

EMC
Symmetrix Remote Data Facility (SRDF)
and SRDF/Asynchronous

PRODUCT GUIDE

P/N 300-001-165
REV A01

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

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

Printed September, 2003

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, AutoIS, Celerra, CLARiiON, CLARAlert, DG, E-Infostructure, HighRoad, Navisphere, PowerPath, ResourcePak, SRDF, VisualSAN, WideSky, and The EMC Effect are registered trademarks and EMC Automated Networked Storage, EMC ControlCenter, EMC Developers Program, EMC Enterprise Storage, EMC Enterprise Storage Network, EMC OnCourse, EMC Proven, EMC Snap, Access Logix, AutoAdvice, Automated Resource Manager, AutoSwap, AVALONidm, C-Clip, CacheStorm, Celerra Replicator, Centera, CentraStar, CLARevent, Connectrix, CopyCross, CopyPoint, CrosStor, 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, SnapView /IP, StorageScope, SupportMate, SymmAPI, SymmEnabler, Symmetrix DMX, TimeFinder, Universal Data Tone, VisualSRM, and where information lives are trademarks of EMC Corporation. All other trademarks used herein are the property of their respective owners.

Preface	xi
Chapter 1 Introducing SRDF	
SRDF Overview	1-2
SRDF/Data Mobility	1-2
SRDF/Asynchronous	1-2
Local Versus Remote Mirroring.....	1-3
Basic SRDF Configuration.....	1-4
SRDF Supported Data Protection Options.....	1-5
SRDF Configuration Using Switched Fibre Channel	1-5
Monitoring and Controlling SRDF.....	1-6
SRDF Volume Types	1-7
Primary (Source) Volumes.....	1-7
Secondary (Target) Volumes.....	1-8
Local Volumes	1-8
Device Groups.....	1-8
Dynamic SRDF Devices	1-9
Dynamic SRDF Groups.....	1-10
Dynamic Spares	1-10
Metadevices.....	1-11
SRDF Director Hardware.....	1-12
ESCON Remote Adapters	1-12
Fibre Channel Remote Adapters	1-12
Symmetrix 8000 GigE Remote Adapters.....	1-13
Symmetrix DMX GigE Remote Adapters	1-13
SRDF Configurations.....	1-14
SRDF Interfamily Connectivity	1-16
SRDF Fully Switched Fabric Connectivity	1-17
PPRC Command Support.....	1-17

Chapter 2 **SRDF Technical Concepts**

SRDF Link Protocols	2-2
SRDF Unidirectional Link Protocol	2-2
SRDF Bidirectional Link Protocol	2-2
SRDF Dual-Directional Link Protocol	2-2
SRDF Link States.....	2-3
Primary Modes of Operation	2-4
Synchronous Mode (Real-Time or Journal 0 Mode).....	2-4
Semi-Synchronous Mode (Journal 1 Mode).....	2-5
Secondary Modes of Operation	2-6
Adaptive Copying Modes	2-6
Invalid Tracks Warning.....	2-8
SRDF System-Level Attributes	2-8
Concurrent RDF	2-9
Consistency Groups	2-10
How a Consistency Group Works.....	2-10
Continuous Processing	2-12
MVS Technical Considerations.....	2-12

Chapter 3 **SRDF Operations**

Write Operations.....	3-2
Write Operations in a Unidirectional or Dual-Directional Protocol	3-2
Write Operations in a Bidirectional Protocol	3-2
Read Operations	3-3
Recovery Operations.....	3-4
Failover (Formerly known as Target Takeover)	3-4
Returning Control to the Local Host	3-4
Recovery for a Large Number of Invalid Tracks (Open Systems Sites Only)	3-5
Business Continuation Using SRDF.....	3-6
Business Continuation Using SRDF and TimeFinder	3-8
Using a BCV as a Primary (Source) Device	3-8
Performing Remote Backups Without a Remote Host.....	3-9
SRDF Multi-Hop.....	3-10
Migrating Data From R1 to a Larger R2 Device.....	3-11
R1/R2 Swap Procedure Modification.....	3-12
Dynamic R1/R2 Swap	3-13

Chapter 4 **SRDF/Asynchronous Operations**

SRDF/Asynchronous Overview	4-2
SRDF/A Benefits	4-2
Known Requirements and Limitations at Initial Release...	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/A States.....	4-11
SRDF/A Mode Transitions and Recovery	4-12
NR State (System Startup)	4-12
Inactive State.....	4-12
Switching to SRDF/Asynchronous Mode.....	4-12
Active State: System Working in Asynchronous Mode, and	
R2 is Consistent	4-13
<i>Dropping Out of Active State</i>	4-14
<i>Deactivating SRDF/A</i>	4-14
SRDF/A Recovery Scenarios.....	4-15
Temporary Link Loss	4-15
Permanent Link Loss	4-15
Primary Symmetrix Global Memory <i>Full</i> Condition	4-16
Failback From R2.....	4-16
 Glossary	 g-1
 Index	 i-1

Figures

1-1	Basic SRDF Configuration	1-4
1-2	Switched SRDF With Multiple Primary and Secondary Devices	1-6
1-3	Dynamic SRDF	1-9
1-4	Two Production Sites and One Recovery Site	1-14
1-5	Data Vaulting Solution	1-15
1-6	Sites Containing Both Primary and Secondary Devices	1-16
1-7	PPRC and GDPS Support	1-18
1-8	SRDF With and Without FarPoint	1-19
2-1	Concurrent RDF Configuration	2-9
2-2	Primary (Source) and Secondary (Target) Relationships	2-10
2-3	Failed Link Between Source 2 and Target 1	2-11
2-4	Sources 1, 2, and 3 in a Consistency Group	2-11
2-5	Failed Link Between Source 2 and Target 1	2-12
3-1	SRDF Business Continuance	3-6
3-2	Establish Operation	3-7
3-3	Restore Operation	3-7
3-4	BCV Functioning As a Primary (Source) SRDF Device	3-9
3-5	SRDF Multi-Hop	3-10
3-6	R1/R2 Swap Procedure Modification	3-12
4-1	SRDF/Asynchronous	4-6
4-2	SRDF/A Locality of Reference Example	4-8
4-3	SRDF/A Allowed State Transitions	4-11

Tables

1-1	Data Protection Options for SRDF Disk Devices	1-5
1-2	SRDF Interfamily Connectivity Support	1-17

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/Asynchronous for users requiring a basic knowledge of SRDF and SRDF/A concepts and operation. Detailed SRDF installation and configuration information is available in the documents listed in *Related Documentation* 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) and how it works.

**Related
Documentation**

Other related EMC documents include:

- ◆ *Symmetrix High Availability Product Guide*
- ◆ *EMC ControlCenter SRDF Manager for UNIX Product Guide*
- ◆ *EMC ControlCenter SRDF Manager for Windows Product Guide*
- ◆ *Symmetrix Data Migration Service Product Guide*
- ◆ *SRDF Host Component for OS/390 and z/OS Product Guide*
- ◆ *SRDF/TimeFinder Manager for AS/400 RISC Product Guide*
- ◆ *SRDF Control Software for AS/400 RISC Product Guide*
- ◆ *EMC Host Connectivity Guides*
- ◆ *Symmetrix Manager Base Component for UNIX Product Guide*
- ◆ *Symmetrix Manager SRDF Component for UNIX Product Guide*
- ◆ *SRDF Host Component for MSP Product Guide*
- ◆ *SRDF Host Component for VOS3 Product Guide*

EMC WideSky Solutions Enabler Symmetrix publications of related interest are:

- ◆ *Symmetrix SRDF CLI Product Guide*
- ◆ *Symmetrix Base Management CLI Product Guide*
- ◆ *Symmetrix TimeFinder CLI Product Guide*
- ◆ *Symmetrix Device Masking CLI Product Guide*
- ◆ *WideSky Solutions Enabler Installation Guide*

For additional information on all Symmetrix-related publications, contact your EMC Sales Representative or refer to the EMC Powerlink web site at:

<http://powerlink.EMC.com>

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.

Palatino, bold	◆ Specific filenames or complete paths ◆ 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

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 Service 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 web site 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 Overview 1-2
- ◆ SRDF Volume Types 1-7
- ◆ SRDF Director Hardware 1-12
- ◆ SRDF Configurations 1-14

SRDF Overview

Symmetrix® Remote Data Facility (SRDF®) is a Symmetrix-based business continuance and disaster recovery solution sold as a separate license by EMC.

In simplest terms, SRDF is a configuration of Symmetrix systems, the purpose of which is to maintain multiple, real-time copies of logical volume data in more than one location. The Symmetrix systems can be in the same room, in different buildings within the same campus, or hundreds of kilometers apart.

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

- ◆ Disaster recovery
- ◆ Recovery from planned outages
- ◆ Remote backup
- ◆ Data center migration
- ◆ Data Replication and Mobility

SRDF/Data Mobility

SRDF/Data Mobility (SRDF/DM) is a special form of SRDF that permits operation in SRDF *Adaptive Copy* mode only and is designed for data replication and/or migration between two or more Symmetrix systems. SRDF/DM transfers data from source volumes to remote volumes quickly 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. Unlike full-function SRDF, SRDF/Data Mobility is not intended for disaster recovery.

SRDF/DM is sold as a separate license by EMC.

SRDF/Asynchronous

SRDF/Asynchronous (SRDF/A) is an SRDF application used for remote replication that allows you to maintain an asynchronous *dependent write consistent* copy of Symmetrix stored data at a remote site. In SRDF/A implementations:

- ◆ The recovery data from the target side lags behind the source by seconds or minutes.

- ◆ There is minimal host impact since the host adapter does not wait for an acknowledgement for background writes. Global Memory is used as a data buffer.
- ◆ It has the potential to more efficiently use network bandwidth due to *locality of reference* (for example, applications tend to write data in proximity of time and space). SRDF/A sends data to a remote site as cycles of dependent write consistent data. The write I/Os coming in are managed in global memory buffers through advanced Enginuity code algorithms and, regardless of how many times a specific block is written to, only the last write is sent over to the target box with its associated Delta Sets (there are two alternating cycles). Therefore, less data needs to be transferred to the target system.

SRDF/Asynchronous provides a protection mode that maintains *dependent write consistency* while balancing data exposure versus host impact and link costs. By balancing performance, data exposure, distance, and link costs, customers will benefit from an implementation that:

- ◆ Decouples local host application activity from the remote copy activity, thus minimizing host performance impact
- ◆ Efficiently utilizes bandwidth due to *locality of reference* concept (that is most, not all, applications tend to write data in proximity of time and place)
- ◆ Is designed for minimal and predictable data loss
- ◆ Uses no host cycles

SRDF/A is an Enginuity solution available at revision 5670 and is only supported with the Symmetrix DMX™ systems. SRDF/A is sold as a separate license by EMC. For more detailed information on SRDF/A, refer to Chapter 4, *SRDF/Asynchronous Operations*.

Local Versus Remote Mirroring

Local mirroring, or RAID 1, is a 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 RDF) reside in a different storage unit.

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 host connects to Symmetrix B with the secondary (target) or R2 volume containing the remotely mirrored data. The Symmetrix systems communicate through SRDF links.

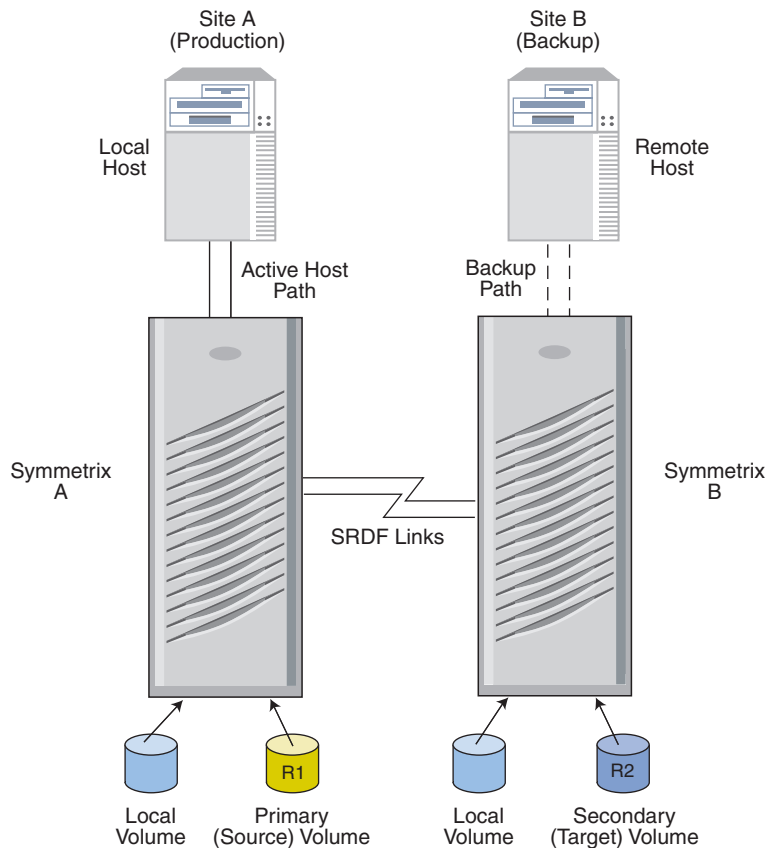


Figure 1-1 Basic SRDF Configuration

For an explanation of local, primary (source), and secondary (target) volumes, refer to *SRDF Volume Types* on page 1-7.

SRDF Supported Data Protection Options

The Symmetrix data protection options ensure a higher level of data protection, recoverability, and availability than the standard Symmetrix availability and reliability features. The options listed in Table 1-1 can be purchased separately and implemented into the Symmetrix SRDF operation.

Table 1-1 Data Protection Options for SRDF Disk Devices

Data Protection Option	Description
Mirroring (RAID Level 1) ^a	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 ^{a b}	Provides higher performance and availability than RAID Level 5 offerings for business online applications. 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.
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. RAID 1 and **either** Parity RAID (3+1) **or** Parity RAID (7+1) can be intermixed in the same Symmetrix system. Parity RAID (3+1) **and** Parity RAID (7+1) **cannot exist** simultaneously in the same Symmetrix system.
- b. SRDF R1, R2 Swap is not supported on Parity RAID devices. Dynamic SRDF (including EMC native PPRC support) does not support Parity RAID devices.

SRDF Configuration Using Switched Fibre Channel

A more complex *switched SRDF* configuration using Fibre Channel switches connected through E-Ports is shown in Figure 1-2 on page 1-6. Note that in this configuration, multiple primary (source) R1 devices are remotely connected through a Switched Area Network (SAN) to multiple secondary (target) R2 devices.

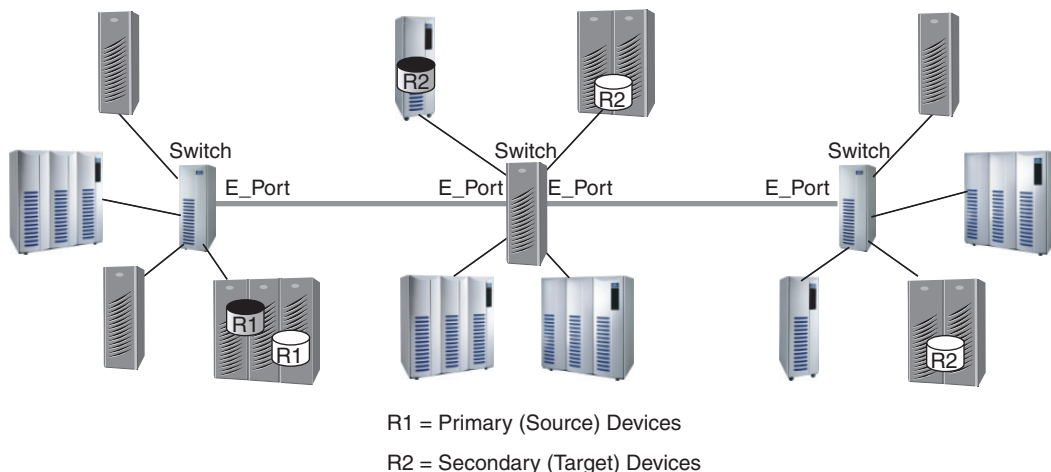


Figure 1-2 Switched SRDF With Multiple Primary and Secondary Devices

Monitoring and Controlling SRDF

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

After SRDF is up and running, you can monitor and control its operation by purchasing the appropriate EMC ControlCenter Base Component software and the appropriate SRDF or TimeFinder Control Option software from EMC.

For information about EMC ControlCenter software, contact your EMC representative.

SRDF Volume Types

SRDF refers to Symmetrix volumes as:

- ◆ Primary or source volumes (R1)
- ◆ Secondary or target volumes (R2)
- ◆ Local volumes
- ◆ Device groups
- ◆ Dynamic SRDF devices
- ◆ Dynamic Spares
- ◆ Metadevices

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

In Symmetrix Enginuity™ microcode version 5x64 and later, 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 or source volumes are also referred to as *R1* volumes. Updates to a primary volume are automatically copied to a secondary (target) volume in the other Symmetrix system.

Primary volumes can be locally protected by:

- ◆ A dynamic spare (refer to *Dynamic Spares* on page 1-10 for more information)
- ◆ Conventional mirroring (the primary volume is then referred to as a *mirrored pair*)
- ◆ RAID-S protection (the primary volume is a RAID-S data volume array)
- ◆ Symmetrix Parity RAID (3+1) and Parity RAID (7+1) protection (the primary volume is a Parity RAID data volume array)

In addition, a primary volume can be paired with a Business Continuity 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 (target) volumes are also referred to as R2 volumes.

As with primary volumes, secondary volumes can be locally protected by:

- ◆ A dynamic spare
- ◆ Conventional mirroring (the secondary volume is then referred to as a mirrored pair)
- ◆ RAID-S protection (the secondary volume is actually a RAID-S data volume array)
- ◆ Symmetrix Parity RAID (3+1) and Parity RAID (7+1) protection (the primary volume is a Parity RAID data volume array)

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 local mirroring, RAID-S, Parity RAID (3+1) and Parity RAID (7+1), or a dynamic spare.

Device Groups

A Symmetrix device is assigned a logical name when the device is added to a *device group*. A device group represents several devices and can be established to provide configuration, status, and performance data on the collective devices within the device group.

For example, a device group can be created to include all the physical devices used in a database application. The device group for the database application includes three devices with logical names DEV001, DEV002, and DEV003. Devices DEV001 and DEV002 are mirrored, and a third device, DEV003, is a RAID-S or Parity RAID device.

Engenuity code 5669 increases the maximum number of definable SRDF groups to 64 from 16. This enhancement allows customers more flexibility when working with recently-added remote data facility features:

- ◆ If you create an SRDF group each time you create a Peer-to-peer Remote Copy (PPRC) group, a sufficient number of SRDF groups must be available.

- ◆ If you have a large storage network with many Symmetrix systems interconnected, more than 16 SRDF groups may be necessary, especially if the data center uses Switched SRDF, where one SRDF director can be associated with more than one SRDF group.
- ◆ Symmetrix systems with Dynamic SRDF may also require more than 16 SRDF groups.

Dynamic SRDF Devices

Since Engenuity code 5568, devices can be configured to be Dynamic SRDF-capable devices. Dynamic SRDF functionality enables you to create, delete, and swap SRDF pairs while the Symmetrix system is in operation. Using Dynamic SRDF technology, you can establish SRDF device pairs from non-SRDF devices, then synchronize and manage them in the same way as configured SRDF pairs. Figure 1-3 shows Dynamic SRDF connections in a switched fabric environment.

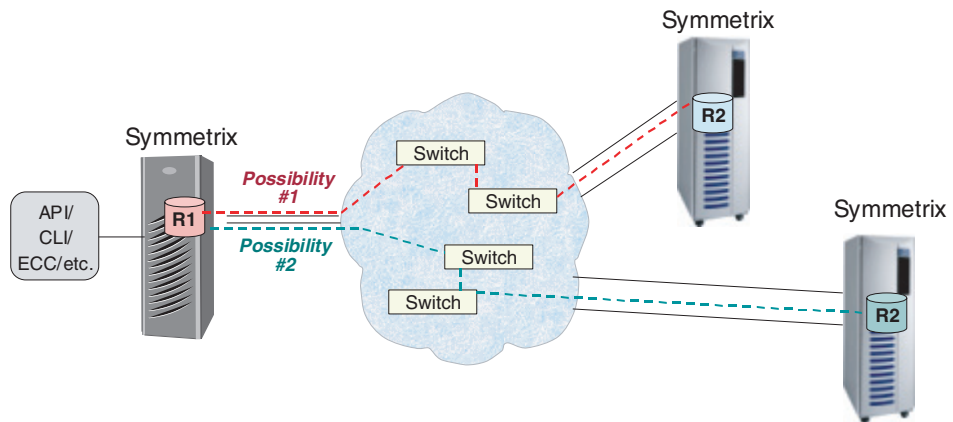


Figure 1-3 Dynamic SRDF

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

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 relationship pair from non-SRDF devices

Dynamic SRDF Groups

For Fibre Channel and Gigabit Ethernet connections, Enginuity code 5669 utilizes characteristics of Switched SRDF and Dynamic SRDF. It leverages the Switched SRDF capability of allowing one SRDF director to support more than one SRDF group and combines that with the Dynamic SRDF capability of allowing on-the-fly creation of SRDF device pairs.

This SRDF feature allows you to dynamically create empty SRDF groups (which has never before been allowed), dynamically associate the group with SRDF directors, then populate it with Dynamic SRDF device pairs. Dynamically reversing these steps is also possible with this feature.

Configurations created through this mechanism are persistent. PPRC is not supported with this feature in Enginuity code 5669.

Dynamic Spares

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 performance of all its disk drives and can detect when a disk is failing.

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 a dynamic spare is invoked, it contains no data; the data from the failed disk must be transferred to the dynamic spare. To do this, a mirror disk (a disk containing a copy of the data on the failed disk) is used. the Symmetrix system copies the data from the *next available mirror* — in other words, if there is a local mirror (a mirror disk within the same Symmetrix system), the Symmetrix system uses the local mirror. If there is only a remote mirror, the Symmetrix system uses the remote mirror.

If a disk fails completely and has no mirror, a dynamic spare cannot be invoked, because there is no way to copy the data from the failed disk. If a disk is failing but has not failed completely, a dynamic spare can be invoked. However, if the failing disk corrupted its data, the dynamic spare will receive the corrupted data. Therefore, it is always best to invoke a dynamic spare using a fully functioning mirror of the failed or failing disk.

In an SRDF configuration, if the failed device has a corresponding device on the other side of the SRDF link, the Symmetrix system containing the failed device automatically informs the other Symmetrix system that a dynamic spare was invoked.

Metadevices

In Enginuity microcode version 5265 and later, you can create a logical device that spans multiple physical devices. This device is known as a *metadevice*.

Windows NT supports only logical drives A through Z. If you have more than 23 physical devices (drives A and B are typically reserved and C is the boot drive), you can create logical drives that span multiple physical devices to take advantage of all potential storage. For example, you could create a drive F that spans two or more physical devices.

In an SRDF configuration, you can create metadevices for both primary and secondary volumes.

SRDF Director Hardware

SRDF director hardware provides the communications transport layer for SRDF data and information exchanges between Symmetrix systems.

SRDF director hardware consists of three types of adapters:

- ◆ ESCON Remote Adapters (RAs)
- ◆ Fibre Channel Remote Adapters (RAFs or RFs)
- ◆ Gigabit Ethernet (GigE) Remote Adapters

ESCON Remote Adapters

ESCON Remote Adapters, known as RAs, are board sets that can provide the link connections, fiber optic protocol support, and communications control between two Symmetrix systems in an SRDF configuration.

The RA board set that sends data across an SRDF link is known as an *RA-1*. An *RA-1* functions like an IBM ESCON host channel interface.

The RA board set that receives data sent across an SRDF link is known as an *RA-2*. An *RA-2* functions like an IBM ESCON storage director interface.

An ESCON SRDF link consists of an *RA-1* board set in one Symmetrix system, an *RA-2* board set in another Symmetrix system, and a fiber cable connecting them.

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 utilized for SRDF. With Symmetrix DMX1000, DMX2000, and DMX3000 models, an optional four-processor ESCON board can be utilized for SRDF.

Fibre Channel Remote Adapters

Enginuity microcode versions 5x66 and later support SRDF Fibre Channel emulation. The Remote Adapter for Fibre (RAF or RF) is a two-port Fibre Channel board set that is Enginuity-configured as the link between Symmetrix systems in an SRDF or Open Systems SDMS configuration.

**Symmetrix 8000
GigE Remote
Adapters**

The Symmetrix 8000 GigE remote adapter enables direct native IP connectivity for SRDF deployment. It provides two 1 Gb/s Ethernet ports and connects via SC connectors (1000Base SX). It requires Enginuity code 5568 or higher. The GigE adapter port must be connected to a device with a port type 1000Base SX. Fibre optic cables can be 50 micron or 62.5 micron.

**Symmetrix DMX
GigE Remote
Adapters**

The Symmetrix DMX GigE remote adapter enables direct native IP connectivity for SRDF deployment. It provides from one-to-four 1 Gb/s Ethernet ports and connects via LC connectors (1000Base SX). It requires Enginuity code 5670 or higher. It supports SRDF traffic to one or more remote Symmetrix systems that also have GigE remote directors installed. Fibre optic cables can be 50 micron or 62.5 micron.

SRDF Configurations

This section provides examples of typical SRDF configurations.

In Figure 1-4, two production sites (A and C) send data across SRDF links to one recovery site (B).

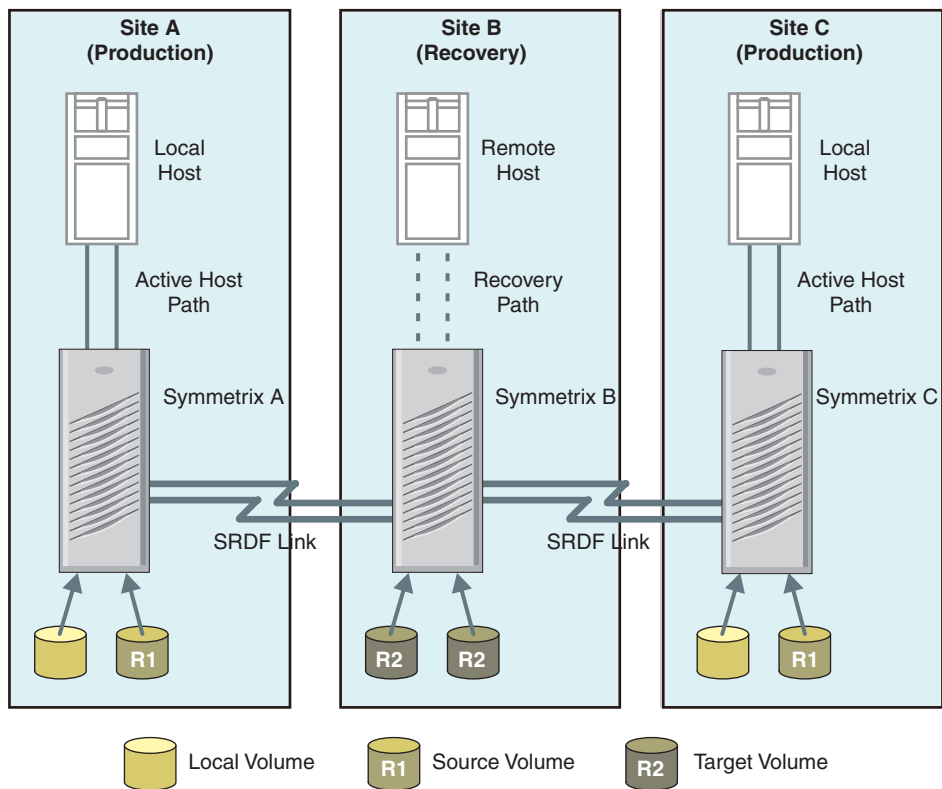


Figure 1-4 Two Production Sites and One Recovery Site

In Figure 1-5, one recovery site (G) provides a data vaulting solution for six production sites (A through F).

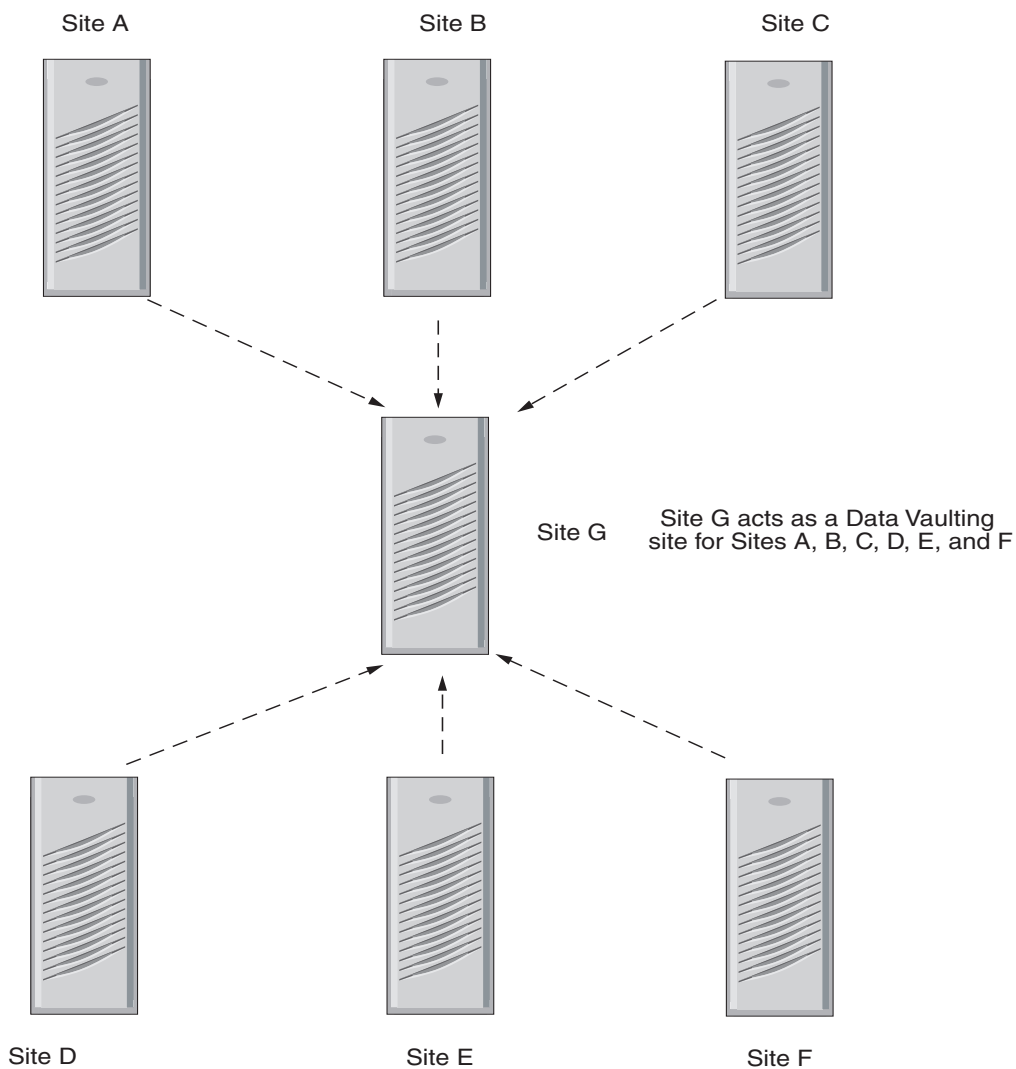


Figure 1-5 Data Vaulting Solution

Figure 1-6 illustrates the versatility of SRDF: some sites have *either* primary (R1) or secondary (R2) devices, while other sites have *both* primary (R1) and secondary (R2) devices.

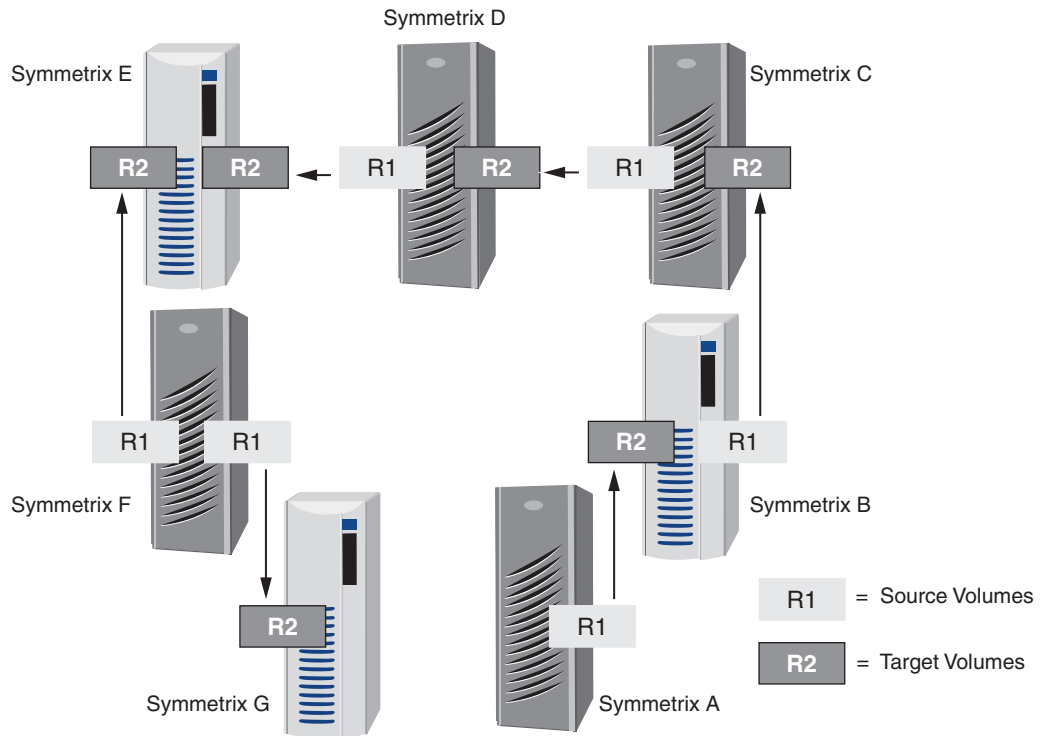


Figure 1-6 Sites Containing Both Primary and Secondary Devices

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 first realized with the release of code 5x68.

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 1-17.

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

a. Use for data migration only from source devices on Symmetrix 3x30/5x30 system models running Enginuity level 5267 to target devices on Symmetrix DMX system models s running Enginuity level 5670.

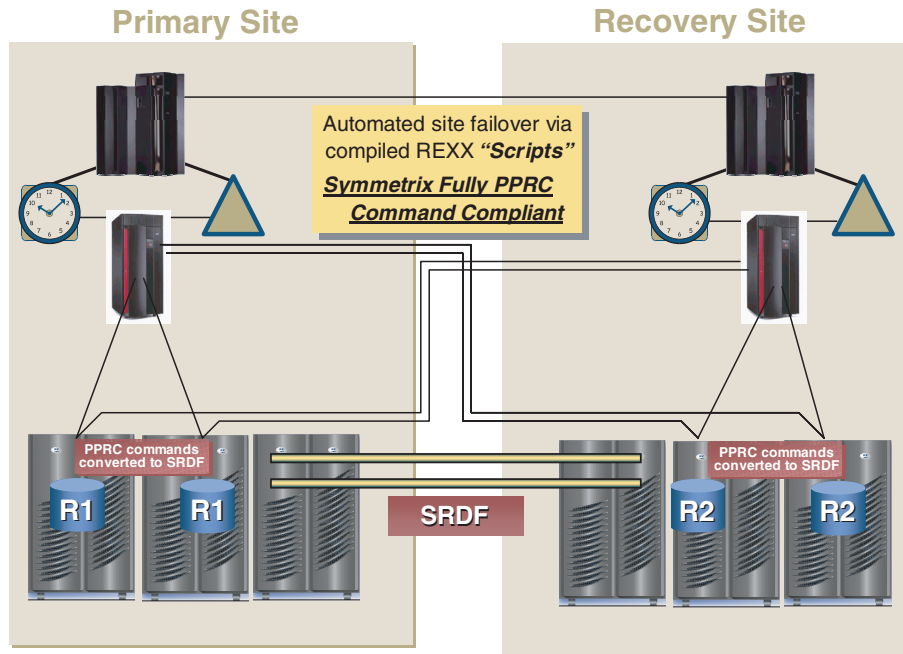
b. Only ESCON and Fibre Channel connections (no FarPoint) are supported.

SRDF Fully Switched Fabric Connectivity

Enginuity code 5x67 and later supports fully switched open SRDF (fibre) connections. Switched SRDF allows all RAs 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 full-duplex. Switched SRDF also enables fewer RAs to be present on the fabric, depending on system performance and redundancy requirements.

PPRC Command Support

Enginuity code 5568 enables the Symmetrix system to support native IBM Peer-to-Peer Remote Copy (PPRC) commands. PPRC is the remote copying solution available with IBM storage systems (Figure 1-7). PPRC is implemented as a subset of the Dynamic SRDF feature. As a result, Symmetrix DMX systems running Enginuity level 5669 (or higher) will support PPRC commands and IBM's Geographically Dispersed Parallel Sysplex (GDPS) solution.



SYM-000092

Figure 1-7 PPRC and GDPS Support

SRDF FarPoint

SRDF FarPoint™ is an SRDF feature used with ESCON extended distance solutions (and ESCON campus solutions with links greater than 15 km) 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 1-8 on page 1-19.

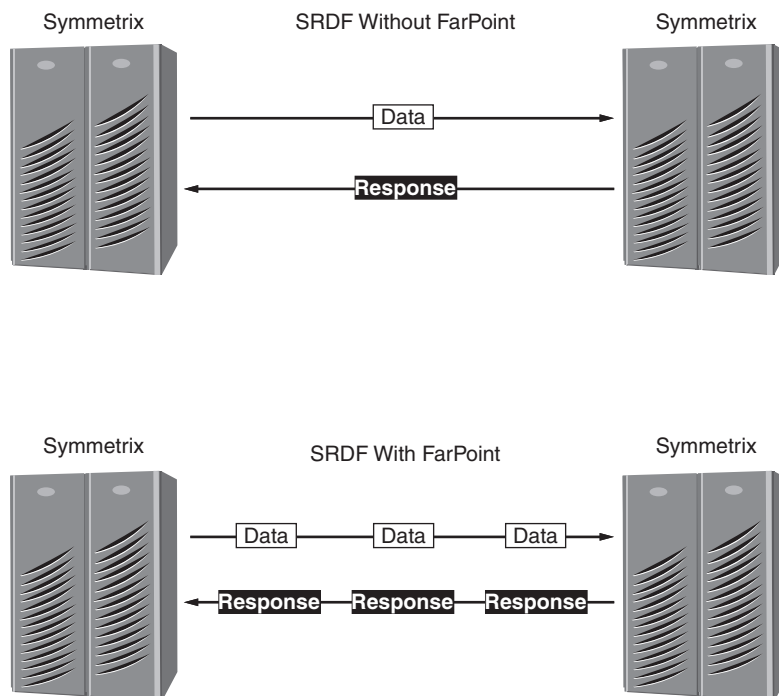


Figure 1-8 SRDF With and Without FarPoint

SRDF With FarPoint

With SRDF FarPoint, a single remote adapter can serially transmit more than one write I/O over the SRDF link (with up to one outstanding I/O per logical volume). 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.

Preserving Synchronization

From the point of view of the host, FarPoint does not change the SRDF protocol: the Symmetrix system still 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 the next write operation. Interaction with the host remains unchanged, so the Journal Zero 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). This solution is slower but safer, which is a priority when dealing with data recovery. In this situation, the write pipeline is cleared before any read operations are performed.

Handling I/O Failures

In a system failure, the host's point of view is always the most important consideration. If failover is necessary, a consistent view of the data exists on the secondary Symmetrix system except for the open write tasks that have not yet been acknowledged to the host. These open write operations may or may not be written on the remote Symmetrix system.

This situation can also occur without FarPoint, except that more tracks may be involved because of the pipelining. The solution to this problem is identical to the solution without FarPoint: the application must keep enough tracking information (such as log files) to allow it to resolve any inconsistencies in the data. Any application that uses logging (such as a database application) should bring the secondary Symmetrix system to a consistent state reflecting the moment of failure.

Performance Impact of FarPoint

A system using FarPoint can improve performance substantially depending on the length of the T3 link. Even on a short link, the FarPoint solution may provide some benefits.

Following are a few notes regarding performance improvements:

- ◆ The service time for each write operation cannot be reduced below twice the propagation time across the link. On very long links you should still expect long response times in `Journal Zero` mode. In `Journal One` mode, where a write operation may be acknowledged to the host without waiting for the status, the response times may still be short.
- ◆ Write operations to a single device are still serialized by the host, so in the case of a journal zero 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.

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

This chapter contains the following topics:

- ◆ SRDF Link Protocols 2-2
- ◆ SRDF Link States 2-3
- ◆ Primary Modes of Operation 2-4
- ◆ Secondary Modes of Operation 2-6
- ◆ Concurrent RDF 2-9
- ◆ Consistency Groups 2-10

SRDF Link Protocols

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

- ◆ Unidirectional
- ◆ Bidirectional
- ◆ Dual-Directional

SRDF Unidirectional Link Protocol

If all primary volumes reside in one Symmetrix system and all secondary (secondary) volumes reside in another Symmetrix system, write operations move in one direction, from primary to secondary. A *unidirectional* link protocol, in which data moves in the same direction over every link in the link group, satisfies this scenario.

SRDF Bidirectional Link Protocol

If each Symmetrix system contains both primary and secondary volumes, write operations move in both directions over the SRDF link group. For an SRDF configuration in which the Symmetrix systems are relatively close together (less than 60 km apart), a *bidirectional* link protocol can be used. With a bidirectional protocol, data moves in two directions over the *same link*.

For information about how write operations occur using a bidirectional link protocol, refer to *Write Operations in a Bidirectional Protocol* on page 3-2).

SRDF Dual-Directional Link Protocol

For extended-distance SRDF configurations (using E1, E3, T1, T3, and ATM links and/or network connections) that require data to move in two directions, a *dual-directional* link protocol is required. With a dual-directional protocol, multiple links are used; some links send data in one direction, while other links send data in the opposite direction.

SRDF Link States

There are four possible states for SRDF links:

The link is enabled only when the Symmetrix Remote Adapter is online and the link is operational.

- ◆ Remote Adapter is in IMPL (initial microprogram load) monitor mode — link disabled
- ◆ Remote Adapter is offline (the operator panel switch on the Remote Adapter is off) — link disabled
- ◆ Remote Adapter is online but the link is offline (damaged or disconnected cable) — link disabled
- ◆ Remote Adapter is online and the link is online (operational) — link enabled

If any link in a group is operational, the primary (*source*) and secondary (*target*) devices in the group can communicate. Primary and secondary devices may be placed in ready or not ready states on the links. If a link fails, the Symmetrix system retains the last known logical states for the affected devices. 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.

Primary Modes of Operation

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

SRDF /Data Mobility is not available in synchronous or semi-synchronous modes, but is supported in Adaptive Copy mode only. Refer to *Adaptive Copying Modes* on page 2-6 for more information.

For information on SRDF/Asynchronous (SRDF/A) refer to Chapter 4, *SRDF/Asynchronous Operations*. SRDF/A is sold as a separate license by EMC.

Synchronous Mode (Real-Time or Journal 0 Mode)

Used mainly for the campus solution, 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 acknowledgment is sent to the local host.

Synchronous mode provides real-time mirroring of data between the local Symmetrix system and the remote Symmetrix system. Data is written simultaneously to global memory of both systems — in real time — before the application I/O is completed, ensuring the highest possible data availability as shown in the following steps:

1. The local Symmetrix system containing the primary or source volume receives a write operation from the user application.
2. The write operation immediately moves 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 acknowledgment 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.

Semi-Synchronous Mode (Journal 1 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 acknowledgment is sent to the local host.

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

Semi-synchronous mode writes data to the local system, completes the I/O, and then synchronizes the data with the remote system as shown in the following steps:

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

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.
3. Read operations continue normally. A second write I/O will not be accepted.
4. Data moves to the remote Symmetrix system containing the secondary volume.
5. The remote Symmetrix system sends an acknowledgement to the local Symmetrix system.
6. When the acknowledgement is received, another write I/O can be accepted.

Secondary Modes of Operation

Secondary modes are used in conjunction with the primary operational modes described in *Primary Modes of Operation* on page 2-4.

SRDF/Data Mobility is available in Adaptive Copy mode only. EMC further recommends that you use SRDF/Data Mobility in *Adaptive Copy Disk Mode* for best performance.

Secondary modes include:

- ◆ Adaptive Copying Modes
 - Adaptive Copy Write-Pending Mode
 - Adaptive Copy Disk Mode
- ◆ Domino Modes
 - Device Domino Mode
 - Link Domino Mode

Adaptive Copying Modes

Adaptive copying 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 source 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 system is reached, regardless of whether the maximum skew value is reached.

The advantage to this mode is that it is typically faster to read data from global memory than from disk. The disadvantage is that global memory is temporarily consumed by the data until it moves to disk.

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 source 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.



CAUTION

Adaptive Copy Disk mode should not be used if the source volumes are not protected by local mirroring, RAID-S, or Parity RAID (3+1) and Parity RAID (7+1).

Domino Modes

Domino modes effectively stop all write operations to both primary and secondary volumes if all mirrors in a primary or secondary mirror group fail, or if all SRDF links in a link group become unavailable. While such a shutdown temporarily halts production processing, domino modes can prevent so-called *rolling disasters*.

There are two types of domino modes: device domino mode and link domino mode.

Device Domino Mode

You set device domino mode at the device level. If this mode is set to Yes for a secondary (target) 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 set link domino mode at the link group level. If this mode is set to Yes for an SRDF link group, and the last remaining link in the link group fails, all primary (source) volumes in the link group become unavailable to their host.

After the problem is corrected, you must reenable the volumes using one of the SRDF control utilities (the Symmetrix service processor or EMC ControlCenter software).

With either domino mode, the appropriate primary (source) volumes are forced off line 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 *Consistency Groups* on page 2-10 for more information).

Invalid Tracks Warning

When enabled for a secondary (target) volume, SRDF issues a warning if you attempt to recover data from that secondary volume when it is not synchronized with its primary (source) volume. In such cases, another form of data recovery (for example, a tape restore) is more appropriate.

SRDF System-Level Attributes

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

- ◆ Force links to go offline after a power-up. This prevents the possibility of a rolling disaster following a power outage.
- ◆ Prevent automatic link recovery after an all-link failure. This feature is useful if you are using Business Continuance Volumes (BCVs) in conjunction with secondary devices. In such cases EMC recommends that you split the BCVs before reactivating the links to ensure a stable database on the BCVs.

Concurrent RDF

Engenuity 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 RDF and is supported in both ESCON and Fibre Channel RDF configurations.

Concurrent RDF requires that each secondary volume operate in the same primary mode, either both synchronous or both semi-synchronous, but allows either (or both) volumes to be placed into adaptive copy mode.

Figure 2-1 shows a concurrent RDF configuration in which one secondary volume is in synchronous mode, while the other is in adaptive copy. 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.

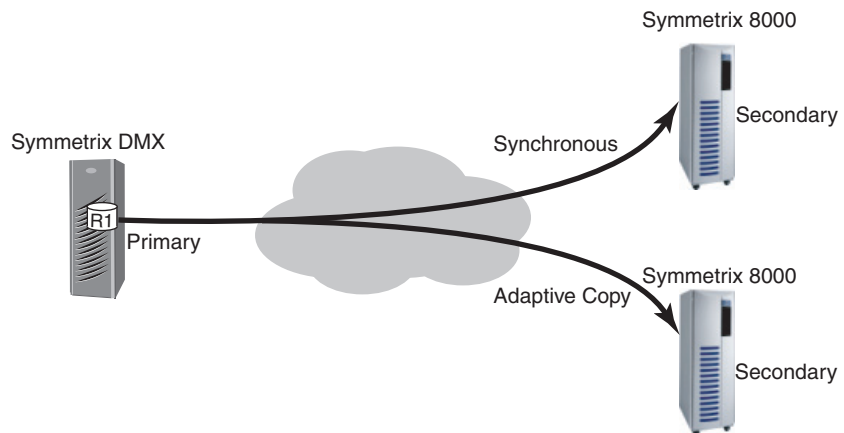


Figure 2-1 Concurrent RDF Configuration

Normal operating rules for SRDF also apply to concurrent configurations. When operating in synchronous mode, ending status for an I/O is not presented until the remote Symmetrix acknowledges receipt of the I/O to the primary Symmetrix. 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 is in synchronous mode and one is in adaptive copy mode, ending status is presented to the host when the synchronous volume acknowledges receipt of the I/O.

Consistency Groups

A *consistency group* is a group of Symmetrix devices specially configured to act in unison to maintain the integrity of a database distributed across multiple SRDF units 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 2-7).

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 2-2.

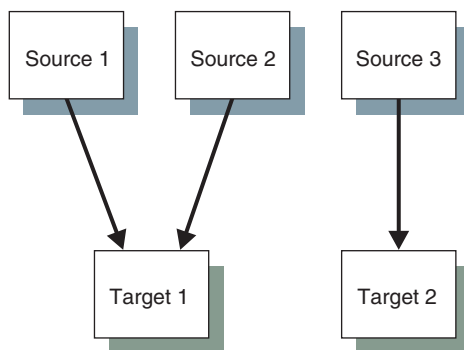


Figure 2-2 Primary (Source) and Secondary (Target) Relationships

Next, assume that the links between Source 2 and Target 1 fail. Without a consistency group, Sources 1 and 3 continue to copy data to the target devices while Source 2 does not (refer to Figure 2-3).

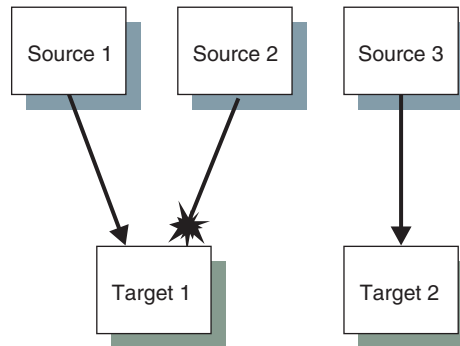


Figure 2-3 Failed Link Between Source 2 and Target 1

The result is that the copy of the database spread across Targets 1 and 2 becomes inconsistent.

However, if Sources 1, 2, and 3, belong to a consistency group, as shown in Figure 2-4, and the link between Source 2 and secondary Target 1 fails, the consistency group automatically stops Sources 1 and 3 from sending data to Targets 1 and 2, as shown in Figure 2-5.

Thus, the consistency of the database copy (spanning Targets 1 and 2) remains intact.

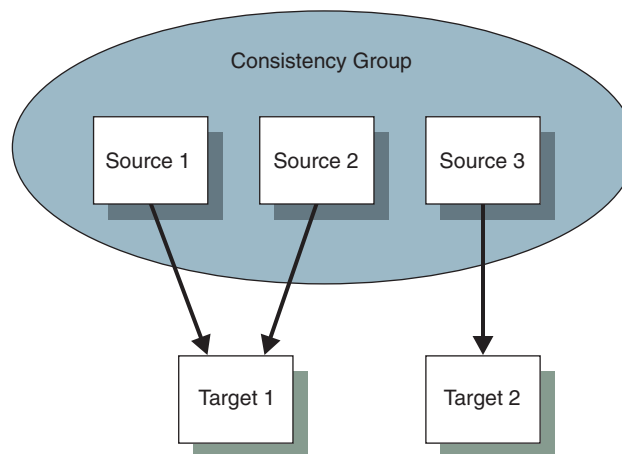


Figure 2-4 Sources 1, 2, and 3 in a Consistency Group

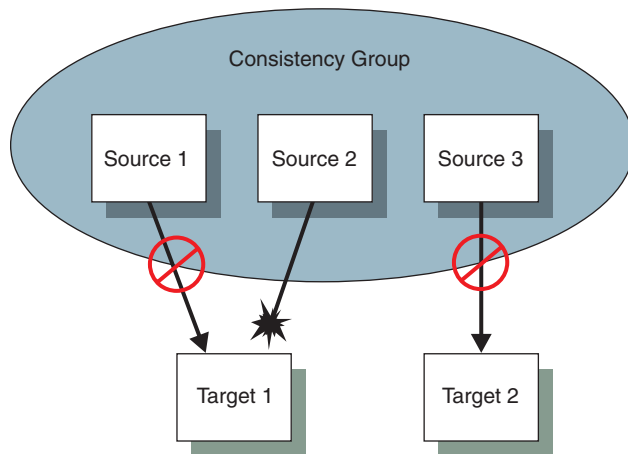


Figure 2-5

Failed Link Between Source 2 and Target 1**Continuous Processing**

I/O to the primary (source) devices in the consistency group can still occur even when the devices are suspended on the 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.

MVS Technical Considerations

You use host-based EMC software (the CGroup utility) 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 devices are initialized, all related volumes and controllers for the database are grouped to form the consistency group. If a write operation on a primary volume fails to reach the corresponding secondary volume, the Symmetrix system performs a unit check and returns the results to the controlling host computer.

A *suspend channel* program must be pre-built for each controller in a consistency group. The host software runs the suspend channel program for each controller in the group, and SRDF suspends the links for all devices in the group.

For more information about consistency groups and the CGroup utility, refer to the EMC *Symmetrix Consistency Group for MVS Product Guide*, P/N 300-999-031.

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

- ◆ Write Operations3-2
- ◆ Read Operations.....3-3
- ◆ Recovery Operations3-4
- ◆ Business Continuance Using SRDF3-6
- ◆ Business Continuance Using SRDF and TimeFinder.....3-8
- ◆ Migrating Data From R1 to a Larger R2 Device 3-11
- ◆ R1/R2 Swap Procedure Modification3-12
- ◆ Dynamic R1/R2 Swap.....3-13

Write Operations

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

Write Operations in a Unidirectional or Dual-Directional Protocol

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 an update 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.

A copy task, rather than a write task, is used for updates to volumes running in an Adaptive Copy mode.

Write Operations in a Bidirectional Protocol

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 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, because of an ESCON protocol limitation. Instead, the RA-2 must ask the RA-1 to *read* the data from the RA-2 (actually, from a 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

Read operations, in which the local host reads data from the remote Symmetrix system, are performed only to recover from a local data availability problem. Several events can cause 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.
- ◆ 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.

In addition to the requested data, the remote Symmetrix system sends the corresponding track ID from the secondary device to describe the transferred data accurately. The remote Symmetrix system uses the CRC byte to ensure data integrity.

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.

Recovery Operations



CAUTION

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

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.

If the remote host has read-only access to the secondary device, it does not interfere with SRDF operations. SRDF uses global memory slot locks on the secondary side to ensure that data reaches global memory on the secondary side.

Failover (Formerly known as Target Takeover)

If the remote host is to be activated to continue production processing, you must initiate a *failover* process. You can use EMC ControlCenter to control the failover process.

Initiating a failover involves changing the states of the secondary volumes to read/write and transferring production processing to the remote host. Updates made to the secondary volumes appear to the secondary volume group 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.

Returning Control to the Local Host

After the local host and the Symmetrix system containing the primary volumes are restored, production processing can resume on the local host. Several steps are required to transfer processing from the remote host back to the local host:

1. Halt processing on the remote host and change the states of the secondary devices to read only.
2. Bring the Symmetrix system on which the primary volumes reside to a ready state. Make sure that the SRDF links are suspended so that data movement across the links does not occur.

3. Use the service processor, an automated script, or EMC ControlCenter software to send invalid track identification information from the secondary volumes to the primary volumes.
4. Enable the SRDF links 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 ControlCenter software to automate or semi-automate this process.

Recovery for a Large Number of Invalid Tracks (Open Systems Sites Only)

If the recovery site (the remote 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) on the primary volumes. In this case, you can resynchronize the primary and secondary volumes while the remote 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 remote host and restart the local host.

Business Continuance Using SRDF

Business continuance refers to practices that enable you to achieve nearly non-stop, 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).

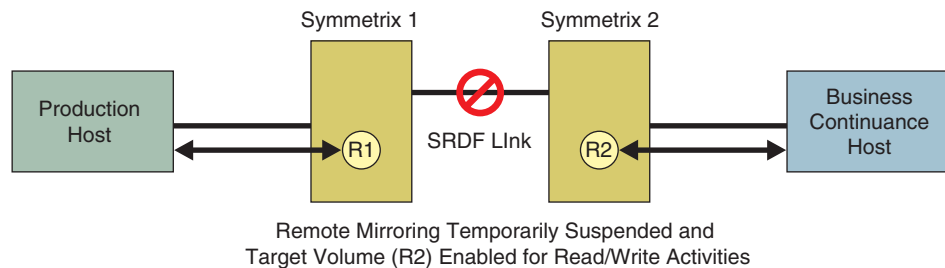


Figure 3-1 SRDF Business Continuance

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 — Euro-currency, new applications.

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, known as an *establish*, propagates any updates made to the primary volumes while the links were suspended, bringing the secondary volumes current with the most recent data

and resynchronizing the secondary volumes with the primary volumes (Figure 3-2).

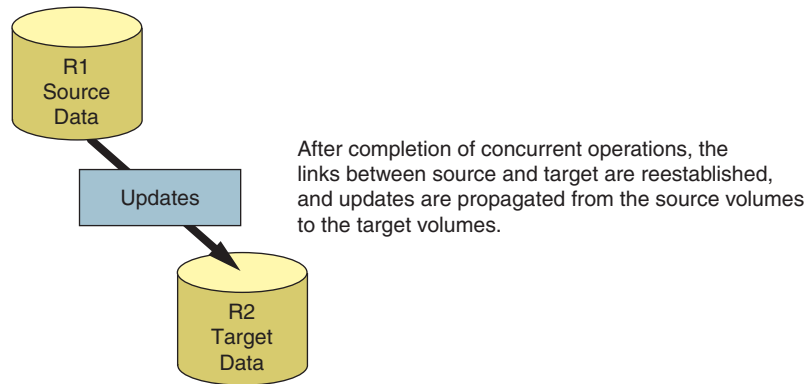


Figure 3-2 Establish Operation

Alternatively, you can perform a *Restore* operation to copy 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.

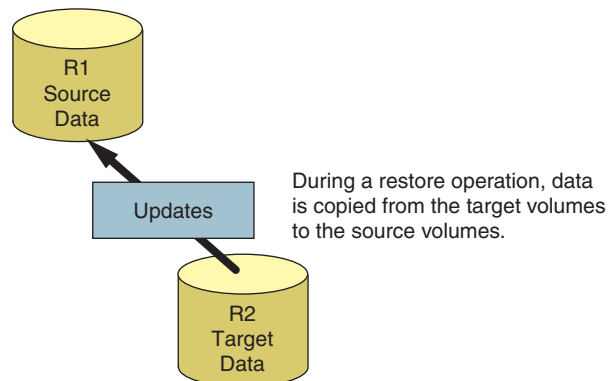


Figure 3-3 Restore Operation

Business Continuance Using SRDF and TimeFinder

SRDF, when used with EMC TimeFinder software, allows you to perform business continuance operations without temporarily suspending remote mirroring. The sections that follow describe how SRDF can be used with TimeFinder for business continuance operations.

If a BCV is paired with a secondary device, and you want to copy (restore) data from the BCV to the secondary device but you do *not* want to propagate the data across the SRDF links to the primary device, you should make sure that the *secondary device is suspended* on the SRDF link. You also should use the appropriate TimeFinder software restore command options to restrict the restore operation to the desired secondary volumes. If, on the other hand, you want the restored data to move across the SRDF links to the primary device automatically, you should make sure that the *secondary device is active* (not suspended) on the SRDF links.

For information about EMC TimeFinder, refer to the documentation (see *Related Documentation* in the *Preface*) or contact your EMC representative.

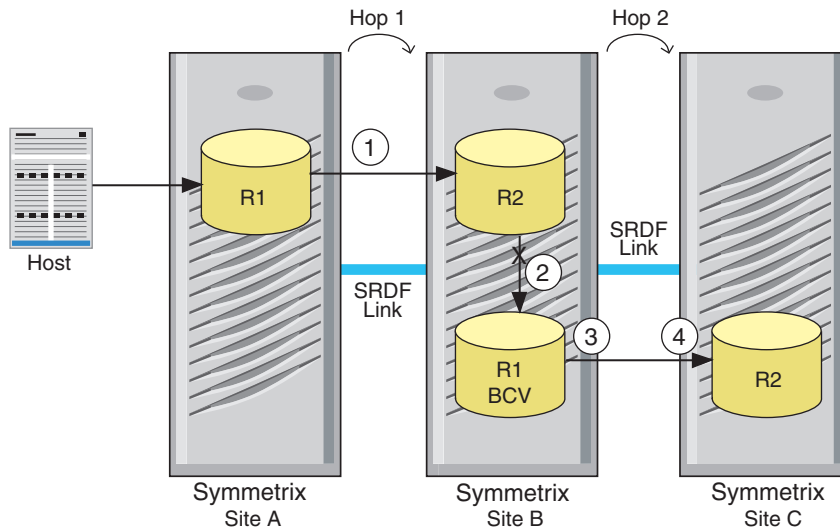
Using a BCV as a Primary (Source) Device

A BCV device may also function as an SRDF primary device if the BCV is not locally mirrored (a BCV R1 device). The SRDF remote mirror connection for the BCV R1 device is suspended while the BCV is paired with a local device.

When a BCV is paired with a primary device, the corresponding secondary device is disabled. When the BCV is split, the secondary device is reenabled and resynchronized with the primary device (Figure 3-4).

Using a BCV as a primary device is useful if you want to duplicate a system remotely.

You can use a BCV as a primary device in Enginuity microcode version 5x64 and later.



1. Data is remotely mirrored to the R2 and R1 BCV devices on the Symmetrix at Site B.
2. The BCV device pair (R2 and R1 BCV) is split.
3. The R2 on the Symmetrix at Site C is resynchronized with the R1 BCV at Site B.
4. Data is mirrored to the R2 device on the Symmetrix at Site C.

Figure 3-4

BCV Functioning As a Primary (Source) SRDF Device

Performing Remote Backups Without a Remote Host

You can issue SRDF and BCV control commands from a host or server connected to the local Symmetrix system. The commands move across the SRDF link to the secondary Symmetrix system. A host or server does not have to be located at the secondary location, allowing you 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.

This functionality is available in Engenuity microcode version 5265.

SRDF Multi-Hop

SRDF multi-hop enables a BCV on the local side of an SRDF configuration to perform an incremental synchronization with a secondary volume on the remote side, copying only changed tracks from the primary BCV to the secondary volume. This process eliminates the performance lag caused by a full synchronization across great distances, making a multitiered SRDF configuration feasible.

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 SRDF extended distance configuration.

Figure 3-5 illustrates SRDF multi-hop.

SRDF multi-hop is available in Enginuity microcode version 5265 or later.

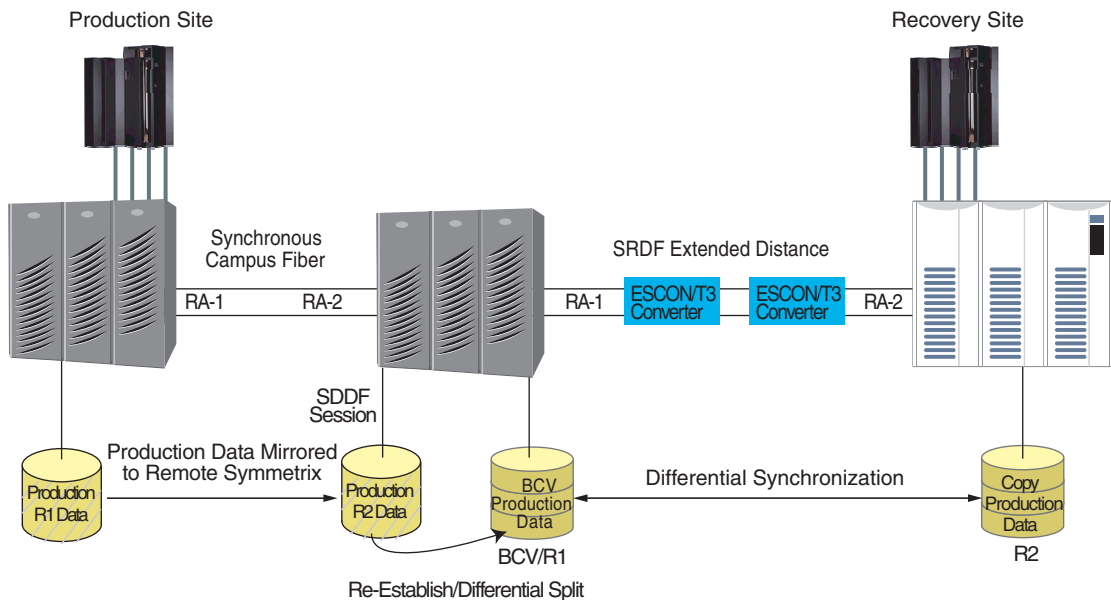


Figure 3-5 SRDF Multi-Hop

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.
- ◆ Concatenated metadevices are not supported, striped metadevices are supported.

R1/R2 Swap Procedure Modification

Introduced in code 5x65, the swap feature enables *switching* sides for source (R1) and target (R2) devices. In case of a source volume failure, this allows the target Symmetrix to assume a source role for the associated logical volumes (Figure 3-6). Codes 5x65 and 5x66 supported this swap through an online configuration change procedure. Code 67 supports the swap through a system call, without needing a new configuration load. This change is the difference between a lengthy (~5 minutes) user-driven procedure and a host command that takes milliseconds.

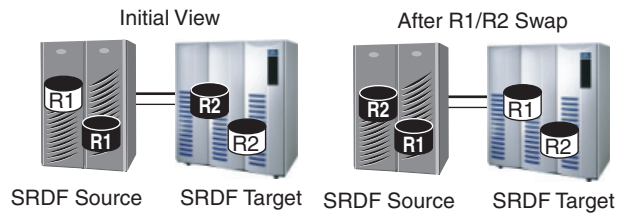


Figure 3-6 R1/R2 Swap Procedure Modification

Dynamic R1/R2 Swap

If you have:

- ◆ An SRDF license and a Configuration Manager license with Enginuity version 5x66 or 5267
- ◆ Or, an SRDF license and a Configuration Manager license with Enginuity version 5567 or higher, and the devices to be swapped are not Dynamic SRDF enabled in your Symmetrix configuration
- ◆ Or, an SRDF license with Symmetrix Enginuity version 5567 or higher and Dynamic SRDF is enabled in your Symmetrix configuration

you can swap the SRDF personality of the SRDF device designations of a specified device group. Source R1 device(s) become target R2 device(s) and target R2 device(s) become source R1 device(s).

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

Dynamic swap is not supported if Enginuity versions 5567 and 5568 are mixed between your local and remote Symmetrix.

R1/R2 swap operations are not supported for Enginuity version 5669 configurations where the R2 device is larger than the R1 device.

Dynamic SRDF swap operations are not supported for SRDF asynchronous devices.

SRDF/Asynchronous Operations

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

- ◆ SRDF/Asynchronous Overview4-2
- ◆ 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 Mode Transitions and Recovery.....4-12
- ◆ SRDF/A Recovery Scenarios.....4-15

SRDF/Asynchronous Overview

Beginning with Enginuity version 5670, Symmetrix supports SRDF/Asynchronous (SRDF/A) for Symmetrix DMX systems. 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 with the database. In the event of a disaster at the R1 site or if SRDFRDF 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:

- ◆ Extended data replication with database consistency
- ◆ Promotes efficient link utilization resulting in lower link bandwidth
- ◆ A consistent point-in-time image on the R2 devices is maintained at all times
- ◆ Supports all current SRDF topologies, including point-to-point and switched fabric
- ◆ 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.

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

Known Requirements and Limitations at Initial Release

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

- ◆ SRDF/A supports only single-Symmetrix-DMX-to-single-Symmetrix-DMX system support.
- ◆ Neither full volume snaps (clones) nor EMC Snap[™] are supported for local replication of the R2 side while SRDF/A is actively in session. The link must be down for the SNAP activate command and duration of the SNAP session.
- ◆ SRDF/A is supported for open systems through CLI only. Mainframe interface is supported through SRDF Host Component V5.2 and later.
- ◆ SRDF/A is unidirectional only in this first release.
- ◆ Support is for only one SRDF group on either side with devices participating in the SRDF/A setup.
 - All SRDF/A devices in this single session must be managed together.
 - SRDF/A will not be allowed if the devices are already enabled for SRDF consistency group protection.
- ◆ R1 and R2 personalities cannot be swapped, except under special configuration procedures. Consult your EMC representative.
- ◆ SRDF/A with Concurrent RDF is not supported.
- ◆ Dynamic SRDF operations is not supported.
- ◆ SRDF/AR (also know as SAR) will not support SRDF/A devices.

For more information on specific restrictions, refer to the *WideSky 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 now 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 *A*. Therefore it must ensure that *A* and *B* are either in the same cycle or that *B* is in a cycle that is processed to the remote side after the cycle which contains *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 remote site (for example, 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 source to the 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 link only once. This approach lowers the 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.

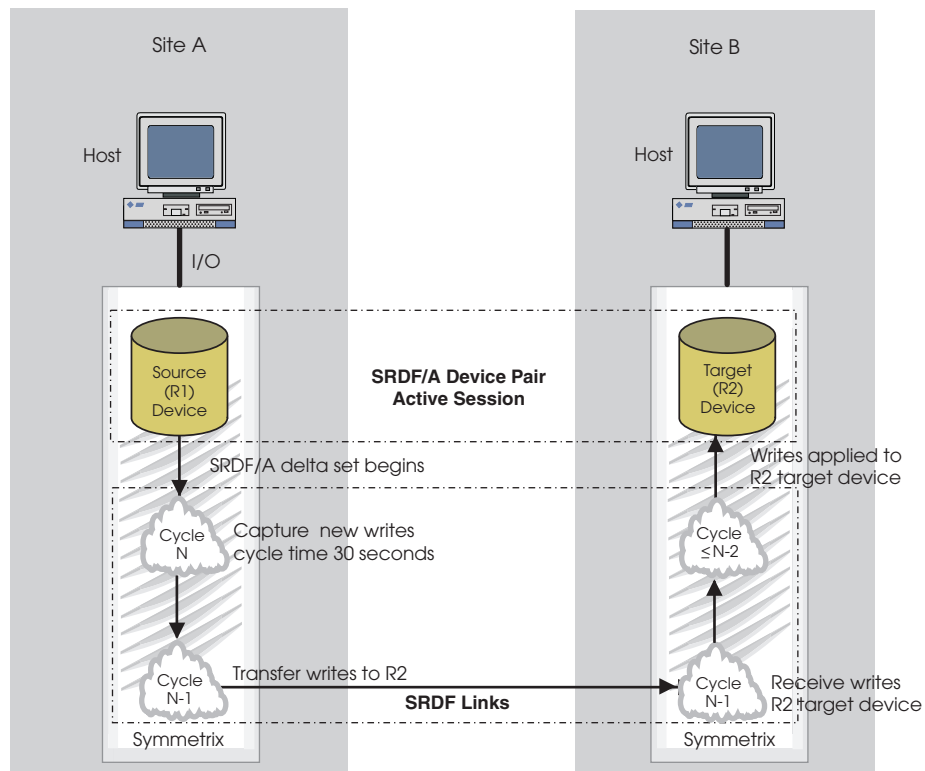


Figure 4-1 SRDF/Asynchronous

Dependent Write Consistency

Dependent write consistency is achieved through the processing of three ordered SRDF/A delta sets (cycles) between the source (R1) and the target (R2). Refer to Figure 4-1 on page 4-6, which depicts the three SRDF/A cycles. 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 active cycle on the R1, overwriting any duplicate tracks intended for data transfer over the link. The cycle is active for a predetermined minimum amount of time, which can be configured on the Symmetrix system. The default time is 30 seconds. After the minimum time has been reached, the cycle data becomes the inactive data and moves into the next cycle position ($N-1$) and begins transferring the delta set over the link into the R2. The new cycle N then begins collecting new writes.

In cycle $N-1$, the delta set is temporarily collected on the R2 side. When the $N-1$ cycle has finished transferring data into the R2, the cycle data moves into the next cycle position ($N-2$) and begins destaging the data to the R2 storage devices. The data is considered committed to the R2 in cycle $N-2$ 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 cycle time.

When the delta set $N-2$ becomes the active cycle on the R2 side, the R1 and R2 are considered to be in the *consistent pair state*, both containing a consistent image of data. The user can verify if the data in a current SRDF/A session has been applied to the R2. Consistency protection can be enabled and disabled for SRDF/A devices.

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 channel write I/O bandwidth), where it is buffered (creating an additional global memory requirement) and sent out the inter-controller 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.

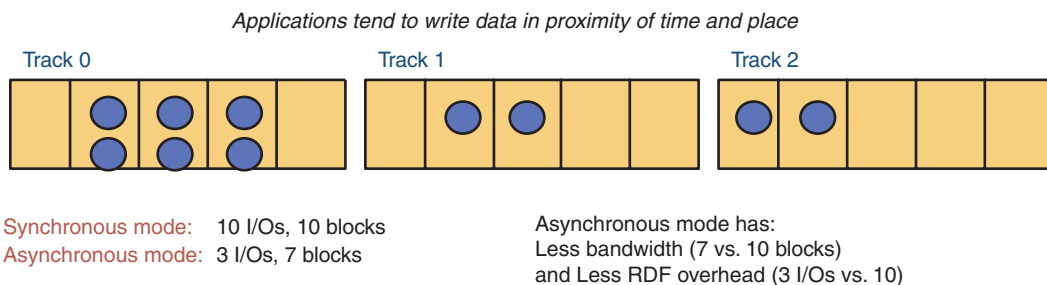


Figure 4-2 SRDF/A Locality of Reference Example

The SRDF/A flow control model, at any single point in time, has three cycles in *motion*, 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 site, the *active* cycle is collecting all new writes and tagging them as belonging to cycle *N*. There is also an *inactive* cycle *N-1* which is not receiving any new data, but is pushing the data it had collected when it was the active cycle to the remote side. These two cycles switch roles from *active* to *inactive* during the cycle switch process, which is driven by the Remote Adapters (RAs), and is initiated by the primary side. These *cycles* are physically implemented as a collection of pointers to global memory slots.

At the remote site there is also an *inactive* cycle (*N-1*), which is receiving data from its partner inactive cycle in the primary site and placing it in global memory. The *active* cycle (*N-2*) at the remote site is *draining* the global memory from a previous cycle that had been received in its entirety and therefore represents a consistent image of the data. It does this by simply marking all tracks as write pending to the R2s, and then lets the normal disk adapter (DA) algorithms destage the data to the disk.

The cycles on the R2 side are physically implemented as linked lists with Cache Only Virtual Devices (COVDs). A COVD is a virtual device with no log device associated with it. COVDs provide an easy way for the SRDF/A function to keep track of the location of global memory slots on behalf of their associated standard devices (the real R2s).

The switching between *active* and *inactive* 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 inactive ($N-1$) cycle.

Once the cycle 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 cycle switch if all conditions for a cycle switch are met.

The frequency of these cycle switches is a user controllable setting. The current minimum cycle time is 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.

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.

Version 1 of SRDF/A is a single Symmetrix solution only.

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*. (The reason for this is that *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*.

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: Devices are NR on the link, no data is moving to the R2 side

Do not confuse the use of *active* and *inactive* in describing the overall state of the SRDF/A feature with the cycle descriptors of *active* and *inactive* used internally within SRDF/A.

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.

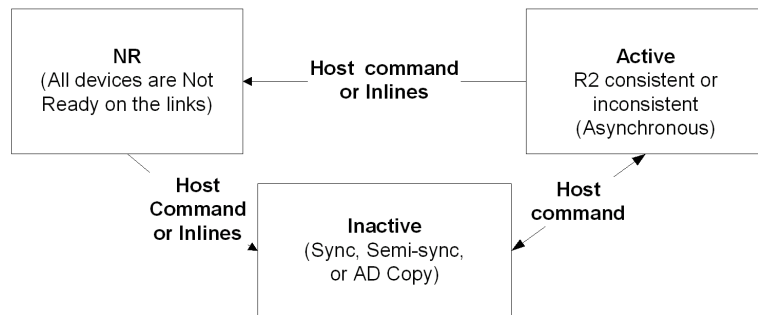


Figure 4-3 SRDF/A Allowed State Transitions

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 4-11.

For specific open systems commands, refer to the *Symmetrix SRDF CLI Product Guide*. For mainframe commands, refer to the *SRDF Host Component for OS/390 and 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 WP/Disk mode). 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 in version 1.

Switching to SRDF/Asynchronous Mode

The user can ask the R1 Symmetrix to switch to SRDF Asynchronous mode.

In the first release, 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, 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 re-sync 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.

A direct transition *to* Adaptive Copy write pending or disk mode, while SRDF/A is active, is not allowed. See *Deactivating SRDF/A* on page 4-14.

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 lost, 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 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 SymDial 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 OS/390 and 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 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 becomes *full*. The inactive and active cycles consume all the available write memory in the Symmetrix.

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.

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

This glossary contains terms related to SRDF. Many of these terms are used in this manual.

A

Adapter Symmetrix circuit board (card) that provides the host interface to the director board. In an SRDF environment, the adapter is called either an RA (ESCON), or an RAF (Fibre Channel).

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

B

Buffer Credits Buffer credits are associated with the 2 KB buffers found in Fibre Channel devices. The number of buffers at the receiving end of a link determines the amount of data that can be loaded and transmitted down the fiber link, before transmission must stop and the data transmission is acknowledged to the sender.

C

Change Tracker

Change Tracker is a tool available to EMC SEs to aid in the estimation of bandwidth required in an SRDF configuration. It's especially useful for sizing SRDF in a Multi-Hop configuration.

The Change Tracker application allows a user to monitor the update activity on Symmetrix devices. Change Tracker consists of two programs; a "collector" program that gathers the update activity statistics, and a "reporter" program that analyzes the data and formats the data into useful reports.

Concurrent RDF

An SRDF feature supported in Enginuity 5567 and later that allows a single primary volume to be remotely mirrored to two secondary volumes concurrently.

D

Dark Fiber / Lit Fiber

Dark Fiber is fiber optic cable that is inactive (dark) until the customer lights it up with their own new traffic. When a customer leases dark fiber, they have full use of the bandwidth on that fiber. In a "lit" (currently used) fiber service, a wavelength or specific protocol such as Gigabit Ethernet has been provided to the customers building or data center. It is already lit up with other traffic since it is a shared fiber. Note that a band in a DWDM fiber could be considered "dark" if it is not yet used.

**Dense Wavelength
Division Multiplexing
(DWDM)**

DWDM assigns one wavelength of light (λ) to each data stream applied. These wavelengths are then multiplexed through a single fiber pair. The process is protocol independent. The channel module locks in on the bit rate assigned. Distance is limited by the power of the laser, typically 80 km or less today. Ratios of 32:1 protected channels are available today for enterprises and 128:1 for carrier class systems.

Device

Uniquely addressable part of the Symmetrix subsystem that consists of a set of access arms, the associated disk surfaces, and the electronic circuitry required to locate, read, and write data. See also *Volume*.

Director

Component in the Symmetrix subsystem that allows the Symmetrix system to transfer data between the host channels and disk devices.

Disk Director

Component in the Symmetrix subsystem that interfaces between cache and the disk devices.

Dynamic Sparring Symmetrix feature that automatically transfers data from a failing disk device to an available spare disk device without affecting data availability. This feature supports all non-mirrored devices in the Symmetrix subsystem.

E

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

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

Enginuity The operating environment for EMC Symmetrix systems that provides high speed, highly available storage management functionality.

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

ESCON Director Device that provides a dynamic switching function and extended link path lengths (with XDF capability) when attaching an ESCON channel to a Symmetrix serial channel interface.

F

Fibre Channel Open Systems High performance serial-I/O channel path, with data rates currently 100MB/s and plans to move to 200MB/s. Transports SCSI commands and data in serial frames, which can be routed in a switched environment. Combines the performance, hardware based functionality and reliability of I/O channels with the distance and routing benefits of serialized network protocols. This forms the basis of the second-generation transport layer for SRDF.

Frame Data packet format in an ESCON environment.

G

Gigabyte (GB) 10^9 bytes.

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

H

Hyper-Volume Extension Ability to define more than one logical volume on a single physical disk device, utilizing the disk's full formatted capacity. These logical volumes are user-selectable in size. The minimum volume size is one cylinder, and the maximum size depends on the disk device capacity and the emulation mode selected.

I

IML Initial microcode program loading.

internet A collection of networks interconnected by a set of routers that allow them to function as a single, large virtual network

Internet (capital "I") The largest internet in the world consisting of large national backbone nets and a myriad of regional and local campus networks all over the world. The Internet uses the Internet Protocol suite.

I/O Device Addressable input/output unit, such as a disk device.

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

ISL Inter-Switch Link. The link between two switches in a fabric.

K

Kilobyte (KB) 1024 bytes.

L

Logical Volume User-defined storage device allowing the user to define a physical disk device as a number of logical volumes.

M

Megabyte (MB) 10^6 bytes.

Mirroring Process in which a Symmetrix system maintains two identical copies of a designated volume on separate disks. Each volume is automatically updated during a write operation. If one disk device fails, the Symmetrix system automatically uses the other disk device.

Mirrored Pair Logical volume with all data recorded twice, once on each of two different physical devices.

P

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

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

S

SCSI Adapter Card in the Symmetrix subsystem that provides the physical interface between the disk director and the disk devices.

SCU_ID For 3880 storage control emulations, this value uniquely identifies the storage director without respect to its selection address. It identifies to the host system, through the EREP, the director detecting the failing subsystem component. This value automatically increments by one for each director installed. The SCU_ID must be a unique number in the host system. It should be an even number and start on a zero boundary.

SONET Synchronous Optical Network Transport. A physical layer technology designed to provide a universal transmission and multiplexing scheme, with transmission rates in the multiple-gigabit

per second range, and a sophisticated operations and management system. SONET is the transport layer between the network protocol and the physical fiber that provides the framing and add/drop multiplexing function. Throughput ranges from OC1, at 51.84 Mb/s to OC256 at 13 gigabits /second (Gb/s). OC192 is currently being deployed for long haul traffic within carrier networks, connecting data centers and local offices.

SSID For 3990 storage control emulations, this value identifies the physical components of a logical direct access storage device (DASD) subsystem. The SSID must be a unique number in the host system. It should be an even number and start on a zero boundary.

Symm 4.0 Symmetrix family that includes models 3330, 3430, 3700, 5330, 5430, and 5700.

Symm 4.8 Symmetrix family that includes models 3630, 3830, 3930, 5630, 5830, and 5930.

Symm 5.0 Symmetrix family that includes models 8430 and 8730.

T

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

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

U

Unit Address Hexadecimal value that uniquely defines a physical I/O device on a channel path in an OS/390 environment.

V

Volume General term referring to a storage device. In the Symmetrix subsystem, a volume corresponds to single disk device.

A

- Adaptive copying mode 2-6
 - copy disk mode 2-7
 - write-pending mode 2-6
- Attributes, system-level 2-8

B

- Backups without a remote host 3-9
- BCV
 - as a primary (source) device 3-8
 - concurrent operations 3-6
 - performing remote backups 3-9
 - SRDF multi-hop 3-10
- Bidirectional link protocol 2-2
- Business continuance operations 3-6 to 3-10

C

- Concurrent RDF 2-9
- Consistency groups 2-10 to 2-12

D

- Data
 - protection guidelines 1-5
- Data Mobility 1-2
- Data protection options 1-5 to??
- Data Vaulting 1-15
- Device groups
 - definition of 1-8
- Domino mode 2-7
 - device domino mode 2-7
 - link domino mode 2-8

- Dual-directional link protocol 2-2
- Dynamic RDF 1-9, 3-13
- Dynamic spares 1-10

E

- EMC ControlCenter software 1-6
- ESCON
 - Remote Adapter 1-12

F

- Failover (target takeover) 3-4
- FarPoint 1-18
 - performance impact 1-21
 - preserving synchronization 1-20
- Fibre Channel
 - Remote Adapter 1-12

I

- Invalid tracks warning 2-8

J

- Journal 0 mode 2-4
- Journal 1 mode 2-5

L

- Link protocols
 - bidirectional 2-2
 - dual-directional 2-2
 - unidirectional 2-2
- Links

- possible states 2-3
- Local mirroring 1-3
- Local volumes 1-8

M

- Meta devices 1-11
- Mirroring
 - local vs. remote 1-3
- Modes of operation
 - primary 2-4 to 2-5
 - real-time 2-4
 - secondary 2-6 to 2-8

P

- Primary volumes 1-7

R

- RA 1-12
- RAF 1-12
- RAID 1 1-3
- Read operations 3-3
- Recovery operations 3-4 to 3-5
- Remote mirroring 1-3
- RF 1-12

S

- Secondary volumes 1-8
- Semi-synchronous mode 2-5
- Source volumes 1-7
- SRDF
 - configurations 1-14 to 1-21
 - FarPoint 1-18
 - hardware 1-12
 - overview 1-2
- SRDF control operations
 - swap 3-13
- SRDF FarPoint, performance impact 1-21
- SRDF/Data Mobility 1-2
- Swap RDF devices 3-13
- Symmetrix Remote Data Facility
 - Refer to SRDF
- Synchronous mode 2-4
- System-level attributes 2-8

T

- Target takeover 3-4
- Target volumes 1-8

U

- Unidirectional link protocol 2-2

V

- Volume types 1-7 to 1-11

W

- Windows NT meta devices 1-11
- Write operations 3-2