

EMC MISSION-CRITICAL BUSINESS CONTINUITY AND DISASTER RECOVERY FOR ORACLE EXTENDED RAC

EMC VPLEX, EMC RecoverPoint, EMC VMAX, EMC VNX, Brocade Networking

- Simplified management for high availability, business continuity, and disaster recovery
- Resilient Oracle RAC deployments
- Active/active data centers

EMC Solutions Group

Abstract

This white paper describes a solution that increases availability for Oracle RAC environments by deploying EMC[®] VPLEX[™] active/active data centers and EMC RecoverPoint[™] continuous remote replication. Storage and networking for the solution is provided by EMC Symmetrix[®] VMAX[®] and EMC VNX[®] arrays and Brocade networking.

July 2012



Copyright © 2012 EMC Corporation. All Rights Reserved.

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

The information in this publication is provided “as is.” EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

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

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

VMware and ESXi are registered trademarks or trademarks of VMware, Inc. in the United States and/or other jurisdictions. Brocade, DCX, MLX, VCS, and VDX are registered trademarks of Brocade Communications Systems, Inc., in the United States and/or in other countries. All other trademarks used herein are the property of their respective owners.

Part Number H10746

Table of contents

Executive summary	5
Business case.....	5
Solution overview	5
Key benefits.....	6
Introduction	7
Purpose	7
Scope	7
Audience.....	7
Terminology.....	7
Solution overview	8
Introduction	8
HA and business continuity solution	9
Disaster recovery solution.....	11
Database and workload profile	12
Hardware resources	12
Software resources	13
EMC VPLEX Metro infrastructure	14
Introduction	14
VPLEX Metro solution configuration.....	17
EMC RecoverPoint configuration	19
Introduction	19
RecoverPoint and VPLEX.....	21
RecoverPoint CRR for VPLEX	23
RecoverPoint configuration	24
Oracle database architecture	28
Introduction	28
Oracle RAC and VPLEX.....	29
Oracle Extended RAC on VPLEX Metro.....	29
Brocade network infrastructure	31
Introduction	31
IP network configuration (Sites A and B).....	31
SAN configuration	32

- EMC storage infrastructure 33**
 - Overview 33
 - EMC Symmetrix VMAX 33
 - EMC VNX5700 33
 - EMC Unisphere 33

- Testing and validation 34**
 - Introduction 34
 - Validate the replica at Site C 35
 - Fail over to the replica at Site C 37
 - Validate the replica at Site A 39
 - Solution characterization data 40

- Conclusion 41**
 - Summary 41
 - Findings 41

- References 42**
 - EMC 42
 - Oracle 42

Executive summary

Business case

Global enterprises demand always-on application and information availability to remain competitive. The EMC solution described in this white paper offers a high-availability (HA) business continuity strategy for mission-critical Oracle Real Application Clusters (RAC) databases, including point-in-time disaster recovery (DR).

Recovery point objectives (RPOs) and recovery time objectives (RTOs) are key metrics when planning a business continuity strategy. They answer two fundamental questions for businesses considering the potential impact of a disaster or failure:

- How much data can we afford to lose (RPO)?
- How fast do we need the system or application to recover (RTO)?

Mission-critical business continuity demands aggressive RPOs and RTOs to minimize data loss and recovery times. The main challenges that the business must consider when designing a business continuity strategy with DR include:

- Minimizing RPO and RTO
- Eliminating single points of failure (SPOFs)—technology, people, processes
- Maximizing available points in time (PITs) for DR
- Reducing infrastructure costs
- Increasing resource utilization

This white paper introduces an EMC solution that addresses all of these challenges for Oracle RAC 11g databases.

Solution overview

EMC[®] VPLEX[™] and EMC RecoverPoint[™] are the primary enabling technologies for the solution.

- VPLEX is a SAN-based federation solution that delivers both local and distributed storage federation. Its breakthrough technology, AccessAnywhere[™], enables the same data to exist in two separate geographical locations, and to be accessed and updated at both locations at the same time.

With the optional VPLEX Witness component, applications continue to be available, with no interruption or downtime, even in the event of disruption at one of the data centers.

- RecoverPoint is replication technology that uses sophisticated journaling techniques and write splitting to provide local and remote replication. It provides fast and easy recovery to any point in time—a granularity not possible with other replication technologies.

VPLEX and RecoverPoint are complementary and integrated:

- VPLEX Metro provides the virtual storage layer that enables an active/active metro data center with 24/7 application availability, no single points of failure, and zero RPOs and RTOs.

- VPLEX Metro now has an integrated RecoverPoint write splitter. This enables the solution to leverage RecoverPoint to replicate Oracle data to a third site and to recover to a specific point in time in the event of data corruption, viruses, human error, or wide-scale physical disasters that affect both VPLEX Metro clusters.

This solution also takes advantage of both VPLEX and RecoverPoint support for heterogeneous storage platforms. EMC Symmetrix® VMAX® 20K and EMC VNX® 5700 arrays provide the enterprise-class storage platform for the two data centers, while an entry-level EMC Symmetrix VMAX 10K provides the storage at the third DR site.

Brocade Ethernet fabrics and MLXe core routers provide seamless networking and Layer 2 extension between the two data centers while Brocade DCX 8510 Backbones provide redundant SAN infrastructure, including fabric extension. The SAN at the DR site is built with Brocade 5100 switches.

EMC VPLEX Metro has passed Oracle's rigorous testing standards and is Oracle certified in a stretch cluster configuration. VPLEX Metro can provide Oracle Real Application Clusters (RAC) customers with an easy-to-deploy, active/active solution, as they transform from single- to dual-site environments. (See [Oracle RAC Technologies Matrix – Linux](#) and [Oracle RAC Technologies Matrix – Generic](#).)

Key benefits

The solution increases the availability of Oracle RAC databases by providing:

- Active/active data centers that support zero RPOs and RTOs and mission-critical business continuity
- Third-site replication and recovery to specific points in time in the event of data corruption, viruses, or human error
- Quick recovery and restart of your Oracle databases and applications

Additional VPLEX Metro benefits include:

- Fully automatic failure handling
- Increased utilization of hardware and software assets:
 - Active/active use of both data centers
 - Automatic load balancing between data centers
 - Zero downtime maintenance
- Simplified deployment of Oracle RAC on Extended Distance Clusters
- Reduced costs by increasing automation and infrastructure utilization

Additional RecoverPoint benefits include:

- DR testing without affecting production or interrupting replication
- Long distance third-site replication to meet regulatory or DR requirements
- Mature and proven technology for high-availability Oracle environments

Introduction

Purpose

This white paper describes a solution that increases availability for Oracle RAC environments by creating active/active data centers in geographically separate locations and enabling remote replication and recovery to specific points in time.

Scope

The scope of the white paper is to:

- Introduce the key enabling technologies for the solution
- Describe the solution architecture and design
- Outline how the key components are configured
- Present the results of the tests performed to validate the solution
- Identify the key business benefits of the solution

Audience

This white paper is intended for Oracle database administrators, storage administrators, IT architects, and technical managers responsible for designing, creating, or managing mission-critical Oracle deployments.

Terminology

This white paper includes the terms in Table 1.

Table 1. Terminology

Term	Description
ASM	Automatic Storage Management
CRR	Continuous remote replication
CNA	Converged network adapter
HA	High availability
HAIP	Highly available virtual IP
HBA	Host bus adapters
ISL	Inter-Switch Link
LAG	Link Aggregation Group
NL-SAS	Nearline SAS
OCR	Oracle Cluster Registry
Oracle Extended RAC	Oracle RAC on Extended Distance Clusters
PIT	Point in time
RAC	Real Application Clusters
RPA	RecoverPoint appliance
RPO	Recovery point objective
RTO	Recovery time objective
SAS	Serial Attached SCSI
VCS	Virtual cluster switch
vLAG	Virtual Link Aggregation Group
vLAN	Virtual LAN

Solution overview

Introduction

Today's businesses are faced with an ever-increasing amount of data that threatens to overwhelm their existing storage management solutions. Data protection is no longer the simple copying of yesterday's changed files to tape. Critical data changes occur throughout the day, and, to protect this data, customers are frequently turning to technology such as continuous remote replication (CRR).

This solution describes how to protect a mission-critical Oracle RAC environment on EMC VPLEX by using EMC RecoverPoint and VPLEX native write-splitting technology to provide CRR and point-in-time recovery.

The solution builds on a previous EMC solution¹ that uses VPLEX Metro, in combination with Oracle RAC on Extended Distance Clusters, to provide these benefits:

- Eliminate all SPOFs in the environment to build a highly available multi-site database
- Provide active/active data centers to enable mission-critical business continuity

Figure 1 shows the main steps in the evolution to the mission-critical business continuity and DR solution described in this white paper. Oracle RAC protects the database within the data center. VPLEX Metro combined with Oracle Extended RAC protects the database across data centers. RecoverPoint provides remote DR protection.

Journey to High Availability with Disaster Recovery

- a. Oracle RAC
- b. Oracle Extended RAC enabled by VPLEX Metro
- c. VPLEX Metro with RecoverPoint continuous remote replication

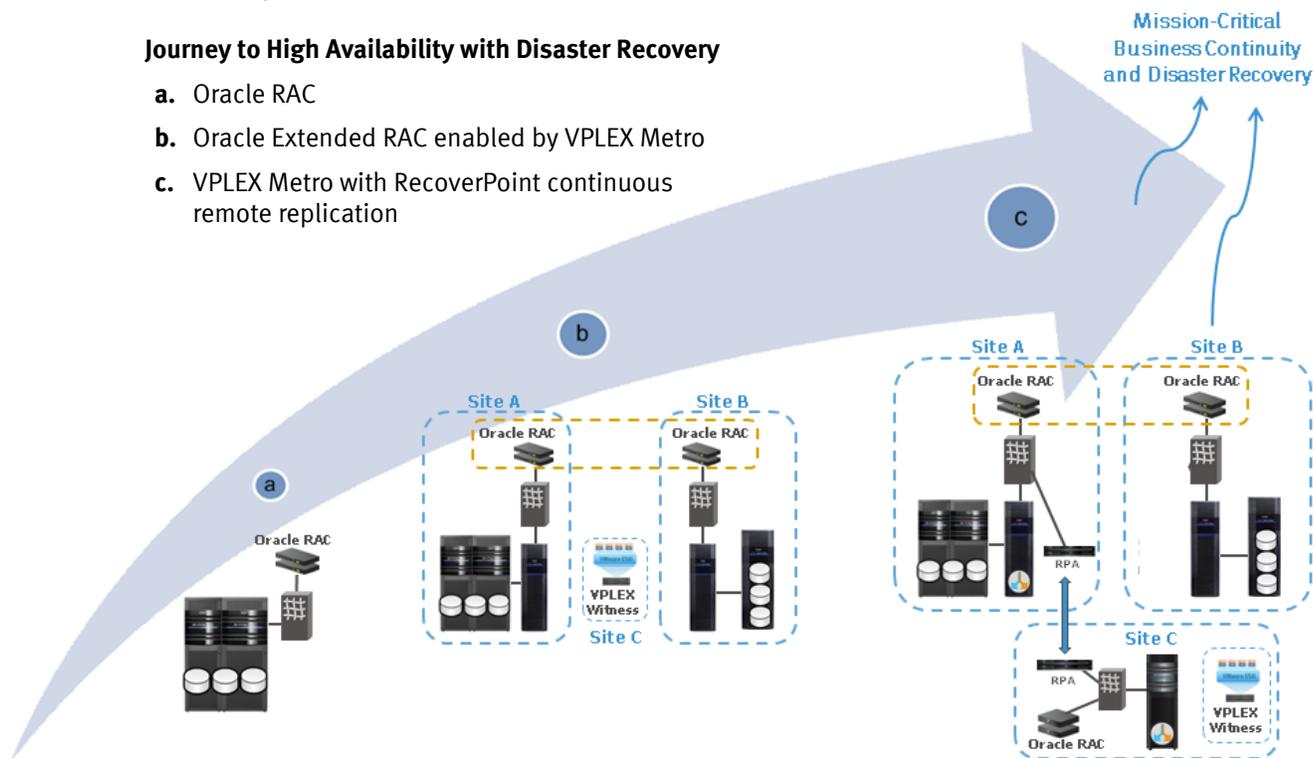


Figure 1. The journey to high availability with disaster recovery: logical view

¹ See the white paper: [EMC Mission-Critical Business Continuity](#)

Data center HA

Deploying an Oracle RAC environment eliminates the database server as a single point of failure and protects the database *within* the data center.

For high availability *between* data centers, the solution combines EMC VPLEX Metro storage virtualization technology with Oracle RAC on Extended Distance Clusters to remove the data center as a single point of failure, as shown in Figure 2. This provides a robust, high-availability and business continuity strategy.

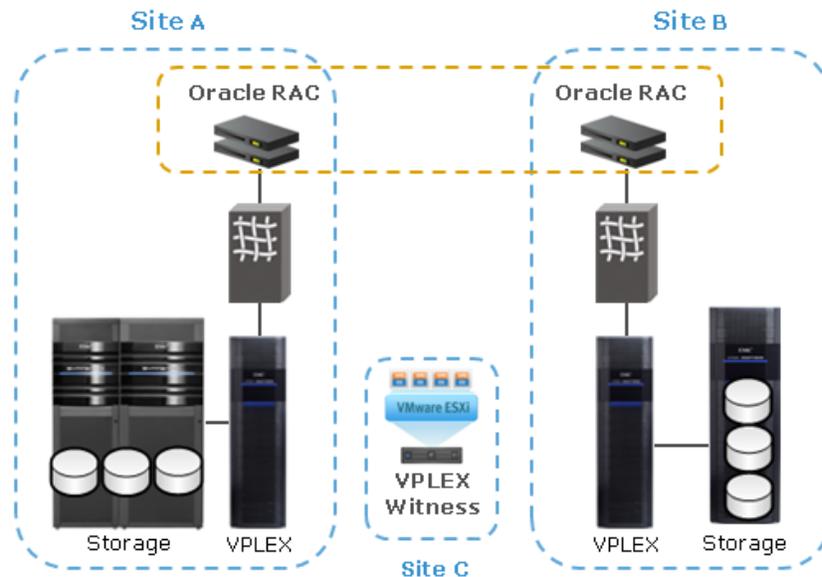


Figure 2. Data center HA

The solution uses VPLEX Witness to monitor connectivity between the two VPLEX clusters and ensure continued availability in the event of an inter-cluster network partition failure or a cluster failure. VPLEX Witness is deployed on a virtual machine at a third, separate failure domain (Site C).

Oracle RAC on Extended Distance Clusters over VPLEX provides these benefits:

- VPLEX simplifies management of Extended Oracle RAC as cross-site high availability is built in at the infrastructure level.
To the Oracle DBA, installation, configuration, and maintenance are the same as for a single site implementation of Oracle RAC.
- VPLEX eliminates the need for host-based mirroring of ASM disks and the host CPU cycles that this consumes.
With VPLEX, ASM disk groups are configured with external redundancy and are protected by VPLEX distributed mirroring.
- Hosts need to connect only to their local VPLEX cluster and I/O is sent only once from that node. However, hosts have full read/write access to the same database at both sites.

With host-based mirroring of ASM disk groups, each write I/O must be sent twice, once to each mirror.

- There is no need to deploy an Oracle voting disk on a third site to act as a quorum device at the application level.
- VPLEX enables you to create consistency groups that will protect multiple databases and/or applications as a unit.

Network HA

In each data center, an Ethernet fabric was built using Brocade virtual cluster switch (VCS) technology, which delivers a self-healing and resilient access layer with all links forwarding. Virtual Link Aggregation Groups (vLAGs) connect the VCS fabrics to the Brocade MLXe core routers that extend the Layer 2 network across the two data centers.

Figure 3 shows the physical architecture of all layers of the solution, including the network components.

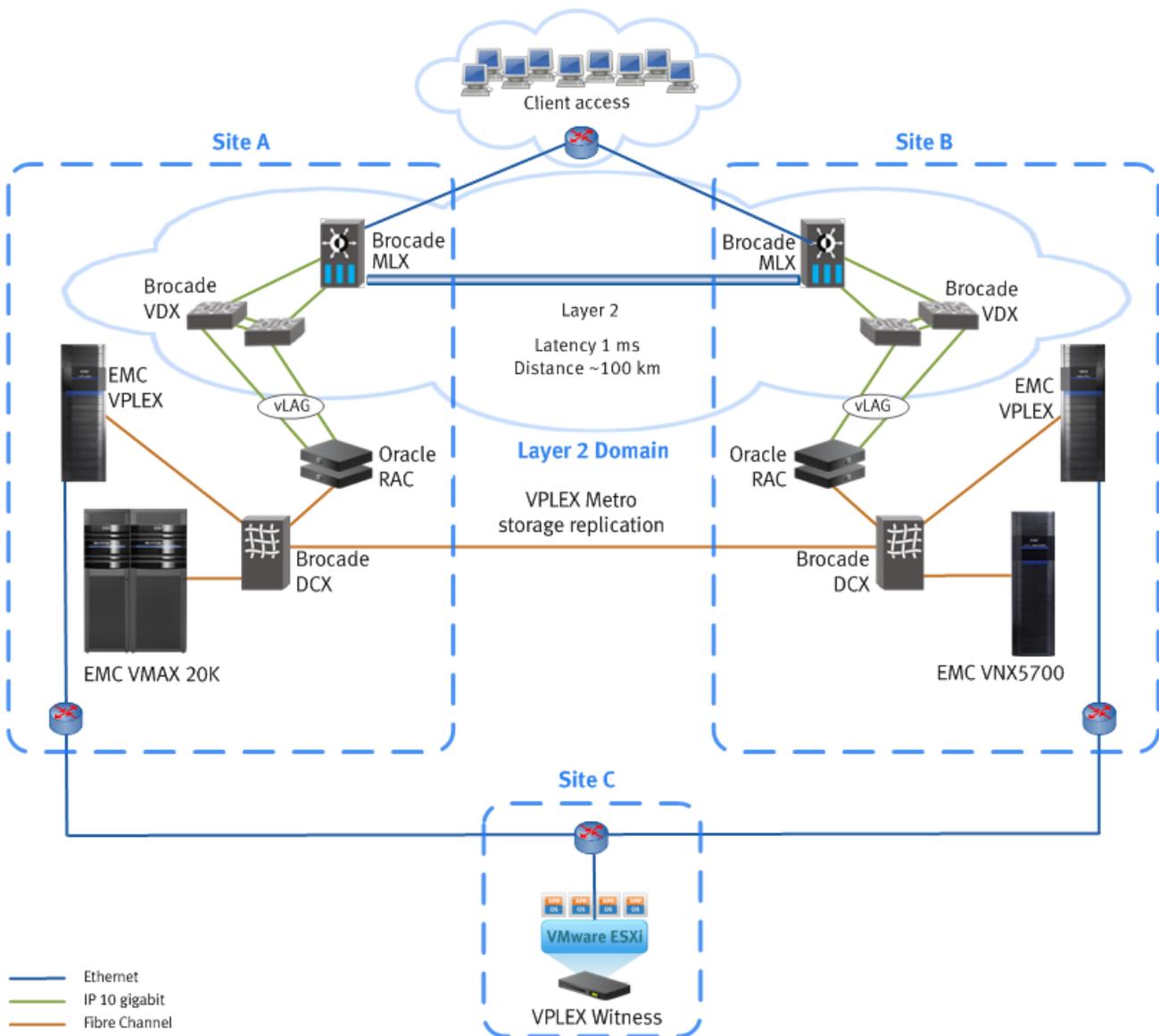


Figure 3. Solution architecture (VPLEX with Oracle Extended RAC)

Disaster recovery solution

This solution adds EMC RecoverPoint into the VPLEX environment to provide remote DR protection for the Oracle RAC database.

Many replication options and site topologies are possible when deploying VPLEX with RecoverPoint. This solution has a three-site topology, created by adding an EMC VMAX 10K storage array to the site where VPLEX Witness is deployed (that is, Site C), as shown in Figure 4. This co-location of VPLEX Witness and the RecoverPoint DR site is not a requirement; it is used in this solution for convenience only.

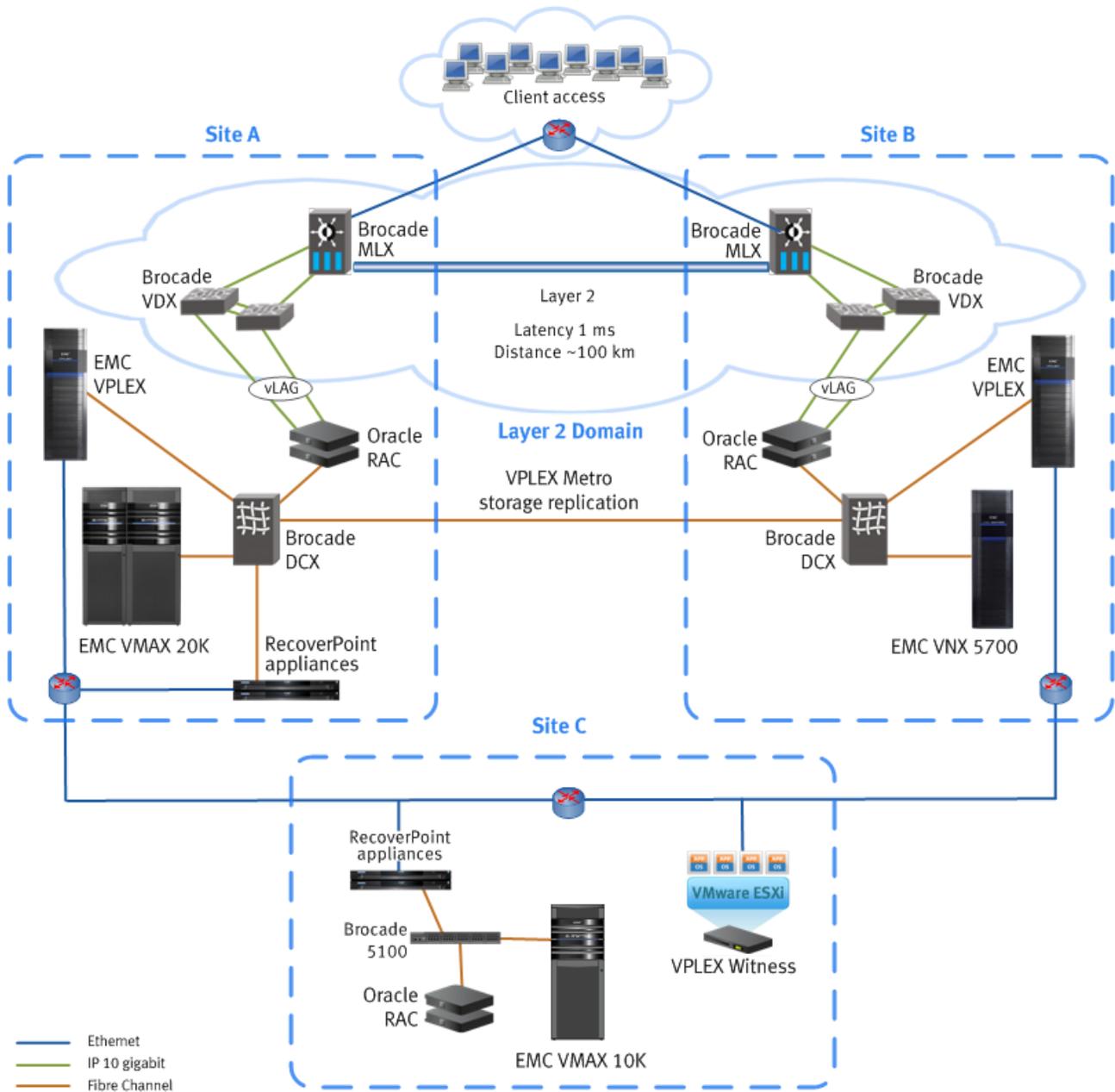


Figure 4. Solution architecture (VPLEX Metro with RecoverPoint)

This three-site topology provides the following benefits:

- With VPLEX enabling active/active data centers with zero RTOs and RPOs, customers can avoid the downtime associated with site loss and storage array failures.
- With RecoverPoint replicating production data to a third site, customers can rapidly recover from application data corruption, viruses, or human error, and from wide-scale physical disasters that affect both VPLEX Metro clusters—for example, flooding, terrorism, earthquakes, and power outages.
- With RecoverPoint journaling technology, customers can recover and restart the Oracle RAC environment to specific points in time, in case the latest image is corrupt.

With RecoverPoint continuous remote replication (CRR), data is replicated from a local production site to a remote DR site. In this solution, the VPLEX at Site A is the local site, while the VMAX at Site C is the remote site—the maximum distance between local and remote sites is 20,000 km.

In the event of Site A failure, users can choose to continue operations at the second VPLEX data center (Site B) while the first site is repaired. RecoverPoint will not be able to track write activity until Site A is restored. When connectivity is restored between Site A and Site B, VPLEX incrementally and nondisruptively rebuilds the copy of the data located at Site A. RecoverPoint then automatically re-synchronizes by replicating all new data at Site A to the remote copy at Site C.

Database and workload profile

Table 2 details the database and workload profile for the solution.

Table 2. Database and workload profile

Profile characteristic	Details
Number of databases	1
Database type	OLTP
Database size	500 GB
Database name	ORARP
Oracle RAC	4 physical nodes (production); 2 physical nodes (DR)
Benchmark profile	Swingbench Order Entry workload

Hardware resources

Table 3 details the hardware resources for the solution.

Table 3. Solution hardware environment

Purpose	Quantity	Configuration
Storage (Site A)	1	EMC Symmetrix VMAX 20K, with: <ul style="list-style-type: none"> • 2 engines • 171 x 450 GB FC drives • 52 x 2 TB SATA drives

Purpose	Quantity	Configuration
Storage (Site B)	1	EMC VNX5700, with: <ul style="list-style-type: none"> 45 x 600 GB SAS drives 54 x 300 GB SAS drives 30 x 2 TB NL-SAS drives
Storage (Site C)	1	EMC Symmetrix VMAX 10K, with: <ul style="list-style-type: none"> 1 engine 32 x 1 TB NL-SAS drives 60 x 450 GB FC drives
Distributed storage federation	2	VPLEX Metro cluster, with: <ul style="list-style-type: none"> 2 x VS2 engines
RecoverPoint appliances	4	GEN4, 2 per site
Oracle RAC database servers	6	4 x eight-core CPUs, 128 GB RAM
VMware ESXi server for VPLEX Witness	2	2 x two-core CPUs, 48 GB RAM
Network switching and routing platform (Sites A and B)	2	Brocade DCX 8510 Backbone, with: <ul style="list-style-type: none"> Fx8-24 FC extension card 2 x 48-port FC Blades with 16 Gb FC line speed support
		Brocade MLXe Router
	4	Brocade VDX 6720 in VCS mode
Network switch (Site C)	1	Brocade 5100

Software resources Table 4 details the software resources for the solution.

Table 4. Solution software environment

Software	Version	Purpose
EMC Enginuity™	5876.82.57	Symmetrix VMAX operating environment
EMC VPLEX GeoSynchrony®	5.1	VPLEX operating environment
EMC VPLEX Witness	5.1	Monitor and arbitrator component for handling VPLEX cluster failure and inter-cluster communication loss
EMC RecoverPoint/EX	3.5 P1 (n.149)	Replication
EMC VNX OE for block	05.31.000.5.715	VNX operating environment
EMC VNX OE for file	7.0.52.1	VNX operating environment
EMC Unisphere® for VPLEX	1.0	VPLEX management software
Oracle Database 11g (with Oracle RAC and Oracle Grid Infrastructure)	Enterprise Edition 11.2.0.3	Oracle database and cluster software

EMC VPLEX Metro infrastructure

Introduction

Overview

This section describes the VPLEX Metro infrastructure for the solution, which comprises the following components:

- EMC VPLEX Metro cluster at each data center (Site A and Site B)
- EMC VPLEX Witness in a separate failure domain (Site C)

EMC VPLEX

EMC VPLEX is a storage virtualization solution for both EMC and non-EMC storage arrays. EMC offers VPLEX in three configurations to address customer needs for high availability and data mobility, as shown in Figure 5:

- VPLEX Local
- VPLEX Metro
- VPLEX Geo

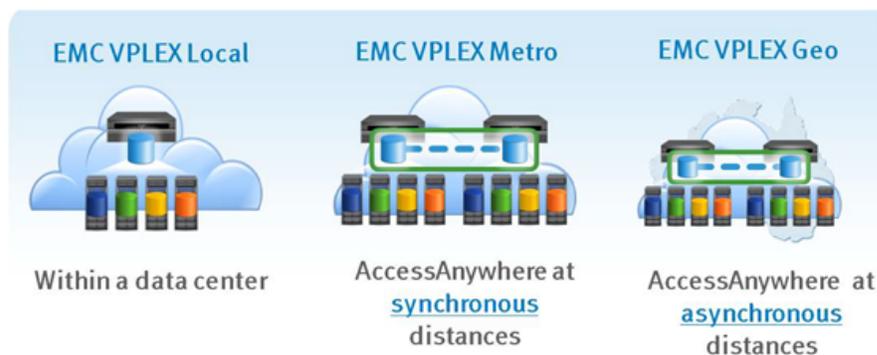


Figure 5. VPLEX topologies

For detailed descriptions of these VPLEX configurations, refer to the documents listed in [References](#) on [page 42](#).

EMC VPLEX Metro

This solution uses VPLEX Metro, which uses a unique clustering architecture to help customers break the boundaries of the data center and allow servers at multiple data centers to have read/write access to shared block storage devices. VPLEX Metro delivers active/active, block-level access to data on two sites within synchronous distances with a round-trip time of up to 5 ms.

EMC VPLEX Witness

VPLEX Witness is an optional external server that acts as a cluster arbiter. It is installed as a virtual machine in a separate failure domain to the VPLEX clusters. VPLEX Witness connects to both VPLEX clusters using a Virtual Private Network (VPN) over the management IP network; it requires a round trip time that does not exceed 1 second.

By reconciling its own observations with information reported periodically by the clusters, VPLEX Witness enables the cluster(s) to distinguish between inter-cluster network partition failures and cluster failures and to automatically resume I/O at the appropriate site.

VPLEX Witness failure handling semantics apply only to distributed volumes within a consistency group and only when the detach rules identify a static preferred cluster for the consistency group (see [VPLEX consistency groups](#) on [page 16](#) for further details).

EMC VPLEX Management Interface

You can manage and administer a VPLEX environment with Unisphere for VPLEX or you can connect directly to a management server and start a VPLEX CLI session.

EMC VPLEX High Availability

VPLEX Metro enables application and data mobility and, when configured with VPLEX Witness, provides a high-availability infrastructure for clustered applications such as Oracle RAC. VPLEX Metro enables you to build an extended or stretch cluster as if it was a local cluster, and removes the data center as a single point of failure. Furthermore, as the data and applications are active at both sites, the solution provides a simple business continuity strategy.

VPLEX logical storage structures

VPLEX encapsulates traditional physical storage array devices and applies layers of logical abstraction to these exported LUNs, as shown in Figure 6.

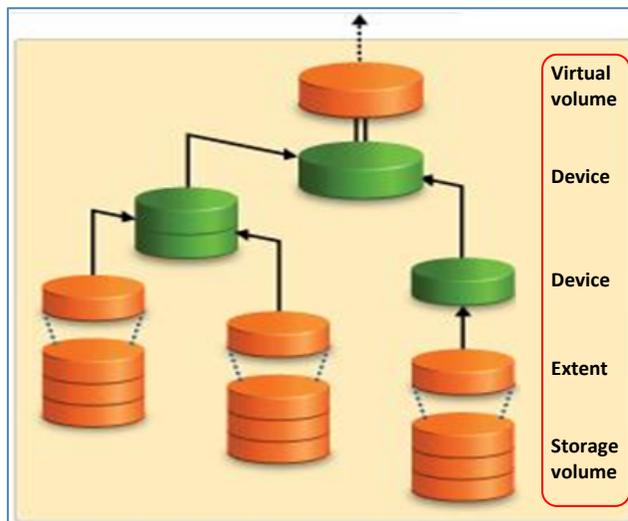


Figure 6. VPLEX logical storage structures

- A *storage volume* is a LUN exported from an array and encapsulated by VPLEX.
- An *extent* is the mechanism VPLEX uses to divide storage volumes and may use all or part of the capacity of the underlying storage volume.
- A *device* encapsulates an extent or combines multiple extents or other devices into one large device with a specific RAID type.

- A *distributed device* is a device that encapsulates other devices from two separate VPLEX clusters.
- At the top layer of the VPLEX storage structures are *virtual volumes*. These are created from a top-level device (a device or distributed device) and always use the full capacity of the top-level device.

Virtual volumes are the elements that VPLEX exposes to hosts using its front-end ports. VPLEX presents virtual volumes to a host, and to RecoverPoint appliances, through a *storage view*.

VPLEX can encapsulate devices across heterogeneous storage arrays, including virtually provisioned thin devices and traditional LUNs.

VPLEX consistency groups

Consistency groups aggregate virtual volumes together so that the same detach rules and other properties can be applied to all volumes in the group. There are two types of VPLEX consistency groups:

- **Synchronous consistency groups**—These are used in VPLEX Local and VPLEX Metro to apply the same detach rules and other properties to a group of volumes in a configuration. This simplifies configuration and administration on large systems.

Synchronous consistency groups use write-through caching (known as synchronous cache mode) and with VPLEX Metro are supported on clusters separated by up to 5 ms of latency. VPLEX Metro sends writes to the back-end storage volumes, and acknowledges a write to the application only when the back-end storage volumes in *both* clusters acknowledge the write.

- **Asynchronous consistency groups**—These are used for distributed volumes in VPLEX Geo, where clusters can be separated by up to 50 ms of latency.

Detach rules

Detach rules determine I/O processing semantics for a consistency group when connectivity with a remote cluster is lost—for example, in the case of a network partitioning or remote cluster failure. Synchronous consistency groups support the following detach rules to determine cluster behavior during a failure:

- Static preference rule identifies a preferred cluster
- No-automatic-winner rule suspends I/O on both clusters

When connectivity is lost between clusters, the configured detach rule is automatically invoked. However, VPLEX Witness can be deployed to override the static preference rule and ensure that the non-preferred cluster remains active if the preferred cluster fails.

VPLEX Metro solution configuration

Storage structures

Figure 7 shows the physical and logical storage structure used by VPLEX Metro in the context of this solution.

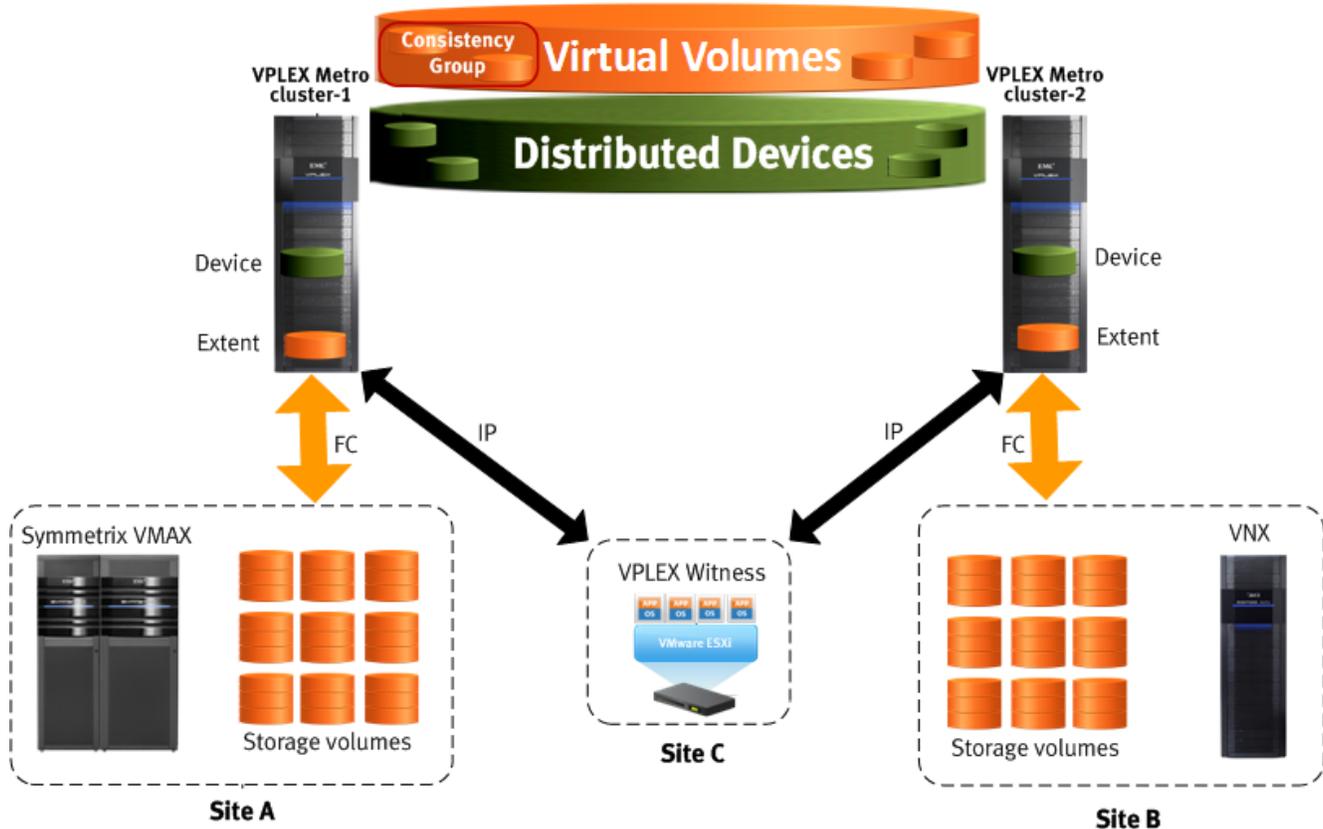


Figure 7. VPLEX physical and logical storage structures for Oracle RAC

There is a one-to-one mapping between storage volumes, extents, and devices at each site. The devices encapsulated at Site A (cluster-1) are virtually provisioned thin devices, while the devices encapsulated at Site B (cluster-2) are traditional LUNs.

To create distributed devices, all cluster-1 devices are mirrored remotely on cluster-2 in a distributed RAID 1 configuration. These distributed devices are encapsulated by virtual volumes, which are then presented to the hosts through storage views.

Consistency groups

Consistency groups are particularly important for databases and their applications. For example:

- **Write-order fidelity**—To maintain data integrity, all Oracle database LUNs (for example, data, control, and log files) should be placed together in a single consistency group.

- **Transactional dependency**—Often multiple databases have transaction dependencies, such as when an application issues transactions to multiple databases and expects the databases to be consistent with each other. All LUNs that require I/O dependency to be preserved should reside in a single consistency group.
- **Application dependency**—Oracle RAC maintains Oracle Cluster Registry (OCR) and voting files within a set of disks that must be accessible to maintain database availability. The database and OCR disks should reside in a single consistency group.

For the solution, a single synchronous consistency group—Extended_Oracle_RAC_CG—contains all the virtual volumes that hold the Oracle ASM disk groups and the OCR and voting files. The detach rule for the consistency group has cluster-1 as the preferred cluster.

A second consistency group—RecoverPoint_system—contains the repository and journal volumes, as shown in Figure 8.

Name	Operational Status	Detach Rule	# Volumes	Visibility	Storage at Clusters
Extended_Oracle_RAC_CG	OK	Winner:cluster-1 (..	37	Both Clusters	Both Clusters
RecoverPoint_system	OK	N/A	2	cluster-1	cluster-1

Selected Consistency Group:
RecoverPoint_system

Virtual Volumes

- SiteA_VP_RPA_JRNL_040B_1_vol
- SiteA_VP_RPA_REPO_040C_1_vol

Figure 8. VPLEX consistency groups for solution

EMC RecoverPoint configuration

Introduction

Overview

RecoverPoint is an advanced, enterprise-class data protection solution designed with the performance, reliability, and flexibility required for enterprise applications in heterogeneous storage and server environments. It provides bi-directional local and remote data replication, point-in-time data recovery, and instant, nondisruptive access to replicated data for DR, testing, reuse, and other purposes.

RecoverPoint data protection options

RecoverPoint provides the following replication options for both physical and virtualized environments:

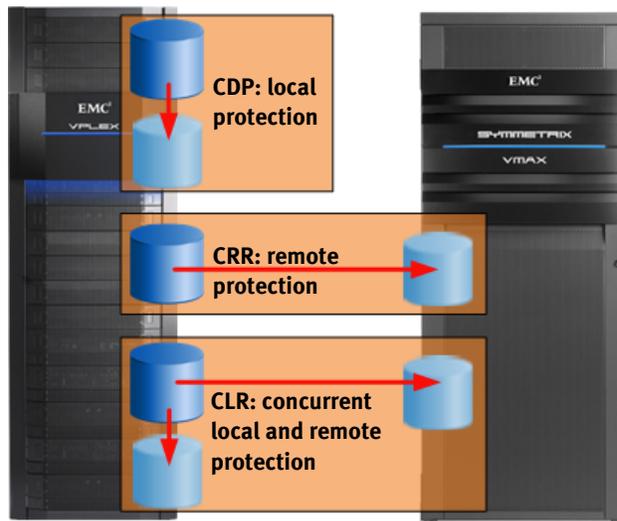


Figure 9. RecoverPoint replication options

- **Continuous data protection (CDP):** CDP continuously captures and stores data modifications and enables point-in-time recovery, with no data loss. It supports both synchronous and asynchronous replication.
- **Continuous remote replication (CRR):** CRR supports synchronous and asynchronous replication between remote sites over FC or a WAN. Asynchronous replication provides crash-consistent protection and recovery to specific points in time, with a small RPO. Synchronous replication is supported when the remote sites are connected through FC and provides a zero RPO.
- **Concurrent local and remote replication (CLR):** CLR is a combination of CRR and CDP and provides concurrent local and remote data protection.

The CDP copy is normally used for operational recovery, while the CRR copy is normally used for disaster recovery.

RecoverPoint appliance

RecoverPoint is appliance-based, which enables it to better support large amounts of information stored across heterogeneous environments. This approach enables RecoverPoint to deliver continuous replication with minimal impact to an application's I/O operations.

RecoverPoint appliances (RPAs) run the RecoverPoint software and manage all aspects of data replication. A cluster of two or more active RPAs is deployed at each RecoverPoint site. If one RPA in a cluster fails, RecoverPoint immediately switches over the functions of that appliance to one or more of the remaining RPAs in the cluster.

For remote replication, RPAs use powerful deduplication, compression, and bandwidth reduction technologies to minimize the use of bandwidth and dramatically reduce the time lag between writing data to storage at the source and target sites.

RecoverPoint splitter

RecoverPoint uses write splitting technology to monitor and split application writes to protected volumes. The write splitter ensures that copies of each write are sent to the original target volume and to the local RPA simultaneously. The RPA then forwards the write to the target replica volume.

RecoverPoint supports host-based, intelligent fabric-based, and array-based splitters. EMC VMAX, VNX, and CLARiiON arrays have integrated RecoverPoint splitters; these enable local and remote replication without the cost or complexity of additional hardware.

Starting with GeoSynchrony 5.1, EMC VPLEX also includes a RecoverPoint splitter. This splitter is designed to interoperate with RecoverPoint 3.5 and higher and enables VPLEX to take advantage of RecoverPoint data protection capabilities in VPLEX Local and VPLEX Metro configurations. Because VPLEX can federate heterogeneous EMC and non-EMC storage arrays, and because VPLEX carries out the write splitting function for these arrays, you can use RecoverPoint to protect data on storage arrays that do not possess native write-splitting capabilities, without additional SAN or host infrastructure requirements.

The VPLEX splitter supports LUNs up to 32 TB in size. In addition, multiple RecoverPoint clusters can share a VPLEX write splitter. This enables up to four RecoverPoint clusters to replicate volumes on a single VPLEX cluster.

RecoverPoint consistency groups

The consistency and write-order fidelity of RecoverPoint replicas are assured by RecoverPoint's use of replication sets and consistency groups. A replication set consists of a production volume and the replica volume(s) to which it is replicating. A RecoverPoint consistency group logically groups replication sets that must be consistent with one another. The consistency group ensures that updates to the production volumes are written to the replicas in consistent write order and that the replicas can always be used to continue working or to restore the production source.

RecoverPoint is unique in its ability to guarantee a consistent copy at the target site under all circumstances, and in its ability to maintain distributed write-order fidelity in multi-host heterogeneous SAN environments.

RecoverPoint replication is policy-driven. A replication policy, based on a particular business need, can be uniquely specified for each consistency group. This policy governs the replication parameters for the consistency group—for example, the RPO and RTO for the consistency group, and its deduplication, data compression, and bandwidth reduction settings.

RecoverPoint journals

Each consistency group has either two or three associated journals. A production journal is always provisioned; this stores system marking information that is used when synchronizing replication volumes at two sites. Depending on the replication type (local, remote, or both), a replica journal is provisioned at the local production site and/or at the remote site. The replica journals store timestamped snapshots of application writes for later recovery to selected points in time. Snapshots stored at a replica journal represent the data that has changed on the production storage since the closing of the previous snapshot.

In synchronous replication, every write is retained in the replica journal—this supports recovery to any point in time. In asynchronous replication, several writes are grouped in a single snapshot—this supports recovery to significant points in time (you can set the snapshot granularity for asynchronous replication to seconds or MBs).

RecoverPoint replica journals also support bookmarks (that is, named snapshots). These are useful for marking particular points in time, such as an event in an application or a point in time to which you wish to fail over. Bookmarks can be created and named manually; or they can be created automatically at regular intervals or in response to a system event.

Repository volume

Each RPA cluster has a dedicated repository volume that stores replication environment information. This volume must be seen only by the RecoverPoint appliances on the same site as itself. In addition, the volume cannot be located on a VPLEX distributed device.

RecoverPoint and VPLEX

RecoverPoint replication for VPLEX

The VPLEX GeoSynchrony 5.1 and RecoverPoint 3.5 software releases bring RecoverPoint operational and DR capabilities to the VPLEX Local and VPLEX Metro products. With the VPLEX write splitter, customers can take full advantage of all RecoverPoint features for VPLEX virtual volumes that reside on the following device types²:

- Local devices
- Metro distributed RAID-1 (synchronous cache mode)

² For the support requirements for RecoverPoint with VPLEX write splitting technology, see the GeoSynchrony 5.1 and RecoverPoint 3.5 release notes.

In the solution, all the production volumes are VPLEX virtual volumes. Because of the one-to-one mapping between the VPLEX storage volumes, extents, and devices, and the RAID 1 VPLEX device geometry (see [VPLEX Metro solution configuration](#) on [page 17](#)), RecoverPoint can replicate the production VPLEX virtual volumes as it would any other source volume.

Figure 10 shows the logical relationships between the host, virtual volumes, and corresponding VPLEX and RecoverPoint components.

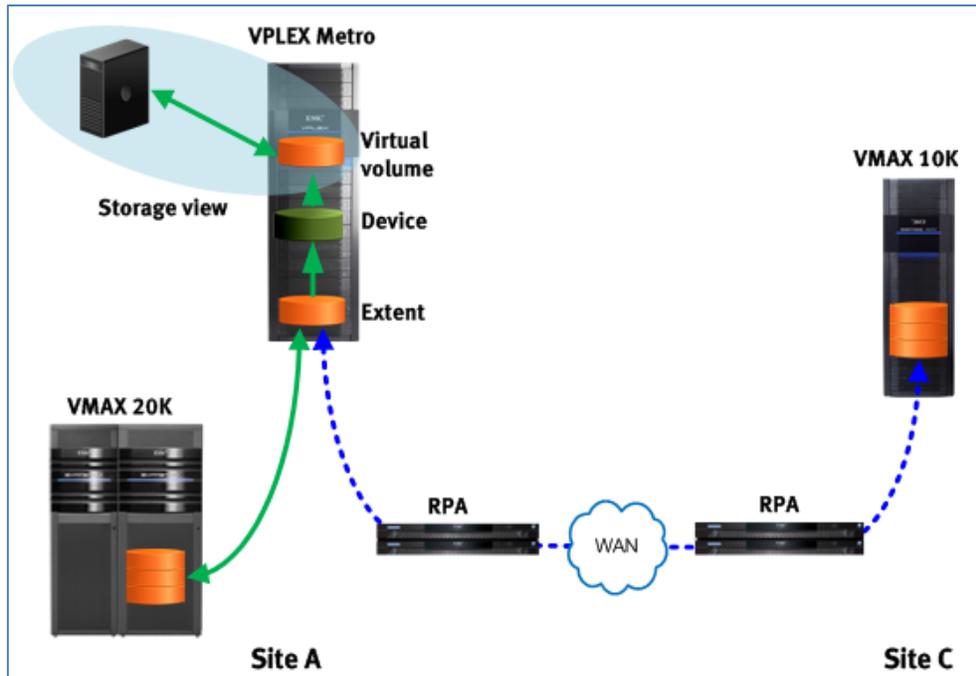


Figure 10. Logical relationship between VPLEX and RecoverPoint components

RecoverPoint and VPLEX consistency groups

Both VPLEX and RecoverPoint consistency groups logically group volumes that must be consistent with each other. For solutions that use the VPLEX write-splitter and RecoverPoint, VPLEX consistency groups are used together with RecoverPoint consistency groups to ensure that volume protection and failure behavior is aligned across both products.

EMC recommends that all volumes in a VPLEX consistency group be configured in a single RecoverPoint consistency group, and that all volumes in a RecoverPoint consistency group should be configured in a single VPLEX consistency group.

RecoverPoint CRR for VPLEX

Figure 11 illustrates the data flow for asynchronous CRR with a VPLEX write splitter. All the production volumes are VPLEX virtual volumes.

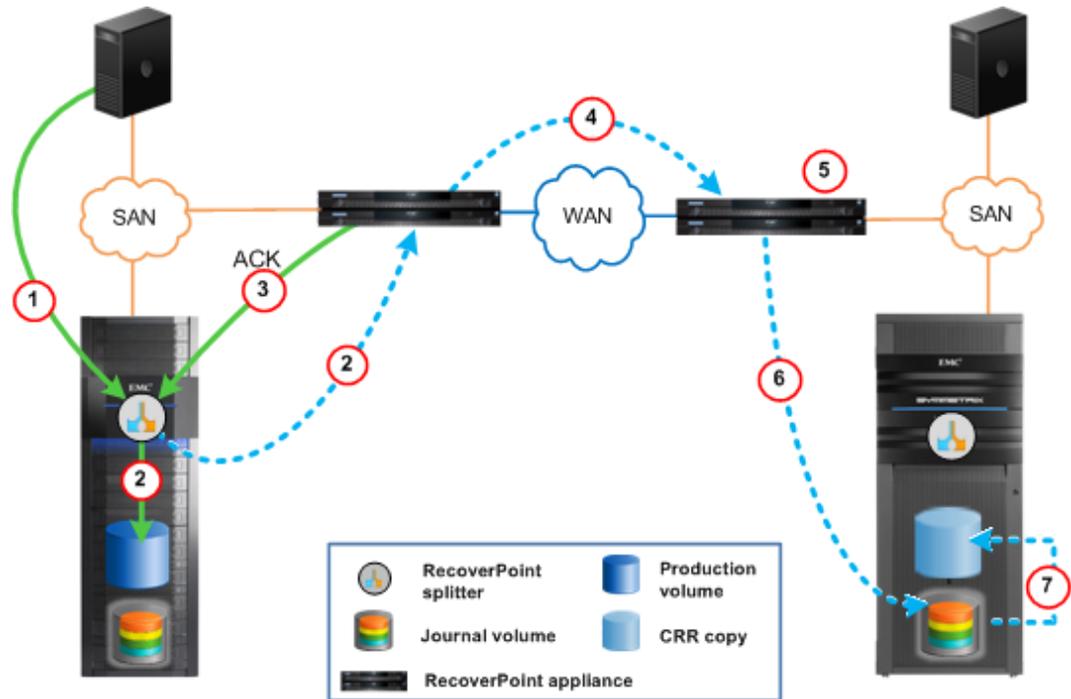


Figure 11. RecoverPoint CRR data flow (asynchronous)

1. The host issues a write to a LUN that is being protected by RecoverPoint. The write is intercepted by the VPLEX splitter.
2. The splitter “splits” the write and sends it to the production volume and to the local RPA.
3. When the RPA receives the write, it immediately acknowledges it back to the splitter. (With synchronous replication, the ACK is delayed until the write has been received by the RPA at the remote site.)
4. When the write is received by the local RPA, it is bundled with other writes, deduplicated to remove redundant blocks, sequenced, and timestamped. The package is then compressed and transmitted with a checksum for delivery to the remote RPA cluster.
5. When the package is received at the remote site, the remote RPA verifies the checksum, to ensure the package was not corrupted in transmission, and uncompresses the data.
6. The RPA then writes the data to the replica journal at the remote site.
7. After the data has been written to the journal volume, it is distributed to the remote replica volumes—write order is preserved during this distribution.

Overview

When integrating VPLEX and RecoverPoint, there are three possible installation paths:

- Install RecoverPoint and VPLEX in a green-field deployment
- Install VPLEX into an existing RecoverPoint environment
- Install RecoverPoint into an existing VPLEX environment

For this solution, VPLEX Metro was already installed, configured, and servicing a four-node Oracle RAC environment that extended across two data centers (Site A and Site B)³. RecoverPoint was added into this existing environment and was configured for CRR to provide disaster recovery on Site C with a VMAX 10K storage array.

Design considerations

With VPLEX GeoSynchrony 5.1 and RecoverPoint 3.5, VPLEX Metro virtual volumes can be replicated locally (CDP), remotely (CRR), or both (CLR). However, there are some considerations:

- RecoverPoint can replicate the VPLEX cluster at only one site of a VPLEX Metro installation. No communication is allowed between RecoverPoint and the second site of a Metro installation.
- The RecoverPoint system repository volume (5.72 GB in size) and journal volumes (minimum size 5 GB; 20 GB for this solution) must not be located on a distributed device on the VPLEX site protected by RecoverPoint.
- Each VPLEX consistency group that is protected by RecoverPoint must have its site preference (winning site) set to the site where the write splitting is taking place.

Install and configure utilities and wizards

For assistance with managing and maintaining VPLEX environments, EMC offers the VPLEX Procedure Generator tool. This provides up-to-date guidance and workflows for VPLEX installations, and includes information on installing and managing RecoverPoint in conjunction with VPLEX. Figure 12 shows the RecoverPoint options provided by the utility.

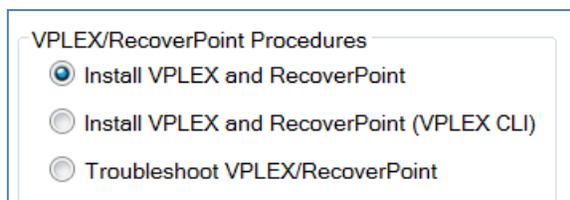


Figure 12. EMC VPLEX Procedure Generator options for RecoverPoint

The EMC Unisphere for VPLEX storage management platform and RecoverPoint Deployment Manager further simplify RecoverPoint installation and management.

RecoverPoint Deployment Manager provides a simple wizard to drive the RecoverPoint installation, as shown in Figure 13.

³ See the white paper: *EMC Mission-Critical Business Continuity for SAP*.

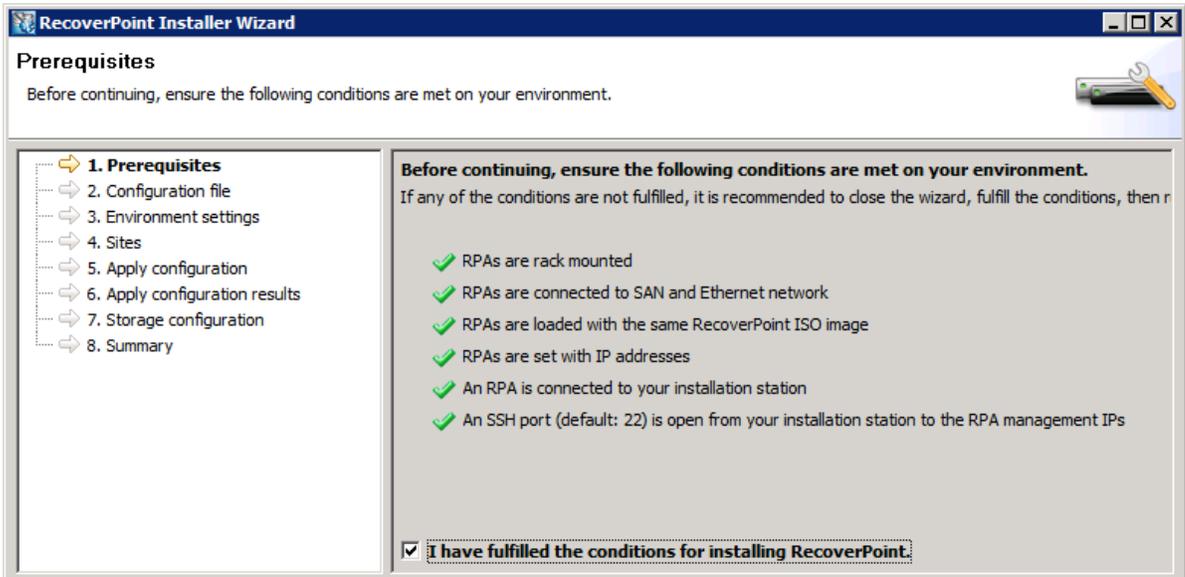


Figure 13. RecoverPoint Installer Wizard

Unisphere for VPLEX enables you to monitor the health of your VPLEX and RecoverPoint systems. It also provides easy access to the RecoverPoint Management Application, as shown in Figure 14.

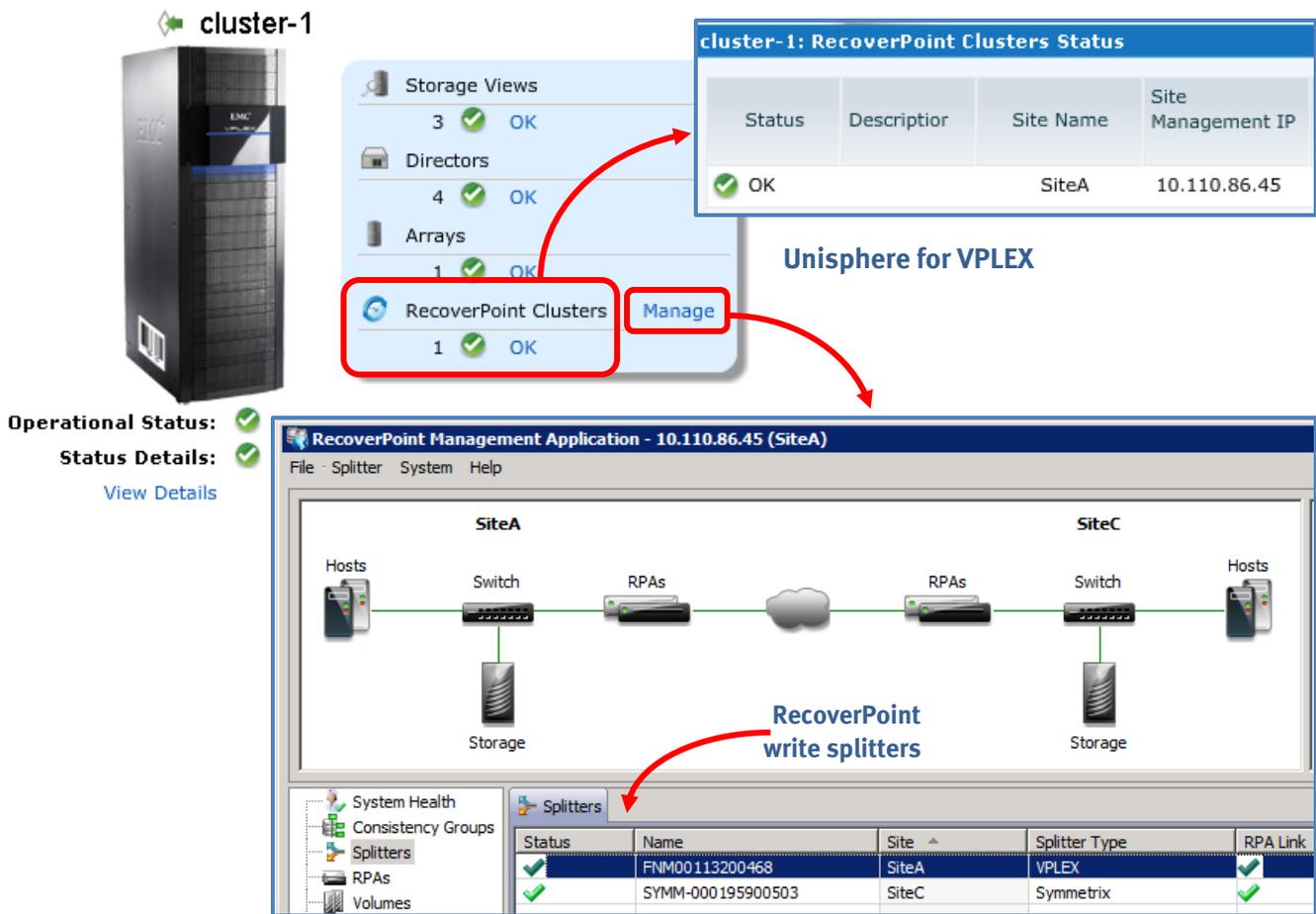


Figure 14. Administering RecoverPoint with Unisphere for VPLEX

RecoverPoint consistency group for solution

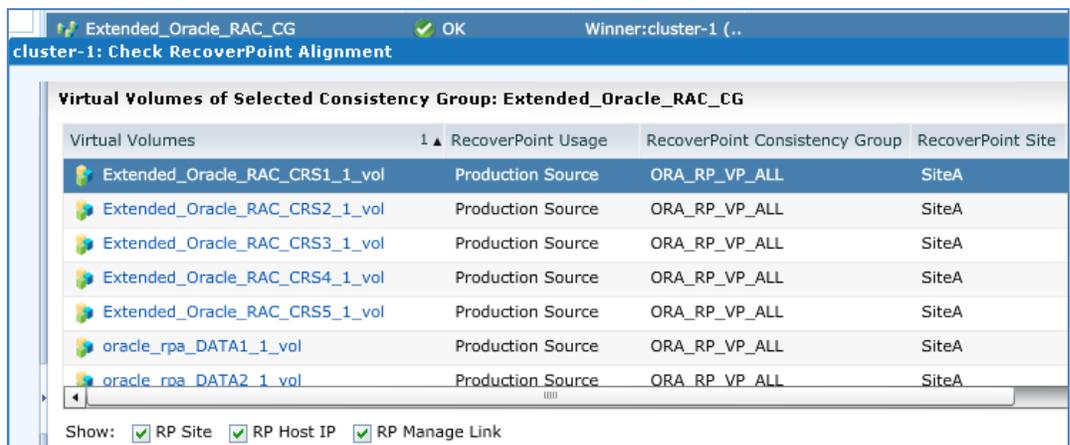
RecoverPoint can replicate only those VPLEX virtual volumes that are members of a VPLEX consistency group with the **recoverpoint-enabled** attribute set to **true**. The number of such consistency groups will vary depending on the configuration and the number of applications being replicated.

For the solution, a single RecoverPoint consistency group (ORA_RP_VP_ALL) was configured to replicate the production volumes on Site A to the replica volumes on Site C. As recommended by EMC, this RecoverPoint consistency group contains the same member volumes as the two VPLEX consistency groups: Extended_Oracle_RAC_CG (production) and RecoverPoint_system (system).

All virtual volumes in a VPLEX consistency group must:

- Have the same RecoverPoint role (production or replica)
- Belong to the same RecoverPoint consistency group

Checking this alignment of VPLEX and RecoverPoint consistency groups is the final step in deploying an application on VPLEX for RecoverPoint replication. You can do this using Unisphere for VPLEX, as shown in Figure 15.



The screenshot shows a window titled 'cluster-1: Check RecoverPoint Alignment' with a status bar indicating 'Extended_Oracle_RAC_CG' and 'OK'. The main content area is titled 'Virtual Volumes of Selected Consistency Group: Extended_Oracle_RAC_CG'. It contains a table with the following data:

Virtual Volumes	RecoverPoint Usage	RecoverPoint Consistency Group	RecoverPoint Site
Extended_Oracle_RAC_CRS1_1_vol	Production Source	ORA_RP_VP_ALL	SiteA
Extended_Oracle_RAC_CRS2_1_vol	Production Source	ORA_RP_VP_ALL	SiteA
Extended_Oracle_RAC_CRS3_1_vol	Production Source	ORA_RP_VP_ALL	SiteA
Extended_Oracle_RAC_CRS4_1_vol	Production Source	ORA_RP_VP_ALL	SiteA
Extended_Oracle_RAC_CRS5_1_vol	Production Source	ORA_RP_VP_ALL	SiteA
oracle_rpa_DATA1_1_vol	Production Source	ORA_RP_VP_ALL	SiteA
oracle_rpa_DATA2_1_vol	Production Source	ORA_RP_VP_ALL	SiteA

At the bottom of the dialog, there are checkboxes for 'Show: RP Site', 'RP Host IP', and 'RP Manage Link', all of which are checked.

Figure 15. Validating consistency group alignment

RecoverPoint CRR requires one replica volume on the remote site for each production volume. The replica volume must be the exact same size as its associated production volume. For the solution, LUNs from the VMAX 10K on Site C were provisioned and zoned to the Site C RecoverPoint appliance.

The RecoverPoint Management Application view in Figure 16 shows the ORA_RP_VP_ALL consistency group being replicated from the Production Source to the Remote Replica, with all writes going to the Site C replica journal and from there to Site C storage.

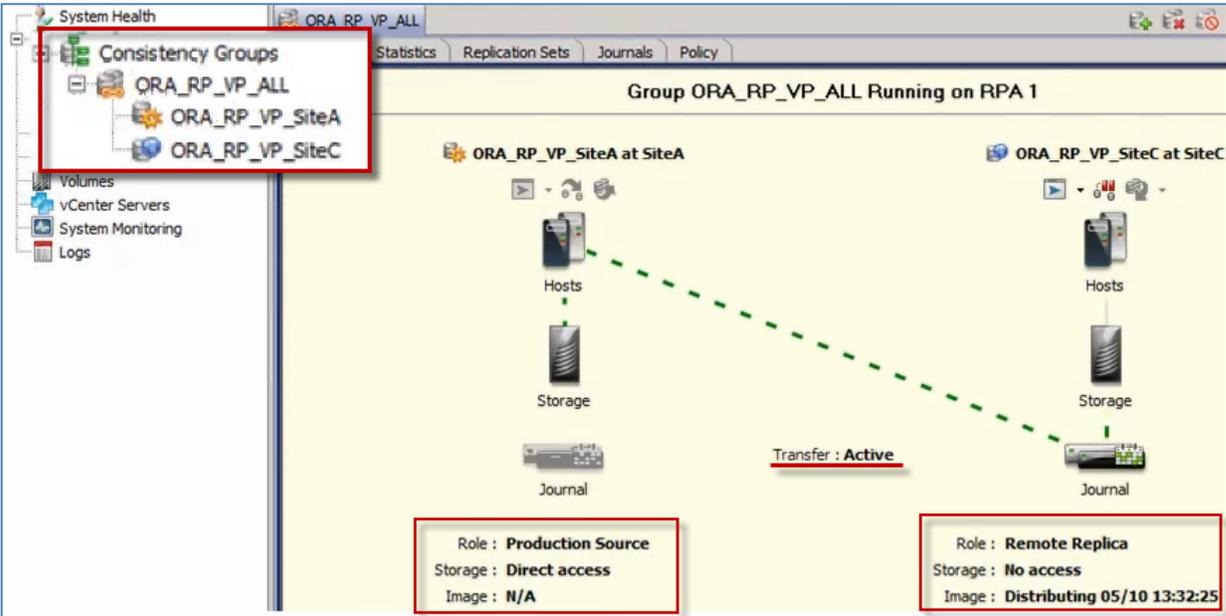


Figure 16. RecoverPoint consistency group

Oracle database architecture

Introduction

Overview

The solution uses these Oracle components and options:

- Oracle Database 11g Release 2 Enterprise Edition
- Oracle Automatic Storage Management (ASM)
- Oracle Clusterware

Oracle Database 11gR2

Oracle Database 11g Release 2 Enterprise Edition delivers industry-leading performance, scalability, security, and reliability on a choice of clustered or single servers running Windows, Linux, or UNIX. It provides comprehensive features for transaction processing, business intelligence, and content management applications.

Oracle ASM

Oracle ASM is an integrated, cluster-aware database file system and disk manager. ASM file system and volume management capabilities are integrated with the Oracle database kernel. In Oracle Database 11gR2, Oracle ASM has also been extended to include support for OCR and voting files to be placed within ASM disk groups.

Oracle Clusterware

Oracle Clusterware is a portable cluster management solution that is integrated with the Oracle database. It provides the infrastructure necessary to run Oracle RAC, including Cluster Management Services and High Availability Services. A non-Oracle application can also be made highly available across the cluster using Oracle Clusterware.

Oracle Grid Infrastructure

In Oracle Database 11gR2, the Oracle Grid Infrastructure combines Oracle ASM and Oracle Clusterware into a single set of binaries, separate from the database software. This infrastructure now provides all the cluster and storage services required to run an Oracle RAC database.

Oracle Real Application Clusters 11g

Oracle RAC is primarily a high-availability solution for Oracle database applications within the data center. It enables multiple Oracle instances to access a single database. The cluster consists of a group of independent servers co-operating as a single system and sharing the same set of storage disks. Each instance runs on a separate server in the cluster. RAC can provide high availability, scalability, fault tolerance, load balancing, and performance benefits, and removes any single point of failure from the database solution.

Oracle RAC on Extended Distance Clusters

Oracle RAC on Extended Distance Clusters (Oracle Extended RAC) is an architecture that allows servers in the cluster to reside in physically separate locations. This removes the data center as a single point of failure.

Oracle Extended RAC enables all nodes within the cluster, regardless of location, to be active. It provides high availability and business continuity during a site or network failure, as follows:

- Storage and data remain available and active on the surviving site.
- Oracle Services load balance and fail over to the Oracle RAC nodes on the surviving site.
- Oracle Transparent Application Failover (TAF) allows sessions to automatically fail over to Oracle RAC nodes on the surviving site.
- Oracle RAC nodes on the surviving site continue to process transactions.

Oracle recommends that the Oracle Extended RAC architecture fits best where the two data centers are relatively close (no more than 100 km apart)⁴. EMC recommends no more than 1 ms latency.

Oracle RAC and VPLEX

Oracle RAC is normally run in a local data center due to the potential impact of distance-induced latency and the relative complexity and overhead of extending Oracle RAC across data centers with host-based mirroring using Oracle ASM. With EMC VPLEX Metro, however, an Oracle Extended RAC deployment, from the Oracle DBA perspective, becomes a standard Oracle RAC install and configuration⁵.

Oracle Extended RAC on VPLEX Metro

A four-node Oracle Extended RAC on VPLEX Metro environment was deployed for this solution, with two RAC nodes deployed on each production site. The Oracle 11g database was deployed on Oracle ASM disk groups configured with external redundancy to make use of the protection offered by VPLEX distributed mirroring. Figure 17 provides a logical representation of this deployment.

⁴ See the Oracle white paper: *Oracle Real Application Clusters (RAC) on Extended Distance Clusters*.

⁵ See the EMC white paper: *Oracle Extended RAC with EMC VPLEX Metro Best Practices Planning*.

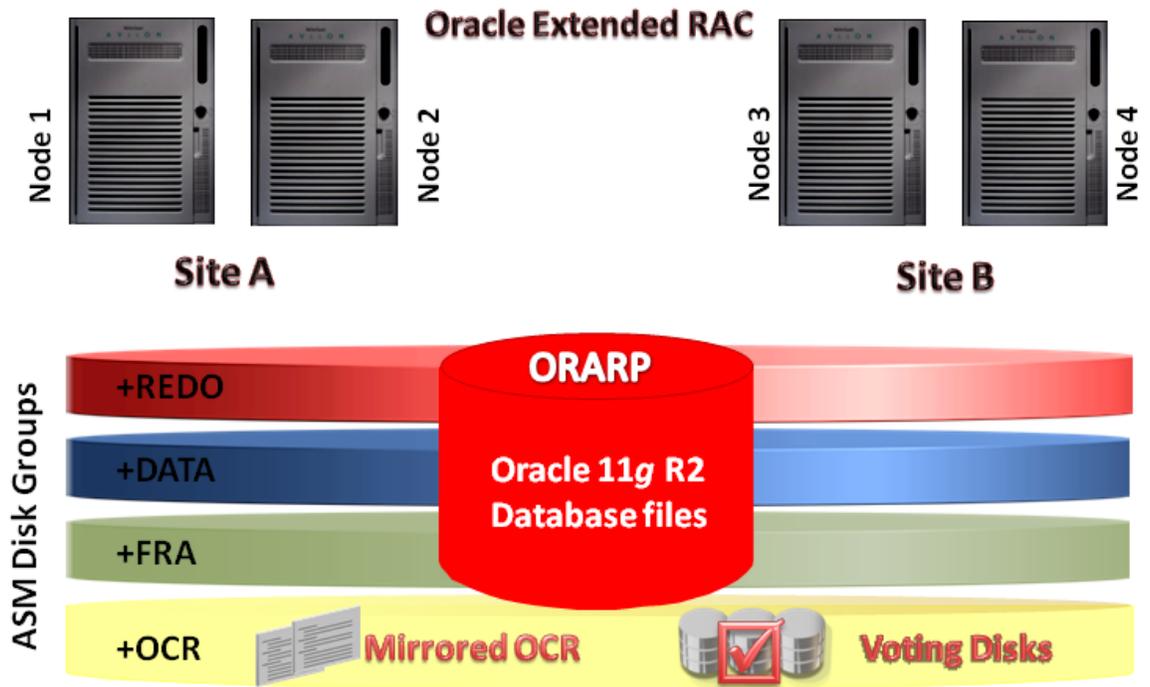


Figure 17. Oracle Extended RAC over EMC VPLEX Metro

On the DR site (Site C), a two-node Oracle RAC cluster was configured with Oracle ASM and a dedicated OCR disk group. Oracle Grid Infrastructure and Oracle Database 11gR2 binaries were installed on each node.

Brocade network infrastructure

Introduction

Overview

This section describes the IP and SAN networks deployed for the solution and the Layer 2 extension between the two data centers at Sites A and B. The network infrastructure is built using these Brocade components:

IP network

- Brocade VDX 6720 Data Center Switches
- Brocade MLX Series routers
- Brocade 1020 CNAs

SAN

- Brocade DCX 8510 Backbones
- Brocade 825 HBAs
- Brocade 5100 Switch

IP network configuration (Sites A and B)

For the solution, the IP network in each data center is built using two Brocade VDX 6720 switches in a VCS configuration. All servers are connected to the network using redundant 10 GbE connections provided by Brocade 1020 CNAs. The two Brocade VDX switches at each site are connected to a Brocade MLX Series router using a Virtual Link Aggregation Group (vLAG). The Brocade MLX Series routers extend the Layer 2 network between the two data centers.

Note A vLAG is a fabric service that enables a Link Aggregation Group (LAG) to originate from multiple Brocade VDX switches. In the same way as a standard LAG, a vLAG uses the Link Aggregation Control Protocol (LACP) to control the bundling of several physical ports together to form a single logical channel.

Oracle RAC relies on a highly-available virtual IP (the HAIP or RAC interconnect) for private network communication. With HAIP, interconnect traffic is load balanced across the set of interfaces identified as the private network. For this solution, a separate vLAN—VLAN 10—is used for the interconnect. VLAN 20 handles all public traffic. All traffic between Site A and Site B is routed through the Brocade MLX routers using multiple ports configured as a LAG. Figure 18 shows the IP network infrastructure in the two data centers.

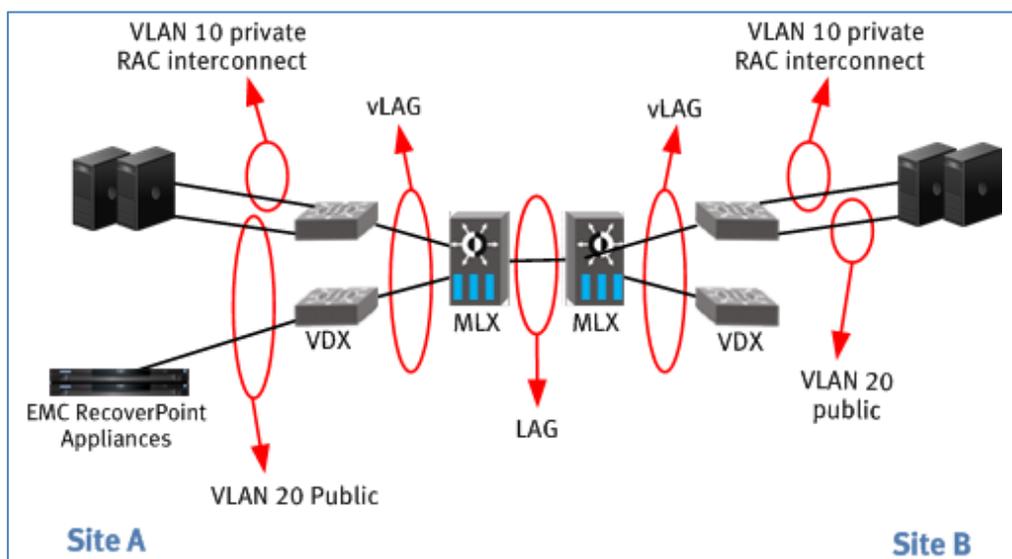


Figure 18. IP networks: Site A and Site B

SAN configuration

The SAN at Sites A and B is built with Brocade DCX 8510 Backbones, as shown in Figure 19. Site C uses Brocade 5100 switches. All servers are connected to the SAN using redundant 8 Gb connections that are provided by Brocade 825 HBAs.

The VPLEX-to-VPLEX connection between the data centers at Sites A and B uses multiple FC connections between the Brocade DCX 8510 Backbones. These are used in active/active mode with failover. Sites A and C communicate with each other over a 1 Gb WAN link.

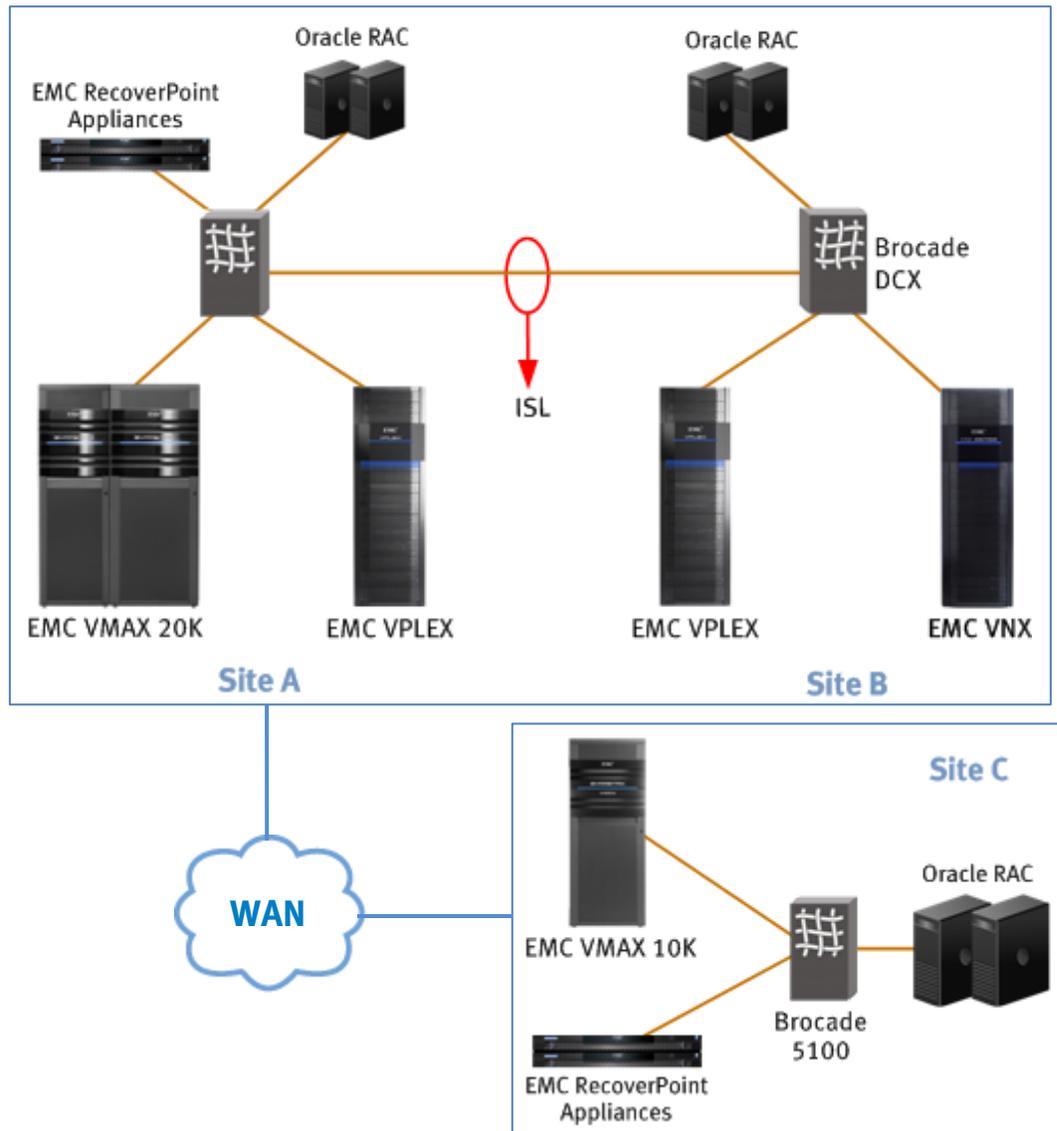


Figure 19. SAN: Sites A, B, and C

EMC storage infrastructure

Overview

This section describes the storage infrastructure for the solution:

- An EMC Symmetrix VMAX 20K array provides the storage platform at Site A.
- An EMC VNX5700 array provides the storage platform at Site B.
- An EMC Symmetrix VMAX 10K array provides the storage platform at Site C.

EMC Symmetrix VMAX

Overview

The EMC Symmetrix VMAX family is a comprehensive range of enterprise storage arrays that are purpose-built to be the foundation for enterprises of all sizes as they transform their IT environment to take advantage of the Hybrid Cloud.

Built on the strategy of simple, intelligent, modular storage, VMAX arrays incorporate a highly scalable Virtual Matrix Architecture that enables them to grow seamlessly and cost-effectively from an entry-level configuration into the world's largest storage system. All VMAX arrays are based on Intel Xeon processors and are optimized for the virtual data center. The EMC Enginuity operating environment provides the intelligence that controls all components in a VMAX array.

VMAX 20K

The VMAX 20K is ideal for high-end configurations that require performance and incremental scalability that enables you to meet your growth requirements by adding VMAX engines and disk drives non-disruptively to the existing frame. This enables true pay-as-you-grow economics for high-growth storage environments.

VMAX 10K

The VMAX 10K is designed for the growing number of IT organizations and service providers with demanding storage requirements and limited resources. It uses the same powerful and revolutionary scale-out Virtual Matrix Architecture as the VMAX 20K to deliver enterprise performance and availability at mid-tier prices.

EMC VNX5700

The VNX5700 is a member of the VNX series next-generation storage platform, which is designed to deliver maximum performance and scalability for mid-tier enterprises, enabling them to dramatically grow, share, and cost-effectively manage multiprotocol file and block systems.

The VNX series uses the Intel Xeon 5600 series processors, which help make it 2 to 3 times faster overall than its predecessor. The VNX quad-core processor supports the demands of advanced storage capabilities such as virtual provisioning, compression, and deduplication. Furthermore, performance of the Xeon 5600 series enables EMC to realize its vision for Fully Automated Storage Tiering (FAST) on the VNX, with optimized performance and capacity.

EMC Unisphere

EMC Unisphere is a storage management platform with an intuitive user interface and support for EMC's complete storage portfolio, including the VMAX family, VPLEX, and the VNX family. Unisphere provides customers with the same EMC-standard look and feel across all these platforms.

Testing and validation

Introduction

Overview

The EMC solutions engineering team began the testing and validation of this solution with an existing VPLEX infrastructure in place. This common infrastructure is shared with other EMC solutions, including that described in the white paper: *EMC Mission-Critical Business Continuity for SAP*.

For this solution, we added EMC RecoverPoint into the VPLEX environment to provide remote DR protection for an Oracle RAC database. The remote infrastructure for RecoverPoint replication was deployed on the site that hosts the VPLEX Witness component of the existing VPLEX environment (that is, Site C).

Test scenarios

The tests carried out for the current solution demonstrate the benefits of adding RecoverPoint to the environment. The test scenarios include:

- Validating that a crash-consistent RecoverPoint image of the database can be recovered to a specific point in time from the replica at Site C and the Oracle RAC instances restarted.
- Validating that production can be switched to Site C in the event of database corruption, viruses, or other problems with the production database.
- Validating that, after failover to Site C, a crash-consistent RecoverPoint image of the database can be recovered to a specific point in time from the replica at Site A (formerly the production source).

For testing purposes, a Swingbench workload was run against the production and replica sites as and when required, as shown in Figure 20.

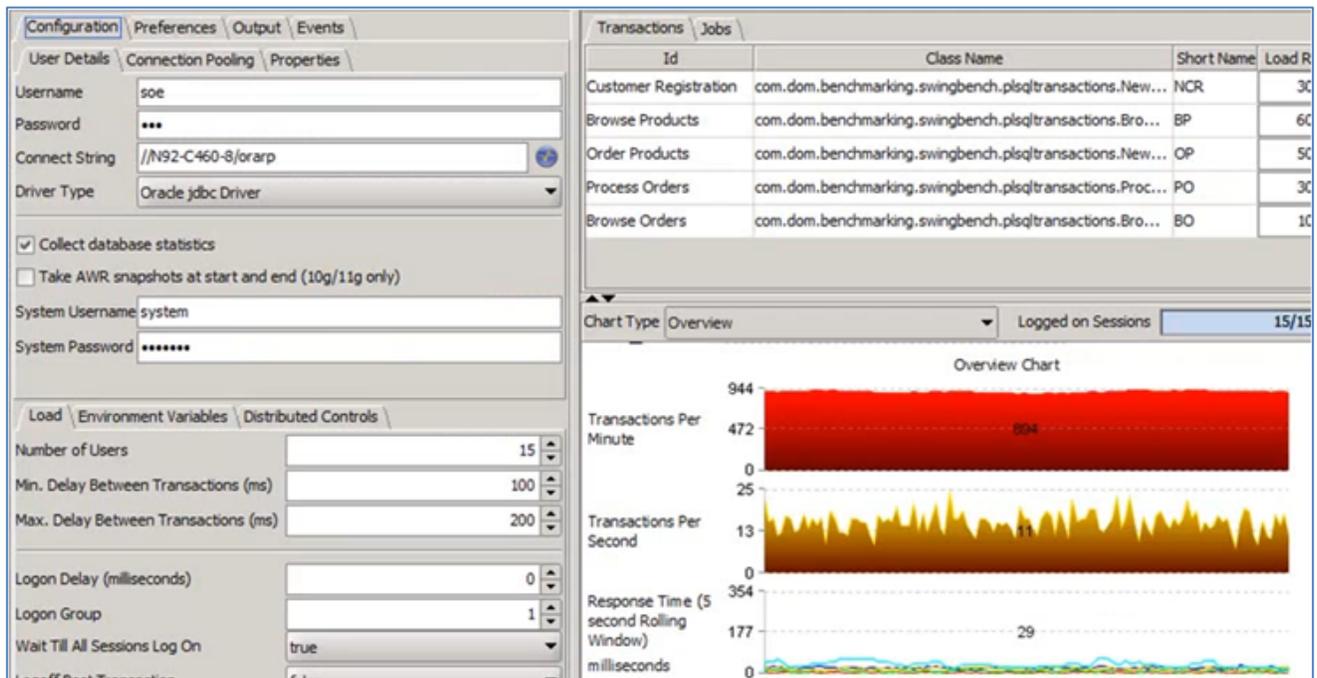


Figure 20. Swingbench workload

Validate the replica at Site C

Overview

This test scenario validates that RecoverPoint can successfully replicate the Oracle RAC database from the VPLEX production environment to the Oracle RAC environment at Site C, and that the database can be recovered at Site C to selected points in time and for a variety of purposes.

The main steps are:

1. Manually create a bookmark.
Note: RecoverPoint continuously creates point-in-time snapshots. For test purposes, a bookmark (named snapshot) is created manually to enable quick and easy recovery to a specific point in time.
2. Enable image access to the bookmarked image at Site C.
3. Recover the database into the RAC environment at Site C.
4. Verify the integrity of the database at Site C.

The screenshot in Figure 21 shows the status of the RecoverPoint environment before these steps were performed.

Site A is the production source; Site C is the remote replica. RecoverPoint is replicating consistency group ORA_RP_VP_ALL to the replica journal and storage at Site C. The local journal at Site A is unused (it is configured so that it can take over the role of replica journal in the event that production needs to fail over to the remote site). The remote replica is not currently enabled for access.

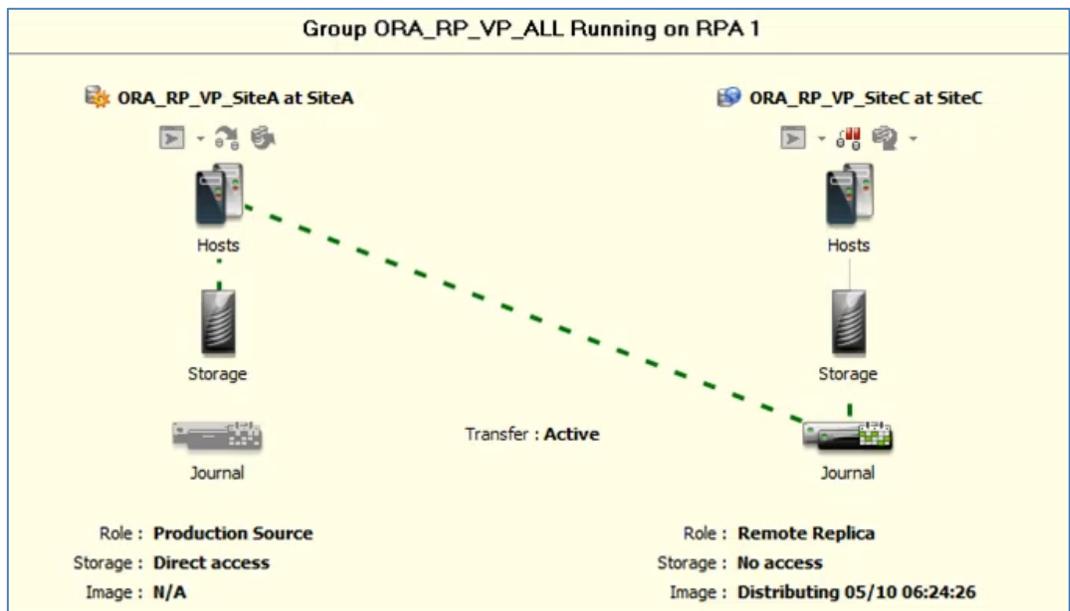


Figure 21. RecoverPoint environment prior to testing

Test procedure

For this test scenario:

1. With Swingbench running load against the Oracle RAC database in the VPLEX environment, insert records into the production database. Figure 22 shows the record count and the timestamp of the last entry. This information was later used to validate the integrity of the data at Site C.

```
oracle@sse-ea-erac-n01:~/scripts> ./rp_DisplayOra.sh
USER is "TEST"

TOTAL RECORD INSERTED: 5000 HIGHER TIMESTAMP INSERTED: 11-MAY-12 07.32.48.684667 AM
oracle@sse-ea-erac-n01:~/scripts>
```

Figure 22. Record count and timestamp at production site (Site A)

2. Using the RecoverPoint Management Application, create a bookmark and name it appropriately for easy identification in the journal, as shown in Figure 23 and Figure 24.

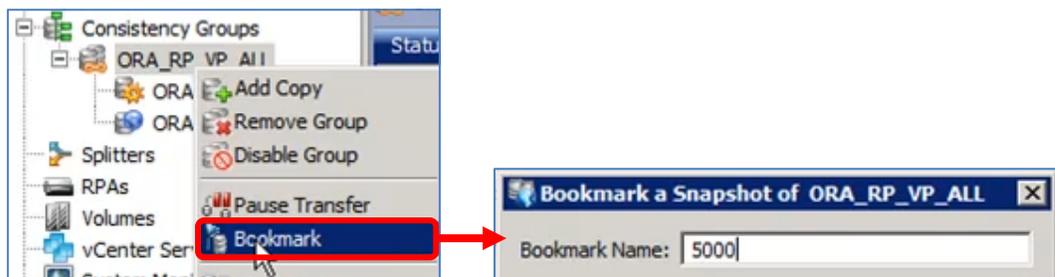


Figure 23. Creating bookmark '5000'

3. Enable host access to the bookmarked image at Site C. You do this by using the **Enable Image Access** option for the CRR replica at Site C and selecting the bookmark created in Step 2, as shown in Figure 24.

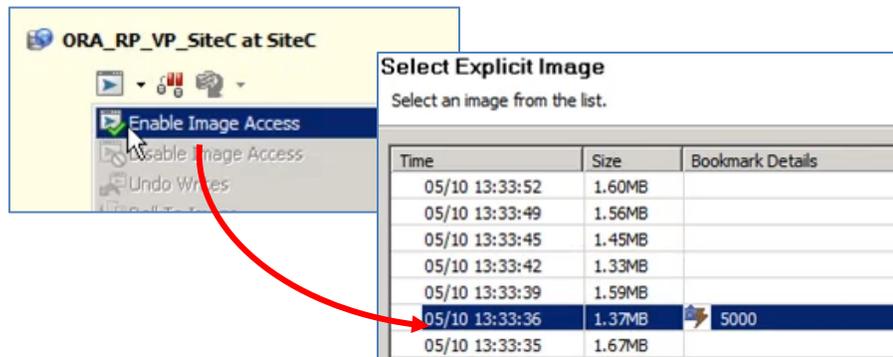


Figure 24. Enabling image access to the bookmarked image

The system rolls the replica storage to the bookmarked point in time and host access to the bookmarked image is then enabled, as shown by the screenshot in Figure 25.

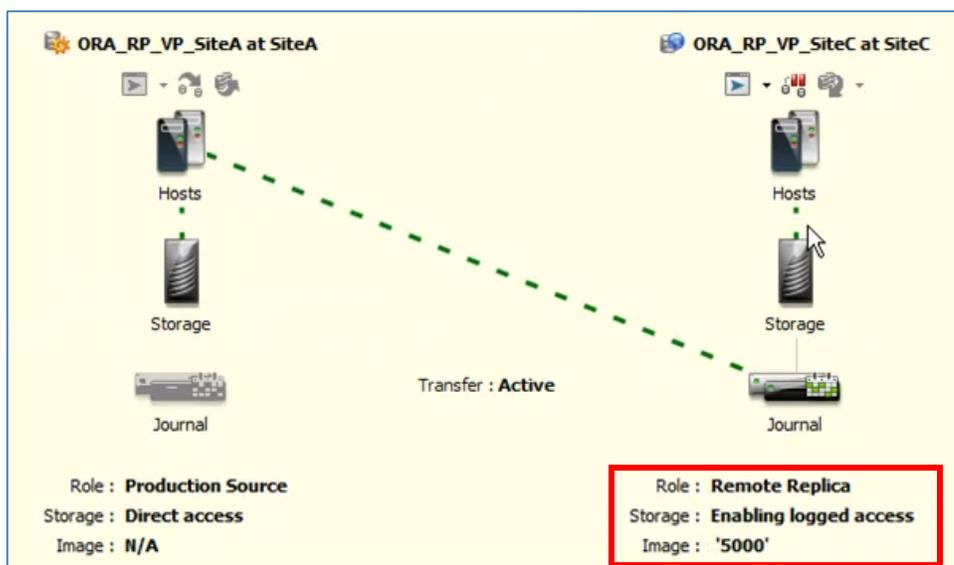


Figure 25. Image access enabled at the remote replica (Site C)

4. Recover the RAC database at Site C so that the bookmarked image can start being used for processing.
5. To verify data integrity at Site C, display the record count and timestamp from the recovered database. Figure 26 shows that these match the values recorded at the production site (see Figure 22).

```
oracle@N92-C460-8:~/scripts> ./rp_DisplayORA.sh
USER is "TEST"

TOTAL RECORD INSERTED: 5000 HIGHER TIMESTAMP INSERTED: 11-MAY-12 07.32.48.684667 AM
oracle@N92-C460-8:~/scripts>
```

Figure 26. Record count and timestamp at the remote replica site (Site C)

Result

Host applications at Site C now have access to the replica volumes, which have been recovered to the bookmarked point in time and the data validated. The Swingbench session on the production database is unaffected throughout the procedure and RecoverPoint replication to Site C continues uninterrupted.

Fail over to the replica at Site C

Overview

This test scenario validates that the Oracle RAC environment can be failed over to the remote replica at Site C in the event of database corruption or other problems in the VPLEX environment. It assumes that the bookmark created in the previous test scenario is a valid image of the database and that recovery of the database into the RAC environment at Site C has already successfully completed.

To fail over to the replica site, you use the **Failover to ORA_RP_VP_SiteC** option for the remote replica at Site C, as shown in Figure 27.

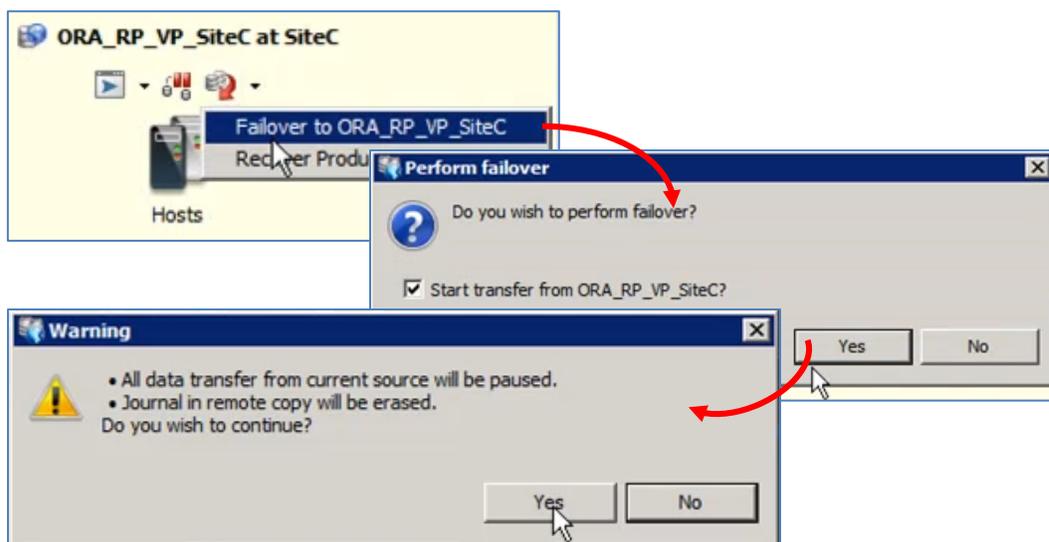


Figure 27. Fail over to Site C

During the failover process, RecoverPoint:

- Promotes the bookmarked image at Site C to become the production source
- Erases all entries in the replica journal at Site C
- Starts replicating to Site A (formerly the production site)

Result

The roles of the former production source and remote replica are switched, as shown by the screenshot in Figure 28. Site C is now the production source and Site A is the remote replica. RecoverPoint is replicating consistency group ORA_RP_VP_ALL to the replica journal and storage at Site A. The remote replica is not currently enabled for access.

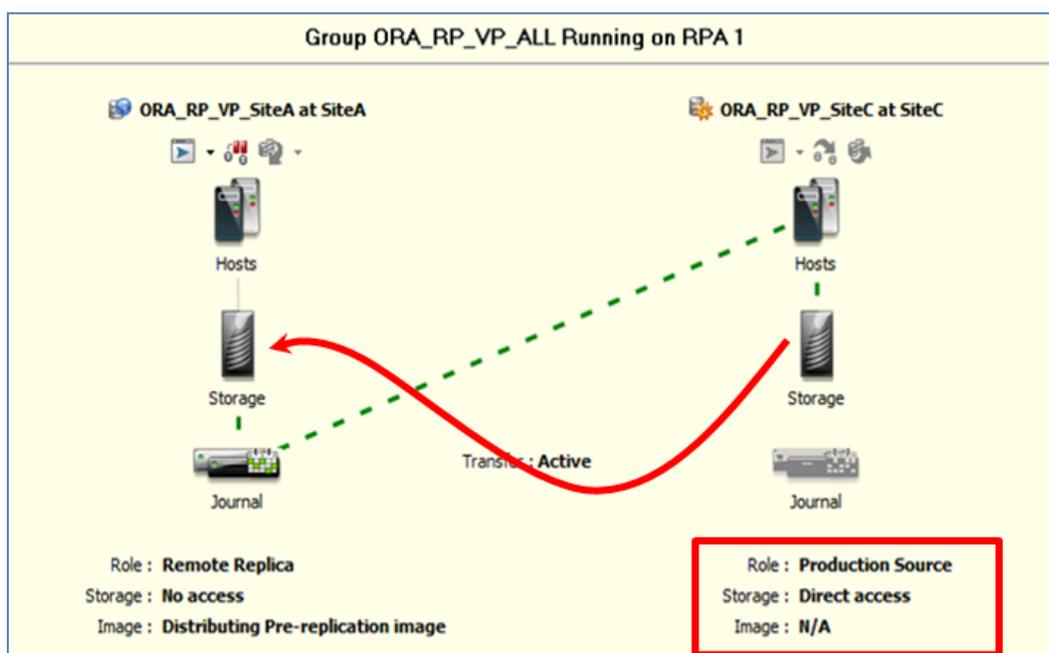


Figure 28. Production failed over to Site C

Validate the replica at Site A

Overview

To validate the replica at the former production site (Site A), you follow the same procedure as for the first test scenario, except with the roles of Site A and Site C reversed.

1. With Swingbench running load against the Oracle RAC database on Site C, insert records into the production database. Figure 29 shows the record count and the timestamp of the last entry. This information was later used to validate the integrity of the data at Site A (now the remote replica).

```
oracle@N92-C460-8:~/scripts> ./rp_DisplayORA.sh
USER is "TEST"

TOTAL RECORD INSERTED: 11000 HIGHER TIMESTAMP INSERTED: 11-MAY-12 07.59.17.291895 AM
oracle@N92-C460-8:~/scripts>
```

Figure 29. Record count and timestamp at replica site (Site A)

2. Using the RecoverPoint Management Application, create a bookmark and name it appropriately for easy identification.
3. Enable host access to the bookmark image at Site A. You do this by using the **Enable Image Access** option for the remote replica at Site A and selecting the bookmark created in Step 2.

The system rolls back to the bookmarked point in time and host access to the bookmarked image is then enabled.

4. Recover the RAC database at Site A so the bookmarked image can start being used for processing.
5. To verify data integrity at Site A, display the record count and timestamp for the recovered database. Figure 30 shows that these match the values recorded at the current production source (see Figure 29).

```
oracle@sse-ea-erac-n01:~/scripts> ./rp_DisplayOra.sh
USER is "TEST"

TOTAL RECORD INSERTED: 11000 HIGHER TIMESTAMP INSERTED: 11-MAY-12 07.59.17.291895 AM
oracle@sse-ea-erac-n01:~/scripts>
```

Figure 30. Record count and timestamp at replica site (Site A)

Result

The roles of the former production source and remote replica have switched and RecoverPoint is now replicating to the former production journal and from that journal to the former production storage. The journal of the former replica at Site C becomes invalid since Site C is now the source.

The bookmark created at the new remote replica site is a valid image and can be used to continue working if any problem arises on the production site. The Swingbench session to the production database at Site C is unaffected throughout the procedure and RecoverPoint commences replicating to the VPLEX virtual volumes.

**Solution
characterization
data**

Table 5 provides characterization data for the functionality tested for this solution.

Table 5. Solution characterization data

Operation	Finding
Failover from Site A to Site C	1 minute 3 seconds
Failover from Site C to Site A	1 minute 25 seconds
RecoverPoint snapshot frequency	4 seconds
Enable latest snapshot	4 mouse clicks
Create a bookmark	3 mouse clicks
Enable bookmark	7 mouse clicks
Initiate a failover	3 mouse clicks

Conclusion

Summary

This solution demonstrates how to leverage EMC VPLEX and EMC RecoverPoint to provide business continuity and disaster recovery for mission-critical Oracle Extended RAC databases in heterogeneous storage environments:

- EMC VPLEX Metro provides a virtual storage layer that enables an active/active Metro data center with 24/7 application availability, no single points of failure, and zero RPOs and RTOs. In the event of server, site, data center, or network failure, the Oracle database remains available.
- EMC RecoverPoint replication technology uses sophisticated journaling and write splitting to enable continuous replication to a third DR site. This enables fast and easy point-in-time recovery.

The three-site topology deployed for this solution provides these main benefits:

- With VPLEX enabling active/active data centers with zero RTOs and RPOs, customers can avoid the downtime associated with site loss and storage array failures.
- With RecoverPoint replicating production data to a third site, customers can rapidly recover from application data corruption, viruses, or human error, and from wide-scale physical disasters that affect both VPLEX Metro clusters.
- With RecoverPoint journaling technology, customers can recover to specific points in time, in case the latest image is corrupt.

In addition, deploying Oracle RAC on Extended Distance Clusters over VPLEX provides these benefits:

- Simplified management of deployment—installation, configuration, and maintenance are the same as for a single site RAC deployment.
- Hosts connect only to their local VPLEX cluster, but have full read/write access to the same database at both sites.
- No need to deploy Oracle voting disk and Clusterware on a third site.
- Eliminates the costly host CPU cycles consumed by ASM mirroring—I/O is sent only once from the host to the local VPLEX.
- Ability to create consistency groups that protect multiple databases and/or applications as a unit.

Findings

The tests performed to validate the functionality of this solution demonstrate that:

- A crash-consistent image of the database can be recovered to a specific point in time from the remote RecoverPoint replica.
- Production can be switched to the remote replica site in the event of database corruption, viruses, or other problems with the production database.
- After failover to the remote site, a crash-consistent image of the database can be recovered to a specific point in time from the RecoverPoint replica at the former production site.

References

EMC

For additional information, see the following EMC documents (available on EMC.com and on the EMC online support website):

- *EMC Mission Critical Business Continuity for SAP*
- *EMC RecoverPoint: Adding Application Recovery To VPLEX Local And Metro*
- *EMC VPLEX GeoSynchrony Release 5.1 Administration Guide*
- *EMC RecoverPoint: Deploying with VPLEX Technical Notes*
- *EMC VPLEX Metro Witness Technology and High Availability*
- *Conditions for Stretched Hosts Cluster Support on EMC VPLEX Metro*
- *Oracle Extended RAC with EMC VPLEX Metro Best Practices Planning*
- *EMC VPLEX with GeoSynchrony 5.0 Configuration Guide*
- *Implementation and Planning Best Practices for EMC VPLEX—Technical Notes*
- *EMC VPLEX with GeoSynchrony 5.0 and Point Releases CLI Guide*
- *EMC Simple Support Matrix for EMC VPLEX and GeoSynchrony*
- *Validating Host Multipathing with EMC VPLEX—Technical Notes*

Oracle

For additional information, see the following Oracle documents:

- *Oracle Real Application Clusters (RAC) on Extended Distance Clusters*
- *Oracle RAC Technologies Matrix—Linux*
(www.oracle.com/technetwork/database/enterprise-edition/tech-generic-linux-new-086754.html)
- *Oracle RAC Technologies Matrix—Generic*
(www.oracle.com/technetwork/database/clustering/tech-generic-unix-new-166583.html)