Dell EMC VPLEX™ GeoSynchrony

Version 6.2

VPLEX Product Guide

Rev 01

DELL EMC

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

CONTENTS

FIGURES

TABLES

Preface

As part of an effort to improve its product lines, Dell 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 Dell EMC technical support professional if a product does not function properly or does not function as described in this document.

Note: This document was accurate at publication time. Go to Dell EMC Online Support (https://www.dell.com/support) to ensure that you are using the latest version of this document.

Purpose

This document is part of the VPLEX documentation set, and describes the VPLEX features and use cases, configuration options, VPLEX software and its upgrade, and the hardware overview.

Audience

This guide is intended for use by customers who wish to understand the software and hardware features of VPLEX, the use cases of VPLEX, product offerings, and the configuration options.

Related documents (available on Dell EMC Online Support) include:

- *VPLEX Release Notes for GeoSynchrony Releases*
- *VPLEX Product Guide*
- *VPLEX Hardware Environment Setup Guide*
- *VPLEX Configuration Worksheet*
- *VPLEX Configuration Guide*
- *VPLEX Security Configuration Guide*
- *VPLEX CLI Reference Guide*
- *VPLEX Administration Guide*
- Unisphere for VPLEX Help
- *VPLEX Element Manager API Guide Version 2 (REST API v2)*
- *VPLEX Open-Source Licenses*
- *VPLEX GPL3 Open-Source Licenses*
- Procedures provided through the SolVe Desktop
- Dell EMC Host Connectivity Guides
- *Dell EMC VPLEX Hardware Installation Guide*
- Various best practices technical notes available on Dell EMC Online Support

Special notice conventions used in this document

Dell 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

Dell EMC uses the following type style conventions in this document:

Table 1 Typographical conventions

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

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

Product information

For documentation, release notes, software updates, or information about Dell EMC products, go to Dell EMC Online Support at https://www.dell.com/support.

Technical support

Go to Dell EMC Online Support and click Support. You will see several options for contacting Dell EMC Technical Support. Note that to open a service request, you must have a valid support agreement. Contact your Dell EMC sales representative for details about obtaining a valid support agreement or with questions about your account.

Online communities

Visit Dell EMC Community Network (DECN) at https://www.dell.com/community/Dell-Community/ct-p/English for peer contacts, conversations, and content on product support and

solutions. Interactively engage online with customers, partners, and certified professionals for all Dell EMC products.

Your comments

Your suggestions will help us continue to improve the accuracy, organization, and overall quality of the user publications. Send your opinions of this document to techpubcomments@emc.com.

CHAPTER 1

Introducing VPLEX

This chapter introduces the Dell EMC VPLEX product family.

VPLEX overview

Dell EMC VPLEX federates data that is located on heterogeneous storage arrays to create dynamic, distributed and highly available data centers.

Use VPLEX to:

- Move data nondisruptively between Dell EMC and other third party storage arrays without any downtime for the host.

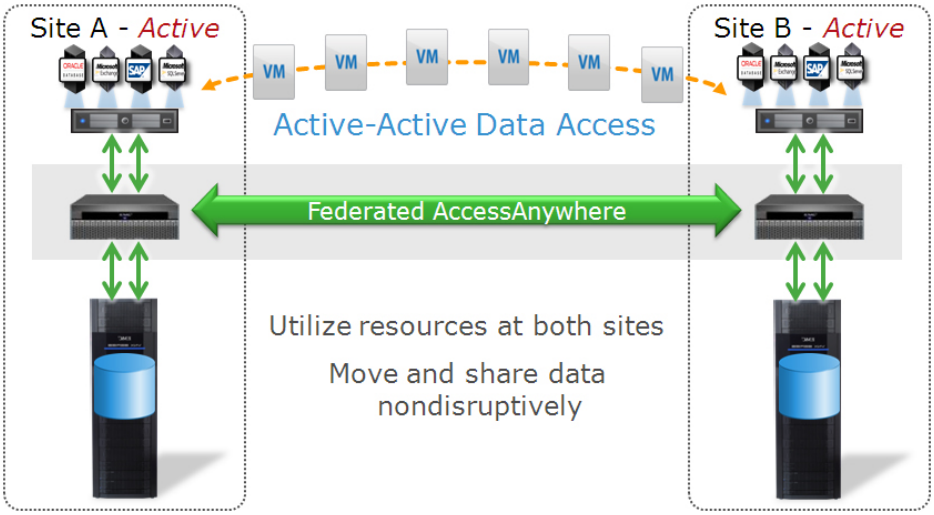
  VPLEX moves data transparently and the virtual volumes retain the same identities and the same access points to the host. There is no need to reconfigure the host.

- Protect data in the event of disasters or failure of components in your data centers.

  With VPLEX, you can withstand failures of storage arrays, cluster components, an entire site failure, or loss of communication between sites (when two clusters are deployed) and still keep applications and data online and available.

With VPLEX, you can transform the delivery of IT to a flexible, efficient, reliable, and resilient service.

Figure 1 VPLEX active-active



VPLEX addresses these two primary IT needs:

- **Mobility**:VPLEX moves applications and data between different storage installations:
  - Within the same data center or across a campus (VPLEX Local)
  - Within a geographical region (VPLEX Metro)
- **Availability:** VPLEX creates high-availability storage infrastructure across these same varied geographies with unmatched resiliency.

VPLEX offers the following unique innovations and advantages:

- VPLEX distributed/federated virtual storage enables new models of application and Data Mobility.

  VPLEX is optimized for virtual server platforms (VMware ESX, Hyper-V, Oracle Virtual Machine, AIX VIOS).

  VPLEX can streamline or accelerate transparent workload relocation over distances, including moving virtual machines.

- Size VPLEX to meet your current needs. Grow VPLEX as your needs grow.

  A VPLEX cluster includes one, two, or four engines.

  Add an engine to an operating VPLEX cluster without interrupting service.

  Add a second cluster to an operating VPLEX cluster without interrupting service.

  The scalable architecture of VPLEX ensures maximum availability, fault tolerance, and performance.
- Every engine in a VPLEX cluster can access all the virtual volumes presented by VPLEX.

  Every engine in a VPLEX cluster can access all the physical storage connected to VPLEX.
- In a Metro configuration, VPLEX AccessAnywhere provides cache-consistent active-active access to data across two VPLEX clusters.

VPLEX pools the storage resources in multiple data centers so that the data can be accessed anywhere. With VPLEX, you can:

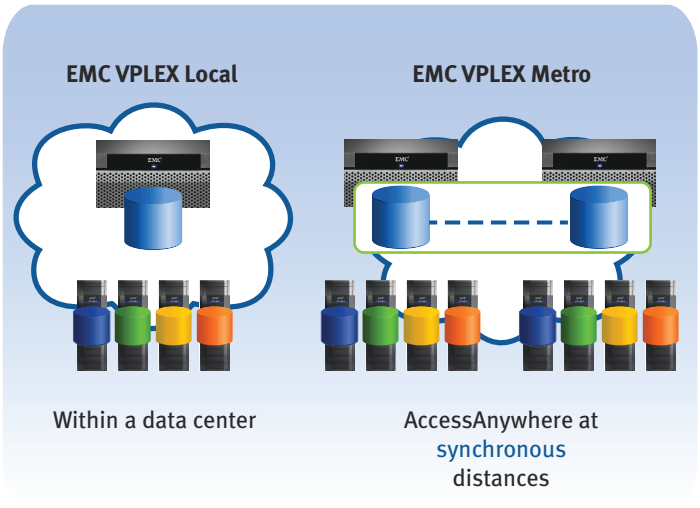
- Provide continuous availability and workload mobility.
- Replace your tedious data movement and technology refresh processes with VPLEX's patented simple, frictionless two-way data exchange between locations.
- Create an active-active configuration for the active use of resources at both sites.
- Provide instant access to data between data centers. VPLEX allows simple, frictionless two-way data exchange between locations.
- Combine VPLEX with virtual servers to enable private and hybrid cloud computing.

VPLEX product family

The VPLEX product family includes:

- VPLEX Local
- VPLEX Metro

Figure 2 VPLEX family: Local and Metro



VPLEX Local

VPLEX Local consists of a single cluster. VPLEX Local:

- Federates Dell EMC and non-Dell EMC storage arrays.

  Federation allows transparent data mobility between arrays for simple, fast data movement and technology refreshes.
- Standardizes LUN presentation and management using simple tools to provision and allocate virtualized storage devices.
- Improves storage utilization using pooling and capacity aggregation across multiple arrays.
- Increases protection and high availability for critical applications.

  Mirrors storage across mixed platforms without host resources.

  Leverage your existing storage resources to deliver increased protection and availability for critical applications.

Deploy VPLEX Local within a single data center.

VPLEX Metro

VPLEX Metro consists of two VPLEX clusters connected by inter-cluster links with not more than 10ms Round Trip Time (RTT). VPLEX Metro:

- Transparently relocates data and applications over distance, protects your data center against disaster.
  Manage all of your storage in both data centers from one management interface.
- Mirrors your data to a second site, with full access at near local speeds.

Deploy VPLEX Metro within a data center for:

- Additional virtual storage capabilities beyond that of a VPLEX Local.
- Higher availability.
  Metro clusters can be placed up to 100 km apart, allowing them to be located at opposite ends of an equipment room, on different floors, or in different fire suppression zones; all of which might be the difference between riding through a local fault or fire without an outage.

Deploy VPLEX Metro between data centers for:

- Mobility: Redistribute application workloads between the two data centers.
- Availability: Applications must keep running in the presence of data center failures.
- Distribution: One data center lacks space, power, or cooling.

Combine VPLEX Metro virtual storage and virtual servers to:

- Transparently move virtual machines and storage across synchronous distances.
- Improve utilization and availability across heterogeneous arrays and multiple sites.

Distance between clusters is limited by physical distance, by host, and by application requirements. VPLEX Metro clusters contain additional I/O modules to enable the inter-cluster WAN communication over IP or Fibre Channel.

VPLEX hardware platforms

VPLEX offers two different hardware platforms: VS2 and VS6. In a VPLEX cluster configuration, all engines must be of the same platform type.

In a VPLEX Metro deployment, both the VPLEX clusters must be of the same generation platform.

Configuration highlights

A VPLEX cluster primarily consists of:

- One, two, or four VPLEX engines.

  Each engine contains two directors.

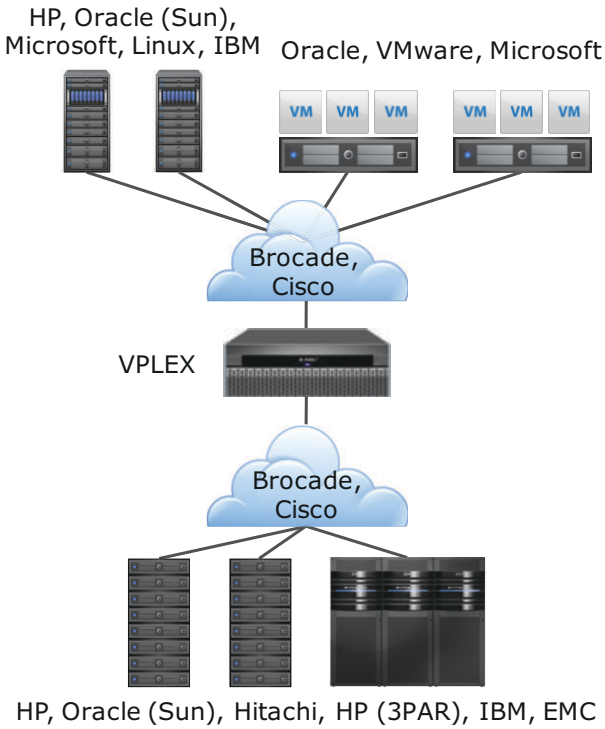
  In a VPLEX cluster, all the engines must be of VS2 platform, or of VS6 platform.

  Dual-engine or quad-engine clusters contain:

  - One pair of Fibre Channel switches on VS2 hardware, and dual InfiniBand switches on VS6 hardware, for the communication between the directors .
  - Two Uninterruptible Power Sources (UPS) for battery power backup of the switches and the management server on VS2 hardware. Two Uninterruptible Power Sources (UPS) for battery power backup of the switches on VS6 hardware.
- A management server that acts as the management interface to other VPLEX components in the cluster. The management servers in VS2 and the VS6 hardware are as follows:
  - VS2 hardware: One management server in a cluster.
  - VS6 hardware: Two management servers, which are called Management Module Control Stations (MMCS-A and MMCS-B), in the first engine. All the remaining engines will have Akula management modules for the management connectivity.

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

Figure 3 Configuration highlights



VPLEX conforms to established world wide naming (WWN) guidelines that can be used for zoning. It also supports Dell EMC storage and arrays from other storage vendors, such as HDS, HP, and IBM. VPLEX provides storage federation for operating systems and applications that support clustered file systems, including both physical and virtual server environments with VMware ESX and Microsoft Hyper-V. The network fabrics from Brocade and Cisco are supported in VPLEX.

Refer to the *Dell EMC Simple Support Matrix, Dell EMC VPLEX and GeoSynchrony*, available at http://elabnavigator.EMC.com under the Simple Support Matrix tab.

Grow your VPLEX without disruption

Deploy VPLEX to meet your current high-availability and data mobility requirements.

Add engines or a second cluster to scale VPLEX as your requirements increase. You can do all the following tasks without disrupting service:

- Add engines to a VPLEX cluster.
- Upgrade engine hardware from VS2 to VS6.
- Convert a VPLEX Local to a VPLEX Metro.
- Upgrade GeoSynchrony.
- Integrate with RecoverPoint.
- Modify an existing association with RecoverPoint.

Management interfaces

In a VPLEX Metro configuration, both clusters can be managed from either management server.

Inside VPLEX clusters, management traffic traverses a TCP/IP based private management network.

In a VPLEX Metro configuration, management traffic traverses a VPN tunnel between the management servers on both clusters.

Web-based GUI

VPLEX's web-based graphical user interface (GUI) provides an easy-to-use point-and-click management interface.

The following figures show the screen to claim storage:

Figure 4 Claim storage using the GUI (for Flash)

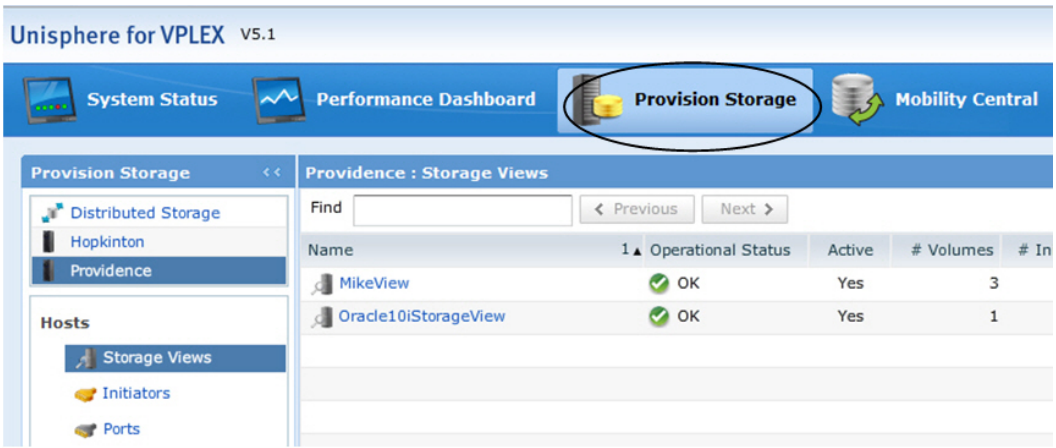
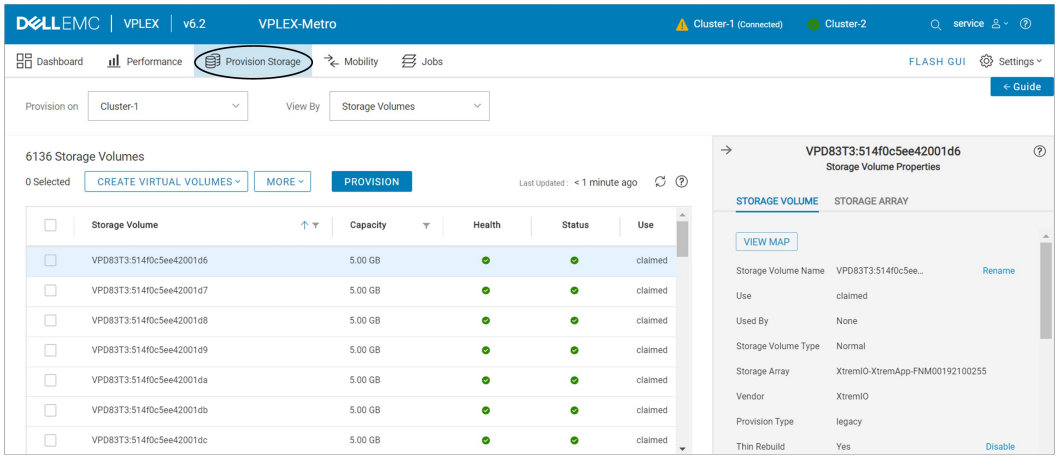


Figure 5 Claim storage using the GUI (for HTML5)



The GUI supports most of the VPLEX operations, and includes Dell EMC Unisphere for VPLEX Online help to assist new users in learning the interface.

VPLEX operations that are not available in the GUI, are supported by the Command Line Interface (CLI), which supports full functionality.

Note: Starting with VPLEX GeoSynchrony 6.2, the Flash GUI is depreciated.

VPLEX CLI

The VPLEX CLI supports all VPLEX operations.

The CLI is divided into command contexts:

- Global commands are accessible from all contexts.
- Other commands are arranged in a hierarchical context tree, and can be executed only from the appropriate location in the context tree.

The following example shows a CLI session that performs the same tasks as shown in Figure 4 on page 18.

Example 1 Claim storage using the CLI:

In the following example, the `claimingwizard` command finds unclaimed storage volumes, claims them as thin storage, and assigns names from a CLARiiON hints file:

```
VPlexcli:/clusters/cluster-1/storage-elements/
storage-volumes> claimingwizard --file /home/service/clar.txt
--thin-rebuild
Found unclaimed storage-volume
VPD83T3:6006016091c50e004f57534d0c17e011 vendor DGC:
claiming and naming clar_LUN82.
Found unclaimed storage-volume
VPD83T3:6006016091c50e005157534d0c17e011 vendor DGC:
claiming and naming clar_LUN84.
Claimed 2 storage-volumes in storage array car
Claimed 2 storage-volumes in total.
VPlexcli:/clusters/cluster-1/storage-elements/storage-volumes>
```

The *Dell EMC VPLEX CLI Guide* provides a comprehensive list of VPLEX commands and detailed instructions on using those commands.

VPLEX Element Manager API

VPLEX Element Manager API uses the Representational State Transfer (REST) software architecture for distributed systems such as the World Wide Web. It allows software developers and other users to use the API to create scripts to run VPLEX CLI commands.

VPLEX Element Manager API supports all VPLEX CLI commands that can be run from the root context.

Note: Starting with VPLEX GeoSynchrony 6.2, the REST API is depreciated. For more details, see REST API v2 guide available on Dell EMC Online Support at https://www.dell.com/support.

VPLEX licensing

Dell EMC VPLEX GeoSynchrony 6.2 requires a new software activation process under the company standard Digital Licensing Platform.

To obtain the proper license key file for installation, visit the Dell EMC Software Licensing Central Website with the License Authorization Code (LAC) associated with your acquisition. Use the **License** menu option in VPLEX UniSphere Graphical User Interface (GUI) to install the license key file to complete the activation for compliance. The license key file identifies the specific entitlement associated with the specific VPLEX cluster for proper activation. From the GUI, you can install, show license, and view the usage reports.

In the case of VPLEX usage consumption change such as adding capacity or new storage array under VPLEX management, obtain a new license key file to reflect the proper software entitlement and activation change.

In the case of VPLEX software non-disruptive upgrade (NDU) to run the latest GeoSynchrony from an older version that did not require the use of electronic license activation by file (that is, any release earlier than GeoSynchrony 6.2), contact Dell EMC Customer Support Services for assistance to complete the conversion with proper license entitlement and activation.

CHAPTER 2

VPLEX use cases

This chapter describes the general features, benefits, and the important use cases of VPLEX.

General use cases and benefits

The following table summarizes the general VPLEX use cases and their benefits.

Table 2 General VPLEX use cases and benefits

General use cases	Benefits
Mobility	• **Migration:** Move data and applications without impact on users. • **Virtual Storage federation:** Achieve transparent mobility and access within a data center and between data centers. • **Scale-out cluster architecture:** Start small and grow larger with predictable service levels.
Availability	• **Resiliency:** Mirror across arrays within a single data center or between data centers without host impact. This increases availability for critical applications. • **Distributed cache coherency:** Automate sharing, balancing, and failover of I/O across the cluster and between clusters whenever possible. • **Advanced data caching:** Improve I/O performance and reduce storage array contention.

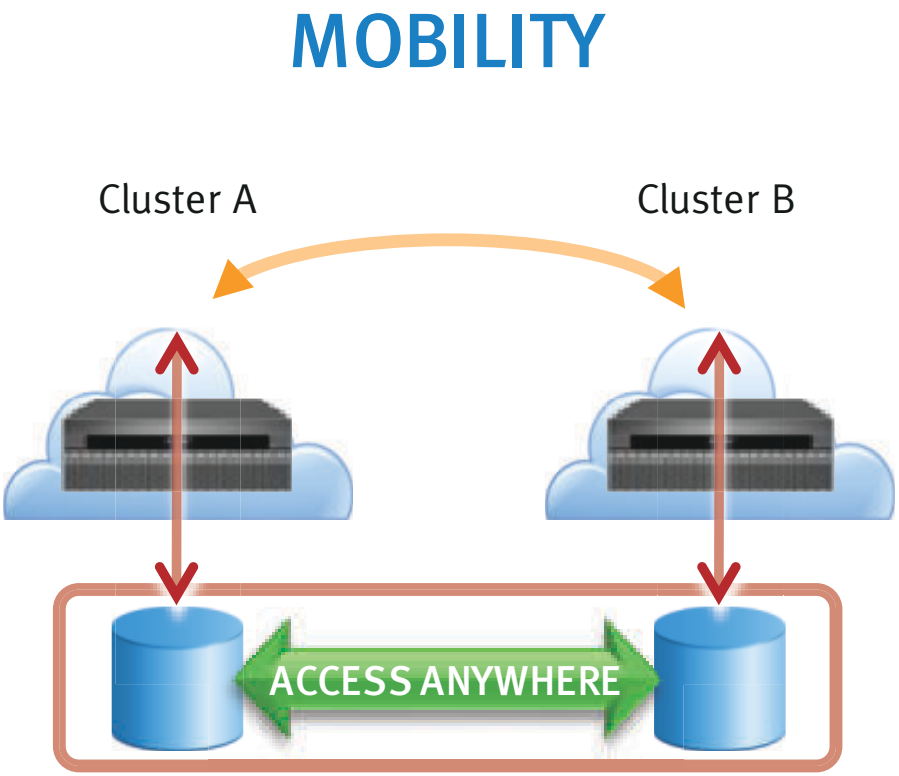
For all VPLEX deployments, GeoSynchrony performs the following:

- Presents storage volumes from back-end arrays to VPLEX engines.
- Federates the storage volumes into hierarchies of VPLEX virtual volumes with user-defined configuration and protection levels.
- Presents virtual volumes to production hosts in the SAN through the VPLEX front-end.
- For VPLEX Metro, presents a global, block-level directory for distributed cache and I/O between VPLEX clusters.

Mobility

Use VPLEX to move data between data centers, relocate a data center or consolidate data, without disrupting host application access to the data.

Figure 6 Moving data with VPLEX



The source and target arrays can be in the same data center (VPLEX Local) or in different data centers separated by up to 10ms (VPLEX Metro). The source and target arrays can be heterogeneous.

When you use VPLEX to move data, the data retains its original VPLEX volume identifier during and after the mobility operation. No change in volume identifiers eliminates application cut over. The application continues to use the same data, though the data has been moved to a different storage array.

There are many types and reasons to move data:

- Move data from a hot storage device.
- Move the data from one storage device to another without moving the application.
- Move operating system files from one storage device to another.
- Consolidate data or database instances.
- Move database instances.
- Move storage infrastructure from one physical location to another.

With VPLEX, you no longer need to spend significant time and resources preparing to move data and applications. You do not have to plan for an application downtime or restart the applications as part of the data movement activity. Instead, a move can be made instantly between sites, over distance, and the data remains online and available during the move without any outage or

downtime. Considerations before moving the data include the business impact, type of data to be moved, site locations, total amount of data, and schedules.

The data mobility feature of VPLEX is useful for disaster avoidance, planned upgrade, or physical movement of facilities. The mobility jobs in VPLEX are as follows:

Table 3 Types of data mobility operations

Extent	Moves data from one extent to another extent (within a cluster).
Device	Moves data from one device to another device (within a cluster and across clusters).
Batch	Moves data using a migration plan file. Create batch migrations to automate routine tasks. • Use batched extent migrations to migrate arrays within the same cluster where the source and destination have the same number of LUNs and identical capacities. • Use batched device migrations to migrate to dissimilar arrays and to migrate devices within a cluster and between the clusters in a VPLEX Metro configuration.

Technology refresh

In typical IT environments, migrations to new storage arrays (technology refreshes) require that the data that is being used by hosts be copied to a new volume on the new array. The host must then be reconfigured to access the new storage. This process requires downtime for the host.

VPLEX makes it easier to replace heterogeneous storage arrays on the back-end. Migrations between heterogeneous arrays can be complicated and may require additional software or functionality. Integrating heterogeneous arrays in a single environment is difficult and requires a staff with a diverse skill set.

When VPLEX is inserted between the front-end and back-end redundant fabrics, VPLEX appears as the target to hosts and as the initiator to storage.

The data resides on virtual volumes in VPLEX and it can be copied nondisruptively from one array to another without any downtime. There is no need to reconfigure the host; the physical data relocation is performed by VPLEX transparently and the virtual volumes retain the same identities and the same access points to the host.

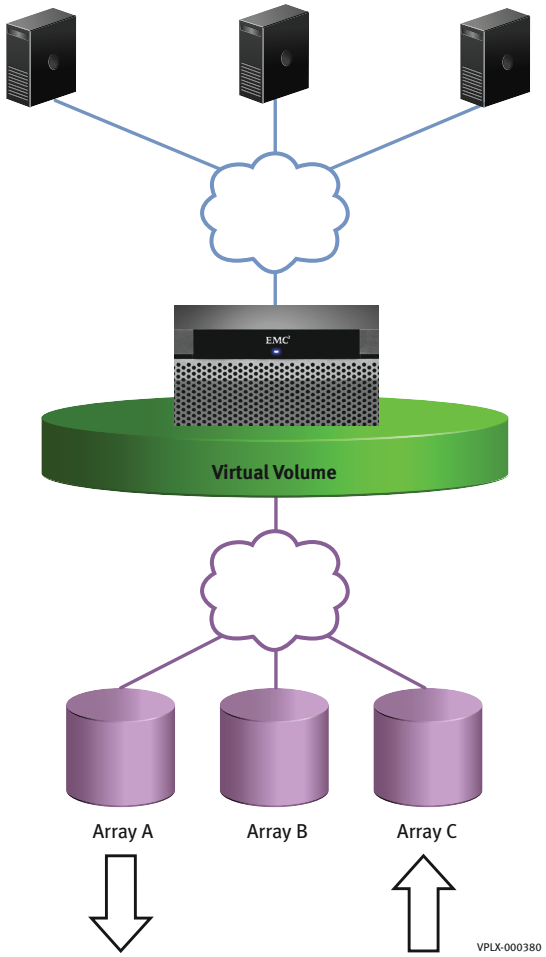
In the following figure, the virtual disk is made up of the disks of Array A and Array B. The site administrator has determined that Array A has become obsolete and should be replaced with a new array. Array C is the new storage array. Using Mobility Central, the administrator:

- Adds Array C into the VPLEX cluster.
- Assigns a target extent from the new array to each extent from the old array.
- Instructs VPLEX to perform the migration.

VPLEX copies data from Array A to Array C while the host continues its access to the virtual volume without disruption.

After the copy of Array A to Array C is complete, Array A can be decommissioned:

Figure 7 VPLEX technology refresh



Because the virtual machine is addressing its data to the abstracted virtual volume, its data continues to flow to the virtual volume without any need to change the address of the data store.

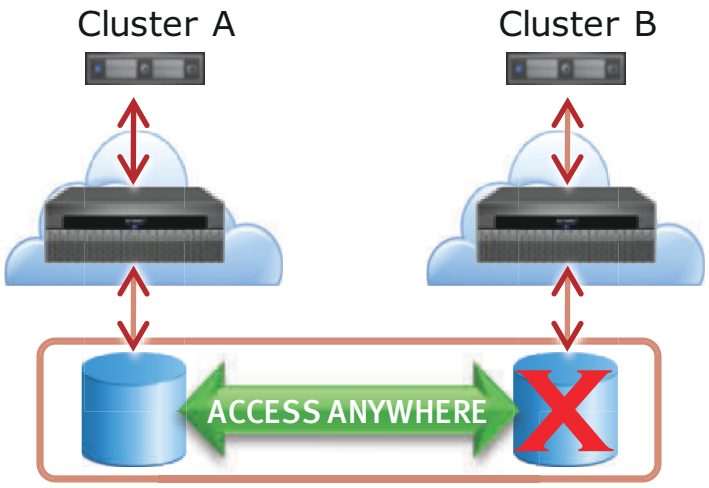
Although this example uses virtual machines, the same is true for traditional hosts. Using VPLEX, the administrator can move data used by an application to a different storage array without the application or server being aware of the change.

This allows you to change the back-end storage arrays transparently, without interrupting I/O.

Availability

VPLEX features allow the highest possible resiliency in the event of an outage. The following figure shows a VPLEX Metro configuration where storage has become unavailable at one of the cluster sites.

Figure 8 High availability infrastructure example



Maintain availability and non-stop access by mirroring across locations.
Eliminate storage operatios nfrom failover.

VPLEX redundancy provides reduced Recovery Time Objective (RTO) and Recovery Point Objective (RPO). Because VPLEX GeoSynchrony AccessAnywhere mirrors all data, applications continue without disruption using the back-end storage at the unaffected site.

With the Federated AccessAnywhere feature of VPLEX, the data remains consistent, online, and always available. VPLEX does not need to ship the entire file back and forth like other solutions. It only sends the changed updates as they are made, greatly reducing bandwidth costs and offering significant savings over other solutions.

To know more about high availability with VPLEX, see Chapter 4 Integrity and resiliency. .

Redundancy with RecoverPoint

Dell EMC RecoverPoint provides comprehensive data protection by continuous replication of host writes. With RecoverPoint, applications can be recovered to any point in time.

Replicated writes can be written to:

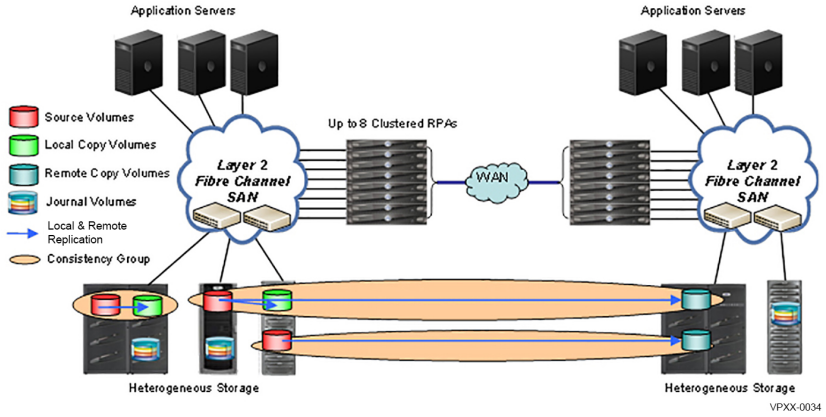
- Local volumes, to provide recovery from operational disasters.
- Remote volumes, to provide recovery from site disasters.
- Both local and remote volumes.

VPLEX GeoSynchrony includes a RecoverPoint splitter. A splitter is software that duplicates application writes so that they are sent to their normally designated volumes and RPAs simultaneously. The splitter is built into VPLEX such that the VPLEX volumes can have their I/O replicated by RecoverPoint Appliances (RPAs) to volumes that are located in VPLEX on one or more heterogeneous storage arrays.

Note: RecoverPoint integration is offered for VPLEX Local and VPLEX Metro configurations.

The VPLEX splitter works with a RecoverPoint Appliance (RPA) to orchestrate the replication of data either remotely or locally, or both.

Figure 9 RecoverPoint architecture



The VPLEX splitter enables VPLEX volumes in a VPLEX Local or VPLEX Metro to mirror I/O to a RecoverPoint Appliance.

RecoverPoint/VPLEX configurations

RecoverPoint can be configured on VPLEX Local or Metro systems as follows:

- VPLEX Local and local protection
- VPLEX Local and local/remote protection
- VPLEX Metro and RecoverPoint local at one site
- VPLEX Metro and RecoverPoint with both local and remote replications

In VPLEX Local systems, RecoverPoint can replicate local volumes.

In VPLEX Metro systems, RecoverPoint can replicate local volumes and distributed RAID 1 volumes.

Virtual volumes can be replicated locally, remotely, or both.

Distances between production sources and replication volumes vary based on the recovery objectives, inter-site bandwidth, latency, and other limitations outlined in the Dell EMC Simple Support Matrix (ESSM) for RecoverPoint.

VPLEX Local and Local Protection

In VPLEX Local with local protection configurations, I/O is split to replica volumes that are located at the same site.

RPAs are deployed with the VPLEX cluster.

This configuration supports unlimited points in time, with granularity up to a single write for local VPLEX virtual volumes. The replica volume can be a VPLEX virtual volume or any other heterogeneous storage supported by RecoverPoint.

Application event aware based rollback is supported for Microsoft SQL, Microsoft Exchange, and Oracle database applications.

Users can quickly return to any point-in-time, to recover from operational disasters.

VPLEX Local and Local/Remote Protection

In VPLEX Local with local/remote protection configurations, I/O is split to replica volumes located both at the site where the VPLEX cluster is located and a remote site.

RPAs are deployed at both sites.

If the local replication site fails, you can recover to any point in time at the remote site. Recovery can be automated through integration with MSCE and VMware SRM.

This configuration can simulate a disaster at the local site to test RecoverPoint disaster recovery features at the remote site.

Application event aware based rollback is supported for Microsoft SQL, Microsoft Exchange, and Oracle database applications.

The remote site can be an independent VPLEX cluster or, the remote site can be an array-based splitter.

VPLEX Metro and RecoverPoint local at one site

In VPLEX Metro/RecoverPoint local replication configurations, I/O is split to replica volumes located at only one VPLEX cluster. RPAs are deployed at one VPLEX cluster.

VPLEX Metro/RecoverPoint local replication configurations support unlimited points in time on VPLEX distributed and local virtual volumes.

Users can quickly return to any point-in-time, in order to recover from operational disasters.

VPLEX Metro and RecoverPoint with both Local and Remote Replication

In VPLEX Metro/RecoverPoint with both local and remote replication configurations, I/O is:

- Written to both VPLEX clusters (as part of normal VPLEX operations).
- Split on one VPLEX cluster to replica volumes located both at the cluster and at a remote site.

RPAs are deployed at one VPLEX cluster and at a third site.

The third site can be an independent VPLEX cluster or the remote site can be an array-based splitter.

This configuration supports unlimited points in time, with granularity up to a single write, for local and distributed VPLEX virtual volumes.

- RecoverPoint Appliances can (and for MetroPoint must) be deployed at each VPLEX cluster in a Metro system. For MetroPoint replication, a different RecoverPoint cluster must be viewing each exposed leg of the VPLEX distributed volume. The two RecoverPoint clusters become the active and standby sites for the MetroPoint group.
- All RecoverPoint protected volumes must be on the preferred cluster, as designated by VPLEX consistency group-level detach rules.
- Customers can recover from operational disasters by quickly returning to any point-in-time on the VPLEX cluster where the RPAs are deployed or at the third site.
- Application event aware based rollback is supported on VPLEX Metro distributed/local virtual volumes for Microsoft SQL, Microsoft Exchange, and Oracle database applications.
- If the VPLEX cluster fails, then the customers can recover to any point in time at the remote replication site. Recovery at the remote site to any point in time can be automated through integration with MSCE and VMware Site Recovery Manager (SRM).
- This configuration can simulate a disaster at the VPLEX cluster to test RecoverPoint disaster recovery features at the remote site.

Shared VPLEX splitter

The VPLEX splitter can be shared by multiple RecoverPoint clusters. This allows data to be replicated from a production VPLEX cluster to multiple RecoverPoint clusters.

Shared RecoverPoint RPA cluster

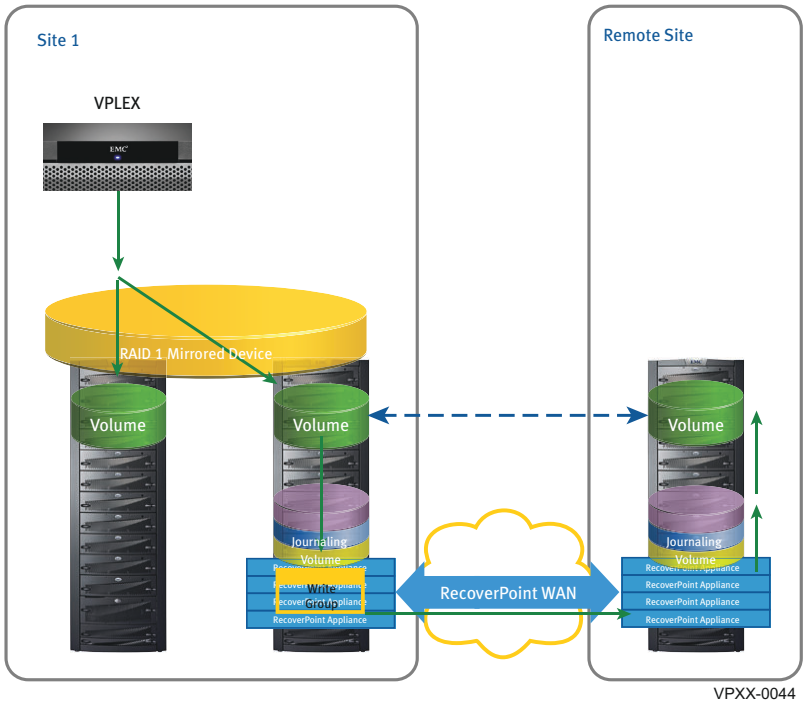
The RecoverPoint cluster can be shared by multiple VPLEX sites.

RecoverPoint replication with CLARiiON

VPLEX and RecoverPoint can be deployed in conjunction with CLARiiON based RecoverPoint splitters, in both VPLEX Local and VPLEX Metro environments.

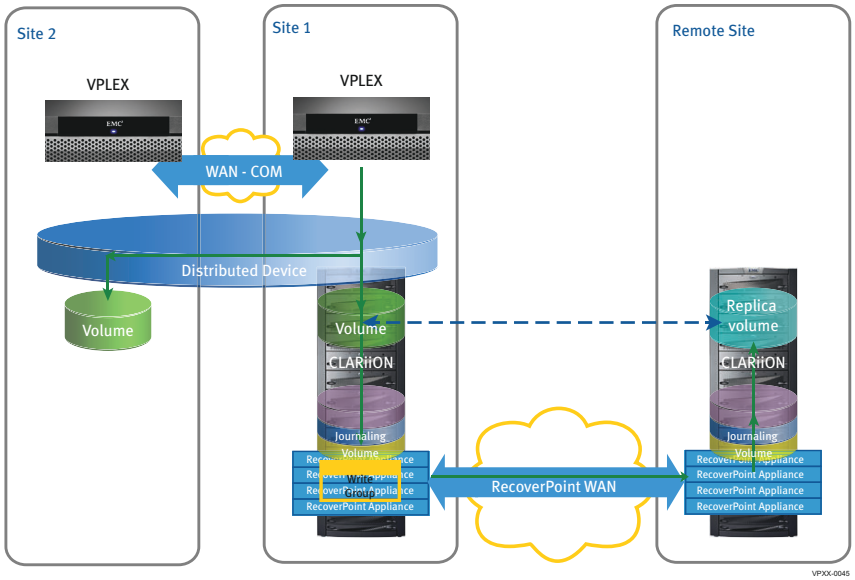
In the configuration depicted below, a host writes to VPLEX Local. Virtual volumes are written to both legs of RAID 1 devices. The VPLEX splitter sends one copy to the usual back-end storage, and one copy across a WAN to a CLARiiON array at a remote disaster recovery site.

Figure 10 Replication with VPLEX Local and CLARiiON



In the configuration depicted below, host writes to the distributed virtual volumes are written to the both legs of the distributed RAID 1 volume. Additionally, a copy of the I/O is sent to the RPA. RPA then distributes to the replica on the CLARiiON array at a remote disaster recovery site:

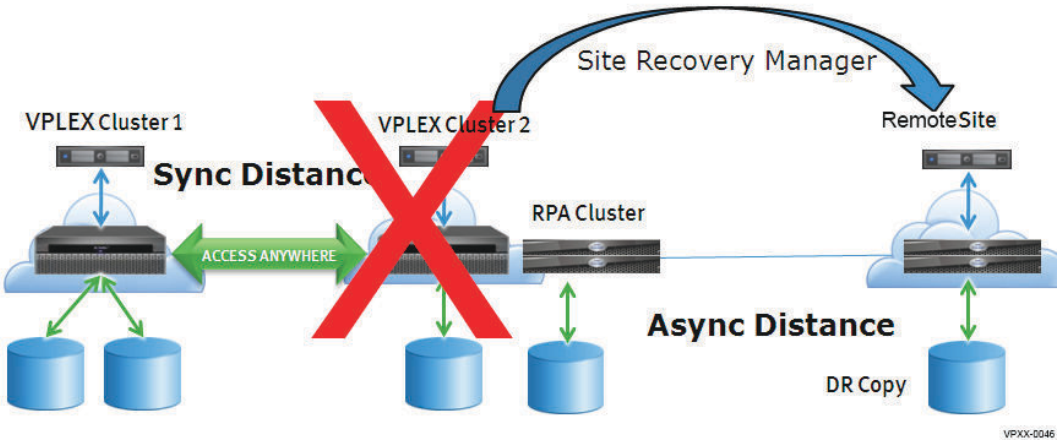
Figure 11 Replication with VPLEX Metro and CLARiiON



vCenter Site Recovery Manager support for VPLEX

With RecoverPoint replication, you can add Site Recovery Manager support to VPLEX.

Figure 12 Support for Site Recovery Manager



When an outage occurs in VPLEX Local or VPLEX Metro configurations, the virtual machines can be restarted at the replication site with automatic synchronization to the VPLEX configuration when the outage is over.

MetroPoint

VPLEX GeoSynchrony configured with RecoverPoint in a VPLEX Metro provides the MetroPoint topology. This MetroPoint topology provides a 3-site or 4-site solution for continuous availability, operational and disaster recovery, and continuous data protection. MetroPoint also supports a 2-site topology with the ability to expand to a third remote site in future.

The MetroPoint topology provides full RecoverPoint protection of both sides of a VPLEX distributed volume across both sides of a VPLEX Metro configuration, maintaining replication and

protection at a consistency group level, even when a link from one side of the VPLEX Metro to the replication site is down.

In MetroPoint, VPLEX Metro and RecoverPoint replication are combined in a fully redundant manner to provide data protection at both sides of the VPLEX Metro and at the replication site. With this solution, data is replicated only once from the active source site to the replication site. The standby source site is ready to pick up and continue replication even under a complete failure of the active source site.

MetroPoint combines the high availability of the VPLEX Metro with redundant replication and data protection of RecoverPoint. MetroPoint protection allows for one production copy of a distributed volume on each Metro site, one local copy at each Metro site, and one remote copy for each MetroPoint consistency group. Each production copy can have multiple distributed volumes.

MetroPoint offers the following benefits:

- Full high availability for data access and protection.
- Continuous data protection and disaster recovery.
- Operational recovery at all three sites for redundancy.
- Efficient data transfer between VPLEX Metro sites and to the remote site.
- Load balancing across replication links and bi-directional replication.
- Out of region data protection with asynchronous replication.
- Any-Point-in-Time operational recovery in the remote site and optionally in each of the local sites. RecoverPoint provides continuous data protection with any-point-in-time recovery.
- Full support for all operating systems and clusters normally supported with VPLEX Metro.
- Support for a large variety of Dell EMC and third-party storage arrays.

The *Dell EMC VPLEX GeoSynchrony Administration Guide* provides you additional information on the MetroPoint topologies, installation, configuration and upgrade of MetroPoint, and the failover scenarios.

CHAPTER 3

Features in VPLEX

This chapter describes the specific features of VPLEX.

VPLEX security features

The operating systems of the VPLEX management server and the directors are based on a Novell SUSE Linux Enterprise Server 11 distribution.

The operating system has been configured to meet Dell EMC security standards by disabling or removing unused services, and protecting access to network services through a firewall.

The VPLEX security features include:

- Role-based access control
- SSH Version 2 to access the management server shell
- Customizable password policies
- LDAP authentication using LDAP v3 protocol
- IPv6 support
- HTTPS to access the VPLEX GUI
- IPSec VPN inter-cluster link in a VPLEX Metro configuration
- IPSec VPN to connect each cluster of a VPLEX Metro to the VPLEX Witness server
- SCP to copy files
- Support for separate networks for all VPLEX cluster communication
- Defined user accounts and roles
- Defined port usage for cluster communication over management server
- Certificate Authority (CA) certificate (default expiration 5 years)
- Two host certificates (default expiration 2 years)
- Third host certificate for optional VPLEX Witness
- External directory server support

  CAUTION The WAN-COM inter-cluster link carries unencrypted user data. To ensure privacy of the data, establish an encrypted VPN tunnel between the two sites.

For more information about security features and configuration see the *Dell EMC VPLEX Security Configuration Guide*.

ALUA

Asymmetric Logical Unit Access (ALUA) routes I/O of the LUN directed to non-active/failed storage processor to the active storage processor without changing the ownership of the LUN.

Each LUN has two types of paths:

- **Active/optimized paths** are direct paths to the storage processor that owns the LUN.

  Active/optimized paths are usually the optimal path and provide higher bandwidth than active/non-optimized paths.
- **Active/non-optimized paths** are indirect paths to the storage processor that does not own the LUN through an interconnect bus.

  I/Os that traverse through the active/non-optimized paths must be transferred to the storage processor that owns the LUN. This transfer increases latency and has an impact on the array.

VPLEX detects the different path types and performs round robin load balancing across the active/optimized paths.

VPLEX supports all three flavors of ALUA:

- **Explicit ALUA** - The storage processor changes the state of paths in response to commands (for example, the Set Target Port Groups command) from the host (the VPLEX backend).

  The storage processor must be explicitly instructed to change a path's state.

  If the active/optimized path fails, VPLEX issues the instruction to transition the active/non-optimized path to active/optimized.

  There is no need to failover the LUN.

- **Implicit ALUA** -The storage processor can change the state of a path without any command from the host (the VPLEX back end).

  If the controller that owns the LUN fails, the array changes the state of the active/non-optimized path to active/optimized and fails over the LUN from the failed controller.

  On the next I/O, after changing the path's state, the storage processor returns a Unit Attention "Asymmetric Access State Changed" to the host (the VPLEX backend).

  VPLEX then re-discovers all the paths to get the updated access states.

- **Implicit/explicit ALUA** - Either the host or the array can initiate the access state change.

Storage processors support implicit only, explicit only, or both.

Provisioning with VPLEX

VPLEX allows easy storage provisioning among heterogeneous storage arrays. Use the web-based GUI to simplify everyday provisioning or create complex devices.

There are three ways to provision storage in VPLEX:

- Integrated storage provisioning (VIAS—VPLEX Integrated Array Services based provisioning)
- EZ provisioning
- Advanced provisioning

All provisioning features are available in the Unisphere for VPLEX GUI.

Integrated storage

Integrated storage refers to storage created through the VPLEX Integrated Services feature. This feature requires the use of Array Management Providers (AMPs) to leverage functionality on the array, specifically on storage pools. If your array functionality includes storage pools and the array is supported for use with VPLEX, you can integrate the array with VPLEX and provision storage from pools on the array through VPLEX. VIAS uses the Storage Management Initiative - Specification (SMI-S) provider and the VPLEX Restful APIs to communicate with the arrays that support integrated services to enable provisioning.

Before provisioning from storage pools, you must first register the AMP that manages the array. Your VPLEX system can include AMPs that manage some or all of the arrays in your environment. An array must be integrated with VPLEX in order to provision storage from pools on the array. Note that you can provision from storage volumes and also on an integrated array.

For more information about registering AMPs and provisioning from storage pools, refer to the provisioning chapter of the *VPLEX Administration Guide*.

Other storage

Other storage refers to storage from arrays that are not integrated with VPLEX through AMPs. Because VPLEX cannot access functionality on the array, you cannot use array functionality such as storage pools. Therefore, you can only provision from storage volumes discovered on the array.

There are two ways to provision from storage volumes: EZ-Provisioning and advanced provisioning.

Support for thin volumes and unmapping

Thin provisioning advertises the VPLEX virtual volumes as thin volumes to the hosts. Thin provisioning dynamically allocates block resources only when they are required. It essentially allows efficient utilization of physical block resources from the storage arrays.

Hosts gather the properties related to the thin provisioning feature of a VPLEX virtual volume and send SCSI commands to free storage block resources that are not in use. If the blocks of the back end storage volumes are free, the blocks can be mapped to other changed regions. Thin provisioning enables dynamic freeing of storage blocks on storage volumes for which thin provisioning is supported.

Note: The *Dell EMC Simplified Support Matrix for VPLEX* provides more information on the supported storage volumes.

VPLEX thin provisioning support includes the following features:

- Discovery of the back-end storage volumes capable for thin provisioning - During the back-end storage volume discovery, VPLEX gathers all thin provisioning related storage volume properties. VPLEX also performs a consistency check on all the properties related to thin-provisioning.
- Reporting thin provisioning enabled VPLEX virtual volumes to hosts - VPLEX shares the details of the thin provisioning-enabled virtual volumes with the hosts.
- Reclaiming the unused storage blocks - Through a command, VPLEX removes the mapping between a deleted virtual machine and its storage volumes and reclaims the storage blocks corresponding to the VMFS blocks used by that virtual machine.
- Handling storage exhaustion - The exhaustion of storage blocks on non-mirrored storage volumes are notified to the host as a space allocation failure. This error notification is posted to the host and the VMware hosts stop the impacted virtual machine.

To prevent potential mapping of all the blocks in the storage volumes that are thin capable, VPLEX uses thin rebuilds. Thin rebuilds can be configured to be set or unset for any claimed storage volume on which VPLEX builds virtual volumes. This property controls how VPLEX does its mirror rebuilding.

The unmap feature reclaims the unused VMFS blocks by removing the mapping between the logical blocks and the physical blocks. This essentially removes the link between a logical block and a physical block that has unknown or unused resources.

Performance monitoring

VPLEX performance monitoring provides a customized view into the performance of your system. You decide which aspects of the system's performance to view and compare.

You can view and assess the VPLEX performance using these methods:

- Unisphere Performance Monitoring Dashboard, which shows real-time performance monitoring data for up to one hour of history.
- Performance statistics collection using the CLI and the API. These methods let you collect and view the statistics, and export them to an external application for analysis.
- Monitoring with Simple Network Management Protocol (SNMP).

Unisphere Performance Monitoring Dashboard

The Unisphere Performance Monitoring Dashboard supports these general categories of performance monitoring:

- Current load monitoring that allows administrators to watch CPU load during upgrades, I/O load across the inter-cluster WAN link, and front-end against the back-end load during data mining or back up.
- Long term load monitoring that collects data for capacity planning and load balancing.
- Object-based monitoring that collects data for the virtual volume.

The Unisphere Performance Monitoring Dashboard is a customized view into the performance of the VPLEX system:

Figure 13 Unisphere Performance Monitoring Dashboard (for Flash)

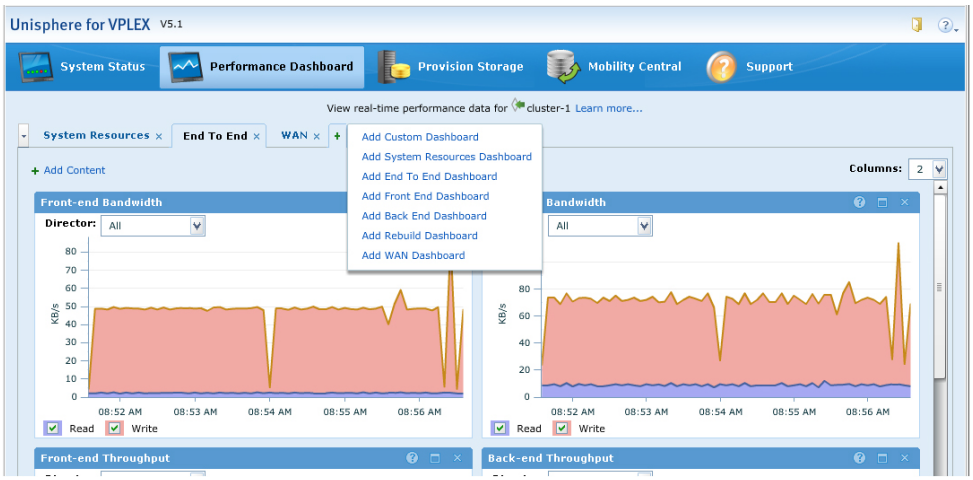
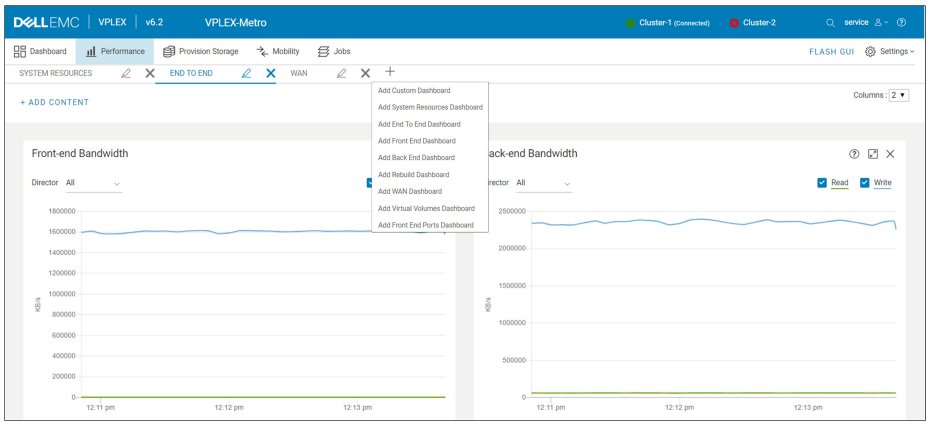


Figure 14 Unisphere Performance Monitoring Dashboard (for HTML5)



You decide which aspects of the system's performance to view and compare:

Figure 15 Unisphere Performance Monitoring Dashboard - select information to view (for Flash)

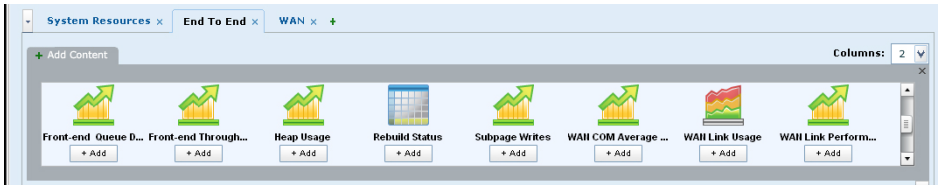
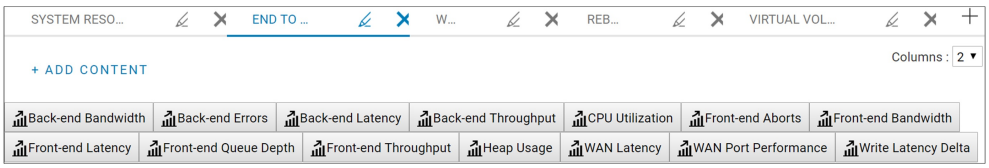


Figure 16 Unisphere Performance Monitoring Dashboard - select information to view (for HTML5)



Performance information is displayed as a set of charts. For example, the following figure shows front-end throughput for a selected director (for Flash) and all directors (for HTML5):

Figure 17 Unisphere Performance Monitoring Dashboard - sample chart (for Flash)

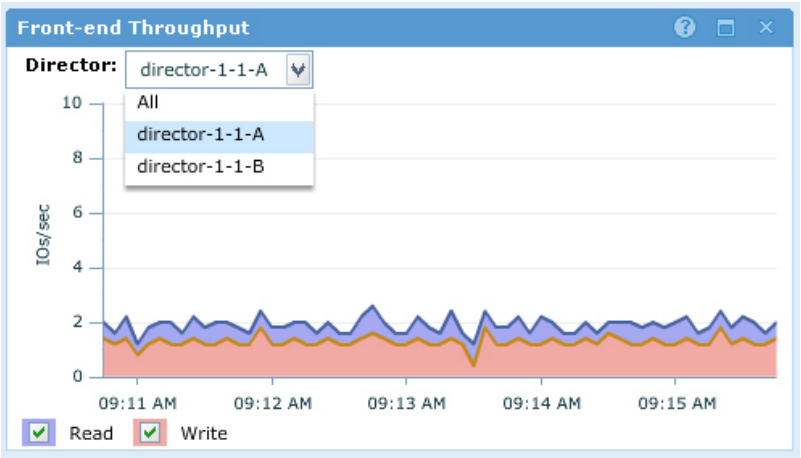
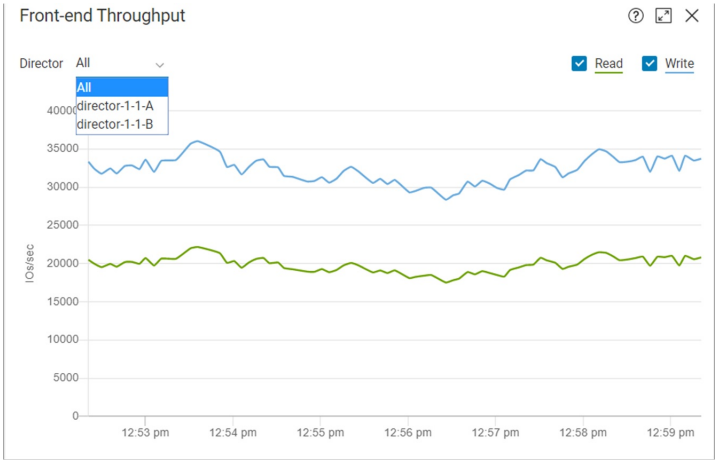


Figure 18 Unisphere Performance Monitoring Dashboard - sample chart (for GUI)



For additional information about the statistics available through the Performance Monitoring Dashboard, see the Dell EMC Unisphere for VPLEX online help available in the VPLEX GUI.

Performance monitoring using the CLI

The CLI supports current load monitoring, long term load monitoring, object base monitoring, and troubleshooting monitoring. The CLI collects and displays performance statistics using:

monitors - Gather the specified statistic from the specified target at the specified interval.

monitor sinks - Direct the output to the desired destination. Monitor sinks include the console, a file, or a combination of the two.

Use the three pre-configured monitors for each director to collect information to diagnose common problems.

Use the CLI to create a toolbox of custom monitors to operate under varying conditions including debugging, capacity planning, and workload characterization. For example:

- Create a performance monitor to collect statistics for CompareAndWrite (CAW) operations, miscompares, and latency for the specified virtual volume on director-1-1-B.
- Add a file sink to send output to the specified directory on the management server.

The *Dell EMC VPLEX Administration Guide* describes the procedure for monitoring VPLEX performance using the CLI.

VPLEX Performance Monitor

The VPLEX Performance Monitor is a stand-alone, customer installable tool that allows you to collect virtual volume metrics from a VPLEX Local or VPLEX Metro system. It allows Storage Administrators to see up to 30 days of historical virtual volume performance data to troubleshoot performance issues and analyze performance trends.

The VPLEX Performance Monitor tool is delivered as an OVA (Open Virtualization Format Archive) file that you deploy as a VMware virtual appliance. The virtual appliance connects to one VPLEX system and collects performance metrics for all virtual volumes that are in storage views. Historical virtual volume metrics are stored in a database within the virtual appliance for 30 days. The virtual appliance has a web application to view the data in charts. The charts show all 30 days of data at once and allows you to zoom in on data down to one minute.

The VPLEX Performance Monitor charts the following key virtual volume metrics:

- Throughput (total read and write IOPS)
- Read Bandwidth (KB/s)
- Write Bandwidth (KB/s)
- Read Latency (usec)
- Write Latency (usec)

Digital licensing

The digital licensing feature of Unisphere allows you to install and list details of a license, as well as usage statistics.

Using the Unisphere GUI, you can perform these tasks relating to product licenses and associated features.

- Install License - Allows you to Install a license on a local cluster in order to use product. The license key is included in the installed product and valid only on a specific cluster. A license can only be used on the cluster on which it is installed.

- Show License Details - Displays features of the installed license or license specified by the license file. Examples of features displayed are Local/Metro Capacity, 1TO1 Frame, 180 Day Migration, and so on.
- Usage Intelligence - Displays details relating to the usage of license features on the local cluster. Examples of usage details displayed are licensed capacity, capacity used, and the alerts when the usage exceeds the licensed capacity for a specific license feature. Usage Intelligence is displayed only for the VPLEX_LOCAL_CAPACITY,VPLEX_METRO_CAPACITY, VPLEX_LOCAL_LARGE_CAPACITY, and VPLEX_METRO_LARGE_CAPACITY licenses.

HTML5 based new GUI

Starting with VPLEX GeoSynchrony 6.2, HTML5 based new GUI is introduced which has Dell Clarity standards, and it is compatible with the latest browsers. The new HTML5 based GUI is based on a new REST API v2 which is available. For more information, see the *GUI Help Pages*.

Note: Starting with VPLEX GeoSynchrony 6.2, the Flash GUI is depreciated.

CHAPTER 4

Integrity and resiliency

This chapter describes how the high availability and the redundancy features of VPLEX provide robust system integrity and resiliency.

About VPLEX resilience and integrity

With VPLEX, you get true high availability. Operations continue and data remains online even when a failure occurs. Within synchronous distances (VPLEX Metro), think of VPLEX as providing disaster avoidance instead of just disaster recovery.

VPLEX Metro provides shared data access between sites. The same data (not a copy), exists at more than one location simultaneously. VPLEX can withstand a component failure, a site failure, or loss of communication between sites and still keep the application and data online and available. VPLEX clusters are capable of surviving any single hardware failure in any subsystem within the overall storage cluster, including host connectivity and memory subsystems. A single failure in any subsystem does not affect the availability or integrity of the data.

VPLEX redundancy creates fault tolerance for devices and hardware components that continue operation as long as one device or component survives. This highly available and robust architecture can sustain multiple device and component failures without disrupting service to I/O.

Failures and events that do not disrupt I/O include:

- Unplanned and planned storage outages
- SAN outages
- VPLEX component failures
- VPLEX cluster failures
- Data center outages

To achieve high availability, you must create redundant host connections and supply hosts with multi path drivers.

Note: In the event of a front-end port failure or a director failure, hosts without redundant physical connectivity to a VPLEX cluster and without multi-pathing software installed could be susceptible to data unavailability.

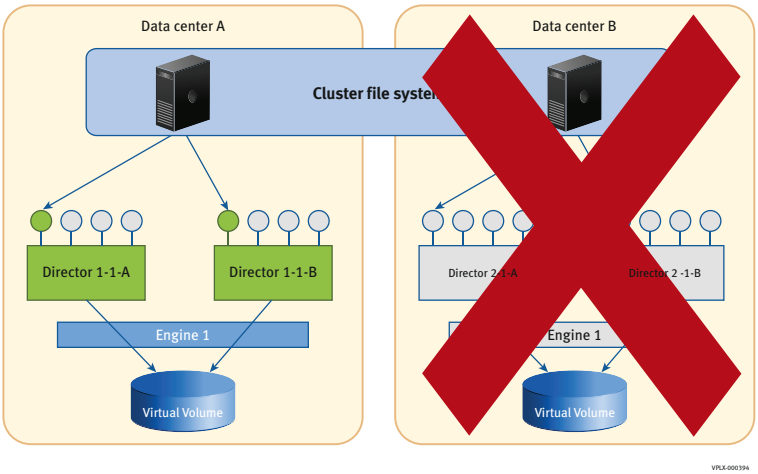
Site distribution

When two VPLEX clusters are connected together with VPLEX Metro, VPLEX gives you shared data access between sites. VPLEX can withstand a component failure, a site failure, or loss of communication between sites and still keep the application and data online and available.

VPLEX Metro ensures that if a data center goes down, or even if the link to that data center goes down, the other site can continue processing the host I/O.

In the following figure, despite a site failure at Data Center B, I/O continues without disruption in Data Center A.

Figure 19 Path redundancy: different sites



Cluster

VPLEX is a true cluster architecture. That is, all components are always available and I/O that enters the cluster from anywhere can be serviced by any node within the cluster, while cache and coherency is maintained for all reads and writes.

As you add more engines to the cluster, you get the added benefits of more cache, increased processing power, and more performance.

A VPLEX cluster provides N–1 fault tolerance, which means that any component failure can be sustained, and the cluster will continue to operate as long as one director survives.

A VPLEX cluster (running either on VS2 hardware or VS6 hardware) consists of redundant hardware components.

A single engine supports two directors. If one director in an engine fails, the second director in the engine continues to service I/O. Similarly, if a VPLEX cluster contains multiple engines, VPLEX can handle more than one failure without disrupting any services as long as quorum (defined by set rules) is not lost.

All hardware resources (CPU cycles, I/O ports, and cache memory) are pooled.

A two-cluster configuration (Metro) offers true high availability. Operations continue and data remains online even if an entire site fails. It also provides a high availability solution with zero recovery point objective (RPO).

Quorum

Quorum refers to the minimum number of directors required for the cluster to service and maintain operations.

There are different quorum rules for a cluster to become operational and start servicing I/Os when it is booting up, also called “gaining quorum.” Different rules for an operational cluster seeing director failures to either continue servicing operations and I/O after failure handling is called “maintaining quorum.” Stopping servicing operations and I/O is called “losing quorum.” These rules are described below:

- **Gaining quorum** - A non-operational VPLEX cluster gains quorum and becomes operational when more than half of the configured directors restart and come in contact with each other. In a single engine cluster, it refers to all the directors.
- **Maintaining quorum** - An operational VPLEX cluster seeing failures will continue operating in the following scenarios:
  - Director failures
    - If less than half of the operational directors with quorum fail.
    - If half of the operational directors with quorum fail, then the remaining directors will check the operational status of the failed directors over the management network and remain alive.

    After recovering from this failure, a cluster can tolerate further similar director failures until only one director is remaining. In a single engine cluster, a maximum of one director failure can be tolerated.
  - Intra-cluster communication failure
    - If there is a split in the middle, that is, half of the operational directors with quorum lose communication with the other half of the directors, and both halves are running, then the directors detect the operational status over the management network and instruct half with the director with the lowest UUID to keep running and the directors without the lowest UUID to operationally stop.
- **Quorum loss** - An operational VPLEX cluster seeing failures stops operating in the following scenarios:
  - If more than half of the operational directors with quorum fail at the same time.
  - If half of the operational directors with quorum fail, and the directors are unable to determine the operation status of the other half of the directors (whose membership includes a low UUID).
  - In a dual or quad engine cluster, if all of the directors loose contact with each other.

Metadata volumes

Meta-volumes store VPLEX metadata, including virtual-to-physical mappings, data about devices, virtual volumes, and system configuration settings.

Metadata is stored in cache and backed up on specially designated external volumes called meta-volumes.

After the meta-volume is configured, updates to the metadata are written to both the cache and the meta-volume when the VPLEX configuration is modified.

Each VPLEX cluster maintains its own metadata, including:

- The local configuration for the cluster.
- Distributed configuration information shared between clusters.

At system startup, VPLEX reads the metadata and loads the configuration information onto each director.

When you make changes to the system configuration, VPLEX writes these changes to the metadata volume.

If VPLEX loses access to the metadata volume, the VPLEX directors continue uninterrupted, using the in-memory copy of the configuration. VPLEX blocks changes to the system until access is restored or the automatic backup meta-volume is activated.

Meta-volumes experience high I/O only during system startup and upgrade.

I/O activity during normal operations is minimal.

Backup metadata volumes

Backup metadata volumes are point-in-time snapshots of the current metadata, and provide extra protection before major configuration changes, refreshes, or migrations.

Backup creates a point-in-time copy of the current in-memory metadata without activating it. You must create a backup metadata volume in any of these conditions:

- As part of an overall system health check before a major migration or update.
- If VPLEX permanently loses access to active meta-volumes.
- After any major migration or update.

Logging volumes

Logging volumes keep track of blocks written:

- During an inter-cluster link outage.
- When one leg of a DR1 becomes unreachable and then recovers.

After the inter-cluster link or leg is restored, the VPLEX system uses the information in logging volumes to synchronize the mirrors by sending only changed blocks across the link.

Logging volumes also track changes during loss of a volume when that volume is one mirror in a distributed device.

CAUTION If no logging volume is accessible, then the entire leg is marked as out-of-date. A full re-synchronization is required once the leg is reattached.

The logging volumes on the continuing cluster experience high I/O during:

- Network outages or cluster failures
- Incremental synchronization

  When the network or cluster is restored, VPLEX reads the logging volume to determine what writes to synchronize to the reattached volume.

There is no I/O activity during normal operations.

Global cache

Memory systems of individual directors ensure durability of user and critical system data. Synchronous systems (write-through cache mode) leverage the back-end array by writing user data to the array. An acknowledgment for the written data must be received before the write is acknowledged back to the host.

High availability and VPLEX hardware

VPLEX supports two types of hardware: VS2 and VS6. The architectural design of the VPLEX hardware environment supports high availability.

The VPLEX hardware is largely designed to withstand technical failures and provide uninterrupted data availability. The critical components in the hardware are redundant to ensure that the failure of a component does not bring the system down.

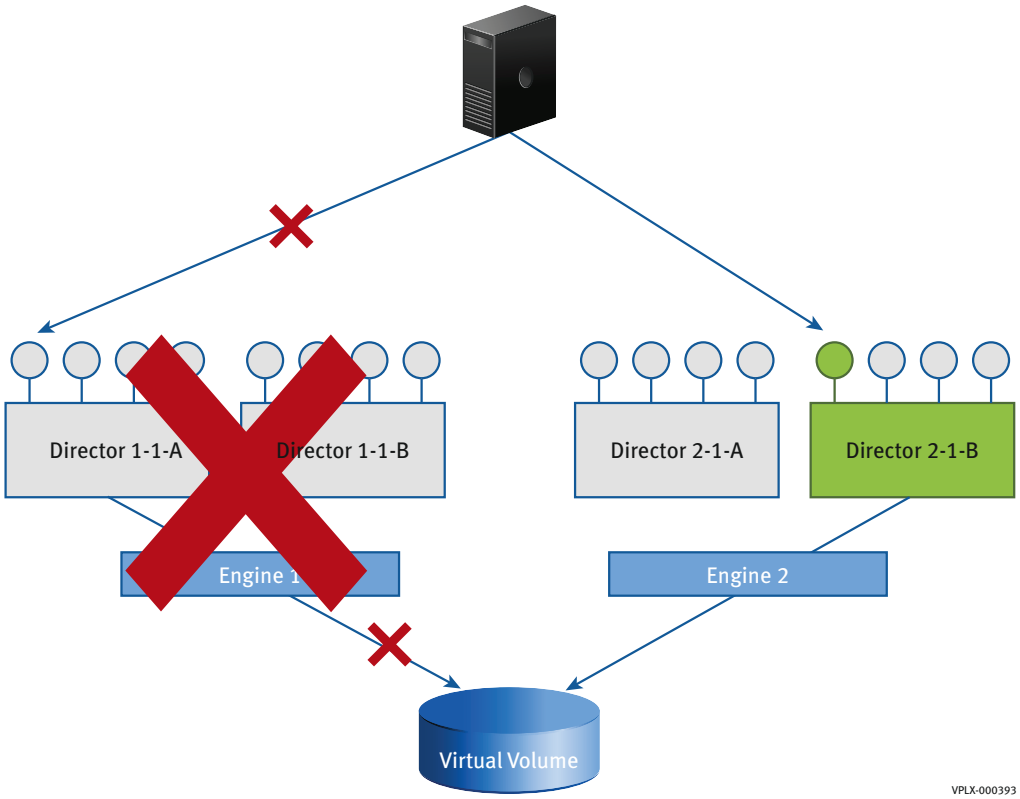
VPLEX engines

A VPLEX engine contains two directors, I/O modules, fans, and redundant power supplies. A VPLEX cluster can have one (single), two (dual), or four (quad) engines. A cluster that has multiple engines uses redundant network switches for the intra-cluster communication. Each switch is backed by a dedicated uninterruptible power supply (UPS). The directors provide redundant front-end and back-end I/O connections. A redundant standby power supply provides battery backup to each engine in the events of power outages.

Note: In a cluster that runs on VS6 hardware, the first engine contains the Management Module Control Stations (MMCS-A and MMCS-B), which is the management entity in VS6 hardware.

In a dual-engine or quad-engine configuration, if one engine goes down, another engine completes the host I/O processing as shown in the following figure.

Figure 20 Path redundancy: different engines



In a VPLEX Metro configuration, multi-pathing software plus volume presentation on different engines yields continuous data availability in the presence of engine failures.

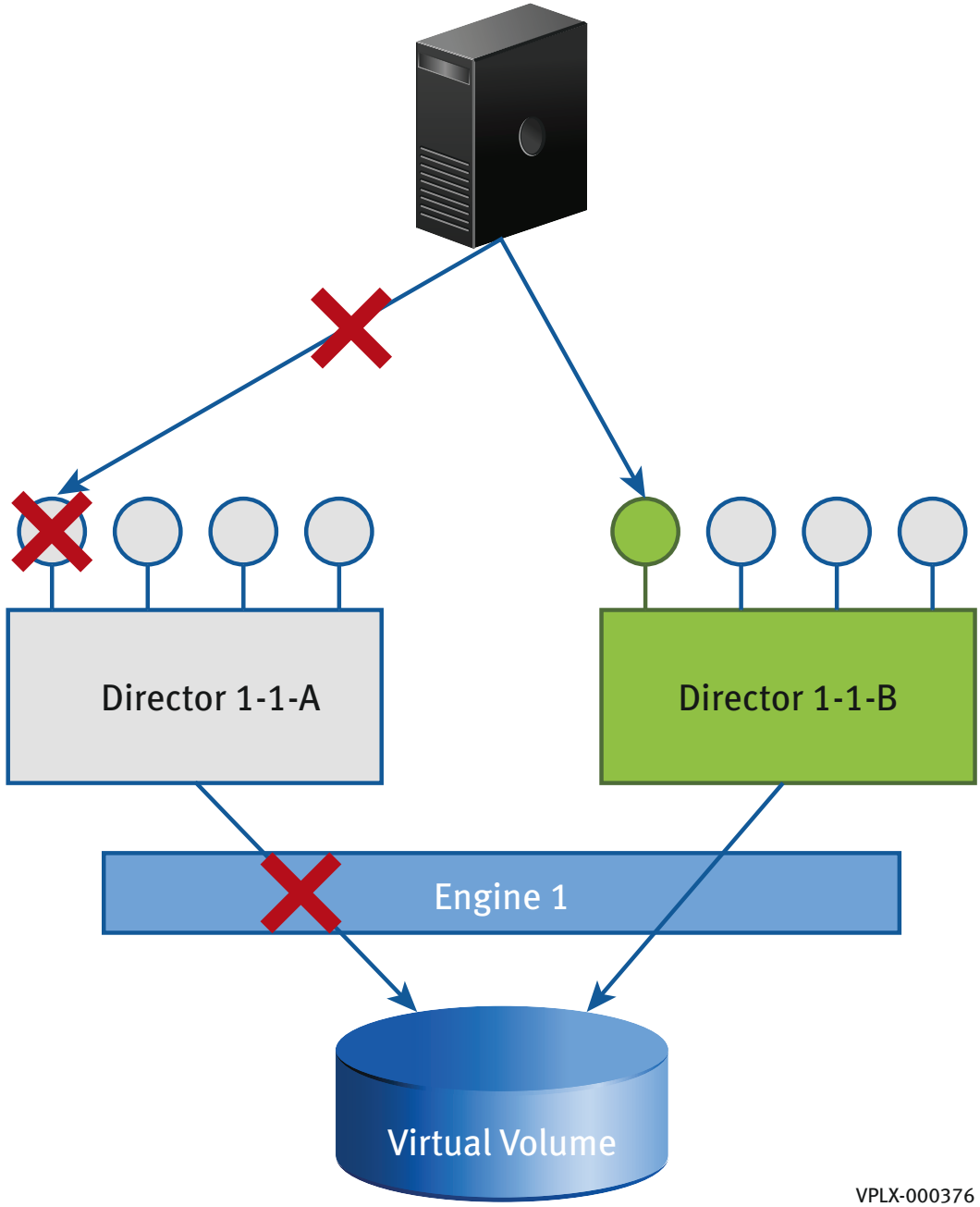
Directors

A VPLEX director is the component that process the I/O requests from the hosts in a VPLEX environment. It interacts with the backend storage arrays for servicing the I/Os.

A director has two I/O modules for servicing I/Os from the arrays; one for the connectivity with the storage arrays on the back end, and another for connecting to the hosts on the front end. The management module in the director is used for management connectivity to the directors and for intra-cluster communication. The local communication module is completely dedicated to intra-cluster communication.

The front-end ports on all directors can provide access to any virtual volume in the cluster. Include multiple front-end ports in each storage view to protect against port failures. When a director port fails, the host multi-pathing software seamlessly fails over to another path through a different port, as shown in the following figure:

Figure 21 Path redundancy: different ports



Combine multi-pathing software plus redundant volume presentation for continuous data availability in the presence of port failures.

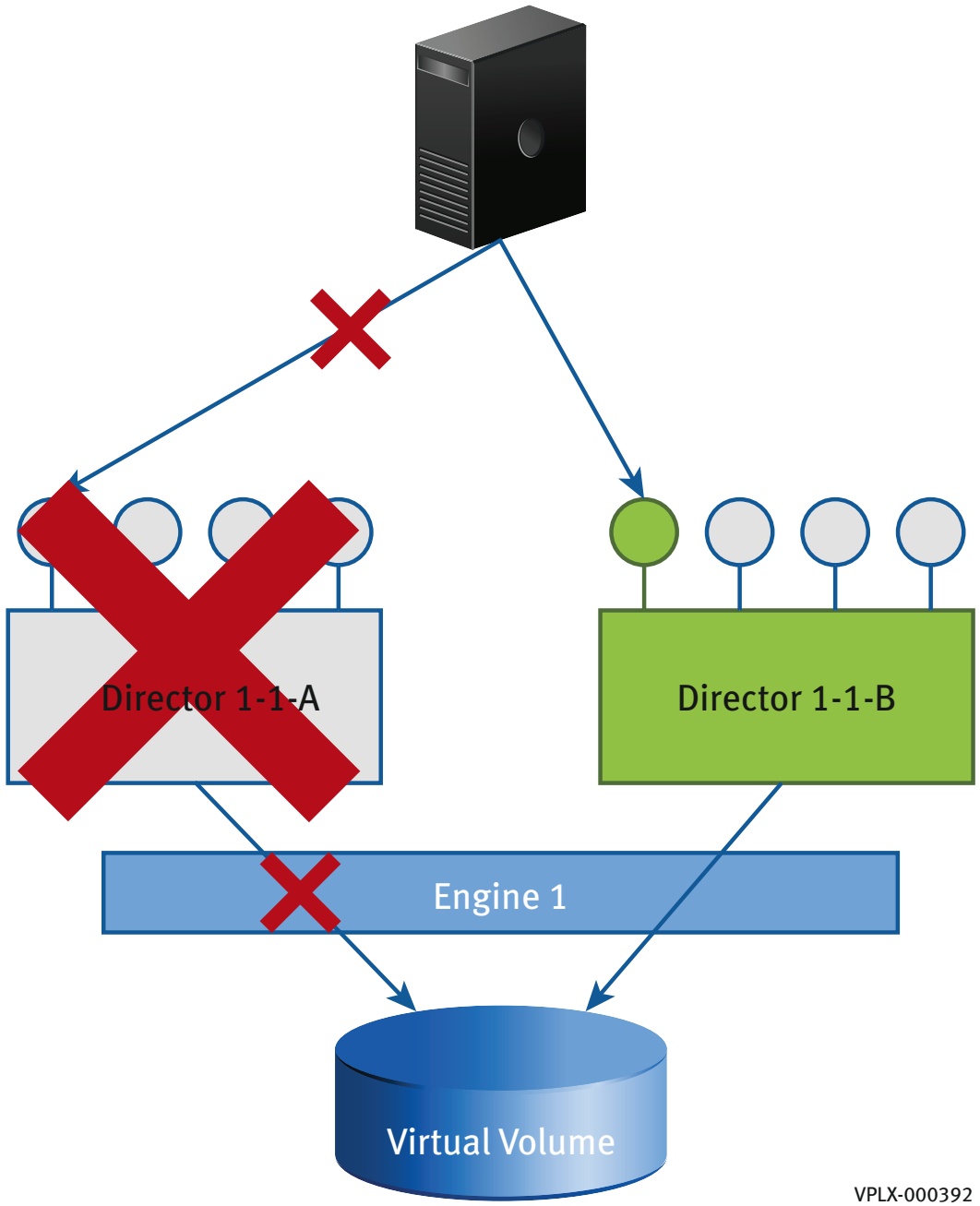
Back-end ports, local COM ports, and WAN COM ports provide similar redundancy for additional resilience.

Each VPLEX engine includes redundant directors. Each director can service I/O for any other director in the cluster due to the redundant nature of the global directory and cache coherency.

If one director in the engine fails, another director continues to service I/O from the host.

In the following figure, Director 1-1-A has failed, but Director 1-1-B services the host I/O that was previously being serviced by Director 1-1-A.

Figure 22 Path redundancy: different directors



Management server

Each VPLEX cluster has one management server. You can manage both clusters in a VPLEX Metro configuration from a single management server. The management server acts as a management interfaces to other VPLEX components in the cluster. Redundant internal network IP interfaces connect the management server to the public network. Internally, the management server is on a dedicated management IP network that provides accessibility to all major components in the cluster.

The larger role of the management server includes:

- Coordinating data collection, VPLEX software upgrades, configuration interfaces, diagnostics, event notifications, and some director-to-director communication.
- Forwarding VPLEX Witness traffic between directors in the local cluster and the remote VPLEX Witness server.

The management servers in VS2 and the VS6 hardware are as follows:

- VS2 hardware: One management server in a cluster
- VS6 hardware: Two management servers, MMCS-A and MMCS-B in the first engine. All the remaining engines will have Akula management modules for the management connectivity.

Switches for communication between directors

The network switches provide high availability and redundant connectivity between directors and engines in a dual-engine or quad-engine VPLEX cluster.

Each network switch is powered by a UPS, and has redundant I/O ports for the communication between the directors. These switches do not connect to the front-end hosts or back-end storage.

The network switches used for the intra-cluster communication (communication between directors) in the VS2 and the VS6 hardware are as follows:

- VS2: One pair of Fibre Channel switches
- VS6: Dual InfiniBand (IB) switches

Power supplies

The VPLEX cluster is connected to your AC power source. Each cluster is equipped with ample uninterruptible power supplies that enable the cluster to ride out power disruptions. The power supply units differ between the VS2 and the VS6 hardware.

The power supply architecture of a VPLEX cluster that runs on VS2 hardware contains:

- Power distribution panels (PDPs) that connect to the AC power source of the cluster and transfer power to the VPLEX components through power distribution units (PDUs). This provides a centralized power interface and distribution control for the power input lines. The PDPs contain manual on/off power switches for their power receptacles.
- PDUs that connect to the AC power source of the cluster to supply power to the VPLEX components through SPS. The PDUs contain circuit breakers, which protects the hardware components from power fluctuations.
- Standby power supply (SPS) units that have sufficient capacity to ride through transient site power failures. A single standby power supply provides enough power for the attached engine to ride through two back-to-back 5-minute losses of power. One SPS assembly (two SPS modules) provides backup power to each engine in the event of an AC power interruption. Each SPS module maintains power for two five-minute periods of AC loss while the engine shuts down.
- Two uninterruptible power supplies, UPS-A and UPS-B for the dual and quad engine clusters. One UPS provides battery backup for the first network switch and the management server, and a second UPS provides battery backup for the other switch. Each UPS module maintains power for two five-minute periods of AC loss while the engine shuts down.

The power supply architecture of a VPLEX cluster that runs on VS6 hardware contains:

- PDUs that connect to the AC power source of the cluster to supply power to the VPLEX components through Power Supply Units (PSU). The PDUs contain circuit breakers, which protect the hardware components from power fluctuations.

- Backup Battery Units (BBU) that have sufficient capacity to ride through transient site power failures. One BBU assembly (two BBU modules) provides backup power to each engine in the event of an AC power interruption.
- Two uninterruptible power supplies, UPS-A and UPS-B for the dual and quad engine clusters. One UPS provides battery backup for the first network switch, and a second UPS provides battery backup for the other switch. Each UPS module maintains power for two five-minute periods of AC loss while the engine shuts down.

High Availability with VPLEX Witness

VPLEX Witness helps multi-cluster VPLEX configurations automate the response to cluster failures and inter-cluster link outages.

VPLEX Witness is an optional component installed as a virtual machine on a customer host.

Note:

- The customer host must be deployed in a separate failure domain from either of the VPLEX clusters to eliminate the possibility of a single fault affecting both a cluster and VPLEX Witness.
- The VPLEX Witness server supports round trip time latency of 1 second over the management IP network.

In a Metro configuration, VPLEX uses rule sets to define how failures are handled. If the clusters lose contact with one another or if one cluster fails, rule sets define which cluster continues operation (the preferred cluster) and which suspends I/O (the non-preferred cluster). This works for most link or cluster failures.

In the case where the preferred cluster fails, all I/O is suspended resulting in data unavailability.

VPLEX Witness observes the state of the clusters, and thus can distinguish between a outage of the inter-cluster link and a cluster failure. VPLEX Witness uses this information to guide the clusters to either resume or suspend I/O.

VPLEX Witness works in conjunction with consistency groups. VPLEX Witness guidance does not apply to local volumes and distributed volumes that are not members of a consistency group.

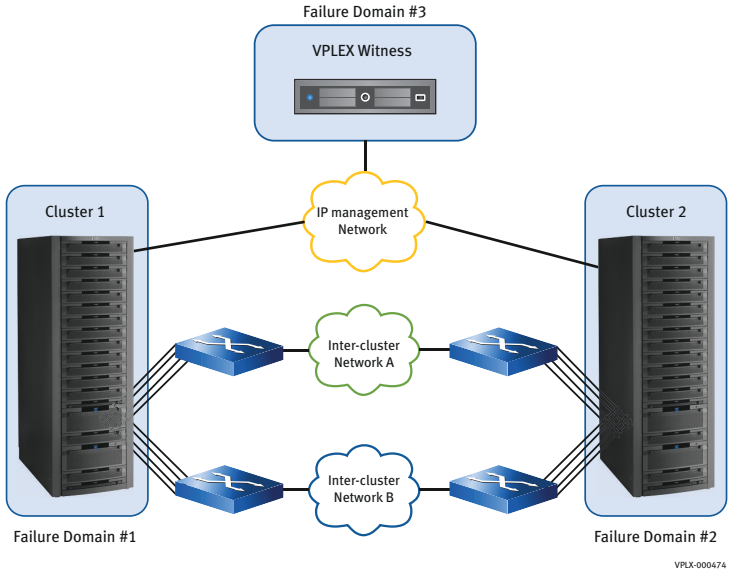
In Metro systems, VPLEX Witness provides seamless zero recovery time objective (RTO) fail-over for storage volumes in synchronous consistency groups.

Combine VPLEX Witness and VPLEX Metro to provide the following features:

- High availability for applications in a VPLEX Metro configuration leveraging synchronous consistency groups (no single points of storage failure).
- Fully automatic failure handling of synchronous consistency groups in a VPLEX Metro configuration (provided these consistency groups are configured with a specific preference).
- Better resource utilization.

The following figure shows a high level architecture of VPLEX Witness. The VPLEX Witness server must reside in a failure domain separate from cluster-1 and cluster-2.

Figure 23 High level VPLEX Witness architecture



The VPLEX Witness server must be deployed in a separate failure domain to both of the VPLEX clusters. This deployment enables VPLEX Witness to distinguish between a site outage and a link and to provide the correct guidance.

VPLEX Metro HA

VPLEX Metro High Availability (HA) configurations consist of a VPLEX Metro system deployed in conjunction with VPLEX Witness. There are two types of Metro HA configurations:

- VPLEX Metro HA can be deployed in places where the clusters are separated by 5 ms latency RTT or less.
- VPLEX Metro HA combined with Cross Connect between the VPLEX clusters and hosts can be deployed where the clusters are separated by 1 ms latency RTT or less.

Metro HA (without cross-connect)

Combine VPLEX Metro HA with host failover clustering technologies such as VMware HA to create fully automatic application restart for any site-level disaster.

VPLEX Metro/VMware HA configurations:

- Significantly reduce the Recovery Time Objective (RTO). In some cases, RTO can be eliminated.
- Ride through any single component failure (including the failure of an entire storage array) without disruption.
- When VMware Distributed Resource Scheduler (DRS) is enabled, distribute workload spikes between data centers, alleviating the need to purchase more storage.
- Eliminate the requirement to stretch the Fiber Channel fabric between sites. You can maintain fabric isolation between the two sites.

In this deployment, virtual machines can write to the same distributed device from either cluster and move between two geographically disparate locations.

If you use VMware Distributed Resource Scheduler (DRS) to automate load distribution on virtual machines across multiple ESX servers, you can move a virtual machine from an ESX server

attached to one VPLEX cluster to an ESX server attached to the second VPLEX cluster, without losing access to the underlying storage.

Metro HA with cross-connect

VPLEX Metro HA with cross-connect (VPLEX's front end ports are cross-connected) can be deployed where the VPLEX clusters are separated by 1 ms latency RTT or less. VPLEX Metro HA combined with cross-connect eliminates RTO for most of the failure scenarios.

Metro HA with cross-connect failure management

This section describes how VPLEX Metro HA with cross-connect rides through failures of hosts, storage arrays, clusters, VPLEX Witness, and the inter-cluster link.

Host failure

If hosts at one site fail, then VMware HA restarts the virtual machines on the surviving hosts. Since surviving hosts are connected to the same datastore, VMware can restart the virtual machines on any of the surviving hosts.

Cluster failure

If a VPLEX cluster fails:

- VPLEX Witness guides the surviving cluster to continue.
- VMware re-routes I/O to the surviving cluster.
- No disruption to I/O.

Storage array failure

If one or more storage arrays at one site fail:

- All distributed volumes continue I/O to the surviving leg.
- No disruption to the VPLEX clusters or the virtual machines.
- I/O is disrupted only to local virtual volumes on the VPLEX cluster attached to the failed array.

VPLEX Witness failure

If VPLEX Witness fails or becomes unreachable (link outage):

- Both VPLEX clusters call home to report that VPLEX Witness is not reachable.
- No disruption to I/O, VPLEX clusters, or the virtual machines.

Inter-cluster link failure

If the inter-cluster link fails:

- VPLEX Witness guides the preferred cluster to continue.
- I/O suspends at the non-preferred cluster.
- VMware re-routes I/O to the continuing cluster.
- No disruption to I/O.

The following table summarizes how VPLEX HA with cross-connect manages failures.

Table 4 How VPLEX Metro HA recovers from failure

Failure description	Failure handling
Host failure (Site 1)	VMware HA software automatically restarts the affected applications at Site 2.

Table 4 How VPLEX Metro HA recovers from failure (continued)

Failure description	Failure handling
VPLEX cluster failure (Site 1)	VPLEX Witness detects the failure and enables all volumes on the surviving cluster.
Inter-cluster link failure	• If the cross-connects use different physical links from those used to connect the VPLEX clusters, applications are unaffected. Every volume continues to be available in one data center or the other. • If the cross-connect links use the same physical links as those used to connect the VPLEX clusters, an application restart is required.
Storage array failure	Applications are unaffected. VPLEX dynamically redirects I/O to the mirrored copy on the surviving array. **Note:** This example assumes that all distributed volumes are also mirrored on the local cluster. If not, then the application remains available because the data can be fetched or sent from or to the remote cluster. However, each read/write operation now incurs a performance cost.
Failure of VPLEX Witness	Both clusters call home. As long as both clusters continue to operate and there is no inter-cluster link partition, applications are unaffected. **CAUTION** If either cluster fails or if there is an inter-cluster link partition, the system is at a risk of data unavailability. If the VPLEX Witness outage is expected to be long, disable the VPLEX Witness functionality to prevent the possible data unavailability.

Metro HA without cross-connect failure management

This section describes the failure scenarios for VPLEX Metro HA without cross-connect.

VPLEX cluster failure

In the event of a full VPLEX cluster outage at one site:

- VPLEX Witness guides the surviving cluster to continue.
- VMware at the surviving cluster is unaffected.
- VMware restarts the virtual machines at the site where the outage occurred, redirecting I/O to the surviving cluster.

  VMware can restart because the second VPLEX cluster has continued I/O without interruption.

Inter-cluster link failure - non-preferred site

If an inter-cluster link outage occurs, the preferred cluster continues, while the non-preferred cluster suspends. If a virtual machine is located at the preferred cluster, there is no interruption of service. If a virtual machine is located at the non-preferred cluster, the storage associated with the virtual machine is suspended. In such a scenario, most guest operating systems will fail. The virtual machine will be restarted at the preferred cluster after a short disruption.

Note: The preferred cluster is determined by consistency group detach rules.

If an inter-cluster link outage occurs:

- VPLEX Witness guides the preferred cluster to continue.
- VMware at the preferred cluster is unaffected.
- VMware restarts the virtual machines at the non-preferred (suspended) cluster, redirecting I/O to the preferred (uninterrupted) cluster.

VMware can restart because the second VPLEX cluster has continued I/O without interruption.

Higher availability

Combine VPLEX Witness with VMware and cross cluster connection to create even higher availability.

VPLEX Metro Hardware

To ensure continuous availability across multiple data centers in a metro region, VPLEX Metro provides an ideal solution with the option of Metro over IP (MetroIP) or Metro over Fibre Channel (MetroFC). These options use different I/O module SLIC hardware to establish WAN communication.

VPLEX provides options to use a VPLEX Metro with a 10 Gb Ethernet SLIC or a Fibre Channel SLIC. The VPLEX Metro over IP enables a quicker deployment option compared to the Fibre Channel over WAN, and is available with the single, dual, and quad configurations. Metro over IP works with both VS2 and VS6 hardware.

CHAPTER 5

VPLEX software and upgrade

This chapter describes the GeoSynchrony software that runs on the VPLEX directors and its non-disruptive upgrade.

GeoSynchrony

GeoSynchrony is the operating system that runs on the VPLEX directors. GeoSynchrony runs on both the VS2 and the VS6 hardware.

GeoSynchrony is:

- Designed for highly available, robust operation in geographically distributed environments
- Driven by real-time I/O operations
- Intelligent about locality of access
- Designed to provide the global directory that supports AccessAnywhere

The following table summarizes features provided by GeoSynchrony and AccessAnywhere:

Table 5 GeoSynchrony AccessAnywhere features

Feature	Description and considerations
Storage volume encapsulation	LUNs on a back-end array can be imported into an instance of VPLEX and used while keeping their data intact.
	Considerations: The storage volume retains the existing data on the device and leverages the media protection and device characteristics of the back-end LUN.
RAID 0	VPLEX devices can be aggregated to create a RAID 0 striped device.
	Considerations: Improves performance by striping I/Os across LUNs.
RAID-C	VPLEX devices can be concatenated to form a new larger device.
	Considerations: A larger device can be created by combining two or more smaller devices.
RAID 1	VPLEX devices can be mirrored within a site.
	Considerations: Withstands a device failure within the mirrored pair. • A device rebuild is a simple copy from the remaining device to the newly repaired device. Rebuilds are done in incremental fashion, whenever possible. • The number of required devices is twice the amount required to store data (actual storage capacity of a mirrored array is 50%). • The RAID 1 devices can come from different back-end array LUNs providing the ability to tolerate the failure of a back-end array.

Table 5 GeoSynchrony AccessAnywhere features (continued)

Feature	Description and considerations
Distributed RAID 1	VPLEX devices can be mirrored between sites.
	Considerations: Provides protection from site disasters and supports the ability to move data between geographically separate locations.
Extents	Storage volumes can be broken into extents and devices can be created from these extents.
	Considerations: Use extents when LUNs from a back-end storage array are larger than the desired LUN size for a host. This provides a convenient way of allocating what is needed while taking advantage of the dynamic thin allocation capabilities of the back-end array.
Migration	Volumes can be migrated non-disruptively to other storage systems.
	Considerations: Use migration for changing the quality of service of a volume or for performing technology refresh operations.
Global Visibility	The presentation of a volume from one VPLEX cluster where the physical storage for the volume is provided by a remote VPLEX cluster.
	Considerations: Use Global Visibility for AccessAnywhere collaboration between locations. The cluster without local storage for the volume will use its local cache to service I/O but non-cached operations incur remote latencies to write or read the data.

Non-disruptive upgrade (NDU)

VPLEX management server software and GeoSynchrony can be upgraded without disruption.

VPLEX hardware can be replaced, the engine count in a cluster increased, and a VPLEX Local can be expanded to VPLEX Metro without disruption.

VPLEX never has to be completely shut down.

Storage, application, and host upgrades

VPLEX enables the easy addition or removal of storage, applications, and hosts.

When VPLEX encapsulates back-end storage, the block-level nature of the coherent cache allows the upgrade of storage, applications, and hosts.

You can configure VPLEX so that all devices within VPLEX have uniform access to all storage blocks.

Increase engine count

When capacity demands increase, VPLEX supports hardware upgrades for single-engine VPLEX systems to dual-engine and dual-engine to quad-engine VPLEX systems.

These upgrades also increase the availability of front-end and back-end ports in the data center.

Software upgrades

VPLEX is fully redundant for:

- Ports
- Paths
- Directors
- Engines

This redundancy allows GeoSynchrony on VPLEX Local and Metro to be upgraded without interrupting host access to storage, it does not require service window or application disruption.

Note: You must upgrade the VPLEX management server software before upgrading GeoSynchrony. Management server upgrades are non-disruptive.

Simple support matrix

Dell EMC publishes storage array interoperability information in a Simple Support Matrix available on Dell EMC Online Support. This information details tested, compatible combinations of storage hardware and applications that VPLEX supports. The Simple Support Matrix can be located at:

https://www.dell.com/support

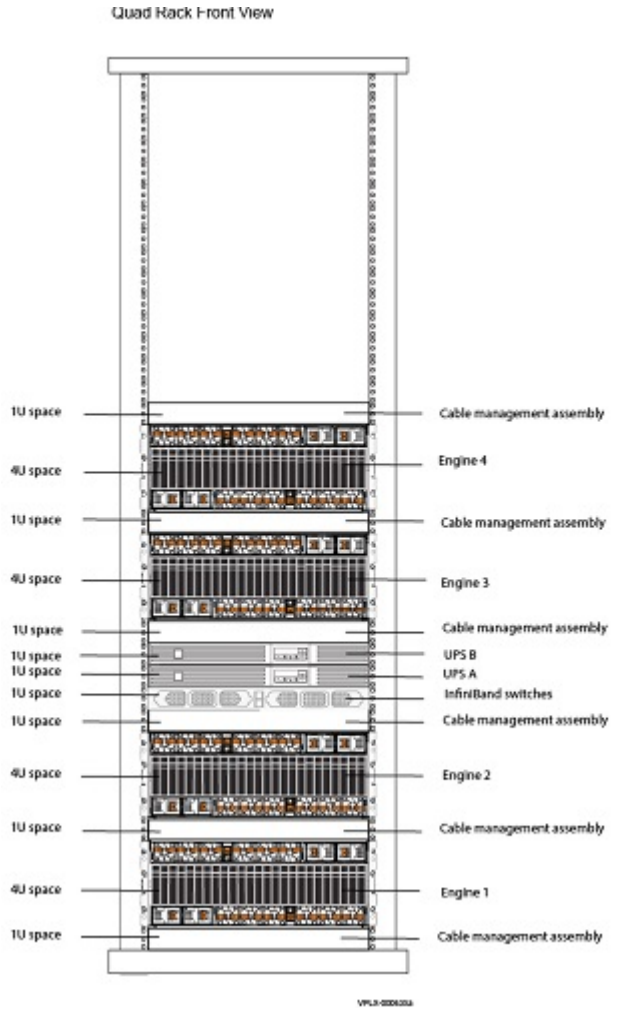
APPENDIX A

VPLEX cluster with VS6 hardware

A VPLEX cluster can include one, two, or four engines. Each VPLEX engine includes two directors. Each director provides front-end and back-end I/O connections.

The following figure shows the front view of a four-engine VPLEX rack.

Figure 24 VS6 Quad Engine cluster - Front view



The following table describes the major components of a VPLEX cluster and their functions.

Table 6 VS6 Hardware components

Component	Description
Engine	A VPLEX VS6 Engine Contains:

Table 6 VS6 Hardware components (continued)

	• Two directors • Two base modules • I/O modules • 20 fans per system • 6 fans per Director • 8 Fans for Drive Bay Cooling (four per base module and four chassis mounted) **Note:** VS6 uses fan filler modules only for the drive bay.
Director	A VS6 director Contains: • Two Haswell/Broadwell Intel Processors • 128GB memory (24 DDR4 DIMM slots available) • One Management module for intra-cluster communication • Five SLiC slots • One 64GB or 80GB M.2 SSD drive for the OS image • Dual Li-Ion battery backup modules per director
Management server - Management Module Control Stations (MMCS)	• MMCS-A and MMCS-B are located in the first engine on the cluster. All the remaining engines will have Akula management modules for the management connectivity. MMCS A is the Management interface to a public network and to the other VPLEX components in the cluster
Communication components	The basic communication components include: • The Local COM InfiniBand fabric for inter-director communications within the rack Dual Mellanox Dingo-V2 InfiniBand switches with FDR10 (40Gbs) support
Cable management assembly	• The cable management assembly is the routing tray for the SLiC cables. • A single engine cluster has two cable management assemblies. • In dual and quad engine clusters, the cable management assemblies between the engines will carry cables from both engines.

Table 6 VS6 Hardware components (continued)

Power supply units	The power supply units contain: • PDUs that connect to the AC power source of the cluster to supply power to the VPLEX components through Power Supply Units (PSU). • Backup Battery Units (BBU) that have sufficient capacity to ride through transient site power failures or to vault their cache when power is not restored. • Two uninterruptible power supplies, UPS-A and UPS-B for the IB switches in a dual or a quad engine cluster.

High availability and VPLEX hardware provides you more information on the VPLEX components and how they support the integrity and resiliency features of VPLEX.

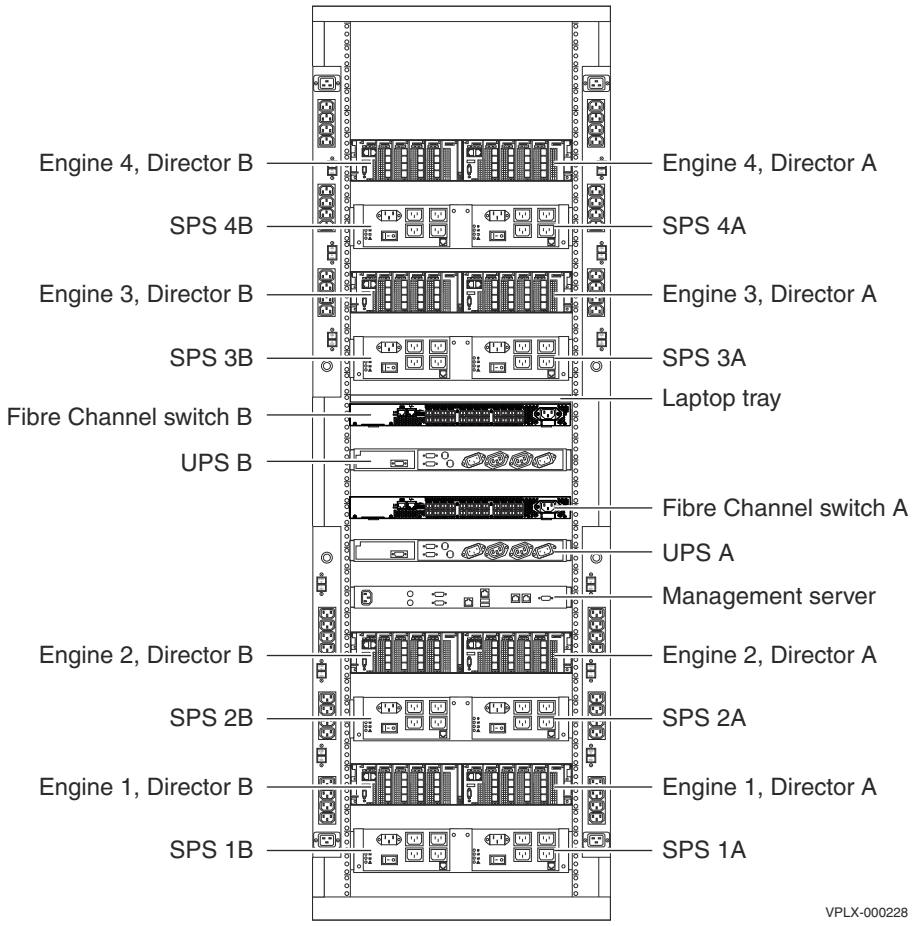
APPENDIX B

VPLEX cluster with VS2 hardware

A VPLEX cluster that runs on VS2 hardware can include one, two, or four engines. Each VPLEX engine includes two directors. Each director provides front-end and back-end I/O connections.

The following figure shows the front view of a four-engine VS2 VPLEX rack.

Figure 25 VS2 Quad Engine cluster



The following table describes the major components of a VPLEX cluster and their functions.

Table 7 VS2 Hardware components

Component	Description
Engine	Contains two directors, with each director providing front-end and back-end I/O connections.
Director	Contains:

Table 7 VS2 Hardware components (continued)

Component	Description
	• Five I/O modules (IOMs) • Management module for intra-cluster communication • Two redundant 400 W power supplies with built-in fans • CPU • Solid-state disk (SSD) that contains the GeoSynchrony operating environment • RAM
Management server	Provides: • Management interface to a public IP network • Management interfaces to other VPLEX components in the cluster • Event logging service
Fibre Channel COM switches (Dual-engine or quad-engine cluster only)	Provides intra-cluster communication support among the directors. This is separate from the storage I/O.
Power subsystem	Power distribution panels (PDPs) connect to the AC power source of the site and transfer power to the VPLEX components through power distribution units (PDUs). This provides a centralized power interface and distribution control for the power input lines. The PDPs contain manual on/off power switches for their power receptacles.
Standby Power Supply (SPS)	One SPS assembly (two SPS modules) provides backup power to each engine in the event of an AC power interruption. Each SPS module maintains power for two five-minute periods of AC loss while the engine shuts down.
Uninterruptible Power Supply (UPS) (Dual-engine or quad-engine cluster only)	One UPS provides battery backup for Fibre Channel switch A and the management server, and a second UPS provides battery backup for Fibre Channel switch B. Each UPS module maintains power for two five-minute periods of AC loss while the engine shuts down.

High availability and VPLEX hardware provides you more information on the VPLEX components and how they support the integrity and resiliency features of VPLEX.

GLOSSARY

A

AccessAnywhere The breakthrough technology that enables VPLEX clusters to provide access to information between clusters that are separated by distance.

active/active A cluster with no primary or standby servers, because all servers can run applications and interchangeably act as backup for one another.

active/passive A powered component that is ready to operate upon the failure of a primary component.

Active Directory A directory service included in most Windows Server operating systems. AD authenticates and authorizes users and computers in a network of Windows domain type.

active mirror A copy of data that is part of a local or remote mirroring service.

asynchronous Describes objects or events that are not coordinated in time. A process operates independently of other processes, being initiated and left for another task before being acknowledged. For example, a host writes data to the blades and then begins other work while the data is transferred to a local disk asynchronously across the WAN.

B

backend port VPLEX director port connected to storage arrays (acts as an initiator).

Backup Battery Unit (BBU) BBUs provides backup power to each engine in the event of an AC power interruption. Each VPLEX engine contains one BBU assembly that contains two BBU modules.

bandwidth The range of transmission frequencies a network can accommodate, expressed as the difference between the highest and lowest frequencies of a transmission cycle. High bandwidth allows fast or high volume transmissions.

bias When a cluster has the bias for a given DR1, it continues to service I/O to volumes on that cluster if connectivity to the remote cluster is lost (due to cluster partition or cluster failure). The bias for a specific volume is determined by the detach rules for the volume, the detach rules for the consistency group (if the volume is a member of a consistency group) and VPLEX Witness (if VPLEX Witness is deployed).

bit A unit of information that has a binary digit value of either 0 or 1.

block The smallest amount of data that can be transferred following SCSI standards, which is traditionally 512 bytes. Virtual volumes are presented to users as a contiguous lists of blocks.

block size The actual size of a block on a device.

Bookmark A label applied to a snapshot so that the snapshot can be explicitly called (identified) during recovery processes (during image access). Bookmarks are created through the CLI or GUI and can be created manually, by the user, or automatically, by the system. Bookmarks created automatically can be created at

pre-defined intervals or in response to specific system events. Parallel bookmarks are bookmarks that are created simultaneously across multiple consistency groups.

byte Memory space used to store eight bits of data.

C

cache Temporary storage for recent writes and recently accessed data. Disk data is read through the cache so that subsequent read references are found in the cache.

cache coherency Managing the cache so that data is not lost, corrupted, or overwritten. With multiple processors, data blocks may have several copies, one in the main memory and one in each of the cache memories. Cache coherency propagates the blocks of multiple users throughout the system in a timely fashion, ensuring that the data blocks do not have inconsistent versions in the different processors caches.

cluster Two or more VPLEX directors forming a single fault-tolerant cluster, deployed as one to four engines.

cluster deployment ID A numerical cluster identifier, unique within a VPLEX cluster. By default, VPLEX clusters have a cluster deployment ID of 1. For multi-cluster deployments, all but one cluster must be reconfigured to have different cluster deployment IDs.

cluster ID The identifier for each cluster in a multi-cluster deployment. The ID is assigned during installation.

clustering Using two or more computers to function together as a single entity. Benefits include fault tolerance and load balancing, which increases reliability and up time.

cluster IP seed The VPLEX IP seed is used to generate the IP addresses used by the internal components of the VPLEX. For more information about components and their IP addresses, see the Dell EMC VPLEX Installation and Setup Guide. Cluster ID is used by the virtualization software (inter director messaging, cluster identification).

COM The intra-cluster communication. The communication used for cache coherency and replication traffic.

command line interface (CLI) An interface that supports the use of typed commands to execute specific tasks.

consistency group A VPLEX structure that groups together virtual volumes and applies the same detach and failover rules to all member volumes. Consistency groups ensures the common application of a set of properties to the entire group. Create consistency groups for sets of volumes that require the same I/O behavior in the event of a link failure. There are two types of consistency groups:

- Synchronous Consistency Groups - Use write-through (synchronous) cache mode to write data to the underlying storage before an acknowledgment is sent to the host. This is dependent on the latency between clusters and the application's tolerance of the latency.
- Asynchronous Consistency Groups - Use write-back (asynchronous) cache mode to write protect data by mirroring it to the memory of another director in the cluster. Data is destaged asynchronously to the back-end storage arrays. Writes are acknowledged once the data has been committed to disk in write order.

continuity of operations (COOP) The goal of establishing policies and procedures to be used during an emergency, including the ability to process, store, and transmit data before and after.

controller A device that controls the transfer of data to and from a computer and a peripheral device.

D

data sharing The ability to share access to the same data with multiple servers regardless of time and location.

detach rule Predefined rules that determine which cluster continues I/O when connectivity between clusters is lost. A cluster loses connectivity to its peer cluster due to cluster partition or cluster failure.

Detach rules are applied at two levels; to individual volumes, and to consistency groups. If a volume is a member of a consistency group, the group detach rule overrides the rule set for the individual volumes. Note that all detach rules may be overridden by VPLEX Witness, if VPLEX Witness is deployed.

device A combination of one or more extents to which you add specific RAID properties. Local devices use storage from only one cluster. In VPLEX Metro and Geo configurations, distributed devices use storage from both clusters.

director A CPU module that runs GeoSynchrony, the core VPLEX software. There are two directors (A and B) in each engine, and each has dedicated resources and is capable of functioning independently.

dirty data The write-specific data stored in the cache memory that is yet to be written to disk.

disaster recovery (DR) The ability to restart system operations after an error, preventing data loss.

discovered array An array that is connected to the SAN and discovered by VPLEX.

disk cache A section of RAM that provides cache between the disk and the CPU. RAMs access time is significantly faster than disk access time. Therefore, a disk-caching program enables the computer to operate faster by placing recently accessed data in the disk cache.

Distributed consistency groups The RecoverPoint consistency group is divided into four segments. Each segment runs on one primary RPA and one to three secondary RPAs.

Distributed consistency groups enable a much higher throughput and IOPS rate, regardless of the amount of data being replicated.

distributed device A RAID 1 device whose mirrors are in different VPLEX clusters.

distributed file system (DFS) Supports the sharing of files and resources in the form of persistent storage over a network.

distributed RAID1 device (DR1) Distributed devices have physical volumes at both clusters in a VPLEX Metro configuration for simultaneous active/active and read/write access using AccessAnywhere.

E

engine Consists of two directors, management modules, and redundant power. Unit of scale for VPLEX configurations. Single = 1 engine, dual = 2 engines, Quad = 4 engines per cluster.

Ethernet A Local Area Network (LAN) protocol. Ethernet uses a bus topology, meaning all devices are connected to a central cable, and supports data transfer rates of between 10 megabits per second and 10 gigabits per second. For example, 100 Base-T supports data transfer rates of 100 Mb/s.

event A log message that results from a significant action initiated by a user or the system.

extent All or a portion (range of blocks) of a storage volume.

F

failover Automatically switching to a redundant or standby device, system, or data path upon the failure or abnormal termination of the currently active device, system, or data path.

fault domain A set of components that share a single point of failure. For VPLEX, the concept that every component of a highly available system is separated, so that if a fault occurs in one domain, it will not result in failure in other domains to which it is connected.

fault tolerance Ability of a system to keep working in the event of hardware or software failure, usually achieved by duplicating key system components.

Fibre Channel (FC) A protocol for transmitting data between computer devices. Longer distance requires the use of optical fiber. However, FC also works using coaxial cable and ordinary telephone twisted pair media. Fibre channel offers point-to-point, switched, and loop interfaces. Used within a SAN to carry SCSI traffic.

Fibre Channel over IP (FCIP) Combines Fibre Channel and Internet protocol features to connect SANs in geographically distributed systems.

field replaceable unit (FRU) A unit or component of a system that can be replaced on site as opposed to returning the system to the manufacturer for repair.

firmware Software that is loaded on and runs from the flash ROM on the VPLEX directors.

firmware Software that is loaded on and runs from the flash ROM on the VPLEX directors.

front end port VPLEX director port connected to host initiators (acts as a target).

G

geographically distributed system A system that is physically distributed across two or more geographically separated sites. The degree of distribution can vary widely, from different locations on a campus or in a city to different continents.

gigabit (Gb or Gbit) 1,073,741,824 (2^30) bits. Often rounded to 10^9.

gigabit Ethernet The version of Ethernet that supports data transfer rates of 1 gigabit per second.

gigabyte (GB) 1,073,741,824 (2^30) bytes. Often rounded to 10^9.

global file system (GFS) A shared-storage cluster or distributed file system.

H

hold provisioning An attribute of a registered array that allows you to set the array as unavailable for further provisioning of new storage.

host bus adapter (HBA) An I/O adapter that manages the transfer of information between the host computers bus and memory system. The adapter performs many low-level interface functions automatically or with minimal processor involvement to minimize the impact on the host processors performance.

I

infiniband A networking standard used for passing data between and among computers. The VS6 hardware uses this protocol for the intra-cluster communication.

input/output (I/O) Any operation, program, or device that transfers data to or from a computer.

internet Fibre Channel protocol (iFCP) Connects Fibre Channel storage devices to SANs or the Internet in geographically distributed systems using TCP.

internet small computer system interface (iSCSI) A protocol that allows commands to travel through IP networks, which carries data from storage units to servers anywhere in a computer network.

intranet A network operating like the World Wide Web but with access restricted to a limited group of authorized users.

J

Journal volumes Volumes that contain data waiting to be distributed to target replica volumes and copies of the data previously distributed to the target volumes. Journal volumes allow convenient rollback to any point in time, enabling instantaneous recovery for application environments.

K

kilobit (Kb) 1,024 (2^10) bits. Often rounded to 10^3.

kilobyte (KB) 1,024 (2^10) bytes. Often rounded to 10^3.

L

latency Amount of time required to fulfill an I/O request.

LDAP Lightweight Directory Access Protocol, an application protocol that accesses and maintains distributed directory information services over an IP network.

load balancing Distributing the processing and communications activity evenly across a system or network so that no single device is overwhelmed. Load balancing is especially important when the number of I/O requests issued is unpredictable.

local area network (LAN) A group of computers and associated devices that share a common communications line and typically share the resources of a single processor or server within a small geographic area.

local device A combination of one or more extents to which you add specific RAID properties. Local devices use storage from only one cluster.

logical unit number (LUN) Virtual storage to which a given server with a physical connection to the underlying storage device may be granted or denied access. LUNs are used to identify SCSI devices, such as external hard drives that are connected to a computer. Each device is assigned a LUN number which serves as the device's unique address.

M

Management Module Control Station (MMCS) The management entity (management server) is VS6 hardware. The first engine in a cluster has two MMCS: A and B. All the remaining engines have Akula management modules for the management connectivity.

megabit (Mb) 1,048,576 (2^20) bits. Often rounded to 10^6.

megabyte (MB) 1,048,576 (2^20) bytes. Often rounded to 10^6.

metavolume A storage volume used by the system that contains the metadata for all the virtual volumes managed by the system. There is one metadata storage volume per cluster.

Metro-Plex Two VPLEX Metro clusters connected within Metro (synchronous) distances, approximately 60 miles or 100 kilometers.

MetroPoint consistency group (Metro group) A consistency group that is protecting a DR1 volume in a MetroPoint topology.

mirroring The writing of data to two or more disks simultaneously. If one of the disk drives fails, the system can instantly switch to one of the other disks without losing data or service. RAID 1 provides mirroring.

mirroring services Mirroring features provided through a storage service profile.

miss An operation where the cache is searched but does not contain the data, so the data instead must be accessed from disk.

N

namespace A set of names recognized by a file system in which all names are unique.

network System of computers, terminals, and databases connected by communication lines.

network architecture Design of a network, including hardware, software, method of connection, and the protocol used.

network-attached storage (NAS) Storage elements connected directly to a network.

network partition When one site loses contact or communication with another site.

Non-distributed consistency groups Transfer data through one primary RPA that is designated by the user during group creation. The policies applied by the consistency group can be modified at any time.

In the event of RPA failure, groups that transfer data through the failed RPA will move to other RPAs in the cluster.

O

Open LDAP Open source implementation of the Lightweight Directory Access Protocol (LDAP).

P

parity checking Checking for errors in binary data. Depending on whether the byte has an even or odd number of bits, an extra 0 or 1 bit, called a parity bit, is added to each byte in a transmission. The sender and receiver agree on odd parity, even parity, or no parity. If they agree on even parity, a parity bit is added that makes each byte even. If they agree on odd parity, a parity bit is added that makes each byte odd. If the data is transmitted incorrectly, the change in parity will reveal the error.

partition A subdivision of a physical or virtual disk, which is a logical entity only visible to the end user and not to any of the devices.

point-in-time (PIT) See the description for Snapshot/PIT.

Production journal volumes Volumes that hold system delta marking information.

Production volumes Volumes that are written to by the host applications. Writes to production volumes are split such that they are sent to both the normally designated volumes and RPAs simultaneously.

Each production volume must be exactly the same size as the replica volume to which it replicates.

R

RAID The use of two or more storage volumes to provide better performance, error recovery, and fault tolerance.

RAID 0 A performance orientated striped or dispersed data mapping technique. Uniformly sized blocks of storage are assigned in regular sequence to all of the arrays disks. Provides high I/O performance at low inherent cost. No additional disks are required. The advantages of RAID 0 are a very simple design and an ease of implementation.

RAID 1 Also called mirroring, this has been used longer than any other form of RAID. It remains popular because of simplicity and a high level of data availability. A mirrored array consists of two or more disks. Each disk in a mirrored array holds an identical image of the user data. RAID 1 has no striping. Read performance is improved since either disk can be read at the same time. Write performance is lower than single disk storage. Writes must be performed on all disks, or mirrors, in the RAID 1. RAID 1 provides very good data reliability for read-intensive applications.

RAID leg A copy of data, called a mirror, that is located at a user's current location.

rebuild The process of reconstructing data onto a spare or replacement drive after a drive failure. Data is reconstructed from the data on the surviving disks, assuming mirroring has been employed.

RecoverPoint Appliance (RPA) Hardware that manages all aspects of data protection for a storage group, including capturing changes, maintaining the images in the journal volumes, and performing image recovery.

RecoverPoint cluster All connected RecoverPoint Appliances on both sides of the replication.

RecoverPoint consistency groups A logical entity consisting of a collection of replication sets grouped together to ensure write order consistency across all the primary volumes of the replication sets.

RecoverPoint consistency groups are used to configure protection policies and set RPO and RTO policies. Consistency group policies determine behavior during failover and recovery. An operation on a RecoverPoint consistency group applies to all of the replication sets within the consistency group, and their associated production, replica, and journal volumes.

RecoverPoint consistency groups monitor their member volumes to ensure consistency and write-order fidelity. If two data sets are dependent on one another (a database and a database log), they should be part of the same RecoverPoint consistency group.

There are three types of RecoverPoint consistency groups:

- Non-distributed consistency groups
- Distributed consistency groups
- MetroPoint consistency groups

RecoverPoint site All RecoverPoint entities that are on one side of the replication.

RecoverPoint volumes The volumes used for RecoverPoint Replication.

The following types of volumes are required in all RecoverPoint configurations:

- Repository volumes
- Production volumes
- Replica volumes
- Journal volumes

Recovery Point Objective (RPO) Recovery Point Objective. The time interval between the point of failure of a storage system and the expected point in the past at which the storage system is capable of recovering customer data. Informally, RPO is maximum amount of data loss that can be tolerated by the application after a failure. The value of the RPO is highly dependent upon the recovery technique used. For example, RPO for backups is typically days, for asynchronous replication minutes, and for mirroring or synchronous replication seconds or instantaneous.

Recovery Time Objective (RTO) Recovery Time Objective. Not to be confused with RPO, RTO is the time duration within which a storage solution is expected to recover from failure and begin servicing application requests. Informally, RTO is the longest tolerable application outage due to a failure of a storage system. RTO is a function of the storage technology. It may measure in hours for backup systems, minutes for a remote replication, and seconds (or less) for a mirroring.

redundancy The duplication of hardware and software components. In a redundant system, if a component fails then a redundant component takes over, allowing operations to continue without interruption.

registered array An array that is registered with VPLEX. Registration is required to make the array available for services-based provisioning. Registration includes connecting to and creating awareness of the array's intelligent features. Only VMAX and VNX arrays can be registered.

reliability The ability of a system to recover lost data.

Remote Direct Memory Access (RDMA) Allows computers within a network to exchange data using their main memories and without using the processor, cache, or operating system of either computer.

Replica journal volumes Volumes that hold:

- Snapshots that are either waiting to be replicated or already distributed to the replica
- Metadata for each image
- Bookmarks

Each replica journal holds as many snapshots as its capacity allows.

After distribution, the oldest snapshot is discarded to make room for the newest snapshot. The number of snapshots in the journal depends on the size of the snapshots and the capacity of the volumes.

replication set When RecoverPoint is deployed, a production source volume and one or more replica volumes to which it replicates.

replica volumes Volumes to which production volumes replicate. In prior releases, the replica volume must be exactly the same size as its production volume. In RecoverPoint (RP) 4.0 and GeoSynchrony release 5.2, RecoverPoint supports a feature called Fake Size, where the replica volume size can be higher than the production volume.

repository volume A volume dedicated to RecoverPoint for each RPA cluster. The repository volume serves all RPAs of the particular RPA cluster and the splitter associated with that cluster. The repository volume stores configuration information about the RPAs and RecoverPoint consistency groups. There is one repository volume per RPA cluster.

restore source This operation restores the source consistency group from data on the copy target.

S

scalability Ability to easily change a system in size or configuration to suit changing conditions, to grow with your needs.

simple network management protocol (SNMP) Monitors systems and devices in a network.

site ID The identifier for each cluster in a multi-cluster VPLEX. By default, in a non-geographically distributed system the ID is 0. In a geographically distributed system, the ID of the first cluster is 1, the next is 2, and so on, each number identifying a physically separate cluster. These identifiers are assigned during installation.

SLES SUSE Linux Enterprise Server (SLES) is a Linux distribution supplied by SUSE and targeted at the business market.

small computer system interface (SCSI) A set of evolving ANSI standard electronic interfaces that allow personal computers to communicate faster and more flexibly than previous interfaces with peripheral hardware such as disk drives, tape drives, CD-ROM drives, printers, and scanners.

Snapshot/PIT A point-in-time copy that preserves the state of data at an instant in time, by storing only those blocks that are different from an already existing full copy of the data.

Snapshots are also referred to as point-in-time (PIT). Snapshots stored at a replica journal represent the data that has changed on the production storage since the closing of the previous snapshot.

splitter Dell EMC RecoverPoint write-splitting technology built into GeoSynchrony starting in 5.1.

storage area network (SAN) A high-speed special purpose network or subnetwork that interconnects different kinds of data storage devices with associated data servers on behalf of a larger network of users.

storage view A combination of registered initiators (hosts), front-end ports, and virtual volumes, used to control a hosts access to storage.

storage volume A Logical Unit Number (LUN) or unit of storage presented by the back end array.

stripe depth The number of blocks of data stored contiguously on each storage volume in a RAID 0 device.

striping A technique for spreading data over multiple disk drives. Disk striping can speed up operations that retrieve data from disk storage. Data is divided into units and distributed across the available disks. RAID 0 provides disk striping.

synchronous Describes objects or events that are coordinated in time. A process is initiated and must be completed before another task is allowed to begin.

For example, in banking, two withdrawals from a checking account that started at the same time must not overlap; therefore, they are processed synchronously.

T

throughput

1. The number of bits, characters, or blocks passing through a data communication system or portion of that system.
2. The maximum capacity of a communications channel or system.
3. A measure of the amount of work performed by a system over a period of time. For example, the number of I/Os per day.

tool command language (TCL) A scripting language often used for rapid prototypes and scripted applications.

transfer size The size of the region in cache used to service data migration. The area is globally locked, read at the source, and written at the target. Transfer size can be as small as 40 K, as large as 128 M, and must be a multiple of 4 K. The default value is 128 K.

A larger transfer size results in higher performance for the migration, but may negatively impact front-end I/O. This is especially true for VPLEX Metro migrations. Set a large transfer-size for migrations when the priority is data protection or migration performance.

A smaller transfer size results in lower performance for the migration, but creates less impact on front-end I/O and response times for hosts. Set a smaller transfer-size for migrations when the priority is front-end storage response time.

transmission control protocol/Internet protocol (TCP/IP) The basic communication language or protocol used for traffic on a private network and the Internet.

U

uninterruptible power supply (UPS) A power supply that includes a battery to maintain power in the event of a power failure.

universal unique identifier (UUID) A 64-bit number used to uniquely identify each VPLEX director. This number is based on the hardware serial number assigned to each director.

V

virtualization A layer of abstraction that is implemented in software that servers use to divide available physical storage into storage volumes or virtual volumes.

virtual volume Unit of storage presented by the VPLEX front end ports to hosts. A virtual volume looks like a contiguous volume, but can be distributed over two or more storage volumes.

W

wide area network (WAN) A geographically dispersed telecommunications network. This term distinguishes a broader telecommunication structure from a local area network (LAN).

world wide name (WWN) A specific Fibre Channel Name Identifier that is unique worldwide and represented by a 64-bit unsigned binary value.

write-back mode A caching technique where the completion of a write request is communicated as soon as the data is in cache, with writes to disk occurring at different times. Write-back is faster than write-through, but risks a system failure before the data is safely written to disk.

write-through mode A caching technique in which the completion of a write request is communicated only after data is written to disk. This is almost equivalent to non-cached systems, but with data protection.

INDEX