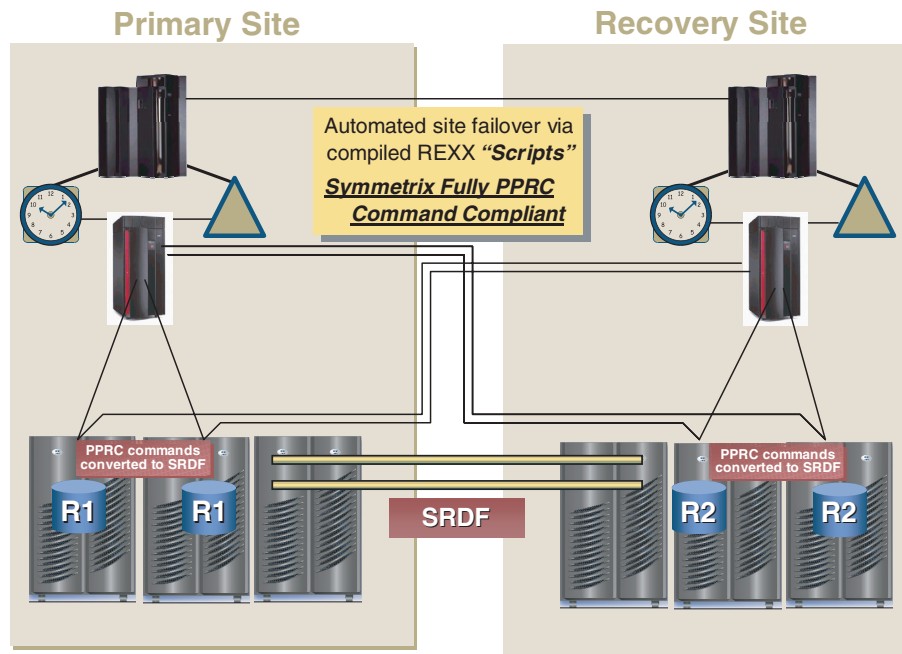
To improve mainframe volume performance, Symmetrix RAID 10 stripes data of a logical device across multiple Symmetrix logical devices. (The feature is analogous to metavolume stripes on open systems.) Four Symmetrix devices (each one-fourth the size of the original mainframe device) appear as one mainframe device to the host. Any four Symmetrix logical devices can be chosen to define a RAID 10 group provided they are the same type (for example, IBM 3390) and have the same mirror configuration. Striping occurs across this group of four devices with a striping unit of one cylinder. Since each member of the stripe group is mirrored, the entire set is protected. RAID 10 uses four pairs of disks in its Symmetrix implementation.

At Enginuity Version 5568, in an SRDF configuration, RAID 10 striped volumes can be remotely paired with non-RAID 10 volumes. The size of the non-RAID 10 volume must equal the sum of all members in the RAID 10 volume. At Enginuity Version 5669, the sum of the R2 volumes can be larger than the sum of the R1 volumes.

PPRC Command Support

Enginuity level 5568 (or higher) enables the Symmetrix system to support native IBM Peer-to-Peer Remote Copy (PPRC) commands through a Symmetrix feature called Compatible Peer. PPRC is the remote copying solution available with IBM storage systems (Figure 2-2).

Compatible Peer implements PPRC version 1, architectural level 2, beginning with Enginuity level 5568. Enginuity level 5671 adds support for PPRC version 1, architectural levels 3 and 4 (Hyper-swap support). As a result, Symmetrix systems will support IBM's Geographically Dispersed Parallel Sysplex (GDPS) solution. Compatible Peer is available on Symmetrix systems with connections to ESCON and/or FICON hosts.



SYM-000092

Figure 2-2 PPRC and GDPS Support

For more information on PPRC support for Symmetrix systems, refer to *GDPS: Installation and Testing Guide*.

SRDF Link Configurations

The links between a given pair of Symmetrix systems in an SRDF configuration can use one of the following methods for transmitting data:

- ◆ Unidirectional
- ◆ Bidirectional
- ◆ Dual-Directional

SRDF Unidirectional Link Configuration

If all primary (source, R1) volumes reside in one Symmetrix system and all secondary (target, R2) volumes reside in another Symmetrix system, write operations move in one direction, from primary to secondary. This is a *unidirectional* configuration, in which data moves in the same direction over every link in the SRDF group.

SRDF Bidirectional Link Configuration

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.

For information about how write operations occur using an ESCON SRDF bidirectional link configuration, refer to *Write Operations in an ESCON Bidirectional Configuration* on page 3-2.

SRDF Dual-Directional Link Configuration

For an ESCON-based extended-distance SRDF configurations (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 is required. With a dual-directional configuration, multiple SRDF groups are used; some groups send data in one direction, while other groups send data in the opposite direction.

SRDF Link and Volume States

SRDF Link States

SRDF links can be in two possible states: online or offline. The 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.

The SRDF link can be offline if one or more of the following occurs:

- ◆ The remote adapter is offline (the operator panel switch on the remote adapter is off)—link disabled.
- ◆ The remote adapter is online but the link is offline (damaged or disconnected cable or other damaged hardware)—link disabled.

Configuration Settings Affecting Device Ready and Link States

If all SRDF links fail, the Symmetrix system retains the last known logical state for the affected devices. This behavior is known as a *pure ready state*. In other words, if a device is in a ready state before a link failure, it is restored to a ready state. If a device is in a not-ready state before a link failure, it is restored in a not-ready state after the link is again made ready. However, all devices in the group will remain not ready on the SRDF link after all links fail if the `prevent automatic recovery after all links fail` setting was set to `Yes` in the Symmetrix configuration file.

The configuration setting, `Force RAs offline after power up`, prevents SRDF *links* from coming online following a Symmetrix power cycle.

Prior to Enginuity level 5669, both configuration settings described above were applied at the Symmetrix system level. At Enginuity level 5669 and later, these configuration settings are applied at the SRDF group level.

The device level pure ready state behavior does not apply to SRDF/A mode of operation. Refer to Chapter 4 for information on SRDF/A operations.

Logical Volume States

This section describes the three device states that an SRDF logical volume (primary, source R1 or secondary, target R2) reflects to the host connected to the Symmetrix unit on which the volumes are located.



CAUTION

Understanding how the three device states of an SRDF logical volume are achieved is fundamental to grasping the concepts of SRDF operation. Make sure you fully comprehend this section before attempting any SRDF operations.

The device states can be:

- ◆ Not Ready
- ◆ Read Only
- ◆ Write Enabled

These device states are determined by a combination of two substates: the SRDF state and the device host interface state, as shown in Figure 2-3.

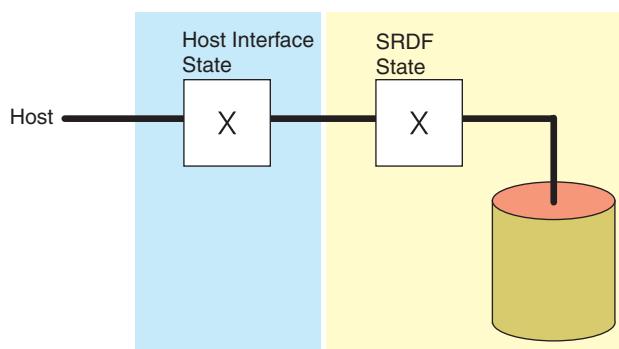


Figure 2-3 SRDF Logical Volume State

In general, you can look at the two states as two layers:

- ◆ An internal layer (SRDF view)—the SRDF state
- ◆ An external layer (host view)—the host interface state

These two layers are configured by different sets of internal Symmetrix parameters, which can be set by EMC host-based control software or `IMPL.bin` file configuration. The state ultimately seen by the host is determined by the combination of these two device states.

Table 2-1 on page 2-15 and Table 2-2 on page 2-15 summarize the device state seen by the host as determined by the individual states. For example, only when a secondary (R2) device is write enabled for both the host interface state and the SRDF state can the host write to the device.

The following sections describes these two substates and how the various device substate combinations determine the actual overall device state presented to the host.

SRDF Volumes States (Symmetrix View)

Primary (R1) Volume States

This section lists the substates a primary (R1) or secondary (R2) volume can have for SRDF operations.

A primary (R1) volume can have the following states listed for SRDF operations:

- ◆ Ready and Read/Write—In this state, the primary (R1) volume is available for read/write operations. This is the default primary (R1) volume state.
- ◆ Not Ready—SRDF volumes can be not ready locally or not ready remotely
 - Locally Not Ready—If the local (R1) volume fails, the host continues to recognize that volume as available for read/write operations as all reads and/or writes continue uninterrupted with the secondary (R2) volume in that remotely mirrored pair.
 - Remotely Not Ready—If R1 volumes are remotely not ready, write updates will not propagate to the secondary volumes. Changes to the R1 volumes are marked invalid as *owed* to the secondary volumes R2.

Secondary (R2) Volume States

A secondary (R2) volume can have one of the three states listed for SRDF operations:

- ◆ Not Ready—In this state, the secondary (R2) volume responds *not ready* to the host for all read and write operations to that volume.
- ◆ Read Only—In this state, the secondary (R2) volume is available for read-only operations. This is the default secondary (R2) volume state.
- ◆ Read/Write—In this state, the secondary (R2) volume is available for read/write operations.

Symmetrix Volume States (Host View)

This section lists the states a primary (R1) or secondary (R2) volume can have for host operations. This represents the volume state as seen by the host interface.

Primary (R1) Volume States

A primary (R1) volume can have one of the three following states. This state is seen by the host connected to the Symmetrix unit in which that volume resides:

- ◆ Write Enabled—In this state, the primary (R1) volume is available for read/write operations. This is the default primary (R1) volume state.
- ◆ Read Only—In this state, the primary (R1) volume responds *device write protected* to the host for all write operations to that volume.
- ◆ SRDF Not Ready—In this state, the primary (R1) volume responds *not ready* to the host for all host accesses to that volume.

Secondary (R2) Volume States

A secondary (R2) volume can have one of the three following states. This state is seen by the host connected to the Symmetrix unit in which that volume resides:

- ◆ Write Enabled—In this state, the secondary (R2) volume is available for read/write operations.
- ◆ Read Only—In this state, the secondary (R2) volume responds *device write protected* to the host for all write operations to that volume. This is the default secondary (R2) volume state.
- ◆ Not Ready—In this state, the secondary (R2) volume responds *not ready* to the host for all host accesses to that volume.

Host Accessibility

The tables in this section describes the accessibility state of the primary (R1) and secondary (R2) volumes to the host connected to the Symmetrix system containing the primary (R1) volumes. The accessibility of the host to a particular Symmetrix volume depends on its state from both the host and SRDF view. Table 2-1 on page 2-15 lists the accessibility for a primary (R1) volume. Table 2-2 on page 2-15 lists the accessibility for a secondary (R2) volume. Refer to the *EMC Solutions Enabler Symmetrix SRDF CLI Product Guide* and the *Symmetrix SRDF Host Component for z/OS Product Guide* for how this information is presented in these operating environments.

Table 2-1 Primary (R1) Volume Accessibility

Host Interface State	SRDF State	Accessibility
Write Enabled	Read/Write	Read/Write
Write Enabled	Not Ready	Depends on secondary (R2) volume availability
Write Disabled	Read/Write	Read Only
Write Disabled	Not Ready	Depends on secondary (R2) volume availability
Not Ready	Read/Write	Unavailable
Not Ready	Not Ready	Unavailable

Table 2-2 Secondary (R2) Volume Accessibility

Host Interface State	SRDF State	Accessibility
Write Enabled	Not Ready	Unavailable
Write Enabled	Read Only	Read Only
Write Enabled	Read/Write	Read/Write
Write Disabled	Not Ready	Unavailable
Write Disabled	Read Only	Read Only
Write Disabled	Read/Write	Read Only
Not Ready	Not Ready	Unavailable
Not Ready	Read Only	Unavailable
Not Ready	Read/Write	Unavailable

Primary Modes of Operation

You select modes of operation for individual devices and manage these modes at the logical volume level.

In an SRDF/Asynchronous (SRDF/A) environment, you select modes of operation for individual devices and manage these modes at the SRDF group level. For information on SRDF/A refer to Chapter 4, SRDF/Asynchronous Operations. SRDF/A is sold as a separate license by EMC.

SRDF /Data Mobility is not available in synchronous or semi-synchronous modes, but is supported in adaptive copy mode only. Refer to *Adaptive Copy Modes* on page 2-19 for more information.

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.

Synchronous mode provides real-time mirroring of data between the local Symmetrix system and the remote Symmetrix system. Data is written to global memory of both systems before the application I/O is completed to the host, ensuring the highest possible data availability. Synchronous mode processing is described in the following steps and illustrated in Figure 2-4.

1. The local Symmetrix system containing the primary (source) volume receives a write operation from the user application.
2. The write I/O operation is propagated to the remote Symmetrix system (containing the secondary or target volume); the local Symmetrix system does not accept additional write operations to the primary volume.
3. The remote Symmetrix system sends an acknowledgement to the local Symmetrix system.
4. The local Symmetrix system sends an I/O complete message to the local host; the local Symmetrix system now accepts additional write operations to the primary volume.

In the mainframe PAV environment, SRDF synchronous mode supports multiple concurrent write I/Os at the volume level.

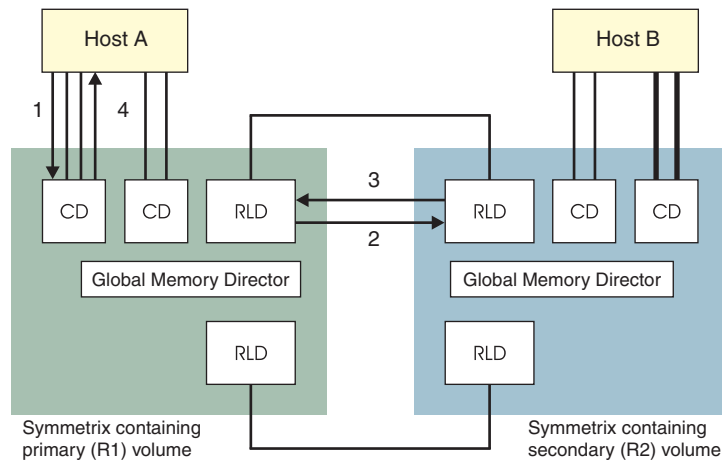


Figure 2-4 Synchronous Mode

Semi-Synchronous Mode

Used mainly for the extended distance 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.

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.

Semi-synchronous mode writes data to the primary device in the local Symmetrix system, completes the I/O, and then synchronizes the data with the secondary device in remote Symmetrix system as shown in the following steps and illustrated in Figure 2-5 on page 2-18.

1. The local Symmetrix system containing the primary volume receives a write operation from the user application.
2. The local Symmetrix system sends an I/O complete message to the local host. The user application now considers the I/O to be complete. Read operations continue normally. A second write I/O will not be accepted.

3. Data from the first write operation is propagated to the remote Symmetrix system containing the secondary volume.
4. The remote Symmetrix system sends an acknowledgement from the first write operation to the local Symmetrix system. When the acknowledgement is received, another write I/O from the host can be accepted.

Because the I/O is completed before synchronizing data with the remote system, the semi-synchronous mode provides an added performance advantage.

In the mainframe environment, semi-synchronous mode is not supported when the Symmetrix host attachment is through FICON connection. In the mainframe PAV environment, semi-synchronous mode is not recommended because it provides no performance benefit. EMC recommends that SRDF users seeking the increased parallelism offered by PAVs run SRDF in synchronous mode.

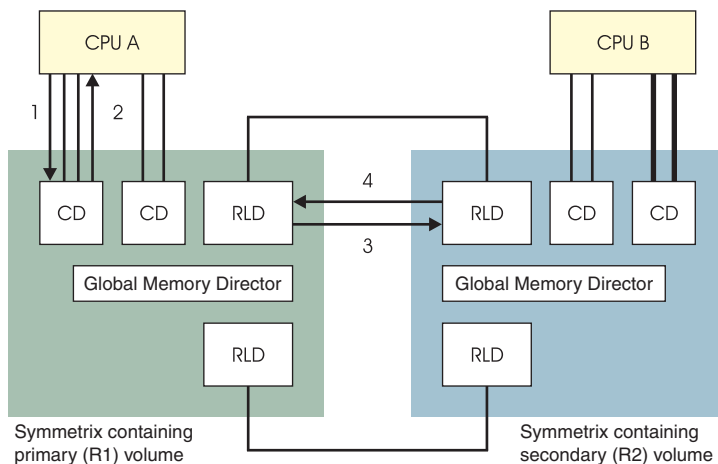


Figure 2-5 Semi-Synchronous Mode

Secondary Modes of Operation

With the SRDF/Synchronous product offering, secondary modes are used in conjunction with the primary operational modes described in *Primary Modes of Operation* on page 2-16. As with the primary modes of operation, you select secondary modes of operation for individual devices and manage these modes at the logical volume level. With SRDF/Data Mobility, only these secondary modes of operation are supported.

Secondary modes include:

- ◆ Adaptive copy write-pending mode
- ◆ Adaptive copy disk mode

Adaptive Copy Modes

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.

There are two adaptive copying modes: adaptive copy write-pending (AW) mode and adaptive copy disk (AD) mode. Both modes allow write tasks to accumulate on the local system before being sent to the remote system.

Adaptive Copy Write-Pending Mode

With adaptive copy write-pending mode, write tasks accumulate in *global memory*. A background process moves, or *destages*, the write-pending tasks to the primary volume and its corresponding secondary volume on the other side of the SRDF link. When the maximum skew value is reached, the primary volume reverts to its primary mode of operation, either synchronous or semi-synchronous, whichever is currently specified. The device remains in the primary mode until the number of tracks to remotely copy becomes less than the maximum skew value.

Adaptive copy write-pending mode reverts to the currently specified primary mode (synchronous or semi-synchronous) if 75 percent of the write-pending limit for the Symmetrix system is reached, regardless of whether the maximum skew value specified for each device is reached.

The advantage to this mode is that it is faster to read data from global memory than from disk, thus improving overall system performance.

An additional advantage is that the unit of transfer across the SRDF link is the updated blocks rather than an entire track, resulting in more efficient use of SRDF link bandwidth.

The disadvantage is that global memory is temporarily consumed by the data until it is transferred across the link.

Adaptive Copy Disk Mode

Adaptive copy disk mode is similar to adaptive copy write-pending mode, except that 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 the skew value is reached, the primary volume reverts to its primary mode of operation, either synchronous or semi-synchronous, whichever is currently specified.

The advantages and disadvantages of this mode are opposite from those of the adaptive copy write-pending mode; that is, while less global memory is consumed it is typically slower to read data from disk than from global memory, additionally, more bandwidth is used because the unit of transfer is the entire track. In addition, because it is slower to read data from disk than global memory, device resynchronization time will increase.



CAUTION

Adaptive copy disk mode should not be used if the primary volumes are not protected by RAID 1, RAID 10, RAID-S, Parity RAID (3+1), Parity RAID (7+1), RAID 5(3+1), or RAID 5 (7+1).

Additional SRDF Modes and Attributes

There are several SRDF modes and attributes that affect SRDF behavior under certain conditions. These attributes are described in the following sections.

Domino Modes

Domino modes effectively stop all write operations to both primary and secondary volumes if all mirrors of a primary or secondary device fail, or if all SRDF links in a link group become unavailable. While such a shutdown temporarily halts production processing, domino modes can prevent data integrity exposure caused by *rolling disasters*.

There are two types of domino modes:

- ◆ Device domino mode
- ◆ Link domino mode.

Device Domino Mode

You can set device domino mode at the device level on primary volumes. If this mode is set to *Yes* on a primary volume, and the secondary volume becomes unavailable to its primary volume for any reason, the primary volume becomes unavailable to its host.

Link Domino Mode

You can set link domino mode at the SRDF group level at either side of the SRDF links. If this mode is set to *Yes* for an SRDF group, and the last remaining link in the SRDF group fails, all primary (source) volumes in the SRDF group become unavailable (*not ready*) to their host.

Once the *not ready* condition is set, you must re-enable the volumes using one of the SRDF control utilities (using the Symmetrix service processor or EMC host-based software).



CAUTION

With either domino mode, the appropriate primary volumes are made *not ready* and all related applications stop. This is an extreme measure. A more moderate measure (if you are using SRDF in a mainframe environment, or an open systems environment with EMC PowerPath® software) is to implement consistency groups (refer to *SRDF/Consistency Groups (SRDF/CG)* on page 2-25 for more information).

Invalid Tracks Attribute

You can only set the invalid tracks attribute on secondary SRDF volumes. When the invalid tracks attribute is enabled, SRDF makes the device not ready to the host if you attempt to access data from that secondary volume when it is not synchronized with its primary volume.

The purpose of the invalid tracks attribute is to inform the user that the data on the secondary devices may not be suitable for use in disaster recovery situations. It is a user's decision to proceed using data from devices that were made not ready by the invalid tracks attribute. In such cases another form of data recovery may be more appropriate. If the user decides to use the existing data on the secondary volume, and the not-ready condition has been set by this attribute, the not-ready condition can be reset by host-based SRDF control software.

SRDF System-Level Attributes

You can set system-level attributes to do the following:

- ◆ **Force RAs offline after powerup**—This attribute applies to all SRDF director types and forces SRDF adapter ports offline following a power outage. This attribute can be used to prevent a data integrity exposure due to a rolling disaster following a power outage.
- ◆ **Prevent automatic link recovery after all links fail**—This attribute prevents SRDF device pairs from automatically resuming their SRDF relationship following a failure of all SRDF links in an SRDF group. This attribute can provide the opportunity to preserve a consistent copy of data on the secondary devices (using Business Continuance Volumes) by ensuring that device level resynchronization does not occur automatically following link restoration.

Prior to Enginuity Version 5568, the Prevent Automatic Recovery attribute is a system level attribute. At Enginuity Version 5568 and higher, the Prevent Automatic Recovery attribute is an SRDF group-level attribute.

Concurrent SRDF

Enginuity 5567 and later 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.

Concurrent SRDF requires that each remote mirror operate in the same primary mode, either both synchronous or both semi-synchronous, but allows either (or both) volumes to be placed into one of the adaptive copy modes.

Figure 2-6 shows a concurrent SRDF configuration in which the primary volume is communicating with one secondary volume in synchronous mode. Concurrently, the same primary volume is communicating with its other secondary volume in one of the adaptive copy modes (adaptive copy write-pending mode or adaptive copy disk mode). Any combination of synchronous/semi-synchronous and adaptive copy is allowed with the exception of one volume operating in synchronous mode and the other operating in semi-synchronous mode.

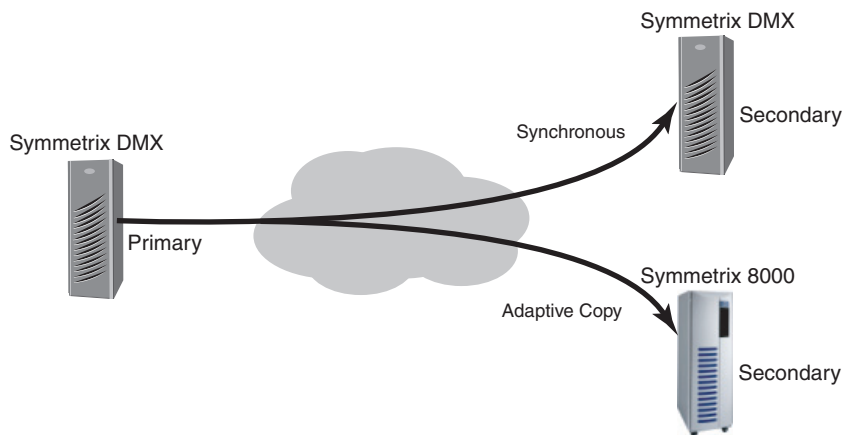


Figure 2-6 Concurrent SRDF Configuration

Normal operating rules for SRDF also apply to concurrent SRDF configurations. When operating in synchronous mode, ending status for an I/O is not presented until the remote Symmetrix system acknowledges receipt of the I/O to the primary Symmetrix system. If both secondary volumes are operating in synchronous mode, ending

status is not presented until both volumes acknowledge receipt of the I/O. If one remote mirror is in synchronous mode and one remote mirror is in adaptive copy mode, ending status is presented to the host when the synchronous volume acknowledges receipt of the I/O.

SRDF/Consistency Groups (SRDF/CG)

SRDF/CG is an SRDF product offering designed to ensure the consistency of the data remotely copied by SRDF/S operations in the event of a rolling disaster.

An SRDF consistency group is a group of Symmetrix devices specially configured to act in unison to maintain the integrity of a database distributed across multiple Symmetrix systems controlled by a mainframe host computer or open systems host computer using EMC PowerPath software.

Another way to ensure the integrity of a remote database is to use domino mode (refer to *Domino Modes* on page 2-21).

How a Consistency Group Works

Assume that you have an SRDF configuration in which three Symmetrix systems contain primary (source) devices, and two additional Symmetrix systems contain secondary (target) devices. The units with primary devices send data to the units with secondary devices as shown in Figure 2-7.

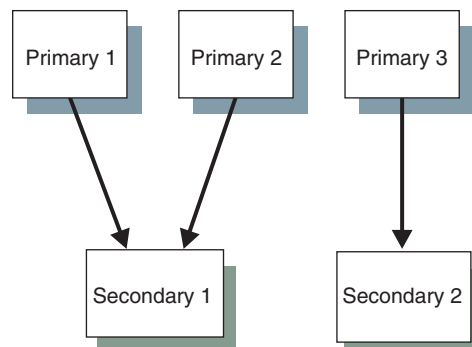


Figure 2-7 Primary and Secondary Relationships

Next, assume that the links between Primary 2 and Secondary 1 fail. Without a consistency group, Primaries 1 and 3 continue to write data to the Secondaries 1 and 2 while Primary 2 does not (Figure 2-8).

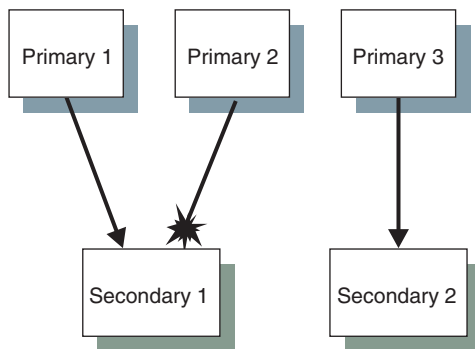


Figure 2-8 Failed Link between Primary 2 and Target 1

The result is that the copy of the data spread across Secondaries 1 and 2 becomes inconsistent.

However, if Primaries 1, 2, and 3, belong to a consistency group, as shown in Figure 2-9, and the link between Primary 2 and Secondary 1 fails, the consistency group automatically stops Primaries 1 and 3 from sending data to Secondaries 1 and 2, as shown in Figure 2-10.

Thus, the consistency of the data (spanning Secondaries 1 and 2) remains intact.

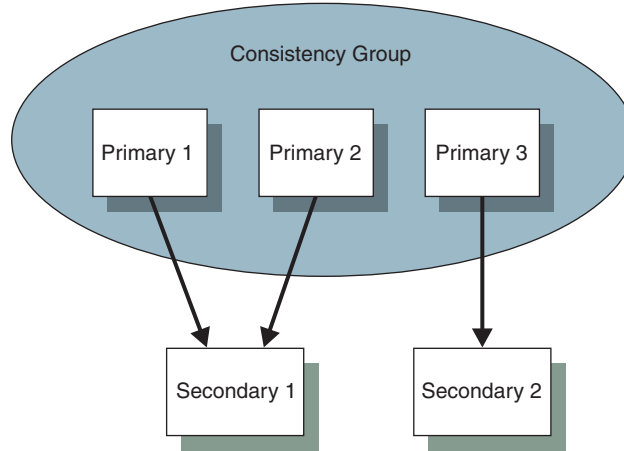


Figure 2-9 Primaries 1, 2, and 3 in a Consistency Group

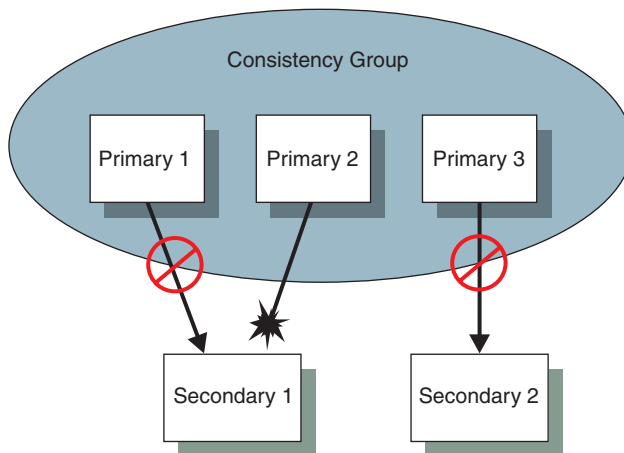


Figure 2-10 Failed Link between Primary 2 and Target 1

Continuous Processing

I/O to the primary devices in the consistency group can still occur even when the devices are not ready on the SRDF links. Such updates are not immediately sent to the remote side. However, they are propagated after the affected links are again operational, and data transfer from the primary devices to the secondary devices resumes.

z/OS Technical Considerations

You use host-based EMC software (Consistency Group for z/OS) to manage and monitor consistency groups within a mainframe environment. The host software identifies the consistency group devices by VOLSER, CUU, or SMS storage group. When the consistency group is initialized, all volumes and related data are enabled with the consistency group attribute. If a write operation on a primary volume fails to reach the corresponding secondary volume, the Symmetrix system responds with a unit check to the host on the failing I/O.

Upon receipt of the unit check, the consistency group utility prevents dependent I/O from being issued by the host and suspends the SRDF relationship for all primary devices in the consistency group.

For more information about consistency groups and the consistency group utility, refer to the EMC *Consistency Group for z/OS Product Guide*, P/N 300-001-069.

SRDF Operations

This chapter describes basic SRDF operations and covers the following topics:

- ◆ Write Operations 3-2
- ◆ Read Operations 3-3
- ◆ Recovery Operations 3-5
- ◆ Business Continuance Using SRDF 3-7
- ◆ Business Continuance Using SRDF and TimeFinder 3-9
- ◆ R1/R2 Swap 3-13
- ◆ Dynamic R1/R2 Swap 3-14
- ◆ Migrating Data From R1 to a Larger R2 Device 3-15

Write Operations

The *write* task is the most common SRDF/S operation. This section describes the write operations for unidirectional, dual-directional, and bidirectional configurations.

Write Operations in a Unidirectional or Dual-Directional Configuration

This process applies to write operations over links that use a unidirectional or dual-directional protocol, in which data moves in only one direction over a given link.

In synchronous or semi-synchronous mode, the local host sends a write I/O to the host adapter which, in turn, sends the I/O to the SRDF queue. The I/O then moves across the SRDF links to the remote Symmetrix system. When it receives the I/O, the remote Symmetrix system returns an acknowledgement to the local Symmetrix system, which relays a message to the host adapter.

Write Operations in an ESCON Bidirectional Configuration

In a bidirectional protocol, data flows in two directions over the same SRDF link. When an RA-2 (normally the receiving end of the SRDF link) must send write data to its corresponding RA-1 (normally the sending end of the SRDF link), the RA-2 cannot simply transmit the data over the SRDF link to the RA-1. This is due to the nature of ESCON protocol and the channel-to-control unit architecture of ESCON-based SRDF. Instead, the RA-2 must ask the RA-1 to *read* the data from a primary volume serviced by the RA-2. When the RA-1 reads the data from the RA-2 volume, the write operation from RA-2 to RA-1 is effectively accomplished.

Read Operations

Primary Volume Read Operations

Read operations from a primary volume behave as normal read operations and are not impacted in any way by SRDF.

Secondary Volume Read Operations

Read operations from a secondary volume can be initiated in response to read I/Os issued by the host attached to the primary Symmetrix system or by the host attached to the secondary Symmetrix system.

Read Operations from the Primary Host

Read operations, in which the primary host reads data from a secondary volume in the remote Symmetrix system, are performed only to recover from a local data availability problem. Several events can cause such a read operation to take place:

- ◆ If data is not in global memory and all of the primary (source) devices are in a not-ready state.
- ◆ If a primary device is in a ready state but the requested track is invalid on the local device.
- ◆ If a disk adapter has a problem accessing a primary device as in the case of a drive timeout or cyclic redundancy check (CRC) error. In these cases, the local Symmetrix system requests the data from the remote Symmetrix system.
- ◆ If a track on a primary device is currently available only from a RAID rebuild, the local Symmetrix system requests the track data from the remote Symmetrix system. This method is faster than accessing the data from the RAID rebuild. The Symmetrix system always performs a full-track read for count-key-data (CKD) and at least a contiguous-block read for fixed block architecture (FBA) data.

In a read operation, the remote Symmetrix system reads data from the secondary device and sends the data across the link to the local Symmetrix system.

If a data availability problem is caused by a CRC error, an invalid track, or a RAID rebuild, data can be destaged (moved from global memory to disk) within the local Symmetrix system if it is in global memory. This process can eliminate the need to read the data from the remote Symmetrix system.

Read Operations from the Secondary Host

SRDF secondary volumes can optionally be made read-only while SRDF is in operation. A host attached to the secondary Symmetrix system may read data from the secondary devices provided the operating system has support for reading data from devices that are hardware write-protected. Such support varies by operating system type. Consult your server operating-system vendor.

If the secondary host has read-only access to the secondary device, it does not interfere with SRDF operations. SRDF ensures that the most current image of data is made available to the secondary host.

Recovery Operations



CAUTION

This section explains recovery in the context of hardware failure. For adequate protection against logical data corruption, EMC recommends that you regularly back up your data.

If the local host or local Symmetrix system fails, and the primary volumes are operating in synchronous mode, the remote Symmetrix system can be ready for operations in minutes. When the failure occurs, the primary and secondary volumes are synchronized within one I/O and there are no invalid tracks on either the primary or secondary side.

Failover to the Secondary Symmetrix System

If a remote host is to be activated to continue production processing, you must initiate a *failover* process. You can use EMC software products such as EMC ControlCenter, Solutions Enabler, or the SRDF Host Component for z/OS to control the failover process.

The failover process involves changing the states of the secondary volumes to read/write and transferring production processing to the secondary host. After the failover has been initiated, updates made to the secondary volumes appear as invalid tracks owed to the primary volumes. Processing continues on the remote host until the local host and local Symmetrix system are operational again and a failback is performed.

Failback to the Primary Symmetrix System

After the primary host and the Symmetrix system containing the primary volumes are again operational, production processing can resume on the primary host. The following steps are required to transfer processing from the secondary host back to the primary host:

1. Halt processing on the secondary host and change the states of the secondary devices to read-only.
2. If a power on or IML is required of the primary Symmetrix system, make sure that the SRDF links are physically disabled to prevent movement of invalid track metadata across the SRDF links.

3. Bring the Symmetrix system on which the primary volumes reside to a ready state. If the SRDF links were physically disabled in the previous step, re-enable the links.
4. Use EMC software products such as EMC ControlCenter, Solutions Enabler, or the SRDF Host Component for z/OS to send invalid track meta data information from the secondary volumes to the primary volumes.
5. Bring the SRDF links online and restart the local host. The primary volumes automatically receive the appropriate data from the secondary volumes, effectively resynchronizing the primary and secondary volumes.

You can use EMC software products such as EMC ControlCenter, Solutions Enabler, or the Stored Procedure Executive (SPE) of SRDF Host Component for z/OS to automate or semi-automate this process.

Recovery for a Large Number of Invalid Tracks

If the recovery site (the secondary host and the Symmetrix system containing the secondary volumes) has handled production processing for a long period of time, there might be a large number of invalid tracks (for example, 500 GB) *owed* to the primary volumes. In this case, you can resynchronize the primary and secondary volumes while the secondary host continues production processing. Then, when there is a relatively small number of invalid tracks on the primary volumes (for example, 50 GB), you can shut down the secondary host and restart the primary host.

This capability is enabled by using the Solutions Enabler or the SRDF Host Component.

Business Continuance Using SRDF

Business continuance refers to practices that enable you to achieve nearly nonstop, 24x7 business operations.

SRDF enables business continuance by allowing you to suspend remote mirroring and temporarily enable the secondary volumes for read/write activity, effectively creating two separate host/storage systems (Figure 3-1).

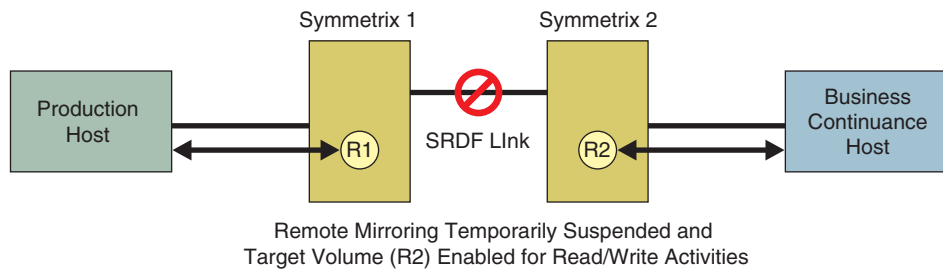


Figure 3-1 SRDF Business Continuance

The following business continuance practices are possible with SRDF:

- ◆ Backups — Avoid taking system offline while nightly backups run.
- ◆ Remote processing — For example, some systems must produce daily billing statements. These statements can be produced at a remote site without interrupting normal processing.
- ◆ Decision Support System (DSS) operations.
- ◆ Application testing.

Concurrent Operations

You can temporarily suspend the SRDF links so that you can read and write data on both the primary and secondary volumes concurrently. This enables you, for example, to run backups on the secondary volumes while production processing continues on the primary volumes (a business continuance practice).

You can then resume the links and copy data from the primary volumes to the secondary volumes.

This last operation propagates any updates made to the primary volumes while the links were suspended and overwrites any changed data on the secondary volumes, bringing both volumes into a synchronized state (Figure 3-2).

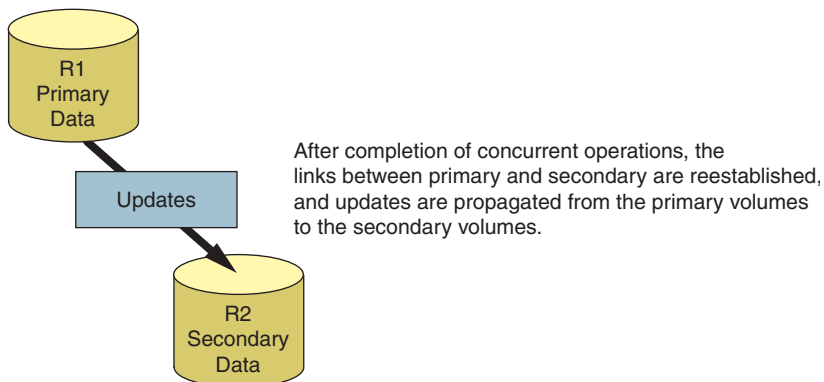


Figure 3-2 Primary-to-Secondary Resynchronization

Alternatively, you can *resynchronize* data in the opposite direction, from secondary to primary. This operation is useful if, for example, you performed application testing on the secondary volumes, production processing was halted on the primary volumes, the testing was successful, and you want to keep the updates.

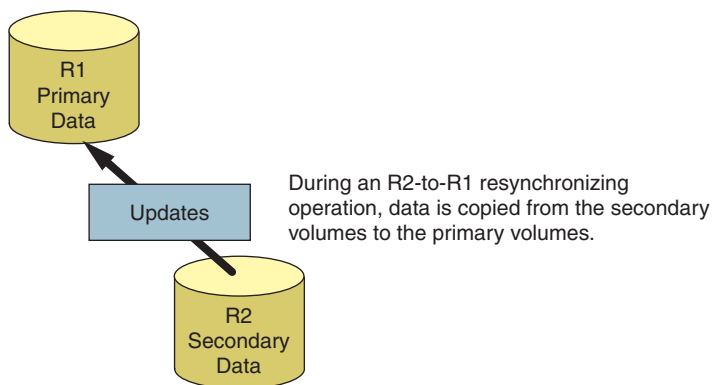


Figure 3-3 Secondary-to-Primary Resynchronization

Business Continuance Using SRDF and TimeFinder

The Symmetrix TimeFinder product offering uses one or more software-controllable, dynamically assignable mirror volumes called business continuance volumes (BCVs). SRDF, when used with EMC TimeFinder software, allows you to perform business continuance operations without temporarily suspending remote mirroring. TimeFinder BCVs can be locally mirrored or remotely mirrored via SRDF. The sections that follow describe how SRDF can be used with TimeFinder for business continuance operations.

The SRDF/Automated Replication solution (introduced on Page 1-5) combines host-based automation with TimeFinder and SRDF functionality. This functionality, described below, provides an automated long distance disaster restart solution.

For information about EMC TimeFinder, refer to the documentation (available on the EMC website: <http://powerlink.emc.com>, then follow these links: **Support, Document Library, Software, TimeFinder**, or contact your EMC representative.

Using TimeFinder/Mirror BCVs with Primary Devices

TimeFinder operations (establish, split, restore) are allowed between a BCV and a primary device. Restore operations to a primary device (R1) will result in the restored data being propagated across the SRDF link if the remote device is not in a suspended state. If the remote device is in a suspended state, restored data will be marked as invalid (*owed*) to the secondary device.

Using a BCV as a Primary (Source) Device

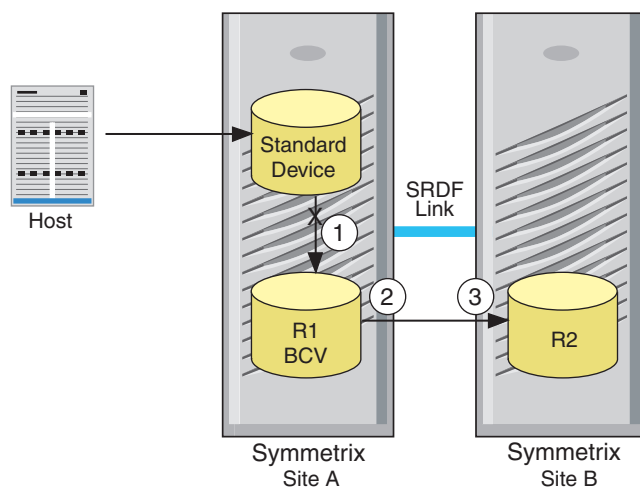
A BCV device may also function as an SRDF primary device. Such a device is called a BCV R1 device. A BCV R1 device may optionally be locally mirrored.

When a BCV R1 is attached to a standard device, the corresponding secondary device is suspended. When the BCV is split, the default behavior is to propagate the changed data to the secondary device. Optionally, data transmission can be suppressed and the BCV R1 device will mark the changed data as invalid (*owed*) to the secondary device, pending the resumption of the BCV R1 to R2 device relationship (Figure 3-4 on page 3-10).

Concurrent SRDF is not supported for BCV R1 devices because a BCV R1 device is not allowed to have two remote mirrors.

Using TimeFinder BCVs with Secondary Devices

TimeFinder operations (establish, split, restore) are allowed between a BCV and a secondary device. By default, restore operations *will not* result in data propagating back to the primary device. Changed data will be marked as invalid (*owed*) to the primary device following a restore operation. If data propagation back to the primary device is desired, TimeFinder software options are available to enable this function.



1. The BCV device is split.
2. The R2 on the Symmetrix at Site B is resynchronized with the R1 BCV at Site A.
3. Data is mirrored to the R2 device on the Symmetrix at Site B.

Figure 3-4

SRDF Single-Hop Configuration (BCV Functioning as a Primary SRDF Device)

Using a BCV as a Secondary (Target) Device

A BCV device may also function as an SRDF secondary device. Such a device is called a BCV R2 device. A BCV R2 device may optionally be locally mirrored.

When a BCV R2 is attached to a standard device, the corresponding primary device is suspended. When the BCV R2 is split, it is again available for SRDF operations.

Host software may restrict functionality of BCV R2 devices. Please refer to the SRDF Host Component or Solution Enabler documentation for information on using BCV R2 devices.

SRDF Remote Command Support

You can issue SRDF and BCV control commands from a host or server connected to the local Symmetrix system and have those commands transmitted across the SRDF link. A host or server does not have to be located at the secondary location, allowing you for example, to create a point-in-time copy of your data at the remote site without needing a host computer or server at the remote site.

SRDF Multi-Hop

SRDF's Multi-hop capability creates the ability to provide disaster recovery protection at great distances without any data loss. A *middle or bunker* site transitions the synchronous replication solution by using TimeFinder in conjunction with SRDF.

SRDF Multi-hop uses a BCV R1 on the secondary side of an SRDF synchronous configuration to perform an incremental synchronization with a secondary volume on a third Symmetrix system. When split, the BCV R1 copies only changed tracks to the secondary volume located in the third Symmetrix system. This process provides two benefits. It eliminates the additional performance impact on synchronous operations across great distances. In addition, the differential resynchronization dramatically reduces bandwidth requirements.

Although the local side of an SRDF multi-hop configuration requires an SRDF license with the ability to do synchronous or semi-synchronous operations, you can use SRDF - Data Mobility (in adaptive copy mode) on Symmetrix systems connected to the SRDF extended distance configuration.

Figure 3-5 on page 3-12 illustrates SRDF multi-hop using ESCON connectivity however, any type of SRDF connection is supported by multi-hop.

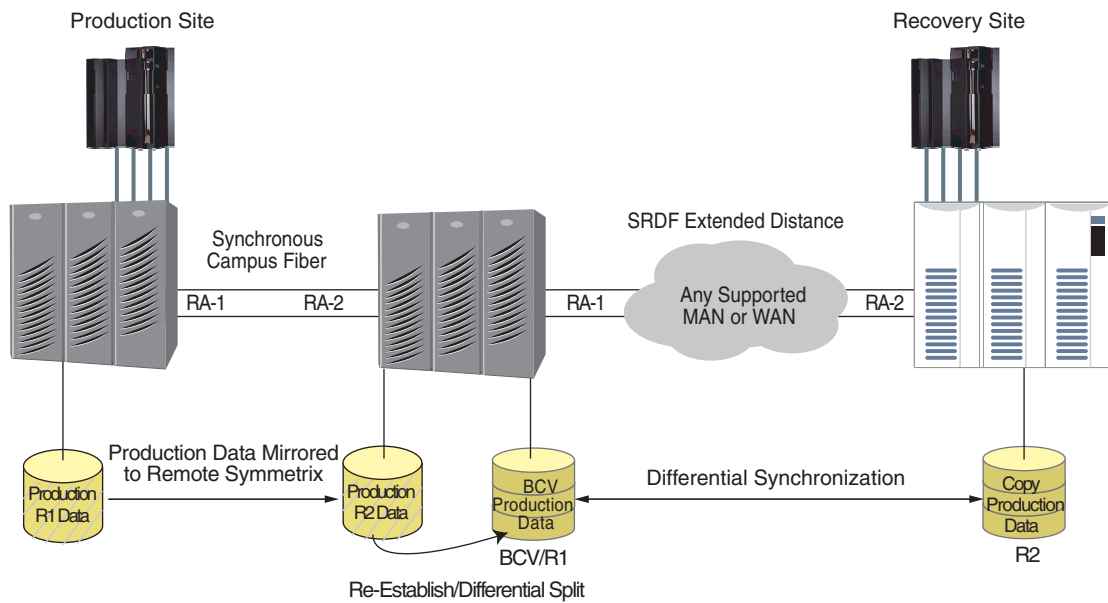


Figure 3-5 SRDF Multi-Hop Configuration

R1/R2 Swap

Dynamic SRDF provides the capability of swapping the R1/R2 personality of an SRDF pair. During a swap, primary R1 devices become secondary R2 devices, and conversely the secondary R2 devices become primary R1 devices. Swapping SRDF devices allows the R2 side to take over operations while retaining a remote mirror on the R1 side. Swapping is especially useful after failing over an application from the R1 side to the R2 side.

Before you can perform an R1/R2 swap, the SRDF pair must be in a suspended state.

After the swap operation is complete, the SRDF pair remains in the suspended state. Incremental or full resynchronization capability is maintained and controlled through EMC SRDF host based control software. For specific procedures for controlling the R1/R2 swap operation, refer to the *SRDF Host Component for z/OS Product Guide* or the *Solutions Enabler SRDF CLI Product Guide*.

R1/R2 Swap Procedure History

When introduced in Enginuity 5x65 and 5x66, the R1/R2 swap feature used an online configuration change procedure to affect the personality swap of the primary (source, R1) and secondary (target, R2) devices. This process was often lengthy, taking several minutes to complete. Beginning with Enginuity 5567, a new capability called Dynamic SRDF was introduced, which dramatically improved the performance of R1/R2 personality swaps (Figure 3-6).

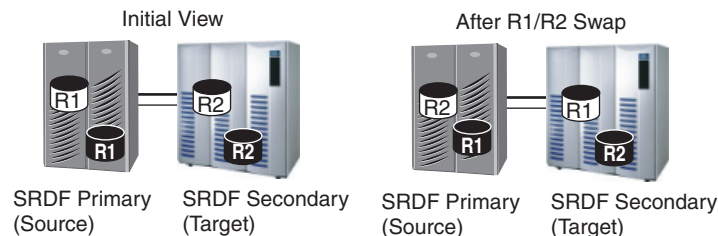


Figure 3-6 R1/R2 Swap Concept

Dynamic R1/R2 Swap

Dynamic R1/R2 swap is available with Enginuity Version 5567 or higher. Dynamic R1/R2 swap is faster than the normal swap. If you have an SRDF license and a Configuration Manager license with Enginuity Version 5567 or higher, you can make the devices SRDF capable. Host software can then be used to create Dynamic SRDF primary and secondary device relationships. Host software can also be used to swap the SRDF personalities of the SRDF devices. Primary (source) R1 device(s) become secondary (target) R2 device(s) and secondary (target) R2 device(s) become primary (source) R1 device(s).

Swaps using Dynamic SRDF perform faster but must be enabled in your Symmetrix system configuration to use this feature.

Dynamic swap is not supported under the following conditions:

- ◆ If Enginuity Versions 5567 and 5568 are mixed between your local and remote Symmetrix system
- ◆ In Enginuity Version 5669 or higher configurations where the R2 device is larger than the R1 device
- ◆ For SRDF/Asynchronous devices
- ◆ When using ESCON RA directors configured with FarPoint
- ◆ For RAID-S, RAID-R, and Parity RAID (3+1) and (7+1) protected devices

Migrating Data From R1 to a Larger R2 Device

Enginuity Version 5669 provides a temporary method for migrating data from an R1 device to a larger R2 device. Data migration to a larger R2 device is only allowed on Symmetrix DMX hardware running Enginuity 5669 or higher.

Certain SRDF operations are blocked when the R2 device is larger than the R1 device:

- ◆ Swap operations are not allowed if the R2 device is larger than the R1 device.
- ◆ Data migrated to a larger R2 device cannot be restored back to the R1 device.
- ◆ Striped metadevices are supported.
- ◆ Concatenated metadevices are not supported with restrictions. For information on these restrictions, refer to the *Solutions Enabler SRDF CLI Product Guide*.

SRDF/Asynchronous Operations

This chapter provides a technical overview of SRDF/Asynchronous (SRDF/A). The following topics are covered:

- ◆ SRDF/Asynchronous Overview 4-2
- ◆ Ensuring Data Consistency: The Dependent Write Principle 4-4
- ◆ How Does SRDF/A Work? 4-6
- ◆ Ensuring Data Consistency: The Cycle Switch Process 4-10
- ◆ SRDF/A States 4-14
- ◆ SRDF/A Mode Transitions and Recovery 4-15
- ◆ SRDF/A Recovery Scenarios 4-18
- ◆ SRDF/A Multi-Session Consistency 4-20
- ◆ SRDF/A Multi-Session Recovery 4-25
- ◆ SRDF/A Multi-Session Failure Scenarios 4-28

SRDF/Asynchronous Overview

Beginning with Enginuity Version 5670, Symmetrix DMX system supports the SRDF/Asynchronous (SRDF/A) product offering. SRDF/A provides a point-in-time image on the target (R2) device, which is only slightly behind the source (R1) device. Data is transferred to the remote Symmetrix system in predefined timed cycles called *delta sets*, which eliminates the redundancy of same track changes being transferred over the link.

SRDF/A provides a long-distance replication solution with minimal impact on performance. This protection level is intended for customers requiring minimal host application impact, who need to maintain a restartable copy of data at the R2 site, which preserves data consistency within the database. In the event of a disaster at the R1 site or if SRDF links are lost during data transfer, a partial delta set of data is discarded. However, a dependent write consistent point-in-time copy of data is retained on the target side.

SRDF/A Benefits

SRDF/A provides the following features and benefits:

- ◆ Supports extended data replication with database consistency
- ◆ Promotes efficient link utilization resulting in lower link bandwidth requirements
- ◆ Maintains a consistent point-in-time image on the secondary (R2) devices at all times
- ◆ Supports all current SRDF topologies (ESCON, FarPoint, point-to-point and switched fabric Fibre Channel, and GigE)
- ◆ Requires no additional hardware, such as switches or routers
- ◆ Supports most open systems hosts and data emulation types supported by the Symmetrix system (FBA and CKD)

For more information on supported hosts, refer to the *EMC Support Matrix* at <http://emc.com> or <http://powerlink.emc.com>, or contact your EMC Sales Representative.

- ◆ Imposes minimal impact on the back-end disk directors
- ◆ Provides a performance response time equivalent to writing to non-SRDF devices
- ◆ Allows failover and failback capability between the R1 and the R2 sites

Requirements and Limitations

Known requirements and limitations for this release of SRDF/A are as follows:

- ◆ When using either full volume Snaps (clones) or EMCSnap™, special considerations must be followed. Refer to the *Solution Enabler SRDF CLI Product Guide* or the *SRDF Host Component for z/OS Product Guide*.
- ◆ SRDF/A is supported for open systems through CLI only. Mainframe interface is supported through SRDF Host Component V5.2 and higher.
- ◆ SRDF/AR (previously known as SAR) will not support SRDF/A devices.

For more information on specific restrictions, refer to the *Solutions Enabler Symmetrix SRDF CLI Product Guide* and the *Symmetrix SRDF Host Component Product Guide*.

Ensuring Data Consistency: The Dependent Write Principle

Any asynchronous remote replication solution, be it host-based or storage-controller based, must ensure that data at the remote site is always consistent so as to enable restartability *and* maintain data integrity.

To achieve this objective, asynchronous replication solutions must honor the logical dependency between write I/Os that are embedded in the I/O stream by the logic of an application, operating system, or database management system. This ensures *dependent write consistency* and is the basic principle behind all of EMC's Consistency Technology solutions (Consistency Group, Consistent Split, Consistent Snap, and SRDF/A).

Understanding this concept is key to understanding how SRDF/A is able to provide consistent data at the remote site at all times.

The definition of a dependent write is a good place to begin understanding how SRDF/A is able to guarantee data consistency. A dependent write is defined as: ***a write I/O that will not be issued by an application until some prior related write I/O has completed.***

A real-world illustration of this concept is the relationship that a Database Management System (DBMS) creates between the I/O that updates a database file and the I/O that writes to the log of the DBMS. No DBMS will ever write changes to the database itself until after it has first written those updates to its log. In this case, the database write is the *dependent* write and the log write is the *predecessor* write. Without this dependent write logic no DBMS could restart with data integrity in the event of a power failure of the CPU on which it was running.

As an example, assume two write I/Os *A* and *B* have a relationship in the application such that I/O *B* will not be issued until I/O *A* has been successfully completed.

Ensuring dependent write consistency means that no failure in the remote replication infrastructure can create the condition where I/O *B* has been replicated to the remote site, but I/O *A* has not been replicated. Such a *missing sequence* condition would destroy the integrity of the data at the remote site.

There are several ways to ensure that this does not occur and that consistency of the remote site is preserved. These are:

- ◆ Ordering write I/Os based on timestamps
- ◆ Ordering write I/O based on sequence numbers
- ◆ Honoring dependent writes

SRDF/A approaches preserving dependent write consistency by honoring the dependent write relationships imbedded in the I/O stream by the application. This makes it an implicitly enterprise-wide, heterogeneous host solution.

The relationship between dependent I/Os is not one of *timing*, but one of *logic*. By respecting the logical relationships between I/Os there is no need for any type of temporal references (timestamps or sequence numbers) to create the consistency. The Symmetrix system has no way of knowing which specific I/Os it receives have dependent relationships, so it must assume that any I/O *B* that arrives after another I/O *A* has been acknowledged as complete to the host, must in fact be dependent on I/O *A*. Therefore, it must ensure that I/Os *A* and *B* are either in the same cycle or that I/O *B* is in a cycle that is processed to the remote side after the cycle that contains I/O *A* is processed.

It is really these *cycles* of I/Os, rather than the I/Os themselves, that are being ordered by SRDF/A. The key to preserving the consistency of data within SRDF/A is to ensure that dependent write relationships are honored during the *cycle switch* process, the frequency of which also determines the *data lag* experienced by the remote site (that is, the amount of data that would be lost in the event of a primary site disaster). The cycle switch process is described in more detail in *Ensuring Data Consistency: The Cycle Switch Process* on page 4-10.

How Does SRDF/A Work?

Symmetrix Ordered Write Processing

Different from traditional ordered write approaches, Symmetrix systems implement asynchronous mode host writes from the primary (source) to the secondary (target) using predetermined timed cycles (delta sets). Each delta set contains groups of I/Os for processing, which are managed for consistency by the Enginuity operating environment. Using cycles of operation, SRDF/A transfers sets of data, one cycle at a time, between the R1 and the R2. If the same track is written to more than one time within an active cycle on the R1 side, SRDF/A sends the update over the SRDF link only once. This approach lowers the required link bandwidth as compared with other ordered write processing approaches, which transfer each and every write separately. Refer to Figure 4-1 for a depiction of the SRDF/A delta sets.

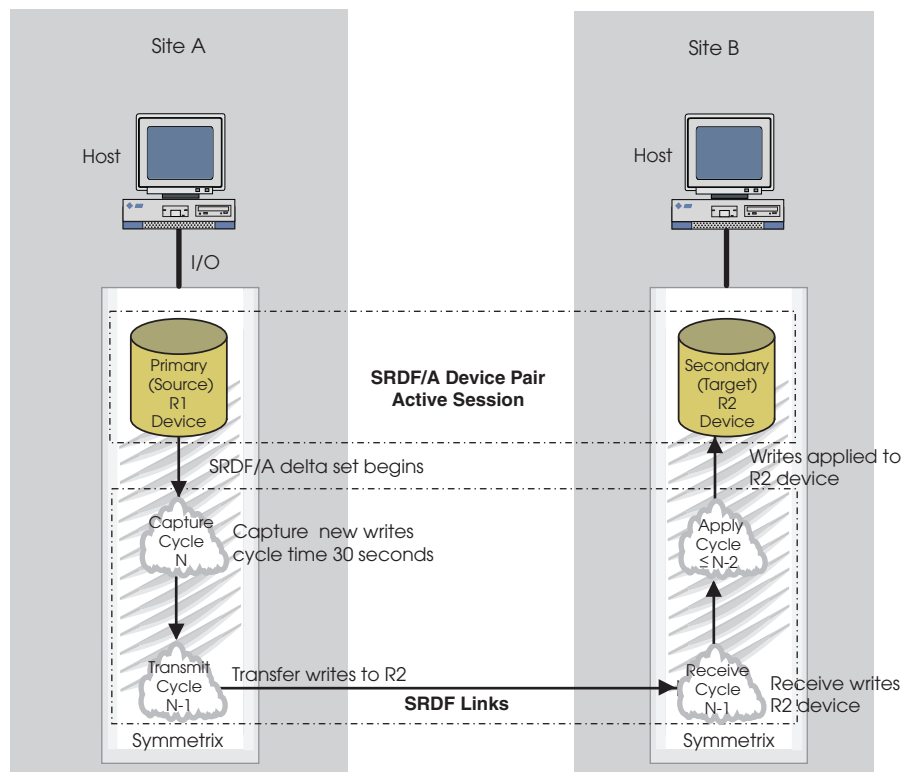


Figure 4-1 SRDF/Asynchronous

Dependent Write Consistency

Dependent write consistency is achieved through the processing of three ordered SRDF/A delta sets (cycles) between the primary (source) R1 and the secondary (target) R2 (Figure 4-1). Dependent write consistency ensures that all writes to the R2 are processed in sequential cycles to maintain a consistent copy of data between the R1 and the R2.

When the first SRDF/A cycle (N) is active, it collects any new writes into the capture cycle on the R1, overwriting any duplicate data intended for data transfer over the link. The cycle is active for a predetermined minimum amount of time (which the EMC Customer Engineer can configure on the Symmetrix system). The default time is 30 seconds. After the minimum time has been reached, the cycle becomes the transmit cycle ($N-1$) and begins transferring the delta set over the SRDF link into the receive cycle ($N-1$) of the Site B (Figure 4-1). The new cycle N then begins collecting new writes.

In receive cycle ($N-1$), the delta set is temporarily collected on the R2 side. When the transmit cycle ($N-1$) has finished transferring data into the R2, the receive cycle data becomes the apply cycle ($N-2$) and marks the cycle data as write-pending to the R2 devices. The data is considered committed to the R2 devices at cycle switch time.

One delta set is dependent upon the other for achieving write consistency. No cycle can begin until the prior one has completed. All data is transferred at the block level.

The cycle is elongated if the write transfer exceeds the set minimal cycle time.

SRDF/A Data Flow Description

At their most basic level, continuous asynchronous copy solutions are data flow control problems. Data flows into the system at a certain rate (the write I/O bandwidth), where it is buffered (creating an additional global memory requirement) and sent out the intercontroller links (outgoing bandwidth) to the remote controller where it is buffered again (additional global memory) before being written to the remote disk. Keeping these resources in balance is the key to a successful implementation.

A key difference between SRDF/A and other approaches is that less data is transferred and the data is handled fewer times. By exploiting the locality of reference phenomenon within an application, data that is updated multiple times in the same cycle is sent across the SRDF/A links only once, and does not have to be duplicated within the global

memory of the sending control unit to preserve data consistency. This is illustrated in Figure 4-2.

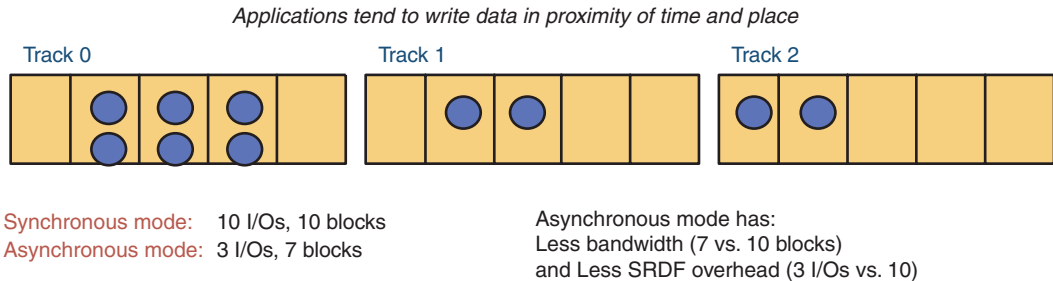


Figure 4-2 SRDF/A Locality of Reference Example

The SRDF/A flow control model, at any single point in time, has three cycles (capture, transmit/receive, and apply) working to achieve the movement of data to the remote site. SRDF/A also handles the data much less than alternative solutions because there is no copying of the data within the global memory for each update as is done in ordered write solutions.

At the primary (R1) site, the *capture* cycle is collecting all new writes and tagging them as belonging to cycle *N*. There is also a *transmit* cycle *N-1* which is not receiving any new data, but is transferring the data it had collected when it was the active cycle to the remote side. The capture cycle switches roles from *capture* to *transmit* during the cycle switch process and a new capture cycle is created. This process, which is driven by the Remote Adapters (RAs), is initiated by the primary side. These *cycles* are physically implemented as a collection of pointers to global memory slots.

At the remote (R2) site, there is also a *receive* cycle (*N-1*), which is receiving data from its partner *transmit* cycle in the primary site and placing it in global memory. The *apply* cycle (*N-2*) at the remote site is marking all the tracks from a previous cycle as write-pending to the secondary devices and then lets the normal disk adapter (DA) algorithms destage the data to the disk.

The switching between Delta Sets on the R2 side is driven by the R1 side issuing a *commit* message to the RAs on the R2 side. This happens once the minimal Delta Set time has elapsed and it has finished sending all the data in its Transmit Delta Set.

The switching from the *receive* to the *apply* cycle on the R2 side is driven by the R1 side issuing a *commit* message to the RAs on the R2 side, once the minimal cycle time has elapsed and it has finished sending all the data in its transmit ($N-1$) cycle.

Once the Delta Set switch is complete on the R2 side, a message is sent back to the R1 side completing the loop and enabling the R1 side to perform its Delta Set switch if all conditions for this switch are met.

The frequency of these cycle switches can be set between five and 30 seconds.

Also, it is important to note that the unit of transfer for each update is at the block level, just as in SRDF/Synchronous mode. It is not at the track level, except for resynchronization scenarios where SRDF/A was suspended and tracks have been marked invalid for later resynchronization.

Ensuring Data Consistency: The Cycle Switch Process

SRDF/A ensures data consistency through the cycle switch process. In a single Symmetrix SRDF/A implementation this consistency must be assured across all volumes in the SRDF group. Since SRDF/A is ensuring the consistency of the data, there is no need for the Consistency Group for z/OS product or PowerPath for open systems to be installed for SRDF/A devices.

Dependent write consistency is ensured within SRDF/A by the host adapter obtaining the active cycle number from a single location in global memory and assigning it to each I/O at the *beginning of the I/O* and retaining that cycle number *even if a cycle switch occurs during the life of that I/O*.

This results in the cycle switch process being *atomic for dependent write sequences*, even though it is not physically an atomic event across a range of volumes. As a result, two I/Os with a dependent relationship between them will either be in the same cycle, or the dependent I/O will be in a subsequent cycle.

The cycle switch algorithm will *not* allow a dependent I/O *B* to be placed in cycle *N-1* when its predecessor I/O *A* is placed in the new cycle *N*. This is because the predecessor I/O *A* will retain its assigned cycle number, even if the active cycle number increments *during A's* execution. This cycle number retention will necessarily result in *B* being assigned to the new cycle *N*. (This is because *B* is only issued by the application after *A's* execution is completed. If the cycle switched during *A's* execution, then *B* must be assigned to the new cycle.)

The new cycle *N* will then be sent to the remote site only after cycle *N-1* has been successfully transmitted, thus preserving dependent write consistency. In no case will *B* ever be transmitted to the remote side ahead of *A*.

Beginning with Engenuity level 5670.56.53, SRDF/A is supported in configurations where there are multiple Symmetrix systems on the R1 side and at 5671 multiple SRDF/A groups are supported. Initially this support is available only for z/OS environments and requires an equal number of Symmetrix systems, each containing a single SRDF/A group, on each side of the configuration.

Achieving data consistency across multiple SRDF/A groups simply requires that the Delta Set switch process described earlier be coordinated among the participating Symmetrix systems, and that

the switch occur during a very brief time period when no host writes are being serviced by the subsystems.

Achieving this requires a single coordination point to drive the cycle switch process in all participating Symmetrix systems. This function is provided by a z/OS host running through ResourcePak Base, version 5.3, called SRDF/A Multi-Session Consistency (MSC) that is managed and controlled by the SRDF Host Component V5.3 (or later) on the mainframe, and through Solutions Enabler in Open Systems environments (new in 5x71). This capability is available in the SRDF/Consistency Groups option.

SRDF Options and Features

Multiple SRDF/A Groups per Symmetrix

Enginuity 5671 expands SRDF/A to allow multiple (up to 64 depending on configuration) SRDF/A groups per Symmetrix array. Implicit in this feature is the ability to define a source for one SRDF/A group and the target to another SRDF/A group in the same Symmetrix DMX. By doing this, you now have the ability to use SRDF/A with bi-directional operations.

Bi-directional operation within a single SRDF/A group is not supported, however bi-directional operation is supported between the same systems in a SRDF/A pair by using multiple SRDF/A groups. All source-to-target data flow operations within a group still remain unidirectional.

Concurrent SRDF/A and SRDF/S from a Single Source Volume

This feature provides the ability to replicate a group of devices in Synchronous mode to one target site and in Asynchronous mode to a secondary extended distance site. The same source device can be replicated synchronously using SRDF/Synchronous mode down one link and asynchronously using SRDF/Asynchronous mode down another link.

Concurrent operations from a single source device, one leg running SRDF/S to a local/regional site and the other running SRDF/A to a far remote site, is new with the Enginuity 5671 release. Performance will be equivalent to a conventional SRDF/S configuration; the addition of the SRDF/A leg does not impose additional performance restrictions.

This feature enables limited multi-site protection capability. In the event of a primary site failure, there would be no data loss since it is synchronously replicated to a local/regional recovery site. In the event of a regional disruption, data has been simultaneously replicated to a recovery site much further away.

The limitation to this solution is that in the event of a loss of the primary site, establishing remote replication between the remaining sites would mean configuration changes and a full synchronization process. To avoid this limitation, EMC has developed SRDF/Star (Refer to the next section, *SRDF/Star*).

SRDF/Star

Available at Engenuity level 5671, SRDF/Star provides advanced multi-site business continuity protection for mainframe environments. It enables concurrent SRDF/S and SRDF/A operations from the same source volumes with the ability to incrementally establish an SRDF/A session between the two remote sites in the event of a primary site outage—a capability only available through SRDF/Star software.

This capability takes the promise of concurrent synchronous and asynchronous operations (from the same source device) to its logical conclusion. SRDF/Star allows you to quickly re-establish protection between the two remote sites in the event of a primary site failure, and then just as quickly restore the primary site when conditions permit.

With SRDF/Star, enterprises can quickly resynchronize the SRDF/S and SRDF/A copies by replicating only the differences between the sessions—allowing for much faster resumption of protected services after a source site failure.

For more information on SRDF/Star, refer to Chapter 5, *SRDF/Star Operations*.

SRDF/A to SRDF/S Mode Change Feature

This feature allows the capability to switch to and from SRDF/S to SRDF/A while ensuring dependent write consistency at the R2 side throughout the process. This capability offers those currently using SRDF/S the option of switching into SRDF/A during high workload periods to minimize performance impacts to applications. It also allows switching in reverse from SRDF/A to SRDF/S (for example, to catch up tracks owed) before then reverting to SRDF/A again.

Dynamic SRDF with SRDF/A

This feature allows adding and removing Dynamic SRDF groups that participate in SRDF/A, as well as adding and removing devices from SRDF/A groups. A concurrent mirror may be added dynamically to a device that is in an SRDF/A group subject to the normal restrictions for dynamic concurrent SRDF. However, only one mirror per device is allowed to be in an SRDF/A group.

SRDF/A States

SRDF/A can be in one of three logical states:

Active: System is running in asynchronous mode

Inactive: Devices run in their basic mode, that is sync/semi-sync/adaptive copy write-pending, or adaptive copy disk mode

NR (Not Ready): Devices are NR on the link, no data is moving to the R2 side

When running in asynchronous mode (*active* state), the R2 side can be either consistent or inconsistent. SRDF will attempt to make the R2 devices consistent as fast as possible. It will declare the R2 consistent once all the updated blocks for remote mirrors of any R1 devices have been transferred to the R2, and the last cycle containing any track resynchronization data is fully copied to global memory and in the N-2 cycle on the R2 side. Figure 4-3 shows the allowed state transitions for SRDF/A.

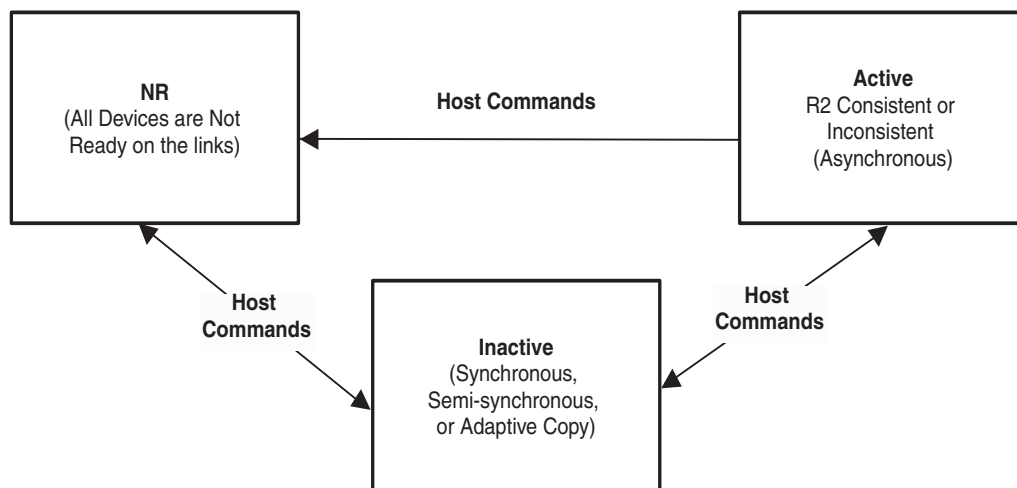


Figure 4-3 SRDF/A Allowed State Transitions

SRDF/A Mode Transitions and Recovery

This section provides a general overview of how SRDF/A moves between the states discussed in *SRDF/A States* on page 4-14.

For specific open systems commands, refer to the *Symmetrix SRDF CLI Product Guide*. For mainframe commands, refer to the *SRDF Host Component for z/OS Product Guide*.

NR State (System Startup)

When an SRDF environment is configured, and the SRDF links come up, all SRDF volumes are in not ready (NR) state by default. This means that all the remote devices on the R1 side are not-ready on the SRDF link. The user can then make all devices ready on the link by issuing a command from the host software. By doing so, the user is switching to the *inactive* SRDF/A state.

Inactive State

In *inactive* state devices are ready on the link, SRDF/A is *inactive*, and all devices work in their basic modes (synchronous, semi-synchronous, or adaptive copy write-pending/disk mode). If SRDF is in the active state, the user may choose to put SRDF/A in an inactive state, by way of a DEACTIVATE command, but doing so will compromise the consistency of the data on the R2 side.

Switching to SRDF/Asynchronous Mode

Through host-based SRDF control software, the user can switch the R1 Symmetrix to SRDF/Asynchronous mode.

SRDF/A is an SRDF group-level feature, meaning that all devices assigned to an SRDF group that is configured to operate in asynchronous mode will operate in asynchronous mode when the SRDF/A state is *active*. Also, only one SRDF/A group of devices may be configured within each Symmetrix system, and the SRDF/A group must be configured for unidirectional SRDF operation.

SRDF verifies that all the devices are ready, and then moves the system to *active* state (on both Symmetrix systems). As a result, cycles are established on the R1 and R2 sides, and the SRDF/A mechanism is enabled.

If the devices were in sync and running SRDF synchronous mode prior to activating SRDF/A, the R2 side is by definition consistent.

If there were invalid tracks to be copied, there is no consistency at the R2 side until the last invalid track has been sent to the remote side by way of the cycle switch mechanism and is in the N-2 cycle. In fact, devices operating in sync, semi-synchronous, or adaptive copy mode will automatically switch to asynchronous mode when SRDF/A is made active, and any outstanding write-pending tracks or invalid tracks marked as a result of adaptive copy mode skew values will be merged into SRDF/A cycles over time (this is similar to how invalid tracks are sent to the R2 side normally).

Once the cycle containing the last invalid track is fully on the R2 side, and in the N-2 cycle, SRDF/A will indicate that the R2 is consistent. The user can determine the consistency state of the R2 side at any time using a host software query command.

Transition from Adaptive Copy WP Mode to SRDF/A

When devices are in adaptive copy write-pending mode, it is helpful to allow them to transition into asynchronous mode without moving first to synchronous mode. When SRDF/A is activated, all adaptive copy write-pending devices are moved into the adaptive copy pending off mode. With SRDF/A active, when the DA scans a device in the pending off mode, rather than creating a separate SRDF queue record, it adds the slot to the active cycle. When there are no slots left write-pending to the SRDF mirror that are not in the SRDF/A cycle, the device can transition out of the pending off mode. Once out of pending off mode, two cycle switches are required for R2 side to be consistent.

Transition from Adaptive Copy Disk Mode to SRDF/A

Transition into SRDF/A mode from adaptive copy disk mode is immediate. Invalid tracks “owed” to the R2 device as a result of adaptive copy disk skew are scheduled as resynchronous operations. SRDF/A will produce a consistent image of data at the R2 after all re-sync operations are complete and two cycle switches have occurred.

SRDF/A to Adaptive Copy

This feature is intended to help the transition between SRDF/A and Adaptive Copy. In previous revisions of SRDF/A, this transition was accomplished by issuing two separate host commands, one to transition the SRDF mirror to synchronous mode, and a second to transition the mirror to Adaptive Copy. This new feature allows the two-stage transition to occur within the Enginuity code as the result of a single host command.

SRDF/MC is available only within single SRDF/A sessions.

Active State: System Working in Asynchronous Mode, and R2 is Consistent

The *active state* is the normal running state of SRDF/A. The R2 always represents a dependent write-consistent image of the data.

Dropping Out of Active State

SRDF/A supports several methods of dropping out of active state into Not Ready state. One option, called DROP, puts the devices in NR state immediately. This results in tracks being converted to invalids on both the R1 and R2 sides of the SRDF relationship. Resuming SRDF would then require resolving the invalids in the normal way via SRDF Host Component for z/OS *REFRESH* and *RFR-RSUM* commands or Solutions Enabler *symrdf establish* commands. (Remember that dropping out of asynchronous mode does not compromise the consistency of the data on the R2. SRDF/A always provides a consistent image of data at the remote site at all times. It is only during a resynchronization activity that data consistency will be compromised, but this can be avoided by using TimeFinder to preserve a copy of the R2 prior to the resynchronization.)

Another option, called PEND-DROP, puts the devices in NR state only at the end of the current *in-process* cycle. Write-pending tracks in the active cycle will be converted to invalids on the R1 side only. By dropping on a cycle boundary, there will be no need to first resolve invalids upon resuming SRDF/A.

Deactivating SRDF/A

SRDF/A mode also offers the option of moving out of active state but leaving the devices ready on the SRDF link, and by default, in their last primary mode of operation (synchronous or semi-synchronous). Since the devices are left ready on the SRDF link, data will continue to flow and the consistency of the data at the R2 side will be compromised.

The actual command syntax to control SRDF/A will be different between the SRDF Host Component for z/OS and the Solutions Enabler commands for open systems (as is true today for normal SRDF control), but the behavior and functionality offered will be the same. For specific open systems commands, refer to the *Symmetrix SRDF CLI Product Guide*. For mainframe commands, refer to the *SRDF Host Component for z/OS Product Guide*.

SRDF/A Recovery Scenarios

Temporary Link Loss

If SRDF/A suffers a temporary loss (<10 seconds) of all links the SRDF/A state will remain active and data will continue to accumulate in global memory. This may result in an elongated cycle, but the remote consistency will not be compromised and the R1-R2 relationships will not be suspended. The amount of time SRDF waits until it declares a link loss permanent is configurable (between 0 and 10 seconds).

Permanent Link Loss

If SRDF/A experiences a permanent link loss, it will drop all the devices on the link to *not ready* state. This will result in all data in the active and inactive R1 side cycles being changed from *write-pending* for the remote mirror to *invalid* on the remote mirror. In addition, any new work that enters the Symmetrix system will result in tracks being marked *invalid*. At the R2 side, the inactive cycle tracks are marked *invalid* on the remote mirror, and the active cycle data completes its commit to local devices.

When the links are restored, normal SRDF recovery procedures are followed. The track tables are OR'd using normal host recovery procedures (REFRESH and RFR-RSUM on the SRDF Host Component for z/OS, or `symrdf establish` in Solutions Enabler for open systems). The data is then resynchronized by sending over the invalid tracks as part of the SRDF/A cycles.

The data on the R2 volumes is always consistent in SRDF/A, even when the links have failed. However, the act of starting a resynchronization activity between the R1 and the R2 will temporarily stop the consistency of the R2 data until the resynchronization is fully completed. For this reason, it is recommended that the user employ a method for preserving the data on the R2 volumes prior to commencing resynchronization. This may be done, for example, using TimeFinder Business Continuance Volumes.

Primary Symmetrix Global Memory Full Condition

It is possible that an imbalance may occur within SRDF/A between the incoming write I/O workload and the outgoing SRDF/A bandwidth such that the global memory in the primary Symmetrix system becomes *full*. The inactive and active cycles consume all the available write memory in the Symmetrix system.

In this situation, the user has a choice as to how SRDF/A will behave:

- ◆ The Symmetrix system can throttle the host at the speed of the links, and keep SRDF/A running. In this case the performance will be equivalent to synchronous mode.
or
- ◆ The Symmetrix system throttles the host for a user specified period of time, and if the condition has not resolved itself at the expiration of this time, then SRDF/A is dropped. The default behavior is to drop SRDF/A immediately when this condition occurs.

To avoid a Memory Full condition, the SRDF/A environment must be properly designed and configured. Factors and variables that may cause an imbalance in the SRDF/A environment may include bandwidth, global memory, and workload allocation for the specific implementations

In an effort to assist users to better manage global memory full conditions, Engenuity 5671 provides a method for limiting the amount of write pending slots (i.e. cache) available to SRDF/A. Currently, SRDF/A cycles can grow until they reach the system write pending limit, at which point you can choose to throttle the host for a given amount of time or to drop SRDF/A immediately. Without this feature, if write pending limits are being exceeded, performance suffers across the entire Symmetrix, not just the SRDF/A application.

In addition, SRDF/A now allows users to associate a priority with each SRDF/A session within a Symmetrix system. As cache resources approach maximum limits sessions are dropped and cache resources are released, according to each session's priority.

Failback from R2

In the event that a disaster occurs on the primary site, the data on the R2 devices represents a dependent write consistent image of data that can be used to restart an environment with minimal data loss. Once the primary site has been repaired, the process for returning to the R1 side uses exactly the same methods as are used for synchronous SRDF failback.

Once the workload has been transferred back to the primary site hosts, SRDF/A can be activated and normal asynchronous mode protection can be resumed.

SRDF/A Multi-Session Consistency

Beginning with Enginuity Version 5670, SRDF/A is supported in configurations where there are multiple Symmetrix systems on the primary side. Initially this support was available only for z/OS environments and required an equal number of Symmetrix systems, each containing a single SRDF/A group, on each side of the configuration. At Enginuity level 5671, SRDF/A Multi-Session Consistency exploits SRDF/A multigroup support enabling a different number of Symmetrix systems at the primary and secondary locations. SRDF/A multigroup support permits fan-in and fan-out configurations (Figure 4-4). Further, SRDF/A is now supported in open systems, z/OS, and mixed environments.

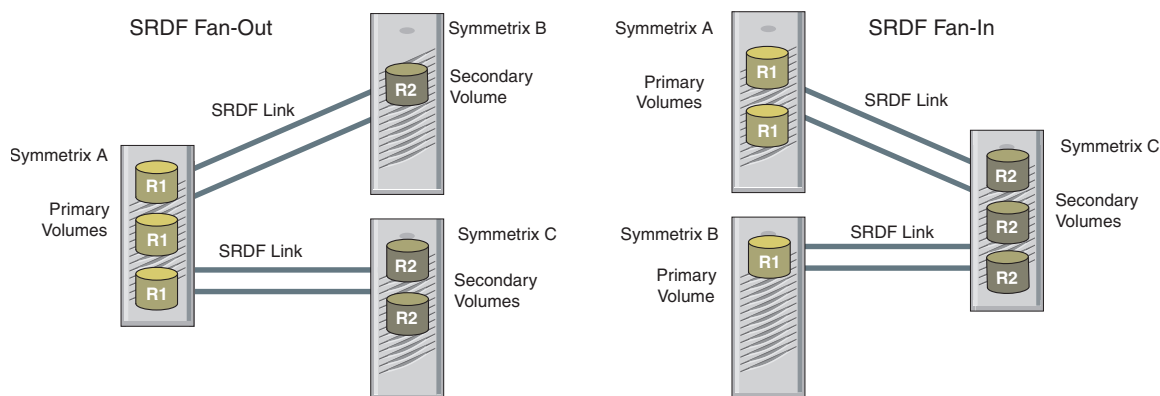


Figure 4-4 SRDF/A Multigroup Fan-In and Fan-Out Configurations

Achieving data consistency across multiple SRDF/A groups simply requires that the cycle switch process described earlier be coordinated among the participating Symmetrix systems, and that the switch occur during a very brief time period when no host writes are being serviced by the subsystems.

Achieving this requires a single coordination point to drive the cycle switch process in all participating Symmetrix systems. This function can be provided by a z/OS host running the Symmetrix Control Facility (SCF) Version 5.4 through an SCF service called SRDF/A MSC that is managed and controlled by the SRDF Host Component V5.2.1 (and later), and ResourcePak Base. In an open systems environment, cycle switching is provided by the Solutions Enabler executing a daemon that implements the MSC function.

SRDF/A Consistency Groups' Multi-Session Consistency: Operational Overview

From a single Symmetrix perspective, I/O is processed exactly the same way in SRDF/CG multi-session mode for SRDF/A, as in single-session mode.

1. The host writes to cycle N (capture cycle) on the R1 side.
2. SRDF/A transfers cycle N-1 (transmit cycle) from R1 to R2.
3. The receive cycle N-1 on the R2 side receives data from the transmit cycle, and the apply cycle N-2 restores data to the R2 devices.

During this process, the status and location of the active and inactive cycles are communicated between the R1 and R2 Symmetrix systems. For example, when the R1 finishes sending cycle N-1, it sends a special indication to R2 to let it know that it has completed the inactive cycle transfer. Similarly, when R2 finishes restoring cycle N - 2, it sends a special indication to let R1 know that its active cycle is empty.

At this point in the process, (that is, when the *transmit* cycle on the R1 side is empty and the *apply* cycle on the R2 side is empty), SRDF/A is ready for a cycle switch. In single-session mode, Enginuity would perform the SRDF/A cycle switch once all the conditions are satisfied.

However, when using SRDF/CG multi-session controls in a mainframe or open systems environment, SRDF/CG indicates its *switch readiness* to the host (mainframe or open systems), which is polling for this condition, and the switch command is issued by the host to all Symmetrix systems once they are in this state.

Entering SRDF/CG for SRDF/A Multi-Session Consistency

For the host to control the cycle switch process, the Symmetrix systems must be aware that they are running in multi-session mode. The prerequisite for entering multi-session mode is that SRDF/A must be active on the Symmetrix system.

Simply activating SRDF/A does not place it in multi-session mode, and conversely exiting multi-session mode does not drop or deactivate SRDF/A, it merely places SRDF/A in normal or *single-session* mode. However, if SRDF/A is dropped or deactivated, then multi-session mode is necessarily terminated and would need to be re-entered once SRDF/A was made active again.

SRDF/A enters multi-session mode upon receipt of a host command. As part of the command to enter multi-session mode, and with each switch command issued thereafter, the host provides a tag to each capture cycle that is retained throughout that cycle's life. This cycle tag is a value that is common across all participating SRDF/A sessions and eliminates the need to synchronize the cycle numbers across them. The cycle tag is the mechanism by which consistency is assured.

Figure 4-5 updates the SRDF/A state diagram from single mode to incorporate multi-session SRDF/A.

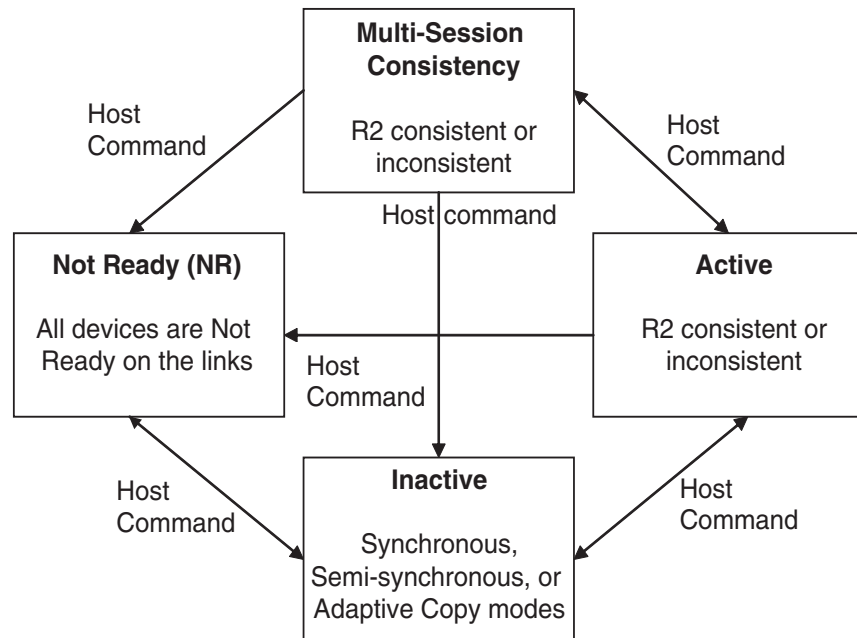


Figure 4-5 Multi-Session Consistency

Performing a Consistent Cycle Switch

Multi-session SRDF/A performs a coordinated cycle switch during a very short window of time when there are no host writes being completed. This time period is referred to as an *SRDF/A window*.

When the host discovers that all systems are ready for a cycle switch, it issues a single command to each Symmetrix system that performs a cycle switch to open the SRDF/A window. The SRDF/A window is implemented as a bit in the SRDF/A state table in global memory where the cycle number is also stored.

The window is implemented as a bit in the SRDF/A state table in global memory where the cycle number is also stored. This table is accessed by the host adapter to obtain the cycle number at the start of each write in single-session mode. In multi-session mode, the host adapter also checks this bit, and if it is on (an open window), it disconnects from the channel and begins polling the bit to see if the host has closed the window. While the window is open, any I/Os that start will be disconnected, and as a result no dependent I/Os will even be issued by any host to any devices in the multi-session group.

The SRDF/A window remains open on each Symmetrix system until the last Symmetrix system in the multi-session group acknowledges to the host that the switch and open command has been processed and a close command has been received. As a result, dependent write consistency across the SRDF/A multi-session group is created, and once the SRDF/A window has been opened and the cycle has successfully switched on all Symmetrix systems, the SRDF/A window can then be closed by the SRDF/A MSC software, allowing all disconnected writes to complete and normal processing to resume.

Enginuity does provide a fail-safe mechanism to ensure that the window will not remain open permanently due to an SRDF/A MSC software failure. SRDF/A will close the window itself if the host has not closed it within 5 seconds.

As part of this *switch and open* operation, SRDF/A MSC assigns a cycle *tag* value to the active cycle. (Note that this cycle tag value is separate from the cycle number assigned internally by SRDF/A.) This cycle tag is carried by the SRDF/A process to the remote side and is used by the host at the recovery site to ensure that only data from the same host cycle is applied to the R2 devices in each Symmetrix system in the event of a disaster.

Once *all* Symmetrix systems have completed a cycle switch, the host issues a command to close the window (turn off the bit in the state table), and all disconnected write I/Os complete.

During this window, read I/Os complete normally to any devices or PAV aliases that have not received a write. The SRDF/A window is an attribute of the SRDF/A group and is checked at the start of each I/O, at no additional overhead, because the host adapter is already obtaining the cycle number from global memory as part of SDRF/A's existing minimal overhead.

SRDF/A Multi-Session Recovery

R1 Side Recovery

When SRDF/A is deactivated or dropped while in multi-session mode, each R1 Symmetrix system starts the same cleanup process as in single-session mode— it discards all I/O from both transmit and capture cycles and marks corresponding tracks invalid to the SRDF mirror.

The host software does not need to perform any special recovery on the R1 side.

R2 Side Recovery

When SRDF/A is deactivated or dropped while in multi-session mode, each R2 Symmetrix system finishes the restore of its apply cycle and discards its receive cycle, if it is not marked as complete. If the R2 Symmetrix system has a complete receive cycle, it puts it “on hold” in cache, awaiting a decision from the host software. The host software must issue a command to either discard or force the restore of the data in the R2 receive cycle based on the rules that follow.

This is where SRDF/A uses the host cycle tags. The host software must use its cycle tags during recovery of receive cycles on R2 systems. It uses the cycle tags to locate parts of the same host cycle scattered among all receive and apply cycles on all R2 systems.

There are three different cases to be considered when SRDF/A has been terminated with respect to the R2:

- ◆ In the first case, all receive cycles on all Symmetrix systems have the same tag and all are marked as complete. In this case, the host software may choose to either commit all receive cycles or discard all of them. The default behavior will be to commit the receive cycles to deliver the most current consistent data to the user.
- ◆ In the second case, all receive cycles on all Symmetrix systems have the same tag, but at least one Symmetrix system does not have a complete receive cycle. The host software **must** discard all receive cycles in this case. The most current consistent data is already on the R2s, and the presence of an incomplete receive cycle indicates that more current consistent data is not available. Enginuity automatically discards incomplete receive cycles.
- ◆ The third case occurs when there are difference cycle tags within the apply and receive groups. In this case, the Symmetrix systems can be divided into two groups. The first group has Symmetrix

systems with apply cycle tags that match the receive cycle tags of the Symmetrix systems from the second group. In other words, Symmetrix systems from the first group have received commit for a certain host cycle, while the Symmetrix systems from the second group have not. In this case, the receive cycles of the Symmetrix systems from the second group are **necessarily** complete and host software **must** force their restore. At the same time, host software must discard receive cycles of the Symmetrix systems from the first group regardless of their completeness.

This third, nontrivial, case is illustrated in Figure 4-6.

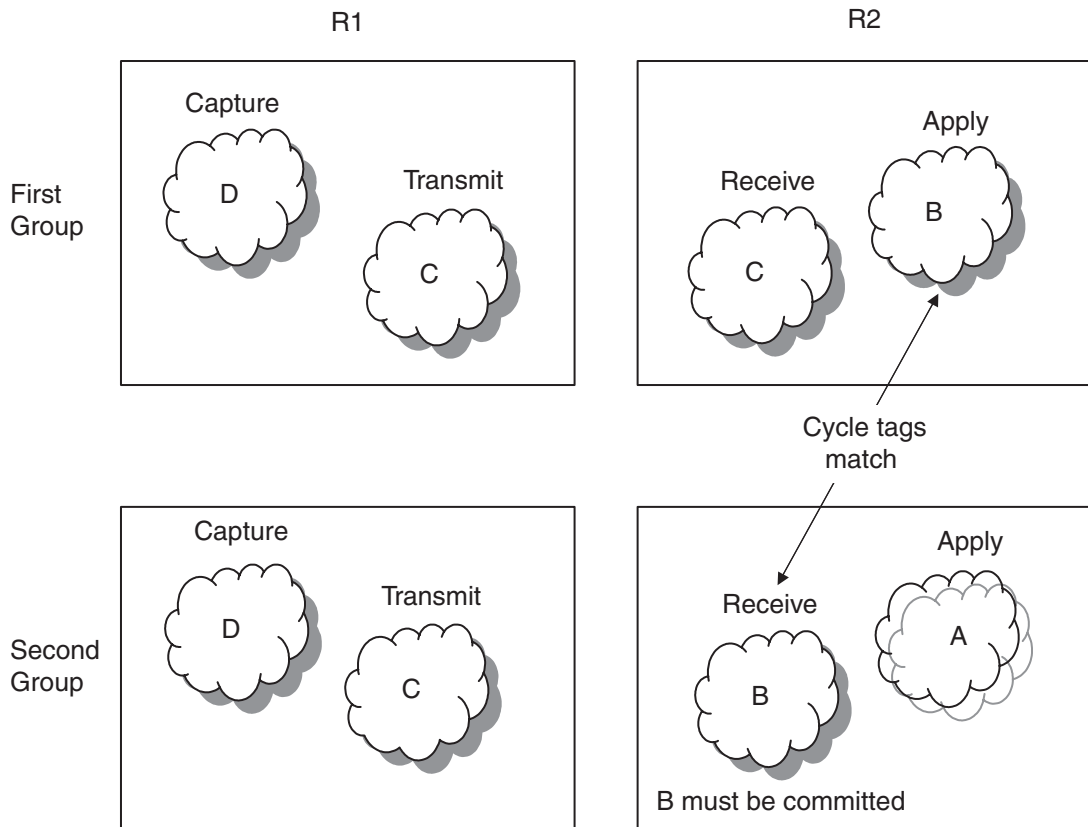


Figure 4-6 Multi-Session R2 Side Recovery

In Figure 4-6, there are two pairs of Symmetrix systems in an SRDF/A multi-session configuration. The letters A, B, C, D represent the cycle tags associated with each cycle.

The R2 Symmetrix system of the top pair received a commit command for host cycle B from its R1 partner as part of the SRDF/A cycle switch process. Cycle B was made the apply cycle and began restoring its data to the R2 devices. It also acknowledged this commit to R1 and therefore, R1 started sending host cycle C to the receive cycle on the R2.

The bottom pair never received the commit message for cycle B, so the bottom R2 Symmetrix system's apply cycle, which belongs to host cycle A, remains empty and its receive cycle, which belongs to host cycle B, is complete. *Remember that **both** R2 side Symmetrix systems must have had a complete receive cycle B and an empty apply cycle A for the cycle switch process that promoted B to the apply cycle on the top pair to even start.* Therefore, it is not possible for cycle B to be incomplete or cycle A to be nonempty on the bottom pair. Consequently, cycle B **must** be made the apply cycle and restored to the R2 devices.

Given the prerequisites to cycle switching in a multi-session configuration, the mere presence of an apply cycle number on one R2 Symmetrix system that matches a receive cycle number on another R2 Symmetrix system necessitates the restoration of the data in that receive cycle.

The functionality for determining the disposition of the apply and receive cycles on the R2 Symmetrix systems in a multi-session SRDF/A configuration is imbedded in the Multi-Session Consistency service of the Symmetrix Control Facility and is invoked by the SRDF Host Component for z/OS.

The recovery logic just described is implemented in SRDF/A MSC and will execute automatically from the host on the R1 side if SRDF links are available on all Symmetrix system pairs.

In the event of a loss of the R1 site, a stand-alone recovery utility is available to perform the recovery processing on the R2 site. This recovery utility must run from a separate z/OS system (i.e., one that is **not** part of the SRDF/A replication process) that has connectivity to all SRDF/A R2 systems. (It is also possible to run the stand-alone utility from the R1 site if SRDF connectivity is in place between all Symmetrix system pairs.)

SRDF/A Multi-Session Failure Scenarios

SRDF/A MSC Host Failure

In the event of a failure of the SRDF/A MSC function in the driving host, SRDF/A multi-session will eventually drop due to the cache full conditions that result from lack of cycle switching.

Tolerance Mode

Tolerance mode is still supported in SRDF/A multi-session mode, but it should be noted that if it is set *on* for any SRDF/A session, a service action on the R2 Symmetrix system will result in inconsistency for the entire multi-session group.

For more information on Tolerance mode, refer to the z/OS Host Component Guide.

SRDF/A Multi-Session Consistency Drop Policy

Should an individual SRDF/A session drop, for example due to a total link failure between a subset of the Symmetrix system pairs, the SRDF/A MSC's behavior is governed by a *drop policy* initialization parameter. The options available are:

- ◆ DROPALL, which is the default and tells SRDF/A MSC to drop *all* SRDF/A sessions in the multi-session group if any one session has dropped.
- ◆ REMOVE_FAILING, which tells SRDF/A MSC to remove the failing session from the actively cycling multi-session group, but SRDF/A MSC will continue cycling the remaining sessions in the multi-session group.



CAUTION

Using this option may result in inconsistent data on the secondary side for the multi-session group.

- ◆ DISABLE, which tells SRDF/A MSC to disable control of the multi-session group. The result of the disabling the SRDF/A MSC function means that the SRDF/A sessions that can continue to run will do so in the default SRDF/A single-session mode and SRDF/A MSC will no longer drive the cycle switch process across the remaining sessions. The failed SRDF/A session(s) will be left in their current state.

**CAUTION**

Using this option may result in inconsistent data on the secondary side for the multi-session group.

SRDF/Star Operations

This chapter provides a technical overview of SRDF/Star. The following topics are covered:

- ◆ SRDF/Star Overview5-2
- ◆ How SRDF/Star Works.....5-5

SRDF/Star Overview

Available at Enginuity level 5671 for mainframe environments, SRDF/Star is a solution that operates in a *concurrent* SRDF configuration (A-to-B and A-to-C) where one remote mirror operates in SRDF/S mode (A-to-B) and the other remote mirror operates in SRDF/A mode (A-to-C).

SRDF/Star provides for rapid re-establishment of cross-site protection in the event of Primary Site (A) failure. Rather than a full resynchronization between sites B and C, SRDF/Star provides a differential B to C synchronization, dramatically reducing the time it takes to remotely protect the new production site. SRDF/Star also provides a mechanism to determine which site (B or C) has the most current data in the event of a rolling disaster that affects site A. In all cases, you maintain the ability to choose both which site to operate from and which site's data to use when recovering from a primary site failure. Figure 5-1 shows a SRDF/Star configuration. It is strongly recommended that all SRDF devices be locally protected and that there is capacity allocated for one replica (BCV, Snap, or clone) in at least one of the remote sites in the configuration.

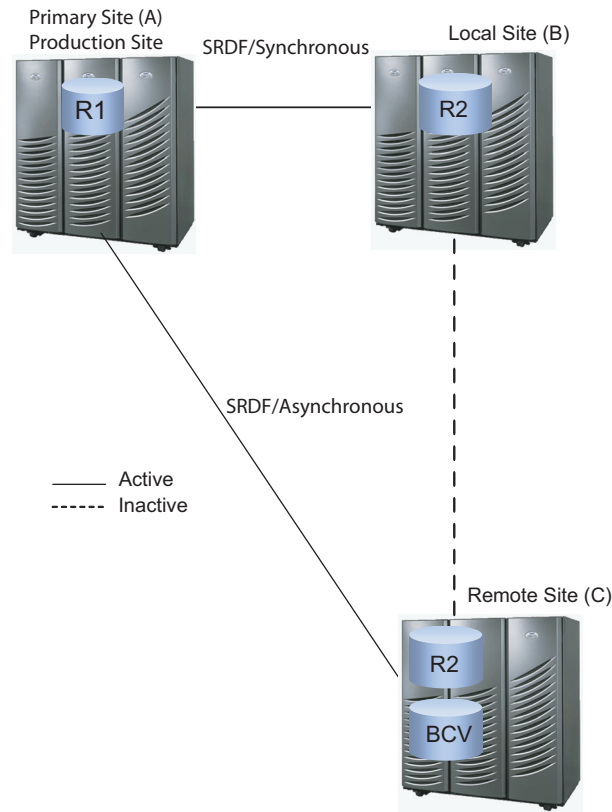


Figure 5-1 Concurrent SRDF Configured for SRDF/Star Support

The concurrent configuration option of SRDF/A offers the ability to restart your environment at long distances with minimal data loss, while simultaneously providing a zero data loss restart capability at a local site. Such a configuration provides protection for both a site disaster and a regional disaster while minimizing performance impact and loss of data.

In a concurrent SRDF/A configuration *without* the SRDF/Star functionality, the loss of the primary A site would normally mean that the long distance replication would stop and data would no longer propagate to the C site. Data at C would continue to age as production was resumed at site B. Resuming SRDF/A between sites B and C would require a full resynchronization to re-enable disaster recovery protection. This is both a time and resource consuming process.

SRDF/Star avoids this full resynchronization by allowing a B-to-C (or C-to-B) resynchronization to be done differentially using host software commands and procedures.

In addition, SRDF/Star allows the SRDF personalities of sites A and Site B to be swapped and the SRDF/A relationship to be transferred to Sites B and C during planned outages with no data movement at all.

SRDF/Star Benefits

SRDF/Star provides the following features and benefits:

- ◆ The ability to maintain business continuance despite the loss of any site in a three-site configuration
- ◆ The ability to resume asynchronous protection between the two secondary sites, *with minimal data movement*, in the event of a primary site failure

Known Requirements and Limitations at this Release

Known requirements and limitations for this release of SRDF/A are as follows:

- ◆ All SRDF/Star SRDF device pairs must be the same geometry and size.
- ◆ Only Fibre or GigE SRDF links are supported.
- ◆ All SRDF Groups (even inactive groups) must be defined prior to entering SRDF/Star mode.
- ◆ SRDF/Star is defined by the MSC Group Definition:

The R2 site changes as a function of Site switchovers.

For more information on specific restrictions, refer to the *Solutions Enabler Symmetrix SRDF CLI Product Guide* and the *Symmetrix SRDF Host Component Product Guide*.

How SRDF/Star Works

The SRDF/Star solution applies to a three Site configuration (A, B, C) providing for differential resynchronization between sites Site B and C and a restart at either site.

- ◆ Normal operation - This refers to function provided by the multi-session consistency task running in the host. MSC manages and manipulates Symmetrix data structures and SRDF features to enable differential B-to-C or C-to-B resynchronization in the SRDF/Star configurations.
- ◆ Planned Switchovers - This refers to host based automation executed at a remote site to effect reconfiguration of SRDF/Star in support of planned site switchovers between the synchronous SRDF sites.
- ◆ Unplanned Failovers - This refers to host based automation executed at the remote site that achieves the reconfiguration and

Descriptions of states and events in this chapter are from the R1 point of view unless otherwise noted. Refer to Figure 5-2 on page 5-7 for all scenarios.

The initial configuration is Site A has R1 devices in a SRDF/S relationship with Site B and in a concurrent SRDF/A relationship with Site C. Consistency Group is required. SRDF/Star requires an active host at Site A. A host is also required at either Site B or Site C for recovery processing following a Site A failure Site A and Site B do not require cross Site clustering be configured for SRDF/Star. Clustering within either Site is supported. BCVs are required at Site B, and Site C. If planned switching between Site A and Site is going to need supported BCVs are also required at Site A. The BCVs retain any consistency images created either during a planned or unplanned event.

SRDF/Star provides processes and procedures to reconfigure SRDF/S and SRDF/A, provide differential resynchronization between Sites B and Site C, and manage planned and unplanned failover from Site A to Site B or Site A to Site C. Restart may occur at either Site B or Site C so resynchronization procedures are presented for each case.

The solution provides a mechanism to determine whether Site C or Site B has most current data. Data resynchronization is possible from either Site B or Site C to the other Site, and is a decision that is made

by the user. SRDF/Star also provides a mechanism to determine when current active R1 cycle (capture) contents reach the active R2 cycle (commit).

Normal operation of SRDF/Star is controlled by the host based Multi-Session Consistency (MSC) task at the R1 Site. MSC performs the session management at the SRDF/S R2 Site and when necessary at the SRDF/A Site C. MSC session management maintains the information needed to perform differential synchronization between Site B and Site C. Other host software including SRDF Host Component, Consistency Groups, support utilities, REXX automation, and documented procedures are used to accomplish resynchronization and manage the reconfigurations. Automation for some basic procedure operations is provided with SRDF/Star and is supported by EMC if unaltered.

Mainframe customers wishing to script additional suggested procedures and/or extend them could use any number of automation products including native REXX or EMCSPE.

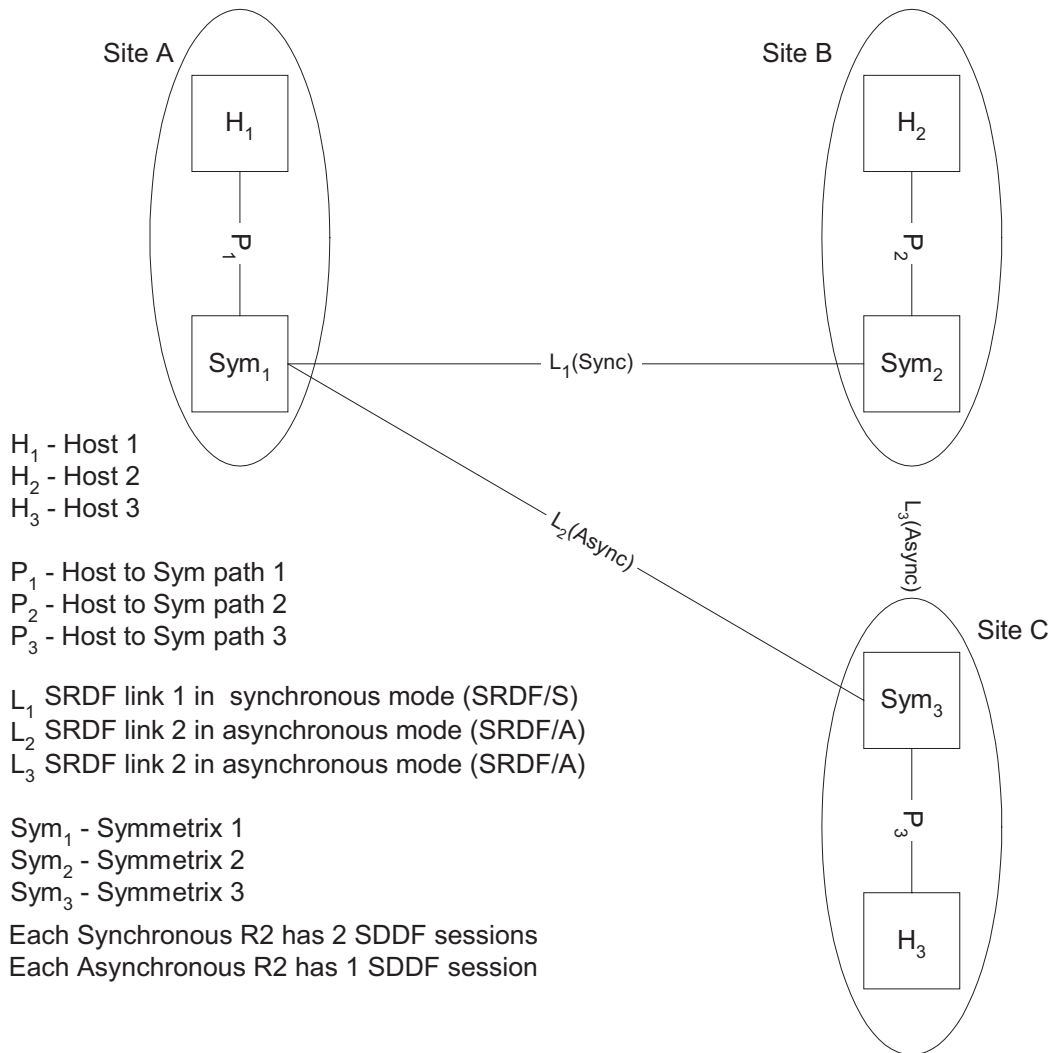


Figure 5-2 SRDF/Star Configuration Reference

SRDF/Star Automation

EMC provides host-based automation for both planned switchover actions and unplanned failover actions. This automation executes in a z/OS host and is written in REXX, using the EMC Stored Procedure Executive automation tool. Detailed descriptions and implementation guidelines can be found in *SRDF/Star for z/OS - An Implementation and Usage Primer*.

A

- adaptive copy
 - disk mode 1-4, 2-20
 - write-pending mode 1-4, 2-19
- adaptive copy mode 1-21, 1-22, 2-19
- asynchronous remote replication 4-4
- ATM links 2-10
- attributes, system-level 2-22

B

- backups without a remote host 3-11
- BCV 2-3
 - as a primary (source) device 3-9, 3-10
 - concurrent operations 3-7
 - performing remote backups 3-11
 - SRDF multi-hop 3-11
- business continuance
 - operations 3-7 to 3-11
- business continuance volumes (BCVs) 2-2, 3-9

C

- channel interface states 2-14
- Concurrent SRDF 1-21, 1-22, 2-6, 2-23
- configurations
 - SRDF 1-16 to 1-25
- connectivity
 - point-to-multipoint 1-20
 - point-to-point 1-20, 1-21
- Consistency Group 4-4
- Consistency Technology solutions 4-4
- Consistent Snap 4-4

- Consistent Split 4-4

D

- data protection options 1-11
- data recovery 2-7
- data vaulting 1-17
- Database Management System 4-4
- database management systems 1-6
- DBMS 1-6, 4-4
- delta set 4-7
- dependent write consistency 4-5, 4-7, 4-10
 - delta sets 4-7
- dependent write operations 1-6
- device states 2-12
- domino mode 2-21
 - device domino mode 2-21
 - link domino mode 2-21
- dual-directional configuration 2-10
- dual-directional link protocol 2-10
- dynamic spare 2-2, 2-3
- Dynamic Sparing 1-11, 2-7
- Dynamic Sparing with mirrored pairs 2-7
- Dynamic Sparing with SRDF 2-7
- Dynamic SRDF 2-3, 2-4, 2-6, 3-14
- dynamic SRDF devices 2-6
- dynamic SRDF groups 2-6

E

- E1, E3, T1, T3, and ATM links 2-10
- EMC ControlCenter software 1-11
- EMC SRDF/Consistency Groups 1-6

Engenuity 1-13, 1-19, 1-21, 1-22, 2-3, 2-6, 2-8, 2-11, 2-23
 ESCON 1-23, 2-4, 2-9, 2-10, 2-23
 ESCON director 1-13
 ESCON remote adapter (RA) 1-12
 ESCON remote director (RA) 1-13
 Ethernet infrastructure 1-22

F

fabric connectivity 1-19
 fully switched 1-19
 failover (secondary Symmetrix system takeover) 3-5
 FarPoint 1-21, 1-23, 1-24, 1-25
 performance impact 1-25
 preserving synchronization 1-24
 Fibre Channel 2-23
 Fibre Channel Director (RF) 1-13
 Fibre Channel directors 2-6
 Fibre Channel point-to-point connection 2-4
 Fibre Channel remote adapter (RF) 1-12
 Fibre Channel remote directors (RF) 1-13
 FICON 2-9

G

GDPS 2-9
 Geographically Dispersed Parallel Sysplex (GDPS) solution 2-9
 Gigabit Ethernet 2-23
 Gigabit Ethernet director (RE) 1-14
 GigE 1-15, 1-22
 GigE connection 2-4
 GigE remote adapter (RE) 1-12
 GigE remote directors (RE) 1-13
 GigE SRDF directors 2-6

H

host accessibility 2-14, 2-15
 host volume manager software 2-7
 hypervolume 2-7
 size 2-7

I

I/O cycles 4-5
 input/output cycles 4-5
 invalid tracks warning 2-22

L

link protocols
 dual-directional 2-10
 unidirectional 2-10
 links
 possible states 2-11
 local mirroring 1-8
 local volumes 2-3
 logical device 2-7, 2-8
 logical path 1-21
 logical paths 1-20
 logical volume states 2-12
 host accessibility 2-14
 host view 2-14
 not ready 2-12
 primary volume 2-13, 2-14
 read only 2-12
 secondary volumes 2-13, 2-14
 SRDF view 2-13
 write enabled 2-12

M

metavolumes 2-7 to 2-8
 mirroring
 local vs. remote 1-8
 mirroring (RAID Level 1) 1-11
 modes of operation
 primary 2-16 to 2-18
 secondary 2-19 to 2-22
 MPCD-GigE 1-15
 Multi-hop 3-11
 multiple RA pairs 1-13
 multiple Symmetrix logical volumes. 2-7
 Multiprotocol Channel Director (MPCD) 1-12, 1-22

N

Native IP 1-22
 not ready state
 primary volumes 2-13, 2-14
 secondary volumes 2-13, 2-14

O

Ordered write operations 4-5

P

Parity RAID 1-11, 2-4
 Parity RAID (3+1) 2-2, 2-3
 Parity RAID (7+1) 2-2, 2-3
 PAV/MA 1-25
 Peer-to-Peer Remote Copy 2-6
 Peer-to-Peer Remote Copy (PPRC) 2-5, 2-9
 Permanent Member Sparing 1-11
 PowerPath 2-25
 PPRC mode 2-5, 2-7, 2-9
 primary (source) volumes 2-2
 primary (source, R1) volumes 2-10
 primary devices 1-18, 1-19, 1-20, 1-22
 primary volume 2-2
 primary volumes 2-2, 2-10
 host accessibility 2-15
 host view 2-14
 not ready state 2-13, 2-14
 read/write state 2-13
 read-only state 2-14
 states 2-13, 2-14
 write-enabled state 2-14

R

RAID 1 1-8, 2-2, 2-3
 RAID 10 2-2, 2-3, 2-7, 2-8
 RAID 5 1-11
 RAID 5 (3+1) 2-2, 2-3
 RAID 5 (7+1) 2-2, 2-3
 RAID-S 2-2, 2-3
 read operations 3-3 to 3-4
 read/write state
 primary volumes 2-13
 secondary volumes 2-13
 read-only state 2-14
 secondary volumes 2-13

recovery operations 3-5 to 3-6
 remote mirroring 1-8

S

SAN 1-19
 secondary (target) volumes 2-2, 2-10
 secondary devices 1-18, 1-19, 1-20
 secondary Symmetrix system takeover 3-5
 secondary volumes 2-3, 2-10
 host accessibility 2-15
 host view 2-14
 not ready state 2-13, 2-14
 read/write state 2-13
 read-only state 2-13, 2-14
 states 2-13, 2-14
 semi-synchronous mode 1-4, 1-22, 2-17
 Solutions Enabler software 1-11
 SRDF
 basic configuration 1-8
 bidirectional configuration 2-10
 concurrent operations 3-7
 configurations 1-16 to 1-25
 control operations
 swap 3-14
 devices 2-4
 director hardware 1-12
 director/adaptor board sets 1-12
 dual-directional configurations 3-2
 FarPoint 1-23
 FarPoint, performance impact 1-25
 groups 2-4, 2-6
 hardware 1-12
 Host Component for z/OS 1-11
 link states 2-11
 monitoring and controlling 1-11
 Multi-hop 3-11
 over Fibre Channel 1-19
 overview 1-2
 switched topology 1-19
 Symmetrix interfamily connectivity 1-9
 topologies 4-2
 unidirectional configurations 3-2
 volumes states 2-13
 SRDF/A 1-2, 1-4, 1-22, 2-11
 active state 4-17
 adaptive copy pending off mode 4-16

- adaptive copy write-pending mode 4-16
- data consistency 4-10
- Deactivating SRDF/A 4-17
- dependent write consistency 4-10
- failback from R2 4-19
- flow control model 4-8
- Multi-session 4-20 to 4-29
- multi-session failure scenarios 4-28
- multisession mode 4-20
- multi-session recovery 4-25
- NR (not ready) state 4-15
- ordered write processing 4-6
- permanent link loss 4-18
- primary Symmetrix global memory full condition 4-19
- Recovery Scenarios 4-18 to 4-19
- resolving the invalid tracks 4-17
- resynchronization 4-9, 4-18
- SRDF/A
 - Inactive State 4-15
- SRDF/A supported hosts 4-2
- SRDF/AR 1-2, 1-5
- SRDF/Asynchronous 1-2, 1-4, 2-5, 4-2
- SRDF/Automated Replication 1-2, 1-5
- SRDF/CG 1-3, 1-6
- SRDF/Cluster Enabler 1-3
- SRDF/Consistency Groups 1-3, 2-25 to 2-28
- SRDF/Data Mobility 1-2, 1-5
- SRDF/DM 1-2, 1-5
- SRDF/S 1-2, 1-3, 2-25
- SRDF/Star 1-2
- SRDF/Synchronous 1-2
- Storage Area Network (SAN) 1-19
- swap SRDF devices 3-14
- switched SRDF topology 1-19
- switched Fibre Channel configuration 1-19
- switched Fibre Channel fabric connection 2-4
- Switched SRDF 1-19, 2-6
- switching to SRDF/Asynchronous Mode 4-15
- Symmetrix
 - RAID 10 2-8
 - Symmetrix 3XXX/5XXX systems 1-22
 - Symmetrix 8000 systems 1-13, 1-22
 - Symmetrix DMX systems 1-4, 1-13, 1-22
 - Symmetrix DMX1000, DMX2000, and DMX3000 systems 1-13
 - Symmetrix Engenuity 1-13, 1-15

- Symmetrix model numbers 1-3
- Symmetrix RAID 10 1-11
- Symmetrix Remote Data Facility
 - Refer to SRDF
- Symmetrix Remote Data Facility/Synchronous 1-3
- synchronous mode 1-22
- system-level attributes 2-22

T

- TCP connections 1-23
- TimeFinder
 - operations (establish, split, restore) 3-10

U

- unidirectional link protocol 2-10
- UNIX 1-11

V

- volume types 2-2 to 2-8

W

- WAN 1-22
- Windows 1-11
- Windows NT meta devices 2-7
- write operations 3-2
- write-enabled state 2-14

Z

- z/OS 1-11