EMC® VMAX® All Flash Product Guide

VMAX 250F, 450F, 850F, 950F
with HYPERMAX OS

REVISION 07

Published May 2017

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

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

Published in the USA.

EMC Corporation
Hopkinton, Massachusetts 01748-9103
1-508-435-1000 In North America 1-866-464-7381
www.EMC.com

Preface

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

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

Note

This document was accurate at publication time. New versions of this document might be released on EMC Online Support (https://support.emc.com). Check to ensure that you are using the latest version of this document.

Purpose

This document outlines the offerings supported on VMAX All Flash 250F, 450F, 850F, 950F arrays running HYPERMAX OS 5977.

Audience

This document is intended for use by customers and EMC representatives.

Related documentation

The following documentation portfolios contain documents related to the hardware platform and manuals needed to manage your software and storage system configuration. Also listed are documents for external components which interact with your VMAX All Flash array.

EMC VMAX All Flash Site Planning Guide for VMAX 250F, 450F, 850F, 950F with HYPERMAX OS

Provides planning information regarding the purchase and installation of a VMAX 250F, 450F, 850F, 950F with HYPERMAX OS.

EMC VMAX Best Practices Guide for AC Power Connections

Describes the best practices to assure fault-tolerant power to a VMAX3 Family array or VMAX All Flash array.

EMC VMAX Power-down/Power-up Procedure

Describes how to power-down and power-up a VMAX3 Family array or VMAX All Flash array.

EMC VMAX Securing Kit Installation Guide

Describes how to install the securing kit on a VMAX3 Family array or VMAX All Flash array.

E-Lab™ Interoperability Navigator (ELN)

Provides a web-based interoperability and solution search portal. You can find the ELN at https://elabnavigator.EMC.com.

SRDF Interfamily Connectivity Information

Defines the versions of HYPERMAX OS and Enginuity that can make up valid SRDF replication and SRDF/Metro configurations, and can participate in Non-Disruptive Migration (NDM).

EMC Unisphere for VMAX Release Notes

Describes new features and any known limitations for Unisphere for VMAX .

EMC Unisphere for VMAX Installation Guide

Provides installation instructions for Unisphere for VMAX.

EMC Unisphere for VMAX Online Help

Describes the Unisphere for VMAX concepts and functions.

EMC Unisphere for VMAX Performance Viewer Online Help

Describes the Unisphere for VMAX Performance Viewer concepts and functions.

EMC Unisphere for VMAX Performance Viewer Installation Guide

Provides installation instructions for Unisphere for VMAX Performance Viewer.

EMC Unisphere for VMAX REST API Concepts and Programmer's Guide

Describes the Unisphere for VMAX REST API concepts and functions.

EMC Unisphere for VMAX Database Storage Analyzer Online Help

Describes the Unisphere for VMAX Database Storage Analyzer concepts and functions.

EMC Unisphere 360 for VMAX Release Notes

Describes new features and any known limitations for Unisphere 360 for VMAX.

EMC Unisphere 360 for VMAX Installation Guide

Provides installation instructions for Unisphere 360 for VMAX.

EMC Unisphere 360 for VMAX Online Help

Describes the Unisphere 360 for VMAX concepts and functions.

EMC Solutions Enabler, VSS Provider, and SMI-S Provider Release Notes

Describes new features and any known limitations.

EMC Solutions Enabler Installation and Configuration Guide

Provides host-specific installation instructions.

EMC Solutions Enabler CLI Reference Guide

Documents the SYMCLI commands, daemons, error codes and option file parameters provided with the Solutions Enabler man pages.

EMC Solutions Enabler Array Controls and Management for HYPERMAX OS CLI User Guide

Describes how to configure array control, management, and migration operations using SYMCLI commands for arrays running HYPERMAX OS.

EMC Solutions Enabler Array Controls and Management CLI User Guide

Describes how to configure array control, management, and migration operations using SYMCLI commands.

EMC Solutions Enabler SRDF Family CLI User Guide

Describes how to configure and manage SRDF environments using SYMCLI commands.

SRDF Interfamily Connectivity Information

Defines the versions of HYPERMAX OS and Enginuity that can make up valid SRDF replication and SRDF/Metro configurations, and can participate in Non-Disruptive Migration (NDM).

EMC Solutions Enabler TimeFinder SnapVX for HYPERMAX OS CLI User Guide

Describes how to configure and manage TimeFinder SnapVX environments using SYMCLI commands.

EMC Solutions Enabler SRM CLI User Guide

Provides Storage Resource Management (SRM) information related to various data objects and data handling facilities.

EMC SRDF/Metro vWitness Configuration Guide

Describes how to install, configure and manage SRDF/Metro using vWitness.

VMAX Management Software Events and Alerts Guide

Documents the SYMAPI daemon messages, asynchronous errors and message events, and SYMCLI return codes.

EMC ProtectPoint Implementation Guide

Describes how to implement ProtectPoint.

EMC ProtectPoint Solutions Guide

Provides ProtectPoint information related to various data objects and data handling facilities.

EMC ProtectPoint File System Agent Command Reference

Documents the commands, error codes, and options.

EMC ProtectPoint Release Notes

Describes new features and any known limitations.

EMC Mainframe Enablers Installation and Customization Guide

Describes how to install and configure Mainframe Enablers software.

EMC Mainframe Enablers Release Notes

Describes new features and any known limitations.

EMC Mainframe Enablers Message Guide

Describes the status, warning, and error messages generated by Mainframe Enablers software.

EMC Mainframe Enablers ResourcePak Base for z/OS Product Guide

Describes how to configure VMAX system control and management using the EMC Symmetrix Control Facility (EMCSCF).

EMC Mainframe Enablers AutoSwap for z/OS Product Guide

Describes how to use AutoSwap to perform automatic workload swaps between VMAX systems when the software detects a planned or unplanned outage.

EMC Mainframe Enablers Consistency Groups for z/OS Product Guide

Describes how to use Consistency Groups for z/OS (ConGroup) to ensure the consistency of data remotely copied by SRDF in the event of a rolling disaster.

EMC Mainframe Enablers SRDF Host Component for z/OS Product Guide

Describes how to use SRDF Host Component to control and monitor remote data replication processes.

EMC Mainframe Enablers TimeFinder SnapVX and zDP Product Guide

Describes how to use TimeFinder SnapVX and zDP to create and manage space-efficient targetless snaps.

EMC Mainframe Enablers TimeFinder/Clone Mainframe Snap Facility Product Guide

Describes how to use TimeFinder/Clone, TimeFinder/Snap, and TimeFinder/CG to control and monitor local data replication processes.

EMC Mainframe Enablers TimeFinder/Mirror for z/OS Product Guide

Describes how to use TimeFinder/Mirror to create Business Continuance Volumes (BCVs) which can then be established, split, re-established and restored from the source logical volumes for backup, restore, decision support, or application testing.

EMC Mainframe Enablers TimeFinder Utility for z/OS Product Guide

Describes how to use the TimeFinder Utility to condition volumes and devices.

EMC GDDR for SRDF/S with ConGroup Product Guide

Describes how to use Geographically Dispersed Disaster Restart (GDDR) to automate business recovery following both planned outages and disaster situations.

EMC GDDR for SRDF/S with AutoSwap Product Guide

Describes how to use Geographically Dispersed Disaster Restart (GDDR) to automate business recovery following both planned outages and disaster situations.

EMC GDDR for SRDF/Star Product Guide

Describes how to use Geographically Dispersed Disaster Restart (GDDR) to automate business recovery following both planned outages and disaster situations.

EMC GDDR for SRDF/Star with AutoSwap Product Guide

Describes how to use Geographically Dispersed Disaster Restart (GDDR) to automate business recovery following both planned outages and disaster situations.

EMC GDDR for SRDF/SQAR with AutoSwap Product Guide

Describes how to use Geographically Dispersed Disaster Restart (GDDR) to automate business recovery following both planned outages and disaster situations.

EMC GDDR for SRDF/A Product Guide

Describes how to use Geographically Dispersed Disaster Restart (GDDR) to automate business recovery following both planned outages and disaster situations.

EMC GDDR Message Guide

Describes the status, warning, and error messages generated by GDDR.

EMC GDDR Release Notes

Describes new features and any known limitations.

EMC z/OS Migrator Product Guide

Describes how to use z/OS Migrator to perform volume mirror and migrator functions as well as logical migration functions.

EMC z/OS Migrator Message Guide

Describes the status, warning, and error messages generated by z/OS Migrator.

EMC z/OS Migrator Release Notes

Describes new features and any known limitations.

EMC ResourcePak for z/TPF Product Guide

Describes how to configure VMAX system control and management in the z/TPF operating environment.

EMC SRDF Controls for z/TPF Product Guide

Describes how to perform remote replication operations in the z/TPF operating environment.

EMC TimeFinder Controls for z/TPF Product Guide

Describes how to perform local replication operations in the z/TPF operating environment.

EMC z/TPF Suite Release Notes

Describes new features and any known limitations.

Special notice conventions used in this document

EMC uses the following conventions for special notices:

⚠ DANGER

Indicates a hazardous situation which, if not avoided, will result in death or serious injury.

⚠ WARNING

Indicates a hazardous situation which, if not avoided, could result in death or serious injury.

⚠ CAUTION

Indicates a hazardous situation which, if not avoided, could result in minor or moderate injury.

NOTICE

Addresses practices not related to personal injury.

Note

Presents information that is important, but not hazard-related.

Typographical conventions

EMC uses the following type style conventions in this document:

Table 1 Typographical conventions used in this content

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

Where to get help

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

Product information

EMC technical support, documentation, release notes, software updates, or information about EMC products can be obtained on the https://support.emc.com site (registration required).

Technical support

To open a service request through the https://support.emc.com site, you must have a valid support agreement. Contact your EMC sales representative for details about obtaining a valid support agreement or to answer any questions about your account.

Additional support options

- Support by Product — EMC offers consolidated, product-specific information on the Web at: https://support.EMC.com/products
  The Support by Product web pages offer quick links to Documentation, White Papers, Advisories (such as frequently used Knowledgebase articles), and Downloads, as well as more dynamic content, such as presentations, discussion, relevant Customer Support Forum entries, and a link to EMC Live Chat.
- EMC Live Chat — Open a Chat or instant message session with an EMC Support Engineer.

eLicensing support

To activate your entitlements and obtain your VMAX license files, visit the Service Center on https://support.EMC.com, as directed on your License Authorization Code (LAC) letter emailed to you.

- For help with missing or incorrect entitlements after activation (that is, expected functionality remains unavailable because it is not licensed), contact your EMC Account Representative or Authorized Reseller.
- For help with any errors applying license files through Solutions Enabler, contact the EMC Customer Support Center.
- If you are missing a LAC letter, or require further instructions on activating your licenses through the Online Support site, contact EMC's worldwide Licensing team at licensing@emc.com or call:
  - North America, Latin America, APJK, Australia, New Zealand: SVC4EMC (800-782-4362) and follow the voice prompts.
  - EMEA: +353 (0) 21 4879862 and follow the voice prompts.

Your comments

Your suggestions help us improve the accuracy, organization, and overall quality of the documentation. Send your comments and feedback to: VMAXContentFeedback@emc.com

Revision history

The following table lists the revision history of this document.

Table 2 Revision history

Revision	Description and/or change	Operating system
07	New content: - RecoverPoint on page 148 - VMAX 950F support (VMAX All Flash 250F, 450F, 850F, and 950F arrays on page 24) - Secure snaps on page 88 - Data at Rest Encryption on page 38	HYPERMAX OS 5977 Q2 2017 SR
06	Revised content: - Power consumption and heat dissipation numbers for the VMAX 250F. - SRDF/Metro array witness overview	HYPERMAX OS 5997.952. 892
05	New content: - VMAX 250F support - Inline compression on page 44 - Mainframe support - Virtual Witness (vWitness) - Non-disruptive-migration	HYPERMAX OS 5997.952. 892
04	Removed "RPQ" requirement from Third Party racking.	HYPERMAX 5977.810.784
03	Updated Licensing appendix.	HYPERMAX 5977.810.784
02	Updated values in the power and heat dissipation specification table.	HYPERMAX OS 5977.691.684 + Q1 2016 Service Pack
01	First release of the VMAX All Flash with EMC HYPERMAX OS 5977 for VMAX 450F, 450FX, 850F, and 850FX.	HYPERMAX OS 5977.691.684 + Q1 2016 Service Pack

CONTENTS

FIGURES

TABLES

CHAPTER 1

VMAX All Flash with HYPERMAX OS

This chapter summarizes VMAX All Flash specifications and describes the features of HYPERMAX OS. Topics include:

Introduction to VMAX All Flash with HYPERMAX OS

VMAX All Flash arrays are engineered to deliver the highest possible flash density by supporting the highest capacity Flash drives.

The power of VMAX All Flash arrays is their flexibility to grow performance and capacity independently to address a massive variety of real world workloads.

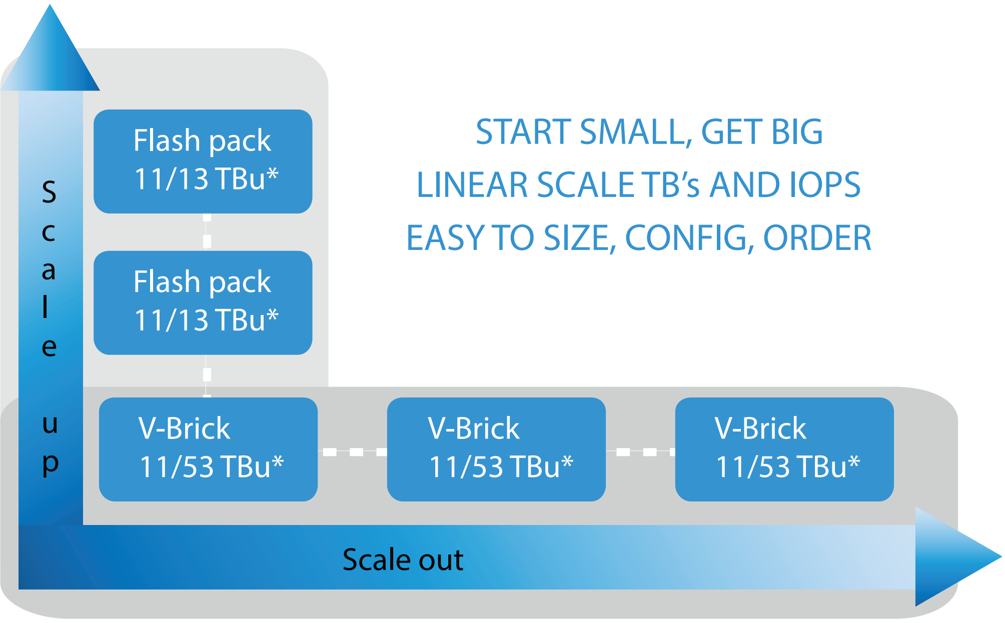
All Flash arrays offer the simplest packaging ever delivered for a VMAX platform. The basic building block is a V-Brick, in open systems arrays; and a zBrick, in mainframe arrays. Depending on the array, this includes:

- An engine with two directors (the redundant data storage processing unit)
- Flash capacity in Drive Array Enclosures (DAEs):
  - VMAX 250F: Two 25-slot DAEs with a minimum base capacity of 13TBu
  - VMAX 450F, VMAX 850F: Two 120-slot DAEs with a minimum base capacity of 53TBu
  - VMAX 950F (open or mixed systems): Two 120-slot DAEs with a minimum base capacity of 53TBu
  - VMAX 950F (mainframe systems): Two 120-slot DAEs with a minimum base capacity of 13TBu
- Multiple software packages are available: F and FX packages for open system arrays and zF and zFX for mainframe arrays.

Customers can scale up the initial configuration by adding 11 TBu (250F) or 13 TBu (450F, 850F, 950F) capacity packs that bundle all required flash capacity and software. In open system arrays, capacity packs are known as Flash capacity packs. In mainframe arrays, capacity packs are known as zCapacity packs. In addition, customers can also scale out the initial configuration by adding additional V-Bricks or zBricks to increase performance, connectivity, and throughput.

Independent and linear scaling of both capacity and performance enables VMAX All Flash to be extremely flexible at addressing varying workloads. For example, the following illustrates scaling opportunities for VMAX 450F, 850F, and 950F open system arrays.

Figure 1 VMAX All Flash scale up and out



* Depending on the VMAX model

VMAX All Flash consists of the following models that combine high scale, low latency, and rich data services.

- VMAX 250F All Flash arrays scale from one to two V-Bricks
- VMAX 450F All Flash arrays scale from one to four V-Bricks/zBricks
- VMAX 850F/950F All Flash arrays scale from one to eight V-Bricks/zBricks

The All Flash arrays:

- Leverage the powerful Dynamic Virtual Matrix Architecture.
- Deliver unprecedented levels of performance and scale. For example, VMAX 950F arrays deliver 6.74M IOPS (RRH) with less than 0.5 ms latency at 150 GB/sec bandwidth. VMAX 250F, 450F, 850F, 950F arrays deliver consistently low response times (< 0.5ms).
- Provide mainframe (VMAX 450F, 850F, 950F) and open systems (including IBM i) host connectivity for mission critical storage needs
- Deliver the power of HYPERMAX OS hypervisor to provide file system storage with eNAS and embedded management services for Unisphere. For more information, refer to Embedded Network Attached Storage on page 36 and Embedded Management on page 35, respectively.
- Offer industry-leading data services such as SRDF remote replication technology with the latest SRDF/Metro functionality, SnapVX local replication services based on SnapVX infrastructure, data protection and encryption, and access to hybrid cloud. For more information, refer to SRDF/Metro on page 137, About TimeFinder on page 84, About CloudArray on page 172, respectively.
- Leverage the latest Flash drive technology in V-Bricks/zBricks and capacity packs of 11 TBu (250F) and 13 TBu (450F, 850F, 950F) to deliver a top-tier diamond service level.

Software packages

VMAX All Flash is available in four models: 250F, 450F, 850F, and 950F. Each model is available with multiple software packages (F/FX for open system arrays, and zF/zFX for mainframe arrays) containing standard and optional features.

Table 3 Symbol legend for VMAX All Flash software features/software package

✓	Standard feature with that model/software package.		✓+	Optional feature with that model/software package.

Table 4 VMAX All Flash software features per model

Software/Feature	VMAX model and software packages									
	250F		450F				850F, 950F			
	F	FX	F	FX	zF	zFX	F	FX	zF	zFX	See:
HYPERMAX OS	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	HYPERMAX OS on page 33
Embedded Management[a]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	Management Interfaces on page 47
Mainframe Essentials Plus					✓	✓			✓	✓	Mainframe Features on page 73
SnapVX	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	About TimeFinder on page 84
AppSync Starter Pack	✓	✓	✓	✓			✓	✓			AppSync on page 55
Compression	✓	✓	✓	✓			✓	✓			Inline compression on page 44
Non-Disruptive Migration	✓	✓	✓	✓			✓	✓			Non-Disruptive Migration overview on page 158
SRDF	✓+	✓	✓+	✓	✓+	✓	✓+	✓	✓+	✓	Remote replication solutions on page 93
SRDF/Metro	✓+	✓	✓+	✓			✓+	✓			SRDF/Metro on page 137
Embedded Network Attached Storage (eNAS)	✓+	✓	✓+	✓			✓+	✓			Embedded Network Attached Storage on page 36
Unisphere 360	✓+	✓	✓+	✓	✓+	✓	✓+	✓	✓+	✓	Unisphere 360 on page 49
ViPR Suite	✓+	✓	✓+	✓			✓+	✓			ViPR suite on page 52
Data at Rest Encryption (D@RE)	✓+	✓	✓+	✓	✓+	✓	✓+	✓	✓+	✓	Data at Rest Encryption on page 38
CloudArray Enabler	✓+	✓	✓+	✓			✓+	✓			CloudArray® for VMAX All Flash on page 171
PowerPath®	✓+	✓[b]	✓+	✓[b]			✓+	✓[b]			PowerPath Migration Enabler on page 164
AppSync Full Suite	✓+	✓+	✓+	✓+			✓+	✓+			AppSync on page 55

Table 4 VMAX All Flash software features per model (continued)

Software/Feature	**VMAX model and software packages**										
	250F		**450F**				**850F, 950F**				
	F	**FX**	**F**	**FX**	**zF**	**zF X**	**F**	**FX**	**zF**	**zF X**	**See:**
ProtectPoint	✓+	✓+	✓+	✓+			✓+	✓+			Backup and restore to external arrays on page 59
AutoSwap and zDP					✓+	✓			✓+	✓	Mainframe SnapVX and zDP on page 90
GDDR					✓+	✓+			✓+	✓+	Geographically Dispersed Disaster Restart (GDDR) on page 51

a. eManagement includes: embedded Unisphere, Solutions Enabler, and SMI-S.

b. The FX package includes 75 PowerPath licenses.. Additional licenses are available separately.

VMAX All Flash 250F, 450F, 850F, and 950F arrays

For open systems, VMAX All Flash arrays range in size from single up to two (250F), four (450F) or eight (850F, 950F) V-Brick systems. For mainframe, VMAX All Flash arrays range in size from a single to four (450F) and up to eight (850F, 950F) zBricks. V-Bricks/zBricks and high capacity disk enclosures are consolidated in the same system bay, providing a dramatic increase in floor tile density.

VMAX All Flash arrays are built on a scalable architecture. Additional capacity is available as Flash Capacity Packs (open system arrays) and zCapacity packs (mainframe arrays). Additional processing power is available as V-Bricks (open system arrays) and zBricks (mainframe arrays).

VMAX All Flash arrays come fully pre-configured from the factory, significantly reducing time to first I/O at installation.

VMAX All Flash array features include:

- All Flash configuration
- For 450F, 850F, and 950F arrays:
  - System bay dispersion of up to 82 feet (25 meters) from the first system bay[1]
  - Each system bay can house either one or two V-Bricks/zBricks

[1] Available through RPQ only.

VMAX All Flash 250F, 450F, 850F, and 950F specifications

The following tables list specifications for each VMAX All Flash model.

Table 5 V-Brick/zBrick specifications

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
Number of V-Bricks/zBricks supported	1 to 2	1 to 4	1 to 8	1 to 8
Engine enclosure	4u	4u	4u	4u
CPU	Intel Xeon E5-2650-v4, 2.2 GHz 12 core	Intel Xeon E5-2650-v2, 2.6 GHz 8 core	Intel Xeon E5-2697-v2, 2.7 GHz 12 core	Intel Xeon E5-2697-v4, 2.3 GHz 18 core
# Cores per CPU/per engine/per system	12/48/96	8/32/128	12/48/384	18/72/576
Dynamic Virtual Matrix Interconnect	Direct Connect 56Gbps per port	InfiniBand Dual Redundant Fabric: 56Gbps per port	InfiniBand Dual Redundant Fabric: 56Gbps per port	InfiniBand Dual Redundant Fabric: 56Gbps per port
Vault				
Vault strategy	Vault to Flash	Vault to Flash	Vault to Flash	Vault to Flash
Vault implementation	2 to 4 NVMe Flash Modules/Engine	4 to 8 Flash Modules/Engine	4 to 8 Flash Modules/Engine	4 to 8 NVMe Flash Modules/Engine

Table 6 Cache specifications

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
Cache: System Min (raw)	512GB	1024GB	1024GB	1024GB
Cache: System Max (raw)	4TB (with 2048GB engine)	8TB (with 2048GB engine)	16TB (with 2048GB engine)	16TB (with 2048GB engine)
Cache: Per-Engine options	512GB, 1024GB, 2048GB	1024GB, 2048GB	1024GB, 2048GB	1024GB, 2048GB

Table 7 Front-end I/O modules

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
Max front-end I/O modules/V-Brick/ zBrick	8	6 (up to 8 on mainframe)	6 (up to 8 on mainframe)	6 (up to 8 on mainframe)
Front-end I/O modules and protocols supported (Optical Ports)	FC: 4 x 8Gbs (FC, SRDF) FC: 4 x 16Gbs (FC, SRDF) 10GbE: 4 x 10GbE (iSCSI, SRDF) GbE: 4 x 1GbE (2 Cu/2 Opt SRDF)	FC: 4 x 8Gbs (FC, SRDF) FC: 4 x 16Gbs (FC, SRDF) iSCSI: 4 x 10GbE (iSCSI) 10GbE: 2 x 10GbE (SRDF) GbE: 4 x 1Gbe (2 Cu/2 Opt SRDF) FICON 4 x 16Gbs (FICON)	FC: 4 x 8Gbs (FC, SRDF) FC: 4 x 16Gbs (FC, SRDF) iSCSI: 4 x 10GbE (iSCSI) 10GbE: 2 x 10GbE (SRDF) GbE: 4 x 1Gbe (2 Cu/2 Opt SRDF) FICON 4 x 16Gbs (FICON)	FC: 4 x 8Gbs (FC, SRDF) FC: 4 x 16Gbs (FC, SRDF) 10GbE: 4 x 10GbE (iSCSI, SRDF) GbE: 4 x 1GbE (2 Cu/2 Opt SRDF) FICON: 4 x 16GBs (FICON)

Table 8 eNAS I/O modules

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
Max number eNAS I/O Modules/ Software Data Mover [a]	2	2	2	2
eNAS I/O modules supported[b]	10GbE: 2 x 10GbE Opt 10GbE: 2 x 10GbE Cu 8Gbs: 4 x 8Gbs FC (Tape BU)	10GbE: 2 x 10GbE Opt 10GbE: 2 x 10GbE Cu GbE: 4 x 1GbE Cu 8Gbs: 4 x 8Gbs FC (Tape BU)	10GbE: 2 x 10GbE Opt 10GbE: 2 x 10GbE Cu GbE: 4 x 1GbE Cu 8Gbs: 4 x 8Gbs FC (Tape BU)	10GbE: 2 x 10GbE Opt 10GbE: 2 x 10GbE Cu 8Gbs: 4 x 8Gbs FC (Tape BU)

a. Maximum number of supported eNAS I/O module types/Data Mover, or support for eight Data Movers on the VMAX 850F are available by request.

b. Quantity one (1) 2 x 10GbE Optical module is the default choice/Data Mover.

Table 9 eNAS Software Data Movers

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
Max Software Data Movers	4 (3 Active + 1 Standby) (4 Data Movers requires minimum 2 V-Bricks)	4 (3 Active + 1 Standby) (4 Data Movers requires minimum 2 V-Bricks/zBricks)	8 (7 Active + 1 Standby) (8 Data Movers requires minimum 4 V-Bricks)	8 (7 Active + 1 Standby) (8 Data Movers requires minimum 4 V-Bricks)
Max NAS capacity/array	1.1 PBu (cache limited)	1.5 PBu	3.5 PBu	3.5 PBu

Table 10 Capacity, drives

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
Max Capacity per Array[a]	1.16PBe	2.3PBe	4.4PBe	4.42PBe
Base capacity per V-Brick	13.2TBu[b]	52.6TBu	52.6TBu	56.6TBu
Base capacity per zBrick	N/A	52.6 TBu	52.6TBu	13.2TBu
Incremental Capacity Blocks	13.2TBu[b]	13.2TBu	13.2TBu	13.2TBu
Max drives per V-Brick/zBrick	50	240	240	240
Max drives per array	100	960	1920	1920
Max drives per system bay	100/200[c]	480	480	480
Min drive count per V-Brick/zBrick	8 + 1 spare	16 + 1 spare	16 + 1 spare	16 + 1 spare

a. Max capacity per array based on over provisioning ratio of 1.0.

b. 13.2TBu is based on RAID 5(7+1) and 11.3 TBu is possible with RAID 5(3+1): 313.2TBu V-Brick and Capacity Block usable capacities are based on RAID 5 (7+1). 11.3TBu base capacity and Capacity Block increments possible with RAID 5(3+1) on VMAX 250F.

c. Two hundred drives are supported in a single cabinet when two systems are packaged in the same rack.

Table 11 Flash Drive specifications

	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
Flash drives supported (2.5")	960GB, 1.92TB, 3.84TB, 7.68TB, 15.36TB	960GB, 1.92TB, 3.84TB	960GB, 1.92TB, 3.84TB	960GB, 1.92TB, 3.84TB, 7.68TB, 15.36TB
Back-end interface	12Gbps SAS	6Gbps SAS	6Gbps SAS	6Gbps SAS
RAID options	RAID 5 (7+1) (Default) RAID 5 (3+1) RAID 6 (6+2)	RAID 5 (7+1) RAID 6 (14+2)	RAID 5 (7+1) RAID 6 (14+2)	RAID 5 (7+1) RAID 6 (14+2)
Mixed RAID Group Support	No	No	No	No
Support for Mixed Drive Capacities	Yes	Yes	Yes	Yes

Table 12 Flash Array Enclosure

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
120 x 2.5" drive DAE	No	Yes	Yes	Yes
25 x 2.5" drive DAE	Yes	No	No	No

Table 13 Cabinet configurations

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
Standard 19" bays	Yes	Yes	Yes	Yes

Table 13 Cabinet configurations (continued)

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
Single V-Brick/zBrick System Bay Configuration	No (Packaging based on Dual V-Bricks, but initial V-Brick in each system bay supported)	No (Packaging based on Dual V-Bricks/zBricks, but initial V-Brick/zBrick in each system bay supported)	No (Packaging based on Dual V-Bricks/zBricks, but initial V-Brick/zBrick in each system bay supported)	No (Packaging based on Dual V-Bricks, but initial V-Brick in each system bay supported)
Dual V-Brick/zBrick System Bay Configuration	Yes (Default packaging)	Yes (Default packaging)	Yes (Default packaging)	Yes (Default packaging)
Third-party rack mount option	Yes	Yes	Yes	Yes

Table 14 Dispersion specifications

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
System bay dispersion	N/A (single floor tile system)	Yes, with RPQ only Up to 82 feet (25m) between System Bay 1 and any other System Bay	Yes, with RPQ only Up to 82 feet (25m) between System Bay 1 and any other System Bay	Yes, with RPQ only Up to 82 feet (25m) between System Bay 1 and any other System Bay

Table 15 Pre-configuration

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
100% thin provisioned	Yes	Yes	Yes	yes

Table 16 Host support

Feature	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
Open systems	Yes	Yes	Yes	Yes
Mainframe (CKD 3380 and 3390 emulation)	No	Yes	Yes	Yes
Mixed open system and mainframe	No	No	No	Yes

Table 17 Supported I/O protocols

I/O protocols	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
8 Gb/s FC Host/SRDF ports				
Maximum/V-Brick	32	24	24	24
Maximum/array	64	96	192	192
16 Gb/s FC Host ports				
Maximum/V-Brick	32	24	24	24
Maximum/array	64	96	192	192
16 Gb/s FICON ports				
Maximum/V-Brick	N/A	32	32	32
Maximum/array	N/A	128	256	256
10GbE iSCSI ports				
Maximum/V-Brick	32	24	24	24
Maximum/array	64	96	192	192
10GbE SRDF ports (Optical)				
Maximum/V-Brick	32	12	12	24
Maximum/array	64	48	96	192
GbE SRDF ports (Optical/Cu)				
Maximum/V-Brick	16/16	12/12	12/12	12/12
Maximum/array	64	48	96	96

Table 18 eNAS ports

I/O protocols	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
10GbE Optical ports				
Maximum ports/Data Mover	4	2	2	4
Maximum ports/array	16	8	16	32
10GbE Copper ports [a]				
Maximum ports/Data Mover	4	2	2	4
Maximum ports/array	16	8	16	32
1GbE Copper ports #GUID-DE3B84E2-228B-4D16-9126-567020A09DD6/DFJDJASFJDSKJFKLJLASDJ				
Maximum ports/Data Mover	N/A	4	4	N/A
Maximum ports/array	N/A	16	32	N/A
8Gb/s Tape Back Up ports #GUID-DE3B84E2-228B-4D16-9126-567020A09DD6/DFJDJASFJDSKJFKLJLASDJ				
Maximum ports/Data Mover	2	2	2	2

Table 18 eNAS ports (continued)

I/O protocols	VMAX 250F	VMAX 450F	VMAX 850F	VMAX 950F
Maximum ports/array	8	8	16	16

a. Available by request.

Flash Drive support

VMAX All Flash arrays supports the latest dual-ported native SAS Flash drives (VMAX 250F supports 12Gb/s drives; 450F, 850F, and 950F support 6Gb/s drives). All Flash drives support two independent I/O channels with automatic failover and fault isolation. VMAX 950F arrays also support mixed FBA and CKD drive configurations.

Check with your EMC sales representative for the latest list of supported drives and types. All capacities are based on 1 GB = 1,000,000,000 bytes. Actual usable capacity may vary depending upon configuration.

Table 19 2.5" Flash drives used in V-Bricks/zBricks and capacity blocks

Capacity type	VMAX 250F, 450F, 850F, 950F			VMAX 250F, 950F	
Nominal capacity (GB)[a]	960	1920	3840	7680	15360
Raw capacity (GB)	960	1920	3840	7680	15360
Open systems formatted capacity (GB)[b]	938.41	1879.64	3761.03	7522.06	15047.65
Mainframe formatted capacity (GB)[c]	940.26	1880.52	3761.80	7523.61	15047.98

a. Additional Drive Capacity Blocks and V-Bricks/zBricks in any given configuration could contain different underlying drive sizes in order to achieve the desired Usable Capacity. This is automatically optimized by the VMAX Sizer Configuration Tool.
b. Open Systems Formatted Capacity is also referred to a TBu in this document.
c. Mainframe not supported on VMAX 250F.

Power consumption

Table 20 Power consumption and heat dissipation

	VMAX 250F		VMAX 450F		VMAX 850F		VMAX 950F	
Maximum power and heat dissipation at <26°C and >35°C [a]	Maximum total power consumption <26°C / >35°C (kVA)	Maximum heat dissipation <26°C / >35°C (Btu/Hr)	Maximum total power consumption <26°C / >35°C (kVA)	Maximum heat dissipation <26°C / >35°C (Btu/Hr)	Maximum total power consumption <26°C / >35°C (kVA)	Maximum heat dissipation <26°C / >35°C (Btu/Hr)	Maximum total power consumption <26°C / >35°C (kVA)	Maximum heat dissipation <26°C / >35°C (Btu/Hr)
System bay 1 Dual V-Brick	4.13 / 5.19	14,090 / 17,698	6.69 / 9.05	22,813 / 30,861	6.94 / 9.30	23,665 / 31,713	7.25 / 9.61	24,712 / 32,760
System bay 2 Dual V-Brick[b]	N/A		6.28 / 8.38	21,415 / 28,576	6.49 / 8.59	22,131 / 29,292	6.80 / 8.90	23,178 / 30,339

a. Power values and heat dissipations shown at >35°C reflect the higher power levels associated with both the battery recharge cycle, and the initiation of high ambient temperature adaptive cooling algorithms. Values at <26°C are reflective of more steady state maximum values during normal operation.
b. Power values for system bay 2 and all subsequent system bays where applicable.

Space and weight requirements

Table 21 Space and weight requirements, VMAX 250F

Bay configurations [a]	Height (in/cm) [b]	Width (in/cm)	Depth[c] (in/cm)	Weight (max lbs/kg)
1 system, 1 V-Brick	75/190	24/61	42 in (106.7 cm)	570/258

Table 21 Space and weight requirements, VMAX 250F (continued)

Bay configurations [a]	Height (in/cm) [b]	Width (in/cm)	Depth[c] (in/cm)	Weight (max lbs/kg)
1 system, 2 V-Bricks, or 2 systems, 1 V-Brick each	75/190	24/61	42 in (106.7 cm)	850/385
2 systems, 2 V-Bricks in one system, 1 V-Brick in other	75/190	24/61	42 in (106.7 cm)	1130/513
2 systems, 2 V-Bricks each system	75/190	24/61	42 in (106.7 cm)	1410/640

a. Clearance for service/airflow is the front at 42 in (106.7 cm) front and the rear at 30 in (76.2 cm).
b. An additional 18 in (45.7 cm) is recommended for ceiling/top clearance.
c. Includes rear door.

Table 22 Space and weight requirements, VMAX 450F, VMAX 850F, VMAX 950F

Bay configurations [a]	Height[b] (in/cm)	Width[c] (in/cm)	Depth[d] (in/cm)	Weight (max lbs/kg)
System bay	75/190	24/61	47/119	1860/844

a. Clearance for service/airflow is the front at 42 in (106.7 cm) front and the rear at 30 in (76.2 cm).
b. An additional 18 in (45.7 cm) is recommended for ceiling/top clearance.
c. Measurement includes .25 in. (0.6 cm) gap between bays.
d. Includes front and rear doors.

Radio frequency interference specifications

Electro-magnetic fields, which include radio frequencies can interfere with the operation of electronic equipment. EMC Corporation products have been certified to withstand radio frequency interference (RFI) in accordance with standard EN61000-4-3. In Data Centers that employ intentional radiators, such as cell phone repeaters, the maximum ambient RF field strength should not exceed 3 Volts /meter.

Table 23 Minimum distance from RF emitting devices

Repeater power level[a]	Recommended minimum distance
1 Watt	9.84 ft (3 m)
2 Watt	13.12 ft (4 m)
5 Watt	19.69 ft (6 m)
7 Watt	22.97 ft (7 m)
10 Watt	26.25 ft (8 m)
12 Watt	29.53 ft (9 m)
15 Watt	32.81 ft (10 m)

a. Effective Radiated Power (ERP)

HYPERMAX OS

This section highlights the features of the HYPERMAX OS.

What's new in HYPERMAX OS 5977 Q2 2017

This section describes new functionality and features provided by HYPERMAX OS 5977 Q2 2017 for VMAX All Flash arrays.

RecoverPoint
HYPERMAX OS 5977 Q2 2017 SR introduces support for RecoverPoint on VMAX storage arrays. RecoverPoint is a comprehensive data protection solution designed to provide production data integrity at local and remote sites. RecoverPoint also provides the ability to recover data from any point in time using journaling technology.

RecoverPoint on page 148 provides more information.

Secure snaps
Secure snaps is an enhancement to the current snapshot technology. Secure snaps prevent administrators or other high-level users from intentionally or unintentionally deleting snapshot data. In addition, secure snaps are also immune to automatic failure resulting from running out of Storage Resource Pool (SRP) or Replication Data Pointer (RDP) space on the array.

Secure snaps on page 88 provides more information.

Support for VMAX 950F
The VMAX 950F All Flash array is designed to meet the needs of high-end enterprise space. VMAX 950F scales from one to eight V-Bricks/zBricks and provides a maximum of 4PB effective capacity.

VMAX All Flash 250F, 450F, 850F, and 950F arrays on page 24 provides more information.

Data at Rest Encryption
Data at Rest Encryption (D@RE) now supports the OASIS Key Management Interoperability Protocol (KMIP) and can integrate with external servers that also support this protocol. This release has been validated to interoperate with the following KMIP-based key managers:

- Gemalto SafeNet KeySecure
- IBM Security Key Lifecycle Manager

Data at Rest Encryption on page 38 provides more information.

Mixed FBA/CKD drive support for VMAX 950F arrays
HYPERMAX OS 5977 Q2 2017 SR introduces support for mixed FBA and CKD drive configurations.

Flash Drive support on page 31 provides more information.

HYPERMAX OS emulations

HYPERMAX OS provides emulations (executables) that perform specific data service and control functions in the HYPERMAX environment. The following table lists the available emulations.

Table 24 HYPERMAX OS emulations

Area	Emulation	Description	Protocol Speed[a]
Back-end	DS	Back-end connection in the array that communicates with the drives, DS is also known as an internal drive controller.	SAS 12 Gb/s (VMAX 250F) SAS 6 Gb/s (VMAX 450F, 850F, and 950F)
	DX	Back-end connections that are not used to connect to hosts. Used by ProtectPoint and Cloud Array.	FC 16 or 8 Gb/s
		ProtectPoint links Data Domain to the array. DX ports must be configured for FC protocol.	
Management	IM	Separates infrastructure tasks and emulations. By separating these tasks, emulations can focus on I/O-specific work only, while IM manages and executes common infrastructure tasks, such as environmental monitoring, Field Replacement Unit (FRU) monitoring, and vaulting.	N/A
	ED	Middle layer used to separate front-end and back-end I/O processing. It acts as a translation layer between the front-end, which is what the host knows about, and the back-end, which is the layer that reads, writes, and communicates with physical storage in the array.	N/A
Host connectivity	FA - Fibre Channel SE - iSCSI EF - FICON [b]	Front-end emulation that: • Receives data from the host (network) and commits it to the array • Sends data from the array to the host/network	FC - 16 or 8 Gb/s SE - 10 Gb/s EF - 16 Gb/s

Table 24 HYPERMAX OS emulations (continued)

Area	Emulation	Description	Protocol Speed[a]
Remote replication	RF - Fibre Channel RE - GbE	Interconnects arrays for Symmetrix Remote Data Facility (SRDF).	RF - 8 Gb/s SRDF RE - 1 GbE SRDF RE - 10 GbE SRDF

a. The 8 Gb/s module auto-negotiates to 2/4/8 Gb/s and the 16 Gb/s module auto-negotiates to 16/8/4 Gb/s using optical SFP and OM2/OM3/OM4 cabling.

b. Only on VMAX 450F, 850F, and 950F arrays.

Container applications

HYPERMAX OS provides an open application platform for running data services. HYPERMAX OS includes a light-weight hypervisor that enables multiple operating environments to run as virtual machines on the storage array.

Application containers are virtual machines that provide embedded applications on the storage array. Each container virtualizes hardware resources required by the embedded application, including:

- Hardware needed to run the software and embedded application (processor, memory, PCI devices, power management)
- VM ports, to which LUNs are provisioned
- Access to necessary drives (boot, root, swap, persist, shared)

Embedded Management

The eManagement container application embeds management software (Solutions Enabler, SMI-S, Unisphere for VMAX) on the storage array, enabling you to manage the array without requiring a dedicated management host.

With eManagement, you can manage a single storage array and any SRDF attached arrays. To manage multiple storage arrays with a single control pane, use the traditional host-based management interfaces, Unisphere for VMAX and Solutions Enabler. To this end, eManagement allows you to link-and-launch a host-based instance of Unisphere for VMAX.

eManagement is typically pre-configured and enabled at the EMC factory, thereby eliminating the need for you to install and configure the application. However, starting with HYPERMAX OS 5977.945.890, eManagement can be added to VMAX arrays in the field. Contact your EMC representative for more information.

Embedded applications require system memory. The following table lists the amount of memory unavailable to other data services.

Table 25 eManagement resource requirements

VMAX All Flash model	CPUs	Memory	Devices supported
VMAX 250F	4	16 GB	200K
VMAX 450F	4	16 GB	200K
VMAX 850F, 950F	4	20 GB	400K

Virtual Machine ports

Virtual machine (VM) ports are associated with virtual machines to avoid contention with physical connectivity. VM ports are addressed as ports 32-63 per director FA emulation.

LUNs are provisioned on VM ports using the same methods as provisioning physical ports.

A VM port can be mapped to one and only one VM.

A VM can be mapped to more than one port.

Embedded Network Attached Storage

Embedded Network Attached Storage (eNAS) is fully integrated into the VMAX All Flash array. eNAS provides flexible and secure multi-protocol file sharings (NFS 2.0, 3.0, 4.0/4.1), CIFS/SMB 3.0) and multiple file server identities (CIFS and NFS serves). eNAS enables:

- File server consolidation/multi-tenancy
- Built-in asynchronous file level remote replication (File Replicator)
- Built-in Network Data Management Protocol (NDMP)
- VDM Synchronous replication with SRDF/S and optional automatic failover manager
  File Auto Recovery (FAR) with optional File Auto Recover Manager (FARM)
- Anti-virus

eNAS provides file data services that enable customers to:

- Consolidate block and file storage in one infrastructure
- Eliminate the gateway hardware, reducing complexity and costs
- Simplify management

Consolidated block and file storage reduces costs and complexity while increasing business agility. Customers can leverage rich data services across block and file storage including storage provisioning, dynamic Host I/O Limits, and Data at Rest Encryption.

eNAS solutions and implementation

The eNAS solution runs on standard array hardware and is typically pre-configured at the factory. In this scenario, EMC provides a one-time setup of the Control Station and Data Movers, containers, control devices, and required masking views as part of the factory eNAS pre-configuration. Additional front-end I/O modules are required to implement eNAS. However, starting with HYPERMAX OS 5977.945.890, eNAS can be added to VMAX arrays in the field. Contact your EMC representative for more information.

eNAS uses the HYPERMAX OS hypervisor to create virtual instances of NAS Data Movers and Control Stations on VMAX All Flash controllers. Control Stations and Data Movers are distributed within the VMAX All Flash based upon the number of engines and their associated mirrored pair.

By default, VMAX All Flash arrays are configured with:

- Two Control Station virtual machines
- Data Mover virtual machines. The number of Data Movers varies by array size:

- VMAX 250F = Two (default), four (maximum, requires two V-Bricks)
- VMAX 450F = Two (default), four (maximum)
- VMAX 850F, 950F = Two (default), four, six, or eight (six and eight configurations only available by RPQ)

All configurations include one standby Data Mover.

eNAS configurations

The storage capacity required for arrays supporting eNAS is the same (~ 680 GB).

The following table lists eNAS configurations and front-end I/O modules.

Table 26 eNAS configurations by array

Component	Description	VMAX 250F	VMAX 450F	VMAX 850F, 950F
Data movers[a] virtual machine	Maximum number	4	4	8[b]
	Max capacity/DM	512 TB	512 TB	512 TB
	Logical cores[c]	12/24	12/24	16/32/48/64[b]
	Memory (GB)[c]	48/96	48/96	48/96/144/192 [b]
	I/O modules (Max)[c]	12	12[d]	24[d]
Control Station virtual machines (2)	Logical cores	2	2	2
	Memory (GB)	8	8	8
NAS Capacity/ Array	Maximum	1.15 PB	1.5 PB	3.5 PB

a. Data movers are added in pairs and must support the same configuration.

b. The 850F and 950F can be configured through Sizer with a maximum of four data movers. However, six and eight data movers can be ordered by RPQ. As the number of data movers increases, the maximum number of I/O cards , logical cores, memory, and maximum capacity also increases.

c. For 2, 4, 6, and 8 data movers, respectively.

d. A single 2-port 10GbE Optical I/O module is configured per Data Mover for initial All-Flash configurations. However, that I/O module can be replaced with a different I/O module (i.e., 4-port 1GbE or 2-port 10GbE copper) using the normal replacement capability that exists with any eNAS Data Mover I/O module. In addition, additional I/O modules can be configured via a I/O module upgrade/add as long as standard rules are followed. (can't exceed more than 3 I/O modules per Data Mover, all I/O modules must be in the same slot on each director on which a Data Mover resides).

Replication using eNAS

The following replication methods are available for eNAS file systems:

- Asynchronous file system level replication using VNX Replicator for File. Refer to *Using VNX Replicator 8.x*.
- Synchronous replication with SRDF/S using File Auto Recovery (FAR) with the optional File Auto Recover Manager (FARM).
- Checkpoint (point-in-time, logical images of a production file system) creation and management using VNX SnapSure.

Refer to *Using VNX SnapSure 8.x*.

Note

SRDF/A, SRDF/Metro, and TimeFinder are not available with eNAS.

eNAS management interface

eNAS block and file storage is managed using the Unisphere for VMAX File Dashboard. Link and launch enables you to run the block and file management GUI within the same session.

The configuration wizard helps you create storage groups (automatically provisioned to the Data Movers) quickly and easily. Creating a storage group creates a storage pool in Unisphere for VNX that can be used for file level provisioning tasks.

Data protection and integrity

HYPERMAX OS provides a suite of integrity checks, RAID options, and vaulting capabilities to ensure data integrity and to protect data in the event of a system failure or power outage.

VMAX All Flash arrays support the following RAID levels at the array level:

- VMAX 250F: RAID5 (7+1) (Default), RAID5 (3+1) and RAID6 (6+2)
- VMAX 450F, 850F, 950F: RAID5 (7+1) and RAID6 (14+2)

Data at Rest Encryption

Securing sensitive data is one of the greatest challenges faced by many enterprises. Increasing regulatory and legislative demands and the constantly changing threat landscape have brought data security to the forefront of IT issues. Several of the most important data security threats are related to protection of the storage environment. Drive loss and theft are primary risk factors. EMC® Data at Rest Encryption (D@RE) protects data confidentiality by adding back-end encryption to the entire array.

D@RE provides hardware-based, on-array, back-end encryption for VMAX arrays by using SAS I/O modules that incorporate AES-XTS inline data encryption. These modules encrypt and decrypt data as it is being written to or read from disk, thus protecting your information from unauthorized access even when disk drives are removed from the array.

D@RE supports either an internal embedded key manager, or an external, enterprise-grade key manager accessible through Key Management Interoperability Protocol (KMIP). The following external key managers are supported:

- SafeNet KeySecure by Gemalto
- IBM Security Key Lifecycle Manager

Note

For supported external key manager and HYPERMAX OS versions, refer to the EMC E-Lab Interoperability Matrix (https://www.emc.com/products/interoperability/elab.htm).

When D@RE is enabled, all configured drives are encrypted, including data drives, spares, and drives with no provisioned volumes. Vault data is encrypted on Flash I/O modules.

D@RE enables:

- Secure replacement for failed drives that cannot be erased.
  For some types of drive failures, data erasure is not possible. Without D@RE, if the failed drive is repaired, data on the drive may be at risk. With D@RE, simply delete the applicable keys, and the data on the failed drive is unreadable.
- Protection against stolen drives.
  When a drive is removed from the array, the key stays behind, making data on the drive unreadable.
- Faster drive sparing.
  The drive replacement script destroys the keys associated with the removed drive, quickly making all data on that drive unreadable.
- Secure array retirement.
  Simply delete all copies of keys on the array, and all remaining data is unreadable.

D@RE is compatible with all array features and all supported drive types or volume emulations. Encryption is a powerful tool for enforcing your security policies. D@RE delivers encryption without degrading performance or disrupting your existing applications and infrastructure.

Enabling D@RE

D@RE is a licensed feature, and is pre-configured and installed at the factory. The process to upgrade an existing array to use D@RE is disruptive and requires re-installing the array, and may involve a full data back up and restore. Before you upgrade, you must plan how to manage any data already on the array. EMC Professional Services offers services to help you upgrade to D@RE.

D@RE components

Embedded D@RE (Figure 2 on page 40) uses the following components, all of which reside on the primary Management Module Control Station (MMCS):

- RSA Embedded Data Protection Manager (eDPM)— Embedded key management platform, which provides onboard encryption key management functions, such as secure key generation, storage, distribution, and audit.
- RSA BSAFE® cryptographic libraries— Provides security functionality for RSA eDPM Server (embedded key management) and the EMC KTP client (external key management).
- Common Security Toolkit (CST) Lockbox— Hardware- and software-specific encrypted repository that securely stores passwords and other sensitive key manager configuration information. The lockbox binds to a specific MMCS.

External D@RE (Figure 3 on page 40) uses the same components as embedded, and adds the following:

- EMC Key Trust Platform (KTP)— Also known as the KMIP Client, this component resides on the MMCS and communicates via the OASIS Key Management Interoperability Protocol (KMIP) with external key managers to manage encryption keys.
- External Key Manager— Provides centralized encryption key management capabilities such as secure key generation, storage, distribution, audit, and enabling Federal Information Processing Standard (FIPS) 140-2 level 3 validation with High Security Module (HSM).
- Cluster/Replication Group— Multiple external key managers sharing configuration settings and encryption keys. Configuration and key lifecycle changes made to one node are replicated to all members within the same cluster or replication group.

Figure 2 D@RE architecture, embedded

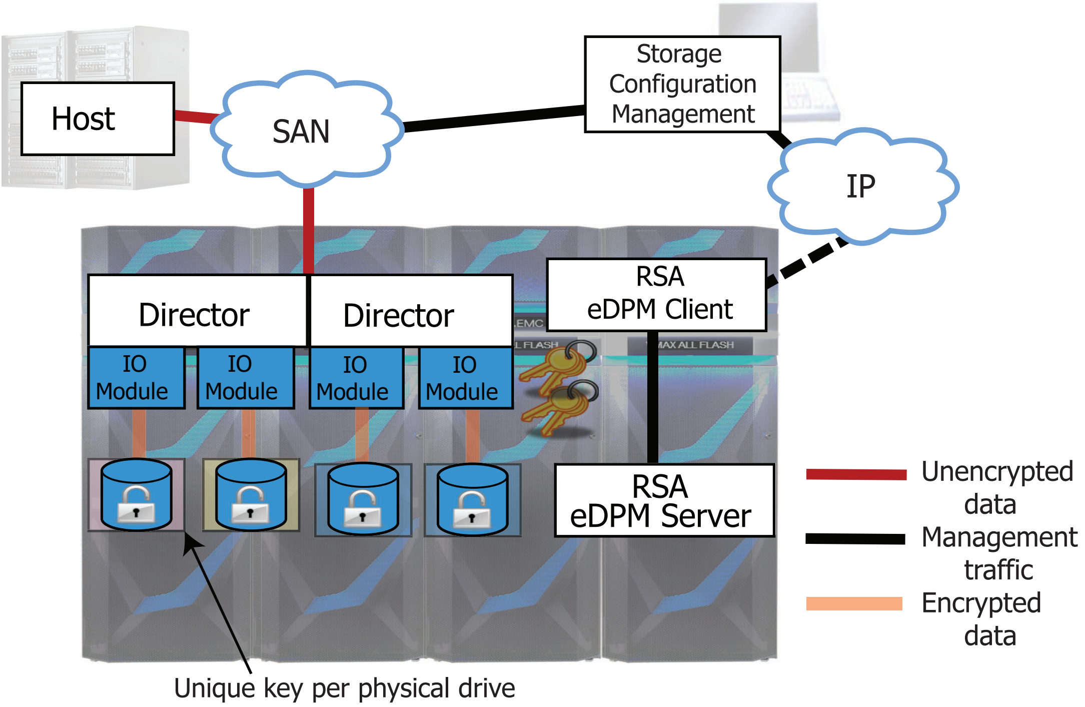
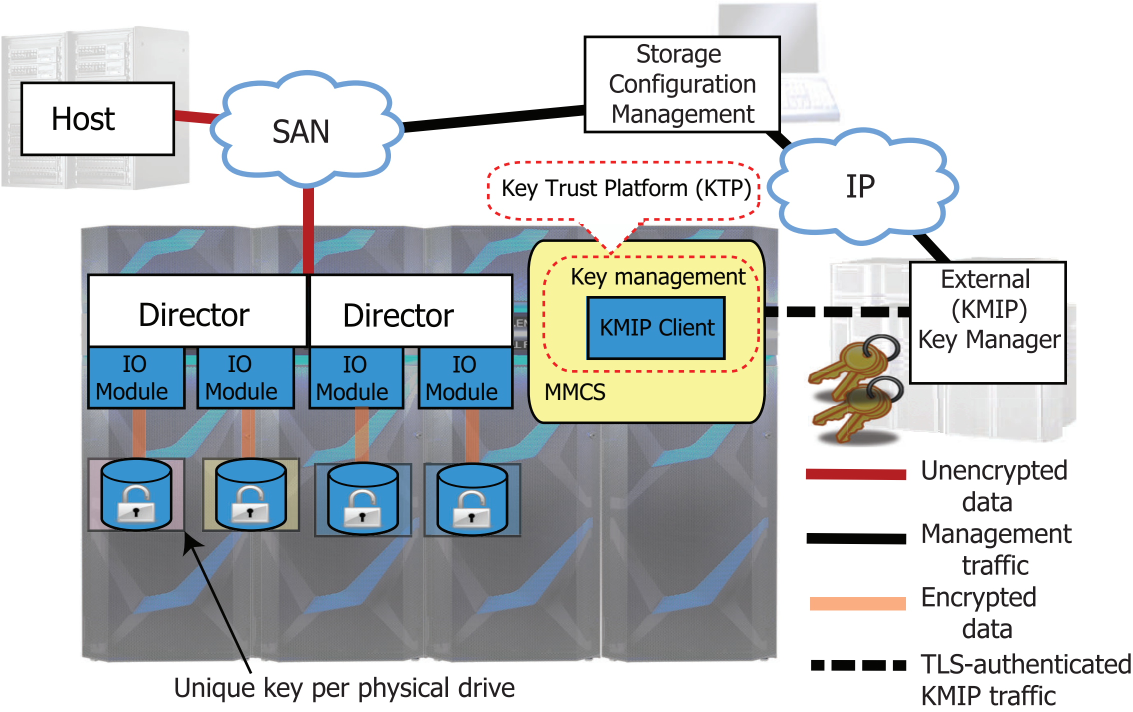


Figure 3 D@RE architecture, external



External Key Managers

D@RE's external, enterprise-grade key management is provided by Gemalto SafeNet KeySecure and IBM Security Key Lifecycle Manager. Keys are generated and distributed using the best practices as defined by industry standards (NIST 800-57 and ISO 11770). With D@RE, there is no need to replicate keys across volume snapshots or remote sites. D@RE external key managers can be used with either a FIPS 140-2 level 3 validated HSM, in the case of Gemalto SafeNet KeySecure, or FIPS 140-2 level 1 validated software, in the case of IBM Security Key Lifecycle Manager.

Encryption keys must be both highly available when they are needed, and tightly secured. Keys, and the information required to use keys (during decryption), must be preserved for the lifetime of the data. This is critical for encrypted data that is kept for many years.

Encryption keys must be accessible. Key accessibility is vital in high-availability environments. D@RE caches the keys locally so that connection to the Key Manager is required only for operations such as the initial installation of the array, replacement of a drive, or drive upgrades.

Key lifecycle events (generation and destruction) are recorded in the VMAX Audit Log.

Key protection

The local keystore file is encrypted with a 256-bit AES key derived from a randomly generated password, and stored in the Common Security Toolkit (CST) Lockbox, which leverages RSA's BSAFE technology. The Lockbox is protected using MMCS-specific stable system values of the primary MMCS. These are the same SSVs that protect Secure Service Credentials (SSC).

Compromising the MMCS's drive or copying Lockbox/keystore files off the array causes the SSV tests to fail. Compromising the entire MMCS only gives an attacker access if they also successfully compromise SSC.

There are no backdoor keys or passwords to bypass D@RE security.

Key operations

D@RE provides a separate, unique Data Encryption Key (DEK) for each drive in the array, including spare drives. The following operations ensure that D@RE uses the correct key for a given drive:

- DEKs stored in the VMAX array include a unique key tag and key metadata when they are wrapped (encrypted) for use by the array.
  This information is included with the key material when the DEK is wrapped (encrypted) for use in the array.
- During encryption I/O, the expected key tag associated with the drive is supplied separately from the wrapped key.
- During key unwrap, the encryption hardware checks that the key unwrapped properly and that it matches the supplied key tag.
- Information in a reserved system LBA (Physical Information Block, or PHIB) verifies the key used to encrypt the drive and ensures the drive is in the correct location.
- During initialization, the hardware performs self-tests to ensure that the encryption/decryption logic is intact.
  The self-test prevents silent data corruption due to encryption hardware failures.

Audit logs

The audit log records major activities on the VMAX All Flash array, including:

- Host-initiated actions
- Physical component changes
- Actions on the MMCS
- D@RE key management events

- Attempts blocked by security controls (Access Controls)

The Audit Log is secure and tamper-proof. Event contents cannot be altered. Users with the Auditor access can view, but not modify, the log.

Data erasure

EMC Data Erasure uses specialized software to erase information on arrays. Data erasure mitigates the risk of information dissemination, and helps secure information at the end of the information lifecycle. Data erasure:

- Protects data from unauthorized access
- Ensures secure data migration by making data on the source array unreadable
- Supports compliance with internal policies and regulatory requirements

Data Erasure overwrites data at the lowest application-addressable level to drives. The number of overwrites is configurable from 3x (the default) to 7x with a combination of random patterns on the selected arrays.

Overwrite is supported. An optional certification service is available to provide a certificate of erasure. Drives that fail erasure are delivered to customers for final disposition.

For individual Flash drives, Secure Erase operations erase all physical flash areas on the drive which may contain user data.

EMC offers the following data erasure services:

- EMC Data Erasure for Full Arrays — Overwrites data on all drives in the system when replacing, retiring or re-purposing an array.
- EMC Data Erasure/Single Drives — Overwrites data on individual drives.
- EMC Disk Retention — Enables organizations that must retain all media to retain failed drives.
- EMC Assessment Service for Storage Security — Assesses your information protection policies and suggests a comprehensive security strategy.

All erasure services are performed on-site in the security of the customer's data center and include a Data Erasure Certificate and report of erasure results.

Block CRC error checks

HYPERMAX OS supports and provide:

- Industry standard T10 Data Integrity Field (DIF) block cyclic redundancy code (CRC) for track formats.
  For open systems, this enables host-generated DIF CRCs to be stored with user data by the arrays and used for end-to-end data integrity validation.
- Additional protections for address/control fault modes for increased levels of protection against faults. These protections are defined in user-definable blocks supported by the T10 standard.
- Address and write status information in the extra bytes in the application tag and reference tag portion of the block CRC.

Data integrity checks

HYPERMAX OS validates the integrity of data they hold at every possible point during the lifetime of the data. From the point at which data enters an array, the data is continuously protected by error detection metadata. This protection metadata is checked by hardware and software mechanisms any time data is moved within the array subsystem, allowing the array to provide true end-to-end integrity checking and protection against hardware or software faults.

The protection metadata is appended to the data stream, and contains information describing the expected data location as well as CRC representation of the actual data contents. The expected values to be found in protection metadata are stored persistently in an area separate from the data stream. The protection metadata is used to validate the logical correctness of data being moved within the array any time the data transitions between protocol chips, internal buffers, internal data fabric endpoints, system cache, and system drives.

Drive monitoring and correction

HYPERMAX OS monitors medium defects by both examining the result of each disk data transfer and proactively scanning the entire disk during idle time. If a block on the disk is determined to be bad, the director:

1. Rebuilds the data in the physical storage, if necessary.
2. Rewrites the data in physical storage, if necessary.

The director also keeps track of each bad block detected on a drive. If the number of bad blocks exceeds a predefined threshold, the array proactively invokes a sparing operation to replace the defective drive, and then automatically alerts EMC Customer Support to arrange for corrective action, if necessary. With the deferred service model, often times immediate action is not required.

Physical memory error correction and error verification

HYPERMAX OS corrects single-bit errors and report an error code once the single-bit errors reach a predefined threshold. In the unlikely event that physical memory replacement is required, the array notifies EMC support, and a replacement is ordered.

Drive sparing and direct member sparing

When HYPERMAX OS 5977 detects a drive is about to fail or has failed, a direct member sparing (DMS) process is initiated. Direct member sparing looks for available spares within the same engine that are of the same block size, capacity and speed, with the best available spare always used.

With direct member sparing, the invoked spare is added as another member of the RAID group. During a drive rebuild, the option to directly copy the data from the failing drive to the invoked spare drive is supported. The failing drive is removed only when the copy process is finished. Direct member sparing is automatically initiated upon detection of drive-error conditions.

Direct member sparing provides the following benefits:

- The array can copy the data from the failing RAID member (if available), removing the need to read the data from all of the members and doing the rebuild. Copying to the new RAID member is less CPU intensive.
- If a failure occurs in another member, the array can still recover the data automatically from the failing member (if available).
- More than one spare for a RAID group is supported at the same time.

Vault to flash

VMAX All Flash arrays initiate a vault operation if the system is powered down, transitions offline, or if environmental conditions, such as the loss of a data center due to an air conditioning failure occur.

Each array comes with Standby Power Supply (SPS) modules. If you lose power, the array uses the SPS power to write the system mirrored cache onto flash storage. Vaulted images are fully redundant; the contents of the system mirrored cache are saved twice to independent flash storage.

The vault operation

When a vault operation is initiated:

- During the save part of the vault operation, the VMAX All Flash array stops all I/O. When the system mirrored cache reaches a consistent state, directors write the contents to the vault devices, saving two copies of the data. The array then completes the power down, or, if power down is not required, remains in the offline state.
- During the restore part of the operation, the array startup program initializes the hardware and the environmental system, and restores the system mirrored cache contents from the saved data (while checking data integrity).

The system resumes normal operation when the SPSes are sufficiently recharged to support another vault. If any condition is not safe, the system does not resume operation and notifies Customer Support for diagnosis and repair. This allows Customer Support to communicate with the array and restore normal system operations.

Vault configuration considerations

The following configuration considerations apply:

- To support vault to flash, the VMAX All Flash arrays require the following number of flash I/O modules:
  - VMAX 250F two to six per engine/V-Brick
  - VMAX 450F four to eight per engine/V-Brick/zBrick
  - VMAX 850F, 950F four to eight per engine/V-Brick/zBrick
- The size of the flash module is determined by the amount of system cache and metadata required for the configuration. For the number of supported Flash I/O modules, refer to Table 5 on page 25.
- The vault space is for internal use only and cannot be used for any other purpose when the system is online.
- The total capacity of all vault flash partitions will be sufficient to keep two logical copies of the persistent portion of the system mirrored cache.

Inline compression

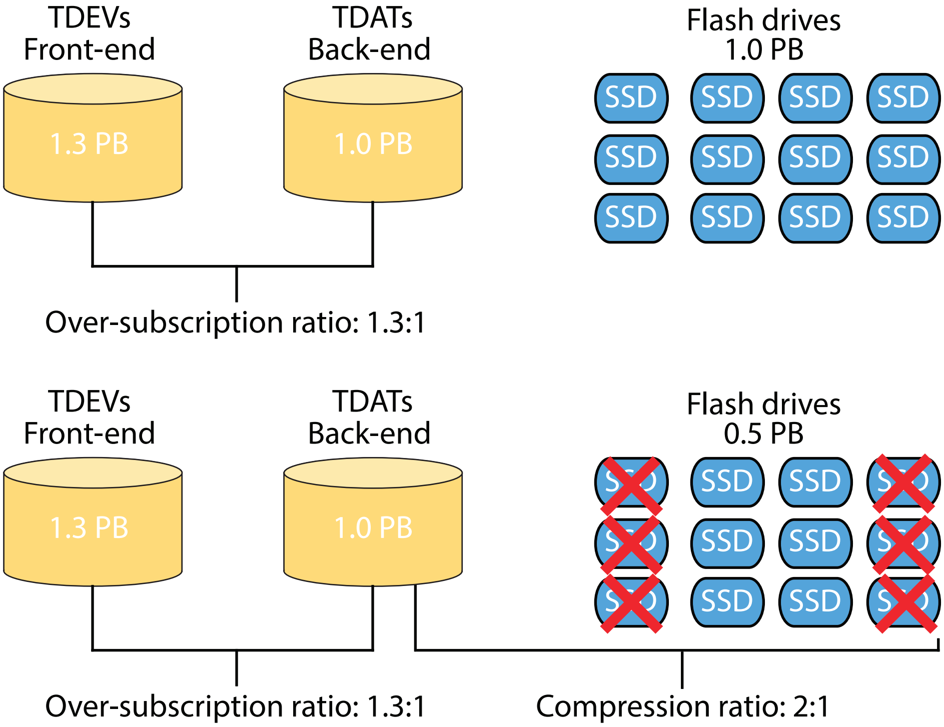
HYPERMAX OS 5977.945.890 introduces support for inline compression on VMAX All Flash arrays. Inline compression compresses data as it is written to flash drives.

Inline compression is a storage group attribute that you enable (default) or disable at the storage group level. When enabled, new I/O to the storage group is compressed when written to disk, while existing data on the storage group starts to compress in the background. After disabling, new I/O is no longer compressed, and existing data will remain compressed until it is written again, at which time it will decompress.

Inline compression and over-subscription complement each other. Over-subscription allows presenting larger than needed devices to hosts without having the physical drives to fully allocate the space represented by the thin devices. Inline compression further reduces the data footprint by increasing the effective capacity of the array. This is illustrated in the following example, where 1.3 PB of host attached devices

(TDEVs) is over-provisioned to 1.0 PB of back-end (TDATs), which reside on 1.0 PB of Flash drives. Following the data compression process, the data blocks are then compressed, by a ratio of 2:1, reducing the number of Flash drives by half. Basically, with compression enabled, the array requires half as many drives to support the same front-end capacity.

Figure 4 Inline compression and over-subscription



While this feature is pre-configured on new VMAX All Flash arrays at the factory, existing VMAX All Flash arrays in the field are eligible for upgrade. Contact your EMC Support Representative for more information.

Other compression-related notes:

- All supported data services, such as SnapVX, SRDF and encryption are supported with compression.
- Open systems (FBA) only (including eNAS). No CKD, including mixed FBA/CKD arrays. Any open system VMAX All Flash array with compression enabled, cannot have CKDs added to it.
- ProtectPoint operations are still supported to Data Domain arrays, and CloudArray can run on a compression-enabled array as long as it is in a separate SRP.
- Compression is enabled/disabled through Solutions Enabler and Unisphere for VMAX.
- Compression efficiency can be monitored on the SRP, storage group, and volume level.
- Activity Based Compression: the most active tracks are held in cache and not compressed until they cool enough to move from cache to disk. This feature helps improve the overall performance of the array while reducing wear on the flash drives.

CHAPTER 2

Management Interfaces

This chapter provides an overview of interfaces to manage arrays. Topics include:

Management interface versions

The following management software supports HYPERMAX OS 5977 Q2 2017 SR:

- Unisphere for VMAX V8.4
- Solutions Enabler V8.4
- Mainframe Enablers V8.2
- GDDR V5.0
- SMI-S V8.4
- SRA V6.3
- VASA Provider V8.4

Unisphere for VMAX

EMC Unisphere for VMAX is a web-based application that allows you to quickly and easily provision, manage, and monitor arrays.

Unisphere allows you to perform the following tasks:

Table 27 Unisphere tasks

Section	Allows you to:
Home	Perform viewing and management functions such as array usage, alert settings, authentication options, system preferences, user authorizations, and link and launch client registrations.
Storage	View and manage storage groups and storage tiers.
Hosts	View and manage initiators, masking views, initiator groups, array host aliases, and port groups.
Data Protection	View and manage local replication, monitor and manage replication pools, create and view device groups, and monitor and manage migration sessions.
Performance	Monitor and manage array dashboards, perform trend analysis for future capacity planning, and analyze data.
Databases	Troubleshoot database and storage issues, and launch Database Storage Analyzer.
System	View and display dashboards, active jobs, alerts, array attributes and licenses.
Support	View online help for Unisphere tasks.

Unisphere for VMAX is also available as Representational State Transfer (REST) API. This robust API allows you to access performance and configuration information, and to provision storage arrays. It can be used in any of the programming environments that support standard REST clients, such as web browsers and programming platforms that can issue HTTP requests.

Workload Planner

Workload Planner displays performance metrics for applications. Use Workload Planner to model the impact of migrating a workload from one storage system to another.

Use Workload Planner to:

- Model proposed new workloads.
- Assess the impact of moving one or more workloads off of a given array running HYPERMAX OS.
- Determine current and future resource shortfalls that require action to maintain the requested workloads.

FAST Array Advisor

The FAST Array Advisor wizard guides you through the steps to determine the impact on performance of migrating a workload from one array to another.

If the wizard determines that the target array can absorb the added workload, it automatically creates all the auto-provisioning groups required to duplicate the source workload on the target array.

Unisphere 360

Unisphere 360 is an on-premise management solution that provides a single window across arrays running HYPERMAX OS at a single site. It allows you to:

- Add a Unisphere server to Unisphere 360 to allow for data collection and reporting of Unisphere management storage system data.
- View the system health, capacity, alerts and capacity trends for your Data Center.
- View all storage systems from all enrolled Unisphere instances in one place.
- View details on performance and capacity.
- Link and launch to Unisphere instances running v8.2 or higher.
- Manage Unisphere 360 users and configure authentication and authorization rules.
- View details of visible storage arrays, including current and target storage

Solutions Enabler

Solutions Enabler provides a comprehensive command line interface (SYMCLI) to manage your storage environment.

SYMCLI commands are invoked from the host, either interactively on the command line, or using scripts.

SYMCLI is built on functions that use system calls to generate low-level I/O SCSI commands. Configuration and status information is maintained in a host database file, reducing the number of inquiries from the host to the arrays.

Use SYMCLI to:

- Configure array software (For example, TimeFinder, SRDF, Open Replicator)

- Monitor device configuration and status
- Perform control operations on devices and data objects

Solutions Enabler is also available as a Representational State Transfer (REST) API. This robust API allows you to access performance and configuration information, and to provision storage arrays. It can be used in any of the programming environments that support standard REST clients, such as web browsers and programming platforms that can issue HTTP requests.

Mainframe Enablers

The EMC Mainframe Enablers are a suite of software components that allow you to monitor and manage arrays running HYPERMAX OS. The following components are distributed and installed as a single package:

- ResourcePak Base for z/OS
  Enables communication between mainframe-based applications (provided by EMC or independent software vendors) and arrays.
- SRDF Host Component for z/OS
  Monitors and controls SRDF processes through commands executed from a host. SRDF maintains a real-time copy of data at the logical volume level in multiple arrays located in physically separate sites.
- EMC Consistency Groups for z/OS
  Ensures the consistency of data remotely copied by SRDF feature in the event of a rolling disaster.
- AutoSwap for z/OS
  Handles automatic workload swaps between arrays when an unplanned outage or problem is detected.
- TimeFinder SnapVX
  With Mainframe Enablers V8.0 and higher, SnapVX creates point-in-time copies directly in the Storage Resource Pool (SRP) of the source device, eliminating the concepts of target devices and source/target pairing. SnapVX point-in-time copies are accessible to the host via a link mechanism that presents the copy on another device. TimeFinder SnapVX and HYPERMAX OS support backward compatibility to traditional TimeFinder products, including TimeFinder/Clone, TimeFinder VP Snap, and TimeFinder/Mirror.
- Data Protector for z Systems (zDP™)
  With Mainframe Enablers V8.0 and higher, zDP is deployed on top of SnapVX. zDP provides a granular level of application recovery from unintended changes to data. zDP achieves this by providing automated, consistent point-in-time copies of data from which an application-level recovery can be conducted.
- TimeFinder/Clone Mainframe Snap Facility
  Produces point-in-time copies of full volumes or of individual datasets. TimeFinder/Clone operations involve full volumes or datasets where the amount of data at the source is the same as the amount of data at the target. TimeFinder VP Snap leverages clone technology to create space-efficient snaps for thin devices.
- TimeFinder/Mirror for z/OS
  Allows the creation of Business Continuance Volumes (BCVs) and provides the ability to ESTABLISH, SPLIT, RE-ESTABLISH and RESTORE from the source logical volumes.
- TimeFinder Utility

Conditions SPLIT BCVs by relabeling volumes and (optionally) renaming and recataloging datasets. This allows BCVs to be mounted and used.

Geographically Dispersed Disaster Restart (GDDR)

GDDR automates business recovery following both planned outages and disaster situations, including the total loss of a data center. Leveraging the VMAX architecture and the foundation of SRDF and TimeFinder replication families, GDDR eliminates any single point of failure for disaster restart plans in mainframe environments. GDDR intelligence automatically adjusts disaster restart plans based on triggered events.

GDDR does not provide replication and recovery services itself, but rather monitors and automates the services provided by other EMC products, as well as third-party products, required for continuous operations or business restart. GDDR facilitates business continuity by generating scripts that can be run on demand; for example, restart business applications following a major data center incident, or resume replication to provide ongoing data protection following unplanned link outages.

Scripts are customized when invoked by an expert system that tailors the steps based on the configuration and the event that GDDR is managing. Through automatic event detection and end-to-end automation of managed technologies, GDDR removes human error from the recovery process and allows it to complete in the shortest time possible.

The GDDR expert system is also invoked to automatically generate planned procedures, such as moving compute operations from one data center to another. This is the gold standard for high availability compute operations, to be able to move from scheduled DR test weekend activities to regularly scheduled data center swaps without disrupting application workloads.

SMI-S Provider

EMC SMI-S Provider supports the SNIA Storage Management Initiative (SMI), an ANSI standard for storage management. This initiative has developed a standard management interface that resulted in a comprehensive specification (SMI-Specification or SMI-S).

SMI-S defines the open storage management interface, to enable the interoperability of storage management technologies from multiple vendors. These technologies are used to monitor and control storage resources in multivendor or SAN topologies.

Solutions Enabler components required for SMI-S Provider operations are included as part of the SMI-S Provider installation.

VASA Provider

The VASA Provider enables VMAX management software to inform vCenter of how VMFS storage, including VVols, is configured and protected. These capabilities are defined by EMC and include characteristics such as disk type, thin or thick provisioning, storage tiering and remote replication status. This allows vSphere administrators to make quick, intelligent, and informed decisions as to virtual machine placement. VASA offers the ability for vSphere administrators to complement their use of plugins and other tools to track how VMAX devices hosting VMFS volume are configured to meet performance and availability needs.

eNAS management interface

eNAS block and file storage is managed using the Unisphere for VMAX File Dashboard. Link and launch enables you to run the block and file management GUI within the same session.

The configuration wizard helps you create storage groups (automatically provisioned to the Data Movers) quickly and easily. Creating a storage group creates a storage pool in Unisphere for VNX that can be used for file level provisioning tasks.

ViPR suite

The EMC ViPR® Suite delivers storage automation and management insights across multi-vendor storage. It helps to improve efficiency and optimize storage resources, while meeting service levels. The ViPR Suite provides self-service access to speed service delivery, reducing dependencies on IT, and providing an easy to use cloud experience.

ViPR Controller

ViPR Controller provides a single control plane for heterogeneous storage systems. ViPR makes a multi-vendor storage environment look like one virtual array.

ViPR uses software adapters that connect to the underlying arrays. ViPR exposes the APIs so any vendor, partner, or customer can build new adapters to add new arrays. This creates an extensible “plug and play” storage environment that can automatically connect to, discover and map arrays, hosts, and SAN fabrics.

ViPR enables the software-defined data center by helping users:

- Automate storage for multi-vendor block and file storage environments (control plane, or ViPR Controller)
- Manage and analyze data objects (ViPR Object and HDFS Services) to create a unified pool of data across file shares and commodity servers
- Create scalable, dynamic, commodity-based block storage (ViPR Block Service)
- Manage multiple data centers in different locations with single sign-on data access from any data center
- Protect against data center failures using active-active functionality to replicate data between geographically dispersed data centers
- Integrate with VMware and Microsoft compute stacks
- Migrate non-ViPR volumes into the ViPR environment (ViPR Migration Services Host Migration Utility)

For ViPR Controller requirements, refer to the *EMC ViPR Controller Support Matrix* on the EMC Online Support website.

ViPR Storage Resource Management

EMC ViPR SRM provides comprehensive monitoring, reporting, and analysis for heterogeneous block, file, and virtualized storage environments.

Use ViPR SRM to:

- Visualize applications to storage dependencies

- Monitor and analyze configurations and capacity growth
- Optimize your environment to improve return on investment

Virtualization enables businesses of all sizes to simplify management, control costs, and guarantee uptime. However, virtualized environments also add layers of complexity to the IT infrastructure that reduce visibility and can complicate the management of storage resources. ViPR SRM addresses these layers by providing visibility into the physical and virtual relationships to ensure consistent service levels.

As you build out your cloud infrastructure, ViPR SRM helps you ensure storage service levels while optimizing IT resources — both key attributes of successful cloud deployments.

ViPR SRM is designed for use in heterogeneous environments containing multi-vendor networks, hosts, and storage devices. The information it collects and the functionality it manages can reside on technologically disparate devices in geographically diverse locations. ViPR SRM moves a step beyond storage management and provides a platform for cross-domain correlation of device information and resource topology, and enables a broader view of your storage environment and enterprise data center.

ViPR SRM provides a dashboard view of the storage capacity at an enterprise level through Watch4net. The Watch4net dashboard view displays information to support decisions regarding storage capacity.

The Watch4net dashboard consolidates data from multiple ProSphere instances spread across multiple locations. It gives you a quick overview of the overall capacity status in your environment, raw capacity usage, usable capacity, used capacity by purpose, usable capacity by pools, and service levels.

The *EMC ViPR SRM Product Documentation Index* provides links to related ViPR documentation.

vStorage APIs for Array Integration

VMware vStorage APIs for Array Integration (VAAI) optimize server performance by offloading virtual machine operations to arrays running HYPERMAX OS.

The storage array performs the select storage tasks, freeing host resources for application processing and other tasks.

In VMware environments, storage arrays supports the following VAAI components:

- Full Copy — (Hardware Accelerated Copy) Faster virtual machine deployments, clones, snapshots, and VMware Storage vMotion® operations by offloading replication to the storage array.
- Block Zero — (Hardware Accelerated Zeroing) Initializes file system block and virtual drive space more rapidly.
- Hardware-Assisted Locking — (Atomic Test and Set) Enables more efficient meta data updates and assists virtual desktop deployments.
- UNMAP — Enables more efficient space usage for virtual machines by reclaiming space on datastores that is unused and returns it to the thin provisioning pool from which it was originally drawn.
- VMware vSphere Storage APIs for Storage Awareness (VASA).

VAAI is native in HYPERMAX OS and does not require additional software, unless eNAS is also implemented. If eNAS is implemented on the array, support for VAAI requires the VAAI plug-in for NAS. The plug-in is downloadable from EMC support.

SRDF Adapter for VMware® vCenter™ Site Recovery Manager

EMC SRDF Adapter is a Storage Replication Adapter (SRA) that extends the disaster restart management functionality of VMware vCenter Site Recovery Manager 5.x to arrays running HYPERMAX OS.

SRA allows Site Recovery Manager to automate storage-based disaster restart operations on storage arrays in an SRDF configuration.

SRDF/Cluster Enabler

Cluster Enabler (CE) for Microsoft Failover Clusters is a software extension of failover clusters functionality. Cluster Enabler allows Windows Server 2008 (including R2), and Windows Server 2012 (including R2) Standard and Datacenter editions running Microsoft Failover Clusters to operate across multiple connected storage arrays in geographically distributed clusters.

SRDF/Cluster Enabler (SRDF/CE) is a software plug-in module to EMC Cluster Enabler for Microsoft Failover Clusters software. The Cluster Enabler plug-in architecture consists of a CE base module component and separately available plug-in modules, which provide your chosen storage replication technology.

SRDF/CE supports:

- Synchronous mode on page 115
- Asynchronous mode on page 115
- Concurrent SRDF solutions on page 99
- Cascaded SRDF solutions on page 100

EMC Product Suite for z/TPF

The EMC Product Suite for z/TPF is a suite of components that monitor and manage arrays running HYPERMAX OS from a z/TPF host. z/TPF is an IBM mainframe operating system characterized by high-volume transaction rates with significant communications content. The following software components are distributed separately and can be installed individually or in any combination:

- SRDF Controls for z/TPF
  Monitors and controls SRDF processes with functional entries entered at the z/TPF Prime CRAS (computer room agent set).
- TimeFinder Controls for z/TPF
  Provides a business continuance solution consisting of TimeFinder SnapVX, TimeFinder/Clone, and TimeFinder/Mirror.
- ResourcePak for z/TPF
  Provides VMAX configuration and statistical reporting and extended features for SRDF Controls for z/TPF and TimeFinder Controls for z/TPF.

SRDF/TimeFinder Manager for IBM i

EMC SRDF/TimeFinder Manager for IBM i is a set of host-based utilities that provides an IBM i interface to EMC's SRDF and TimeFinder.

This feature allows you to configure and control SRDF or TimeFinder operations on arrays attached to IBM i hosts, including:

- SRDF:
  - Configure, establish and split SRDF devices, including:
    - SRDF/A
    - SRDF/S
    - Concurrent SRDF/A
    - Concurrent SRDF/S
- TimeFinder:
  - Create point-in-time copies of full volumes or individual data sets.
  - Create point-in-time snaphots of images.

Extended features

EMC SRDF/TimeFinder Manager for IBM i extended features provides support for the IBM independent ASP (IASP) functionality.

IASPs are sets of switchable or private auxiliary disk pools (up to 223) that can be brought online/offline on an IBM i host without affecting the rest of the system.

When combined with SRDF/TimeFinder Manager for IBM i, IASPs let you control SRDF or TimeFinder operations on arrays attached to IBM i hosts, including:

- Display and assign TimeFinder SnapVX devices.
- Execute SRDF or TimeFinder commands to establish and split SRDF or TimeFinder devices.
- Present one or more target devices containing an IASP image to another host for business continuance (BC) processes.

Access to extended features control operations include:

- From the SRDF/TimeFinder Manager menu-driven interface.
- From the command line using SRDF/TimeFinder Manager commands and associated IBM i commands.

AppSync

EMC AppSync offers a simple, SLA-driven, self-service approach for protecting, restoring, and cloning critical Microsoft and Oracle applications and VMware environments. After defining service plans, application owners can protect, restore, and clone production data quickly with item-level granularity by using the underlying EMC replication technologies. AppSync also provides an application protection monitoring service that generates alerts when the SLAs are not met.

AppSync supports the following applications and storage arrays:

- Applications — Oracle, Microsoft SQL Server, Microsoft Exchange, and VMware VMFS and NFS datastores and File systems.
- Replication Technologies—SRDF, SnapVX, RecoverPoint, XtremIO Snapshot, VNX Advanced Snapshots, VNXe Unified Snapshot, and ViPR Snapshot.

Note

For VMAX All Flash arrays, AppSync is available in a starter bundle. The AppSync Starter Bundle provides the license for a scale-limited, yet fully functional version of AppSync. For more information, refer to the *AppSync Starter Bundle with VMAX All Flash Product Brief* available on the EMC Online Support Website.

CHAPTER 3

Open Systems Support

This chapter introduces the open systems features supported on VMAX All Flash arrays.

Topics include:

HYPERMAX OS support for open systems

HYPERMAX OS supports FBA device emulations for open systems and D910 for IBM i.

Any logical device manager software installed on a host can be used with the storage devices.

HYPERMAX OS increases scalability limits from previous generations of arrays, including:

- Maximum device size is 64TB
- Maximum host addressable devices is 64,000/array
- Maximum storage groups, port groups, and masking views is 64,000/array
- Maximum devices addressable through each port is 4,000
  HYPERMAX OS does not support meta devices, thus it is much more difficult to reach this limit.

For more information on provisioning storage in an open systems environment, refer to Open Systems-specific provisioning on page 80.

For the most recent information, consult the EMC Support Matrix in the E-Lab Interoperability Navigator at http://elabnavigator.emc.com.

Backup and restore to external arrays

EMC ProtectPoint integrates primary storage on storage arrays running HYPERMAX OS and protection storage for backups on an EMC Data Domain system.

ProtectPoint provides block movement of the data on application source LUNs to encapsulated Data Domain LUNs for incremental backups.

Application administrators can use the ProtectPoint workflow to protect database applications and associated application data.

The ProtectPoint solution uses Data Domain and HYPERMAX OS features to provide protection:

On the Data Domain system:

- vdisk services
- FastCopy

On the storage array:

- FAST.X (tiered storage)
- SnapVX

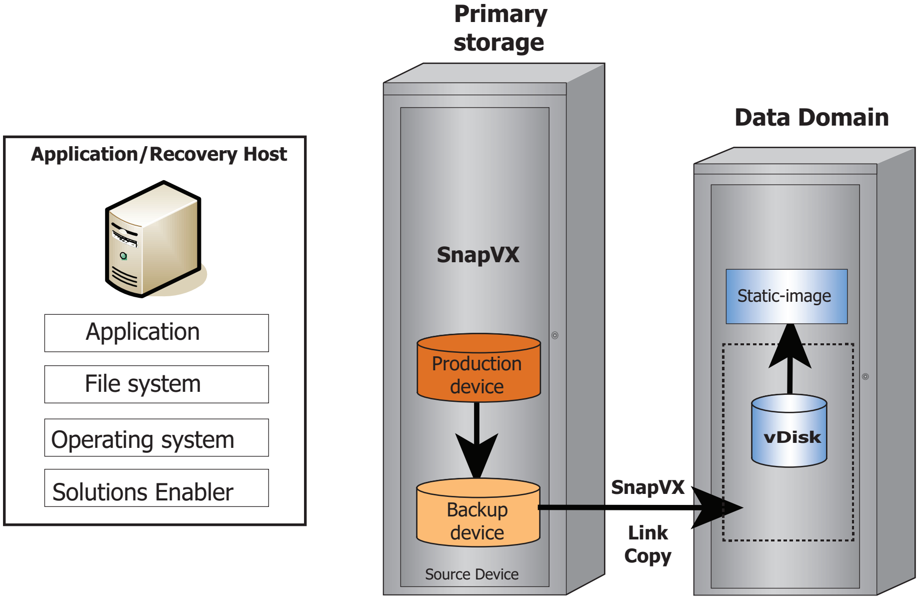
The combination of ProtectPoint and the storage array-to-Data Domain workflow enables the Application Administrator to:

- Back up and protect data
- Retain and replicate copies
- Restore data
- Recover applications

Data movement

The following image shows the data movement in a typical ProtectPoint solution. Data moves from the Application/Recovery (AR) Host to the primary array, and then to the Data Domain system.

Figure 5 ProtectPoint data movement



The Storage administrator configures the underlying storage resources on the primary storage array and the Data Domain system. With this storage configuration information, the Application administrator triggers the workflow to protect the application.

Note

Before triggering the workflow, the Application administrator must put the application in hot back-up mode. This ensures that an application-consistent snapshot is preserved on the Data Domain system.

Application administrators can select a specific backup when restoring data, and make that backup available on a selected set of primary storage devices.

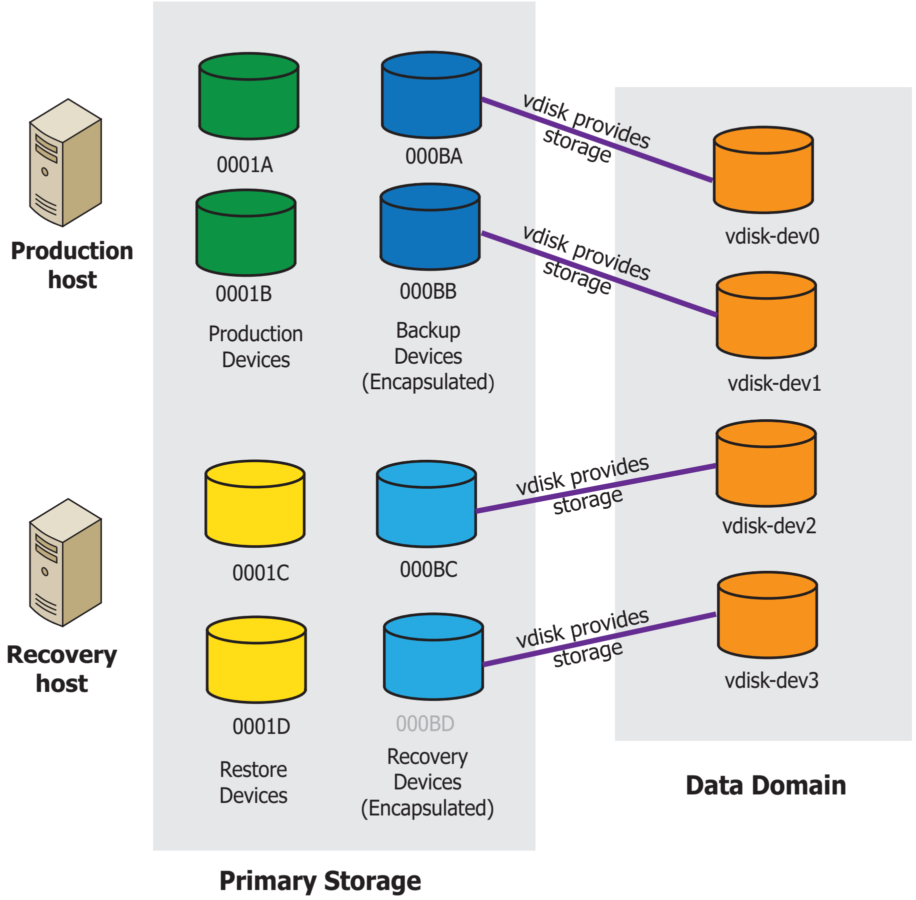
Operations to restore the data and make the recovery or restore devices available to the recovery host must be performed manually on the primary storage through EMC Solutions Enabler. The ProtectPoint workflow provides a copy of the data, but not any application intelligence.

Typical site topology

The ProtectPoint solution requires both IP network (LAN or WAN) and Fibre Channel (FC) Storage Area Network (SAN) connectivity.

The following image shows a typical primary site topology.

Figure 6 Typical RecoverPoint backup/recovery topology



ProtectPoint solution components

This section describes the connections, hosts, devices in a typical ProtectPoint solution.

The following table lists requirements for connecting components in the ProtectPoint solution.

Table 28 ProtectPoint connections

Connected Components	Connection Type
Primary Application Host to primary VMAX array	FC SAN
Primary Application Host to primary Data Domain system	IP LAN
Primary Recovery Host to primary VMAX array	FC SAN
Primary Recovery Host to primary Data Domain system	IP LAN
Primary VMAX array to primary Data Domain system	FC SAN
Secondary Recovery Host to secondary VMAX array (optional)	FC SAN
Secondary Recovery Host to secondary Data Domain system (optional)	IP LAN

Table 28 ProtectPoint connections (continued)

Connected Components	Connection Type
Secondary VMAX array to secondary Data Domain system (optional)	FC SAN
Primary Application Host to secondary Data Domain system (optional)	IP WAN
Primary Data Domain system to secondary Data Domain system (optional)	IP WAN

The following list describes the hosts and devices in a ProtectPoint solution:

Production Host

The host running the production database application. The production host sees only the production VMAX All Flash devices.

Recovery Host

The host available for database recovery operations. The recovery host can include direct access to:

- A backup on the recovery devices (vDisk devices encapsulated through FAST.X), or
- Access to a backup copy of the database on the restore devices (native VMAX All Flash devices).

Production Devices

Host devices available to the production host where the database instance resides. Production devices are the source devices for the TimeFinder/SnapVX operations that copy the production data to the backup devices for transfer to the Data Domain.

Restore Devices

Native VMAX All Flash devices used for full LUN-level copy of a backup to a new set of devices is desired. Restore devices are masked to the recovery host.

Backup Devices

Targets of the TimeFinder/SnapVX snapshots from the production devices. Backup devices are VMAX All Flash thin devices created when the Data Domain vDisk backup LUNs are encapsulated.

Recovery Devices

VMAX All Flash devices created when the Data Domain vDisk recovery LUNs are encapsulated. Recovery devices are presented to the recovery host when the Application administrator performs an object-level restore of specific database objects.

ProtectPoint and traditional backup

The ProtectPoint workflow can provide data protection in situations where more traditional approaches cannot successfully meet the business requirements. This is often due to small or non-existent backup windows, demanding recovery time objective (RTO) or recovery point objective (RPO) requirements, or a combination of both.

Unlike traditional backup and recovery, ProtectPoint does not rely on a separate process to discover the backup data and additional actions to move that data to backup storage. Instead of using dedicated hardware and network resources, ProtectPoint uses existing application and storage capabilities to create point-in-time copies of large data sets. The copies are transported across a storage area network (SAN) to Data Domain systems to protect the copies while providing deduplication to maximize storage efficiency.

ProtectPoint minimizes the time required to protect large data sets, and allows backups to fit into the smallest of backup windows to meet demanding RTO or RPO requirements.

Basic backup workflow

In the basic backup workflow, data is transferred from the primary storage array to the Data Domain system. ProtectPoint manages the data flow. The actual movement of the data is done by SnapVX.

The ProtectPoint solution enables the Application Administrator to take the snapshot on the primary storage array with minimal disruption to the application.

Note

The Application Administrator must ensure that the application is in an appropriate state before initiating the backup operation. This ensures that the copy or backup is application-consistent.

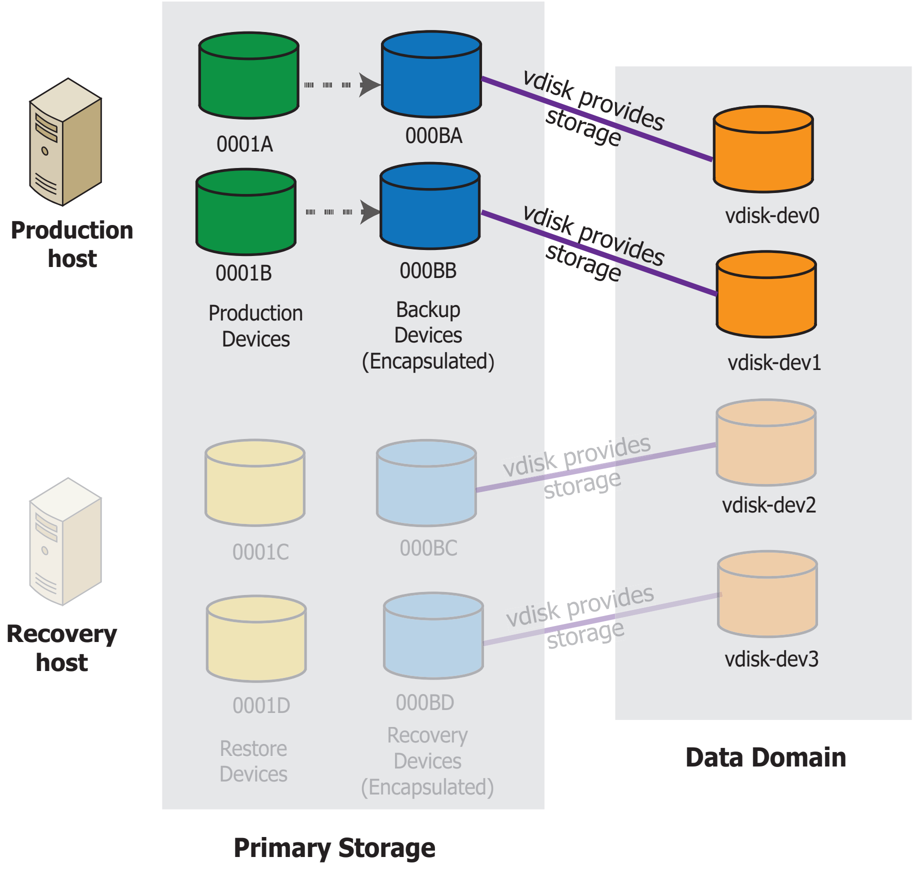
In a typical operation:

- The Application Administrator uses ProtectPoint to create a snapshot.
- ProtectPoint moves the data to the Data Domain system.
- The primary storage array keeps track of the data that has changed since the last update to the Data Domain system, and only copies the changed data.
- Once all the data captured in the snapshot has been sent to the Data Domain system, the Application Administrator can create a static-image of the data that reflects the application-consistent copy initially created on the primary storage array.

This static-image and its metadata are managed separately from the snapshot on the primary storage array, and can used as the source for additional copies of the backup. Static-images that are complete with metadata are called backup images. ProtectPoint creates one backup image for every protected LUN. Backup images can be combined into backup sets that represent an entire application point-in-time backup.

The following image illustrates the basic backup workflow.

Figure 7 Basic backup workflow



1. On the Application Host, the Application Administrator puts the database in hot backup mode.
2. On the primary storage array, ProtectPoint creates a snapshot of the storage device.
   The application can be taken out of hot backup mode when this step is complete.
3. The primary storage array analyzes the data and uses FAST.X to copy the changed data to an encapsulated Data Domain storage device.
4. The Data Domain creates and stores a backup image of the snapshot.

Basic restore workflow

There are two types of restoration:

Object-level restoration

One or more database objects are restored from a snapshot.

Full-application rollback restoration

The application is restored to a previous point-in-time. There are two types of recovery operations:

- A restore to the production database devices seen by the production host.
- A restore to the restore devices which can be made available to the recovery host.

For either type of restoration, the Application Administrator selects the backup image to restore from the Data Domain system.

Object-level restoration

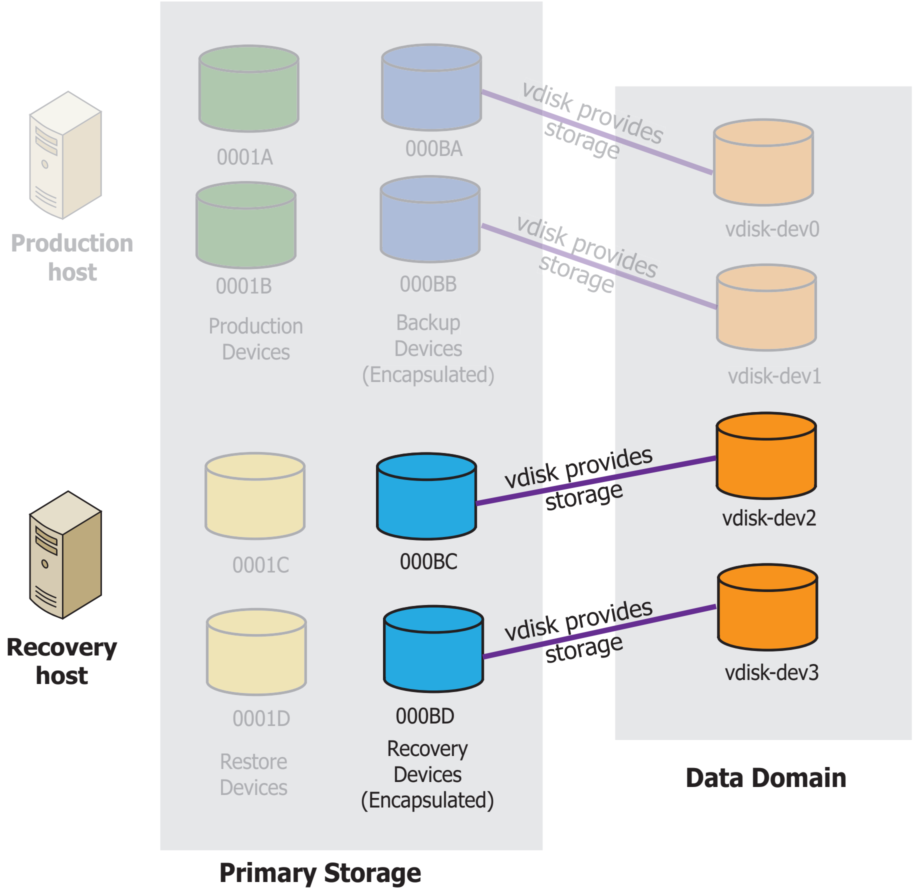
For object-level restoration, the Application Administrator:

- Selects the backup image on the Data Domain system
- Performs a restore of a database image to the recovery devices,

The Storage Administrator masks the recovery devices to the AR Host for an object-level restore.

The following image shows the object-level restoration workflow.

Figure 8 Object-level restoration workflow



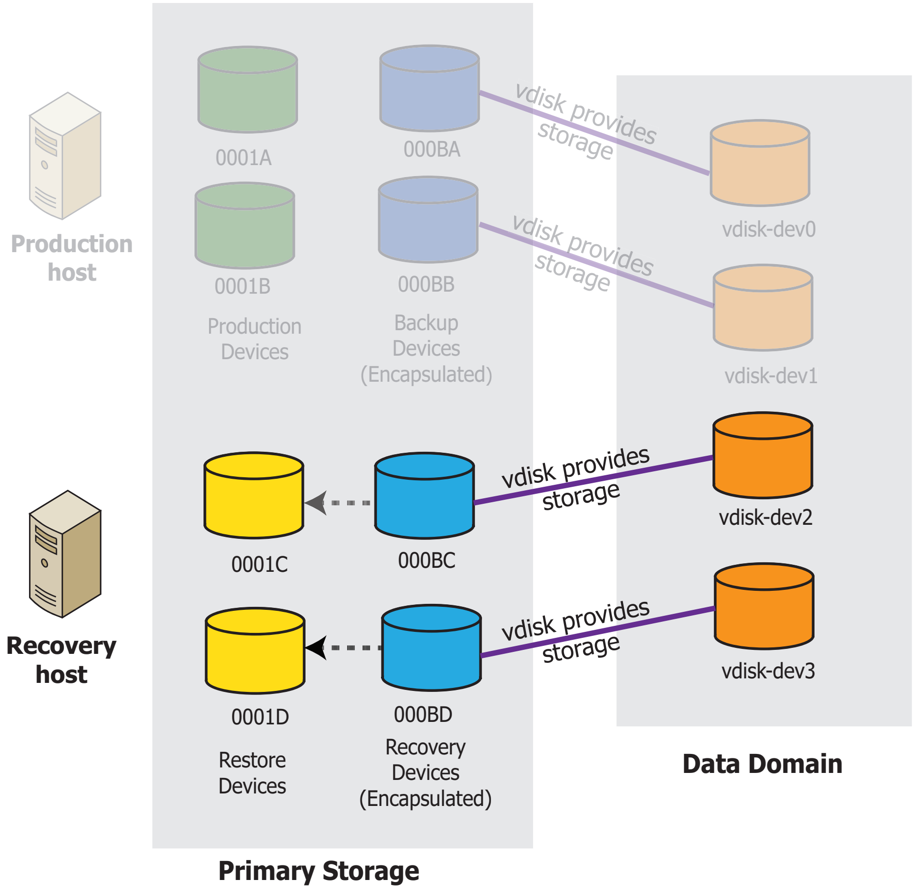
1. The Data Domain system writes the backup image to the encapsulated storage device, making it available on the primary storage array.
2. The Application Administrator mounts the encapsulated storage device to the recovery host, and uses OS- and application-specific tools and commands to restore specific objects.

Full-application rollback restoration

For a full-application rollback restoration, after selecting the backup image on the Data Domain system, the Storage Administrator performs a restore to the primary storage restore or production devices, depending on which devices need a restore of the full database image from the chosen point in time. Unlike object-level restoration, full-application rollback restoration requires manual SnapVX operations to complete the restore process. To make the backup image available on the primary storage array, the Storage Administrator must create a snapshot between the encapsulated Data Domain recovery devices and the restore/production devices, and then initiate the link copy operation.

The following image shows the full application rollback restoration workflow.

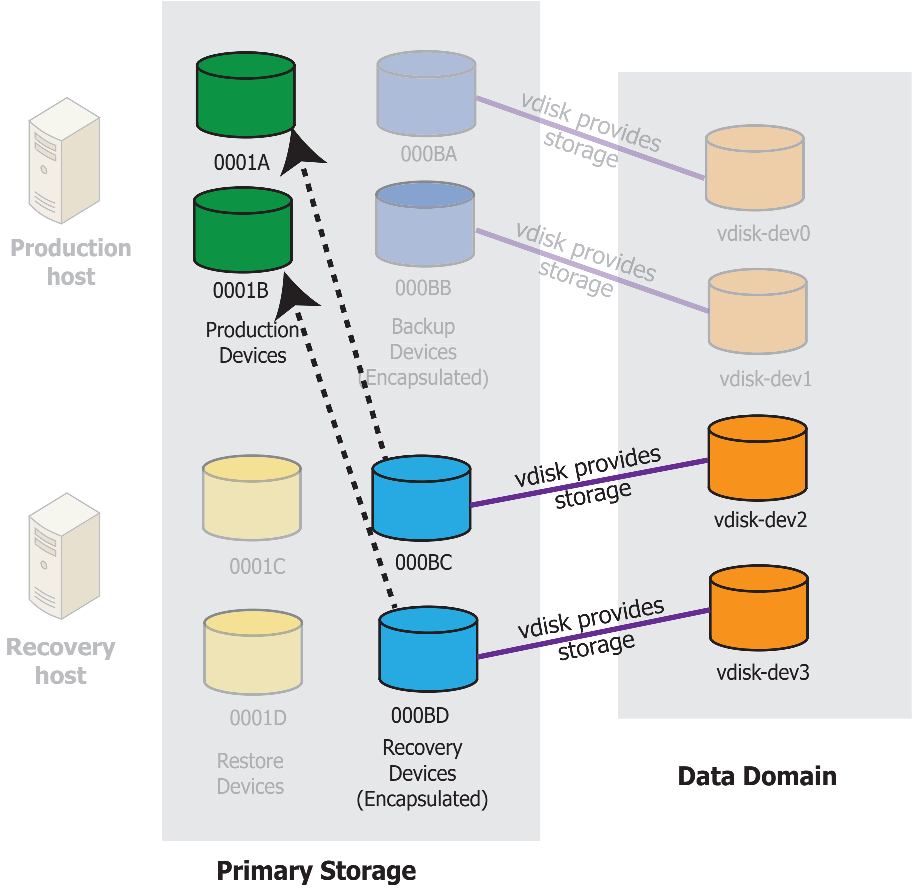
Figure 9 Full-application rollback restoration workflow



1. The Data Domain system writes the backup image to the encapsulated storage device, making it available on the primary storage array.
2. The Application Administrator creates a SnapVX snapshot of the encapsulated storage device and performs a link copy to the primary storage device, overwriting the existing data on the primary storage.
3. The restored data is presented to the Application Host.

The following image shows a full database recovery to product devices workflow. The workflow is the same as a full-application rollback restoration with the difference being the link copy targets.

Figure 10 Full database recovery to production devices



VMware Virtual Volumes

Storage arrays running HYPERMAX OS support VMware Virtual Volumes (VVols). VVols are a new storage object developed by VMware to simplify management and provisioning in virtualized environments. With VVols, the management process moves from the LUN (data store) level to the virtual machine (VM) level. This level of granularity allows VMware and cloud administrators to assign specific storage attributes to each VM, according to its performance and storage requirements.

VVol components

To support management capabilities of VVols, the storage/vCenter environment requires the following:

- EMC VMAX VASA Provider – The VASA Provider (VP) is a software plug-in that uses a set of out-of-band management APIs (VASA version 2.0). The VASA Provider exports storage array capabilities and presents them to vSphere through the VASA APIs. VVols are managed by way of vSphere through the VASA Provider APIs (create/delete) and not with the Unisphere for VMAX user interface or Solutions Enabler CLI. After VVols are setup on the array, Unisphere and Solutions Enabler only support VVol monitoring and reporting.
- Storage Containers (SC) – Storage containers are chunks of physical storage used to logically group VVols. SCs are based on the grouping of Virtual Machine Disks (VMDKs) into specific Service Levels. SC capacity is limited only by hardware capacity. At least one SC per storage system is required, but multiple SCs per array are allowed. SCs are created and managed on the array by the Storage Administrator. Unisphere and Solutions Enabler CLI support management of SCs.
- Protocol Endpoints (PE) – Protocol endpoints are the access points from the hosts to the array by the Storage Administrator. PEs are compliant with FC and replace the use of LUNs and mount points. VVols are "bound" to a PE, and the bind and unbind operations are managed through the VP APIs, not with the Solutions Enabler CLI. Existing multi-path policies and NFS topology requirements can be applied to the PE. PEs are created and managed on the array by the Storage Administrator. Unisphere and Solutions Enabler CLI support management of PEs.

Table 29 VVol architecture component management capability

Functionality	Component
VVol device management (create, delete)	VASA Provider APIs / Solutions Enabler APIs
VVol bind management (bind, unbind)	
Protocol Endpoint device management (create, delete)	Unisphere/Solutions Enabler CLI
Protocol Endpoint-VVol reporting (list, show)	
Storage Container management (create, delete, modify)	
Storage container reporting (list, show)	

VVol scalability

The following details the VVol scalability limits:

Table 30 VVol-specific scalability

Requirement	Value
Number of VVols/Array	64,000
Number of Snapshots/Virtual Machine[a]	12
Number of Storage Containers/Array	16
Number of Protocol Endpoints/Array	1/ESXi Host
Maximum number of Protocol Endpoints/ Array	1,024
Number of arrays supported /VP	1
Number of vCenters/VP	2
Maximum device size	16 TB

[a] VVol Snapshots can only be managed through vSphere. They cannot be created through Unisphere or Solutions Enabler.

VVol workflow

Before you begin

Install and configure following EMC applications:

- Unisphere for VMAX V8.2 or higher
- Solutions Enabler CLI V8.2 or higher
- VASA Provider V8.2 or higher

For instructions on installing Unisphere and Solutions Enabler, refer to their respective installation guides. For instructions on installing the VASA Provider, refer to the *EMC VMAX VASA Provider Release Notes*.

The steps required to create a VVol-based virtual machine are broken up by role:

Procedure

1. The VMAX Storage Administrator, uses either Unisphere for VMAX or Solutions Enabler to create and present the storage to the VMware environment:
   a. Create one or more storage containers on the storage array. This step defines how much storage and from which Service Level the VMware user can provision.
   b. Create Protocol Endpoints and provision them to the ESXi hosts.
2. The VMware Administrator, uses the vSphere Web Client to deploy the VM on the storage array:
   a. Add the VASA Provider to the vCenter. This allows vCenter to communicate with the storage array.
   b. Create VVol datastore from the storage container.

c. Create the VM Storage policies.
d. Create the VM in the VVol datastore, selecting one of the VM storage policies.

CHAPTER 4

Mainframe Features

This chapter describes mainframe-specific functionality provided with VMAX arrays.

HYPERMAX OS support for mainframe

VMAX 450F, 850F, and 950F arrays can be ordered with the zF and zFX software packages to support mainframe.

VMAX arrays provide the following mainframe support for CKD:

- Support for mixed FBA and CKD drive configurations. Flash Drive support on page 31 provides more information.
- Support for 64, 128, 256 FICON single and multi mode ports, respectively
- Support for CKD 3380/3390 and FBA devices
- Mainframe (FICON) and OS FC/iSCSI connectivity
- High capacity flash drives
- 16 Gb/s FICON host connectivity
- Support for Forward Error Correction, Query Host Access, and FICON Dynamic Routing
- T10 DIF protection for CKD data along the data path (in cache and on disk) to improve performance for multi-record operations
- D@RE external key managers:
  - Gemalto SafeNet KeySecure
  - IBM Security Key Lifecycle Manager

Data at Rest Encryption on page 38 provides more information.

IBM z Systems functionality support

VMAX arrays support the latest IBM z Systems enhancements, ensuring that the VMAX can handle the most demanding mainframe environments. VMAX arrays support:

- zHPF, including support for single track, multi track, List Prefetch, bi-directional transfers, QSAM/BSAM access, and Format Writes
- zHyperWrite
- Non-Disruptive State Save (NDSS)
- Compatible Native Flash (Flash Copy)
- Concurrent Copy
- Multi-subsystem Imaging
- Parallel Access Volumes
- Dynamic Channel Management (DCM)
- Dynamic Parallel Access Volumes/Multiple Allegiance (PAV/MA)
- Peer-to-Peer Remote Copy (PPRC) SoftFence
- Extended Address Volumes (EAV)
- Persistent IU Pacing (Extended Distance FICON)
- HyperPAV
- PDS Search Assist

- Modified Indirect Data Address Word (MIDAW)
- Multiple Allegiance (MA)
- Sequential Data Striping
- Multi-Path Lock Facility
- HyperSwap

Note

VMAX can participate in a z/OS Global Mirror (XRC) configuration only as a secondary.

IBM 2107 support

When VMAX arrays emulate an IBM 2107, they externally represent the array serial number as an alphanumeric number in order to be compatible with IBM command output. Internally, VMAX arrays retain a numeric serial number for IBM 2107 emulations. HYPERMAX OS handles correlation between the alphanumeric and numeric serial numbers.

Logical control unit capabilities

The following table lists logical control unit (LCU) maximum values:

Table 31 Logical control unit maximum values

Capability	Maximum value
LCUs per director slice (or port)	255 (within the range of 00 to FE)
LCUs per VMAX split[a]	255
Splits per VMAX array	16 (0 to 15)
Devices per VMAX split	65,280
LCUs per VMAX array	512
Devices per LCU	256
Logical paths per port	2,048
Logical paths per LCU per port (see Table 32 on page 76)	128
VMAX system host address per VMAX array (base and alias)	64K
I/O host connections per VMAX engine	32

[a] A VMAX split is a logical partition of the VMAX system, identified by unique devices, SSIDs, and host serial number. The maximum VMAX system host address per array is inclusive of all splits.

The following table lists the maximum LPARs per port based on the number of LCUs with active paths:

Table 32 Maximum LPARs per port

LCUs with active paths per port	Maximum volumes supported per port	VMAX maximum LPARs per port
16	4K	128
32	8K	64
64	16K	32
128	32K	16
255	64K	8

Disk drive emulations

When VMAX arrays are configured to mainframe hosts, the data recording format is Extended CKD (ECKD). The supported CKD emulations are 3380 and 3390.

Cascading configurations

Cascading configurations greatly enhance FICON connectivity between local and remote sites by using switch-to-switch extensions of the CPU to the FICON network. These cascaded switches communicate over long distances using a small number of high-speed lines called interswitch links (ISLs). A maximum of two switches may be connected together within a path between the CPU and the VMAX array.

Use of the same switch vendors is required for a cascaded configuration. To support cascading, each switch vendor requires specific models, hardware features, software features, configuration settings, and restrictions. Specific IBM CPU models, operating system release levels, host hardware, and HYPERMAX levels are also required.

For the most up-to-date information about switch support, consult the EMC Support Matrix (ESM), available through E-Lab™ Interoperability Navigator (ELN) at http://elabnavigator.emc.com.

CHAPTER 5

Provisioning

This chapter provides an overview of storage provisioning. Topics include:

Thin provisioning

VMAX All Flash arrays are pre-configured at the factory with thin provisioning pools ready for use. Thin provisioning improves capacity utilization and simplifies storage management. Thin provisioning enables storage to be allocated and accessed on demand from a pool of storage that services one or many applications. LUNs can be "grown" over time as space is added to the data pool with no impact to the host or application. Data is widely striped across physical storage (drives) to deliver better performance than standard provisioning.

Note

DATA devices (TDATs) are provisioned/pre-configured/created while the host addressable storage devices TDEVs are created by either the customer or customer support, depending on the environment.

Thin provisioning increases capacity utilization and simplifies storage management by:

- Enabling more storage to be presented to a host than is physically consumed
- Allocating storage only as needed from a shared thin provisioning pool
- Making data layout easier through automated wide striping
- Reducing the steps required to accommodate growth

Thin provisioning allows you to:

- Create host-addressable thin devices (TDEVs) using Unisphere for VMAX or Solutions Enabler
- Add the TDEVs to a storage group
- Run application workloads on the storage groups

When hosts write to TDEVs, the physical storage is automatically allocated from the default Storage Resource Pool.

Pre-configuration for thin provisioning

VMAX All Flash arrays are custom-built and pre-configured with array-based software applications, including a factory pre-configuration for thin provisioning that includes:

- *Data devices (TDAT)* — an internal device that provides physical storage used by thin devices.
- *Virtual provisioning pool* — a collection of data devices of identical emulation and protection type, all of which reside on drives of the same technology type and speed. The drives in a data pool are from the same disk group.
- *Disk group*— a collection of physical drives within the array that share the same drive technology and capacity. RAID protection options are configured at the disk group level. EMC strongly recommends that you use one or more of the RAID data protection schemes for all data devices.

Table 33 RAID options

RAID	Provides the following	Configuration considerations
RAID 5	Distributed parity and striped data across all drives in the RAID group. Options include: • RAID 5 (3 + 1) — Consists of four drives with parity and data striped across each device. • RAID-5 (7 + 1) — Consists of eight drives with data and parity striped across each device.	• RAID-5 (3 + 1) provides 75% data storage capacity. Only available with VMAX 250F arrays. • RAID-5 (7 + 1) provides 87.5% data storage capacity. • Withstands failure of a single drive within the RAID-5 group.
RAID 6	Striped drives with double distributed parity (horizontal and diagonal). The highest level of availability options include: • RAID-6 (6 + 2) — Consists of eight drives with dual parity and data striped across each device. • RAID-6 (14 + 2) — Consists of 16 drives with dual parity and data striped across each device.	• RAID-6 (6 + 2) provides 75% data storage capacity. Only available with VMAX 250F arrays. • RAID-6 (14 + 2) provides 87.5% data storage capacity. • Withstands failure of two drives within the RAID-6 group.

- *Storage Resource Pools* — one (default) Storage Resource Pool is pre-configured on the array. This process is automatic and requires no setup. You cannot modify Storage Resource Pools, but you can list and display their configuration. You can also generate reports detailing the demand storage groups are placing on the Storage Resource Pools.

Thin devices (TDEVs)

Note

VMAX All Flash arrays support only thin devices.

Thin devices (TDEVs) have no storage allocated until the first write is issued to the device. Instead, the array allocates only a minimum allotment of physical storage from the pool, and maps that storage to a region of the thin device including the area targeted by the write.

These initial minimum allocations are performed in small units called thin device extents. The device extent for a thin device is 1 track (128 KB).

When a read is performed on a device, the data being read is retrieved from the appropriate data device to which the thin device extent is allocated. Reading an area of a thin device that has not been mapped does not trigger allocation operations. Reading an unmapped block returns a block in which each byte is equal to zero.

When more storage is required to service existing or future thin devices, data devices can be added to existing thin storage groups.

Thin device oversubscription

A thin device can be presented for host use *before* mapping all of the reported capacity of the device.

The sum of the reported capacities of the thin devices using a given pool can exceed the available storage capacity of the pool. Thin devices whose capacity exceeds that of their associated pool are "oversubscribed".

Over-subscription allows presenting larger than needed devices to hosts and applications without having the physical drives to fully allocate the space represented by the thin devices.

Open Systems-specific provisioning

HYPERMAX host I/O limits for open systems

On open systems, you can define host I/O limits and associate a limit with a storage group. The I/O limit definitions contain the operating parameters of the input/output per second and/or bandwidth limitations.

When an I/O limit is associated with a storage group, the limit is equally divided among all the directors in the masking view associated with the storage group. All devices in that storage group share that limit.

When applications are configured, you can associate the limits with storage groups that contain a list of devices. A single storage group can only be associated with one limit and a device can only be in one storage group that has limits associated.

Up to 4096 host I/O limits can be defined.

Consider the following when using host I/O limits:

- Cascaded host I/O limits controlling parent and child storage groups limits in a cascaded storage group configuration.
- Offline and failed director redistribution of quota that supports all available quota to be available instead of losing quota allocations from offline and failed directors.
- Dynamic host I/O limits support for dynamic redistribution of steady state unused director quota.

Auto-provisioning groups on open systems

You can auto-provision groups on open systems to reduce complexity, execution time, labor cost, and the risk of error.

Auto-provisioning groups enables users to group initiators, front-end ports, and devices together, and to build masking views that associate the devices with the ports and initiators.

When a masking view is created, the necessary mapping and masking operations are performed automatically to provision storage.

After a masking view exists, any changes to its grouping of initiators, ports, or storage devices automatically propagate throughout the view, automatically updating the mapping and masking as required.

Auto-provisioning group components

The components of an auto-provisioning group are as follows:

Initiator group

A logical grouping of Fibre Channel initiators. An initiator group is limited to either a parent, which can contain other groups, or a child, which contains one initiator role. Mixing of initiators and child name in a group is not supported.

Port group

A logical grouping of Fibre Channel front-end director ports.
The maximum ports in a port group is 32.

Storage group

A logical grouping of thin devices. LUN addresses are assigned to the devices within the storage group when the view is created if the group is either cascaded or stand alone.

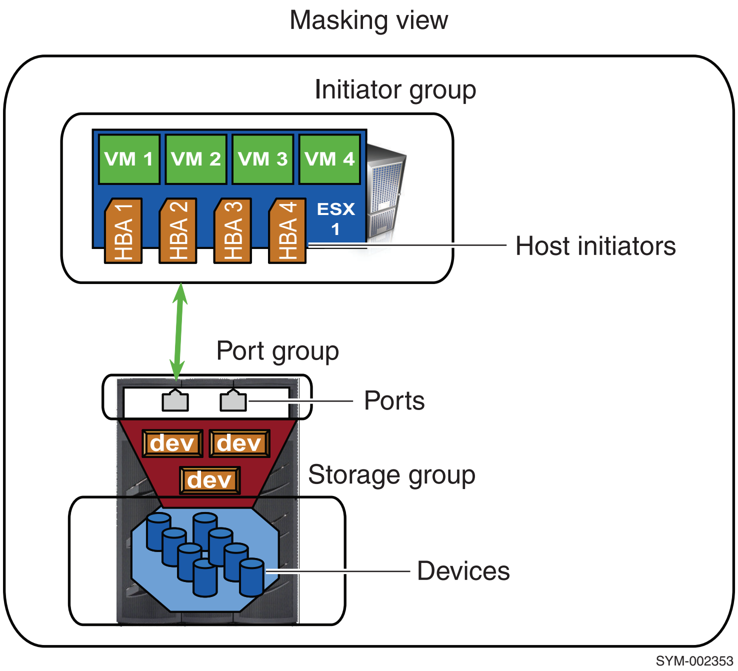
Cascaded storage group

A parent storage group comprised of multiple storage groups (parent storage group members) that contain child storage groups comprised of devices. By assigning child storage groups to the parent storage group members and applying the masking view to the parent storage group, the masking view inherits all devices in the corresponding child storage groups.

Masking view

An association between one initiator group, one port group, and one storage group. When a masking view is created, the group within the view is a parent, the contents of the children are used. For example, the initiators from the children initiator groups and the devices from the children storage groups. Depending on the server and application requirements, each server or group of servers may have one or more masking views that associate a set of thin devices to an application, server, or cluster of servers.

Figure 11 Auto-provisioning groups



CloudArray as an external tier

VMAX All Flash can be fully integrated with the market leading CloudArray storage solution for the purposes of migration. By enabling this technology, customers can seamlessly archive older application workloads out to the cloud, freeing up valuable Flash capacity for newer workloads. Once the older applications are archived out to the cloud, they will be directly available for retrieval at any time.

Manage the CloudArray configuration using the CloudArray management console (setup, cache encryption, monitoring) and the traditional management interfaces (Unisphere for VMAX, Solutions Enabler, API).

CHAPTER 6

Native local replication with TimeFinder

This chapter describes local replication features. Topics include:

About TimeFinder

EMC TimeFinder delivers point-in-time copies of volumes that can be used for backups, decision support, data warehouse refreshes, or any other process that requires parallel access to production data.

Previous VMAX families offered multiple TimeFinder products, each with their own characteristics and use cases. These traditional products required a target volume to retain snapshot or clone data.

Starting with HYPERMAX OS, TimeFinder introduced TimeFinder SnapVX which provides the best aspects of the traditional TimeFinder offerings, combined with increased scalability and ease-of-use.

TimeFinder SnapVX dramatically decreases the impact of snapshots and clones:

- For snapshots, this is done by using redirect on write technology (ROW).
- For clones, this is done by storing changed tracks (deltas) directly in the Storage Resource Pool of the source device - sharing tracks between snapshot versions and also with the source device, where possible.

There is no need to specify a target device and source/target pairs. SnapVX supports up to 256 snapshots per volume. Users can assign names to individual snapshots and assign an automatic expiration date to each one.

With SnapVX, a snaphot can be accessed by *linking* it to a host accessible volume (known as a target volume). Target volumes are standard VMAX All Flash TDEVs. Up to 1024 target volumes can be linked to the snapshots of the source volumes. The 1024 links can all be to the same snapshot of the source volume, or they can be multiple target volumes linked to multiple snapshots from the same source volume.

Note

A target volume may be linked only to one snapshot at a time.

Snapshots can be cascaded from linked targets, and targets can be linked to snapshots of linked targets. There is no limit to the number of levels of cascading, and the cascade can be broken.

SnapVX links to targets in the following modes:

- Nocopy Mode (Default): SnapVX does not copy data to the linked target volume but still makes the point-in-time image accessible through pointers to the snapshot. The point-in-time image will not be available after the target is unlinked because some target data may no longer be associated with the point-in-time image.
- Copy Mode: SnapVX copies all relevant tracks from the snapshot's point-in-time image to the linked target volume to create a complete copy of the point-in-time image that will remain available after the target is unlinked.

If an application needs to find a particular point-in-time copy among a large set of snapshots, SnapVX enables you to link and relink until the correct snapshot is located.

Interoperability with legacy TimeFinder products

TimeFinder SnapVX and HYPERMAX OS provide backward compatibility to legacy replication products by emulating legacy TimeFinder and IBM FlashCopy replication products. You can run your legacy replication scripts/jobs on VMAX All Flash arrays running TimeFinder SnapVX and HYPERMAX OS without altering them.

TimeFinder SnapVX emulates the following legacy replication products:

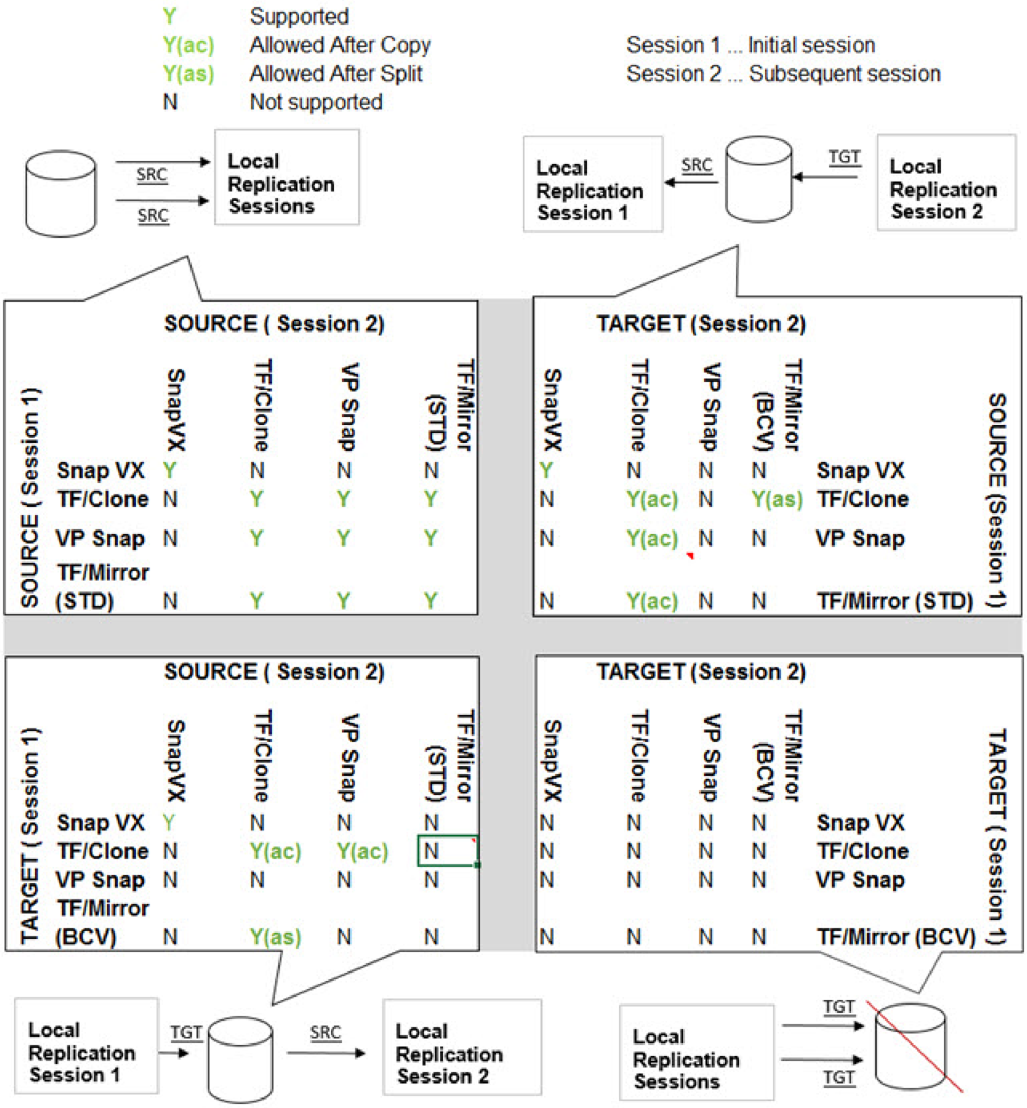
FBA devices	Mainframe (CKD devices)
TimeFinder/Clone	TimeFinder/Clone
TimeFinder/Mirror	TimeFinder/Mirror
TimeFinder VP Snap	TimeFinder Snap
	EMC Dataset Snap
	IBM FlashCopy (Full Volume and Extent Level)

Interoperability between TimeFinder SnapVX and legacy TimeFinder and IBM FlashCopy products depends on:

- The device role in the local replication session.
  A CKD or FBA device can be the source or the target in a local replication session. Different rules apply to ensure data integrity when concurrent local replication sessions run on the same device.
- The management software (Solutions Enabler/Unisphere for VMAX or Mainframe Enablers) used to control local replication.
  - Solutions Enabler and Unisphere for VMAX do not support interoperability between SnapVX and other local replication session on FBA or CKD devices.

    Figure 12 on page 86 provides detailed local replication interoperability support for FBA devices by using open systems management software (Solutions Enabler, Unisphere for VMAX).
  - Mainframe Enablers (MFE) support interoperability between SnapVX and other local replication sessions.
    Figure 13 on page 87 provides detailed interoperability information for CKD devices managed by using Mainframe Enablers.

Figure 12 Local replication interoperability, FBA devices



Y	Supported
Y(ac)	Allowed After Copy
Y(as)	Allowed After Split
N	Not supported

Session 1 ... Initial session
Session 2 ... Subsequent session

SOURCE (Session 1) \ SOURCE (Session 2)	SnapVX	TF/Clone	VP Snap	TF/Mirror (STD)
Snap VX	Y	N	N	N
TF/Clone	N	Y	Y	Y
VP Snap	N	Y	Y	Y
TF/Mirror (STD)	N	Y	Y	Y

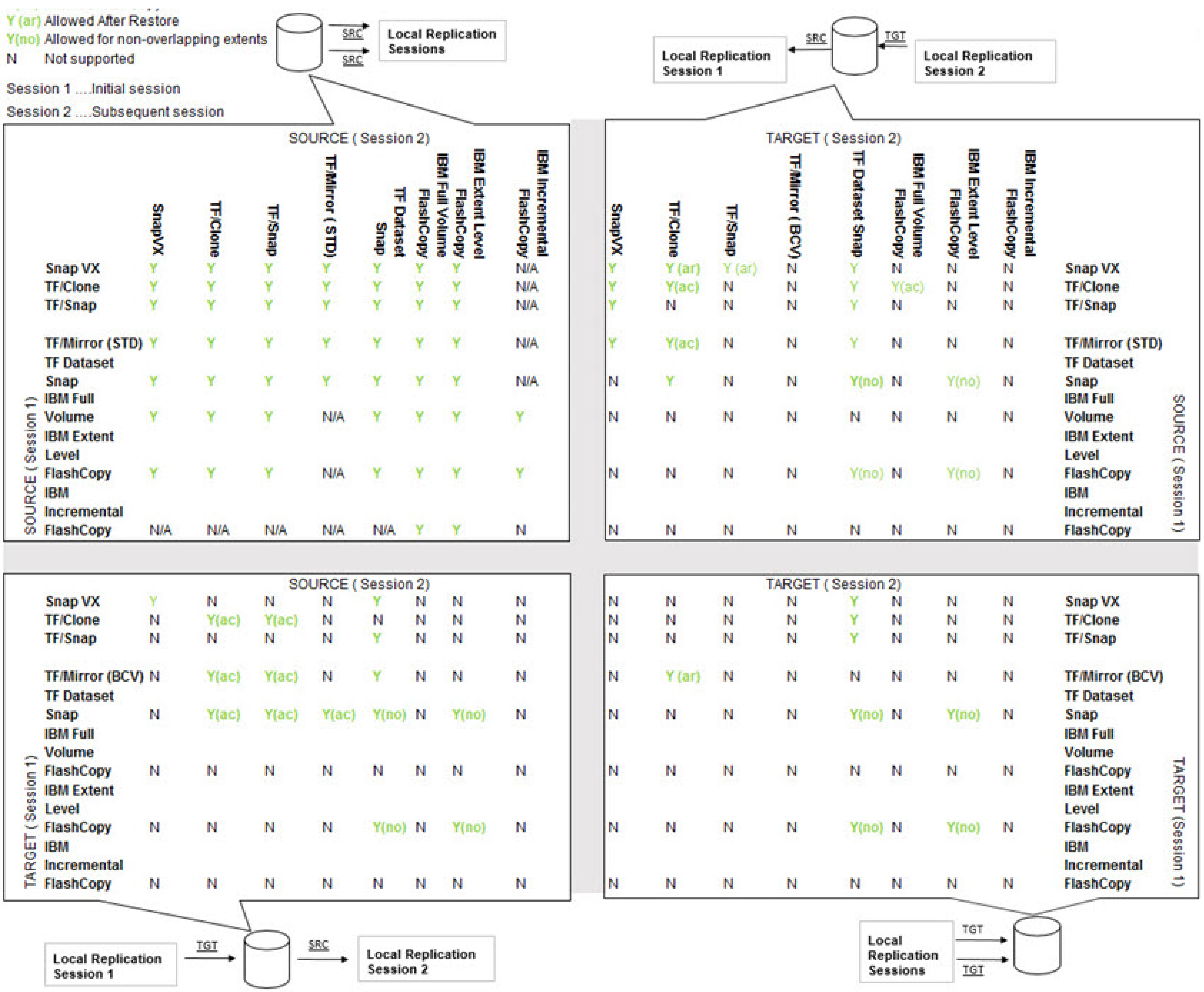
SOURCE (Session 1) \ TARGET (Session 2)	SnapVX	TF/Clone	VP Snap	TF/Mirror (BCV)
Snap VX	Y	N	N	N
TF/Clone	N	Y(ac)	N	Y(as)
VP Snap	N	Y(ac)	N	N
TF/Mirror (STD)	N	Y(ac)	N	N

TARGET (Session 1) \ SOURCE (Session 2)	SnapVX	TF/Clone	VP Snap	TF/Mirror (STD)
Snap VX	Y	N	N	N
TF/Clone	N	Y(ac)	Y(ac)	N
VP Snap	N	N	N	N
TF/Mirror (BCV)	N	Y(as)	N	N

TARGET (Session 1) \ TARGET (Session 2)	SnapVX	TF/Clone	VP Snap	TF/Mirror (BCV)
Snap VX	N	N	N	N
TF/Clone	N	N	N	N
VP Snap	N	N	N	N
TF/Mirror (BCV)	N	N	N	N

SRC
SRC
Local Replication Sessions
Local Replication Session 1
SRC
TGT
Local Replication Session 2
Local Replication Session 1
TGT
SRC
Local Replication Session 2
Local Replication Sessions
TGT
TGT

Figure 13 Local replication interoperability, CKD devices



SRC
SRC
Local Replication Sessions
Local Replication Session 1
SRC
TGT
Local Replication Session 2
Local Replication Session 1
TGT
SRC
Local Replication Session 2
Local Replication Sessions
TGT
TGT

Y (ar) Allowed After Restore
Y(no) Allowed for non-overlapping extents
N Not supported

Session 1 ....Initial session
Session 2 ....Subsequent session

SOURCE (Session 1) / SOURCE (Session 2)

	SnapVX	TF/Clone	TF/Snap	TF/Mirror (STD)	TF Dataset Snap	IBM Full Volume FlashCopy	IBM Extent Level FlashCopy	IBM Incremental FlashCopy
Snap VX	Y	Y	Y	Y	Y	Y	Y	N/A
TF/Clone	Y	Y	Y	Y	Y	Y	Y	N/A
TF/Snap	Y	Y	Y	Y	Y	Y	Y	N/A
TF/Mirror (STD)	Y	Y	Y	Y	Y	Y	Y	N/A
TF Dataset Snap	Y	Y	Y	Y	Y	Y	Y	N/A
IBM Full Volume	Y	Y	Y	N/A	Y	Y	Y	Y
IBM Extent Level FlashCopy	Y	Y	Y	N/A	Y	Y	Y	Y
IBM Incremental FlashCopy	N/A	N/A	N/A	N/A	N/A	Y	Y	N

SOURCE (Session 1) / TARGET (Session 2)

	SnapVX	TF/Clone	TF/Snap	TF/Mirror (BCV)	TF Dataset Snap	IBM Full Volume FlashCopy	IBM Extent Level FlashCopy	IBM Incremental FlashCopy
Snap VX	Y	Y (ar)	Y (ar)	N	Y	N	N	N
TF/Clone	Y	Y(ac)	N	N	Y	Y(ac)	N	N
TF/Snap	Y	N	N	N	Y	N	N	N
TF/Mirror (STD)	Y	Y(ac)	N	N	Y	N	N	N
TF Dataset Snap	N	Y	N	N	Y(no)	N	Y(no)	N
IBM Full Volume	N	N	N	N	N	N	N	N
IBM Extent Level FlashCopy	N	N	N	N	Y(no)	N	Y(no)	N
IBM Incremental FlashCopy	N	N	N	N	N	N	N	N

TARGET (Session 1) / SOURCE (Session 2)

	SnapVX	TF/Clone	TF/Snap	TF/Mirror (STD)	TF Dataset Snap	IBM Full Volume FlashCopy	IBM Extent Level FlashCopy	IBM Incremental FlashCopy
Snap VX	Y	N	N	N	Y	N	N	N
TF/Clone	N	Y(ac)	Y(ac)	N	N	N	N	N
TF/Snap	N	N	N	N	Y	N	N	N
TF/Mirror (BCV)	N	Y(ac)	Y(ac)	N	Y	N	N	N
TF Dataset Snap	N	Y(ac)	Y(ac)	Y(ac)	Y(no)	N	Y(no)	N
IBM Full Volume FlashCopy	N	N	N	N	N	N	N	N
IBM Extent Level FlashCopy	N	N	N	N	Y(no)	N	Y(no)	N
IBM Incremental FlashCopy	N	N	N	N	N	N	N	N

TARGET (Session 1) / TARGET (Session 2)

	SnapVX	TF/Clone	TF/Snap	TF/Mirror (BCV)	TF Dataset Snap	IBM Full Volume FlashCopy	IBM Extent Level FlashCopy	IBM Incremental FlashCopy
Snap VX	N	N	N	N	Y	N	N	N
TF/Clone	N	N	N	N	Y	N	N	N
TF/Snap	N	N	N	N	Y	N	N	N
TF/Mirror (BCV)	N	Y (ar)	N	N	N	N	N	N
TF Dataset Snap	N	N	N	N	Y(no)	N	Y(no)	N
IBM Full Volume FlashCopy	N	N	N	N	N	N	N	N
IBM Extent Level FlashCopy	N	N	N	N	Y(no)	N	Y(no)	N
IBM Incremental FlashCopy	N	N	N	N	N	N	N	N

Targetless snapshots

TimeFinder SnapVX management interfaces enable you to take a snapshot of an entire VMAX All Flash Storage Group with a single command. With this in mind, VMAX All Flash supports up to 64K storage groups, which is enough even in the most demanding environment for one per application. The storage group construct already exists in the majority of cases as they are created for masking views. Timefinder SnapVX is able to utilize this already existing structure reducing the administration required to maintain the application and its replication environment.

Creation of SnapVX snapshots does not require you to preconfigure any additional volumes, which reduces the cache footprint of SnapVX snapshots and simplifies implementation. Snapshot creation and automatic termination can easily be scripted.

In the following example, a snapshot is created with a 2 day retention. This command can be scheduled to run in as part of a script to create multiple versions of the snapshot, each one sharing tracks where possible with each other and the source devices. Use a cron job or scheduler to run the snapshot script on a schedule to create up to 256 snapshots of the source volumes; enough for a snapshot every 15 minutes with 2 days of retention:

```
symsnapvx -sid 001 -sg StorageGroup1 -name sg1_snap establish -ttl -
delta 2
```

If a restore operation is required, any of the snapshots created by the example above can be specified.

When the storage group transitions to a restored state, the restore session can be terminated. The snapshot data is preserved during the restore process and can be used again should the snapshot data be required for a future restore.

Secure snaps

Introduced with HYPERMAX OS 5977 Q2 2017 SR, secure snaps is an enhancement to the current snapshot technology. Secure snaps prevent administrators or other high-level users from intentionally or unintentionally deleting snapshot data. In addition, Secure snaps are also immune to automatic failure resulting from running out of Storage Recourse Pool (SRP) or Replication Data Pointer (RDP) space on the array.

When creating a secure snapshot, you assign it an expiration date/time either as a delta from the current date or as an absolute date. Once the expiration date passes, and if the snapshot has no links, HYPERMAX OS automatically deletes the snapshot. Prior to its expiration, Administrators can only extend the expiration date - they cannot shorten the date or delete the snapshot. If a secure snapshot expires, and it has a volume linked to it, or an active restore session, the snapshot is not deleted; however, it is no longer considered secure.

Note

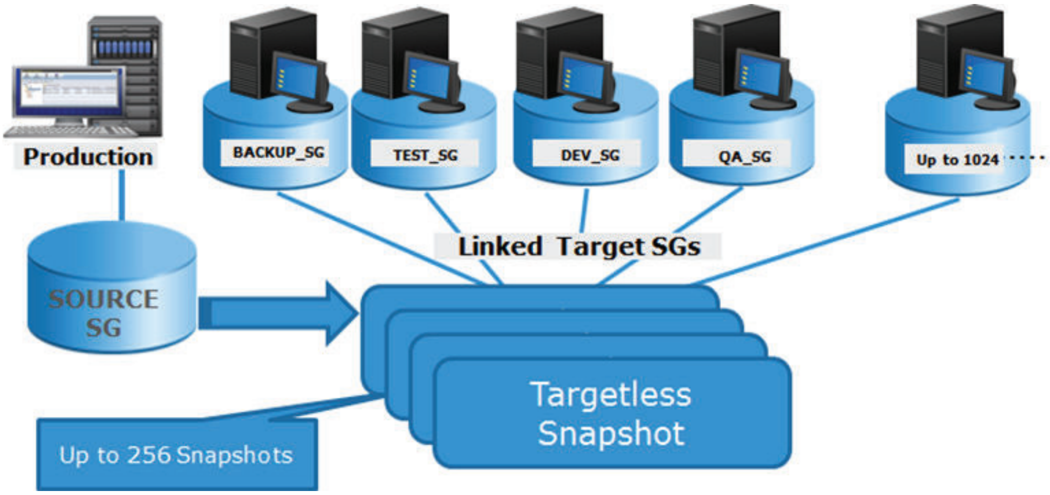
Secure snapshots may only be terminated after they expire or by customer-authorized EMC support. Refer to Knowledgebase article 498316 for additional information.

Provision multiple environments from a linked target

Use SnapVX to provision multiple test, development environments using linked snapshots. To access a point-in-time copy, create a link from the snapshot data to a host mapped target device.

Each linked storage group can access the same snapshot, or each can access a different snapshot version in either no copy or copy mode. Changes to the linked volumes do not affect the snapshot data. To roll back a test development environment to the original snapshot image, perform a relink operation.

Figure 14 SnapVX targetless snapshots



Note

Target volumes must be unmounted before issuing the relink command to ensure that the host operating system does not cache any filesystem data. If accessing through VPLEX, ensure that you follow the procedure outlined in the technical note *EMC VPLEX: LEVERAGING ARRAY BASED AND NATIVE COPY TECHNOLOGIES*, available on support.emc.com

---

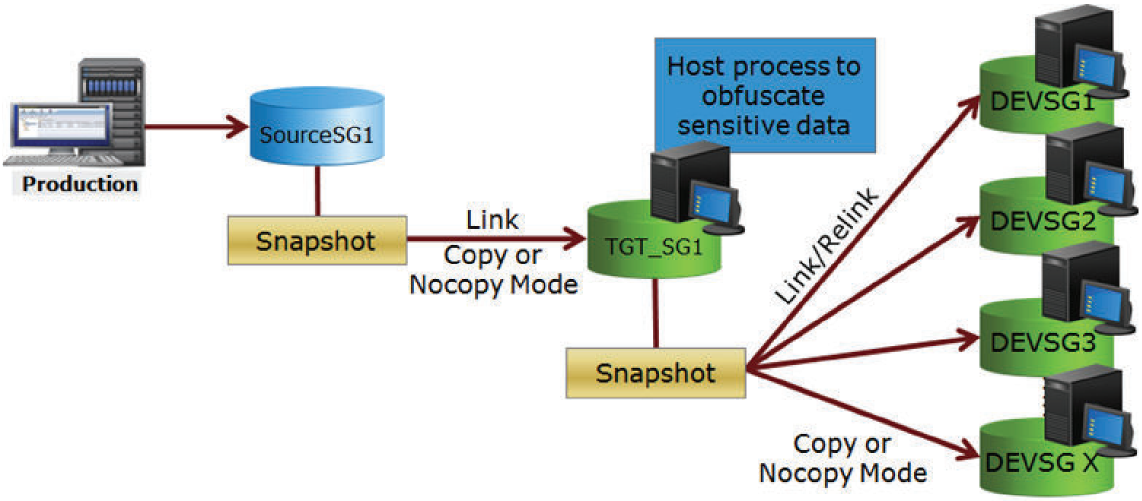
Once the relink is complete, volumes can be remounted.

Snapshot data is unchanged by the linked targets, so the snapshots can also be used to restore production data.

Cascading snapshots

Presenting sensitive data to test or development environments often requires that sensitive data be obfuscated before it is presented to any test or development hosts. Use cascaded snapshots to support obfuscation, as shown in the following image.

Figure 15 SnapVX cascaded snapshots



If no change to the data is required before presenting it to the test or development environments, there is no need to create a cascaded relationship.

Accessing point-in-time copies

To access a point-in time-copy, you must create a link from the snapshot data to a host mapped target device. The links may be created in Copy mode for a permanent copy on the target device, or in NoCopy mode for temporary use. Copy mode links create full-volume, full-copy clones of the data by copying it to the target device's Storage Resource Pool. NoCopy mode links are space-saving snapshots that only consume space for the changed data that is stored in the source device's Storage Resource Pool.

HYPERMAX OS supports up to 1,024 linked targets per source device.

Note

When a target is first linked, all of the tracks are undefined. This means that the target does not know where in the Storage Resource Pool the track is located, and host access to the target must be derived from the SnapVX metadata. A background process eventually defines the tracks and updates the thin device to point directly to the track location in the source device's Storage Resource Pool.

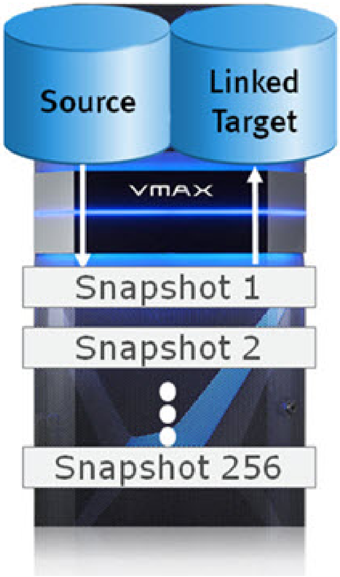
Mainframe SnapVX and zDP

Data Protector for z Systems (zDP) is a mainframe software solution that is deployed on top of SnapVX on VMAX All Flash arrays. zDP delivers the capability to recover from logical data corruption with minimal data loss. zDP achieves this by providing multiple, frequent, consistent point-in-time copies of data in an automated fashion from which an application level recovery can be conducted, or the environment restored to a point prior to the logical corruption.

By providing easy access to multiple different point-in-time copies of data (with a granularity of minutes), precise remediation of logical data corruption can be performed using application-based recovery procedure. zDP results in minimal data loss compared to the previous method of restoring data from daily or weekly backups.

As shown in Figure 16 on page 91, zDP enables you to create and manage multiple point-in-time snapshots of volumes. A snapshot is a pointer-based, point-in-time image of a single volume. These point-in-time copies are created using the SnapVX feature of HYPERMAX OS. SnapVX is a space-efficient method for making volume level snapshots of thin devices and consuming additional storage capacity only when updates are made to the source volume. There is no need to copy each snapshot to a target volume as SnapVX separates the capturing of a point-in-time copy from its usage. Capturing a point-in-time copy does not require a target volume. Using a point-in-time copy from a host requires linking the snapshot to a target volume. You can make multiple snapshots (up to 256) of each source volume.

Figure 16 zDP operation



These snapshots share allocations to the same track image whenever possible while ensuring they each continue to represent a unique point-in-time image of the source volume. Despite the space efficiency achieved through shared allocation to unchanged data, additional capacity is required to preserve the pre-update images of changed tracks captured by each point-in-time snapshot.

zDP implementation is a two-stage process — the planning phase and the implementation phase.

- The planning phase is done in conjunction with your EMC representative who has access to tools that can help size the capacity needed for zDP if you are currently a VMAX All Flash user.
- The implementation phase utilizes the following methods for z/OS:
  - A batch interface that allows you to submit jobs to define and manage zDP.
  - A zDP run-time environment that executes under SCF to create snapsets.

  For details on zDP usage, refer to the *TimeFinder SnapVX and zDP Product Guide*. For details on zDP usage in z/TPF, refer to the *TimeFinder Controls for z/TPF Product Guide*.

CHAPTER 7

Remote replication solutions

This chapter describes EMC's remote replication solutions. Topics include:

Native remote replication with SRDF

The EMC Symmetrix Remote Data Facility (SRDF) family of products offers a range of array based disaster recovery, parallel processing, and data migration solutions for VMAX Family systems, including:

- HYPERMAX OS for VMAX All Flash 250F, 450F, 850F, and 950F arrays
- HYPERMAX OS for VMAX 100K, 200K, and 400K arrays
- Enginuity for VMAX 10K, 20K, and 40K arrays

SRDF replicates data between 2, 3 or 4 arrays located in the same room, on the same campus, or thousands of kilometers apart. Replicated volumes may include a single device, all devices on a system, or thousands of volumes across multiple systems.

SRDF disaster recovery solutions use “active, remote” mirroring and dependent-write logic to create consistent copies of data. Dependent-write consistency ensures transactional consistency when the applications are restarted at the remote location. You can tailor your SRDF solution to meet various Recovery Point Objectives/ Recovery Time Objectives.

Using only SRDF, you can create complete solutions to:

- Create real-time (SRDF/S) or dependent-write-consistent (SRDF/A) copies at 1, 2, or 3 remote arrays.
- Move data quickly over extended distances.
- Provide 3-site disaster recovery with zero data loss recovery, business continuity protection and disaster-restart.

You can integrate SRDF with other EMC products to create complete solutions to:

- Restart operations after a disaster with zero data loss and business continuity protection.
- Restart operations in cluster environments. For example Microsoft Cluster Server with Microsoft Failover Clusters.
- Monitor and automate restart operations on an alternate local or remote server.
- Automate restart operations in VMware environments.

SRDF operates in the following modes:

- **Synchronous mode (SRDF/S)** maintains a real-time copy at arrays located within 200 kilometers. Writes from the production host are acknowledged from the local array when they are written to cache at the remote array.
- **Asynchronous mode (SRDF/A)** maintains a dependent-write consistent copy at arrays located at unlimited distances. Writes from the production host are acknowledged immediately by the local array, thus replication has no impact on host performance. Data at the remote array is typically only seconds behind the primary site.
- **SRDF/Metro** makes R2 devices Read/Write accessible to a host (or multiple hosts in clusters). Hosts write to both the R1 and R2 sides of SRDF device pairs, and SRDF/Metro ensures that each copy remains current and consistent. This feature is only for FBA volumes on arrays running HYPERMAX OS 5977.691.684 or higher. To manage this feature requires version 8.1 or higher of Solutions Enabler/ Unisphere for VMAX.
- **Adaptive copy mode** moves large amounts of data quickly with minimal host impact. Adaptive copy mode does not provide restartable data images at the

secondary site until no new writes are sent to the R1 device and all data has finished copying to the R2.

SRDF 2-site solutions

The following table describes SRDF 2-site solutions.

Table 34 SRDF 2-site solutions

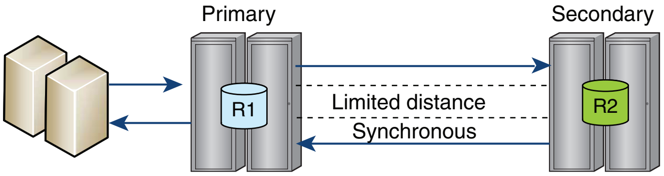
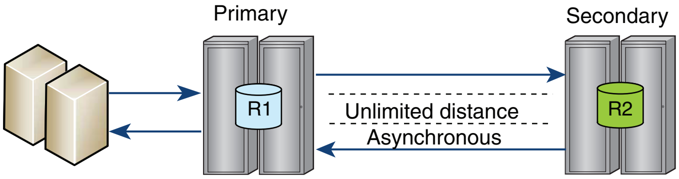
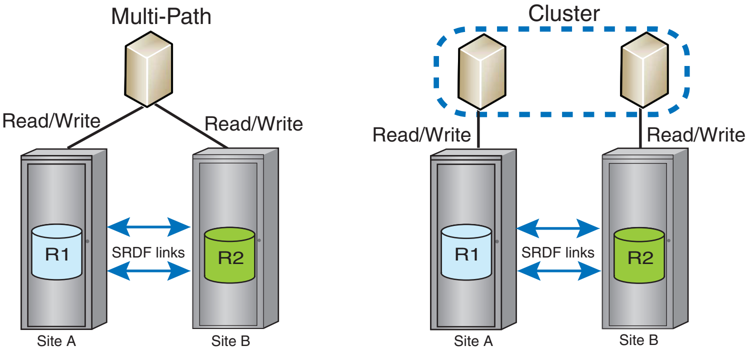
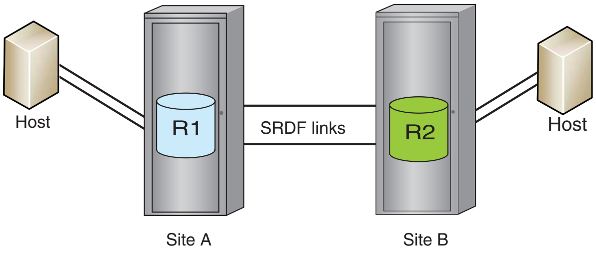
Solution highlights	Site topology
SRDF/Synchronous (SRDF/S) Maintains a real-time copy of production data at a physically separated array. - No data exposure - Ensured consistency protection with SRDF/ Consistency Group - Up to 125 miles (200 km) between arrays See: Write operations in synchronous mode on page 118.	 Primary Secondary R1 R2 Limited distance Synchronous
SRDF/Asynchronous (SRDF/A) Maintains a dependent-write consistent copy of the data on a remote secondary site. The copy of the data at the secondary site is seconds behind the primary site. - RPO seconds before the point of failure - Unlimited distance See: Write operations in asynchronous mode on page 119.	 Primary Secondary R1 R2 Unlimited distance Asynchronous
SRDF/Metro Host or hosts (cluster) read and write to both R1 and R2 devices. Each copy is current and consistent. Write conflicts between the paired SRDF devices are managed and resolved. Up to 125 miles (200 km) between arrays See: SRDF/Metro on page 137.	 Multi-Path Read/Write Read/Write R1 SRDF links R2 Site A Site B Cluster Read/Write Read/Write R1 SRDF links R2 Site A Site B
SRDF/Data Mobility (SRDF/DM) This example shows an SRDF/DM topology and the I/O flow in adaptive copy mode. - The host write I/O is received in cache in Site A - The host emulation returns a positive acknowledgment to the host - The SRDF emulation transmits the I/O across the SRDF links to Site B	

Table 34 SRDF 2-site solutions (continued)

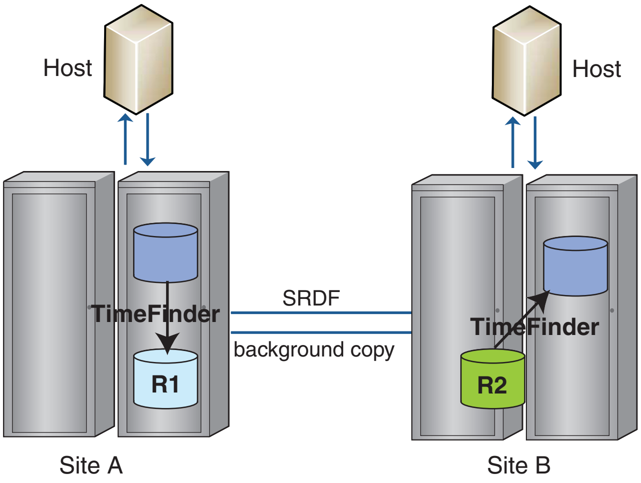
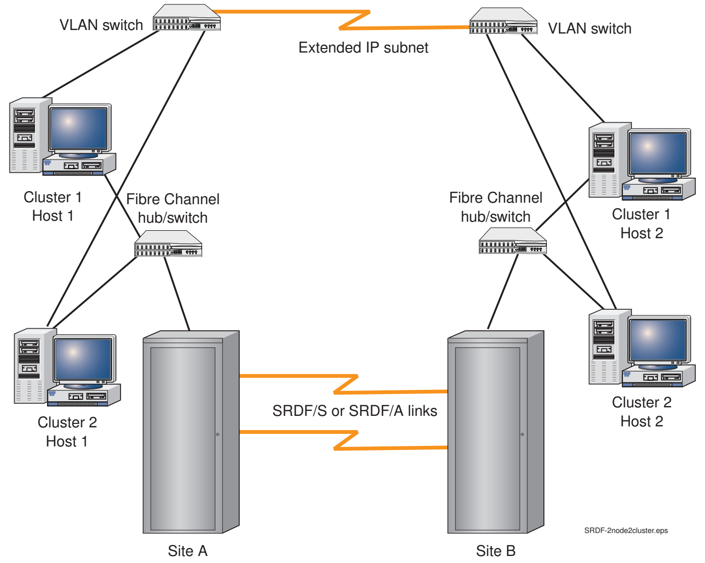
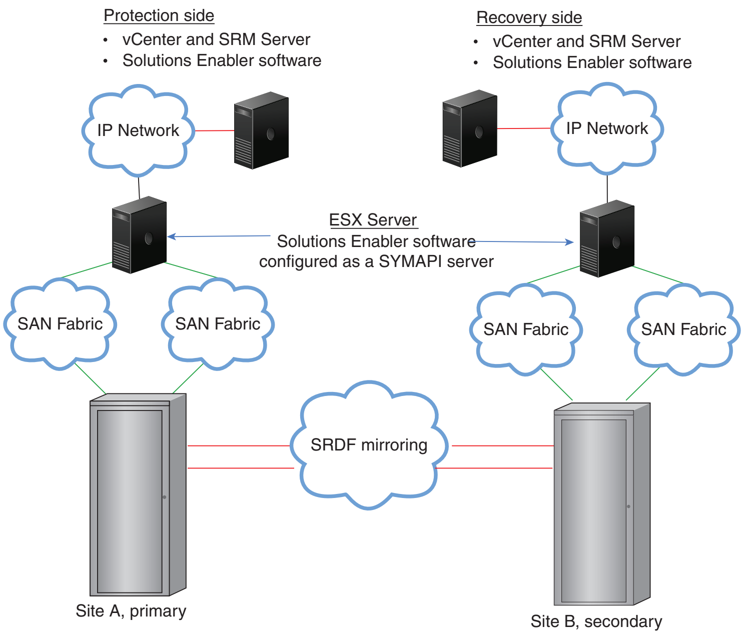
Solution highlights	Site topology
- Once data is written to cache in Site B, the SRDF emulation in Site B returns a positive acknowledgment to Site A Operating Notes: - The maximum skew value set at the device level in SRDF/DM solutions must be equal or greater than 100 tracks - SRDF/DM is only for data replication or migration, not for disaster restart solutions See: Adaptive copy modes on page 115.	**Note** Data may be read from the drives to cache before it is transmitted across the SRDF links, resulting in propagation delays.
SRDF/Automated Replication (SRDF/AR) - Combines SRDF and TimeFinder to optimize bandwidth requirements and provide a long-distance disaster restart solution. - Operates in 2-site solutions that use SRDF/DM in combination with TimeFinder. See: SRDF/AR on page 152.	 Host Host TimeFinder R1 SRDF background copy TimeFinder R2 Site A Site B
SRDF/Cluster Enabler (CE) - Integrates SRDF/S or SRDF/A with Microsoft Failover Clusters (MSCS) to automate or semi-automate site failover. - Complete solution for restarting operations in cluster environments (MSCS with Microsoft Failover Clusters) - Expands the range of cluster storage and management capabilities while ensuring full protection of the SRDF remote replication. For more information, see *EMC SRDF/Cluster Enabler Plug-in Product Guide*.	 VLAN switch Extended IP subnet VLAN switch Cluster 1 Host 1 Fibre Channel hub/switch Fibre Channel hub/switch Cluster 1 Host 2 Cluster 2 Host 1 SRDF/S or SRDF/A links Cluster 2 Host 2 SRDF-2node2cluster.eps Site A Site B

Table 34 SRDF 2-site solutions (continued)

Solution highlights	Site topology
SRDF and VMware Site Recovery Manager Completely automates storage-based disaster restart operations for VMware environments in SRDF topologies. - The EMC SRDF Adapter enables VMware Site Recovery Manager to automate storage-based disaster restart operations in SRDF solutions. - Can address configurations in which data are spread across multiple storage arrays or SRDF groups. - Requires that the adapter is installed on each array to facilitate the discovery of arrays and to initiate failover operations. - Implemented with: ■ SRDF/S ■ SRDF/A ■ SRDF/Star ■ TimeFinder For more information, see: - *Using EMC SRDF Adapter for VMware Site Recovery Manager Tech Book* - *EMC SRDF Adapter for VMware Site Recovery Manager Release Notes*	 Protection side • vCenter and SRM Server • Solutions Enabler software Recovery side • vCenter and SRM Server • Solutions Enabler software IP Network IP Network ESX Server Solutions Enabler software configured as a SYMAPI server SAN Fabric SAN Fabric SAN Fabric SAN Fabric SRDF mirroring Site A, primary Site B, secondary

SRDF multi-site solutions

The following table describes SRDF multi-site solutions.

Table 35 SRDF multi-site solutions

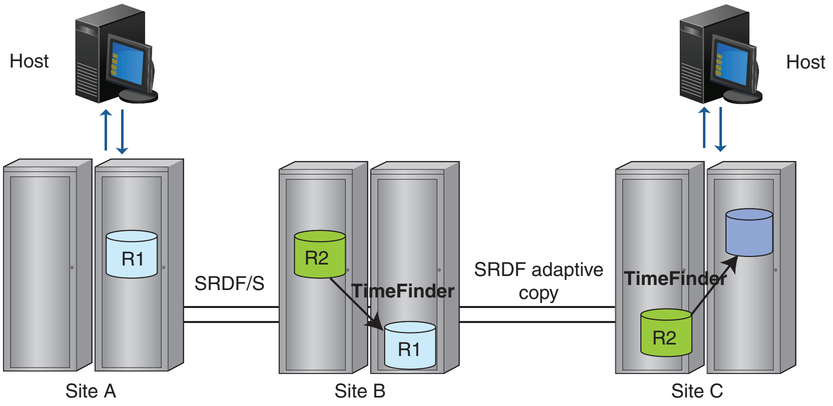
Solution highlights	Site topology
SRDF/Automated Replication (SRDF/AR) - Combines SRDF and TimeFinder to optimize bandwidth requirements and provide a long-distance disaster restart solution. - Operates in 3-site solutions that use a combination of SRDF/S, SRDF/DM, and TimeFinder.	 Host Host R1 SRDF/S R2 TimeFinder R1 SRDF adaptive copy TimeFinder R2 Site A Site B Site C

Table 35 SRDF multi-site solutions (continued)

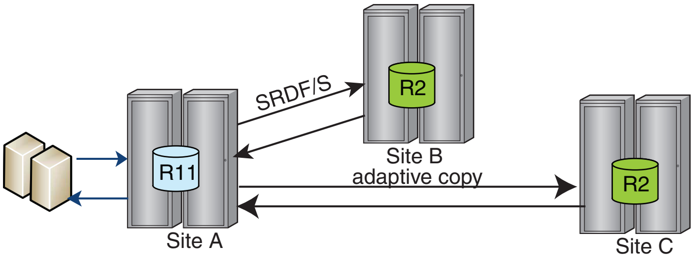
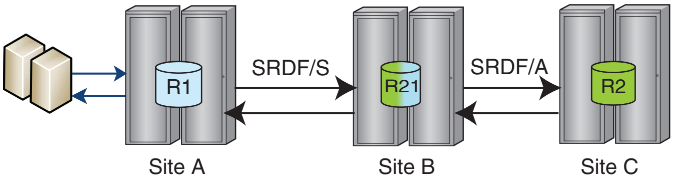
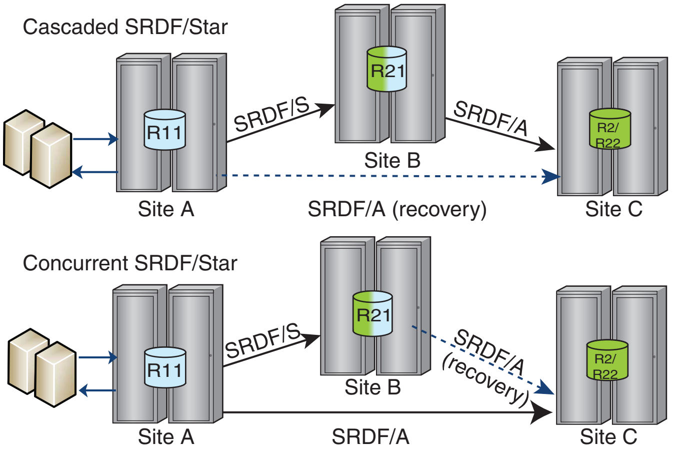
Solution highlights	Site topology
See: SRDF/AR on page 152.	
Concurrent SRDF 3-site disaster recovery and advanced multi-site business continuity protection. - Data on the primary site is concurrently replicated to 2 secondary sites. - Replication to remote site can use SRDF/S, SRDF/A, or adaptive copy See: Concurrent SRDF solutions on page 99.	 SRDF/S R2 Site B adaptive copy R11 Site A R2 Site C
Cascaded SRDF 3-site disaster recovery and advanced multi-site business continuity protection. - Data on the primary site is synchronously mirrored to a secondary (R21) site, and then asynchronously mirrored from the secondary (R21) site to a tertiary (R2) site. - First "hop" is SRDF/S. Second hop is SRDF/A. See: Cascaded SRDF solutions on page 100.	 R1 SRDF/S R21 SRDF/A R2 Site A Site B Site C
SRDF/Star 3-site data protection and disaster recovery with zero data loss recovery, business continuity protection and disaster-restart. - Available in 2 configurations: ■ Cascaded SRDF/Star ■ Concurrent SRDF/Star - Differential synchronization allows rapid reestablishment of mirroring among surviving sites in a multi-site disaster recovery implementation. - Implemented using SRDF consistency groups (CG) with SRDF/S and SRDF/A.	 Cascaded SRDF/Star R21 R11 SRDF/S SRDF/A R2/ R22 Site B Site A SRDF/A (recovery) Site C Concurrent SRDF/Star R21 SRDF/S R11 SRDF/A (recovery) R2/ R22 Site B Site A SRDF/A Site C

Table 35 SRDF multi-site solutions (continued)

Solution highlights	Site topology
See: SRDF/Star solutions on page 101.	

Concurrent SRDF solutions

Concurrent SRDF is a 3-site disaster recovery solution using R11 devices that replicate to two R2 devices. The two R2 devices operate independently but concurrently using any combination of SRDF modes:

- Concurrent SRDF/S to both R2 devices if the R11 site is within synchronous distance of the two R2 sites.
- Concurrent SRDF/A to sites located at extended distances from the workload site.

You can restore the R11 device from either of the R2 devices. You can restore both the R11 and one R2 device from the second R2 device.

Use concurrent SRDF to replace an existing R11 or R2 device with a new device. To replace an R11 or R2, migrate data from the existing device to a new device using adaptive copy disk mode, and then replace the existing device with the newly populated device.

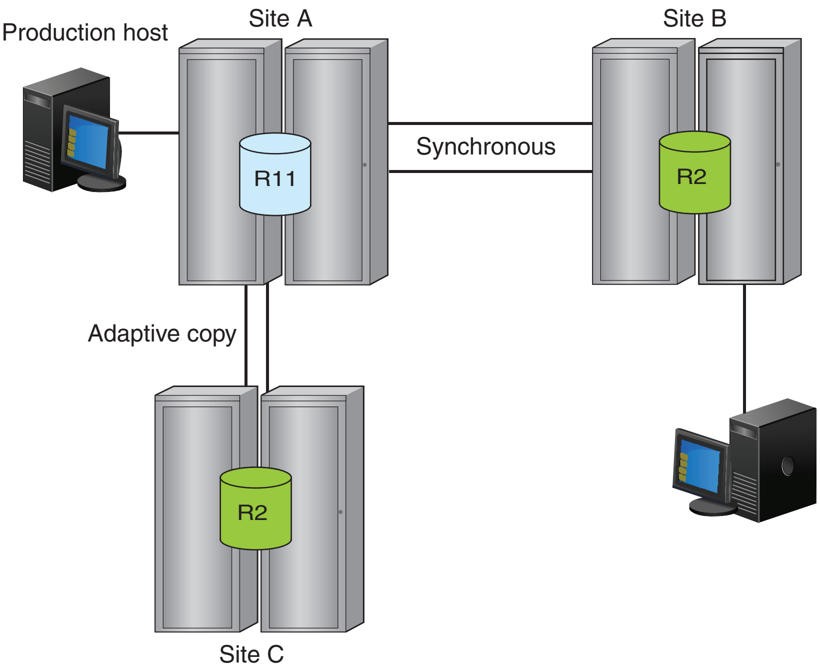
Concurrent SRDF can be implemented with SRDF/Star. SRDF/Star solutions on page 101 describes concurrent SRDF/Star.

Concurrent SRDF topologies are supported on Fibre Channel and Gigabit Ethernet.

The following image shows:

- The R11 -> R2 in Site B in synchronous mode.
- The R11 -> R2 in Site C in adaptive copy mode:

Figure 17 Concurrent SRDF topology



Concurrent SRDF/S with Enginuity Consistency Assist

If both legs of a concurrent SRDF configuration are SRDF/S, you can leverage the independent consistency protection feature. This feature is based on Enginuity Consistency Assist (ECA) and enables you to manage consistency on each concurrent SRDF leg independently.

If consistency protection on one leg is suspended, consistency protection on the other leg can remain active and continue protecting the primary site.

Cascaded SRDF solutions

Cascaded SRDF provides a zero data loss solution at long distances in the event that the primary site is lost.

In cascaded SRDF configurations, data from a primary (R1) site is synchronously mirrored to a secondary (R21) site, and then asynchronously mirrored from the secondary (R21) site to a tertiary (R2) site.

Cascaded SRDF provides:

- Fast recovery times at the tertiary site.
- Tight integration with TimeFinder product family.
- Geographically dispersed secondary and tertiary sites.

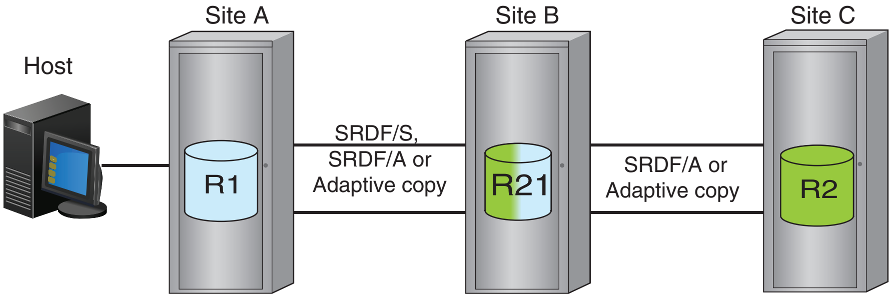
If the primary site fails, cascaded SRDF can continue mirroring, with minimal user intervention, from the secondary site to the tertiary site. This enables a faster recovery at the tertiary site.

Both the secondary and the tertiary site can be failover sites. Open systems solutions typically fail over to the tertiary site.

Cascaded SRDF can be implemented with SRDF/Star. Cascaded SRDF/Star on page 103 describes cascaded SRDF/Star.

The following image shows a cascaded SRDF topology.

Figure 18 Cascaded SRDF topology



SRDF/Star solutions

SRDF/Star is a disaster recovery solution that consists of three sites; primary (production), secondary, and tertiary. The secondary site synchronously mirrors the data from the primary site, and the tertiary site asynchronously mirrors the production data.

In the event of an outage at the primary site, SRDF/Star allows you to quickly move operations and re-establish remote mirroring between the remaining sites. When conditions permit, you can quickly rejoin the primary site to the solution, resuming the SRDF/Star operations.

SRDF/Star operates in concurrent and cascaded environments that address different recovery and availability objectives:

- Concurrent SRDF/Star — Data is mirrored from the primary site concurrently to two R2 devices. Both the secondary and tertiary sites are potential recovery sites. Differential resynchronization is used between the secondary and the tertiary sites.
- Cascaded SRDF/Star — Data is mirrored first from the primary site to a secondary site, and then from the secondary to a tertiary site. Both the secondary and tertiary sites are potential recovery sites. Differential resynchronization is used between the primary and the tertiary site.

Differential synchronization between two remote sites:

- Allows SRDF/Star to rapidly reestablish cross-site mirroring in the event of the primary site failure.
- Greatly reduces the time required to remotely mirror the new production site.

In the event of a rolling disaster that affects the primary site, SRDF/Star helps you determine which remote site has the most current data. You can select which site to operate from and which site's data to use when recovering from the primary site failure.

If the primary site fails, SRDF/Star allows you to resume asynchronous protection between the secondary and tertiary sites, with minimal data movement.

SRDF/Star for open systems

Solutions Enabler controls, manages, and automates SRDF/Star in open systems environments. Session management is required at the production site.

Host-based automation is provided for normal, transient fault, and planned or unplanned failover operations.

EMC Solutions Enabler Symmetrix SRDF CLI Guide provides detailed descriptions and implementation guidelines.

In cascaded and concurrent configurations, a restart from the asynchronous site may require a wait for any remaining data to arrive from the synchronous site. Restarts from the synchronous site requires no wait unless the asynchronous site is more recent (the latest updates need to be brought to the synchronous site).

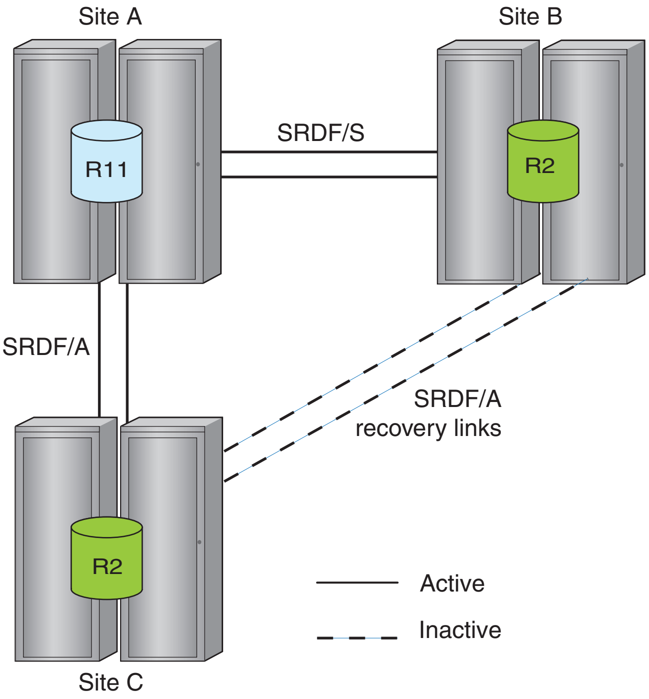
Concurrent SRDF/Star

In concurrent SRDF/Star solutions, production data on R11 devices replicates to two R2 devices in two remote arrays.

In the following image:

- Site B is a secondary site using SRDF/S links from Site A.
- Site C is a tertiary site using SRDF/A links from Site A.
- The (normally inactive) recovery links are SRDF/A between Site C and Site B.

Figure 19 Concurrent SRDF/Star

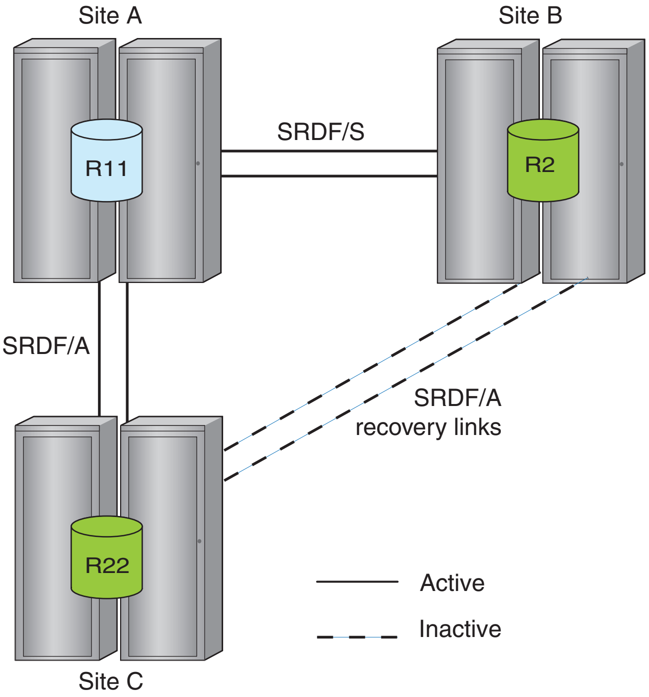


Concurrent SRDF/Star with R22 devices

SRDF supports concurrent SRDF/Star topologies using concurrent R22 devices. R22 devices have two SRDF mirrors, only one of which is active on the SRDF links at a given time. R22 devices improve the resiliency of the SRDF/Star application, and reduce the number of steps for failover procedures.

The following image shows R22 devices at Site C.

Figure 20 Concurrent SRDF/Star with R22 devices



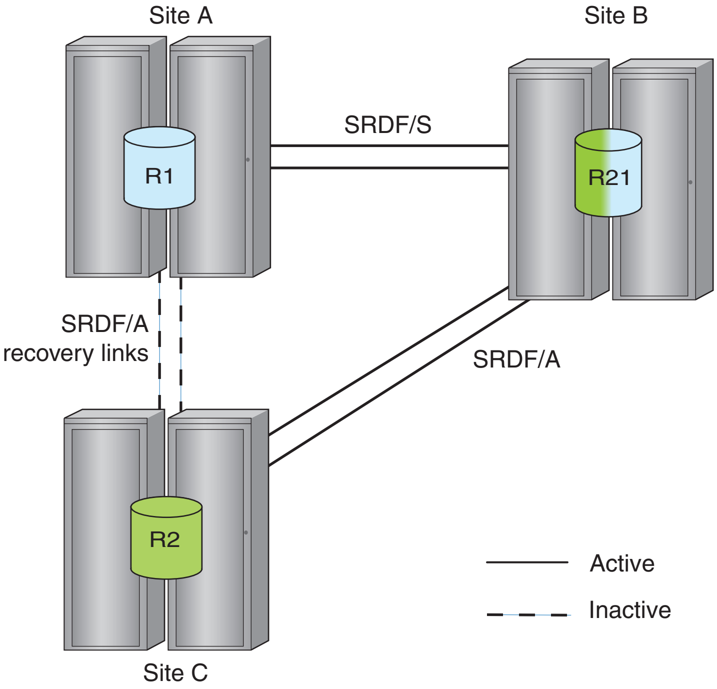
Cascaded SRDF/Star

In cascaded SRDF/Star solutions, the synchronous secondary site is always more current than the asynchronous tertiary site. If the synchronous secondary site fails, the cascaded SRDF/Star solution can incrementally establish an SRDF/A session between primary site and the asynchronous tertiary site.

Cascaded SRDF/Star can determine when the current active R1 cycle (capture) contents reach the active R2 cycle (apply) over the long-distance SRDF/A links. This minimizes the amount of data that must be moved between Site B and Site C to fully synchronize them.

The following image shows a basic cascaded SRDF/Star solution.

Figure 21 Cascaded SRDF/Star

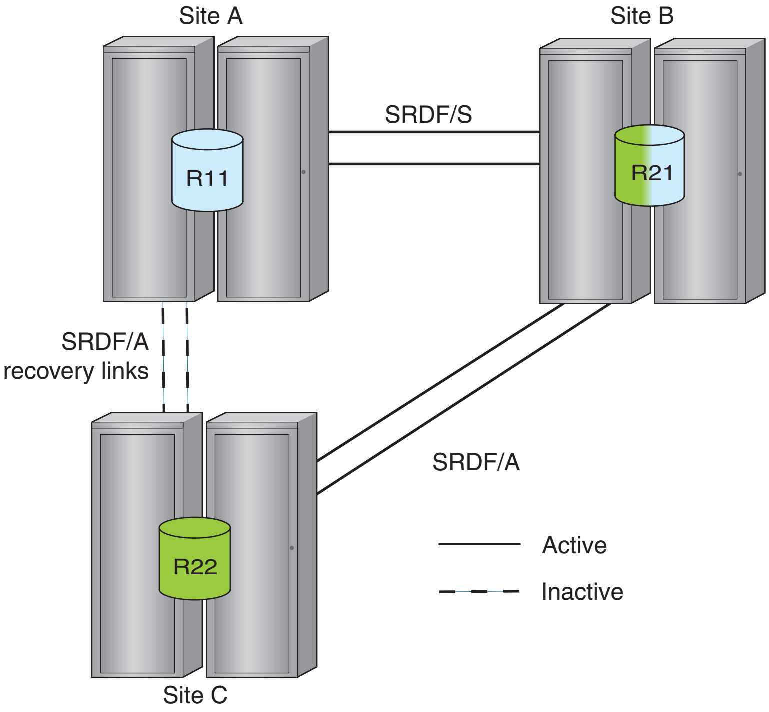


Cascaded SRDF/Star with R22 devices

You can use R22 devices to pre-configure the SRDF pairs required to incrementally establish an SRDF/A session between Site A and Site C in case Site B fails.

The following image shows cascaded R22 devices in a cascaded SRDF solution.

Figure 22 R22 devices in cascaded SRDF/Star



In cascaded SRDF/Star configurations with R22 devices:

- All devices at the production site (Site A) must be configured as concurrent (R11) devices paired with R21 devices (Site B) and R22 devices (Site C).
- All devices at the synchronous site in Site B must be configured as R21 devices.
- All devices at the asynchronous site in Site C must be configured as R22 devices.

Requirements/restrictions
Cascaded and Concurrent SRDF/Star configurations (with and without R22 devices) require the following:

- All SRDF/Star device pairs must be of the same geometry and size.
- All SRDF groups including inactive ones must be defined and operational prior to entering SRDF/Star mode.
- It is strongly recommended that all SRDF devices be locally protected and that each SRDF device is configured with TimeFinder to provide local replicas at each site.

SRDF four-site solutions for open systems

The four-site SRDF solution for open systems host environments replicates FBA data by using both concurrent and cascaded SRDF topologies.

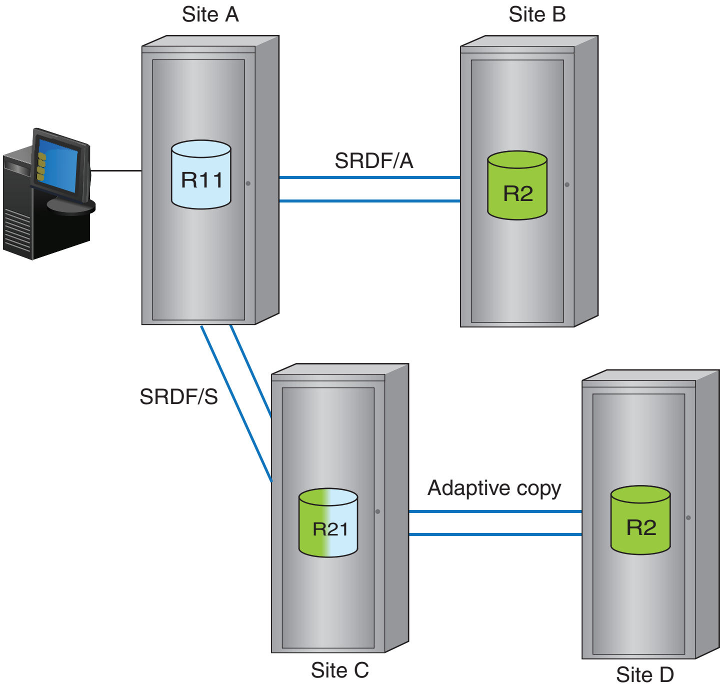
Four-site SRDF is a multi-region disaster recovery solution with higher availability, improved protection, and less downtime than concurrent or cascaded SRDF solutions.

Four-site SRDF solution offers multi-region high availability by combining the benefits of concurrent and cascaded SRDF solutions.

If two sites fail because of a regional disaster, a copy of the data is available, and you have protection between the remaining two sites. You can create a four-site SRDF topology from an existing 2-site or 3-site SRDF topology. Four-site SRDF can also be used for data migration.

The following image shows an example of the four-site SRDF solution.

Figure 23 Four-site SRDF



Interfamily compatibility

SRDF supports connectivity between different operating environments and arrays. Arrays running HYPERMAX OS can connect to legacy arrays running older operating environments. In mixed configurations where arrays are running different versions, SRDF features of the lowest version are supported.

VMAX All Flash arrays can connect to:

- VMAX 250F, 450F, 850F, and 950F arrays running HYPERMAX OS
- VMAX 100K, 200K, and 400K arrays running HYPERMAX OS
- VMAX 10K, 20K, and 40K arrays running Enginuity 5876 with an Enginuity ePack

Note

When you connect between arrays running different operating environments, limitations may apply. Information about which SRDF features are supported, and applicable limitations for 2-site and 3-site solutions is available in the *SRDF Interfamily Connectivity Information*.

This interfamily connectivity allows you to add the latest hardware platform/operating environment to an existing SRDF solution, enabling technology refreshes.

Different operating environments offer different SRDF features.

SRDF supported features

The following table lists the SRDF features supported on each hardware platform and operating environment.

Table 36 SRDF features by hardware platform/operating environment

Feature	**Enginuity 5876**		**HYPERMAX OS 5977**	
	VMAX 40K, VMAX 20K	**VMAX 10K**	**VMAX3**	**VMAX 250F, 450F, 850F, 950F**
Max. SRDF devices/SRDF emulation (either Fibre Channel or GigE)	64K	8K	64K	64K
Max. SRDF groups/array	250	32	250	250
Max. SRDF groups/SRDF emulation instance (either Fibre Channel or GigE)	64	32	250[ab]	250[cd]
Max. remote targets/port	64	64	16K/SRDF emulation (either Fibre Channel or GigE)	16K/SRDF emulation (either Fibre Channel or GigE)
Max. remote targets/SRDF group	N/A	N/A	512	512
Fibre Channel port speed	2/4/8 Gb/s 16 Gb/s on 40K	2/4/8/16 Gb/s	16 Gb/s	16 Gb/s
GbE port speed	1 /10 Gb/s	1 /10 Gb/s	1 /10 Gb/s	1 /10 Gb/s
Min. SRDF/A Cycle Time	1 sec, 3 secs with MSC	1 sec, 3 secs with MSC	1 sec, 3 secs with MSC	1 sec, 3 secs with MSC
SRDF Delta Set Extension	Supported	Supported	Supported	Supported
Transmit Idle	Enabled	Enabled	Enabled	Enabled
Fibre Channel Single Round Trip (SiRT)	Enabled	Enabled	Enabled	Enabled
GigE SRDF Compression				
Software	Supported • VMAX 20K • VMAX 40K: Enginuity 5876.82.57 or higher	Supported	Supported	Supported
Fibre Channel SRDF Compression				
Software	Supported • VMAX 20K • VMAX 40K: Enginuity 5876.82.57 or higher	Supported	Supported	Supported
IPv6 and IPsec				
IPv6 feature on 10 GbE	Supported	Supported	Supported	Supported
IPsec encryption on 1 GbE ports	Supported	Supported	N/A	N/A

Table 36 SRDF features by hardware platform/operating environment (continued)

a. If both arrays are running HYPERMAX OS, up to 250 RDF groups can be defined across all of the ports on a specific RDF director, or up to 250 RDF groups can be defined on 1 port on a specific RDF director.
b. A port on the array running HYPERMAX OS connected to an array running Enginuity 5876 supports a maximum of 64 RDF groups. The director on the HYPERMAX OS side associated with that port supports a maximum of 186 (250 – 64) RDF groups.
c. If both arrays are running HYPERMAX OS, up to 250 RDF groups can be defined across all of the ports on a specific RDF director, or up to 250 RDF groups can be defined on 1 port on a specific RDF director.
d. A port on the array running HYPERMAX OS connected to an array running Enginuity 5876 supports a maximum of 64 RDF groups. The director on the HYPERMAX OS side associated with that port supports a maximum of 186 (250 – 64) RDF groups.

HYPERMAX OS and Enginuity compatibility

Arrays running HYPERMAX OS cannot create a device that is exactly the same size as a device with an odd number of cylinders on an array running Enginuity 5876. In order to support the full suite of features:

- SRDF requires that R1 and R2 devices in a device pair be the same size.
- TimeFinder requires that source and target devices are the same size.

Track size for FBA devices increased from 64Kb in Enginuity 5876 to 128Kb in HYPERMAX OS.

HYPERMAX OS introduces a new device attribute, Geometry Compatible Mode (GCM). A device with GCM set is treated as half a cylinder smaller than its true configured size, enabling full functionality between HYPERMAX OS and Enginuity 5876 for SRDF, TimeFinder SnapVX, and TimeFinder emulations (TimeFinder/Clone, TimeFinder VP Snap, TimeFinder/Mirror), and ORS.

The GCM attribute can be set in the following ways:

NOTICE

Do not set GCM on devices that are mounted and under Local Volume Manager (LVM) control.

- Automatically on a target of an SRDF or TimeFinder relationship if the source is either a 5876 device with an odd number of cylinders, or a 5977 source that has GCM set.
- Manually using Base Controls interfaces. The *EMC Solutions Enabler SRDF Family CLI User Guide* provides additional details.

SRDF device pairs

An SRDF device is a logical device paired with another logical device that resides in a second array. The arrays are connected by SRDF links.

Encapsulated Data Domain devices used for ProtectPoint cannot be part of an SRDF device pair.

R1 and R2 devices

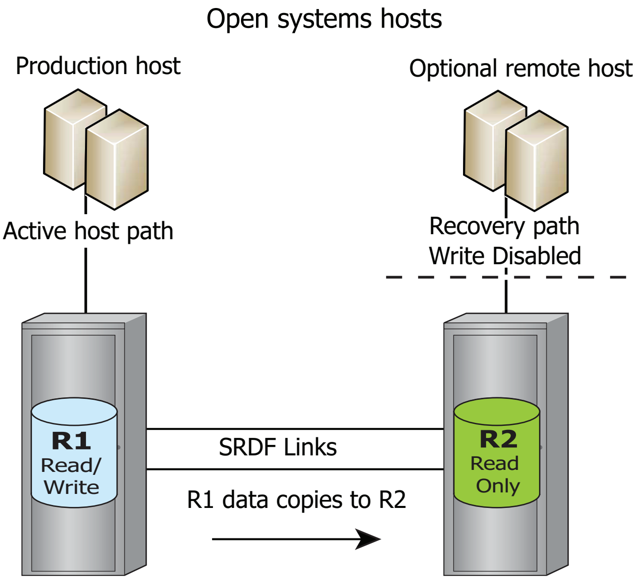
R1 devices are the member of the device pair at the source (production) site. R1 devices are generally Read/Write accessible to the host.

R2 devices are the members of the device pair at the target (remote) site. During normal operations, host I/O writes to the R1 device are mirrored over the SRDF links to the R2 device. In general, data on R2 devices is not available to the host while the SRDF relationship is active. In SRDF synchronous mode, an R2 device can be in Read Only mode that allows a host to read from the R2.

In a typical open systems host environment:

- The production host has Read/Write access to the R1 device.
- A host connected to the R2 device has Read Only (Write Disabled) access to the R2 device.

Figure 24 R1 and R2 devices



Invalid tracks

Invalid tracks are tracks that are not synchronized, that is, they are tracks that are “owed” between the two devices in an SRDF pair.

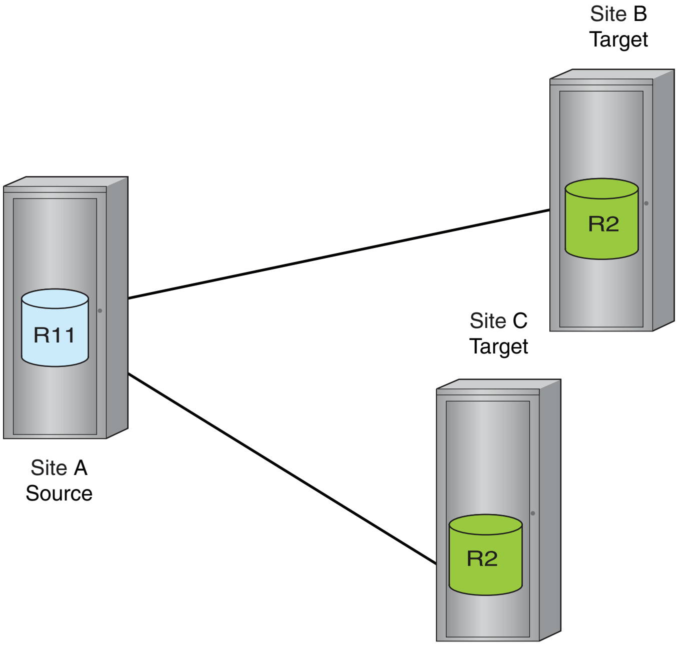
R11 devices

R11 devices operate as the R1 device for two R2 devices. Links to both R2 devices are active.

R11 devices are typically used in SRDF/Concurrent solutions where data on the R11 site is mirrored to two secondary (R2) arrays.

The following image shows an R11 device in an SRDF/Concurrent Star solution.

Figure 25 R11 device in concurrent SRDF



R21 devices

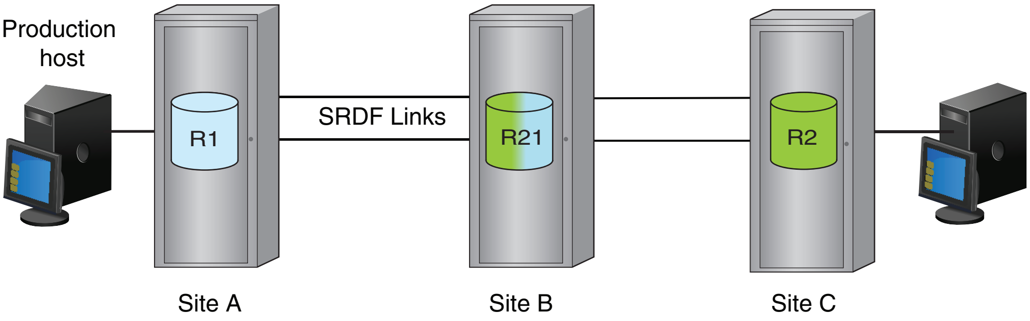
R21 devices operate as:

- R2 devices to hosts connected to array containing the R1 device, and
- R1 device to hosts connected to the array containing the R2 device.

R21 devices are typically used in cascaded 3-site solutions where:

- Data on the R1 site is synchronously mirrored to a secondary (R21) site, and then
- Synchronously mirrored from the secondary (R21) site to a tertiary (R2) site:

Figure 26 R21 device in cascaded SRDF



When the R1->R21->R2 SRDF relationship is established, no host has write access to the R21 device.

Note

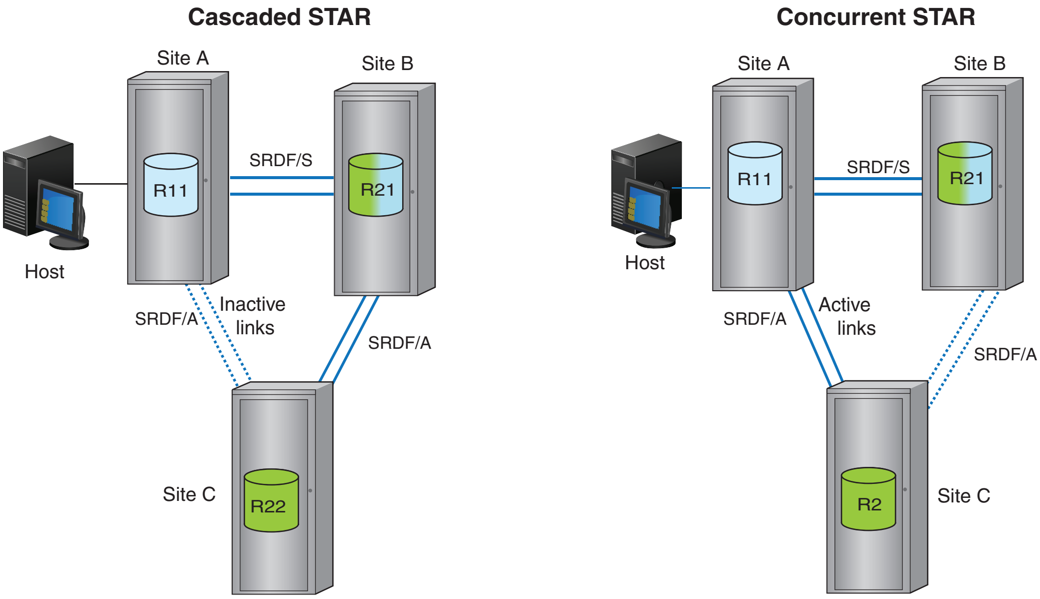
Diskless R21 devices are not supported on arrays running HYPERMAX OS.

R22 devices

R22 devices:

- Have two R1 devices, only one of which is active at a time.
- Are typically used in cascaded SRDF/Star and concurrent SRDF/Star solutions to decrease the complexity and time required to complete failover and failback operations.
- Let you recover without removing old SRDF pairs and creating new ones.

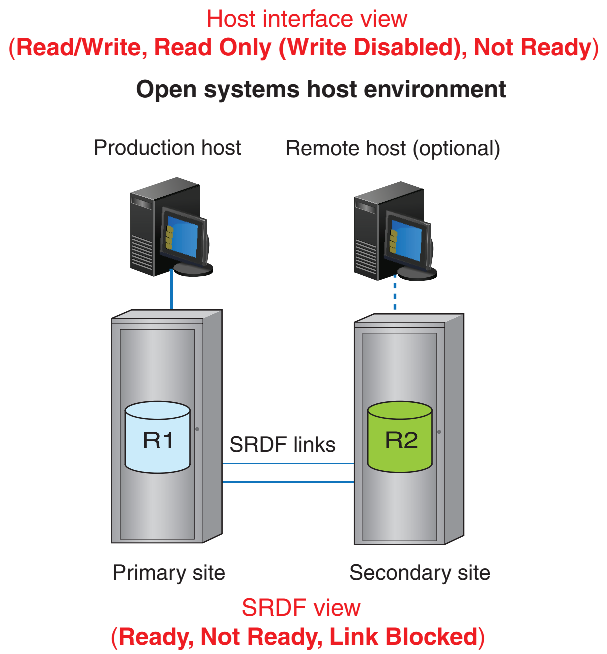
Figure 27 R22 devices in cascaded and concurrent SRDF/Star



SRDF device states

An SRDF device’s state is determined by a combination of two views; host interface view and SRDF view, as shown in the following image.

Figure 28 Host interface view and SRDF view of states


Host interface view
(Read/Write, Read Only (Write Disabled), Not Ready)
Open systems host environment
Production host
Remote host (optional)
R1
SRDF links
R2
Primary site
Secondary site
SRDF view
(Ready, Not Ready, Link Blocked)

Host interface view

The host interface view is the SRDF device state as seen by the host connected to the device.

R1 device states

An R1 device presents one of the following states to the host connected to the primary array:

- Read/Write (Write Enabled)—The R1 device is available for Read/Write operations. This is the default R1 device state.
- Read Only (Write Disabled)—The R1 device responds with Write Protected to all write operations to that device.
- Not Ready—The R1 device responds Not Ready to the host for read and write operations to that device.

R2 device states

An R2 device presents one of the following states to the host connected to the secondary array:

- Read Only (Write Disabled)—The secondary (R2) device responds Write Protected to the host for all write operations to that device.
- Read/Write (Write Enabled)—The secondary (R2) device is available for read/write operations. This state is possible in recovery or parallel processing operations.
- Not Ready—The R2 device responds Not Ready (Intervention Required) to the host for read and write operations to that device.

SRDF view

The SRDF view is composed of the SRDF state and internal SRDF device state. These states indicate whether the device is available to send data across the SRDF links, and able to receive software commands.

R1 device states

An R1 device can have the following states for SRDF operations:

- Ready—The R1 device is ready for SRDF operations.
  The R1 device is able to send data across the SRDF links.

  True even if local mirror(s) of the R1 device are Not Ready for I/O operations.
- Not Ready (SRDF mirror Not Ready)—The R1 device is Not Ready for SRDF operations.

Note

When the R2 device is placed into a Read/Write state to the host, the corresponding R1 device is automatically placed into the SRDF mirror Not Ready state.

R2 device states

An R2 device can have the following states for SRDF operations:

- Ready—The R2 device receives the updates propagated across the SRDF links and can accept SRDF host-based software commands.
- Not Ready—The R2 device cannot accept SRDF host-based software commands, but can still receive updates propagated from the primary array.
- Link blocked (LnkBlk) — Applicable only to R2 SRDF mirrors that belong to R22 devices.
  One of the R2 SRDF mirrors cannot receive writes from its associated R1 device.
  In normal operations, one of the R2 SRDF mirrors of the R22 device is in this state.

R1/R2 device accessibility

Accessibility of a SRDF device to the host depends on both the host and the array view of the SRDF device state.

Table 37 on page 113 and Table 38 on page 114 list host accessibility for R1 and R2 devices.

Table 37 R1 device accessibility

Host interface state	SRDF state	Accessibility
Read/Write	Ready	Read/Write
	Not Ready	Depends on R2 device availability
Read Only	Ready	Read Only
	Not Ready	Depends on R2 device availability
Not Ready	Any	Unavailable

Table 38 R2 device accessibility

Host interface state	SRDF R2 state	Accessibility
Write Enabled (Read/ Write)	Ready	Read/Write
	Not Ready	Read/Write
Write Disabled (Read Only)	Ready	Read Only
	Not Ready	Read Only
Not Ready	Any	Unavailable

Dynamic device personalities

SRDF devices can dynamically swap “personality” between R1 and R2. After a personality swap:

- The R1 in the device pair becomes the R2 device, and
- The R2 becomes the R1 device.

Swapping R1/R2 personalities allows the application to be restarted at the remote site without interrupting replication if an application fails at the production site. After a swap, the R2 side (now R1) can control operations while being remotely mirrored at the primary (now R2) site.

An R1/R2 personality swap is not supported:

- If the R2 device is larger than the R1 device.
- If the device to be swapped is participating in an active SRDF/A session.
- In SRDF/EDP topologies diskless R11 or R22 devices are not valid end states.
- If the device to be swapped is the target device of any TimeFinder or EMC Compatible flash operations.

SRDF modes of operation

SRDF modes of operation address different service level requirements and determine:

- How R1 devices are remotely mirrored across the SRDF links.
- How I/Os are processed.
- When the host receives acknowledgment of a write operation relative to when the write is replicated.
- When writes “owed” between partner devices are sent across the SRDF links.

The mode of operation may change in response to control operations or failures:

- The primary mode (synchronous or asynchronous) is the configured mode of operation for a given SRDF device, range of SRDF devices, or an SRDF group.
- The secondary mode is adaptive copy. Adaptive copy mode moves large amounts of data quickly with minimal host impact. Adaptive copy mode does not provide restartable data images at the secondary site until no new writes are sent to the R1 device and all data has finished copying to the R2.

Use adaptive copy mode to synchronize new SRDF device pairs or to migrate data to another array. When the synchronization or migration is complete, you can revert to the configured primary mode of operation.

Synchronous mode

SRDF/S maintains a real-time mirror image of data between the R1 and R2 devices over distances of ~200 km or less.

Host writes are written simultaneously to both arrays in real time before the application I/O completes. Acknowledgments are not sent to the host until the data is stored in cache on both arrays.

Refer to Write operations in synchronous mode on page 118 and SRDF read operations on page 127 for more information.

Asynchronous mode

SRDF/Asynchronous (SRDF/A) maintains a dependent-write consistent copy between the R1 and R2 devices across any distance with no impact to the application.

Host writes are collected for a configurable interval into “delta sets”. Delta sets are transferred to the remote array in timed cycles.

SRDF/A operations vary depending on whether the SRDF session mode is single or multi-session with Multi Session Consistency (MSC) enabled:

- For single SRDF/A sessions, cycle switching is controlled by Enginuity. Each session is controlled independently, whether it is in the same or multiple arrays.
- For multiple SRDF/A sessions in MSC mode, multiple SRDF groups are in the same SRDF/A MSC session. Cycle switching is controlled by SRDF host software to maintain consistency.

Refer to SRDF/A MSC cycle switching on page 122 for more information.

Adaptive copy modes

Adaptive copy modes:

- Transfer large amounts of data without impact on the host.
- Transfer data during data center migrations and consolidations, and in data mobility environments.
- Allow the R1 and R2 devices to be out of synchronization by up to a user-configured maximum skew value. If the maximum skew value is exceeded, SRDF starts the synchronization process to transfer updates from the R1 to the R2 devices
- Are secondary modes of operation for SRDF/S. The R1 devices revert to SRDF/S when the maximum skew value is reached and remain in SRDF/S until the number of tracks out of synchronization is lower than the maximum skew.

There are two types of adaptive copy mode:

- Adaptive copy disk on page 115
- Adaptive copy write pending on page 116

Note

Adaptive copy write pending mode is not supported when the R1 side of an SRDF device pair is on an array running HYPERMAX OS.

Adaptive copy disk

In adaptive copy disk mode, write requests accumulate on the R1 device (not in cache). A background process sends the outstanding write requests to the corresponding R2 device. The background copy process scheduled to send I/Os from the R1 to the R2 devices can be deferred if:

- The write requests exceed the maximum R2 write pending limits, or
- The write requests exceed 50 percent of the primary or secondary array write pending space.

Adaptive copy write pending

In adaptive copy write pending mode, write requests accumulate in cache on the primary array. A background process sends the outstanding write requests to the corresponding R2 device.

Adaptive copy write-pending mode reverts to the primary mode if the device, cache partition, or system write pending limit is near, regardless of whether the maximum skew value specified for each device is reached.

Domino modes

Under typical conditions, when one side of a device pair becomes unavailable, new data written to the device is marked for later transfer. When the device or link is restored, the two sides synchronize.

Domino modes force SRDF devices into the Not Ready state to the host if one side of the device pair becomes unavailable.

Domino mode can be enabled/disabled at:

- Device level (domino mode) – If the R1 device cannot successfully mirror data to the R2 device, the next host write to the R1 device causes the device to become Not Ready to the host connected to the primary array.
- SRDF group level (link domino mode) – If the last available link in the SRDF group fails, the next host write to any R1 device in the SRDF group causes all R1 devices in the SRDF group become Not Ready to their hosts.

Link domino mode is set at the SRDF group level and only impacts devices where the R1 is on the side where it is set.

SRDF groups

SRDF groups define the relationships between the local SRDF instance and the corresponding remote SRDF instance.

All SRDF devices must be assigned to an SRDF group. Each SRDF group communicates with its partner SRDF group in another array across the SRDF links. Each SRDF group points to one (and only one) remote array.

An SRDF group consists of one or more SRDF devices, and the ports over which those devices communicate. The SRDF group shares CPU processing power, ports, and a set of configurable attributes that apply to all the devices in the group, including:

- Link Limbo and Link Domino modes
- Autolink recovery
- Software compression
- SRDF/A:
  - Cycle time
  - Session priority
  - Pacing delay and threshold

Note

SRDF/A device pacing is not supported in HYPERMAX OS.

Starting in HYPERMAX OS, all SRDF groups are dynamic.

Moving dynamic devices between SRDF groups

You can move dynamic SRDF devices between groups in SRDF/S, SRDF/A and SRDF/A MSC solutions without incurring a full synchronization. This incremental synchronization reduces traffic on the links when you:

- Transition to a different SRDF topology and require minimal exposure during device moves.
- Add new SRDF devices to an existing SRDF/A group and require fast synchronization with the existing SRDF/A devices in the group.

Director boards, links, and ports

SRDF links are the logical connections between SRDF groups and their ports. The ports are physically connected by cables, routers, extenders, switches and other network devices.

Note

Two or more SRDF links per SRDF group are required for redundancy and fault tolerance.

The relationship between the resources on a director (CPU cores and ports) varies depending on the operating environment.

HYPERMAX OS

On arrays running HYPERMAX OS:

- The relationship between the SRDF emulation and resources on a director is configurable:
  - One director/multiple CPU cores/multiple ports
  - Connectivity (ports in the SRDF group) is independent of compute power (number of CPU cores). You can change the amount of connectivity without changing compute power.
- Each director has up to 12 front end ports, any or all of which can be used by SRDF. Both the SRDF Gigabit Ethernet and SRDF Fibre Channel emulations can use any port.
- The data path for devices in an SRDF group is not fixed to a single port. Instead, the path for data is shared across all ports in the group.

Mixed configurations: HYPERMAX OS and Enginuity 5876

For configurations where one array is running Enginuity 5876, and the second array is running HYPERMAX OS, the following rules apply:

- On the 5876 side, an SRDF group can have the full complement of directors, but no more than 16 ports on the HYPERMAX OS side.

- You can connect to 16 directors using one port each, 2 directors using 8 ports each or any other combination that does not exceed 16 per SRDF group.

SRDF consistency

Many applications (in particular, DBMS), use dependent write logic to ensure data integrity in the event of a failure. A dependent write is a write that is not issued by the application unless some prior I/O has completed. If the writes are out of order, and an event such as a failure, or a creation of a point in time copy happens at that exact time, unrecoverable data loss may occur.

An SRDF consistency group (SRDF/CG) is comprised of SRDF devices with consistency enabled.

SRDF consistency groups preserve the dependent-write consistency of devices within a group by monitoring data propagation from source devices to their corresponding target devices. If consistency is enabled, and SRDF detects any write I/O to a R1 device that cannot communicate with its R2 device, SRDF suspends the remote mirroring for all devices in the consistency group before completing the intercepted I/O and returning control to the application.

In this way, SRDF/CG prevents a dependent-write I/O from reaching the secondary site if the previous I/O only gets as far as the primary site.

SRDF consistency allows you to quickly recover from certain types of failure or physical disasters by retaining a consistent, DBMS-restartable copy of your database.

SRDF consistency group protection is available for both SRDF/S and SRDF/A.

SRDF write operations

This section describes SRDF write operations.

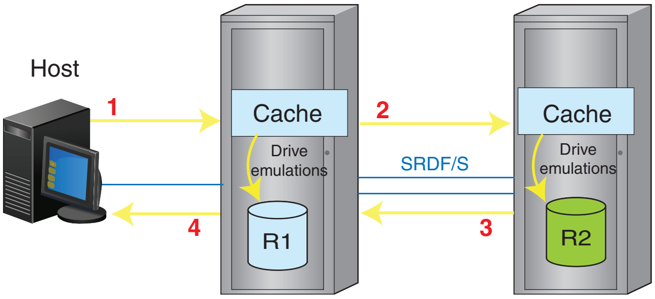
Write operations in synchronous mode

In synchronous mode, data must be successfully written to cache at the secondary site before a positive command completion status is returned to the host that issued the write command.

The following image shows the steps in a synchronous write operation:

1. The local host sends a write command to the local array.
   The host emulations write data to cache and create a write request.
2. SRDF emulations frame updated data in cache according to the SRDF protocol, and transmit it across the SRDF links.
3. The SRDF emulations in the remote array receive data from the SRDF links, write it to cache and return an acknowledgment to SRDF emulations in the local array.
4. The SRDF emulations in the local array forward the acknowledgment to host emulations.

Figure 29 Write I/O flow: simple synchronous SRDF



Write operations in asynchronous mode

In asynchronous mode (SRDF/A), host write I/Os are collected into delta sets on the primary array and transferred in cycles to the secondary array.

SRDF/A sessions behave differently depending on:

- Whether they are managed individually (Single Session Consistency (SSC)) or as a consistency group (Multi Session Consistency (MSC)).
  - In Single Session Consistency (SSC) mode, the SRDF group is managed individually, with cycle switching controlled by Enginuity or HYPERMAX OS. SRDF/A cycles are switched independently of any other SRDF groups on any array in the solution. Cycle switching in asynchronous mode on page 120 provides additional details.
  - In Multi Session Consistency (MSC) mode, the SRDF group is part of a consistency group spanning all associated SRDF/A sessions. Cycle switching is coordinated to provide dependent-write consistency across multiple sessions, which may also span arrays. Cycle switching controlled by SRDF host software. SRDF/A cycles are switched for all SRDF groups in the consistency group at the same time. SRDF/A MSC cycle switching on page 122 provides additional details.
- The number of transmit cycles supported at the R1 side. Enginuity 5876 supports only a single cycle. HYPERMAX OS supports multiple cycles queued to be transferred.

SRDF sessions can be managed individually or as members of a group.

In asynchronous mode, I/Os are collected into delta sets. Data is processed using 4 cycle types that capture, transmit, receive and apply delta sets:

- **Capture cycle**—Incoming I/O is buffered in the capture cycle on the R1 side. The host receives immediate acknowledgment.
- **Transmit cycle**—Data collected during the capture cycle is moved to the transmit cycle on the R1 side.
- **Receive cycle**—Data is received on the R2 side.
- **Apply cycle**—Changed blocks in the delta set are marked as invalid tracks and destaging to disk begins.
  A new receive cycle is started.

The start of the next capture cycle and the number of cycles on the R1 side vary depending on the version of the operating environment on the array participating in the SRDF/A solution.

- **HYPERMAX OS—Multi-cycle mode**—If both arrays in the solution are running HYPERMAX OS, SRDF/A operates in multi-cycle mode. There can be 2 or more cycles on the R1, but only 2 cycles on the R2 side:
  - On the R1 side:
    - One Capture
    - One or more Transmit
  - On the R2 side:
    - One Receive
    - One Apply

  Cycle switches are decoupled from committing delta sets to the next cycle. When the preset Minimum Cycle Time is reached, the R1 data collected during the capture cycle is added to the transmit queue and a new R1 capture cycle is started. There is no wait for the commit on the R2 side before starting a new capture cycle.

  The transmit queue holds cycles waiting to be transmitted to the R2 side. Data in the transmit queue is committed to the R2 receive cycle when the current transmit cycle and apply cycle are empty.

  Queuing allows smaller cycles of data to be buffered on the R1 side and smaller delta sets to be transferred to the R2 side.

  The SRDF/A session can adjust to accommodate changes in the solution. If the SRDF link speed decreases or the apply rate on the R2 side increases, more SRDF/A cycles can be queued the R1 side.

  Multi-cycle mode increases the robustness of the SRDF/A session and reduces spillover into the DSE storage pool.
- **Enginuity 5876**—If either array in the solution is running Enginuity 5876, SRDF/A operates in legacy mode. There are 2 cycles on the R1 side, and 2 cycles on the R2 side:
  - On the R1 side:
    - One Capture
    - One Transmit
  - On the R2 side:
    - One Receive
    - One Apply

  Each cycle switch moves the delta set to the next cycle in the process.

  A new capture cycle cannot start until the transmit cycle completes its commit of data from the R1 side to the R2 side.

  Cycle switching can occur as often as the preset Minimum Cycle Time, but it can also take longer since it is dependent on both the time it takes to transfer the data from the R1 transmit cycle to the R2 receive cycle and the time it takes to destage the R2 apply cycle.

Cycle switching in asynchronous mode

The number of capture cycles supported at the R1 side varies depending on whether one or both the arrays in the solution are running HYPERMAX OS.

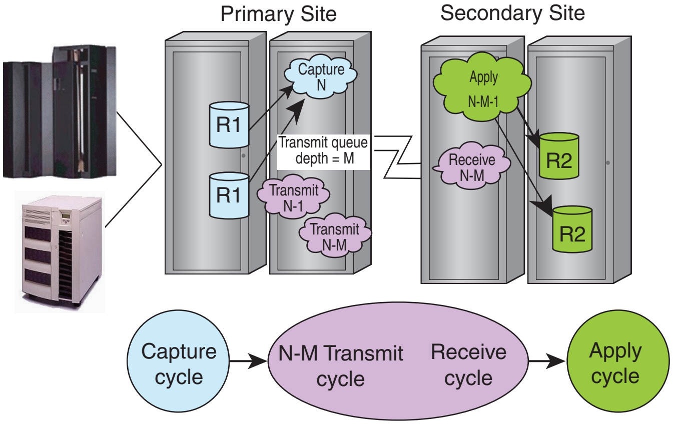
HYPERMAX OS

SRDF/A SSC sessions where both arrays are running HYPERMAX OS have one or more Transmit cycles on the R1 side (multi-cycle mode).

The following image shows multi cycle mode:

- Multiple cycles (one capture cycle and multiple transmit cycles) on the R1 side, and
- Two cycles (receive and apply) on the R2 side.

Figure 30 SRDF/A SSC cycle switching – multi-cycle mode



In multi-cycle mode, each cycle switch creates a new capture cycle (N) and the existing capture cycle (N-1) is added to the queue of cycles (N-1 through N-M cycles) to be transmitted to the R2 side by a separate commit action.

Only the data in the last transmit cycle (N-M) is transferred to the R2 side during a single commit.

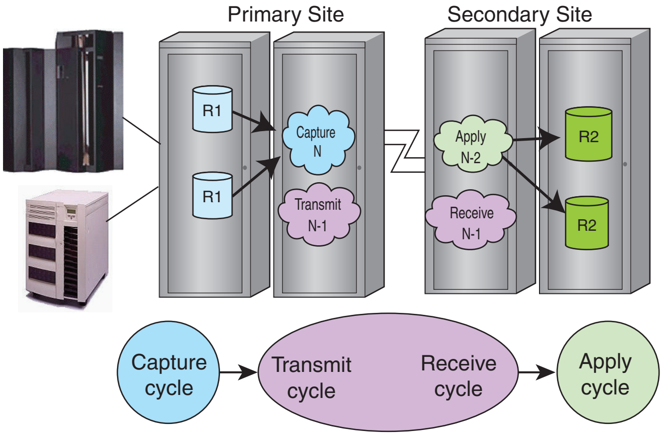
Enginuity 5773 through 5876

SRDF/A SSC sessions that include an array running Enginuity 5773 through 5876 have one Capture cycle and one Transmit cycle on the R1 side (legacy mode).

The following image shows legacy mode:

- 2 cycles (capture and transmit) on the R1 side, and
- 2 cycles (receive and apply) on the R2 side

Figure 31 SRDF/A SSC cycle switching – legacy mode



In legacy mode, the following conditions must be met before an SSC cycle switch can take place:

- The previous cycle's transmit delta set (N-1 copy of the data) must have completed transfer to the receive delta set on the secondary array.
- On the secondary array, the previous apply delta set (N-2 copy of the data) is written to cache, and data is marked write pending for the R2 devices.

SSC cycle switching in concurrent SRDF/A

In single session mode, cycle switching on both legs of the concurrent SRDF topology typically occurs at different times.

Data in the Capture and Transmit cycles may differ between the two SRDF/A sessions.

SRDF/A MSC cycle switching

SRDF/A MSC:

- Coordinates the cycle switching for all SRDF/A sessions in the SRDF/A MSC solution.
- Monitors for any failure to propagate data to the secondary array devices and drops all SRDF/A sessions together to maintain dependent-write consistency.
- Performs MSC cleanup operations (if possible).

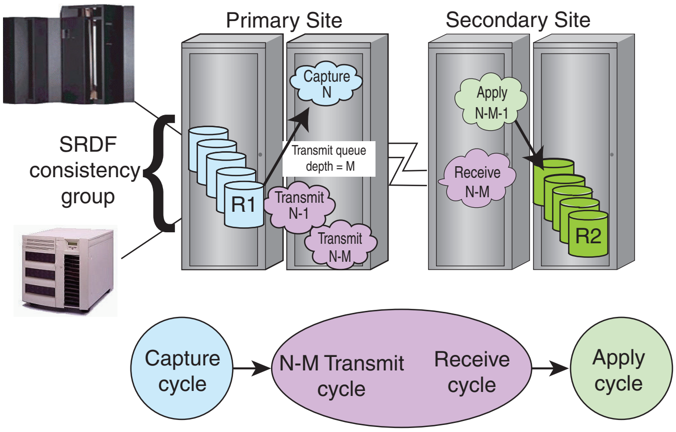
HYPERMAX OS
SRDF/A MSC sessions where both arrays are running HYPERMAX OS have two or more cycles on the R1 side (multi-cycle mode).

Note

If either the R1 side or R2 side of an SRDF/A session is running HYPERMAX OS, Solutions Enabler 8.x or later is required to monitor and manage MSC groups.

The following image shows the cycles on the R1 side (one capture cycle and multiple transmit cycles) and 2 cycles on the R2 side (receive and apply) for an SRDF/A MSC session when both of the arrays in the SRDF/A solution are running HYPERMAX OS.

Figure 32 SRDF/A MSC cycle switching – multi-cycle mode



SRDF cycle switches all SRDF/A sessions in the MSC group at the same time. All sessions in the MSC group have the same:

- Number of cycles outstanding on the R1 side
- Transmit queue depth (M)

In SRDF/A MSC sessions, Enginuity or HYPERMAX OS performs a coordinated cycle switch during a window of time when no host writes are being completed.

MSC temporarily suspends writes across all SRDF/A sessions to establish consistency.

Like SRDF/A cycle switching, the number of cycles on the R1 side varies depending on whether one or both the arrays in the solution are running HYPERMAX OS.

SRDF/A MSC sessions that include an array running Enginuity 5773 to 5876 have only two cycles on the R1 side (legacy mode).

In legacy mode, the following conditions must be met before an MSC cycle switch can take place:

- The primary array's transmit delta set must be empty.
- The secondary array's apply delta set must have completed. The N-2 data must be marked write pending for the R2 devices.

Write operations in cascaded SRDF

In cascaded configurations, R21 devices appear as:

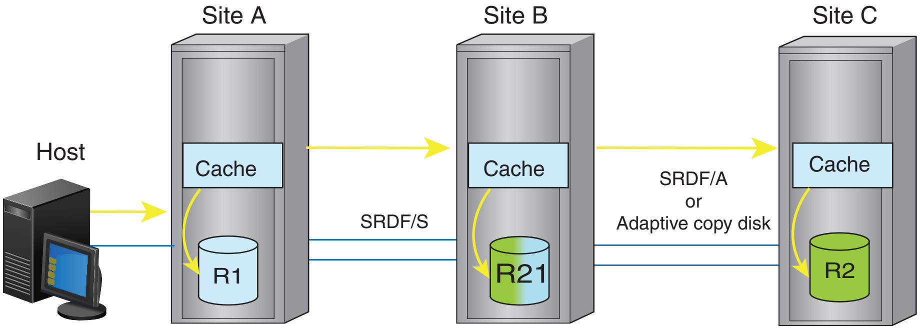
- R2 devices to hosts connected to R1 array
- R1 device to hosts connected to the R2 array

I/O to R21 devices includes:

- Synchronous I/O between the production site (R1)and the closest (R21) remote site.
- Asynchronous or adaptive copy I/O between the synchronous remote site (R21) and the tertiary (R2) site.
- You can Write Enable the R21 to a host so that the R21 behaves like an R2 device. This allows the R21 -> R2 connection to operate as R1 -> R2, while the R1 -> R21 connection is automatically suspended. The R21 begins tracking changes against the R1.

The following image shows the synchronous I/O flow in a cascaded SRDF topology.

Figure 33 Write commands to R21 devices



When a write command arrives to cache in Site B:

- The SRDF emulation at Site B sends a positive status back across the SRDF links to Site A (synchronous operations), and
- Creates a request for SRDF emulations at Site B to send data across the SRDF links to Site C.

SRDF/A cache management

Unbalanced SRDF/A configurations or I/O spikes can cause SRDF/A solutions to use large amounts of cache. Transient network outages can interrupt SRDF sessions. An application may write to the same record repeatedly. This section describes the SRDF/A features that address these common problems.

Tunable cache

You can set the SRDF/A maximum cache utilization threshold to a percentage of the system write pending limit for an individual SRDF/A session in single session mode and multiple SRDF/A sessions in single or MSC mode.

When the SRDF/A maximum cache utilization threshold or the system write pending limit is exceeded, the array exhausts its cache.

By default, the SRDF/A session drops if array cache is exhausted. You can keep the SRDF/A session running for a user-defined period. You can assign priorities to sessions, keeping SRDF/A active for as long as cache resources allow. If the condition is not resolved at the expiration of the user-defined period, the SRDF/A session still drops.

Use the features described below to prevent SRDF/A from exceeding its maximum cache utilization threshold.

SRDF/A cache data offloading

If the system approaches the maximum SRDF/A cache utilization threshold, DSE offloads some or all of the delta set data. DSE can be configured/enabled/disabled independently on the R1 and R2 sides.

Note

EMC recommends that DSE be configured the same on both sides.

DSE works in tandem with group-level write pacing to prevent cache over-utilization during spikes in I/O or network slowdowns.

Resources to support offloading vary depending on the version of Enginuity running on the array.

HYPERMAX OS

HYPERMAX OS offloads data into a Storage Resource Pool. One or more Storage Resource Pools are pre-configured before installation and used by a variety of functions. DSE can use a Storage Resource Pool pre-configured specifically for DSE, or if no such pool exists, DSE can use the default Storage Resource Pool. All SRDF groups on the array use the same Storage Resource Pool for DSE. DSE requests allocations from the Storage Resource Pool only when DSE is activated.

The Storage Resource Pool used by DSE is sized based on your SRDF/A cache requirements. DSE is automatically enabled.

Enginuity 5876

Enginuity 5876 offloads data to a DSE pool that you configure. You must configure a separate DSE pool for each device emulation type (FBA, IBM i, CKD3380 or CKD3390).

- In order to use DSE, each SRDF group must be explicitly associated with a DSE pool.
- By default, DSE is disabled.
- When TimeFinder/Snap sessions are used to replicate either R1 or R2 devices, you must create two separate preconfigured storage pools: DSE and Snap pools.

Mixed configurations: HYPERMAX OS and Enginuity 5876

If the array on one side of an SRDF device pair is running HYPERMAX OS and the other side is running a Enginuity 5876 or earlier, the SRDF/A session runs in Legacy mode.

- DSE is disabled by default on both arrays.
- EMC recommends that you enable DSE on both sides.

Transmit Idle

During short-term network interruptions, the transmit idle state describes that SRDF/A is still tracking changes but is unable to transmit data to the remote side.

Write folding

Write folding improves the efficiency of your SRDF links.

When multiple updates to the same location arrive in the same delta set, the SRDF emulations send the only most current data across the SRDF links.

Write folding decreases network bandwidth consumption and the number of I/Os processed by the SRDF emulations.

Write pacing

SRDF/A write pacing reduces the likelihood that an active SRDF/A session drops due to cache exhaustion. Write pacing dynamically paces the host I/O rate so it does not exceed the SRDF/A session's service rate, preventing cache overflow on both the R1 and R2 sides.

Use write pacing to maintain SRDF/A replication with reduced resources when replication is more important for the application than minimizing write response time.

You can apply write pacing at the group level, or at the device level for individual RDF device pairs that have TimeFinder/Snap or TimeFinder/Clone sessions off the R2 device.

Group-level pacing

SRDF/A group-level pacing paces host writes to match the SRDF/A session's link transfer rate. When host I/O rates spike, or slowdowns make transmit or apply cycle times longer, group-level pacing extends the host write I/O response time to match slower SRDF/A service rates.

When DSE is activated for an SRDF/A session, host-issued write I/Os are paced so their rate does not exceed the rate at which DSE can offload the SRDF/A session's cycle data to the DSE Storage Resource Pool.

Group-level pacing behavior varies depending on whether the maximum pacing delay is specified or not specified:

- If the maximum write pacing delay is not specified, SRDF adds up to 50 milliseconds to the host write I/O response time to match the speed of either the SRDF links or the apply operation on the R2 side, whichever is slower.
- If the maximum write pacing delay is specified, SRDF adds up to the user-specified maximum write pacing delay to keep the SRDF/A session running.

Group-level pacing balances the incoming host I/O rates with the SRDF link bandwidth and throughput capabilities when:

- The host I/O rate exceeds the SRDF link throughput.
- Some SRDF links that belong to the SRDF/A group are lost.
- Reduced throughput on the SRDF links.
- The write-pending level on an R2 device in an active SRDF/A session reaches the device write-pending limit.
- The apply cycle time on the R2 side is longer than 30 seconds and the R1 capture cycle time (or in MSC, the capture cycle target).

Group-level pacing can be activated by configurations or activities that result in slow R2 operations, such as:

- Slow R2 physical drives resulting in longer apply cycle times.
- Director sparing operations that slow restore operations.
- I/O to the R2 array that slows restore operations.

Note

On arrays running Enginuity 5876, if the space in the DSE pool runs low, DSE drops and group-level SRDF/A write pacing falls back to pacing host writes to match the SRDF/A session's link transfer rate.

Device-level (TimeFinder) pacing

HYPERMAX OS

SRDF/A device-level write pacing is not supported or required for asynchronous R2 devices in TimeFinder or TimeFinder SnapVX sessions if either array in the configuration is running HYPERMAX OS, including:

- R1 HYPERMAX OS - R2 HYPERMAX OS
- R1 HYPERMAX OS - R2 Enginuity 5876

- R1 Enginuity 5876 - R2 HYPERMAX OS

Enginuity 5773 to 5876
SRDF/A device-level pacing applies a write pacing delay for individual SRDF/A R1 devices whose R2 counterparts participate in TimeFinder copy sessions.

SRDF/A group-level pacing avoids high SRDF/A cache utilization levels when the R2 devices servicing both the SRDF/A and TimeFinder copy requests experience slowdowns.

Device-level pacing avoids high SRDF/A cache utilization when the R2 devices servicing both the SRDF/A and TimeFinder copy requests experience slowdowns.

Device-level pacing behavior varies depending on whether the maximum pacing delay is specified or not specified:

- If the maximum write pacing delay is not specified, SRDF adds up to 50 milliseconds to the overall host write response time to keep the SRDF/A session active.
- If the maximum write pacing delay is specified, SRDF adds up to the user-defined maximum write pacing delay to keep the SRDF/A session active.

Device-level pacing can be activated on the second hop (R21 -> R2) of a cascaded SRDF and cascaded SRDF/Star, topologies.

Device-level pacing may not take effect if all SRDF/A links are lost.

Write pacing and Transmit Idle

Host writes continue to be paced when:

- All SRDF links are lost, and
- Cache conditions require write pacing, and
- Transmit Idle is in effect.

Pacing during the outage is the same as the transfer rate prior to the outage.

SRDF read operations

Read operations from the R1 device do not usually involve SRDF emulations:

- For read “hits” (the production host issues a read to the R1 device, and the data is in local cache), the host emulation reads data from cache and sends it to the host.
- For read “misses” (the requested data is not in cache), the drive emulation reads the requested data from local drives to cache.

Refer to Read operations from R2 devices on page 128 for more information.

Read operations if R1 local copy fails

In SRDF/S, SRDF/A, and adaptive copy configurations, SRDF devices can process read I/Os that cannot be processed by regular logical devices. If the R1 local copy fails, the R1 device can still service the request as long as its SRDF state is Ready and the R2 device has good data.

SRDF emulations help service the host read requests when the R1 local copy is not available as follows:

- The SRDF emulations bring data from the R2 device to the host site.
- The host perceives this as an ordinary read from the R1 device, although the data was read from the R2 device acting as if it was a local copy.

HYPERMAX OS

Arrays running HYPERMAX OS cannot service SRDF/A read I/Os if DSE has been invoked to temporarily place some data on disk.

Read operations from R2 devices

Reading data from R2 devices directly from a host connected to the R2 is not recommended, because:

- SRDF/S relies on the application's ability to determine if the data image is the most current. The array at the R2 side may not yet know that data currently in transmission on the SRDF links has been sent.
- If the remote host reads data from the R2 device while a write I/O is in transmission on the SRDF links, the host will not be reading the most current data.

EMC strongly recommends that you allow the remote host to read data from the R2 devices while in Read Only mode only when:

- Related applications on the production host are stopped.
- The SRDF writes to the R2 devices are blocked due to a temporary suspension/ split of the SRDF relationship.

SRDF recovery operations

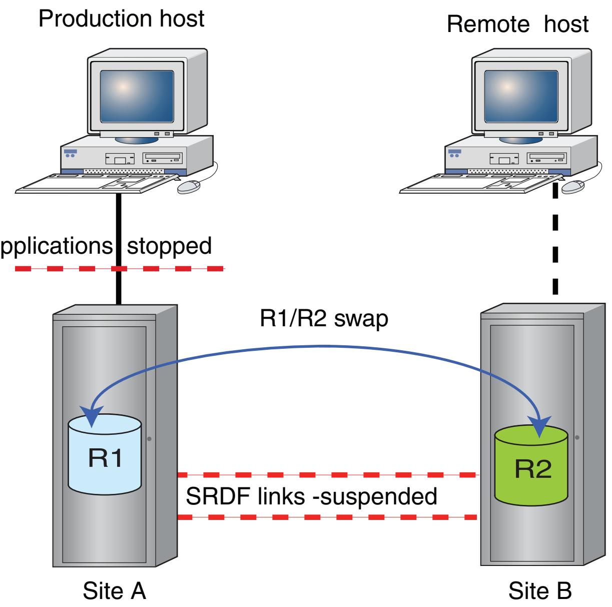
This section describes recovery operations in 2-site SRDF configurations.

Planned failover (SRDF/S)

A planned failover moves production applications from the primary site to the secondary site in order to test the recovery solution, upgrade or perform maintenance at the primary site.

The following image shows a 2-site SRDF configuration before the R1 <-> R2 personality swap:

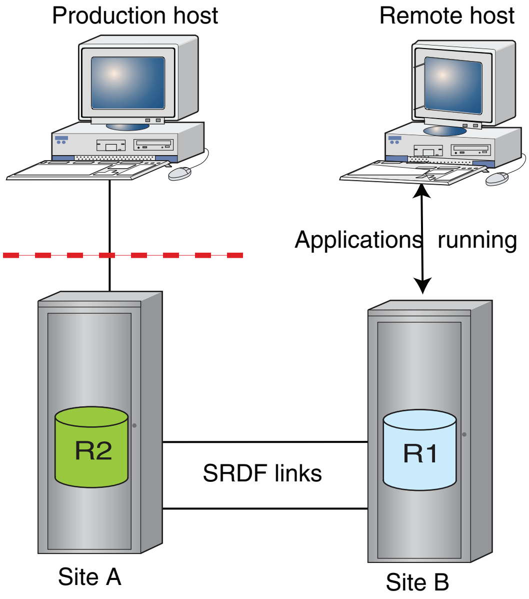
Figure 34 Planned failover: before personality swap



- Applications on the production host are stopped.
- SRDF links between Site A and Site B are suspended.
- If SRDF/CG is used, consistency is disabled.

The following image shows a 2-site SDRF configuration after the R1 <-> R2 personality swap.

Figure 35 Planned failover: after personality swap



When the maintenance, upgrades or testing procedures are complete, you can repeat the same procedure to return production to Site A.

Unplanned failover

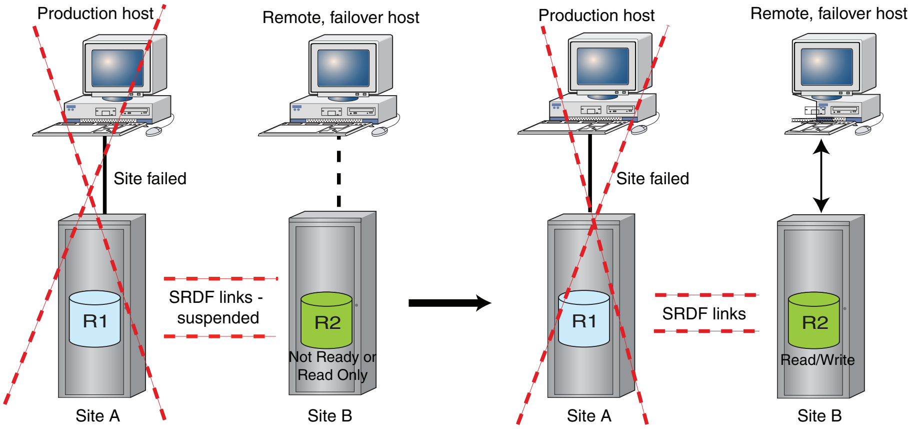
An unplanned failover moves production applications from the primary site to the secondary site after an unanticipated outage at the primary site, and the primary site is not available.

Failover to the secondary site in a simple configuration can be performed in minutes. You can resume production processing as soon as the applications are restarted on the failover host connected to Site B.

Unlike the planned failover operation, an unplanned failover resumes production at the secondary site, but without remote mirroring until Site A becomes operational and ready for a failback operation.

The following image shows failover to the secondary site after the primary site fails.

Figure 36 Failover to Site B, Site A and production host unavailable.



Failback to the primary array

After the primary host and array containing the primary (R1) devices are again operational, an SRDF failback allows production processing to resume on the primary host.

Recovery for a large number of invalid tracks

If the R2 devices have handled production processing for a long period of time, there may large numbers of invalid tracks owed to the R1 devices. SRDF control software can resynchronize the R1 and R2 devices while the secondary host continues production processing. Once there is a relatively small number of invalid tracks owed to the R1 devices, the failback process can be initiated.

Temporary link loss

In SRDF/A configurations, if a temporary loss (10 seconds or less) of all SRDF/A links occurs, the SRDF/A state remains active and data continues to accumulate in global memory. This may result in an elongated cycle, but the secondary array dependent-write consistency is not compromised and the primary and secondary array device relationships are not suspended.

Transmit Idle on page 125 can keep SRDF/A in an active state during all links lost conditions.

In SRDF/S configurations, if a temporary link loss occurs, writes are stalled (but not accumulated) in hopes that the SRDF link comes back up, at which point writes continue.

Reads are not affected.

Note

Switching to SRDF/S mode with the link limbo parameter configured for more than 10 seconds could result in an application, database, or host failure if SRDF is restarted in synchronous mode.

Permanent link loss (SRDF/A)

If all SRDF links are lost for more than link limbo or Transmit Idle can manage:

- All of the devices in the SRDF group are set to a Not Ready state.
- All data in capture and transmit delta sets is changed from write pending for the R1 SRDF mirror to invalid for the R1 SRDF mirror and is therefore owed to the R2 device.
- Any new write I/Os to the R1 device are also marked invalid for the R1 SRDF mirror.
  These tracks are owed to the secondary array once the links are restored.

When the links are restored, normal SRDF recovery procedures are followed

- Metadata representing the data owed is compared and merged based on normal host recovery procedures.
- Data is resynchronized by sending the owed tracks as part of the SRDF/A cycles.

Data on non-consistency exempt devices on the secondary array is always dependent-write consistent in SRDF/A active/consistent state, even when all SRDF links fail. Starting a resynchronization process compromises the dependent-write consistency until the resynchronization is fully complete and two cycle switches have occurred.

For this reason, it is important to use TimeFinder to create a gold copy of the dependent-write consistent image on the secondary array.

SRDF/A session cleanup (SRDF/A)

When an SRDF/A single session mode is dropped, SRDF:

- Marks new incoming writes at the primary array as being owed to the secondary array.
- Discards the capture and transmit delta sets, and marks the data as being owed to the secondary array. These tracks are sent to the secondary array once SRDF is resumed, as long as the copy direction remains primary-to-secondary.
- Marks and discards only the receive delta set at the secondary array, and marks the data is as tracks owed to the primary array.
- Marks and discards only the receive delta set at the secondary array, and marks the data is as tracks owed to the primary array.

Note

It is very important to capture a gold copy of the dependent-write consistent data on the secondary array R2 devices prior to any resynchronization. Any resynchronization compromises the dependent-write consistent image. The gold copy can be stored on a remote set of BCVs or Clones.

Failback from R2 devices (SRDF/A)

If a disaster occurs on the primary array, data on the R2 devices represents an older dependent-write consistent image and can be used to restart the applications.

After the primary array has been repaired, you can return production operations to the primary array by following procedures described in SRDF recovery operations on page 129.

If the failover to the secondary site is an extended event, the SRDF/A solution can be reversed by issuing a personality swap. SRDF/A can continue operations until a planned reversal of direction can be performed to restore the original SRDF/A primary and secondary relationship.

After the workload has been transferred back to the primary array hosts, SRDF/A can be activated to resume normal asynchronous mode protection.

Migration using SRDF/Data Mobility

Data migration is a one-time movement of data, typically of production data on an older array to a new array. Migration is distinct from replication in that once the data is moved, it is accessed only at the target.

You can migrate data between thick devices (also known as fully-provisioned or standard devices) and thin devices (also known as TDEVs). Once the data migration process is complete, the production environment is typically moved to the array to which the data was migrated.

Note

Before you begin, verify that your specific hardware models and Enginuity or HYPERMAX OS versions are supported for migrating data between different platforms.

In open systems host environments, use Solutions Enabler to reduce migration resynchronization times while replacing either the R1 or R2 devices in an SRDF 2-site topology.

When you connect between arrays running different versions, limitations may apply. For example, migration operations require the creation of temporary SRDF groups. Older versions of the operating environment support fewer SRDF groups. You must verify that the older array has sufficient unused groups to support the planned migration.

Migrating data with concurrent SRDF

In concurrent SRDF topologies, you can non-disruptively migrate data between arrays along one SRDF leg while remote mirroring for protection along the other leg.

Once the migration process completes, the concurrent SRDF topology is removed, resulting in a 2-site SRDF topology.

Replacing R2 devices with new R2 devices

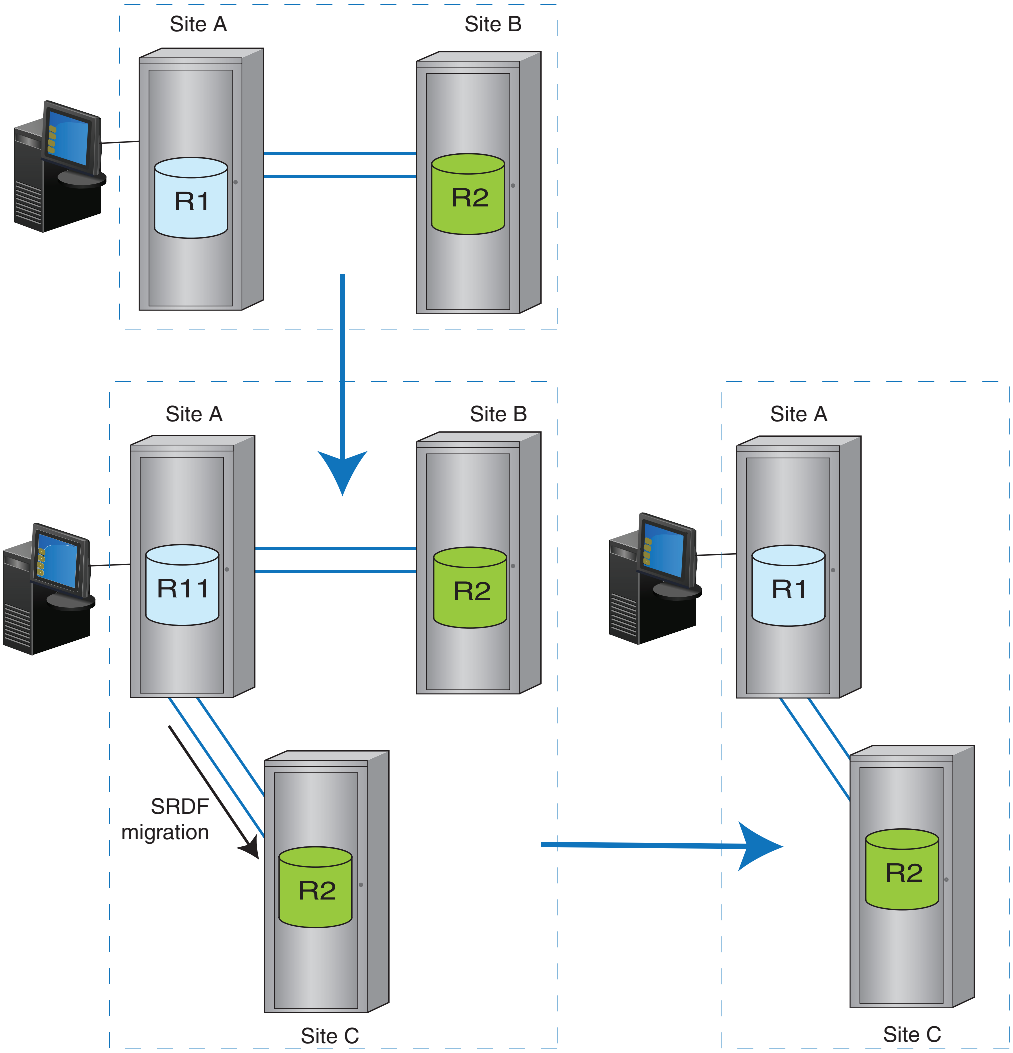
You can manually migrate data as shown in the following image, including:

- Initial 2-site topology
- The interim 3-site migration topology
- Final 2-site topology

After migration, the original primary array is mirrored to a new secondary array.

EMC support personnel are available to assist with the planning and execution of your migration projects.

Figure 37 Migrating data and removing the original secondary array (R2)



Replacing R1 devices with new R1 devices

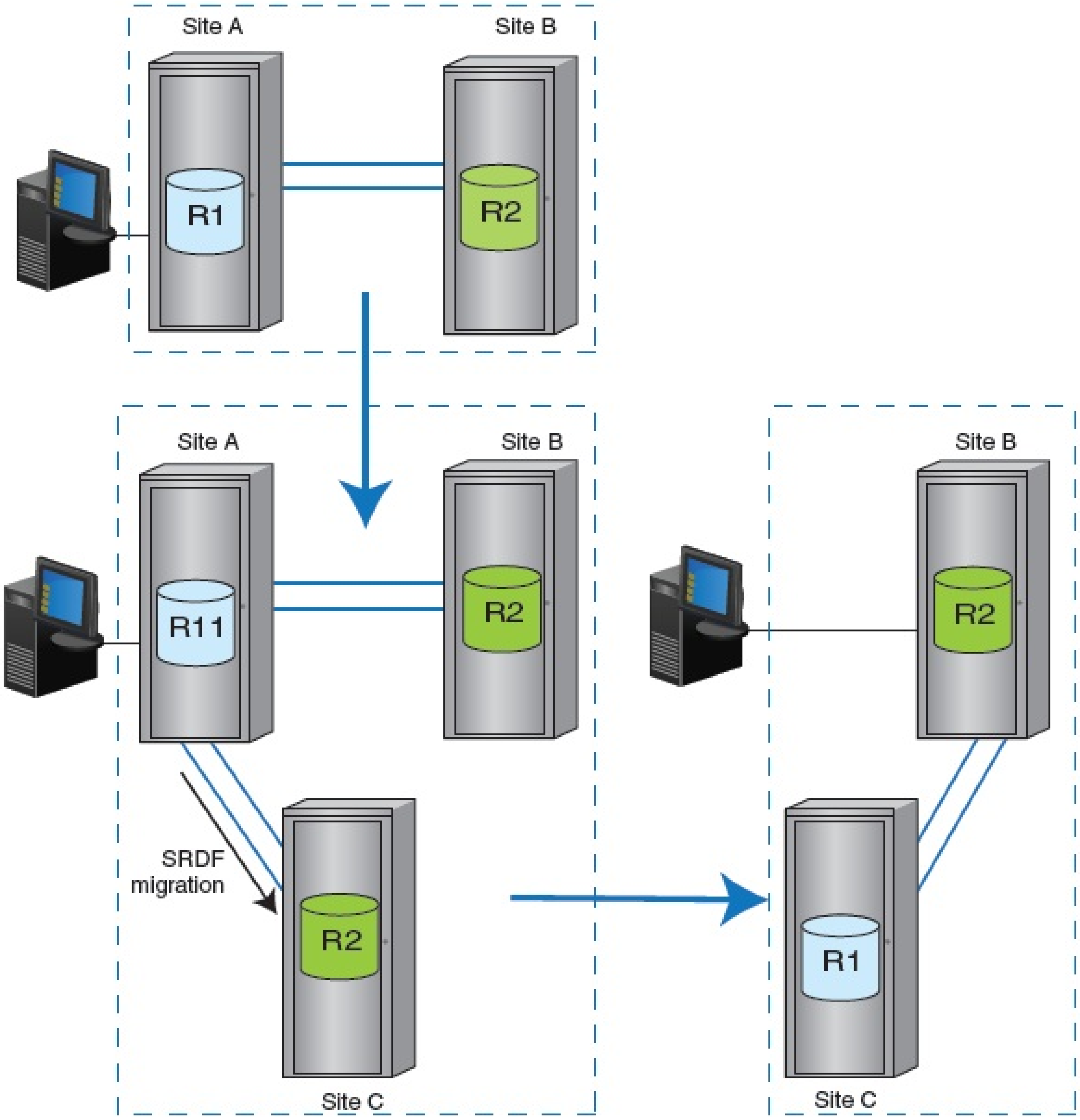
The following image shows replacing the original R1 devices with new R1 devices, including:

- Initial 2-site topology
- The interim 3-site migration topology
- Final 2-site topology

After migration, the new primary array is mirrored to the original secondary array.

EMC support personnel are available to assist with the planning and execution of your migration projects.

Figure 38 Migrating data and replacing the original primary array (R1)



Replacing R1 and R2 devices with new R1 and R2 devices

You can use the combination of concurrent SRDF and cascaded SRDF to replace both R1 and R2 devices at the same time.

Note

Before you begin, verify that your specific hardware models and Enginuity or HYPERMAX OS versions are supported for migrating data between different platforms.

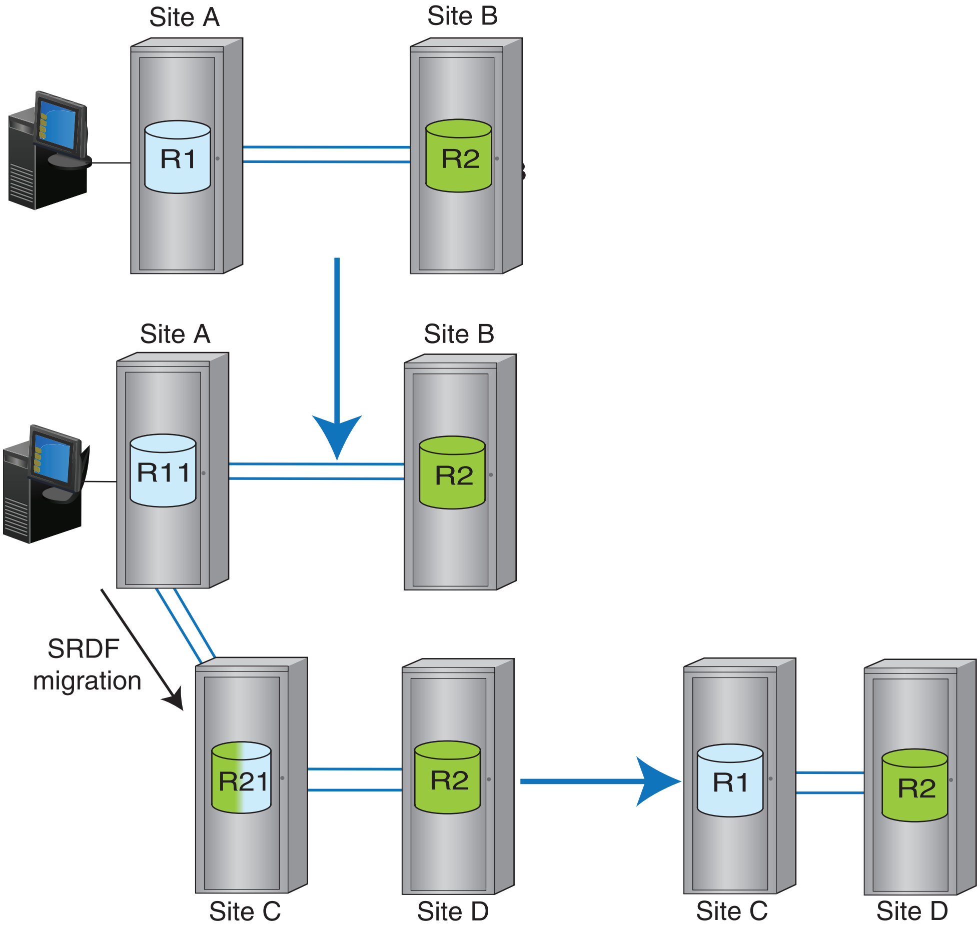
The following image shows an example of replacing both R1 and R2 devices with new R1 and R2 devices at the same time, including:

- Initial 2-site topology

- Migration process
- The final topology

EMC support personnel is available to assist with the planning and execution of your migration projects.

Figure 39 Migrating data and replacing the original primary (R1) and secondary (R2) arrays



Migration-only SRDF

In some of the cases, you can migrate your data with full SRDF functionality, including disaster recovery and other advanced SRDF features.

In cases where full SRDF functionality is not available, you can move your data across the SRDF links using migration-only SRDF.

The following table lists SRDF common operations and features and whether they are supported in SRDF groups during SRDF migration-only environments.

Table 39 Limitations of the migration-only mode

SRDF operations or features	Whether supported during migration
R2 to R1 copy	Only for device rebuild from un-rebuildable RAID group failures.

Table 39 Limitations of the migration-only mode (continued)

SRDF operations or features	**Whether supported during migration**
Failover, failback, domino	Not supported
SRDF/Star	Not supported
SRDF/A features: (DSE, Consistency Group, ECA, MSC)	Not supported
Dynamic SRDF operations: (Create/delete/move SRDF pairs, R1/R2 personality swap)	Not supported
TimeFinder operations	Only on R1
Online configuration change or upgrade	• If online upgrade or configuration changes affect the group or devices being migrated, migration must be suspended prior to the upgrade or configuration changes. • If the changes do not affect the migration group, they are allowed without suspending migration.
Out-of-family Non-Disruptive Upgrade (NDU)	Not supported

SRDF/Metro

In traditional SRDF, R1 devices are Read/Write accessible. R2 devices are Read Only/ Write Disabled.

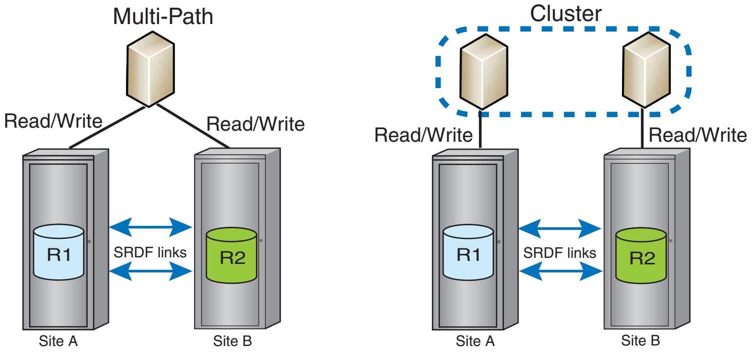
In SRDF/Metro configurations:

- R2 devices are Read/Write accessible to hosts.
- Hosts can write to both the R1 and R2 side of the device pair.
- R2 devices assume the same external device identity (geometry, device WWN) as their R1.

This shared identity causes the R1 and R2 devices to appear to hosts(s) as a single virtual device across the two arrays.

SRDF/Metro can be deployed with either a single multi-pathed host or with a clustered host environment.

Figure 40 SRDF/Metro



Hosts can read and write to both the R1 and R2 devices.

For single host configurations, host I/Os are issued by a single host. Multi-pathing software directs parallel reads and writes to each array.

For clustered host configurations, host I/Os can be issued by multiple hosts accessing both sides of the SRDF device pair. Each cluster node has dedicated access to an individual storage array.

In both single host and clustered configurations, writes to the R1 or R2 devices are synchronously copied to the paired device. Write conflicts are resolved by the SRDF/Metro software to maintain consistent images on the SRDF device pairs. The R1 device and its paired R2 device appear to the host as a single virtualized device.

SRDF/Metro is managed using either Solutions Enabler 8.1 or higher or Unisphere for VMAX 8.1 or higher.

SRDF/Metro requires a license on both arrays.

Storage arrays running HYPERMAX OS can simultaneously support SRDF groups configured for SRDF/Metro operations and SRDF groups configured for traditional SRDF operations.

Key differences SRDF/Metro

- In SRDF/Metro configurations:
  - R2 device is Read/Write accessible to the host.
  - Host(s) can write to both R1 and R2 devices.
  - Both sides of the SRDF device pair appear to the host(s) as the same device.
  - The R2 device assumes the personality of the primary R1 device (geometry, device WWN, etc.).
  - Two additional RDF pair states:
    - ActiveActive for configurations using the Witness options (Array and Virtual)
    - ActiveBias for configurations using bias

  Note

  R1 and R2 devices should not be presented to the cluster until they reach one of these 2 states and present the same WWN.

- All device pairs in an SRDF/Metro group are managed together for all supported operations, with the following exceptions:

    - If all the SRDF device pairs are Not Ready (NR) on the link, createpair operations can add devices to the group if the new device pairs are created Not Ready (NR) on the link.
    - If all the SRDF device pairs are Not Ready (NR) on the link, deletepair operations can delete a subset of the SRDF devices in the SRDF group.
- In the event of link or other failures, SRDF/Metro provides the following methods for determining which side of a device pair remains accessible to the host:
    - **Bias option**: Device pairs for SRDF/Metro are created with a new attribute – `use_bias`. By default, the createpair operation sets the bias to the R1 side of the pair. That is, if the device pair becomes Not Ready (NR) on the RDF link, the R1 (bias side) remains accessible to the host(s), and the R2 (non-bias side) is inaccessible to the host(s).
      When all RDF device pairs in the RDF group have reached the ActiveActive or ActiveBias pair state, bias can be changed (so the R2 side of the device pair remains accessible to the host). Bias on page 141 provides more information.
    - **Witness option**: A designated Witness monitors SRDF on each array and the SRDF links between them. In the event of a failure, the Witness can determine the nature of the failure, and arbitrate which side of the device pair becomes the non-bias side (inaccessible to hosts) and which side becomes the bias side (remains accessible to hosts). The Witness method allows for intelligently choosing on which side to continue operations when the bias-only method may not result in continued host availability to a surviving non-biased array.
      The Witness option is the default.

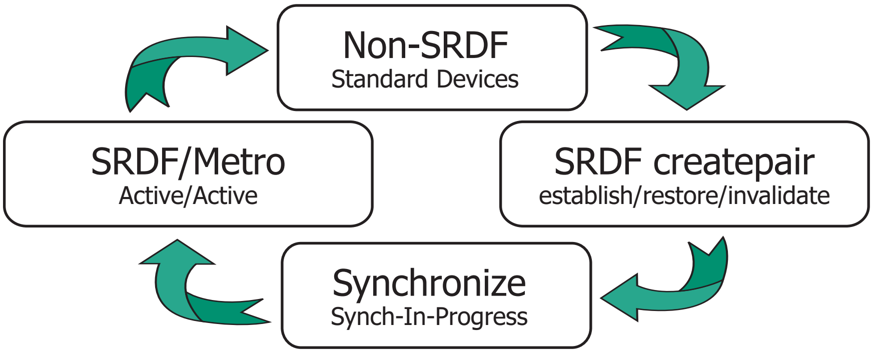
      SRDF/Metro provides two types of Witnesses, Array and Virtual:

        - **Witness Array** : HYPERMAX OS or Enginuity on a third array monitors SRDF/Metro, determines the type of failure, and uses the information to choose one side of the device pair to remain R/W accessible to the host. The Witness option requires two SRDF groups: one between the R1 array and the Witness array and the other between the R2 array and the Witness array.
          Array Witness on page 142 provides more information. The component on the Array Witness .
        - **Virtual Witness option**: Introduced with HYPERMAX OS 5977 Q3 2016 SR, vWitness provides the same functionality as the Witness Array option, only it is packaged to run in a virtual appliance, not on the array.
          Virtual Witness (vWitness) provides more information.

SRDF/Metro life cycle

The life cycle of an SRDF/Metro configuration begins and ends with an empty SRDF group and a set of non-SRDF devices, as shown in the following image.

Figure 41 SRDF/Metro life cycle



The life cycle of an SRDF/Metro configuration includes the following steps and states:

- **Create device pairs** in an empty SRDF group.
  Create pairs using the new -rdf_metro option to indicate that the new SRDF pairs will operate in an SRDF/Metro configuration.

  If all the SRDF device pairs are Not Ready (NR) on the link, the createpair operation can be used to add more devices into the SRDF group.
- **Make the device pairs Read/Write** (RW) on the SRDF link.
  Use the -establish or the -restore options to make the devices Read/Write (RW) on the SRDF link.

  Alternatively, use the -invalidate option to create the devices without making them Read/Write (RW) on the SRDF link.
- **Synchronize** the device pairs.
  When the devices in the SRDF group are Read/Write (RW) on the SRDF link, invalid tracks begin synchronizing between the R1 and R2 devices.

  Direction of synchronization is controlled by either an establish or a restore operation.
- **Activate SRDF/Metro**
  Device pairs transition to the ActiveActive pair state when:

  - Device federated personality and other information is copied from the R1 side to the R2 side.
  - Using the information copied from the R1 side, the R2 side sets its identify as an SRDF/Metro R2 when queried by host I/O drivers.
  - R2 devices become accessible to the host(s).

  When all SRDF device pairs in the group transition to the ActiveActive state,host(s) can discover the R2 devices with federated personality of R1 devices. SRDF/Metro manages the SRDF device pairs in the SRDF group. A write to either side of the SRDF device pair completes to the host only after it is transmitted to the other side of the SRDF device pair, and the other side has acknowledged its receipt.
- **Add/remove devices** to/from an SRDF/Metro group.
  The group must be in either Suspended or Partitioned state to add or remove devices.

  Use the deletepair operation to delete all or a subset of device pairs from the SRDF group. Removed devices return to the non-SRDF state.

  Use the createpair operation to add additional device pairs to the SRDF group.

Use the removepair and movepair operations to remove/move device pairs.

If all device pairs are removed from the group, the group is no longer controlled by SRDF/Metro. The group can be re-used either as a SRDF/Metro or non-Metro group.

- **Deactivate SRDF/Metro**
  If all devices in an SRDF/Metro group are deleted, that group is no longer be part of an SRDF/Metro configuration.

  You can use the createpair operation to re-populate the RDF group, either for SRDF/Metro or for non-Metro.

SRDF/Metro resiliency

If a SRDF/Metro device pair becomes Not Ready (NR) on the SRDF link, SRDF/Metro must respond by choosing one side of the device pair to remain accessible to hosts, while making the other side of the device pair inaccessible. This response to lost connectivity between the two sides of a device pair in an SRDF/Metro configuration is called *bias*.

Initially, the R1 side specified by the createpair operation is the *bias side*. That is, if the device pair becomes NR, the R1 (bias side) side remains accessible (RW) to hosts, and the R2 (*non-bias side*) is made inaccessible (NR) to hosts. Bias can be changed once all the device pairs in the SRDF/Metro group have reached the ActiveActive pair state. The bias side is represented as R1 and the non-bias side is represented as R2.

- During the createpair operation, bias defaults to the R1 device. After device creation, bias side can be changed from the default (R1) to the R2 side
- The initial bias device will be exported as the R1 in all external displays and commands.
- The initial non-bias device will be exported as the R2 in all external displays and commands.
- Changing the bias changes the SRDF personalities of the two sides of the SRDF device pair.

The following sections explain the methods SRDF/Metro provides for determining which side of a device pair is the winner in case of a replication failure.

Bias

In an SRDF/Metro configuration, HYPERMAX OS uses the link between the two sides of each device pair to ensure consistency of the data on the two sides. If the device pair becomes Not Ready (NR) on the RDF link, HYPERMAX chooses the bias side of the device pair to remain accessible to the hosts, while making the non-bias side of the device pair inaccessible. This prevents data inconsistencies between the two sides of the RDF device pair.

Note

Bias applies only to RDF device pairs in an SRDF/Metro configuration.

When adding device pairs to an SRDF/Metro group (createpair operation), HYPERMAX configures the R1 side of the pair as the bias side.

For example, in Solutions Enabler, use the `-use_bias` option to specify that the R1 side of devices are the bias side when the Witness options are not used. For example, to create SRDF/Metro device pairs and make them RW on the link without a Witness array:

```
symrdf -f /tmp/device_file -sid 085 -rdfg 86 establish -use_bias
```

If the Witness options are not used, the `establish` and `restore` commands also require the `use_bias` option.

When the SRDF/Metro devices pairs are configured to use bias, their pair state is ActiveBias.

Bias can be changed when all device pairs in the SRDF/Metro group have reached the ActiveActive or ActiveBias pair state.

Array Witness

When using the Array Witness method, SRDF/Metro uses a third "witness" array to determine the bias side. The witness array runs one of the following operating environments:

- Enginuity 5876 with ePack containing fixes to support SRDF N-x connectivity
- HYPERMAX OS 5977.810.784 with ePack containing fixes to support SRDF N-x connectivity
- HYPERMAX OS 5977.945.890 or later

In the event of a failure, the witness decides which side of the Metro group remains accessible to hosts, giving preference to the bias side. The Array Witness method allows for choosing on which side to continue operations when the Device Bias method may not result in continued host availability to a surviving non-biased array.

The Array Witness must have SRDF connectivity to both the R1-side array and R2-side array.

SRDF remote adapters (RA's) are required on the witness array with applicable network connectivity to both the R1 side and R2 side arrays.

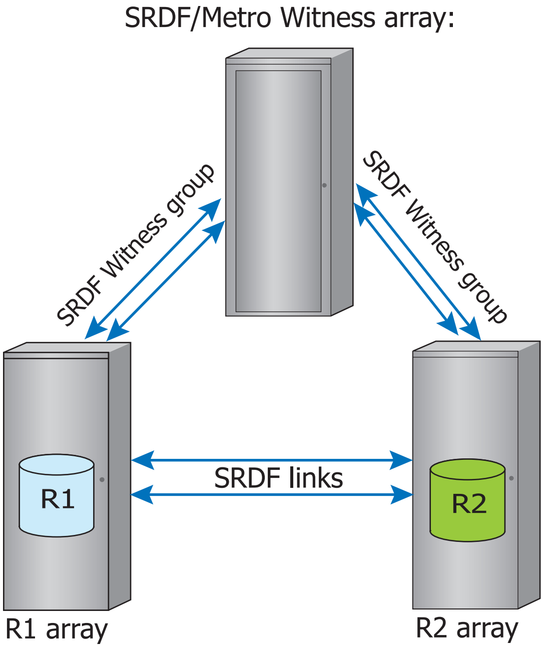
For complete redundancy, there can be multiple witness arrays. If the auto configuration process fails and no other applicable witness arrays are available, SRDF/Metro uses the Device Bias method.

The Array Witness method requires 2 SRDF groups; one between the R1 array and the witness array, and a second between the R2 array and the witness array:

Note

A Witness group is not allowed to contain devices.

Figure 42 SRDF/Metro Array Witness and groups



Solutions Enabler checks that the Witness groups exist and are online when carrying out establish or restore operations. SRDF/Metro determines which witness array an SRDF/Metro group is using, so there is no need to specify the Witness. Indeed, there is no ability to specify the Witness.

When the witness array is connected to both the SRDF/Metro paired arrays, the configuration enters Witness Protected state.

When the Array Witness method is in operation, the state of the device pairs is ActiveActive.

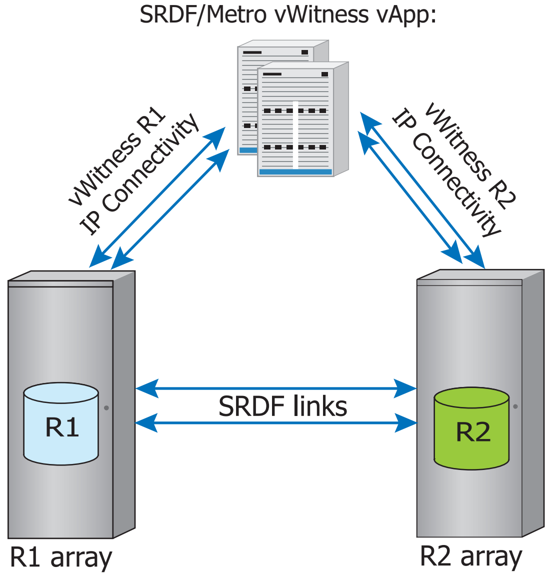
If the witness array becomes inaccessible from both the R1 and R2 arrays, HYPERMAX OS sets the R1 side as the bias side, the R2 side as the non-bias side, and the state of the device pairs becomes ActiveBias.

Virtual Witness (vWitness)

Virtual Witness (vWitness) is an additional resiliency option introduced in HYPERMAX OS 5977.945.890 and Solutions Enabler or Unisphere for VMAX V8.3. vWitness has similar capabilities to the Array Witness method, except that it is packaged to run in a virtual appliance (vApp) on a VMware ESX server, not on an array.

The vWitness and Array Witness options are treated the same in the operating environment, and can be deployed independently or simultaneously. When deployed simultaneously, SRDF/Metro favors the Array Witness option over the vWitness option, as the Array Witness option has better availability. For redundancy, you can configure up to 32 vWitnesses.

Figure 43 SRDF/Metro vWitness vApp and connections



The management guests on the R1 and R2 SRDF/Metro managed arrays maintain multiple IP connections to redundant vWitness virtual appliances. The IP connections use TLS/SSL to ensure secure connectivity between vWitness instances and the arrays.

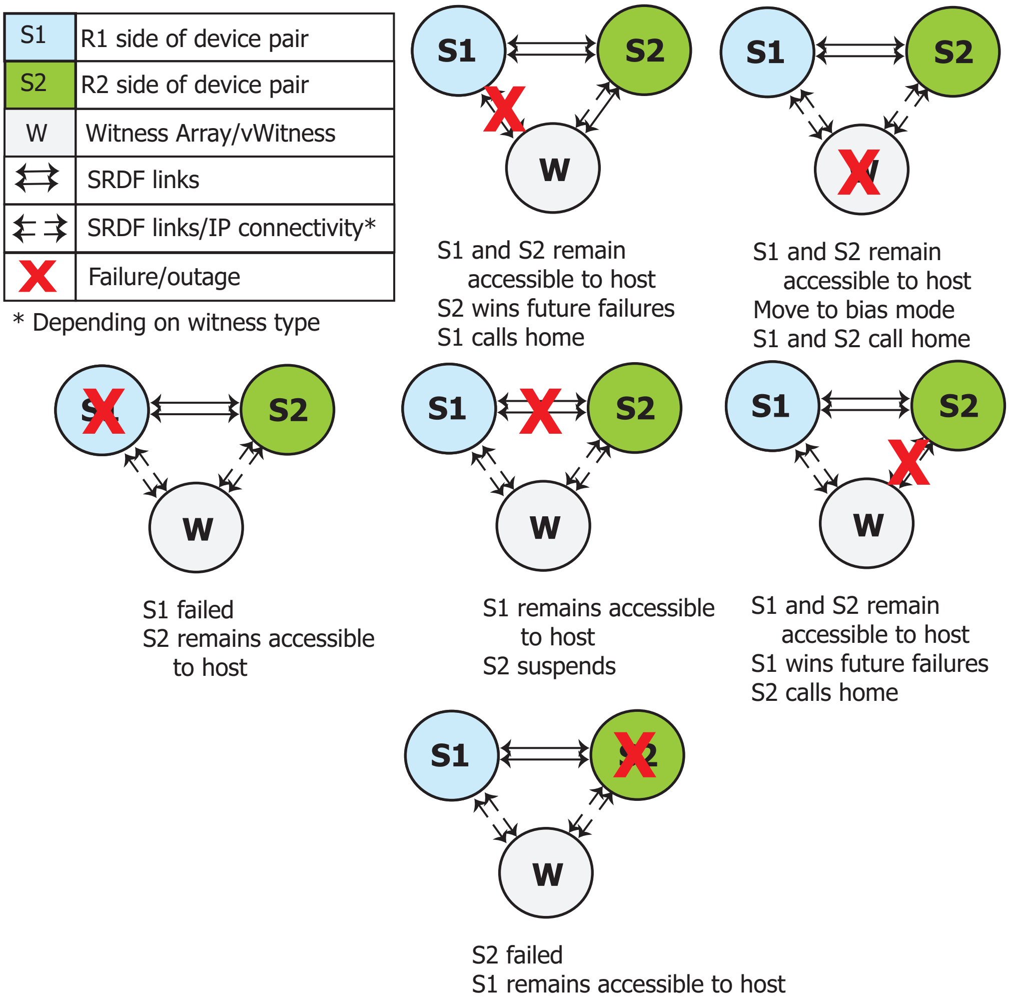
Once you have established IP connectivity to the arrays, you can use the Solutions Enabler or Unisphere for VMAX to perform the following:

- Add a new vWitness to the configuration. This will not affect any existing vWitnesses. Once the vWitness is added, it is enabled for participation in the vWitness infrastructure.
- Query the state of a vWitness configuration.
- Suspend a vWitness. If the vWitness is currently servicing an SRDF/Metro session, this operation requires a force flag. This puts the SRDF/Metro session in an unprotected state until it renegotiates with another witness, if available.
- Remove a vWitness from the configuration. Once removed, SRDF/Metro will break the connection with vWitness. You can only remove vWitnesses that are not currently servicing active SRDF/Metro sessions.

Witness failure scenarios

This section depicts various single and multiple failure behaviors for SRDF/Metro when the Witness option (Array or vWitness) is used.

Figure 44 SRDF/Metro Witness single failure scenarios



S1 and S2 remain accessible to host
S2 wins future failures
S1 calls home

S1 and S2 remain accessible to host
Move to bias mode
S1 and S2 call home

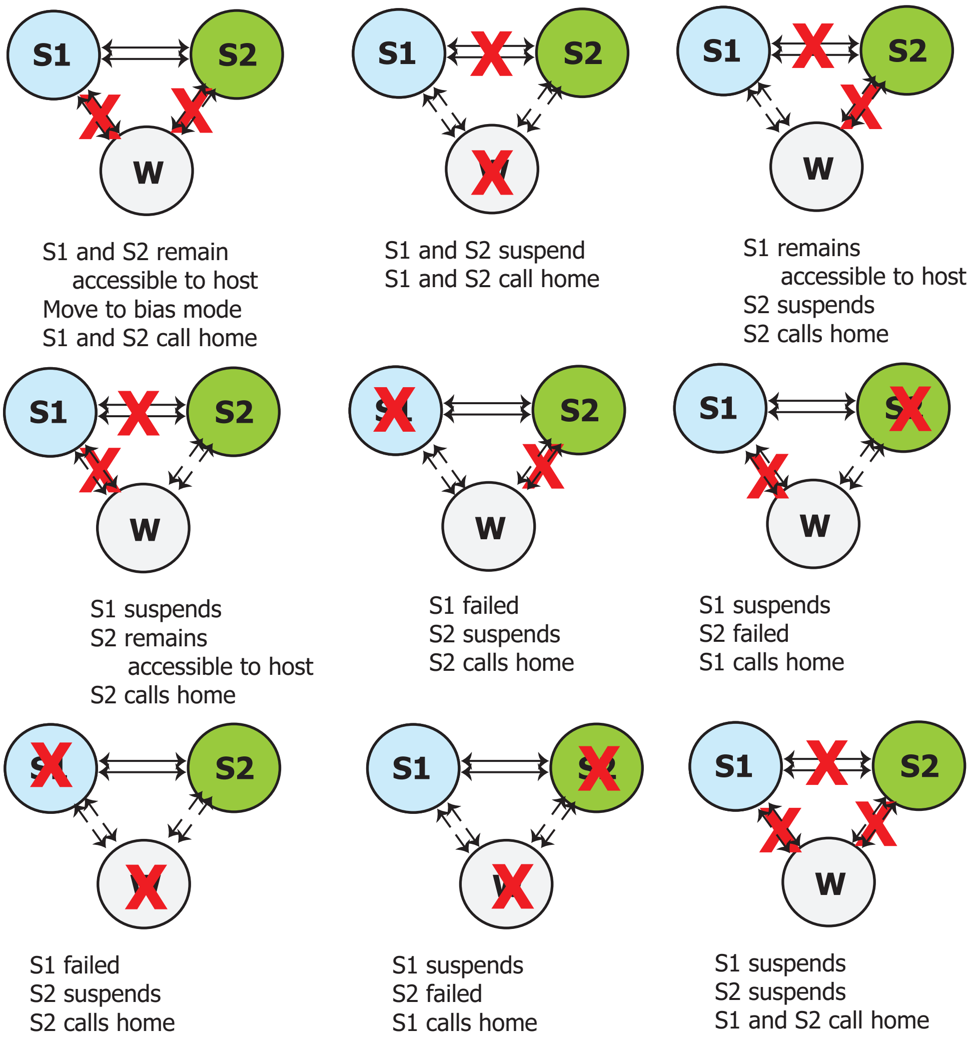
S1 failed
S2 remains accessible to host

S1 remains accessible to host
S2 suspends

S1 and S2 remain accessible to host
S1 wins future failures
S2 calls home

S2 failed
S1 remains accessible to host

Figure 45 SRDF/Metro Witness multiple failure scenarios



S1 and S2 remain accessible to host
Move to bias mode
S1 and S2 call home

S1 and S2 suspend
S1 and S2 call home

S1 remains accessible to host
S2 suspends
S2 calls home

S1 suspends
S2 remains accessible to host
S2 calls home

S1 failed
S2 suspends
S2 calls home

S1 suspends
S2 failed
S1 calls home

S1 failed
S2 suspends
S2 calls home

S1 suspends
S2 failed
S1 calls home

S1 suspends
S2 suspends
S1 and S2 call home

Deactivate SRDF/Metro

To terminate a SRDF/Metro configuration, simply remove all the device pairs (deletepair) in the SRDF group.

Note

The devices must be in Suspended state in order to perform the deletepair operation.

When all the devices in the SRDF/Metro group have been deleted, the group is no longer part of an SRDF/Metro configuration.

NOTICE

The deletepair operation can be used to remove a subset of device pairs from the group. The SRDF/Metro configuration terminates only when the last pair is removed.

Delete one side of a SRDF/Metro configuration

To remove devices from only one side of a SRDF/Metro configuration, use the half_deletepair operation to terminate the SRDF/Metro configuration at one side of the SRDF group.

The half_deletepair operation may be performed on all or on a subset of the SRDF devices on one side of the SRDF group.

Note

The devices must be in Suspended or Partitioned SRDF pair state to perform the half_deletepair operation.

After the half_deletepair operation:

- The devices on the side where the half-deletepair operation was performed are no longer SRDF devices.
- The devices at the other side of the SRDF group retain their configuration as SRDF/Metro

If all devices are deleted from one side of the SRDF group, that side of the SRDF group is no longer part of the SRDF/Metro configuration.

Restore native personality to a federated device

Devices in SRDF/Metro configurations have federated personalities. When a device is removed from an SRDF/Metro configuration, the device personality can be restored to its original native personality.

The following restrictions apply to restoring the native personality of a device which has federated personality as a result of a participating in a SRDF/Metro configuration:

- Requires HYPERMAX OS 5977.691.684 or higher.
- The device must be unmapped and unmasked.
- The device must have a federated WWN.
- The device must not be an SRDF device.
- The device must not be a ProtectPoint device.

SRDF/Metro restrictions

The following restrictions and dependencies apply to SRDF/Metro configurations:

- Both the R1 and R2 side must be running HYPERMAX OS 5977.691.684 or greater.
- Only non-SRDF devices can become part of an SRDF/Metro configuration.
- The R1 and R2 must be identical in size.
- Devices cannot have Geometry Compatibility Mode (GCM) or User Geometry set.
- Online device expansion is not supported.
- createpair -establish, establish, restore, and suspend operations apply to all devices in the SRDF group.
- Control of devices in an SRDF group which contains a mixture of R1s and R2s is not supported.

Interaction restrictions

The following restrictions apply to SRDF device pairs in an SRDF/Metro configuration with TimeFinder and Open Replicator (ORS):

- Open Replicator is not supported.
- Devices cannot be BCVs.
- Devices cannot be used as the target of the data copy when the SRDF devices are RW on the SRDF link with either a SyncInProg or ActiveActive SRDF pair state.
- A snapshot does not support restores or re-links to itself.

RecoverPoint

HYPERMAX OS 5977 Q2 2017 introduces support for RecoverPoint on VMAX storage arrays. RecoverPoint is a comprehensive data protection solution designed to provide production data integrity at local and remote sites. RecoverPoint also provides the ability to recover data from a point in time using journaling technology.

The primary reasons for using RecoverPoint are:

- Remote replication to heterogeneous arrays
- Local and remote data corruption protection
- Disaster recovery
- Secondary device repurposing
- Data migrations

RecoverPoint systems support local and remote replication of data that applications are writing to SAN-attached storage. The systems use existing Fibre Channel infrastructure to integrate seamlessly with existing host applications and data storage subsystems. For remote replication, the systems use existing Fibre Channel connections to send the replicated data over a WAN, or use Fibre Channel infrastructure to replicate data aysnchronously. The systems provide failover of operations to a secondary site in the event of a disaster at the primary site.

Previous implementations of RecoverPoint relied on a splitter to track changes made to protected volumes. This implementation relies on a cluster of RecoverPoint nodes, provisioned with one or more RecoverPoint storage groups, leveraging SnapVX technology, on the VMAX array. Volumes in the RecoverPoint storage groups are visible to all the nodes in the cluster, and available for replication to other storage arrays.

Recoverpoint allows data replication for up to 8,000 LUNs per RecoverPoint cluster and up to eight different RecoverPoint clusters attached to one array. Supported array types include VMAX All Flash, VMAX3, VMAX, VNX, VPLEX, and XtremIO.

RecoverPoint is licensed and sold separately.

Remote replication using eNAS

File Auto Recovery (FAR) allows you to manually failover or move a virtual Data Mover (VDM) from a source eNAS system to a destination eNAS system. The failover or move leverages block-level Symmetrix Remote Data Facility (SRDF) synchronous replication, so it invokes zero data loss in the event of an unplanned operation. This feature consolidates VDMs, file systems, file system checkpoint schedules, CIFS servers, networking, and VDM configurations into their own separate pools. This feature works for a recovery where the source is unavailable. For recovery support in the event of an unplanned failover, an option is provided to recover and clean up the source system and make it ready as a future destination

The manually initiated failover and reverse operations can be performed using EMC File Auto Recovery Manager (FARM). FARM allows you to automatically failover a selected sync-replicated VDM on a source eNAS system to a destination eNAS system. FARM also allows you to monitor sync-replicated VDMs and to trigger automatic failover based on Data Mover, File System, Control Station, or IP network unavailability that would cause the NAS client to lose access to data.

CHAPTER 8

Blended local and remote replication

This chapter describes TimeFinder integration with SRDF.

SRDF and TimeFinder

TimeFinder is a local replication solution that non-disruptively creates point-in-time copies of critical data. You can configure backup sessions, initiate copies, and terminate TimeFinder operations using host-based TimeFinder software.

TimeFinder is tightly integrated with SRDF solutions. You can use TimeFinder and SRDF products to complement each other when you require both local and remote replication. For example, you can use TimeFinder to create local gold copies of SRDF devices for recovery operations and for testing disaster recovery solutions.

The key benefits of TimeFinder integration with SRDF include:

- Remote controls simplify automation—Use EMC host-based control software to transfer commands across the SRDF links. A single command from the host to the primary array can initiate TimeFinder operations on both the primary and secondary arrays.
- Consistent data images across multiple devices and arrays—SRDF/CG guarantees that a dependent-write consistent image of production data on the R1 devices is replicated across the SRDF links.

You can use TimeFinder/CG in an SRDF configuration to create dependent-write consistent local and remote images of production data across multiple devices and arrays.

Note

The SRDF/A single session solution guarantees dependent-write consistency across the SRDF links and does not require SRDF/CG. SRDF/A MSC mode requires host software to manage consistency among multiple sessions.

Note

Some TimeFinder operations are not supported on devices protected by SRDF. For more information, refer to the *Solutions Enabler SnapVX Product Guide*.

R1 and R2 devices in TimeFinder operations

You can use TimeFinder to create local replicas of R1 and R2 devices. The following rules apply:

- You can use R1 devices and R2 devices as TimeFinder source devices.
- R1 devices can be the target of TimeFinder operations as long as there is no host accessing the R1 during the operation.
- R2 devices can be used as TimeFinder target devices if SRDF replication is not active (writing to the R2 device). To use R2 devices as TimeFinder target devices, you must first suspend the SRDF replication session.

SRDF/AR

SRDF/AR combines SRDF and TimeFinder to provide a long-distance disaster restart solution. SRDF/AR can be deployed in 2-site or 3-site solutions:

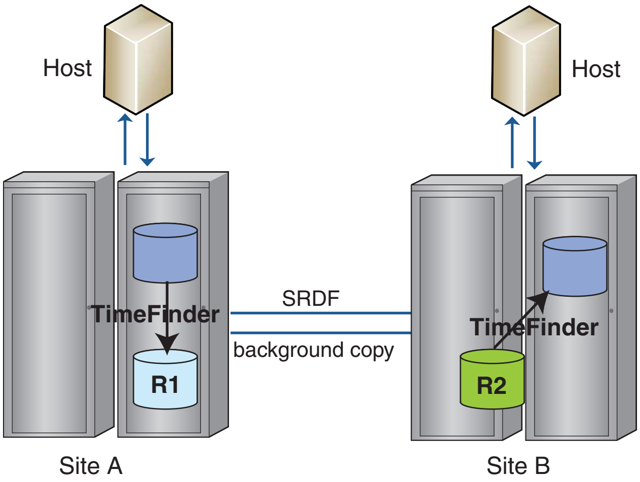
- In 2-site solutions, SRDF/DM is deployed with TimeFinder.
- In 3-site solutions, SRDF/DM is deployed with a combination of SRDF/S and TimeFinder.

The time to create the new replicated consistent image is determined by the time that it takes to replicate the deltas.

SRDF/AR 2-site solutions

The following image shows a 2-site solution where the production device (R1) on the primary array (Site A) is also a TimeFinder target device:

Figure 46 SRDF/AR 2-site solution



In the 2-site solution, data on the SRDF R1/TimeFinder target device is replicated across the SRDF links to the SRDF R2 device.

The SRDF R2 device is also a TimeFinder source device. TimeFinder replicates this device to a TimeFinder target device. You can map the TimeFinder target device to the host connected to the secondary array at Site B.

In the 2-site solution, SRDF operations are independent of production processing on both the primary and secondary arrays. You can utilize resources at the secondary site without interrupting SRDF operations.

Use SRDF/AR 2-site solutions to:

- Reduce required network bandwidth using incremental resynchronization between the SRDF target sites.
- Reduce network cost and improve resynchronization time for long-distance SRDF implementations.

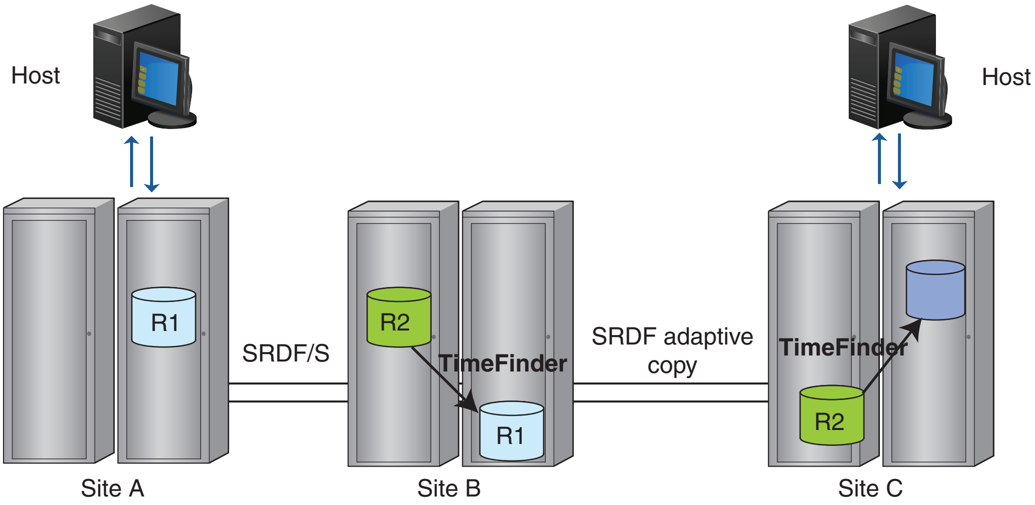
SRDF/AR 3-site solutions

SRDF/AR 3-site solutions provide a zero data loss solution at long distances in the event that the primary site is lost.

The following image shows a 3-site solution where:

- Site A and Site B are connected using SRDF in synchronous mode.
- Site B and Site C are connected using SRDF in adaptive copy mode.

Figure 47 SRDF/AR 3-site solution



If Site A (primary site) fails, the R2 device at Site B provides a restartable copy with zero data loss. Site C provides an asynchronous restartable copy.

If both Site A and Site B fail, the device at Site C provides a restartable copy with controlled data loss. The amount of data loss is a function of the replication cycle time between Site B and Site C.

SRDF and TimeFinder control commands to R1 and R2 devices for all sites can be issued from Site A. No controlling host is required at Site B.

Use SRDF/AR 3-site solutions to:

- Reduce required network bandwidth using incremental resynchronization between the secondary SRDF target site and the tertiary SRDF target site.
- Reduce network cost and improve resynchronization time for long-distance SRDF implementations.
- Provide disaster recovery testing, point-in-time backups, decision support operations, third-party software testing, and application upgrade testing or the testing of new applications.

Requirements/restrictions

In a 3-site SRDF/AR multi-hop solution, SRDF/S host I/O to Site A is not acknowledged until Site B has acknowledged it. This can cause a delay in host response time.

TimeFinder and SRDF/A

In SRDF/A solutions, device-level pacing:

- Prevents cache utilization bottlenecks when the SRDF/A R2 devices are also TimeFinder source devices.
- Allows R2 or R22 devices at the middle hop to be used as TimeFinder source devices. Device-level (TimeFinder) pacing on page 126 provides more information.

Note

Device-level write pacing is not required in configurations that include Enginuity 5876 and HYPERMAX OS.

TimeFinder and SRDF/S

SRDF/S solutions support any type of TimeFinder copy sessions running on R1 and R2 devices as long as the conditions described in R1 and R2 devices in TimeFinder operations on page 152 are met.

CHAPTER 9

Data Migration

This chapter describes data migration solutions.

Topics include:

Overview

Data migration is a one-time movement of data from a source to a target. Typical examples are data center refreshes where data is moved off an old array after which the array is retired or re-purposed. Data migration is *not* data movement due to replication (where the source data is accessible after the target is created) or data mobility (where the target is continually updated).

After a data migration operation, applications that access the data must reference the data at the new location.

To plan a data migration, consider the potential impact on your business, including:

- Type of data to be migrated
- Site location(s)
- Number of systems and applications
- Amount of data to be moved
- Business needs and schedules

Data migration solutions for open systems environments

This section explains the data migration features available for open system environments.

Non-Disruptive Migration overview

Non-Disruptive Migration (NDM) provides a method for migrating data from a source array to a target array across a metro distance, typically within a data center, without application host downtime. NDM requires a VMAX array running Enginuity 5876 with required ePack (source array), and an array running HYPERMAX OS 5977.811.784 or higher (target array). Consult with Dell EMC for required ePack for source arrays running Enginuity 5876. In addition, refer to the NDM support matrix available on eLab Navigator for array operating system version support, host support, and multipathing support for NDM operations.

If regulatory or business requirements for DR (disaster recovery) dictate the use of SRDF/S during migration, contact Dell EMC for required ePacks for SRDF/S configuration.

The NDM operations involved in a typical migration are:

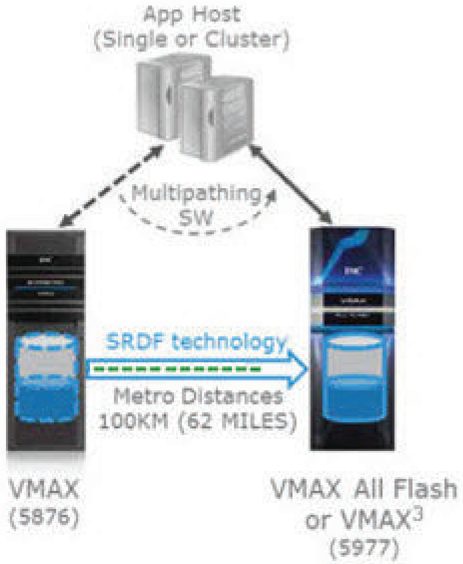
- Environment setup – Configures source and target array infrastructure for the migration process.
- Create – Duplicates the application storage environment from source array to target array.
- Cutover – Switches the application data access form the source array to the target array and duplicates the application data on the source array to the target array.
- Commit – Removes application resources from the source array and releases the resources used for migration. Application permanently runs on the target array.
- Enviroment remove –Removes the migration infrastructure created by the environmental setup.

Some key features of NDM are:

- Simple process for migration:
  1. Select storage group to migrate.
  2. Create the migration session.
  3. Discover paths to the host.
  4. Cutover storage group to VMAX3 or VMAX All Flash array.
  5. Monitor for synchronization to complete.
  6. Commit the migration.
- Allows for inline compression on VMAX All Flash array during migration.
- Maintains snapshot and disaster recovery relationships on source array, but are not migrated.
- Allows for non-disruptive revert to source array.
- Allows up to 16 concurrent migration sessions.
- Requires no license since it is part of HYPERMAX OS.
- Requires no additional hardware in the data path.

The following graphic shows the connections required between the host (single or cluster) and the source and target array, and the SRDF connection between the two arrays.

Figure 48 Non-Disruptive Migration zoning



The App host connection to both arrays uses FC, and the SRDF connection between arrays uses FC – GigE .

It is recommended that migration controls run from a control host and not the application host. The control host should have visibility to both the source array and target array.

The following devices and components are not supported with NDM:

- CKD devices, IBM i devices
- eNAS data
- ProtectPoint, FAST.X, and CloudArray relationships and associated data

Environmental requirements for Non-Disruptive Migration

The following configurations are required for a successful data migration:

Array configuration

- The target array must be running HYPERMAX OS 5977.811.784 or higher. This includes VMAX3 Family arrays and VMAX All Flash arrays.
- The source array must be a VMAX array running Enginuity 5876 with required ePack (contact Dell EMC for required ePack).
- SRDF is used for data migration, so zoning of SRDF ports between the source and target arrays is required. Note that an SRDF license is not required, as there is no charge for NDM.
- The NDM RDF group is configured with a minimum of two paths on different directors for redundancy and fault tolerance. If more paths are found up to eight paths will be configured.
- If SRDF is not normally used in the migration environment, it may be necessary to install and configure RDF directors and ports on both the source and target arrays and physically configure SAN connectivity.

Host configuration

- It is recommended to run NDM commands from a control host (a host separate from the application host).
- Both the source and the target array should be visible to the controlling host that runs the migration commands.
- If the application and NDM commands need to run on the same host, several gatekeeper devices must be provided to control the array. In addition, in the `daemon_options` file the gatekeeper use (`gk_use`) option must be set for dedicated use only, as follows:
  1. In the `/var/symapi/config/daemon_options` file, add the line `storapid:gk_use=dedicated_only`
  2. Save the file.
  3. Run the command `# storedaemon action storapid -cmd reload` to activate the new options setting.

Note

A `gkselect`file, that lists gatekeeper devices is recommended. For more information on the gkselect file, refer to EMC Solutions Enabler Installation and Configuration Guide.

Pre-migration rules and restrictions for Non-Disruptive Migration

In addition to general configuration requirements of the migration environment, the following conditions are evaluated by Solutions Enabler prior to starting a migration.

- A Storage Group is the data container that is migrated, and the following requirements apply to a storage group and its devices:
  - Storage groups must have masking views. All devices within the storage group on the source VMAX must be visible only through a masking view. The device must be mapped to a port that is part of the masking view.

  - Multiple masking views on the storage group using the same initiator group are only allowed if port groups on the target array already exist for each masking view, and the ports in the port groups are selected.
  - Storage groups must be parent or standalone storage groups. A child storage group with a masking view on the child storage group is not supported.
  - Gatekeeper devices in the storage group are not migrated to the target array.
  - Devices must not be masked or mapped as FCoE ports, iSCSI ports, or non-ACLX enabled ports.
  - Devices cannot be in storage groups that are masked.
- For objects that may already exist on the target array, the following restrictions apply:
  - The names of the storage groups (parent and/or children) to be migrated must not exist on the target array.
  - The names of masking views to be migrated must not exist on the target array.
  - The names of the initiator groups to be migrated may exist on the target array. However, the initiator groups on the target array must have the exact same initiators, child groups and port flags as the initiator groups to be migrated. Port flags that are not supported on the VMAX3 arrays are ignored. If an IG on the target array has the same WWNs used on the source array, then the IG name on the target array must be exactly the same as the IG name on the source array. If the initiators are already in an IG on the target array, the operation will be blocked unless the IG on the target array has the same name as the source array, and the IG must have the exact same, initiators, child groups and port flags as the source array. In addition, the consistent lun flag setting on the source array IG must also match the IG flag setting on the target array.
  - The names of the port groups to be migrated may exist on the target array, provided that the groups on the target array have the initiators logged into at least one port in the port group.
- The status of the target array must be as follows:
  - If a target-side Storage Resource Pool (SRP) is specified for the migration that SRP must exist on the target array.
  - The SRP to be used for target-side storage must have enough free capacity to support the migration.
  - If compression is enabled for the storage group to be migrated, it must be supported by the SRP on the target array.
  - The target side must be able to support the additional devices required to receive the source-side data.
  - All initiators provisioned to an application on the source array must also be logged into ports on the target array.
- Only FBA devices are supported (Celerra and D910 are not supported) and the following restrictions apply:
  - Cannot have user geometry set, non-birth identity, or the BCV attribute.
  - Cannot be encapsulated, a Data Domain device, or a striped meta device with different size members.

  - Must be dynamic SRDF R1 and SRDF R2 (DRX) capable and be R1 or non-RDF devices, but cannot be R2 or concurrent RDF devices, or part of a Star Consistency Group.

- Devices in the storage group to be migrated can have TimeFinder sessions and/or they can be R1 devices. The migration controls evaluates the state of these devices to determine if the control operation can proceed.
- The devices in the storage group cannot be part of another migration session.

Migration infrastructure - RDF device pairing

RDF device pairing is done during the create operation, with the following actions occurring on the device pairs.

- NDM creates RDF device pairs, in a DM RDF group, between devices on the source array and the devices on the target array.
- Once device pairing is complete NDM controls the data flow between both sides of the migration process.
- Once the migration is complete, the RDF pairs are deleted when the migration is committed.
- Other RDF pairs may exist in the DM RDF group if another migration is still in progress.

Due to differences in device attributes between the source and target array, the following rules apply during migration:

- Any source array device that has an odd number of cylinders is migrated to a device on the target array that has Geometry Compatibility Mode (GCM).
- Any source array meta device is migrated to a non-meta device on the target array.

About Open Replicator

Open Replicator enables copying data (full or incremental copies) from qualified arrays within a storage area network (SAN) infrastructure to or from arrays running HYPERMAX OS. Open Replicator uses the Solutions Enabler SYMCLI `symrcopy` command.

Use Open Replicator to migrate and back up/archive existing data between arrays running HYPERMAX OS and third-party storage arrays within the SAN infrastructure without interfering with host applications and ongoing business operations.

Use Open Replicator to:

- Pull from source volumes on qualified remote arrays to a volume on an array running HYPERMAX OS.
- Perform online data migrations from qualified storage to an array running HYPERMAX OS with minimal disruption to host applications.

NOTICE

Open Replicator cannot copy a volume that is in use by SRDF or TimeFinder.

Open Replicator operations

Open Replicator includes the following terminology:

Control

The recipent array and its devices are referred to as the control side of the copy operation.

Remote

The donor EMC arrays or third-party arrays on the SAN are referred to as the remote array/devices.

Hot

The Control device is Read/Write online to the host while the copy operation is in progress.

Note

Hot push operations are not supported on arrays running HYPERMAX OS.

Cold

The Control device is Not Ready (offline) to the host while the copy operation is in progress.

Pull

A pull operation copies data to the control device from the remote device(s).

Push

A push operation copies data from the control device to the remote device(s).

Pull operations

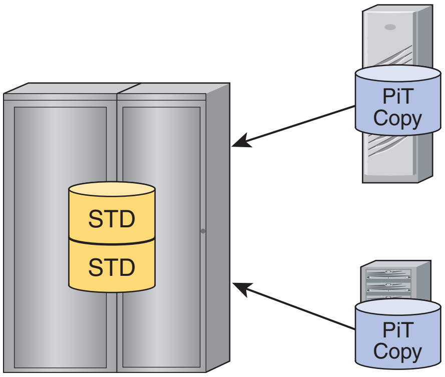
Arrays running HYPERMAX OS support up to 512 pull sessions.

For pull operations, the volume can be in a live state during the copy process. The local hosts and applications can begin to access the data as soon as the session begins, even before the data copy process has completed.

These features enable rapid and efficient restoration of remotely vaulted volumes and migration from other storage platforms.

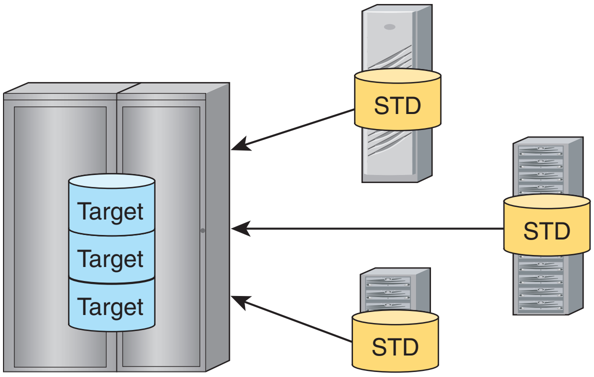
Copy on First Access ensures the appropriate data is available to a host operation when it is needed. The following image shows an Open Replicator hot pull.

Figure 49 Open Replicator hot (or live) pull



The pull can also be performed in cold mode to a static volume. The following image shows an Open Replicator cold pull.

Figure 50 Open Replicator cold (or point-in-time) pull



PowerPath Migration Enabler

EMC PowerPath is host-based software that provides automated data path management and load-balancing capabilities for heterogeneous server, network, and storage deployed in physical and virtual environments. PowerPath includes a migration tool called PowerPath Migration Enabler (PPME). PPME enables non-disruptive or minimally disruptive data migration between storage systems or within a single storage system.

PPME allows applications continued data access throughout the migration process. PPME integrates with other technologies to minimize or eliminate application downtime during data migration.

PPME works in conjunction with underlying technologies, such as Open Replicator, SnapVX, and Host Copy.

Note

PowerPath Multipathing must be installed on the host machine.

The following documentation provides additional information:

- *EMC Support Matrix PowerPath Family Protocol Support*
- *EMC PowerPath Migration Enabler User Guide*

Data migration using SRDF/Data Mobility

SRDF/Data Mobility (DM) uses SRDF's adaptive copy mode to transfer large amounts of data without impact to the host.

SRDF/DM supports data replication or migration between two or more arrays running HYPERMAX OS. Adaptive copy mode enables applications using the primary volume to avoid propagation delays while data is transferred to the remote site. SRDF/DM can be used for local or remote transfers.

Refer to Migration using SRDF/Data Mobility on page 132.

Migrating data with concurrent SRDF

In concurrent SRDF topologies, you can non-disruptively migrate data between arrays along one SRDF leg while remote mirroring for protection along the other leg.

Once the migration process completes, the concurrent SRDF topology is removed, resulting in a 2-site SRDF topology.

Replacing R2 devices with new R2 devices

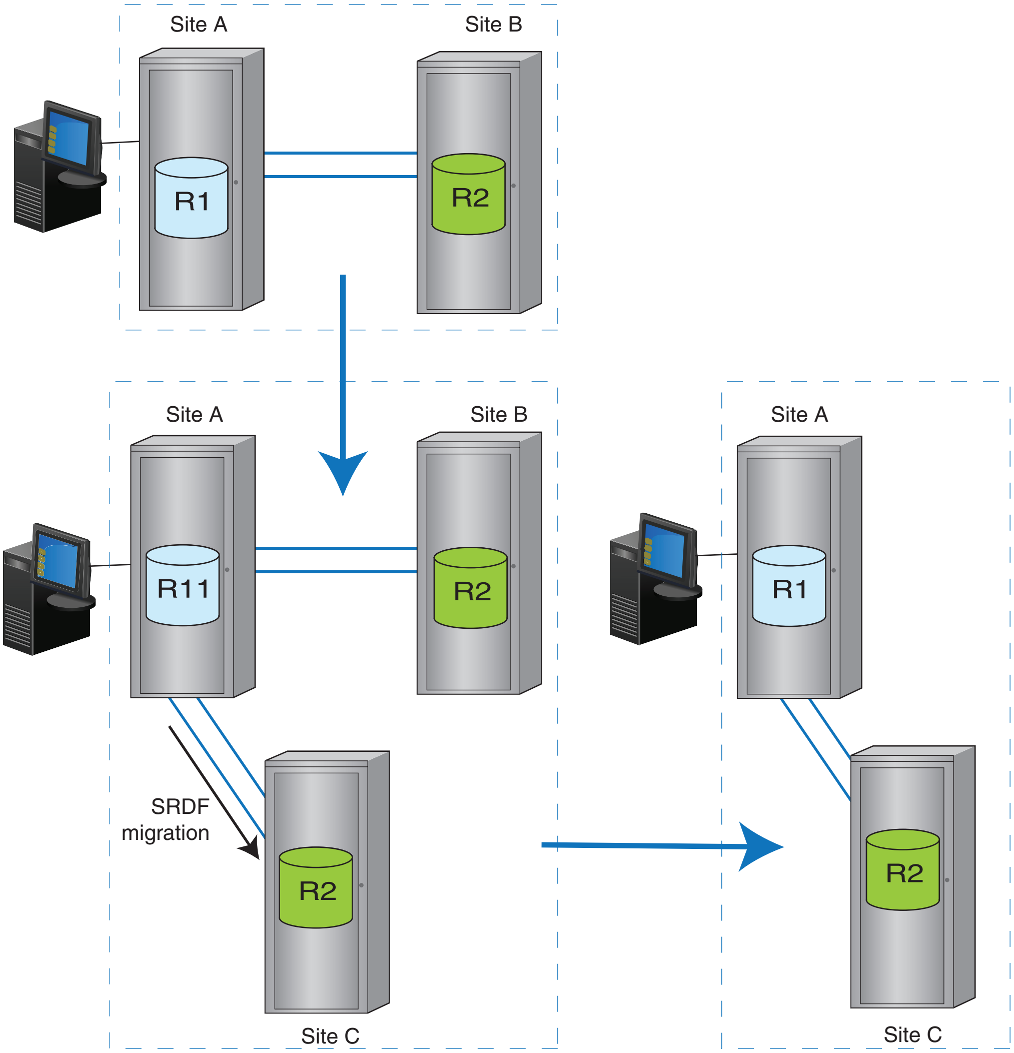
You can manually migrate data as shown in the following image, including:

- Initial 2-site topology
- The interim 3-site migration topology
- Final 2-site topology

After migration, the original primary array is mirrored to a new secondary array.

EMC support personnel are available to assist with the planning and execution of your migration projects.

Figure 51 Migrating data and removing the original secondary array (R2)



Replacing R1 devices with new R1 devices

The following image shows replacing the original R1 devices with new R1 devices, including:

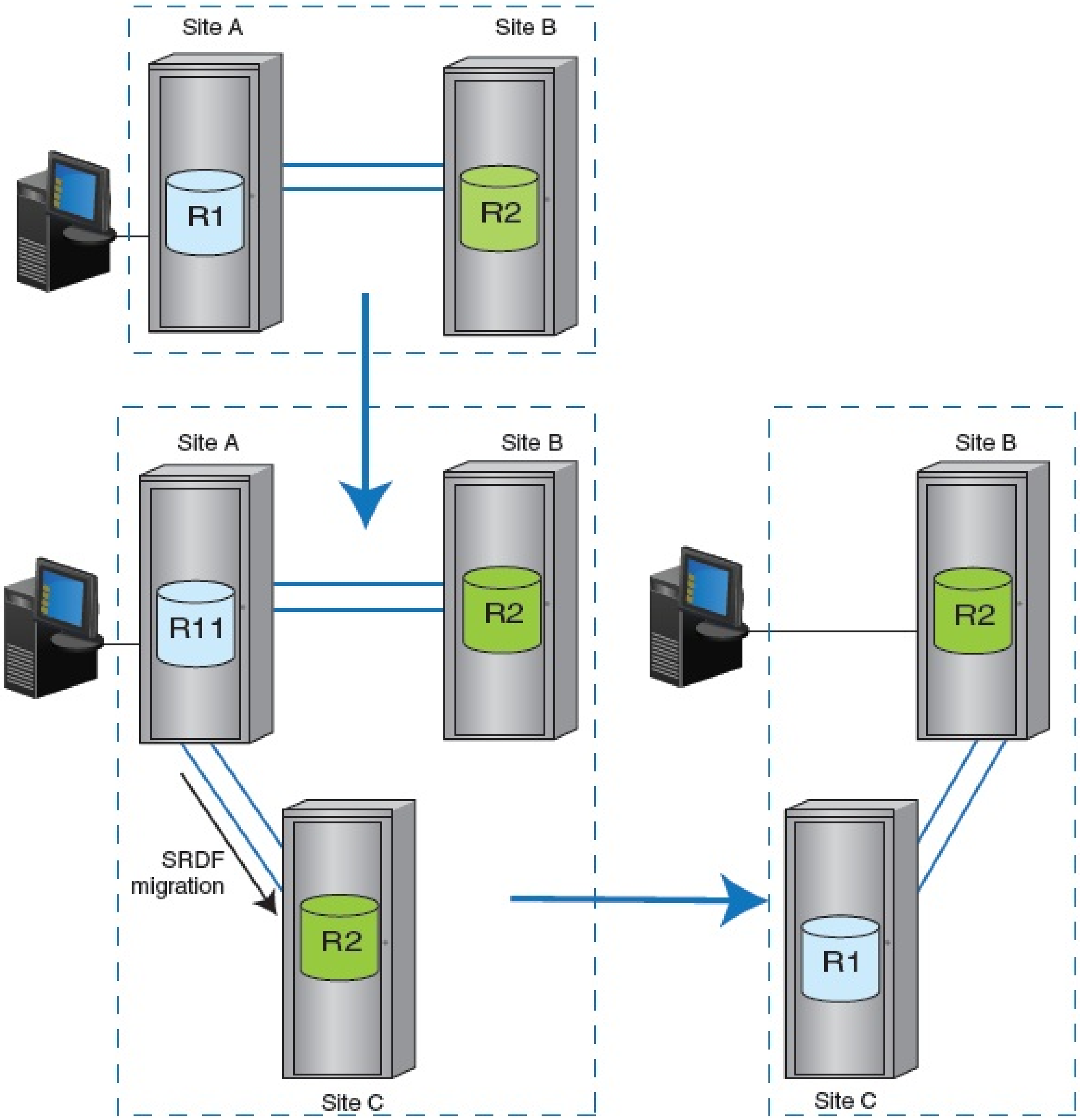
- Initial 2-site topology
- The interim 3-site migration topology

- Final 2-site topology

After migration, the new primary array is mirrored to the original secondary array.

EMC support personnel are available to assist with the planning and execution of your migration projects.

Figure 52 Migrating data and replacing the original primary array (R1)



Replacing R1 and R2 devices with new R1 and R2 devices

You can use the combination of concurrent SRDF and cascaded SRDF to replace both R1 and R2 devices at the same time.

Note

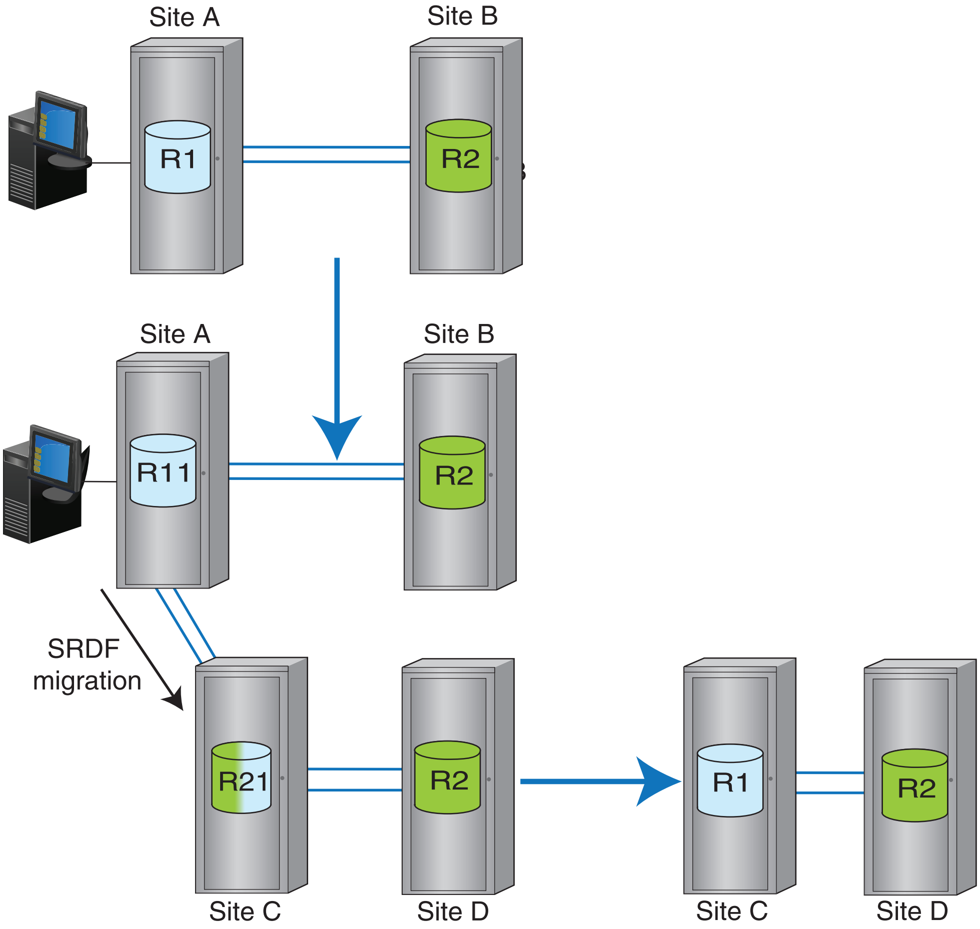
Before you begin, verify that your specific hardware models and Enginuity or HYPERMAX OS versions are supported for migrating data between different platforms.

The following image shows an example of replacing both R1 and R2 devices with new R1 and R2 devices at the same time, including:

- Initial 2-site topology
- Migration process
- The final topology

EMC support personnel is available to assist with the planning and execution of your migration projects.

Figure 53 Migrating data and replacing the original primary (R1) and secondary (R2) arrays



Space and zero-space reclamation

Space reclamation reclaims unused space following a replication or migration activity from a regular device to a thin device in which software tools, such as Open Replicator and Open Migrator, copied-all-zero, unused space to a target thin volume.

Space reclamation deallocates data chunks that contain all zeros. Space reclamation is most effective for migrations from standard, fully provisioned devices to thin devices.

Space reclamation is non-disruptive and can be executed while the targeted thin device is fully available to operating systems and applications.

Zero-space reclamations provides instant zero detection during Open Replicator and SRDF migration operations by reclaiming all-zero space, including both host-unwritten extents (or chunks) and chunks that contain all zeros due to file system or database formatting.

Solutions Enabler and Unisphere for VMAX can be used to initiate and monitor the space reclamation process.

Data migration solutions for mainframe environments

For mainframe environments, z/OS Migrator provides non-disruptive migration from any vendor storage to VMAX arrays. z/OS Migrator can also migrate data from one VMAX array to another. With z/OS Migrator, you can:

- Introduce new storage subsystem technologies with minimal disruption of service.
- Reclaim z/OS UCBs by simplifying the migration of datasets to larger volumes (combining volumes).
- Facilitate data migration while applications continue to run and fully access data being migrated, eliminating application downtime usually required when migrating data.
- Eliminate the need to coordinate application downtime across the business, and eliminate the costly impact of such downtime on the business.
- Improve application performance by facilitating the relocation of poor performing datasets to lesser used volumes/storage arrays.
- Ensure all metadata always accurately reflects the location and status of datasets being migrated.

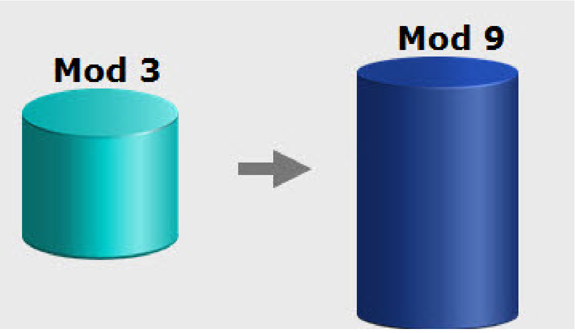
Note

Refer to the *z/OS Migrator Product Guide* for detailed product information.

Volume migration using z/OS Migrator

EMC z/OS Migrator is a host-based data migration facility that performs traditional volume migrations as well as host-based volume mirroring. Together, these capabilities are referred to as the volume mirror and migrator functions of z/OS Migrator.

Figure 54 z/OS volume migration



Volume level data migration facilities move logical volumes in their entirety. z/OS Migrator volume migration is performed on a track for track basis without regard to the logical contents of the volumes involved. Volume migrations end in a volume swap which is entirely non-disruptive to any applications using the data on the volumes.

Volume migrator

Volume migration provides host-based services for data migration at the volume level on mainframe systems. It provides migration from third-party devices to VMAX devices as well as migration between VMAX devices.

Volume mirror

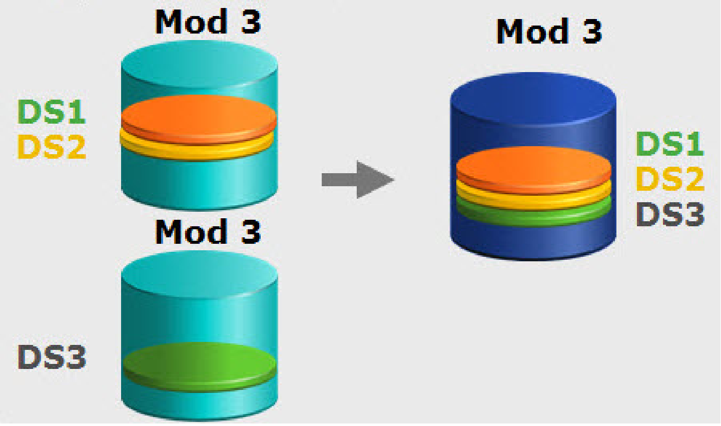
Volume mirroring provides mainframe installations with volume-level mirroring from one VMAX device to another. It uses host resources (UCBs, CPU, and channels) to monitor channel programs scheduled to write to a specified primary volume and clones them to also write to a specified target volume (called a mirror volume).

After achieving a state of synchronization between the primary and mirror volumes, Volume Mirror maintains the volumes in a fully synchronized state indefinitely, unless interrupted by an operator command or by an I/O failure to a Volume Mirror device. Mirroring is controlled by the volume group. Mirroring may be suspended consistently for all volumes in the group.

Dataset migration using z/OS Migrator

In addition to volume migration, z/OS Migrator provides for logical migration, that is, the migration of individual datasets. In contrast to volume migration functions, z/OS Migrator performs dataset migrations with full awareness of the contents of the volume, and the metadata in the z/OS system that describe the datasets on the logical volume.

Figure 55 z/OS Migrator dataset migration



Thousands of datasets can either be selected individually or wild-carded. z/OS Migrator automatically manages all metadata during the migration process while applications continue to run.

CHAPTER 10

CloudArray® for VMAX All Flash

This chapter provides an overview of CloudArray® for VMAX All Flash. Topics include:

About CloudArray

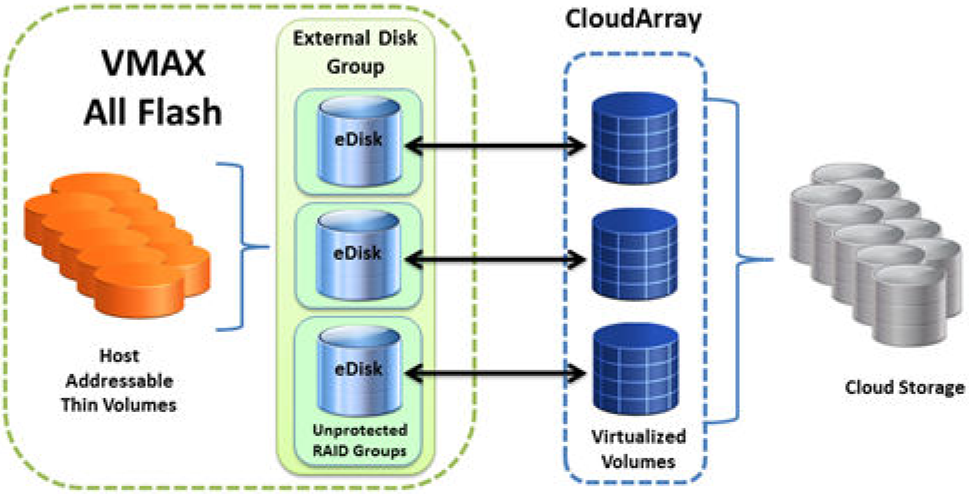
EMC CloudArray is a storage software technology that integrates cloud-based storage into traditional enterprise IT environments. Traditionally, as data volumes increase, organizations must choose between growing the storage environment, supplementing it with some form of secondary storage, or simply deleting cold data. CloudArray combines the resource efficiency of the cloud with on-site storage, allowing organizations to scale their infrastructure and plan for future data growth. CloudArray makes cloud object storage look, act, and feel like local storage, seamlessly integrating with existing applications, giving a virtually unlimited tier of storage in one easy package. By connecting storage systems to high-capacity cloud storage, CloudArray enables a more efficient use of high performance primary arrays while leveraging the cost efficiencies of cloud storage.

CloudArray offers a rich set of features to enable cloud integration and protection for VMAX All Flash data:

- CloudArray's local drive caching ensures recently accessed data is available at local speeds without the typical latency associated with cloud storage.
- CloudArray provides support for more than 20 different public and private cloud providers, including Amazon, EMC ECS, Google Cloud, and Microsoft Azure.
- 256-bit AES encryption provides security for all data that leaves CloudArray, both in-flight to and at rest in the cloud.
- File and block support enables CloudArray to integrate the cloud into the storage environment regardless of the data storage level.
- Data compression and bandwidth scheduling reduce cloud capacity demands and limit network impact.

The following figure illustrates a typical CloudArray deployment for VMAX All Flash.

Figure 56 CloudArray deployment for VMAX All Flash



CloudArray physical appliance

The physical appliance supplies the physical connection capability from the VMAX All Flash to cloud storage using Fibre Channel controller cards. FAST.X presents the physical appliance as an external device.

The CloudArray physical appliance is a 2U server that consists of:

- Up to 40TB usable local cache (12x4TB drives in a RAID-6 configuration)
- 192GB RAM
- 2x2 port 8Gb Fibre Channel cards configured in add-in slots on the physical appliance

Cloud provider connectivity

CloudArray connects directly with more than 20 public and private cloud storage providers. CloudArray converts the cloud's object-based storage to one or more local volumes.

Dynamic caching

CloudArray addresses bandwidth and latency issues typically associated with cloud storage by taking advantage of local storage, called cache. The disk-based cache provides local performance for active data and serves as a buffer for read-write operations. Each volume can be associated with its own, dedicated cache, or can operate off of a communal pool. The amount of cache assigned to each volume can be individually configured. A volume's performance depends on the amount of data kept locally in the cache and the type of disk used for the cache.

For more information on CloudArray cache and configuration guidelines, see the *EMC CloudArray Best Practices* whitepaper on EMC.com.

Security and data integrity

CloudArray employs both in-flight and at-rest encryptions to ensure data security. Each volume can be encrypted using 256-bit AES encryption prior to replicating to the cloud. CloudArray also encrypts the data and metadata separately, storing the different encryption keys locally to prevent any unauthorized access.

CloudArray's encryption is a critical component in ensuring data integrity. CloudArray segments its cache into cache pages and, as part of the encryption process, generates and assigns a unique hash to each cache page. The hash remains with the cache page until that page is retrieved for access by a requesting initiator. When the page is decrypted, the hash must match the value generated by the decryption algorithm. If the hash does not match, then the page is declared corrupt. This process helps prevent any data corruption from propagating to an end user.

Administration

CloudArray is configured using a browser-based graphical user interface. With this interface administrators can:

- Create, modify, or expand volumes, file shares and caches

- Monitor and display CloudArray health, performance, and cache status
- Apply software updates
- Schedule and configure snapshots and bandwidth throttling

CloudArray also utilizes an online portal that enables users to:

- Download CloudArray licenses and software updates
- Configure alerts and access CloudArray product documentation
- Store a copy of the CloudArray configuration file for disaster recovery retrieval

APPENDIX A

Mainframe Error Reporting

This appendix describes mainframe environmental errors.

Error reporting to the mainframe host

HYPERMAX OS can detect the following error types to the mainframe host in the VMAX storage systems:

- Data Check — HYPERMAX OS detected an error in the bit pattern read from the disk. Data checks are due to hardware problems when writing or reading data, media defects, or random events.
- System or Program Check — HYPERMAX OS rejected the command. This type of error is indicated to the processor and is always returned to the requesting program.
- Overrun — HYPERMAX OS cannot receive data at the rate it is transmitted from the host. This error indicates a timing problem. Resubmitting the I/O operation usually corrects this error.
- Equipment Check — HYPERMAX OS detected an error in hardware operation.
- Environmental — HYPERMAX OS internal test detected an environmental error. Internal environmental tests monitor, check, and report failures of the critical hardware components. They run at the initial system power-up, upon every software reset event, and at least once every 24 hours during regular operations.

If an environmental test detects an error condition, it sets a flag to indicate a pending error and presents a unit check status to the host on the next I/O operation. The test that detected the error condition is then scheduled to run more frequently. If a device-level problem is detected, it is reported across all logical paths to the device experiencing the error. Subsequent failures of that device are not reported until the failure is fixed.

If a second failure is detected for a device while there is a pending error-reporting condition in effect, HYPERMAX OS reports the pending error on the next I/O and then the second error.

Enginuity reports error conditions to the host and to the EMC Customer Support Center. When reporting to the host, Enginuity presents a unit check status in the status byte to the channel whenever it detects an error condition such as a data check, a command reject, an overrun, an equipment check, or an environmental error.

When presented with a unit check status, the host retrieves the sense data from the VMAX array and, if logging action has been requested, places it in the Error Recording Data Set (ERDS). The EREP (Environment Recording, Editing, and Printing) program prints the error information. The sense data identifies the condition that caused the interruption and indicates the type of error and its origin. The sense data format depends on the mainframe operating system. For 2105, 2107, or 3990 controller emulations, the sense data is returned in the SIM format.

SIM severity reporting

HYPERMAX OS supports SIM severity reporting that enables filtering of SIM severity alerts reported to the multiple virtual storage (MVS) console.

- All SIM severity alerts are reported by default to the EREP (Environmental Record Editing and Printing program).
- ACUTE, SERIOUS, and MODERATE alerts are reported by default to the MVS console.

The following table lists the default settings for SIM severity reporting.

Table 40 SIM severity alerts

Severity	Description
SERVICE	No system or application performance degradation is expected. No system or application outage has occurred.
MODERATE	Performance degradation is possible in a heavily loaded environment. No system or application outage has occurred.
SERIOUS	A primary I/O subsystem resource is disabled. Significant performance degradation is possible. System or application outage may have occurred.
ACUTE	A major I/O subsystem resource is disabled, or damage to the product is possible. Performance may be severely degraded. System or application outage may have occurred.
REMOTE SERVICE	EMC Customer Support Center is performing service/ maintenance operations on the system.
REMOTE FAILED	The Service Processor cannot communicate with the EMC Customer Support Center.

Environmental errors

The following table lists the environmental errors in SIM format for HYPERMAX OS 5977 or higher.

Note

All listed severity levels can be modified via SymmWin.

Table 41 Environmental errors reported as SIM messages

Hex code	Severity level	Description	SIM reference code
04DD	MODERATE	MMCS health check error	24DD
043E	MODERATE	An SRDF Consistency Group was suspended.	E43E
044D	MODERATE	An SRDF path was lost.	E44D
044E	SERVICE	An SRDF path is operational after a previous failure.	E44E
0461	NONE	The M2 is resynchronized with the M1 device. This event occurs once the M2 device is brought back to a Ready state. [a]	E461
0462	NONE	The M1 is resynchronized with the M2 device. This event occurs once the M1 device is brought back to a Ready state. [a]	E462

Table 41 Environmental errors reported as SIM messages (continued)

Hex code	Severity level	Description	SIM reference code
0463	SERIOUS	One of the back-end directors failed into the IMPL Monitor state.	2463
0465	NONE	Device resynchronization process has started. [a]	E465
0467	MODERATE	The remote storage system reported an SRDF error across the SRDF links.	E467
046D	MODERATE	An SRDF group is lost. This event happens, for example, when all SRDF links fail.	E46D
046E	SERVICE	An SRDF group is up and operational.	E46E
0470	ACUTE	OverTemp condition based on memory module temperature.	2470
0471	ACUTE	The Storage Resource Pool has exceeded its upper threshold value.	2471
0473	SERIOUS	A periodic environmental test (env_test9) detected the mirrored device in a Not Ready state.	E473
0474	SERIOUS	A periodic environmental est (env_test9) detected the mirrored device in a Write Disabled (WD) state.	E474
0475	SERIOUS	An SRDF R1 remote mirror is in a Not Ready state.	E475
0476	SERVICE	Service Processor has been reset.	2476
0477	REMOTE FAILED	The Service Processor could not call the EMC Customer Support Center (failed to call home) due to communication problems.	1477
047A	MODERATE	AC power lost to Power Zone A or B.	247A
047B	MODERATE	Drop devices after RDF Adapter dropped.	E47B
01BA 02BA 03BA	ACUTE	Power supply or enclosure SPS problem.	24BA

Table 41 Environmental errors reported as SIM messages (continued)

Hex code	Severity level	Description	SIM reference code
04BA			
047C	ACUTE	The Storage Resource Pool has Not Ready or Inactive TDATs.	247C
047D	MODERATE	Either the SRDF group lost an SRDF link or the SRDF group is lost locally.	E47D
047E	SERVICE	An SRDF link recovered from failure. The SRDF link is operational.	E47E
047F	REMOTE SERVICE	The Service Processor successfully called the EMC Customer Support Center (called home) to report an error.	147F
0488	SERIOUS	Replication Data Pointer Meta Data Usage reached 90-99%.	E488
0489	ACUTE	Replication Data Pointer Meta Data Usage reached 100%.	E489
0492	MODERATE	Flash monitor or MMCS drive error.	2492
04BE	MODERATE	Meta Data Paging file system mirror not ready.	24BE
04CA	MODERATE	An SRDF/A session dropped due to a non-user request. Possible reasons include fatal errors, SRDF link loss, or reaching the maximum SRDF/A host-response delay time.	E4CA
04D1	REMOTE SERVICE	Remote connection established. Remote control connected.	14D1
04D2	REMOTE SERVICE	Remote connection closed. Remote control rejected.	14D2
04D3	MODERATE	Flex filter problems.	24D3
04D4	REMOTE SERVICE	Remote connection closed. Remote control disconnected.	14D4
04DA	MODERATE	Problems with task/threads.	24DA
04DB	SERIOUS	SYMPL script generated error.	24DB
04DC	MODERATE	PC related problems.	24DC

Table 41 Environmental errors reported as SIM messages (continued)

Hex code	Severity level	Description	SIM reference code
04E0	REMOTE FAILED	Communications problems.	14E0
04E1	SERIOUS	Problems in error polling.	24E1
052F	None	A sync SRDF write failure occurred.	E42F
3D10	SERIOUS	A SnapVX snapshot failed.	E410

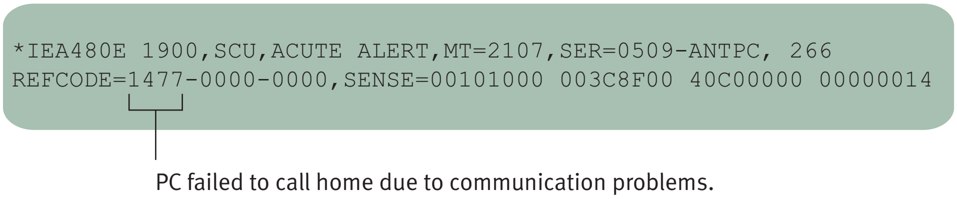
a. EMC recommendation: NONE.

Operator messages

Error messages

On z/OS, SIM messages are displayed as IEA480E Service Alert Error messages. They are formatted as shown below:

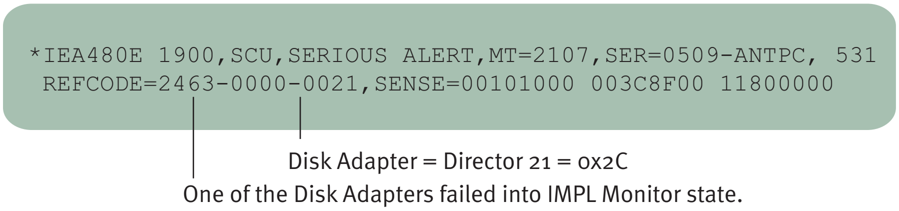
Figure 57 z/OS IEA480E acute alert error message format (call home failure)



```
*IEA480E 1900,SCU,ACUTE ALERT,MT=2107,SER=0509-ANTPC, 266
REFCODE=1477-0000-0000,SENSE=00101000 003C8F00 40C00000 00000014
```

PC failed to call home due to communication problems.

Figure 58 z/OS IEA480E service alert error message format (Disk Adapter failure)

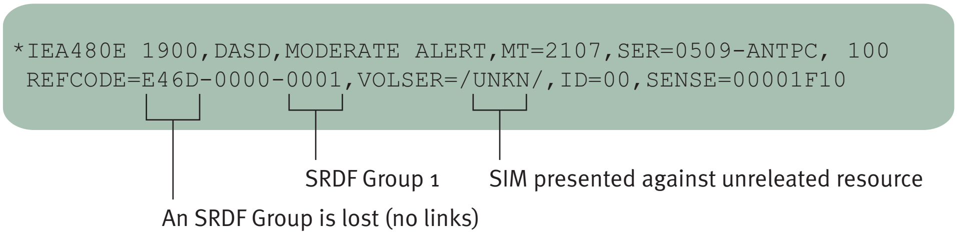


```
*IEA480E 1900,SCU,SERIOUS ALERT,MT=2107,SER=0509-ANTPC, 531
 REFCODE=2463-0000-0021,SENSE=00101000 003C8F00 11800000
```

Disk Adapter = Director 21 = 0x2C

One of the Disk Adapters failed into IMPL Monitor state.

Figure 59 z/OS IEA480E service alert error message format (SRDF Group lost/SIM presented against unrelated resource)



```
*IEA480E 1900,DASD,MODERATE ALERT,MT=2107,SER=0509-ANTPC, 100
 REFCODE=E46D-0000-0001,VOLSER=/UNKN/,ID=00,SENSE=00001F10
```

SRDF Group 1

SIM presented against unreleated resource

An SRDF Group is lost (no links)

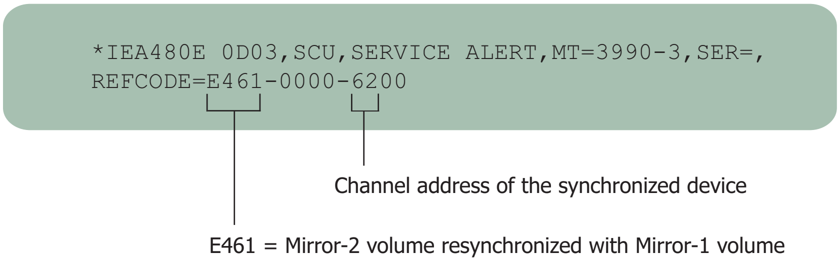
Event messages

The VMAX array also reports events to the host and to the service processor. These events are:

- The mirror-2 volume has synchronized with the source volume.
- The mirror-1 volume has synchronized with the target volume.
- Device resynchronization process has begun.

On z/OS, these events are displayed as IEA480E Service Alert Error messages. They are formatted as shown below:

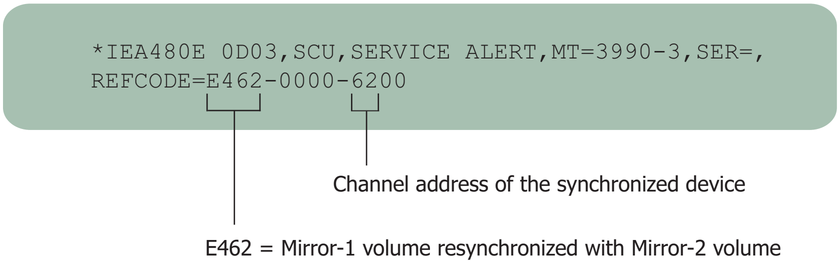
Figure 60 z/OS IEA480E service alert error message format (mirror-2 resynchronization)



```
*IEA480E 0D03,SCU,SERVICE ALERT,MT=3990-3,SER=,
REFCODE=E461-0000-6200
```

Channel address of the synchronized device
E461 = Mirror-2 volume resynchronized with Mirror-1 volume

Figure 61 z/OS IEA480E service alert error message format (mirror-1 resynchronization)



```
*IEA480E 0D03,SCU,SERVICE ALERT,MT=3990-3,SER=,
REFCODE=E462-0000-6200
```

Channel address of the synchronized device
E462 = Mirror-1 volume resynchronized with Mirror-2 volume

APPENDIX B

Licensing

This appendix provides an overview of licensing on arrays running HYPERMAX OS.

Topics include:

eLicensing

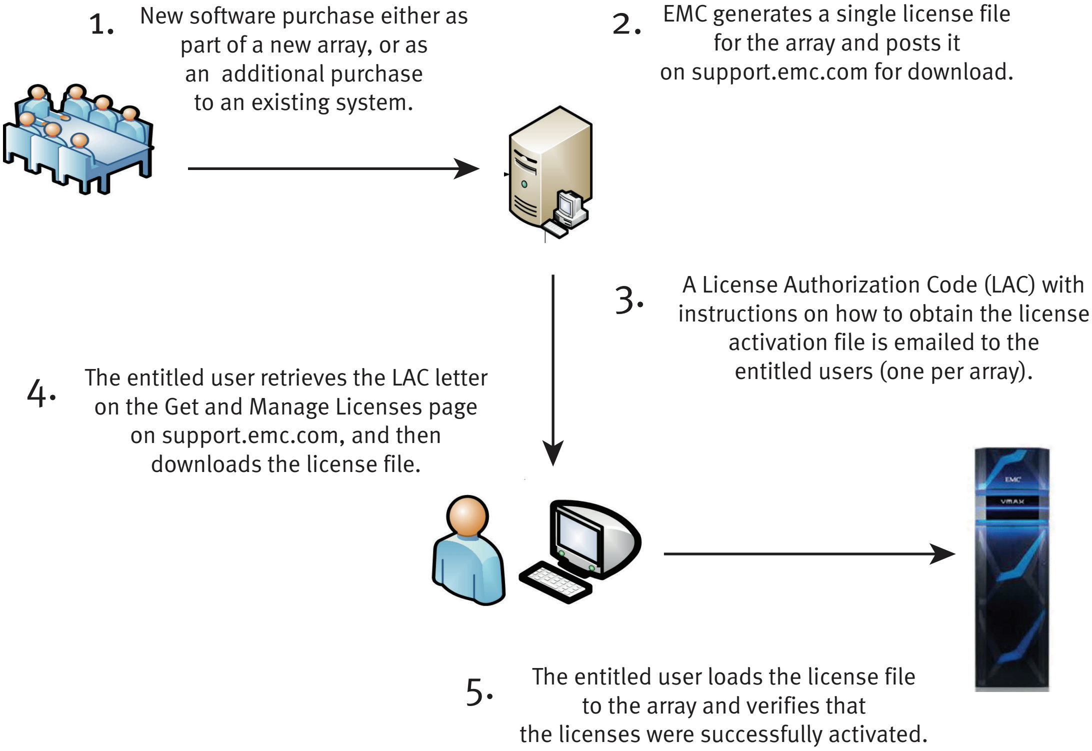
Arrays running HYPERMAX OS use *Electronic Licenses* (eLicenses).

Note

For more information on eLicensing, refer to EMC Knowledgebase article 335235 on the EMC Online Support website.

You obtain license files from EMC Online Support, copy them to a Solutions Enabler or a Unisphere for VMAX host, and push them out to your arrays. The following figure illustrates the process of requesting and obtaining your eLicense.

Figure 62 eLicensing process


1. New software purchase either as part of a new array, or as an additional purchase to an existing system.
2. EMC generates a single license file for the array and posts it on support.emc.com for download.
3. A License Authorization Code (LAC) with instructions on how to obtain the license activation file is emailed to the entitled users (one per array).
4. The entitled user retrieves the LAC letter on the Get and Manage Licenses page on support.emc.com, and then downloads the license file.
5. The entitled user loads the license file to the array and verifies that the licenses were successfully activated.

Note

To install array licenses, follow the procedure described in the *Solutions Enabler Installation Guide* and *Unisphere for VMAX online Help*.

Each license file fully defines all of the entitlements for a specific system, including the license type and the licensed capacity. To add a feature or increase the licensed capacity, obtain and install a new license file.

Most array licenses are array-based, meaning that they are stored internally in the system feature registration database on the array. However, there are a number of licenses that are host-based.

Array-based eLicenses are available in the following forms:

- An *individual license* enables a single feature.

- A *license suite* is a single license that enables multiple features. License suites are available only if all features are enabled.
- A *license pack* is a collection of license suites that fit a particular purpose.

To view effective licenses and detailed usage reports, use Solutions Enabler, Unisphere for VMAX, Mainframe Enablers, Transaction Processing Facility (TPF), or IBM i platform console.

Capacity measurements

Array-based licenses include a *capacity licensed* value that defines the scope of the license. The method for measuring this value depends on the license's *capacity type* (Usable or Registered).

Not all product titles are available in all capacity types, as shown below.

Table 42 VMAX All Flash product title capacity types

Usable	Registered	Other
All F software package titles	ProtectPoint	PowerPath (if purchased separately)
All FX software package titles		Events and Retention Suite
All zF software package titles		
All zFX software package titles		
RecoverPoint		

Usable capacity

Usable Capacity is defined as the amount of storage available for use on an array. The usable capacity is calculated as the sum of all Storage Resource Pool (SRP) capacities available for use. This capacity does not include any external storage capacity.

Registered capacity

Registered capacity is the amount of user data that will be managed or protected by each particular product title. It is independent of the type or size of the disks in the array.

The methods for measuring registered capacity depends on whether the licenses are part of a bundle or individual.

Registered capacity licenses

Registered capacity is measured according to the following:

- ProtectPoint
  - The registered capacity of this license is the sum of all DataDomain encapsulated devices that are link targets. When there are TimeFinder sessions present on an array with only a ProtectPoint license and no TimeFinder license, the capacity is calculated as the sum of all DataDomain encapsulated devices

with link targets and the sum of all TimeFinder allocated source devices and delta RDPs.

Open systems licenses

This section details the licenses available in an open system environment.

License suites

The following table lists the license suites available in an open systems environment.

Table 43 VMAX All Flash license suites

License suite	Includes	Allows you to	With the command
All Flash F	• HYPERMAX OS • Priority Controls • OR-DM • Unisphere for VMAX • FAST • SL Provisioning • Workload Planner • Database Storage Analyzer • Unisphere for File	Create time windows	symoptmz symtw
		• Add disk group tiers to FAST policies • Enable FAST • Set the following FAST parameters: ■ Swap Non-Visible Devices ■ Allow Only Swap ■ User Approval Mode ■ Maximum Devices to Move ■ Maximum Simultaneous Devices ■ Workload Period ■ Minimum Performance Period • Add virtual pool (VP) tiers to FAST policies • Set the following FAST VP-specific parameters: ■ Thin Data Move Mode ■ Thin Relocation Rate ■ Pool Reservation Capacity • Set the following FAST parameters: ■ Workload Period	symfast

Table 43 VMAX All Flash license suites (continued)

License suite	Includes	Allows you to	With the command
		▪ Minimum Performance Period	
		Perform SL-based provisioning	symconfigure symsg symcfg
	AppSync	Manage protection and replication for critical applications and databases for Microsoft, Oracle and VMware environments.	
	• • TimeFinder/Snap • TimeFinder/SnapVX • SnapSure	Create new native clone sessions	symclone
		Create new TimeFinder/ Clone emulations	symmir
		• Create new sessions • Duplicate existing sessions	symsnap
		• Create snap pools • Create SAVE devices	symconfigure
		• Perform SnapVX Establish operations • Perform SnapVX snapshot Link operations	symsnapvx
All Flash FX	All Flash F Suite	Perform tasks available in the All Flash F suite.	
	• SRDF • SRDF/Asynchronous • SRDF/Synchronous • SRDF/STAR • Replication for File	• Create new SRDF groups • Create dynamic SRDF pairs in Adaptive Copy mode	symrdf
		• Create SRDF devices • Convert non-SRDF devices to SRDF • Add SRDF mirrors to devices in Adaptive Copy mode	symconfigure

Table 43 VMAX All Flash license suites (continued)

License suite	Includes	Allows you to	With the command
		Set the dynamic-SRDF capable attribute on devices Create SAVE devices	
		• Create dynamic SRDF pairs in Asynchronous mode • Set SRDF pairs into Asynchronous mode	symrdf
		• Add SRDF mirrors to devices in Asynchronous mode Create RDFA_DSE pools Set any of the following SRDF/A attributes on an SRDF group: ■ Minimum Cycle Time ■ Transmit Idle ■ DSE attributes, including: – Associating an RDFA-DSE pool with an SRDF group DSE Threshold DSE Autostart ■ Write Pacing attributes, including: – Write Pacing Threshold – Write Pacing Autostart – Device Write Pacing exemption – TimeFinder Write Pacing Autostart	symconfigure

Table 43 VMAX All Flash license suites (continued)

License suite	Includes	Allows you to	With the command
		• Create dynamic SRDF pairs in Synchronous mode • Set SRDF pairs into Synchronous mode	symrdf
		Add an SRDF mirror to a device in Synchronous mode	symconfigure
	D@RE	Encrypt data and protect it against unauthorized access unless valid keys are provided. This prevents data from being accessed and provides a mechanism to quickly shred data.	
	SRDF/Metro	• Place new SRDF device pairs into an SRDF/ Metro configuration. • Synchronize device pairs.	
	VIPR Suite (Controller and SRM)	Automate storage provisioning and reclamation tasks to improve operational efficiency.	

Individual licenses

These items are available for arrays running HYPERMAX OS and are not included in any of the license suites:

Table 44 Individual licenses for open systems environment

License	Allows you to	With the command
ProtectPoint	Store and retrieve backup data within an integrated environment containing arrays running HYPERMAX OS and Data Domain arrays.	
RecoverPoint	Protect data integrity at local and remote sites, and recover data from a point in time using journaling technology.	

Ecosystem licenses

These licenses do not apply to arrays:

Table 45 Individual licenses for open systems environment

License	Allows you to
PowerPath	Automate data path failover and recovery to ensure applications are always available and remain operational.
Events and Retention Suite	• Protect data from unwanted changes, deletions and malicious activity. • Encrypt data where it is created for protection anywhere outside the server. • Maintain data confidentiality for selected data at rest and enforce retention at the file-level to meet compliance requirements. • Integrate with third-party anti-virus checking, quota management, and auditing applications.