White Paper

EMC VPLEX 5.0 ARCHITECTURE GUIDE

Abstract

This white paper explains the hardware and software architecture of the EMC® VPLEX™ series with EMC GeoSynchrony™. This paper will be of particular interest to system, application, database, and storage architects and anyone interested in deploying solutions on the EMC VPLEX platform. The value and necessity of all features are highlighted in sufficient detail to allow a reader of general technical experience to assimilate the material.

April 2011

Part Number h8232

Table of Contents

Executive summary

Today's data center administrators are faced with the increasing challenges of doing more with less. Data centers are under pressure as the result of the ever-increasing demand for more and more compute and storage capacity. Pressures of the economy mean fewer dollars for staff and equipment. These pressures have led to widespread adoption of virtualization as a technique for maximizing the efficiency of physical resources for both compute and storage. This trend has supported the recent explosive growth and adoption of cloud computing and has changed the approach that data center administrators are taking to address their growth demands.

The cloud computing and service provider model offers the data center administrator a means of satisfying dynamic bursts in demand without needing to purchase and build out all of the physical equipment needed to satisfy peak demands within their own facilities. Increasingly, business are rebuilding their own IT services around virtualization and retooling around this technology by adopting such techniques as virtual desktop infrastructure (VDI). Through this they gain the benefits of cost reduction through better efficiency but also benefit from the capabilities of dynamic instantiation and centralized management.

The EMC® VPLEX™ family extends these capabilities even further by providing the data center with the benefits of **mobility**, **availability**, and **collaboration** through VPLEX **AccessAnywhere** virtual storage, the breakthrough block-storage technology that enables a single copy of data to be shared, accessed, and relocated over distance. VPLEX 5.0 extends and enhances the capabilities of VPLEX 4.0 virtual storage, adding VPLEX Geo to VPLEX Local and VPLEX Metro in the VPLEX product family. VPLEX Geo extends the reach of VPLEX AccessAnywhere storage, supporting round-trip time (RTT) inter-site latencies of up to 50 ms through asynchronous communication.

The VPLEX family removes physical barriers within, across, and between data centers. VPLEX is the first platform in the world that delivers both local and distributed federation. **Local federation** provides the transparent cooperation of physical elements within a site. **Distributed federation** extends access between two locations across distance.

The combination of virtual servers and EMC Virtual Storage enables entirely new ways to solve IT problems and introduce new models of computing, allowing users to:

- Move applications and their data between data centers without disruption
- Provide continuous operations in the presence of site disasters
- Collaborate over distance with shared data
- Perform batch process in low-cost energy locations
- Enable workload balancing and relocation across sites
- Aggregate data centers and deliver "24 x forever" availability

EMC VPLEX overview

EMC VPLEX represents the next-generation architecture for data mobility and information access. This architecture is based on EMC's 20-plus years of expertise in designing, implementing, and perfecting enterprise-class intelligent cache and distributed data protection solutions.

VPLEX is a solution for federating **EMC and non-EMC storage**. VPLEX resides between the servers and heterogeneous storage assets and has unique characteristics in its architecture:

- **Scale-out clustering hardware** lets you start small and grow big with predictable service levels
- **Advanced data caching** utilizes large-scale SDRAM cache to improve performance and reduce I/O latency and array contention
- **Distributed cache coherence** provides automatic sharing, balancing, and failover of I/O within and between VPLEX clusters
- **A consistent view** of one or more LUNs between VPLEX clusters separated either by a few feet within a data center or across asynchronous RTT distances enables new models of high availability, workload mobility, and collaboration.

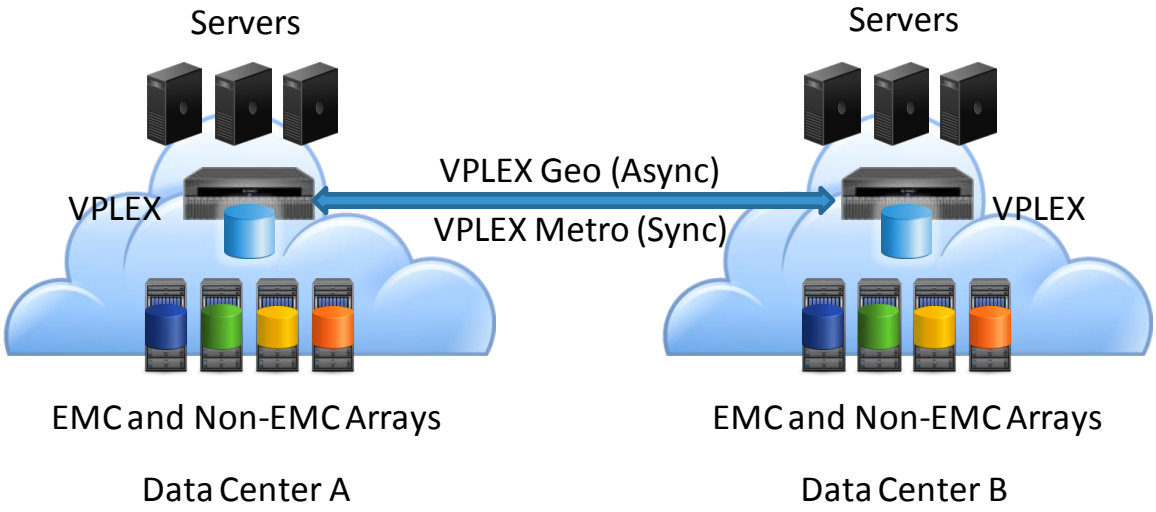

Figure 1. A multicluster deployment of VPLEX

EMC VPLEX family overview

The EMC VPLEX family today consists of:

- **VPLEX Local** for managing data mobility and access within the data center using a single VPLEX cluster.
- **VPLEX Metro** for mobility and access across two locations separated by inter-site RTT of up to 5 ms. VPLEX Metro uses two VPLEX clusters and includes the unique

capability where a remote VPLEX Metro cluster can present LUNs without the need for physical storage for those LUNs at the remote cluster. It also supports synchronous distributed volumes that mirror data between the two clusters using write-through caching.

- **VPLEX Geo**, which also uses two VPLEX clusters, for access between two sites over extended asynchronous distances with RTT latencies up to 50 ms. VPLEX Geo distributed volumes support AccessAnywhere distributed mirroring using write-back caching.

At the highest level, VPLEX has unique capabilities that customers value.

- First, VPLEX is *distributed*, because it can connect multiple sites together over distance, allowing secure and consistent collaboration across distributed users.
- Next, VPLEX is *dynamic*, because it is a single interface for multivendor storage and it delivers dynamic data mobility, which is being able to move applications and data in real time, with no outage required.
- And finally, VPLEX is *smart*, because its unique AccessAnywhere technology can present and keep the same data consistent within and between sites, even across distance.

VPLEX addresses three distinct customer requirements:

- **Mobility**: The ability to move applications and data across different storage installations—within the same data center, across a campus, or within a geographical region. And now, with VPLEX Geo, users can move data across even greater distances.
- **Availability**: The ability to create high-availability storage infrastructure across these same varied geographies with unmatched resiliency.
- **Collaboration**: The ability to provide efficient real-time data collaboration over distance for such "big data" applications as video, geographic/oceanographic research, and others.

VPLEX architecture

VPLEX 5.0 introduces both a new hardware platform, VPLEX VS2, and a new software release, GeoSynchrony™ 5.0. The VPLEX GeoSynchrony 5.0 software release supports both the VPLEX VS1 hardware platform introduced in VPLEX 4.0 as well as VS2.

VPLEX VS2 hardware platform

A VPLEX VS2 system with GeoSynchrony 5.0 is composed of one or two VPLEX clusters: one cluster for VPLEX Local systems and two clusters for VPLEX Metro and VPLEX Geo systems. These clusters provide the VPLEX AccessAnywhere capabilities.

Each VPLEX cluster consists of:

- A VPLEX Management Console
- One, two, or four engines
- One standby power supply for each engine

In configurations with more than one engine, the cluster also contains:

- A pair of Fibre Channel switches
- An uninterruptible power supply for each Fibre Channel switch

VPLEX Metro and VPLEX Geo systems optionally include a Witness. The Witness is implemented as a virtual machine and is deployed in a separate fault domain from two VPLEX clusters. The Witness is used to improve application availability in the presence of site failures and inter-cluster communication loss.

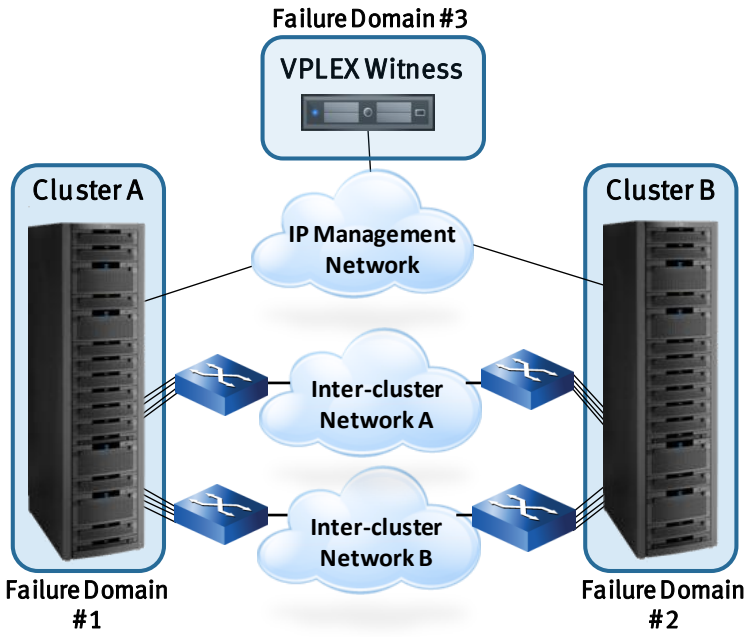

Figure 2. A VPLEX system with the VPLEX Witness

VPLEX Management Console

The VPLEX Management Console is a 1U server in the VPLEX cabinet. This server provides the management interfaces to VPLEX—hosting the VPLEX web server process that serves the VPLEX GUI and REST-based web services interface, as well as the command line interface (CLI) service.

In the VPLEX Metro and VPLEX Geo configurations, the VPLEX Management Consoles of each cluster are inter-connected using a virtual private network (VPN) that allows for remote cluster management from a local VPLEX Management Console. When the system is deployed with a VPLEX Witness, the VPN is extended to include the Witness as well.

VPLEX VS2 Engine

A VPLEX VS2 Engine is a chassis containing two directors, redundant power supplies, fans, I/O modules, and management modules. The directors are the workhorse components of the system and are responsible for processing I/O requests from the hosts, serving and maintaining data in the distributed cache, providing the virtual-to-physical I/O translations, and interacting with the storage arrays to service I/O.

A VPLEX VS2 Engine has 10 I/O modules, with five allocated to each director. Each director has one four-port 8 Gb/s Fibre Channel I/O module used for front-end SAN (host) connectivity and one four-port 8 Gb/s Fibre Channel I/O module used for back-end SAN (storage array) connectivity. Each of these modules has 40 Gb/s effective PCI bandwidth to the CPUs of their corresponding director. A third I/O module, called the WAN COM module, is used for inter-cluster communication. Two variants of this module are offered, one four-port 8 Gb/s Fibre Channel module and one two-port 10 Gb/s Ethernet module. The fourth I/O module provides two ports of 8 Gb/s Fibre Channel connectivity for intra-cluster communication. The fifth I/O module for each director is reserved for future use.

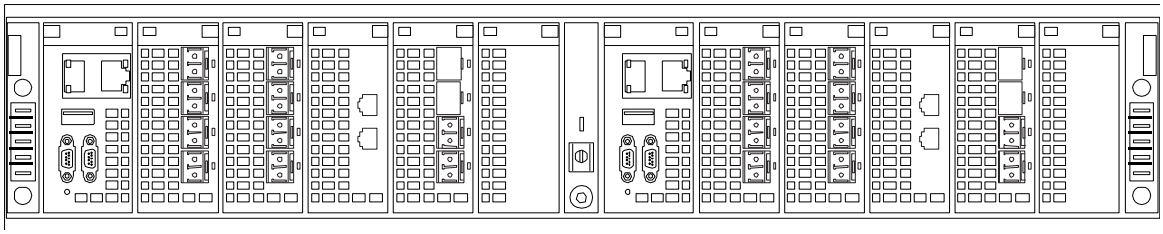

Figure 3. A VPLEX VS2 Engine and its components

The engine uses N+1 cooling and power. Cooling is accomplished through two independent fans for each director, four fans total in the entire enclosure. The fans are integrated into the power supplies and provide front-to-rear cooling. The engine enclosure houses four redundant power supplies that are each capable of providing full power to the chassis. Redundant management modules provide IP connectivity to the directors from the management console that is provided with each cluster. Two private IP subnets provide redundant IP connectivity between the directors of a cluster and the cluster's management console.

Each engine is supported by a redundant standby power supply unit that provides power to ride through transient power-loss and support write-cache vaulting.

Clusters containing two or more engines are fitted with a pair of Fibre Channel switches that provide redundant Fibre Channel connectivity that support intra-cluster communication between the directors. Each Fibre Channel switch is backed by a dedicated uninterruptible power supply (UPS) that provides support for riding through transient power loss.

VPLEX Witness

For VPLEX Metro and Geo an optional component called the VPLEX Witness can be deployed at a third location to improve data availability in the presence of cluster

failures and inter-cluster communication loss. The VPLEX Witness is implemented as a virtual machine and requires a VMware® ESX® server for its operation.

VPLEX GeoSynchrony 5.0

The GeoSynchrony operating environment for VPLEX provides the intelligence that controls all components in an EMC VPLEX system. GeoSynchrony is an intelligent, multitasking, locality-aware operating environment that controls the data flow for virtual storage. It is wholly devoted to virtual storage operations and optimized for mobility, availability, and collaboration. GeoSynchrony was designed for highly available, robust operation in geographically distributed environments. It is driven by real-time I/O operations and is intelligent about locality of access. The GeoSynchrony software uses locality information to migrate cached user and management data transparently within the system to optimize performance.

VPLEX virtualization

VPLEX AccessAnywhere virtual storage in GeoSynchrony 5.0 provides many virtualization capabilities that allow the data center administrator to address their mobility, availability, and collaboration needs.

Table 1. AccessAnywhere capabilities

Virtualization capability	Provides the following	Configuration 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.	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.	Improves performance by striping I/Os across LUNs.
RAID-C	VPLEX devices can be concatenated to form a new larger device.	Provides a means of creating a larger device by composing two or more smaller devices.
RAID 1	VPLEX devices can be mirrored within a site.	Withstands a device failure within the mirrored pair. A device rebuild is a simple copy from the remaining device to the newly repaired device. The number of required devices is twice the amount required to store data (usable storage capacity of a mirrored array is 50 percent). The RAID 1 devices can come from different back-end array LUNs providing the ability to

		tolerate the failure of a back-end array.
Distributed RAID 1	VPLEX devices can be mirrored between sites.	Provides protection from site disasters and supports the ability to move data between geographically separate locations.
Disk slicing	Storage volumes can be partitioned and devices created from these partitions.	Often used when large LUNs are claimed from a back-end storage array. The back-end storage volumes may be thin provisioned. This provides a convenient means 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.	Use for changing the quality of service of a volume or for performing technology refresh operations.
Remote export	The presentation of a volume from one VPLEX cluster where the physical storage for the volume is provided by a remote VPLEX cluster.	Use 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 will incur remote latencies to write or read the data.

VPLEX global distributed cache

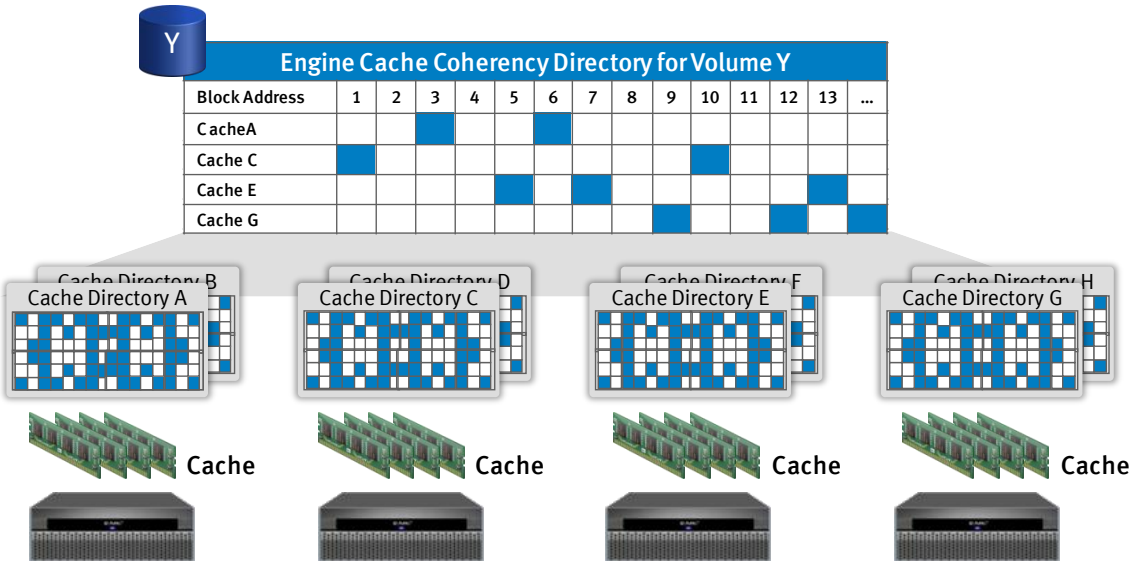

Figure 4. VPLEX distributed cache

The individual memory systems of each VPLEX director are combined to form the VPLEX distributed cache. Data structures within these memories in combination with distributed algorithms achieve the coherency and consistency guarantees provided by VPLEX virtual storage. This guarantee ensures that the I/O behavior observed by hosts accessing VPLEX storage is consistent with the behavior of a traditional disk. The VPLEX distributed algorithms are designed to minimize inter-director messaging and take advantage of I/O locality in the placement of key data structures.

The design is truly distributed: Any director within a cluster is able to service an I/O request for a virtual volume served by that cluster. Each director within a cluster is exposed to the same set of physical storage volumes from the back-end arrays and has the same virtual-to-physical storage mapping metadata for its volumes. The distributed design extends across VPLEX Metro and Geo systems to provided cache coherency and consistency for the global system. This ensures that a host accesses to a distributed volume always receive the most recent consistent data for that volume.

VPLEX cache modes

VPLEX Local and Metro both use the **write-through** cache mode. With write-through caching as a write request is received from a host to a virtual volume, the data is written through to the back-end storage volume(s) that map to the volume. When the array(s) acknowledge this data, an acknowledgement is then sent back from VPLEX to the host indicating a successful write. This provides an especially strong guarantee of data durability in the case of a distributed mirror where the back-end storage volumes supporting the mirror can be placed in different data centers.

This cache mode is not appropriate, however, for VPLEX Geo systems where the VPLEX clusters can be deployed up to 50 ms RTT apart. The latency of synchronous write-through operation would not be acceptable to most applications at this distance. Instead, VPLEX Geo uses **write-back** caching to achieve data durability without requiring synchronous operation. In this cache mode VPLEX accepts host writes into cache and places a protection copy in the memory of another local director before acknowledging the data to the host. The data is then destaged asynchronously to the back-end storage arrays. Cache vaulting logic within VPLEX stores any unwritten cache data onto local SSD storage in the event of a power failure.

The VPLEX RAID 1 and distributed RAID 1 features coupled with VPLEX's distributed coherent cache are the core technologies that deliver local and distributed mobility, availability, and collaboration.

VPLEX mobility

VPLEX provides direct support for data mobility both within and between data centers near and far and enables application mobility, data center relocation, and consolidation.

Data mobility is the relocation of data from one location (the source) to another (the target), after which the data is subsequently accessed only via the target. By contrast **data replication** enables applications to continue to access the source data after the

target copy is created. Similarly, data mobility is different from **data mirroring**, which transparently maintains multiple copies of the data for the purposes of data protection.

During and after a data mobility operation, applications continue to access the data using its original VPLEX volume identifier. This avoids the need to point applications to a new data location or change the configuration of their storage settings, effectively eliminating the need for application cutover.

There are many types and reasons for data mobility:

- Moving data from one storage device to another
- Moving applications from one storage device to another
- Moving operating system files from one storage device to another
- Consolidating data or database instances
- Moving database instances
- Moving data centers containing storage infrastructure from one physical location to another

The non-disruptive nature of VPLEX data mobility operations helps to simplify the planning and execution factors that would normally be considered when performing a disruptive migration. It is still important to consider some of these factors, however, when performing data mobility between data centers and increasing the distance between an application and its data. Considerations include the business impact and the type of data to be moved, site location(s), number of systems and applications, performance requirements, and total amount of data, as well as time considerations and schedules.

Mobility with the VPLEX Migration Wizard

The VPLEX GUI supports the ability to easily move the physical location of virtual storage while providing continuous access to this storage by the host. Using this wizard the operator first displays and selects the virtual volumes to move, and the wizard then displays a collection of candidate storage volumes and allows the operator to select and allocate new virtual storage for each volume. Once selected VPLEX automates the process of moving the data to its new location. Throughout the process the volume will retain its volume identity and continuous access is maintained to the data from the host.

Availability

Business continuity is extended by the VPLEX Local RAID 1 technology, allowing applications to continue processing in the presence of array failures and maintenance operations. The distributed RAID 1 technology extends this protection further, allowing clustered active/active applications to leverage the AccessAnywhere capabilities of VPLEX to ride through site disasters. The VPLEX Witness software

ensures data availability with zero recovery time objective (RTO) and zero recovery point objective (RPO) for clustered applications using a VPLEX Metro deployment.

Local VPLEX mirroring

VPLEX RAID 1 devices provide a local full copy RAID 1 mirror of a device independent of the host and operating system, application, and database. This mirroring capability allows VPLEX to transparently protect applications from back-end storage array failure and maintenance operations. In write-through cache mode, provided by VPLEX Local and VPLEX Metro, writes to a VPLEX virtual volume created from a RAID 1 device will have their data acknowledged by the back-end array supporting each leg of the mirror before VPLEX acknowledges the host. This establishes very strong durability guarantees on the data. In write-back cache mode, provided by VPLEX Geo, a VPLEX director acknowledges host writes only after protecting the data by placing a copy of the data in the memory of a peer director. Later the data is destaged by writing it to both legs of the RAID 1 device.

Distributed VPLEX mirroring

VPLEX Metro and VPLEX Geo support distributed mirroring that protects the data of a virtual volume by mirroring it between the two VPLEX clusters.

- **Distributed RAID 1 volumes with write-through caching:** Maintains a real-time synchronized mirror of a VPLEX virtual volume between the two clusters of the VPLEX system, providing a RPO of zero data loss and concurrent access to the volume through either cluster. In this cache mode the VPLEX clusters must be deployed within synchronous distance of each other with an inter-cluster message RTT of 5 milliseconds or less.
- **Distributed RAID 1 volumes with write-back caching:** Maintains a near-real-time synchronized mirror of a VPLEX virtual volume between two VPLEX clusters, providing a RPO that could be as short as a few seconds.
- **VPLEX consistency groups:** Ensures application-dependent write consistency of application data on VPLEX distributed RAID 1 volumes within the VPLEX system in the event of a disaster, providing for a business point of consistency for disaster restart for all identified applications associated with a business function.

Collaboration

When customers have tried to build collaboration across distance with the traditional solutions, they normally have to save the entire file at one location and then send it to another site using FTP. This is slow, can incur heavy bandwidth costs for large files, or even small files that move regularly, and negatively impacts productivity because the other sites can sit idle while they wait to receive the latest data from another site. If teams decide to do their own work independent of each other, then the data set quickly becomes inconsistent, as multiple people are working on it at the same time and are unaware of each other's most recent changes. Bringing all of the changes together in the end is time-consuming and costly, and grows more complicated as the data set gets larger.

But fortunately, this is also a unique, valuable, and highly differentiated opportunity for VPLEX.

With VPLEX, the same data can be accessible to all users at all times—even if they are at different sites. The data is literally shared, not copied, so that a change made in one site shows up right away at the other site. This is a huge benefit for customers who have always had to rely on shipping large log files and data sets back and forth across sites, then wait for updates to be made in another location before they could resume working on it again.

VPLEX makes the data available in both locations, and because VPLEX is smart, it doesn't need to ship the entire file back and forth like other solutions—it only sends the change updates as they are made, thus greatly reducing bandwidth costs and offering significant savings over other solutions. Deploying VPLEX in conjunction with third-party WAN optimization solutions can deliver even greater benefits. And with VPLEX's AccessAnywhere, the data remains consistent, online, and available—always.

Migration into and out of VPLEX

Non-disruptive migration from non-virtualized storage into VPLEX storage is supported through host migration utilities provided by host-based volume managers and services such as EMC's PowerPath® Migration Enabler (PPME) with host copy or VMware's Storage vMotion®. These migration methods use the host or server to read the data from its current physical location and copy the data into new virtual storage provided by VPLEX. Likewise, the same procedure can be used to move the data from VPLEX to a new physical location when removing the virtual storage layer is desired. Of course, once storage has been virtualized by VPLEX, moving its physical location within the data center or between data centers is easy, fast, and non-disruptive to applications.

VPLEX also supports a fast offline method for migrating from physical to virtual storage and from virtual to physical storage. The physical-to-virtual method involves fully encapsulating the physical device in a VPLEX storage volume and creating a virtual volume from this encapsulation. The virtual volume takes on a new volume identity, but its contents remain the same and any host-based volume labels are preserved. When mapped, masked, and zoned to the host the volume becomes available again for use without any need to move the physical location of the data. This makes the process very quick and resource-efficient, allowing downtime to be minimized for the migration. Once virtualized, all of the VPLEX capabilities for virtual storage are available to the volume such as mirror-protecting the volume between two arrays, or making the volume available to a remote data center. VPLEX does not alter the contents of the volume by adding headers or other such metadata, so the process of de-encapsulating a volume to make the transition from virtual to physical storage is equally straight-forward, and is essentially the inverse of the encapsulation method.

VPLEX System Management Service

The VPLEX System Management Service provides the interfaces used to configure and manage a VPLEX system. This software is loaded directly onto the VPLEX Management Console and allows any web browser with proper security credentials to manage the system from anywhere in the enterprise. The software even allows the administrator to manage the remote cluster of a VPLEX Metro or VPLEX Geo system via the System Management Software running on the local VPLEX Management Console.

VPLEX support for Virtual Provisioning

VPLEX leverages the Virtual Provisioning™ (also known as thin provisioning) capabilities of the back-end storage arrays for its virtual storage; virtual volumes that are constructed from virtually provisioned storage volume are themselves virtually provisioned. That is, the actual storage allocated on a back-end storage volume is a function of the written portions of the storage volume, rather than the advertised capacity of the storage volume.

When establishing a RAID 1 or distributed RAID 1 device, the synchronization method can be set to detect and suppress zero valued data regions. This reduces the amount of time to synchronize a new mirror leg and can significantly reduce the volume of data sent over remote links. This synchronization setting is available for both virtually and fully provisioned storage volumes.

Summary

Binding service-level agreements (SLAs) commit IT organizations to deliver stipulated, measurable support metrics such as application performance, end-user response time, and system availability. Even in the absence of such SLAs, IT executives universally recognize that downtime can have disastrous ramifications in lost revenue, dissatisfied customers, and missed opportunities. VPLEX's superior clustering architecture complements and extends the high-availability characteristics provided by your back-end storage arrays by adding the ability to mirror data between arrays, and between data centers both near and far. VPLEX systems are the logical choice when your virtual storage layer must provide the most uncompromising levels of data and system availability.

VPLEX Local, Metro, and Geo

VPLEX supports three different configurations to suit different needs. The next few sections describe these configurations and when they should be used.

VPLEX Local

VPLEX Local systems are supported in single, dual, or quad configurations consisting of one, two, or four engines, respectively, yielding systems that provide two, four, or eight directors. At the time of purchase, a sizing tool can be used to determine the proper configuration for a given deployment. This tool takes into consideration the

type and amount of I/O activity that occur in the target environment and matches this to the appropriate configuration.

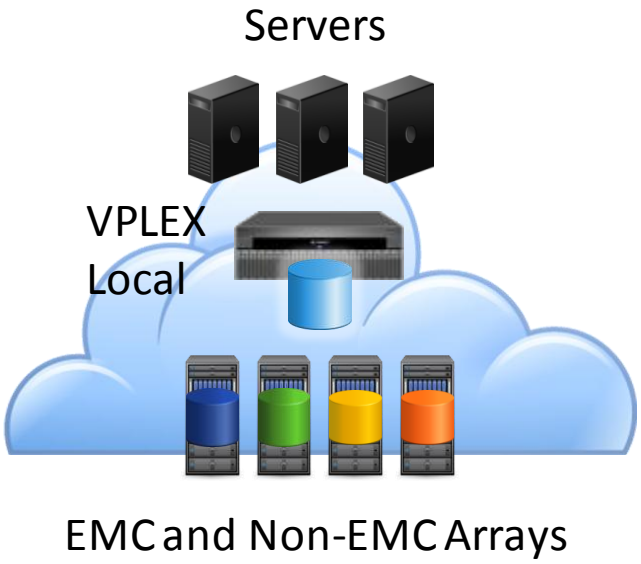

Figure 5. Example of a VPLEX Local single deployment

When to use a VPLEX Local deployment

VPLEX Local is appropriate when the virtual storage capabilities of workload mobility, workload availability, and simplified storage management are desired within a single data center and the scaling capacity of VPLEX Local is sufficient to meet the needs of this data center. If a larger scale is needed, consider deploying VPLEX Metro, or consider deploying multiple independent instances of VPLEX Local.

VPLEX Metro

VPLEX Metro systems contain two clusters, each cluster having one, two, or four engines. The clusters in a VPLEX Metro deployment need not have the same number of engines. For example, a VPLEX Metro system could be composed of one cluster with two engines and the other with four.

The two clusters of a VPLEX Metro must be deployed within synchronous communication distance of each other (about 5 ms of RTT communication latency). VPLEX Metro systems are often deployed to span between two data centers that are close together (roughly up to 100 km or 60 miles apart), but they can also be deployed within a single data center for applications requiring a high degree of local availability.

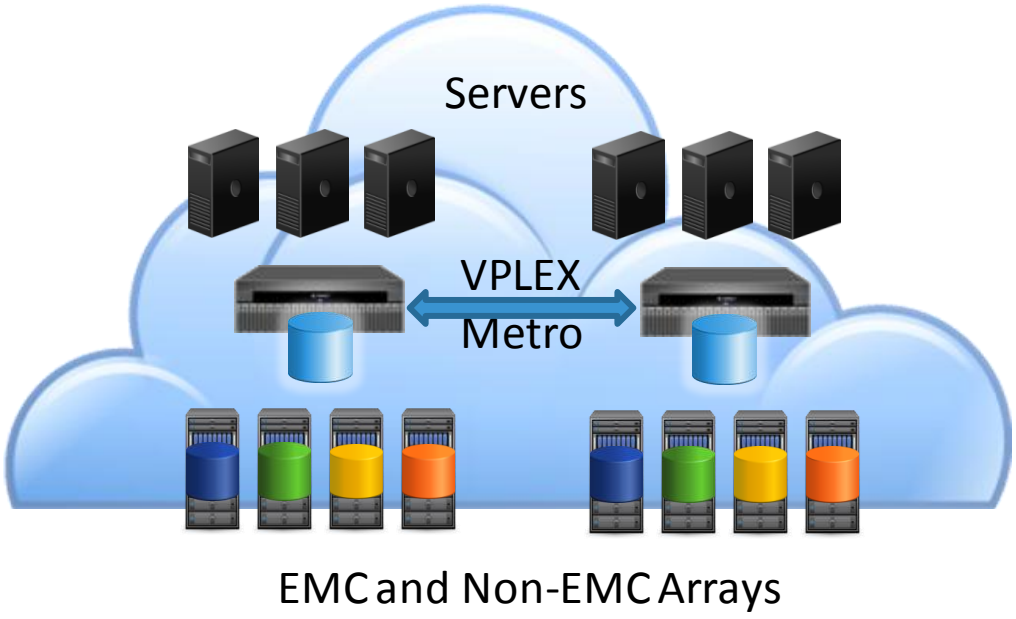

Figure 6. Example of a VPLEX Metro deployment within a data center

VPLEX Witness

The optional VPLEX Witness is recommended to improve application availability and data access in the presence of VPLEX cluster failures or the total loss of inter-cluster communication.

VPLEX virtual volumes can be mirrored between the VPLEX clusters, allowing a host to have access to the data through either cluster. This provides added resiliency in the case of an entire cluster failure. In such a deployment, on a per-consistency group basis, one cluster is designated as the preferred cluster for data availability. Should the redundant communication between the VPLEX clusters be lost, but connectivity with the VPLEX Witness retained, the VPLEX Witness will indicate to the clusters that the preferred cluster should continue providing service to the volumes in the consistency group. In this situation, the non-preferred cluster will stop servicing the volumes, until such time as the link is restored, and the mirrors are re-established. Should the preferred cluster of a consistency group fail, the VPLEX Witness will indicate this failure to the non-preferred cluster, which will continue to provide access to the volumes in the group. Likewise, in the event of the failure of the non-preferred cluster, the Witness will direct the preferred cluster to continue to service the volumes. This prevents a partition between the two clusters from allowing the state of the volumes to diverge; this avoids the well-known split-brain problem.

The use of the Witness is recommended since it increases the overall availability of data in the presence of these failures. When the VPLEX Witness is not deployed, the system will suspend I/O to a volume when that volume's preferred cluster fails.

In VPLEX 5.0 Witness functionality applies only to distributed volumes that are placed in consistency groups. Distributed volumes that are not placed in a consistency

group have their own independent bias settings. These volumes will have their I/O suspended when their preferred cluster fails.

It is very important to deploy the VPLEX Witness into a failure domain that is independent of each of the failure domains containing the two VPLEX clusters, to ensure that a single failure impacts no more than one of these entities.

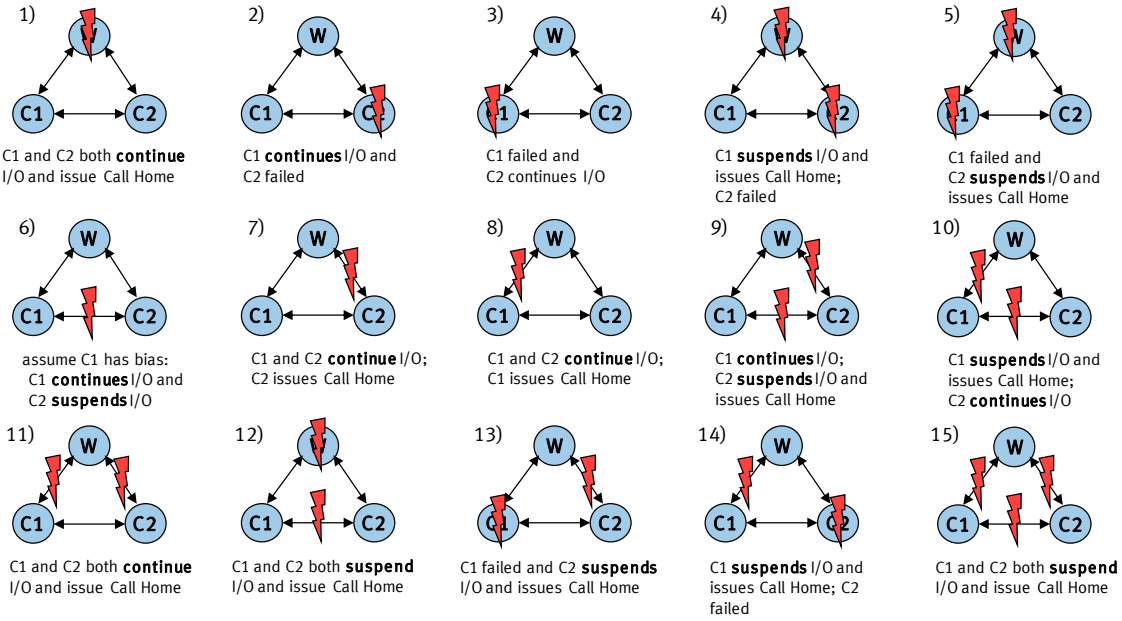

Figure 7. Failure states with the VPLEX Witness

When to use a VPLEX Metro deployment within a data center

Deploying VPLEX Metro within a data center is appropriate when the virtual storage capabilities of mobility, availability, and simplified storage management are desired within a single data center and more scaling is needed beyond that of a VPLEX Local solution or when additional availability is desired.

VPLEX Metro provides the following additional availability benefits over VPLEX Local:

- The two clusters of a VPLEX Metro can be separated by a distance with up to 5 ms RTT latency. This provides excellent flexibility for deployment within a data center and allows the two clusters to be deployed at separate ends of a machine room or on different floors to provide better fault isolation between the clusters. For example, this allows the clusters to be placed in different fire suppression zones, which can mean the difference between riding through a localized fault such as a contained fire and a total system outage.
- The VPLEX Witness allows a VPLEX system to provide continuous availability to data in the presence of total cluster failures and loss of the redundant communication links between clusters. Figure 7 shows how availability is improved with a VPLEX Witness beyond the increased availability of a VPLEX Metro with no Witness.

VPLEX Metro deployment between data centers

In a VPLEX Metro system deployed between two data centers, hosts typically connect only to their local cluster. Clustered applications can have, for example, one set of application servers deployed in data center A, and another set deployed in data center B for added mobility, availability, and collaboration benefits. As described in the previous section, it is important to understand the role that the VPLEX Witness plays in preserving data availability in the presence of cluster failures and inter-cluster communication loss.

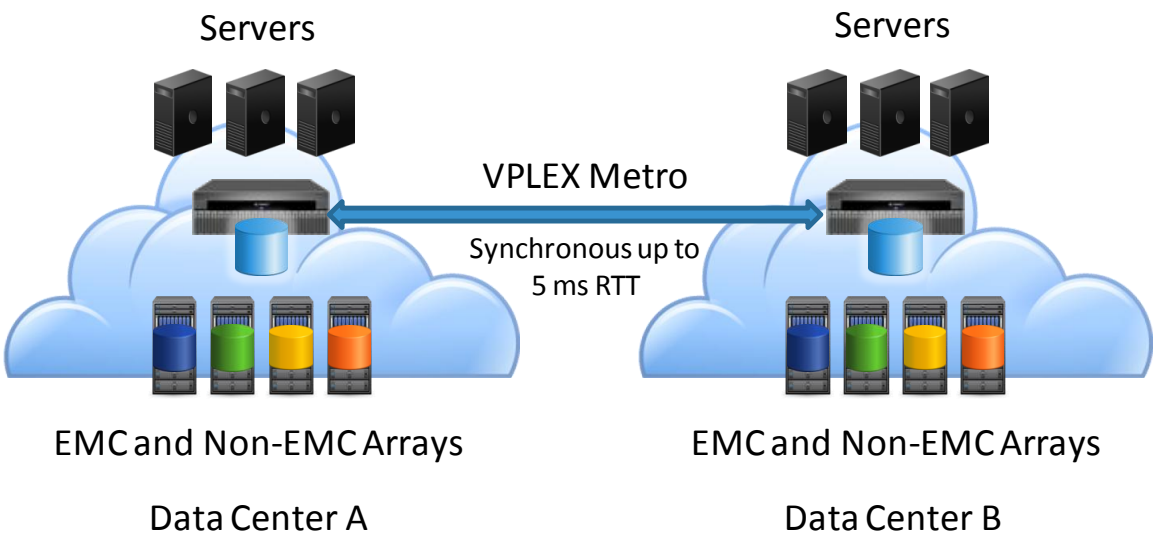

Figure 8. Example of a VPLEX Metro deployment between data centers

When to use a VPLEX Metro deployment between data centers

A deployment of VPLEX Metro between two data centers is recommended for:

- **Mobility:** One wants to redistribute application workloads between the two data centers.
- **Availability:** An application needs to keep running in the presence of data center failures.
- **Collaboration:** Applications in one data center need to access data in the other data center.
- **Distribution:** One data center has run out of space, power, or cooling.

VPLEX Geo

VPLEX Geo, the newest member of the VPLEX family, enables AccessAnywhere between two sites over *asynchronous* distances up to 50 ms RTT. VPLEX Geo now allows for workload mobility and collaboration at greater distances. Like VPLEX Metro, VPLEX Geo systems also contain two clusters, each cluster having one, two, or four engines, and the clusters need not have the same number of engines.

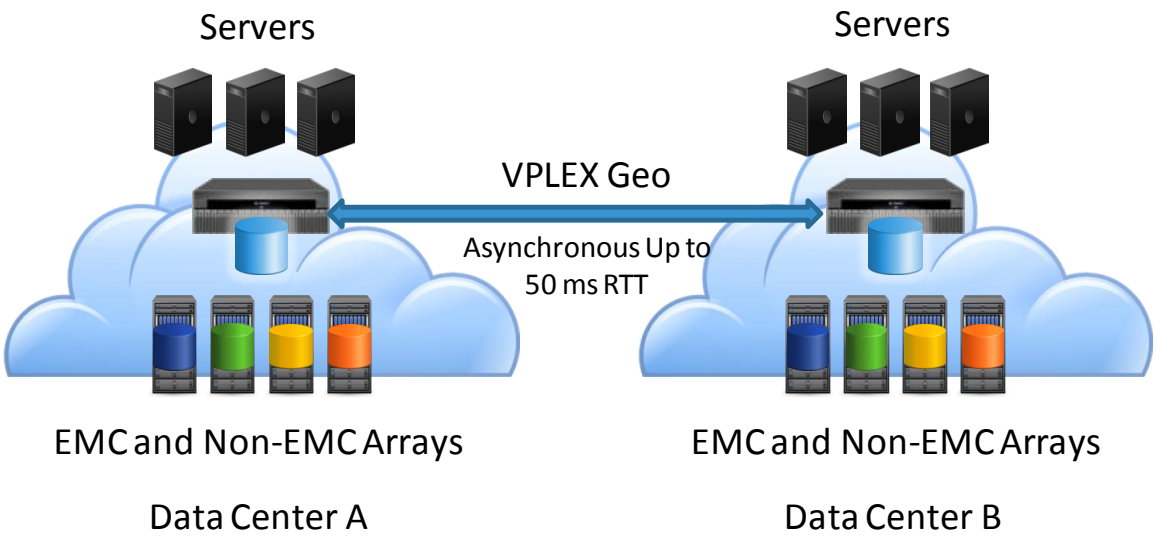

Figure 9. Example of a VPLEX Geo deployment between data centers

VPLEX Geo uses an IP-based protocol over the Ethernet WAN COM link that connects the two VPLEX clusters. This protocol is built on top of the UDP Data Transfer (UDT) protocol, which runs on top of the Layer 3 User Datagram Protocol (UDP). UDT was designed for high-bandwidth data transfer and is well suited to the VPLEX WAN COM transport.

Availability and system integrity

VPLEX system hardware is the most reliable virtual storage system hardware in the industry. However, all hardware is subject to occasional failures. VPLEX provides unique methods that proactively detect and prevent failures from impacting customer operations. Additionally the VPLEX clustering architecture provides component redundancy even in the presence of individual failures to support the availability needs of your most demanding mission-critical applications.

In the next few sections we study different faults that can occur in a data center and look at how VPLEX can be used to add additional availability to applications, allowing their workload to ride through these fault conditions. The following classes of faults and service events are considered:

- Storage array outages (planned and unplanned)
- SAN outages
- VPLEX component failures
- VPLEX cluster failures
- Host failures
- Data center outages

Storage array outages

To overcome both planned and unplanned storage array outages, VPLEX supports the ability to mirror the data of a virtual volume between storage volumes from different arrays using a RAID 1 device. Should one array incur an outage, either planned or unplanned, the VPLEX system will be able to continue processing I/O on the surviving mirror leg. Upon restoration of the failed storage volume, the data from the surviving volume is resynchronized to the recovered leg.

Best practices

For critical data, EMC recommends mirroring data onto storage volumes that are provided by separate arrays.

For the best performance, these storage volumes should be configured identically and be provided by the same type of array.

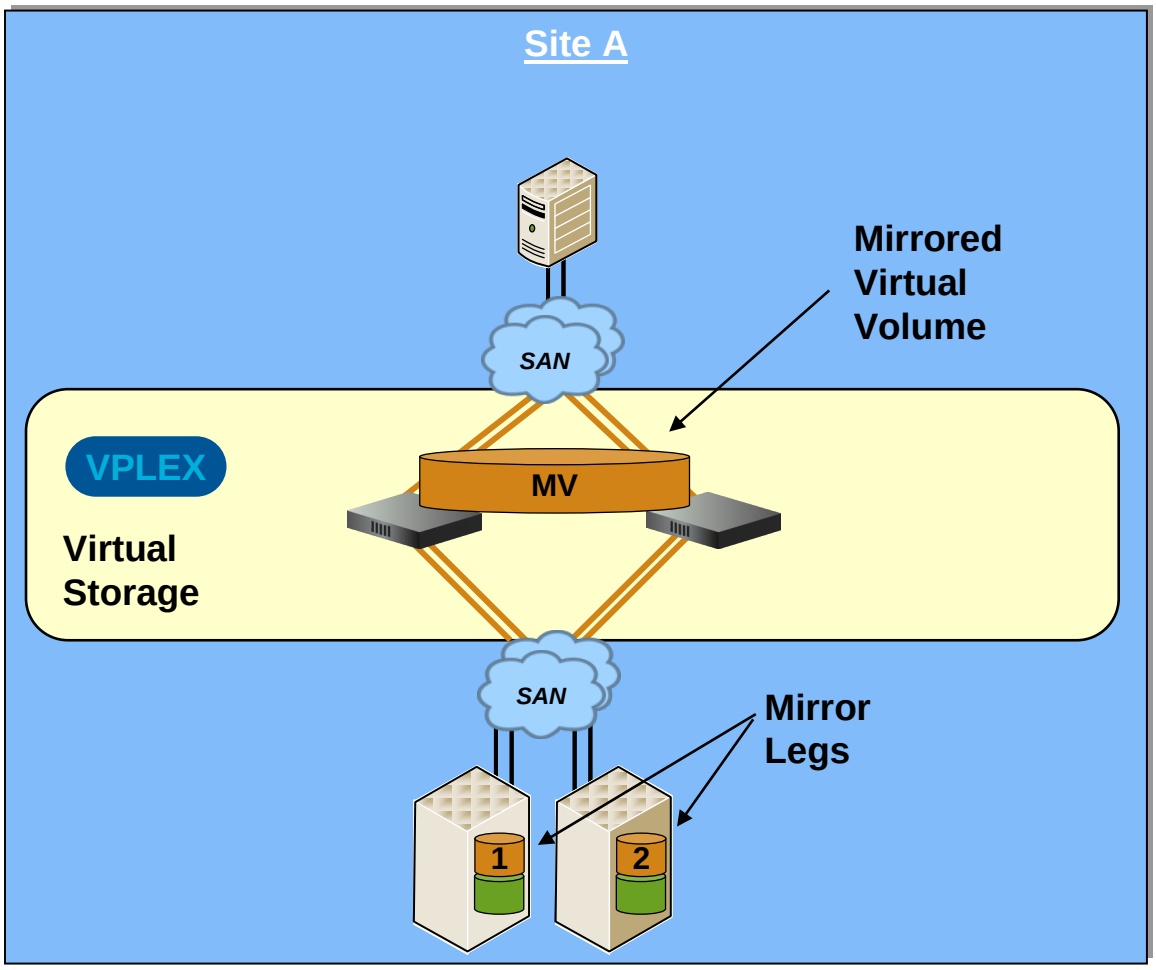

Figure 10. Example of RAID 1 mirroring to protect against array outages

SAN outages

When a pair of redundant Fibre Channel fabrics is used with VPLEX, VPLEX directors should be connected to both fabrics both for the front-end (host-side) connectivity, as

well as for the back-end (storage array side) connectivity. This deployment, along with the isolation of the fabrics, allows the VPLEX system to ride through failures that take out an entire fabric and allows the system to provide continuous access to data through this type of fault. Hosts must also be connected to both fabrics and use multipathing software to ensure continuous data access in the presence of such failures.

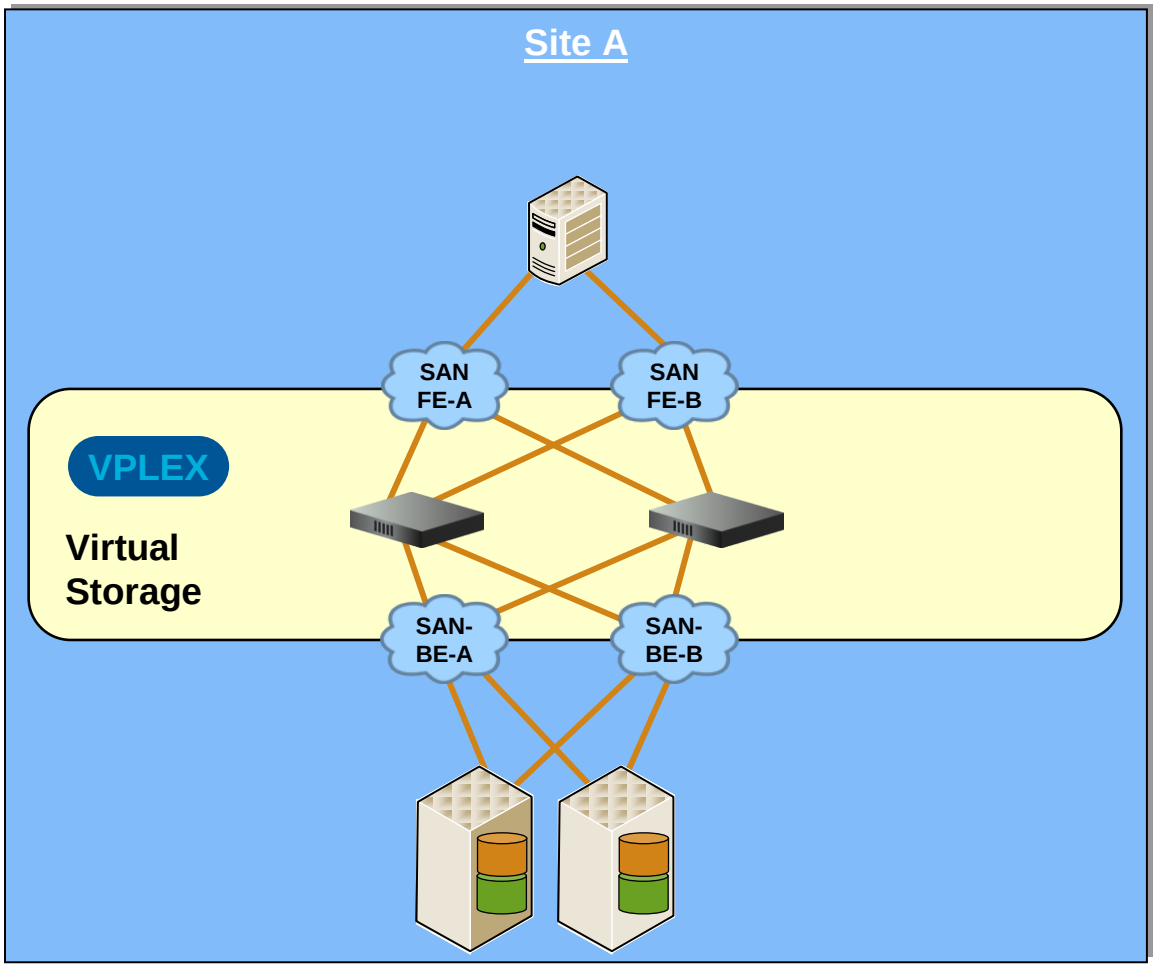

Figure 11. Recommended use of a dual-fabric deployment

Best practice

It is recommended that I/O modules be connected to redundant A and B fabrics.

VPLEX component failures

All critical processing components of a VPLEX system use at a minimum pair-wise redundancy to maximize data availability. This section describes how VPLEX component failures are handled and the best practices that should be used to allow applications to tolerate these failures.

All component failures that occur within a VPLEX system are reported through events that call back to EMC Support to ensure timely response and repair of these fault conditions.

Fibre Channel port failure

VPLEX supports redundant paths for all communications, via redundant ports on the front end, back end, and within and between clusters, allowing communication to continue in the presence of port failures. This redundancy allows multipathing software in the host servers to retransmit and redirect I/O around path failures that occur as a result of port failures or other events in the SAN that lead to path loss.

VPLEX uses its own multipathing logic to maintain redundant paths to back-end storage from each director. This allows VPLEX to ride through port failures on the back-end VPLEX ports as well as on the back-end fabrics and the array ports that connect the physical storage to VPLEX.

The small form-factor pluggable (SFP) transceivers that are used for connectivity to VPLEX are serviceable field replaceable units (FRUs).

Best practices

- Ensure there is a path from each host to at least one front-end port on director A and at least one front-end port on director B. When the VPLEX system has two or more engines, ensure that the host has at least one "A-side" path in one engine and at least one "B-side" in a separate engine. For maximum availability, each host can have a path to at least one front-end port on every director.
- Use multipathing software on the host servers to ensure timely response and continuous I/O in the presence of path failures.
- Ensure that each host has a path to each virtual volume through each fabric.
- Ensure that the LUN mapping and masking for each storage volume presented from a storage array to VPLEX present the volumes out of at least two ports from the array on at least two different fabrics and connects to at least two different ports serviced by two different back-end I/O modules of each director within a VPLEX cluster.
- Ensure that the fabric zoning provides hosts redundant access to the VPLEX front-end ports and provides VPLEX the redundant access to the array ports.

I/O module failure

I/O modules within VPLEX serve dedicated roles. Each VPLEX director has two front-end I/O modules, two back-end I/O modules, and one COM I/O module used for intra- and inter-cluster connectivity. Each I/O module is a serviceable FRU. The following sections describe the behavior of the system and best practices for maximizing availability in the presence of these failures.

FE I/O module

Should an FE I/O module fail, all paths connected to this I/O module will be disrupted and fail. The best practices on page 24 should be followed to ensure that hosts have a redundant path to their data.

During the removal and replacement of an I/O module, the affected director will restart.

BE I/O module

Should a BE I/O module fail, all paths connected to this I/O module will be disrupted and fail. The best practices on page 24 should be followed to ensure that each director has a redundant path to each storage volume through a separate I/O module.

During the removal and replacement of an I/O module, the affected director will restart.

COM I/O module

Should the COM I/O module of a director fail, the director will reset and all service provided from the director will stop. The best practices on page 24 ensure that each host has redundant access to its virtual storage through multiple directors, so the reset of a single director will not cause the host to lose access to its storage.

During the removal and replacement of an I/O module, the affected director will restart.

Director failure

A director failure causes the loss of all service from that director. Each VPLEX Engine has a pair of directors for redundancy. VPLEX clusters containing two or more engines benefit from the additional redundancy provided by the additional directors. Each director within a cluster is capable of presenting the same storage. The best practices on page 24 allow a host to ride through director failures by placing redundant paths to their virtual storage through ports provided by different directors. The combination of multipathing software on the hosts and redundant paths through different directors of the VPLEX system allows the host to ride through the loss of a director.

In a multiengine system, a host can maintain access to its data in the unlikely event that multiple directors should fail by having paths to its virtual storage provided by each director in the system.

Each director is a serviceable FRU.

Engine power supply and fan failure

The VPLEX Engine power supplies are fully redundant and no loss of service or function is incurred in the presence of a single power supply failure. Cooling is provided by fans that are integrated with the power supplies. The VPLEX VS2 enclosure contains four fans, and proper cooling can be sustained with three

operational units. Each power supply/fan unit is a serviceable FRU and can be removed and replaced with no disruption to the system.

Intra-cluster IP subnet failure

Each VPLEX cluster has a pair of private local IP subnets that connect the directors to the management console. These subnets are used for management traffic as well as for protection against intra-cluster partitioning. Link loss on one of these subnets can result in the inability of some members to communicate with other members on that subnet; this results in no loss of service or manageability due to the presence of the redundant subnet.

Intra-cluster Fibre Channel switch failure

Each VPLEX cluster with two or more engines uses a pair of dedicated Fibre Channel switches for intra-cluster communication between the directors within the cluster. Two redundant Fibre Channel fabrics are created with each switch serving a different fabric. The loss of a single Fibre Channel switch results in no loss of processing or service.

VPLEX Engine failure

In VPLEX clusters containing two or more engines, the unlikely event of an engine failure will result in the loss of service from the directors within this engine, but virtual volumes serviced by the directors in other surviving engines will remain available. The best practices on page 24 describe the best practice of placing redundant paths to a virtual volume on directors from different engines in multiengine VPLEX clusters.

Standby power supply failure

Each VPLEX Engine is supported by a pair of SPSs that provide a hold-up time of five minutes, allowing the system to ride through transient power loss up to 30 seconds. For power loss conditions lasting longer than 30 seconds, the engine will vault its dirty cache data and shut down. A single SPS provides enough power for the attached engine. VPLEX provides a pair of SPSs for high availability.

Each SPS is a FRU and can be replaced with no disruption to the services provided by the system. The recharge time for an SPS is up to 5.5 hours and the batteries in the SPS are capable of supporting two sequential five-minute outages.

Inter-cluster link failure

Each director in a VPLEX system has two links dedicated to inter-cluster communication. Each of these links should be configured (for example, zoned) to provide paths to each director in the remote cluster. In this manner, full connectivity between directors remains available even in the presence of a failure of a single link. Should one director lose both links, all inter-cluster I/O will be suspended between the two clusters to preserve write-order fidelity semantics and ensure that the remote site maintains a recoverable image. When this occurs, distributed volumes will be

fractured and access to each volumes restricted to one site (to prevent divergent data).

Access to these devices is governed by user-configured rules, referred to as *detach rules*. These rules govern which VPLEX cluster should continue to allow I/O for a given distributed volume. These rules can be configured on a per-device and per-consistency-group basis, allowing some volumes to remain available on one cluster, and other volumes to remain available on the other cluster. When the deployment contains a VPLEX Witness, the Witness will detect the partition between the two VPLEX clusters, and communicate to the clusters that a partition has occurred. This information instructs each cluster to use the bias settings of the detach rules to identify the cluster that should continue servicing each distributed volume and consistency group.

Once the link failures have been repaired, I/O between the clusters is restored and resynchronization tasks started to restore distributed volumes. These actions will take place automatically, or volumes can be configured to require manual resumption of I/O should coordination with server actions be required. I/O to the devices can take place immediately without needing to wait for the resynchronization tasks to complete.

Metadata volume failure

VPLEX maintains its configuration state, referred to as *metadata*, on storage volumes provided by storage arrays on the SAN. Each VPLEX cluster maintains its own metadata, which describes the local configuration information for this cluster as well as any distributed configuration information shared between clusters. It is strongly recommended that the metadata volume for each cluster be configured with multiple back-end storage volumes provided by different storage arrays of the same type. The data protection capabilities provided by these storage arrays, such as RAID 1 and RAID 5, should be used to ensure the integrity of the system's metadata. Additionally, it is highly recommended that backup copies of the metadata be made whenever configuration changes are made to the system.

VPLEX uses this persistent metadata upon a full system boot and loads the configuration information onto each director. When changes to the system configuration are made, these changes are written out to the metadata volume. Should access to the metadata volume be interrupted, the VPLEX directors will continue to provide their virtualization services using the in-memory copy of the configuration information. Should the storage supporting the metadata device remain unavailable, a new metadata device should be configured. Once a new device has been assigned and configured the in-memory copy of the metadata device maintained by the cluster will then be recorded out onto the new metadata device.

The ability to perform configuration changes is suspended when access to the persistent metadata device is not available.

Dirty region log failure

VPLEX Metro and VPLEX Geo use dirty region logs (DRL) for the following situations:

- To record information about which regions of a fractured distributed mirror have been updated while a mirror leg is detached. This information is kept for each such detached mirror leg.
- To record information about which regions of a local mirror leg could not be written due to inaccessibility of the backend storage.

Should one of these DRL volumes become inaccessible, the directors will record the entire leg as out of date and will require a full resynchronization of this leg of the volume once it is reattached to the mirror.

VPLEX Management Console failure

Each VPLEX cluster has a dedicated management console that connects to the local data center's management IP network and provides management access to the system. In VPLEX Metro and VPLEX Geo environments the management consoles of each cluster are inter-connected by a VPN that traverses the inter-data-center IP network to provide remote management and distributed resource configuration. When a VPLEX Witness is deployed this VPN is extended with point-to-point tunnels from each management console to the VPLEX Witness. In deployments without a VPLEX Witness, I/O processing of the VPLEX directors has no dependency upon the management consoles, hence, the loss of a management console will not interrupt the I/O processing and virtualization services provided by VPLEX. When the VPLEX Witness is deployed, connectivity with the Witness will be lost in the event of a VPLEX Management Console failure, in such a situation, the Witness will consider this cluster to be out of service. Should the remote cluster fail at this point, local access to distributed consistency groups will be suspended. Likewise in this state, should WAN COM communication be lost with the remote cluster from this cluster, the VPLEX Witness will direct the remote cluster to continue to service distributed consistency groups.

Uninterruptible power supply failure

In VPLEX clusters containing two or more engines, a pair of Fibre Channel switches supports intra-cluster communication between the directors in these engines. Each switch has a dedicated UPS that provides backup power in the case of transient power loss. The UPS units will allow the Fibre Channel switches to continue operating for up to five minutes following the loss of power. The lower UPS in the rack also provides backup power to the management console.

VPLEX cluster failure

VPLEX Metro supports two forms of distributed devices: distributed volumes and remote volumes. Distributed volumes provide synchronized copies (mirrors) of the volume's data in each cluster. The mirrored volume appears and behaves as a single volume and acts in a similar manner to a virtual volume that uses a RAID 1 device, but

with the added value that each cluster maintains a copy of the data. Remote volumes provide access to a virtual volume whose data resides in one cluster. Remote volumes, like distributed volumes, are able to take advantage of the VPLEX distributed coherent cache and its prefetch algorithms to provide better performance than a SAN-extension solution. VPLEX Geo supports distributed volumes but not remote volumes, since the long latencies for remote access would introduce too much delay for most applications.

Detach rules define the behavior of distributed volumes in situations where one cluster loses communication with the other cluster. For distributed volumes that are placed within a consistency group, the detach rule is applied at the consistency group level, and the rule applies to all volumes in the group. For distributed volumes that are outside of a consistency group, the rule is applied to the individual volume.

Detach rules indicate which cluster should detach its mirror leg (removing it from service) in the presence of communication loss between the two clusters. These rules effectively define a winning leg should the clusters lose communication with each other.

There are two conditions that can cause the clusters to lose communication: one is inter-cluster link failures discussed in Inter-cluster link failure and the other is a cluster failure. This section describes the latter class of failure. The behavior in this scenario depends on whether the system is a VPLEX Metro or a VPLEX Geo and whether the system is deployed with a VPLEX Witness or not.

VPLEX Metro without the VPLEX Witness

In the presence of cluster failure, any distributed volume that had a detach rule that identified the surviving site as the winning site will continue to have I/O serviced on the surviving leg of the device. Those volumes whose detach rules declared this site to detach in the case of communication loss will have their I/O suspended. Due to the inability to distinguish a cluster failure from a link failure, this behavior is designed to preserve the integrity of the data on these distributed devices. In this situation, I/O can be manually resumed by the administrator.

There are two failure cases for remote virtual volumes. First, should the cluster that supplies the physical media for the virtual volume fail, the remote virtual volume will be completely inaccessible. For the second scenario, should the remote cluster fail (the cluster with no physical media for this volume), access to the virtual volume will remain available from the hosting cluster (the cluster with the physical data).

VPLEX Metro with the VPLEX Witness

In the presence of cluster failure, the surviving cluster will consult the VPLEX Witness and determine that its peer has failed; the surviving cluster will then continue to service I/O to its distributed volumes. Any bias setting is overridden in this state, since a cluster failure has been detected.

For remote virtual volumes, should the cluster that supplies the physical media for the virtual volume fail, the remote virtual volume will be completely inaccessible. If

the remote cluster should fail (the cluster with no physical media for this volume), access to the virtual volume will remain available from the hosting cluster (the cluster with the physical data).

VPLEX Geo without the VPLEX Witness

In the presence of cluster failure, any distributed volume that had a detach rule that identified the surviving site as the winning site will continue to service the volume if and only if the winner cluster is able to determine that the remote cluster was not active for this volume at the time of the failure. If the winning cluster cannot make this determination, the volume will be suspended to allow the administrator to make a choice between resuming access to the volume by resetting the volume's state to the last consistent state, or repairing the failure and resuming access once the failure has been repaired.

If the detach rule did not indicate bias for the surviving site, the volume will be suspended and administrative action will need to be taken to make a choice between resuming access to the volume by resetting the volume's state to the last consistent state (which may be the current state if the remote site was inactive at the time of the failure) or repairing the failure and resuming access to the volume at that time. If the detach rule does indicate a bias, the volumes in the consistency group are suspended and the data is rolled back to the last consistent point. The administrator must explicitly resume the volumes after restarting the applications that use these volumes.

VPLEX Geo with the VPLEX Witness

In the presence of cluster failure, the VPLEX Witness will indicate to the surviving cluster that its peer is considered out of service. If the volume is using the active detach rule and the surviving cluster is able to determine that the remote cluster was not active for this volume at the time of the failure, it will continue service to the volume automatically. If the surviving cluster cannot make this determination, the volume will be suspended to allow the administrator to make a choice between resuming access to the volume by resetting the volume's state to the last consistent state, or repairing the failure and resuming access once the failure has been repaired. If a detach rule with bias is used, the volumes in the consistency group are suspended and the data is rolled back to the last consistent point. The administrator must explicitly resume the volumes after restarting the applications that use these volumes.

Host failures

While this is not a capability provided by VPLEX, host-based clustering is an important technique for maximizing workload resiliency. With host-based clustering, an application can continue providing service in the presence of host failures using either an active/active or an active/passive processing model. When combined with the capabilities of VPLEX listed previously, the result is an infrastructure capable of providing very high degrees of availability.

Data center outages

VPLEX processing in the presence of data center outages behaves in the same manner as described in VPLEX cluster failure. A data center outage can be tolerated with no loss of data or access to distributed volumes for VPLEX Metro deployments when the VPLEX Witness is used and it is deployed into an independent failure domain from each of the VPLEX clusters. When this deployment is combined with failover logic for host clusters, this provides infrastructure that is able to restore service operations quickly, sometimes with no loss of application service, even in the presence of an unplanned data center outage.

Some data center outages are caused by the loss of power in the data center. VPLEX uses standby and UPSs to overcome transient losses of power lasting 30 seconds or less. This must be combined with similar supporting infrastructure for the hosts, network equipment, and storage arrays for a comprehensive solution for tolerating transient power loss. For power-loss conditions lasting longer than 30 seconds, VPLEX will stop providing virtualization service, and the write-back caching mechanism of VPLEX Geo will vault any dirty cache data. Volumes using write-through caching have their data protected from power loss by the back-end storage arrays supporting those volumes.

Non-disruptive upgrades

VPLEX Local and Metro support the ability to non-disruptively upgrade the system from one software version to the next. This upgrade takes place by dividing the system into two sets of directors, the A directors, referred to as the first upgraders, and the B directors, referred to as the second upgraders. The non-disruptive upgrade process requires host-based multipathing software to have a path from each host to both a first upgrader and a second upgrader. During the upgrade process, the first upgraders are first brought offline; this causes a path outage to the hosts on the A directors. Host-based multipathing software then issues I/O to the surviving path(s) on the B directors. The first upgraders are then reimaged and reset to use the new VPLEX image. Once operational, the second upgraders are halted, and host-based multipathing software then fails over to the first upgraders. The second upgraders are then reimaged and reset to rejoin the system and re-establish full redundant path connectivity.

VPLEX Geo uses a similar process but first requires that the two VPLEX clusters be disconnected and upgraded individually using the process described previously. The clusters are then rejoined once the individual clusters have been upgraded. During the upgrade, access to distributed volumes is allowed on only one cluster, and update tracking is performed using DRLs during this process so that the distributed volumes can be quickly resynchronized using VPLEX's incremental resync technology.

VPLEX distributed cache and protection and redundancy

VPLEX utilizes the individual director's memory systems to ensure durability of user and critical system configuration data. User data is made durable in one of two ways depending on the cache mode used for the data. The VPLEX write-through cache

mode leverages the durability properties of a back-end array by writing user data to the array and obtaining an acknowledgement for the written data before itself acknowledging the write back to the host. The VPLEX write-back cache ensures data durability by receiving user data into the cache memory of the director that received the I/O, then placing a protection copy of this data on a peer director before acknowledging the write to the host. This ensures the data is protected in two independent memories. The data is later destaged to back-end storage arrays that provide the physical storage media.

Global distributed cache protection from power failure

In the event of a power failure, VPLEX will copy non-destaged user data and critical metadata to the local solid state storage devices (SSDs) that are connected to each director.

User data in cache is protected if power is lost using standby power supplies. The VPLEX system uses SSDs local to each director to destage user data (dirty cache) and critical metadata during a sudden power-down or an unexpected power outage. This activity is called **vaulting**.

Vaulted images are fully redundant where the contents of dirty cache memory are saved twice to independent SSDs. VPLEX then completes the power-down sequence. Once power is restored, the VPLEX system startup program initializes the hardware and the environmental system, checks the data integrity of each vault, and restores the cache memory contents from the SSD vault devices. Upon completion of this recovery processing the distributed cache and critical metadata contents are restored.

The system resumes normal operation when the standby power supplies are sufficiently recharged to support another vault. If any condition is not safe, the system will call customer support for diagnosis and repair, and will remain in a suspended state. This allows EMC Customer Support to communicate with the VPLEX system and restore normal system operations.

Under normal conditions, the SPS batteries can support two consecutive vaults; this ensures that on power restore after the first power failure, that the system will be able to resume I/O immediately, and that it can still vault if there is a second power failure, enabling customer operations to resume without risking data loss.

EMC Remote Support

Through the VPLEX Management Console, the GeoSynchrony operating environment for VPLEX proactively monitors all end-to-end I/O operations for errors and faults. Through this monitoring the VPLEX GeoSynchrony software reports errors and other events of interest to EMC's remote support capabilities, which include remote notification and remote diagnostics and repair. Remote notification enables EMC to monitor the health of the VPLEX. If operational statistics fall outside a well-defined set of tolerances, or if certain error conditions are encountered, the VPLEX Management Console will automatically contact a support center to report its findings. Additionally,

EMC periodically establishes proactive remote support connections to verify that the system is responding and able to communicate with EMC. When an EMC support engineer is assigned to a service request or support ticket, they remotely access one of the VPLEX Management Consoles of the system in question to gather operational data and logs.

Summary

VPLEX provides extensive internal hardware and software redundancy that not only ensures high availability of the VPLEX services but that further improves upon the workload resiliency of the surrounding infrastructure. When combined with the best practices of host-based clustering, multipathing, fabric redundancy, storage media protection, and standby power infrastructure the resulting solution provides a solid foundation for ensuring that virtual storage provides a robust solution to the availability of storage.

Management and operations

The VPLEX system provides a management abstraction whereby back-end storage resources are virtualized to provide increased capabilities of mobility, availability, and collaboration. Inherent in the VPLEX architecture is the ability to both scale up from one to two to four engines per VPLEX cluster, but also to scale out from a VPLEX Local configuration with one VPLEX cluster to a VPLEX Metro or VPLEX Geo configuration with two VPLEX clusters. Scaling up a VPLEX system increases the IOPS, bandwidth capacity, and availability of the system, while scaling out increases the reach of VPLEX, further extending its mobility, availability, and collaboration capabilities.

To simplify management and operations in VPLEX system environments, VPLEX offers two primary interactive management interfaces as well as a programmatic interface that uses a REST-based protocol.

- **VPLEX GUI** is an intuitive, web-based interface used to monitor, configure, and control a VPLEX system. The VPLEX GUI provides a web interface to many of the VPLEX CLI operations. The GUI is supported by a web server that runs on the VPLEX Management Console of each VPLEX cluster. From the GUI, the operator is able to control either cluster and perform tasks such as provisioning virtual storage to a system local to either cluster as well as to create and manage distributed AccessAnywhere volumes that are available through each of the clusters of the VPLEX system.
- **VPLEX CLI** is an interactive context-based user interface that offers full control over the capabilities of the VPLEX system. The interface supports smart context-sensitive command completion, help, and pattern matching (globbing) that makes using the CLI fast and efficient for the interactive user. Like the GUI, the CLI is able to manage and monitor virtual storage for both the local and remote VPLEX clusters.

- **VPLEX Element Manager API** offers a REST-based interface to the VPLEX system supporting the full capabilities of the VPLEX CLI to program through this web-based interface. This provides the ability to develop task automation for services performed by VPLEX as well as the orchestration of higher-level functions for which VPLEX plays a part.

EMC Services

EMC Global Services

Maximize VPLEX benefits with EMC Global Services A

EMC delivers a full complement of services for EMC VPLEX system hardware and software to ensure that your VPLEX system performs as expected in your environment, while minimizing risk to your business and budget. Expert planning, design, and implementation services help you quickly realize the value of your hardware and software in your environment, no matter how simple or complex. After implementation, EMC's data migration services can help you plan, design, and safely migrate your critical data over any distance to your new system. EMC will also help integrate your new system into your information architecture and applications, such as Microsoft Exchange and SQL Server, Oracle databases and applications, and SAP, and manage your new environment when it is complete. Extensively trained professional services personnel and project management teams, leveraging EMC's extensive storage deployment best practices and guided by our proven methodology, accelerate the business results you need without straining the resources you have.

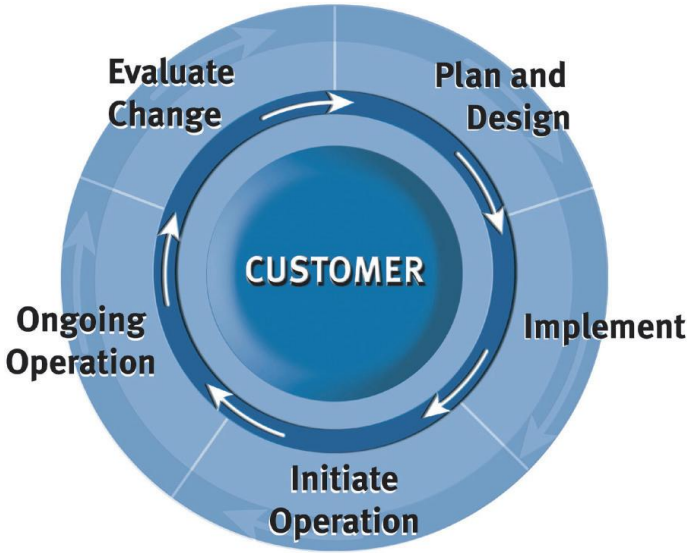

Figure 12. IT Lifecycle—EMC Global Services delivers results throughout the IT lifecycle: Plan, Build, Manage, and Support.

More than ever, organizations are demanding maximum value from their information assets and investments. EMC Global Services provides the strategic guidance and

technology expertise organizations need to address their business and information infrastructure challenges. EMC Global Services is committed to exceptional total customer experience through service excellence. Staffed with over 14,000 professional- and support-service experts worldwide, plus a global network of alliance partners, EMC Global Services delivers proven methodologies, instills industry best practices, and shares industry-leading experience and knowledge. EMC Global Services features rigorously tested and qualified solutions that reduce risk, lower costs, and speed time-to-value. Call upon EMC Global Services to address the full spectrum of requirements across the information lifecycle including strategic, advisory, architectural, implementation, management, and support services.

Proven methodologies

Every EMC Global Services engagement is guided by a proven methodology, the Global Delivery Model (GDM). The GDM ensures rapid, flawless implementation in every engagement around the world. EMC is committed to complete customer satisfaction. To this end, all EMC Global Services technologists undergo comprehensive training and certification in the industry's most advanced storage technology and implementation methodologies. As the leader in networked storage, EMC has the expertise and project management skills to ensure maximum value and minimal disruption during any networked storage engagement.

EMC Consulting Services

EMC Consulting helps customers deploy virtualized data center architectures and next-generation information infrastructures, integrating VPLEX systems into data center strategies to meet growing business demands for flexibility and improved service levels. We help customers understand the people, process, and technology dimensions of virtual data centers, including:

- Managing ROI and benefit expectations of business and application stakeholders
- Enhancing the IT service catalog to express the benefits of VPLEX systems
- Developing reference architectures and aligning application requirements based on tiers of service
- Designing and deploying operational processes optimized for virtual data centers

As part of the world's leading developer and provider of information infrastructure technology and solutions, EMC Consulting transforms information into business results. Our 2,700 consultants around the world bring an information-focused mix of industry, business, and technology expertise to solve today's toughest challenges. We use field-tested tools, proven methodologies, best practices, and industry standards to minimize risk and optimize time-to-value in our engagements.

EMC Residency Services

Residency Services provides experienced, specialized information infrastructure professionals at your site for a defined period of time to furnish the skills, technical knowledge, and the expertise you need to assist in day-to-day infrastructure

operations, management, and support. Services span the information lifecycle—from everyday operational storage tasks like provisioning and problem management to developing a long-term, ongoing information infrastructure strategy. Residency Services is designed to yield operational benefits while reducing costs. EMC infrastructure professionals leverage EMC's best practices library, EMC Knowledgebase, and extensive storage/application expertise to materially improve your infrastructure's operation. EMC's Residency Services portfolio includes Operational Residencies, Technology Residencies, Support Residencies, and Managed Residencies. Residents are skilled in technology areas such as SAN, NAS, CAS, open systems, Microsoft environments, and virtualization, as well as backup, recovery, and archive.

Benefits of EMC's Residency Services include:

- Improved operational efficiency
- Greater and faster return on storage and information assets
- Narrowed staff, skill, and/or experience gaps without additional headcount
- Increased internal customer satisfaction
- Expanded management and support
- Improved planning and operational insight

EMC service and support

All EMC storage platforms are backed by the world's leading services and support organization. The EMC support infrastructure includes more than 4,700 technical experts and more than 80 strategic authorized services network partners serving customers in more than 75 countries, with more than 35 strategically located support centers delivering "follow-the-sun" support.

Remote support

EMC VPLEX systems are equipped with automatic phone-home capabilities, so EMC service experts can monitor a system 24x7. By dialing back into the EMC system, they can take action quickly, analyzing events and abnormalities and resolving most issues before they affect business. Advanced remote support means a proactive and preemptive approach unmatched in the industry.

Software support

An all-inclusive software support and maintenance program ensures optimum availability of mission-critical information. EMC software specialists provide 24x7 telephone support to meet the needs of the most complex multivendor environment. Other EMC e-services like EMC Powerlink® and Knowledgebase make information, solutions, and software upgrades instantly accessible.

Online support with Powerlink

EMC provides online information to customers and partners through our Powerlink web portal. The latest information, specifications, white papers, customer bulletins, and much more can be found here.

Post-sale warranty and product support

Post-sale warranty coverage of VPLEX systems includes EMC's basic three-year hardware and 90-day software warranty plan with 24x7 coverage. Post-warranty service offerings include 24x7 coverage, technical support, and service and maintenance contracts.

Worldwide organization and local support

The EMC Customer Support Center, headquartered in the United States, directly supports EMC hardware and software products. Use the following numbers to contact EMC and obtain technical support:

- U.S.: (800) 782-4362 (SVC-4EMC)
- Canada: (800) 543-4782 (543-4SVC)
- Worldwide: 1 + (508) 497-7901 (or contact the nearest EMC office)

Conclusion

VPLEX 5.0 extends and enhances the capabilities of VPLEX 4.0 virtual storage with the addition of VPLEX Geo to the VPLEX product family. VPLEX Geo extends the reach of VPLEX AccessAnywhere storage, supporting RTT inter-site latencies of up to 50 ms through asynchronous communication.