White Paper

EMC MISSION-CRITICAL BUSINESS CONTINUITY FOR SAP

EMC VPLEX, Symmetrix VMAX, VNX, VMware vSphere HA, Brocade Networking, Oracle RAC, SUSE Linux Enterprise

- Simplified management for high availability and business continuity
- Resilient mission-critical SAP deployments
- Active/active data centers

EMC Solutions Group

Abstract

This white paper describes the transformation of a traditional SAP deployment to a mission-critical business continuity solution with active/active data centers. The solution is enabled by EMC® VPLEX™ Metro, EMC Symmetrix® VMAX™, EMC VNX™, VMware vSphere® HA, Oracle RAC, Brocade networking, and SUSE Linux Enterprise Server for SAP Applications.

June 2012

EMC² PROVEN SOLUTIONS

BROCADE

SUSE

SAP Global Partner

vmware PARTNER TECHNOLOGY ALLIANCE

intel Xeon inside

EMC²

Part Number H9542.1

Table of contents

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 business continuity and high-availability strategy for mission-critical applications such as SAP ERP.

Recovery point objectives (RPOs) and recovery time objectives (RTOs) are key metrics when planning a mission-critical business continuity strategy. They answer two fundamental questions that businesses pose when they consider 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 for SAP demands aggressive RPOs and RTOs to minimize data loss and recovery times. The main challenges that the business must consider when designing such a strategy include:

- Minimizing RPO and RTO
- Eliminating single points of failure (SPOFs)—technology, people, processes
- Maximizing resource utilization
- Reducing infrastructure costs
- Managing the complexity of integrating, maintaining, and testing multiple point solutions

This white paper introduces an EMC solution that addresses all these challenges for SAP ERP applications with an Oracle Real Applications Clusters (RAC) 11*g* database layer.

The solution demonstrates an innovative active/active deployment model for data centers up to 100 km apart. This transforms the traditional active/passive disaster recovery (DR) model to a highly available business continuity solution, with 24/7 application availability, no single points of failure, and near-zero RTOs and RPOs.

Solution overview

EMC® VPLEX™ Metro is the primary enabling technology for the solution. VPLEX Metro is a storage area network-based (SAN) 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 VPLEX Witness added to the solution, applications continue to be available, with no interruption or downtime, even in the event of disruption at one of the data centers.

The white paper demonstrates how the following technologies create this innovative business continuity solution:

- EMC VPLEX Metro provides the virtual storage layer that enables an active/active Metro data center.
- EMC VPLEX Witness supports continuous application availability, even in the event of disruption at one of the data centers.

- EMC Symmetrix® VMAX™ and EMC VNX™ arrays, with proven five 9s availability, support for Fully Automated Storage Tiering (FAST), and a choice of replication technologies, provide the enterprise-class storage platform for the solution.
- Migrating from a single instance database to Oracle RAC on Extended Distance Clusters removes single points of failure at the database layer, across distance.
- VMware vSphere® virtualizes the SAP application components and eliminates these as single points of failure. VMware® High Availability (HA) protects the virtual machines in the case of physical server and operating system failures.
- SUSE Linux Enterprise Server for SAP Applications, with SUSE Linux Enterprise High Availability Extension and SAP Enqueue Replication Server (ERS), protects the SAP central services (message server and enqueue server) across two cluster nodes to eliminate these services as single points of failure.
- Brocade Ethernet fabrics and MLXe core routers provide seamless networking and Layer2 extension between sites.
- Brocade DCX 8510 Backbones provide redundant SAN infrastructure, including fabric extension.

Key benefits

The solution increases the availability for SAP applications by:

- Eliminating single points of failure at all layers in the environment to build a distributed and highly available SAP system
- Providing active/active data centers that support near-zero RPOs and RTOs and mission-critical business continuity

Additional 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 SAP high-availability management
- Simplified deployment of Oracle RAC on Extended Distance Clusters
- Reduced costs by increasing automation and infrastructure utilization

Introduction

Purpose

This white paper describes a solution that increases availability for SAP applications by creating active/active data centers in geographically separate locations and eliminating single points of failure at all layers in the environment.

In SAP environments, business disruption can result from technical, logical, or logistical failures. This solution addresses business continuity from the technical perspective.

Scope

The scope of the white paper is to:

- Introduce the key enabling technologies
- Describe the solution architecture and design
- Describe how the key components are configured
- Describe the steps used to convert an Oracle single instance database to a four-node Oracle RAC cluster on Oracle Automatic Storage Management (ASM)
- Present the results of the tests performed to demonstrate the elimination of single points of failure at all layers in the environment
- Identify the key business benefits of the solution

Audience

This white paper is intended for SAP Basis Administrators, Oracle DBAs, storage administrators, IT architects, and technical managers responsible for designing, creating, and managing mission-critical SAP applications in 24/7 landscapes.

Terminology

This white paper includes the terms in Table 1.

Table 1. Terminology

Term	Description
ABAP	SAP Advanced Business Application Programming
ACFS	Oracle ASM Cluster File System
ASCS	ABAP SAP Central Services
ASM	Oracle Automatic Storage Management
CIFS	Common Internet File System
CNA	Converged network adapter
CRM	Cluster resource manager
DI	Dialog instance
DPS	Dynamic Path Selection
DRS	VMware vSphere Distributed Resource Scheduler
dvSwitch	vSphere distributed switch
DWDM	Dense wavelength division multiplexing
ERP	Enterprise resource planning
ERS	Enqueue Replication Server
FAST VP	Fully Automated Storage Tiering for Virtual Pools
FCoE	Fibre Channel over Ethernet
FEC	Forward Error Correction

Term	Description
FRA	Flash Recovery Area
HA	High availability
HAIP	Highly available virtual IP
HBA	Host bus adapters
IDES	SAP Internet Demonstration and Evaluation System
ISL	Inter-Switch Link
LACP	Link Aggregation Control Protocol
LAG	Link Aggregation Group
LLDP	Link Layer Discovery Protocol
LUW	Logical unit of work
MCT	Multi-Chassis Trunking
MPLS	Multi-Protocol Label Switching
MPP	Multipathing plug-in
NAS	Network-attached storage
NFS	Network File System
NL-SAS	Nearline SAS (Serial Attached SCSI)
OCR	Oracle Cluster Registry
Oracle Extended RAC	Oracle RAC on Extended Distance Clusters
RAC	Real Application Clusters
RFC	Remote function call
RMAN	Oracle Recovery Manager
RPO	Recovery point objective
RTO	Recovery time objective
SAN	Storage area network
SBD	STONITH block device
SFP	Small Form-Factor Pluggable
SLES	SUSE Linux Enterprise Server
SLE HAE	SUSE Linux Enterprise High Availability Extension
SMT	Subscription Management Tool
SPOF	Single point of failure
STONITH	Shoot The Other Node In The Head
TAF	Transparent Application Failover
ToR	Top-of-Rack
VCS	Virtual cluster switch
vLAG	Virtual Link Aggregation Group
vLAN	Virtual LAN
VMDK	Virtual disk
VMFS	Virtual Machine File System
VMHA	VMware High Availability
VNX OE	VNX Operating Environment
VPLS	Virtual Private LAN Service
VPN	Virtual private network
VRF	Virtual Routing and Forwarding
VSI	Virtual Storage Integrator

Solution overview

Introduction

SAP implementations – the challenge and the solution

Traditional SAP implementations have several single points of failure (SPOFs), including:

- Central Services
- Enqueue server*
- Message server*
- Database server
- Single site deployment
- Local disk storage

* In this solution, the enqueue and message servers are implemented as services within the ABAP SAP Central Services (ASCS) instance.

This white paper presents a solution for increasing availability for SAP applications. The architecture and components of the solution create an active/active clustered solution for the entire SAP stack to enhance reliability and availability while simplifying the deployment and management of the environment. This provides the following benefits:

- Eliminates single points of failure at all layers in the environment to build a highly available SAP system
- Provides active/active data centers to enable mission-critical business continuity

Figure 1 illustrates the single points of failure in a SAP environment and the solution components used to address them.

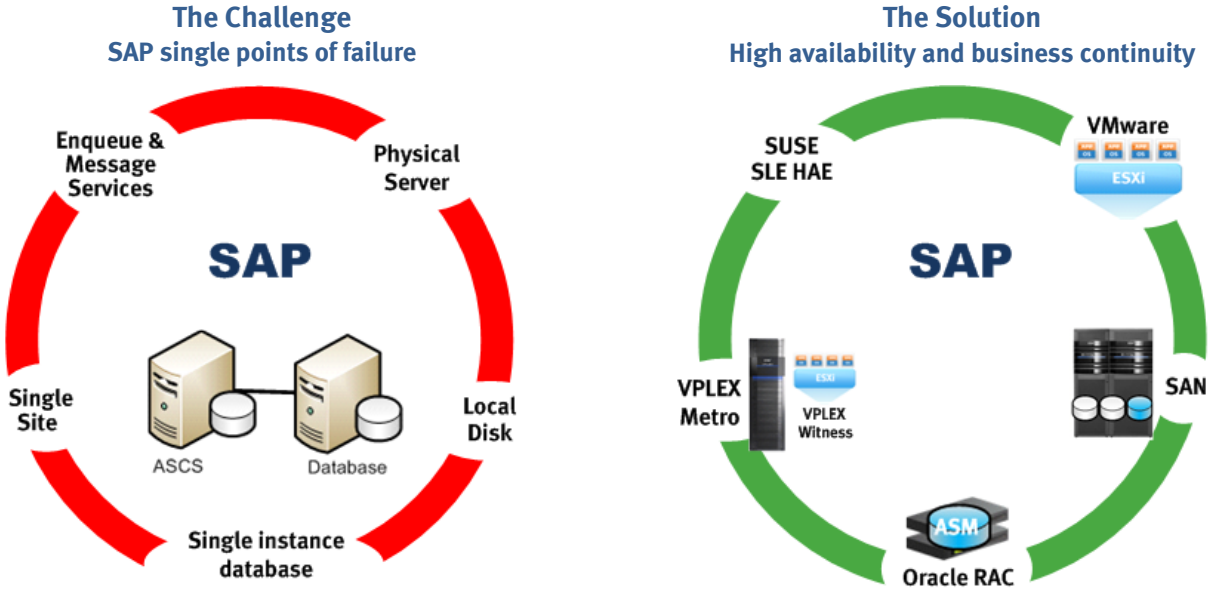

Figure 1. SAP implementations – the challenge, the solution

Solution architecture

The following sections describe the solutions implemented at each layer of the environment to provide high availability and business continuity, as shown in Figure 2.

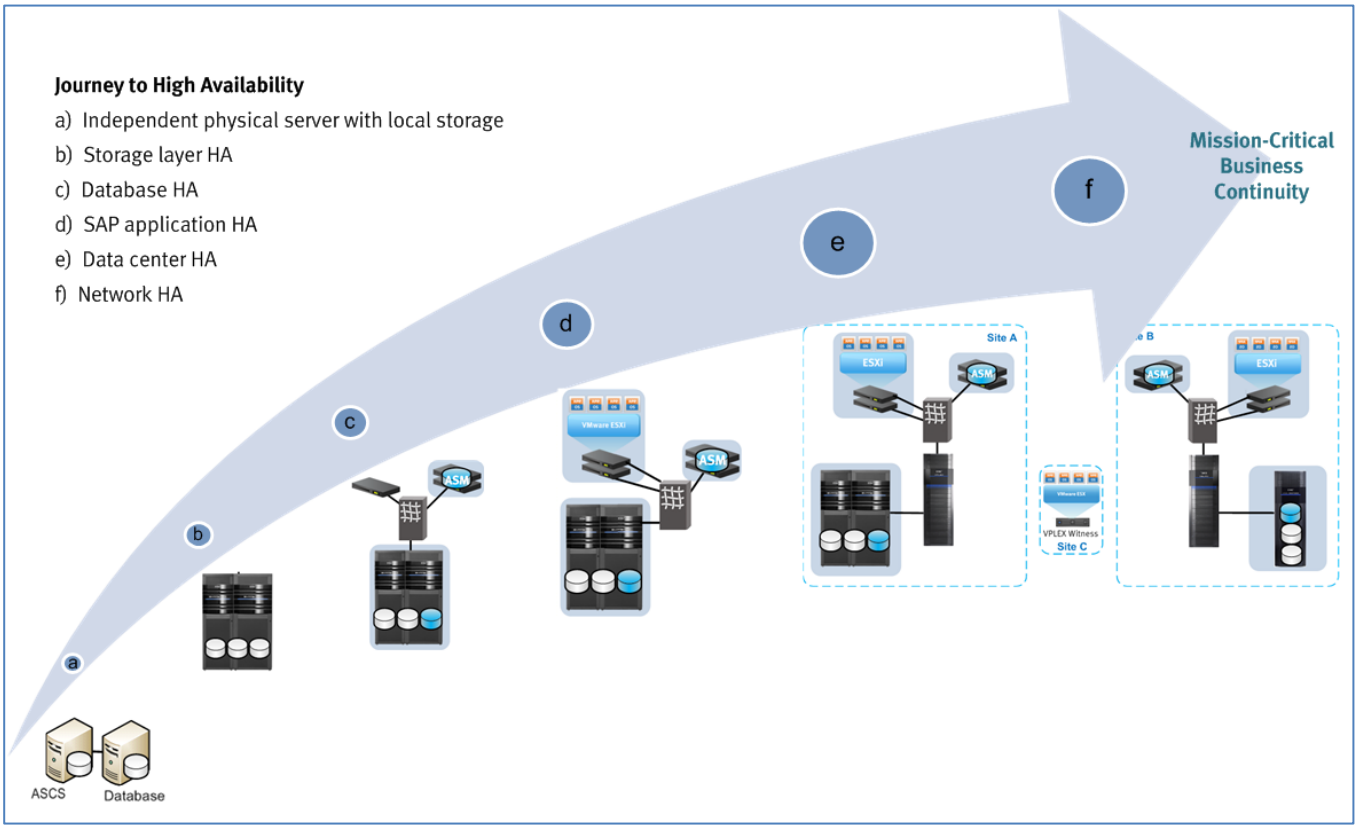

Figure 2. The journey to high availability – logical view

Storage layer high availability

All the storage required by each server in the environment was moved to enterprise-class storage arrays (Symmetrix VMAX and VNX). Brocade 8510 Backbones were deployed to provide a redundant SAN fabric for storage access.

This takes advantage of the proven five 9s uptime provided by the arrays and the SAN Backbones, including their advanced manageability and business continuity features.

Figure 3. Storage HA

Database high availability

The database server is the data repository for the SAP application For this solution, the backend database server was converted from an Oracle single instance database to a four-node Oracle RAC database on Oracle ASM. This eliminates the database server as a single point of failure.

Figure 4. Database HA

SAP application high availability

The SAP application servers were fully virtualized using VMware ESXi™ 5.0. Each of the SAP virtual machines was deployed using SUSE Linux Enterprise Server for SAP Applications 11 SP1 as the guest operating system.

SUSE Linux Enterprise High Availability Extension and SAP Enqueue Replication Server (ERS) were also deployed to protect both the SAP message server and enqueue server. This eliminates the ABAP SAP Central Services (ASCS) as a single point of failure.

Figure 5. SAP application HA

Data center high availability

The high-availability cluster solution described above protects SAP *within* the data center. For high availability *between* data centers, the solution uses EMC VPLEX Metro storage virtualization technology, as shown in Figure 6. The unique VPLEX Metro Access Anywhere active/active clustering technology allows read/write access to distributed volumes across synchronous distances. By mirroring data across locations, VPLEX enables users at both locations to access the same information at the same time.

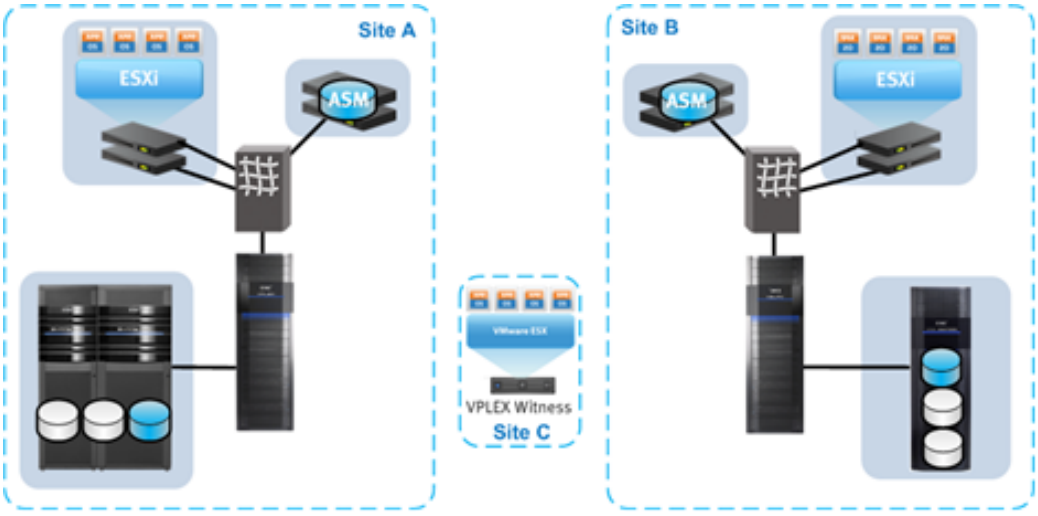

Figure 6. Data center HA

This solution combines VPLEX Metro with SUSE Linux Enterprise High Availability Extension (at the operating system layer) and Oracle RAC (at the database layer) to remove the data center as a single point of failure and provide a robust business continuity strategy for mission-critical applications.

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 exactly 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 to their local VPLEX cluster only 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.

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).

Network high availability

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

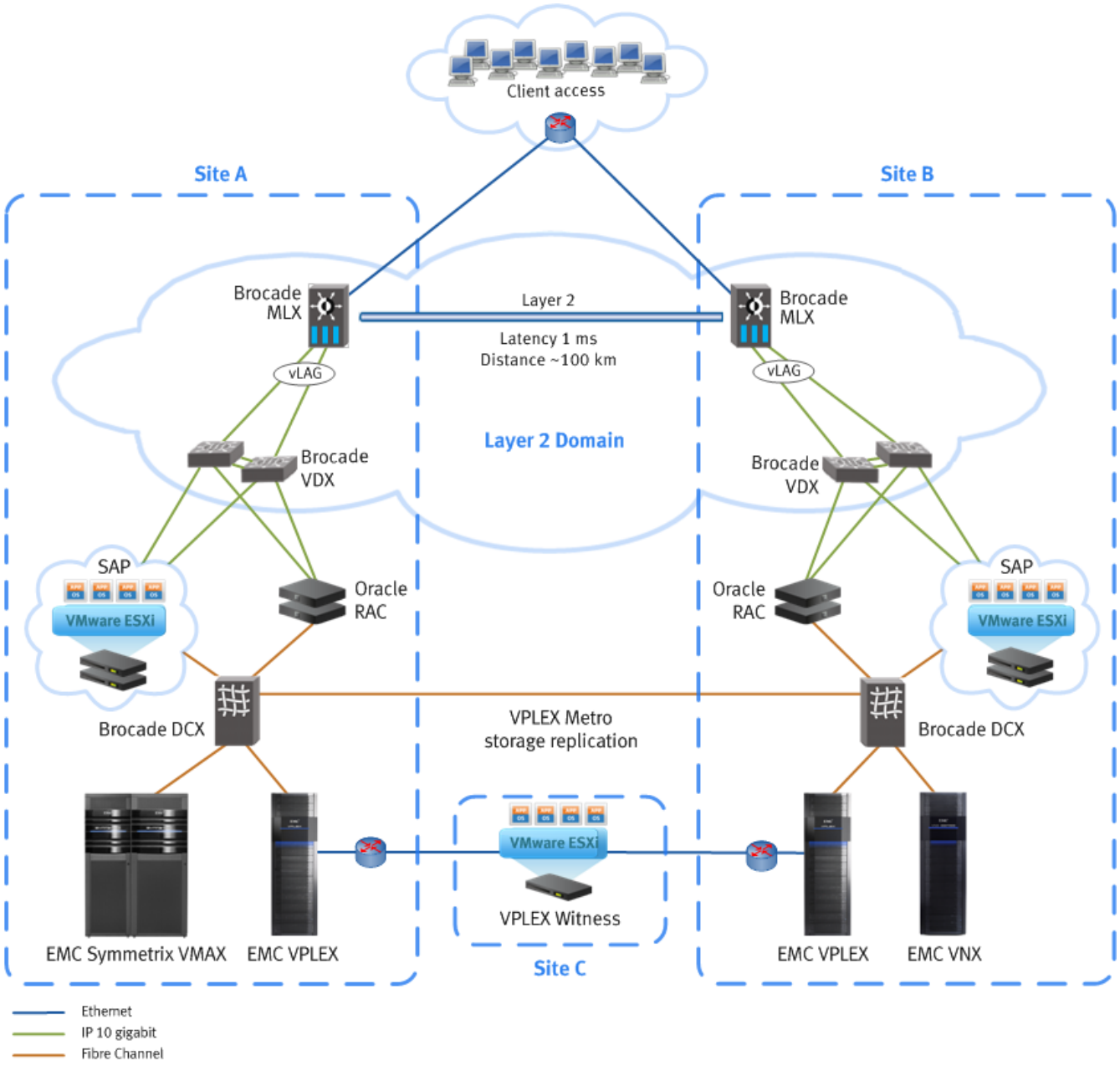

Figure 7. Solution architecture

Protection layers

Table 2 summarizes the high availability (HA) layers that the solution uses to eliminate single points of failure.

Table 2. Local high availability

Local high availability		
Protection	**Site**	**Components protected**
VMware HA and VMware DRS	Sites A & B	SAP virtual machines
SUSE Linux Enterprise HAE and SAP Enqueue Replication Server	Sites A & B	SAP enqueue server, SAP message server
Multiple SAP Dialog Instances	Sites A & B	SAP work processes (DIA, UPD, UP2, SPO)
VMware	Sites A, B, & C	Server virtualization
Oracle RAC	Sites A & B	Oracle database
Oracle Clusterware	Sites A & B	SAP shared file system
Oracle ACFS	Sites A & B	SAP Oracle Home, SAP global directory, SAP transport directory, SAP ASCS directory
EMC Symmetrix VMAX	Site A	Local storage, RAID, multipath
EMC VNX	Site B	Local storage, RAID, multipath

VPLEX Metro then extends the high availability with a clustering architecture that breaks the boundaries of the data center and allows servers at multiple data centers to have read/write access to shared block storage devices. This data center transformation takes traditional high availability to a new level of mission-critical business continuity.

Figure 8 illustrates this high-availability design, with VPLEX Witness and Cross-Cluster Connect deployed to provide the highest level of resilience.

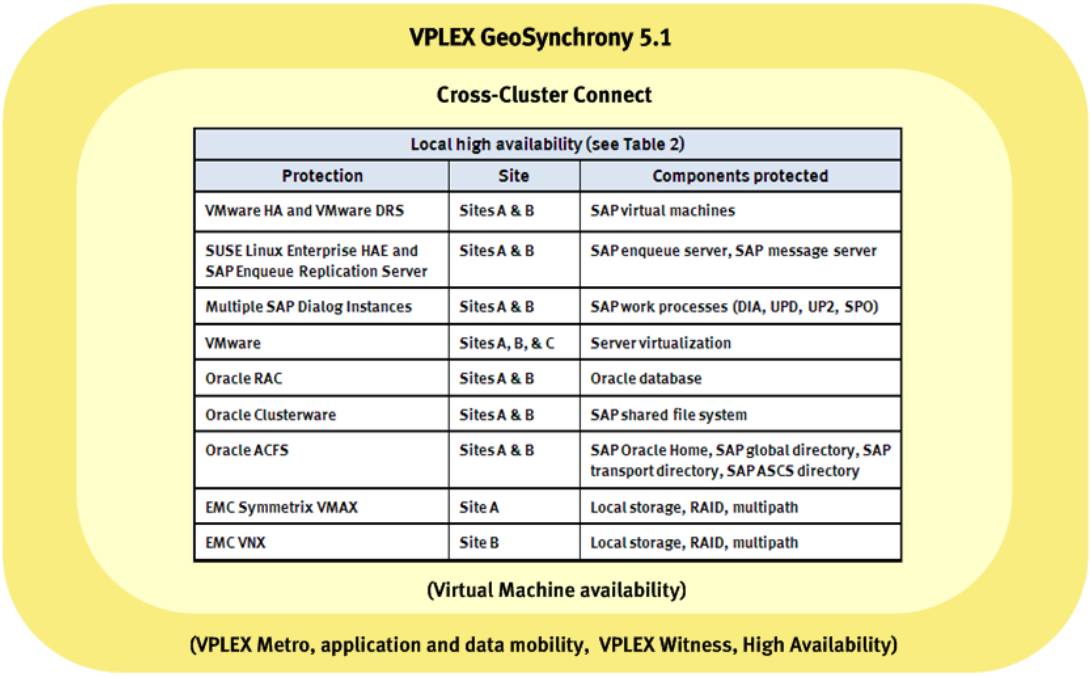

Figure 8. Local HA, with VPLEX enabling multisite business continuity

Each of the technologies shown in Figure 8 is explored in more detail in the relevant sections of the white paper.

Database and workload profile

Table 3 details the database and workload profile for the solution.

Table 3. Database and workload profile

Profile characteristic	Details
Number of databases	1
Database type	SAP OLTP
Database size	500 GB
Database name	VSE
Oracle RAC	4 physical nodes
Workload profile	SAP custom order-to-cash processes

Hardware resources

Table 4 details the hardware resources for the solution.

Table 4. Solution hardware environment

Purpose	Quantity	Configuration
Storage (Site A)	1	EMC Symmetrix VMAX, with: • 2 engines • 171 x 450 GB FC drives • 52 x 2 TB SATA drives
Storage (Site B)	1	EMC VNX5700, with: • 30 x 2 TB NL-SAS drives • 79 x 600 GB SAS drives
Distributed storage federation	2	VPLEX Metro cluster, with: • 2 x VS2 engines
Oracle RAC database servers	4	4 x eight-core CPUs, 128 GB RAM
VMware ESXi servers for SAP	4	2 x four-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	2	Brocade DCX 8510 Backbone, with: • 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

Software resources

Table 5 details the software resources for the solution.

Table 5. Solution software environment

Software	Version	Purpose
EMC Enginuity™	5875.198.148	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 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™	1.1	VNX management software
SUSE Linux Enterprise Server for SAP Applications, including SUSE Linux Enterprise High Availability Extension	11 SP1	Operating system for all servers in the environment
VMware vSphere	5.0	Hypervisor hosting all virtual machines
Oracle Database 11*g* (with Oracle RAC and Oracle Grid Infrastructure)	Enterprise Edition 11.2.0.3	Oracle database and cluster software
SAP ERP	6.04	SAP ERP IDES system

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 9:

- VPLEX Local
- VPLEX Metro
- VPLEX Geo

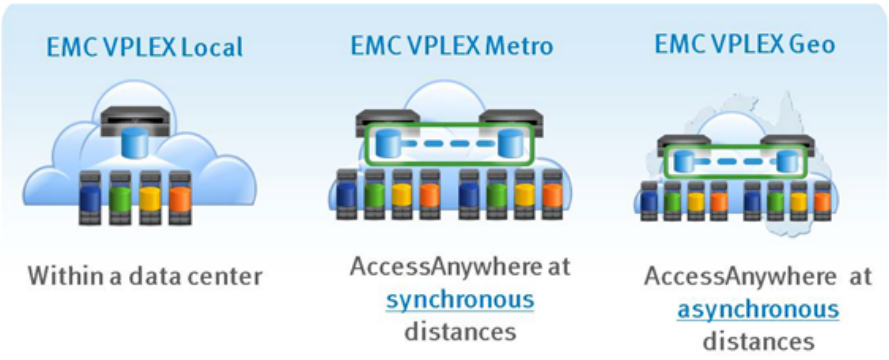

Figure 9. VPLEX topologies

For detailed descriptions of these VPLEX configurations, refer to the documents listed in References on page 73.

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 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 19 for further details).

EMC VPLEX Management Interface

You can manage and administer a VPLEX environment with the web-based VPLEX Management Console or you can connect directly to a management server and start a VPlexcli session (VPLEX command line interface).

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.

An even higher degree of availability can be achieved by using a VPLEX Cross-Cluster Connect configuration. In this case, each host is connected to the VPLEX clusters at both sites. This ensures that, in the unlikely event of a full VPLEX cluster failure, the host has an alternate path to the remaining VPLEX cluster.

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 10.

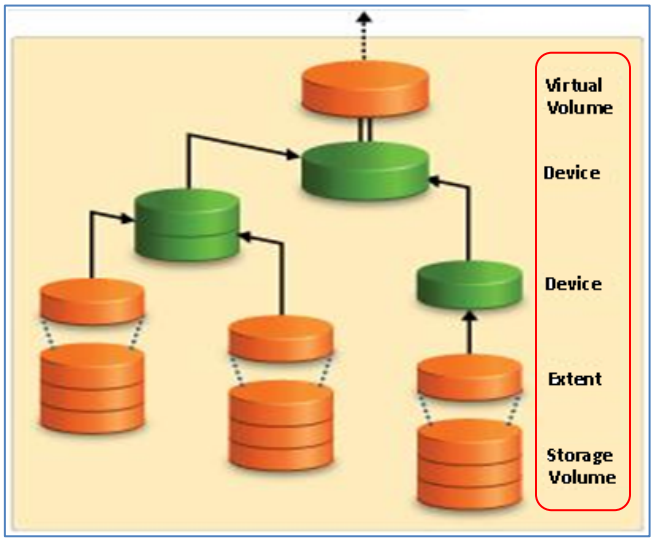

Figure 10. VPLEX logical storage structures

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 a virtual volume to a host 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 consistency group:

- **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 are predefined rules that 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 a detach rule is set it is always invoked when connectivity is lost between clusters. 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 10 shows the physical and logical storage structure used by VPLEX Metro in the context of this solution.

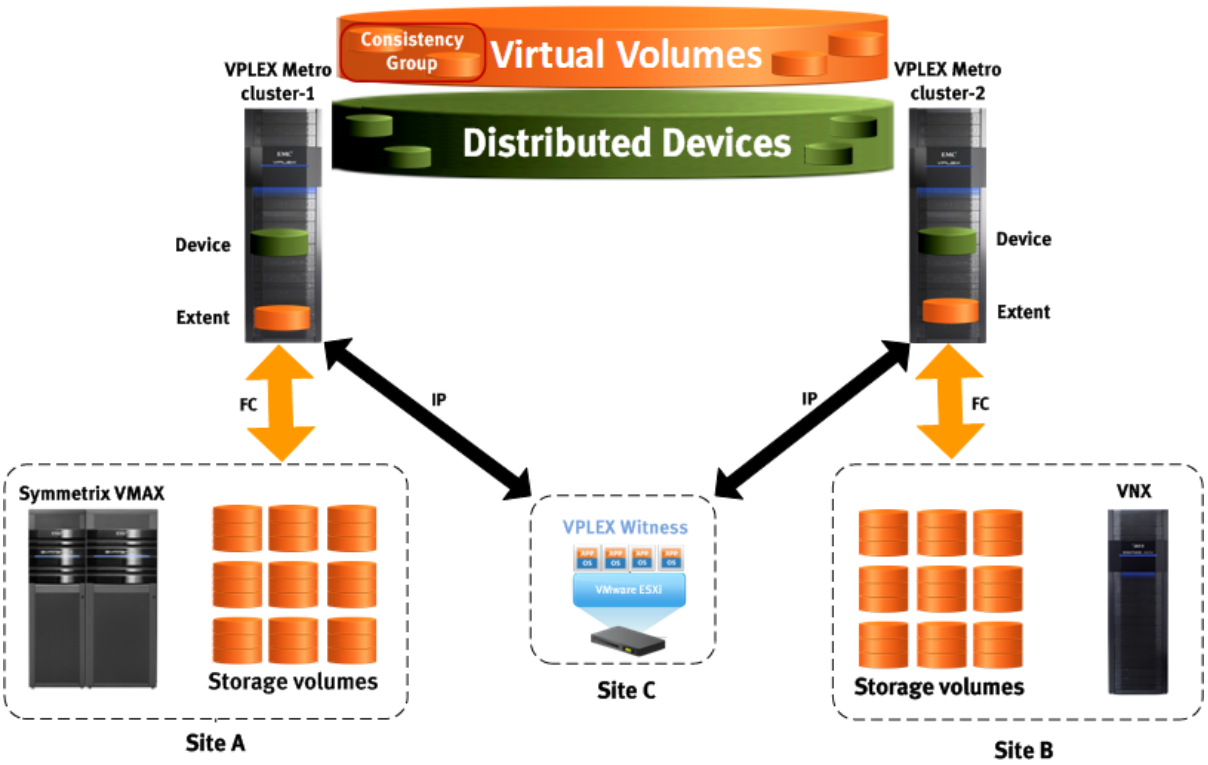

Figure 11. VPLEX physical and logical storage structures for solution

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.

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

Consistency group

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 11*g* database binaries, 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.

Configuration process

For the solution, the VPLEX Metro logical storage structures are configured as follows (Figure 12 to Figure 16 show extracts from the configuration wizards provided by the VPLEX Management Console):

- **Storage volume**—A storage volume is a LUN exported from an array and encapsulated by VPLEX. Figure 12 shows several storage volumes created at Site A, as displayed in the VPLEX Management Console.

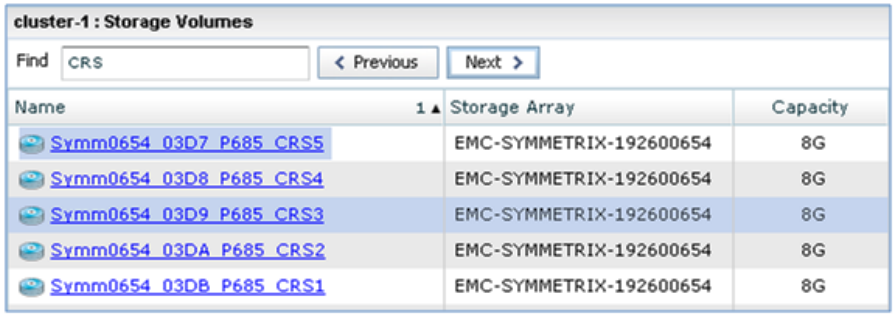

Figure 12. EMC VPLEX storage volumes (Site A)

- **Extent**—In the solution, there is a one-to-one mapping between extents and storage volumes, as shown in Figure 12 and Figure 13.

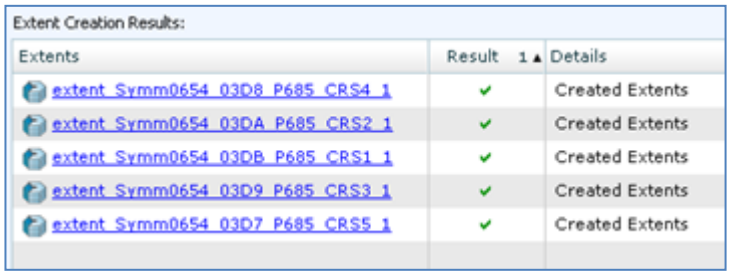

Figure 13. EMC VPLEX Extent Creation wizard

- **Device**—In the solution, there is a one-to-one mapping between devices and extents. Figure 14 shows the option used to configure this one-to-one mapping.

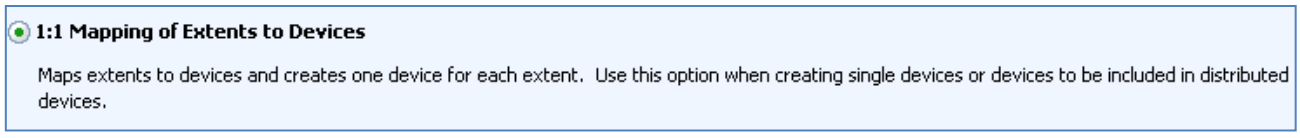

Figure 14. EMC VPLEX Device Creation wizard

- **Distributed device**—In the solution, the distributed devices were created by mirroring a device remotely in a distributed RAID 1 configuration, as shown in Figure 15.

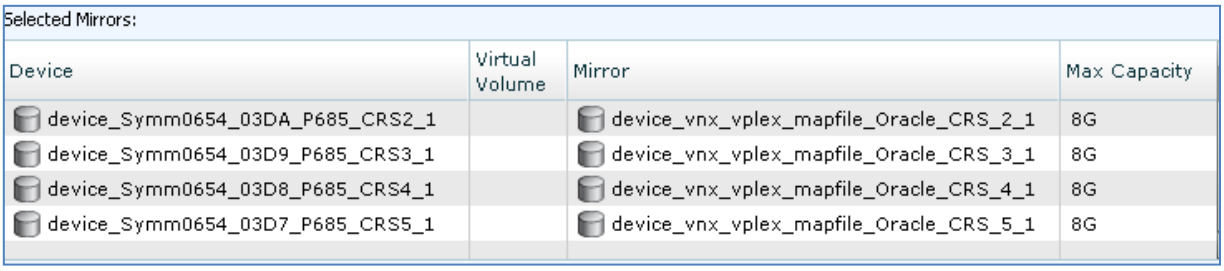

Figure 15. EMC VPLEX Device Creation wizard

- **Virtual volume**—In the solution, all top-level devices are distributed devices. These devices are encapsulated by virtual volumes, which VPLEX presents to the hosts through storage views. The storage views define which hosts access which virtual volumes on which VPLEX ports.
- **Consistency group**—Figure 16 shows the consistency group created for the solution—Extended_Oracle_RAC_CG.

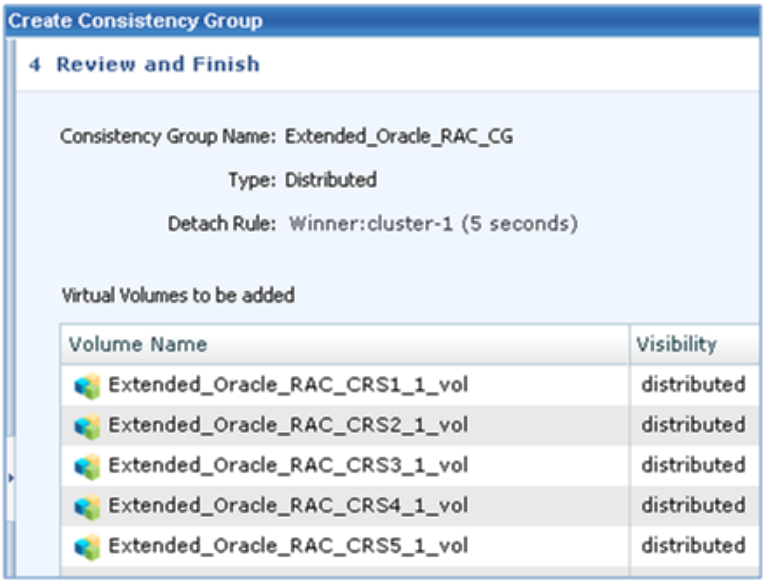

Figure 16. EMC VPLEX Create Consistency Group wizard

VPLEX Witness configuration

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. This is considered a VPLEX Metro HA configuration as storage availability is ensured at the surviving site.

VPLEX Witness is deployed at a third, separate failure domain (Site C) and connected to the VPLEX clusters at Site A and Site B. Site C is located at a distance of less that 1 second latency from Sites A and B.

When a VPLEX Witness has been installed and configured, the VPLEX Management Console displays the status of cluster witness components, as shown in Figure 17.

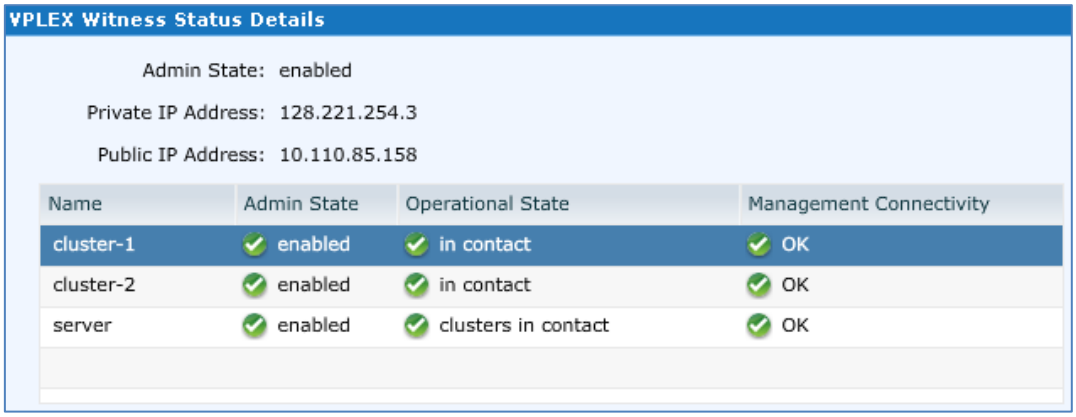

Figure 17. EMC VPLEX Witness components and status

VPLEX performance monitoring

VPLEX 5.1 delivers enhanced performance monitoring through the performance monitoring dashboard. This dashboard provides a customizable view into the performance of the VPLEX system and enables you to view and compare different aspects of system performance, down to the director level.

Many different metrics are currently available, including:

- Front-end Latency chart
- Front-end Bandwidth chart
- Front-end Throughput chart
- CPU Utilization chart
- Rebuild Status chart
- WAN Link Performance chart
- Back-end Latency chart

Figure 18 shows the front-end and CPU performance on cluster-1 (the Site A VPLEX) when Oracle statistics were gathered on the SAP VSE database.

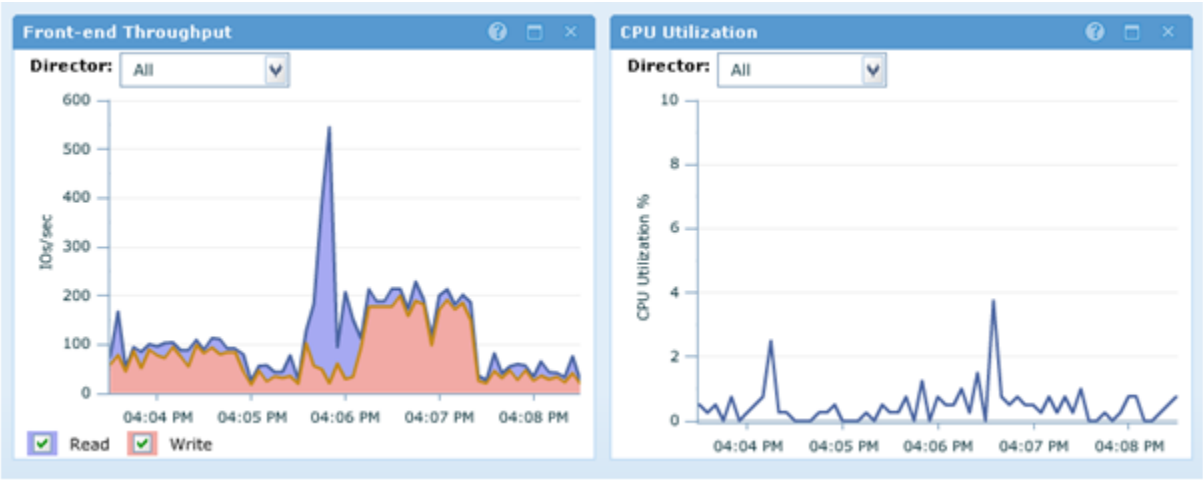

Figure 18. VPLEX performance monitoring dashboard

VMware virtualized infrastructure

Introduction

Overview

For the solution, the SAP application servers are fully virtualized using VMware vSphere 5. This section describes the virtualization infrastructure, which uses these components and options:

- VMware vSphere 5.0
- VMware vCenter™ Server
- VMware vSphere vMotion®
- VMware vSphere High Availability (HA)
- VMware vSphere Distributed Resource Scheduler™ (DRS)
- EMC PowerPath®/VE for VMware vSphere Version 5.7
- EMC Virtual Storage Integrator for VMware vSphere Version 5.1

VMware vSphere 5

VMware vSphere 5 is the industry's most complete, scalable, and powerful virtualization platform, with infrastructure services that transform IT hardware into a high-performance shared computing platform, and application services that help IT organizations deliver the highest levels of availability, security, and scalability.

VMware vCenter Server

VMware vCenter is the centralized management platform for vSphere environments, enabling control and visibility at every level of the virtual infrastructure.

VMware vSphere vMotion

VMware vSphere vMotion is VMware technology that supports live migration of virtual machines across servers with no disruption to users or loss of service.

Storage vMotion is VMware technology that enables live migration of a virtual machine's storage without any interruption in the availability of the virtual machine. This allows the relocation of live virtual machines to new datastores.

VMware vSphere High Availability

VMware vSphere High Availability (HA) is a vSphere component that provides high availability for any application running in a virtual machine, regardless of its operating system or underlying hardware configuration.

VMware vSphere Distributed Resource Scheduler

VMware vSphere Distributed Resource Scheduler (DRS) dynamically and automatically balances load distribution and virtual machine placement across multiple ESXi servers.

EMC PowerPath/VE

EMC PowerPath/VE for VMware vSphere delivers PowerPath multipathing features to optimize VMware vSphere virtual environments. PowerPath/VE installs as a kernel module on the ESXi host and works as a multipathing plug-in (MPP) that provides enhanced path management capabilities to ESXi hosts.

EMC Virtual Storage Integrator for VMware vSphere

EMC Virtual Storage Integrator (VSI) for VMware vSphere is a plug-in to the VMware vSphere client that provides a single management interface for managing EMC storage within the vSphere environment. VSI provides a unified and flexible user experience that allows each feature to be updated independently, and new features to be introduced rapidly in response to changing customer requirements.

When PowerPath/VE is installed on an ESXi host, VSI presents important multipathing details for devices, such as the load-balancing policy, the number of active paths, and the number of dead paths.

VMware deployments on VPLEX Metro

EMC VPLEX Metro delivers concurrent access to the same set of devices at two physically separate locations and thus provides the active/active infrastructure that enables geographically stretched clusters based on VMware vSphere. The use of Brocade Virtual Link Aggregation Group (vLAG) technology enables extension of VLANs, and hence subnets, across different physical data centers.

By deploying VMware vSphere features and components together with VPLEX Metro, the following functionality can be achieved:

- **vMotion**—The ability to live migrate virtual machines between sites in anticipation of planned events such as hardware maintenance.
- **Storage vMotion**—The ability to migrate a virtual machine's storage without any interruption in the availability of the virtual machine. This allows the relocation of live virtual machines to new datastores.
- **VMware DRS**—Automatic load distribution and virtual machine placement across sites through the use of DRS groups and affinity rules.
- **VMware HA**—A VPLEX Metro environment configured with VPLEX Witness is considered a VPLEX Metro HA configuration, as it ensures storage availability at the surviving site in the event of a site-level failure. Combining VPLEX Metro HA with a host failover clustering technology such as VMware HA provides automatic application restart for any site-level disaster. Figure 19 illustrates this HA architecture.

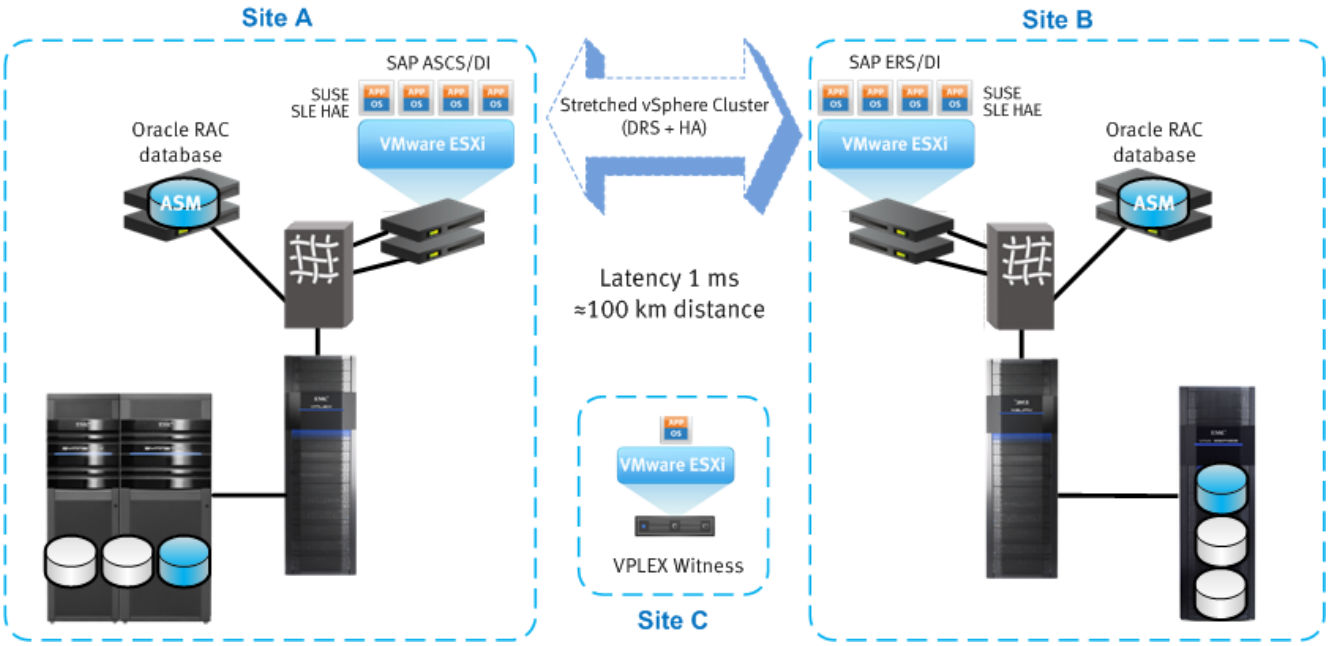

Figure 19. VMware HA with VPLEX Witness – logical view

- **VPLEX Metro HA Cross-Cluster Connect**—Protection of the VMware HA cluster can be further increased by adding a cross-cluster connect between the local VMware ESXi servers and the VPLEX cluster on the remote site.

  Local data unavailability events, which VMware vSphere 5.0 does not recognize, can occur when there is not a full site outage. Cross-connecting vSphere environments to VPLEX clusters protects against this and ensures that failed virtual machines automatically move to the surviving site.

  VPLEX Cross-Cluster Connect is available for up to 1 ms of distance-induced latency.

  This solution uses VPLEX Metro HA with Cross-Cluster Connect to maximize the availability of the VMware virtual machines, as shown in Figure 20.[1]

[1] For detailed information, see the EMC TechBook: *EMC VPLEX Metro Witness Technology and High Availability*.

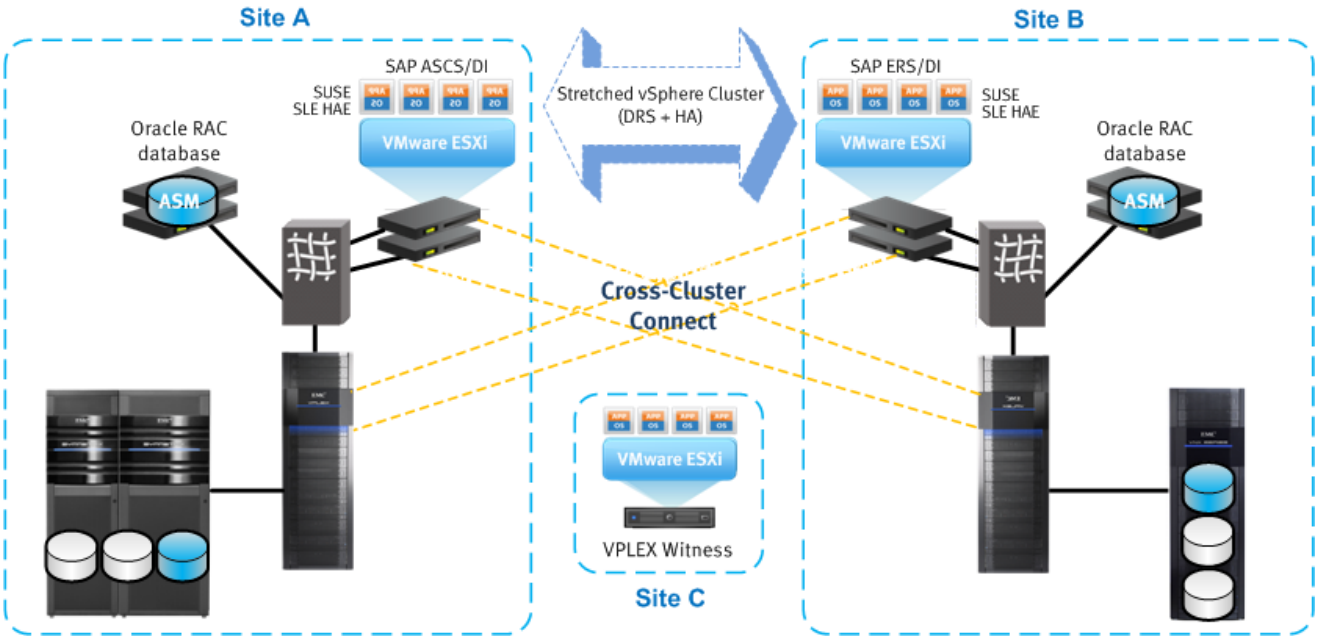

Figure 20. VMware HA with VPLEX Witness and Cross-Cluster Connect – logical view

VMware stretched cluster configuration

VMware and EMC support a stretched cluster configuration that includes ESXi hosts from multiple sites[2]. For the solution, a single vSphere cluster is stretched between Site A and Site B by using a distributed VPLEX virtual volume with VMware HA and VMware DRS. There are four hosts in the cluster, two at each site. VPLEX Metro HA Cross-Cluster Connect provides increased resilience to the configuration.

In vCenter, it is easy to view the configuration of this cluster—SiteAandSiteB—and the features enabled for it, as shown in Figure 21. This view also shows the memory, CPU, and storage resources available to the cluster.

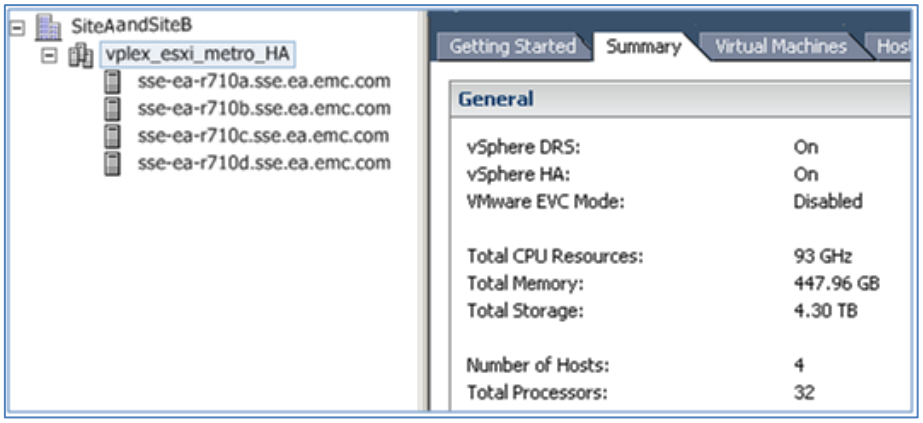

Figure 21. vSphere cluster with HA and DRS enabled

[2] For detailed requirements and scenarios, see the *VMware Knowledge Base article 1026692: Using VPLEX Metro with VMware HA*

Each ESXi server is configured with two 10 GbE physical adapters to provide network failover and high performance. A vSphere distributed switch (dvSwitch)[3] provides a single, common switch across all hosts. The 10 GbE physical adapters (also referred to as uplink adapters) are assigned to the dvSwitch.

Two distributed port groups are assigned to the dvSwitch:

- **dvPortGroupSiteAB**—for virtual machine network traffic
- **Management Network**—for VMkernel traffic and, in particular, vMotion traffic

Figure 22 shows the dvSwitch configuration. As both vSphere 5.0 distributed switches and Brocade VCS switches support Link Layer Discovery Protocol (LLDP), the properties of the associated physical switches can also be easily identified from vCenter.

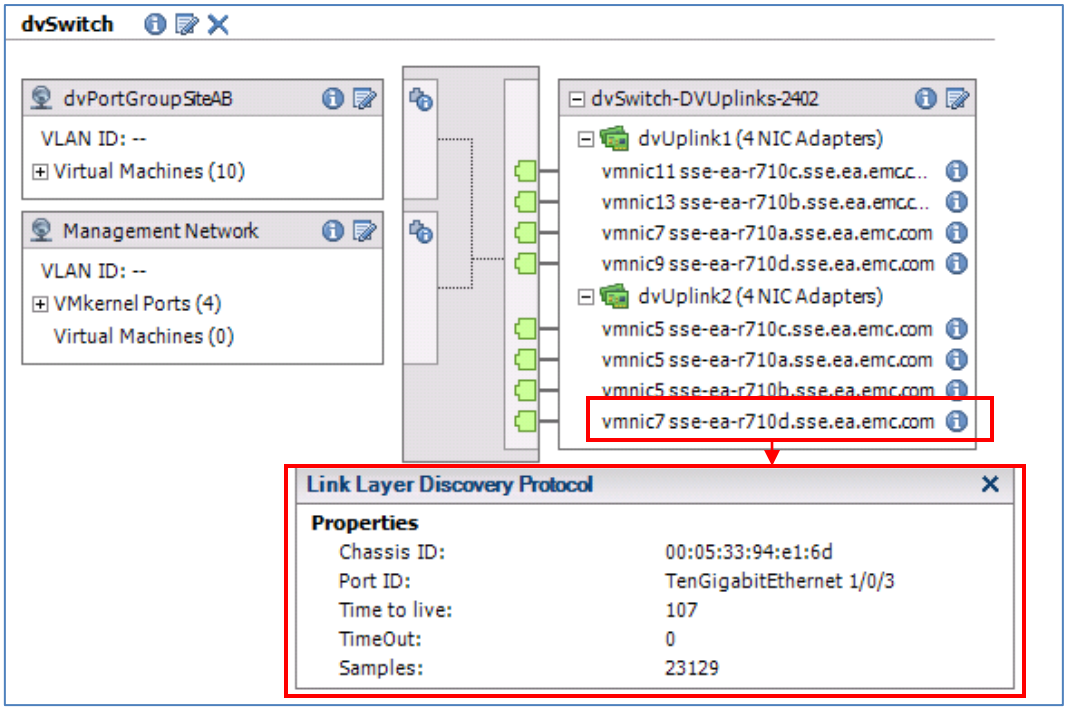

Figure 22. dvSwitch configuration and LLDP detail

Datastore EXT_SAP_VPLEX_DS01 was created on a 1 TB VPLEX distributed volume and presented to the ESXi hosts in the stretch cluster. All virtual machines were migrated to this datastore, using Storage vMotion, because they need to share virtual disks or to be able to vMotion between sites. Figure 23 shows the configuration details for the datastore.

[3] A dvSwitch provides a network configuration that spans all member hosts and allows virtual machines to maintain consistent network configuration as they migrate between hosts. For further information, see the *VMware vSphere Networking ESXi 5.0* document.

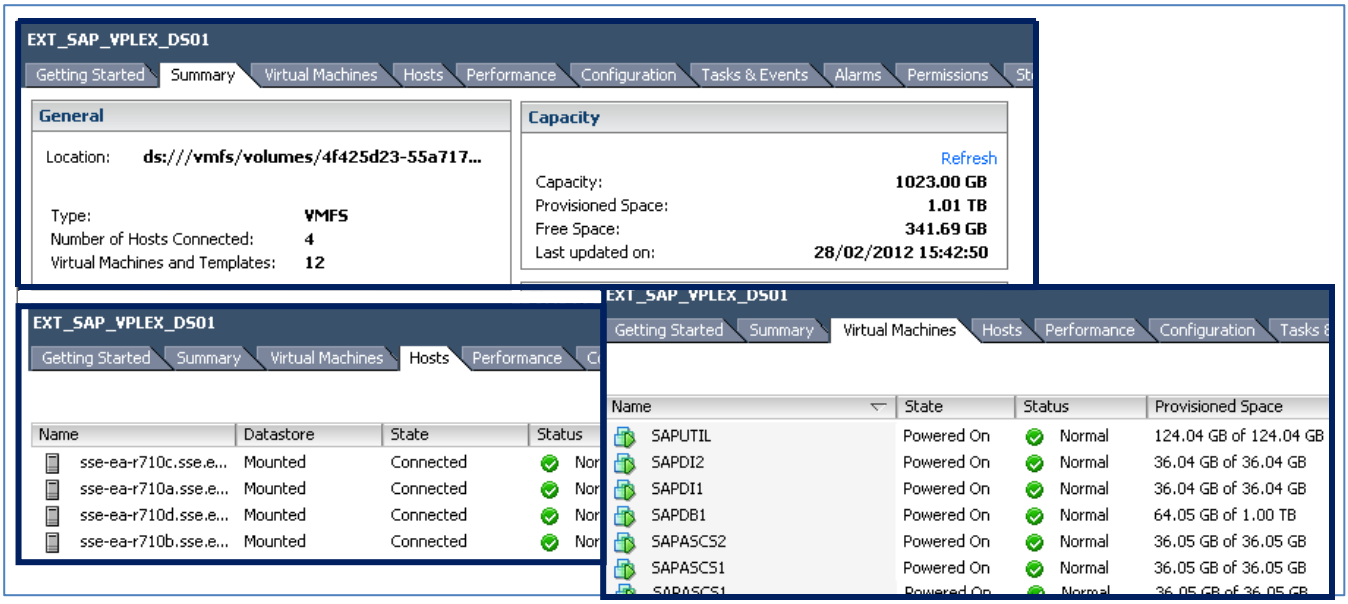

Figure 23. Datastore EXT_SAP_VPLEX_DS01 and associated hosts and virtual machines

VMware vSphere HA configuration

Enabling VMware vSphere HA and VMware vSphere DRS

vSphere HA leverages multiple ESXi hosts, configured as a cluster, to provide rapid recovery from outages and cost-effective high availability for applications running in virtual machines.[4] vSphere HA protects application availability in the following ways:

- It protects against a server failure by restarting the virtual machines on other ESXi servers within the cluster.
- It protects against application failure by continuously monitoring a virtual machine and resetting it in the event of guest OS failure.

For the solution, both vSphere HA and DRS were enabled, as shown in Figure 24.

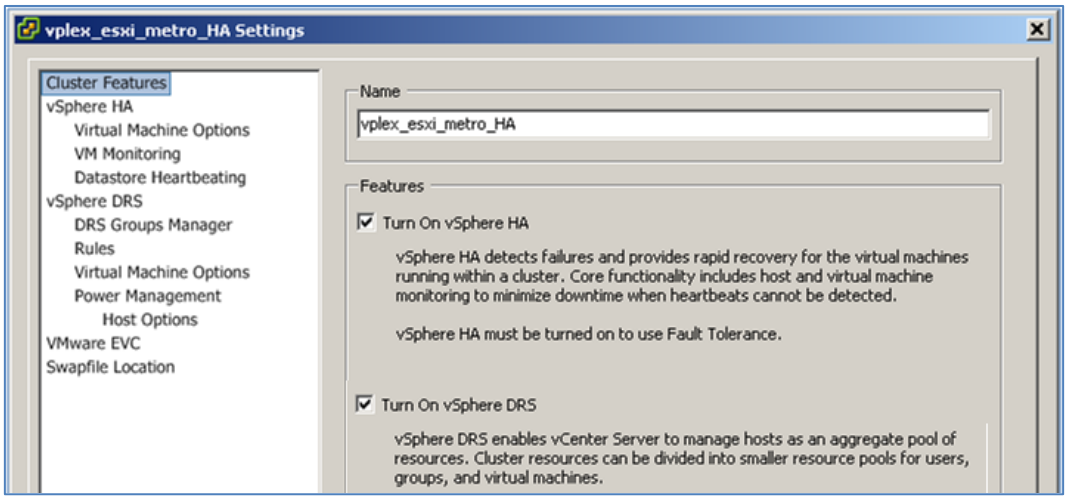

Figure 24. vSphere HA wizard

[4] For further information on vSphere HA, see the *VMware vSphere Availability ESXi 5.0* document.

VM Monitoring

VM Monitoring was configured to restart individual virtual machines if their heartbeat is not received within 60 seconds.

Virtual machine restart options

The VM Restart Priority option for the four SAP virtual machines was set to High. This ensures that these virtual machines are powered on first in the event of an outage. Figure 25 shows this setting and the Host Isolation Response setting (default).

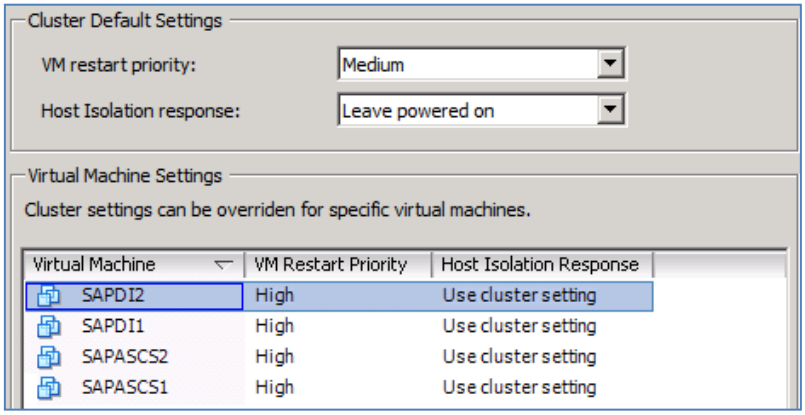

Figure 25. VM Restart Priority and Host Isolation Response settings

Datastore heartbeating

When you create a vSphere HA cluster, a single host is automatically elected as the master host. The master host monitors the state of all protected virtual machines and of the slave hosts. When the master host cannot communicate with a slave host, it uses datastore heartbeating to determine whether the slave host has failed, is in a network partition, or is network isolated.

To meet vSphere HA requirements for datastore heartbeating, a second datastore—EXT_SAP_VPLEX_HA_HB—was created on a 20 GB VPLEX distributed volume and presented to all the ESXi hosts, as shown in Figure 26. In a production environment, vCenter automatically selects two or more datastores for this purpose, based on host visibility.

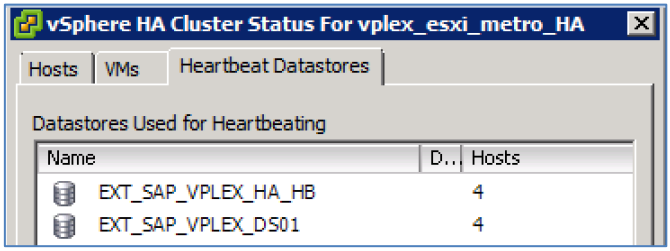

Figure 26. vSphere HA Cluster Status – heartbeat datastores

VMware vSphere DRS configuration

VMware DRS host groups and virtual machine groups

DRS host groups and virtual machine groups simplify management of ESXi host resources. These features were not required for this solution.

VMware DRS affinity rules

DRS uses affinity rules to control the placement of virtual machines on hosts within a cluster. DRS provides two types of affinity rule:

- A VM-Host affinity rule specifies an affinity relationship between a group of virtual machines and a group of hosts.
- A VM-VM affinity rule specifies whether particular virtual machines should run on the same host or be kept on separate hosts.

Table 6 and Figure 27 show the VM-VM affinity rule that the solution uses.

Table 6. VMware DRS affinity rule

VM-VM affinity rule	
SAPASCS - Separate	Keep virtual machines **SAPASCS1** and **SAPASCS2** on separate hosts.

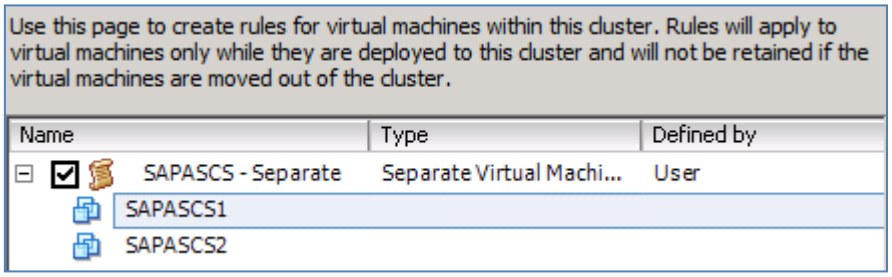

Figure 27. DRS VM-VM affinity rule for vplex_esxi_metro_HA cluster

EMC Virtual Storage Integrator and VPLEX

EMC Virtual Storage Integrator (VSI) provides enhanced visibility into VPLEX directly from the vCenter GUI. The Storage Viewer and Path Management features are accessible through the EMC VSI tab, as shown in Figure 28.

In the solution, VPLEX distributed volumes host the EXT_SAP_VPLEX_DS01 Virtual Machine File System (VMFS) datastore, and Storage Viewer provides details of the datastore's virtual volumes, storage volumes, and paths.

As shown in Figure 28, the LUNS that make up the datastore are four 256 GB distributed RAID 1 VPLEX Metro volumes that are accessible via PowerPath.

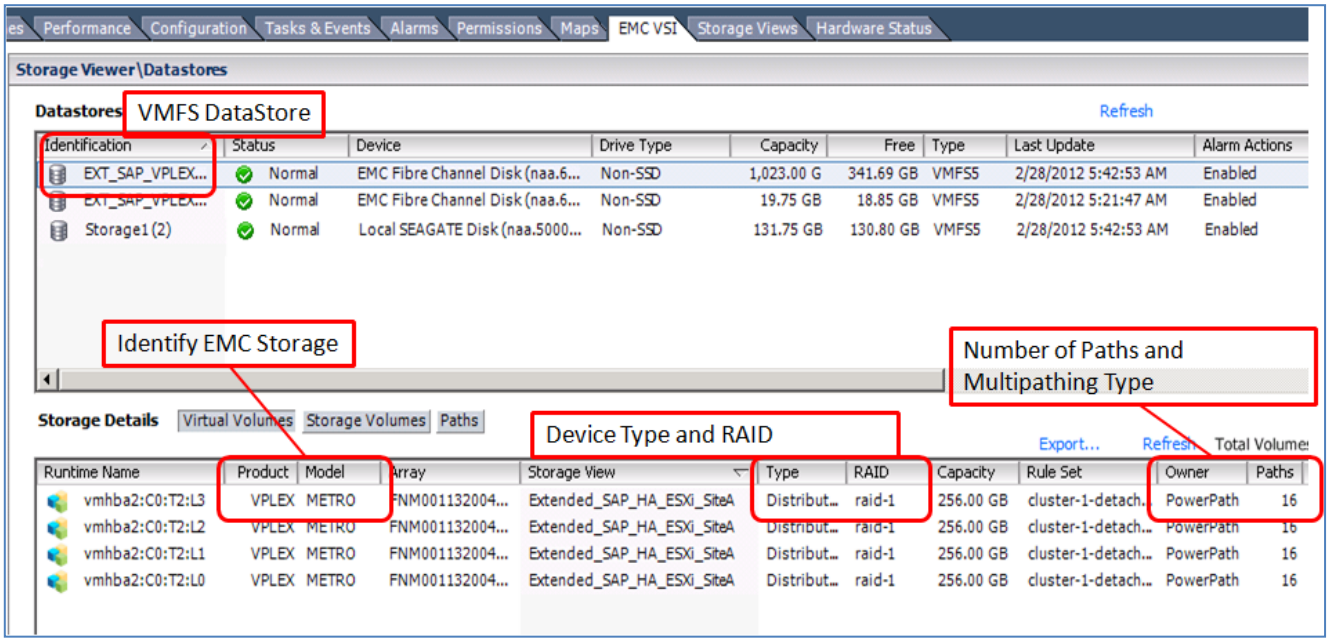

Figure 28. VSI Storage Viewer – datastores

SAP system architecture

Introduction

Overview

This section describes the SAP system architecture deployed for the solution in the two data centers. The SAP application layer uses these SAP and SUSE components:

SAP application

- SAP Enhancement Package 4 for SAP ERP 6.0 IDES
- SAP NetWeaver Application Server for ABAP 7.01
- SAP Enqueue Replication Server

Operating system

- SUSE Linux Enterprise Server for SAP Applications 11 SP1
- SUSE Linux Enterprise High Availability Extension

The SAP system runs in a hybrid environment, with SAP services on virtual machines and the database on physical servers. All SAP instances are installed on VMware vSphere virtual machines with SUSE Linux Enterprise Server for SAP Applications as the operating system. The underlying database is a physical Oracle RAC database on ASM. The VMware and Oracle environments are described in separate sections of the white paper (see VMware virtualized infrastructure and Oracle database).

SAP ERP 6.0

SAP ERP 6.0, powered by the SAP NetWeaver technology platform, is a world-class, fully-integrated enterprise resource planning (ERP) application that fulfills the core business needs of midsize companies and large enterprises across all industries and market sectors. SAP ERP 6.0 delivers a comprehensive set of integrated, cross-functional business processes and can serve as a solid business process platform that supports continued growth, innovation, and operational excellence.

SAP IDES (Internet Demonstration and Evaluation System) supports demos, testing, and functional evaluation based on preconfigured data and clients. IDES contains application data for various business scenarios, with business processes in that are designed to reflect real-life business requirements and have access to many realistic characteristics. This solution uses IDES to represent a model company for testing purposes.

SUSE Linux Enterprise Server for SAP Applications

SUSE Linux Enterprise Server is a highly reliable, scalable, and secure server operating system that is built to power physical, virtual, and cloud applications. It is a preferred Linux platform for SAP.

SUSE Linux Enterprise Server for SAP Applications, based on the newest SUSE Linux Enterprise Server technology, is optimized for all mission-critical SAP NetWeaver software solutions and appliances. SAP and SUSE validate and certify SUSE Linux Enterprise Server for SAP Applications jointly to eliminate potential software incompatibilities. This partnership tightly integrates the application workload with the operating system and eliminates the possibility of incompatibilities when patches are applied to either the applications or the operating system.

SUSE Linux Enterprise High Availability Extension

SUSE Linux Enterprise Server for SAP Applications includes SUSE Linux Enterprise High Availability Extension, which offers high-availability service and application clustering, file systems and clustered file systems, network-attached storage (NAS), network file systems, volume managers, SAN and drivers, and the means to manage of all these components working together. SUSE Linux Enterprise High Availability Extension provides an integrated clustering solution for physical and virtual Linux deployments, enabling the implementation of highly available Linux clusters and eliminating single points of failure.

SAP system configuration

SAP system architecture

The solution implements a high-availability SAP system architecture, as shown in Figure 29.

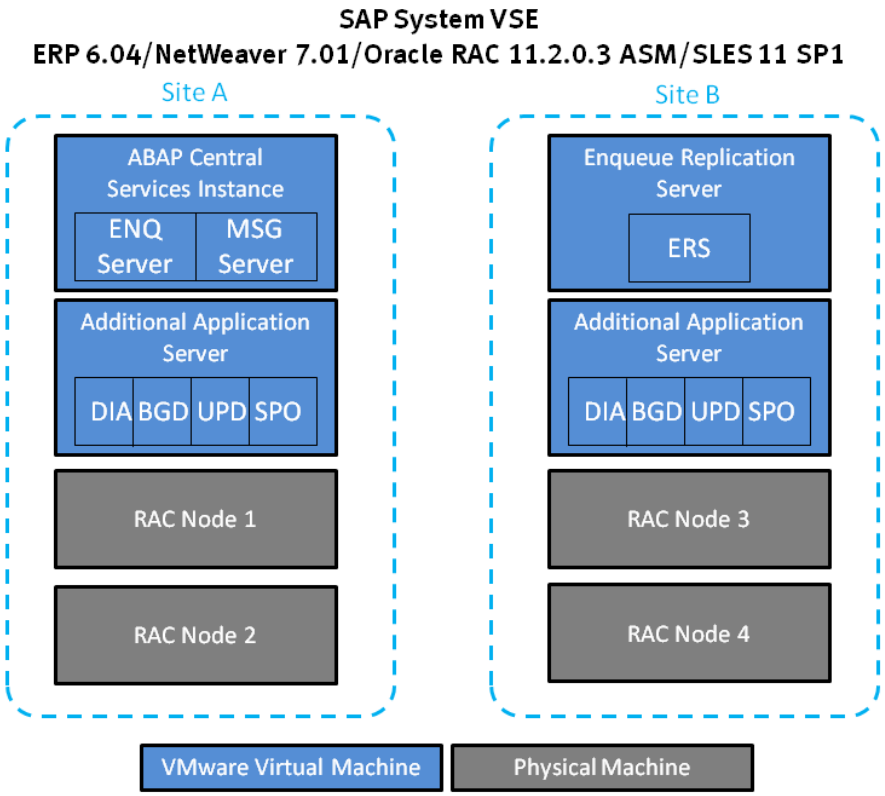

Figure 29. SAP system architecture

The enqueue server and message server are decoupled from the Central Instance and implemented as services within the ASCS instance[5]. SAP ERS is installed as part of the HA architecture to provide zero application lock loss and further protect the

[5] The enqueue server manages logical locks, its objective being to minimize the duration of a database lock. Unlike database locks, an SAP lock can exist across several database LUWs. The message server informs all servers (instances) in an SAP system of the existence of the other servers. Other clients (for example, SAPlogon and RFC clients with load balancing) can also contact it for information about load balancing.

enqueue server[6]. Two dialog instances are installed to provide redundant work processes such as dialog (DIA), background (BGD), update (UPD), spool (SPO), and gateway.

Key design considerations

The SAP system deployed for the solution implements these key design features:

- The ASCS instance is installed with a virtual hostname (SAPVIPE), to decouple it from the virtual machine hostname.
- The ERS instance is installed with a different instance number (01), to avoid future confusion when both ASCS and ERS are under cluster control.
- SAP patches, parameters, basis settings, and load balancing settings are all installed and configured according to the SAP installation guide and the SAP Notes listed on page 73.
- VMware best practices for SAP are adopted in this solution[7].
- SAP update processes (UPD/UP2) are configured on the additional application server instances.
- SAP ASCS instance profile, ERS instance and start profiles, and dialog instance profiles are updated with ERS configurations. See Appendix – Sample configurations for sample configurations.
- SAP shared file systems, including /sapmnt/‹SID› (available to all SAP instances) and /usr/sap/‹SID›/ASCS00 (available to SAP cluster nodes, ASCS instance, and ERS instance), are stored on Oracle ASM Cluster File System (ACFS) and mounted as Network File System (NFS) shares on the SAP virtual machines. These shared file systems are presented as a highly available NFS resource that is managed by Oracle Clusterware.
- Some IDES functionality—for example, synchronization with the external GTS system—is deactivated to eliminate unnecessary external interfaces that are outside the scope of the solution.
- The storage for the entire SAP environment is encapsulated and virtualized for this solution. The storage is distributed across the two sites and made available to the SAP servers through VPLEX Metro.

SUSE Linux Enterprise High Availability Extension configuration

SAP virtual machine architecture with SUSE Linux Enterprise High Availability Extension

The solution uses SUSE Linux Enterprise High Availability Extension to protect the central services (message server and enqueue server) across two cluster nodes built on VMware virtual machines. VMware High Availability (VMHA) protects the virtual machines. Figure 30 shows this architecture.

---

[6] SAP ERS provides a replication mechanism for the enqueue server by holding a copy of the locking table within its shared memory segment. ERS installation for Linux is not part of the standard SAPInst process. For installation instructions, see the SAP Enqueue Replication Server help portal on help.sap.com.

[7] For full details, see: *SAP Solutions on VMware: Best Practices Guide.*

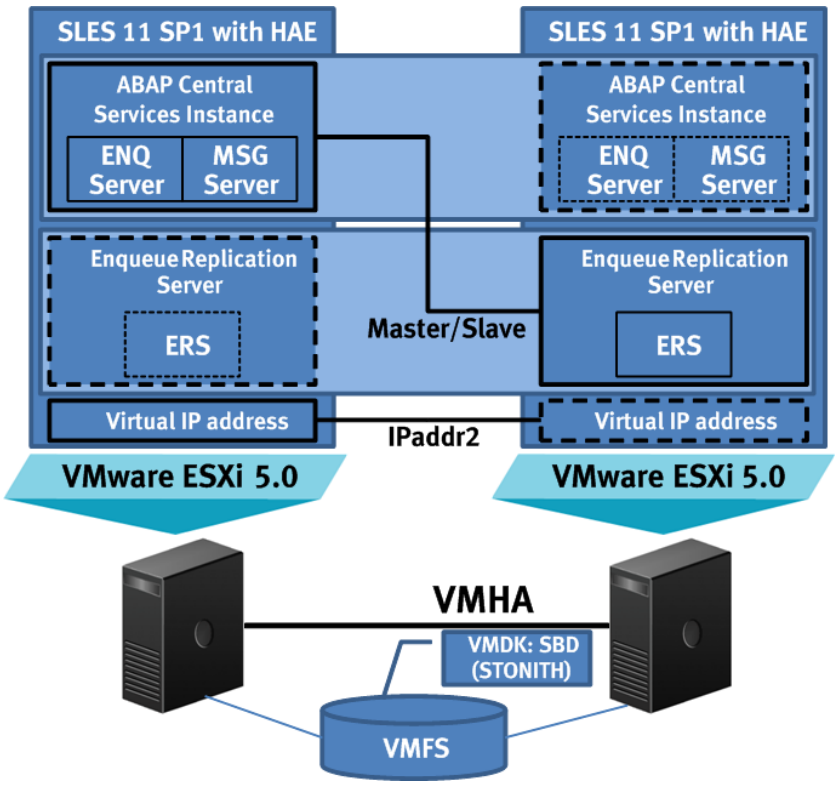

Figure 30. SAP ASCS cluster architecture with SUSE Linux Enterprise HAE

The key components of SUSE Linux Enterprise High Availability Extension implemented in this solution include:

- OpenAIS[8]/Corosync[9]—a high-availability cluster manager that supports multinode failover
- Resource agents (virtual IP address, master/slave, and SAPInstance) to monitor and control the availability of resources
- High-availability GUI and various command line tools

Table 7 shows the configuration of the SAP virtual machines.

Table 7. SAP virtual machines

VM role	Quantity	vCPUs	Memory (GB)	OS bootdisk (GB)	VM name
SAP ASCS	1	2	4	32	SAPASCS1
SAP ERS	1	2	4	32	SAPASCS2
SAP AAS	2	2	4	32	SAPDI1
		2	4	32	SAPDI2

[8] OpenAIS is an open implementation of the Application Interface Specification (AIS) provided by the Service Availability Forum (SA Forum).

[9] The Corosync Cluster Engine is a group communication system with additional features for implementing high availability within applications.

Installation and configuration process

The SUSE white paper *Running SAP NetWeaver on SUSE Linux Enterprise Server with High Availability – Simple Stack* describes how to install and configure the SUSE software and SAP NetWeaver.

Appendix – Sample configurations provides a sample configuration file that supports the features and functionality validated by this solution. You should consider the time values (timeout, intervals, and so on) here as "initial" values to be fine tuned and optimized for your particular environment.

For the solution, SUSE Linux Enterprise High Availability Extension was installed and configured using YaST and Pacemaker GUI. Here is a summary of the installation and configuration process:

1. Set up an internal SMT Server (for security purposes) to update all software packages to the latest versions.
2. In the YaST Software Management module, select **Patterns › High Availability** to install the High Availability Extension, as shown in Figure 31.

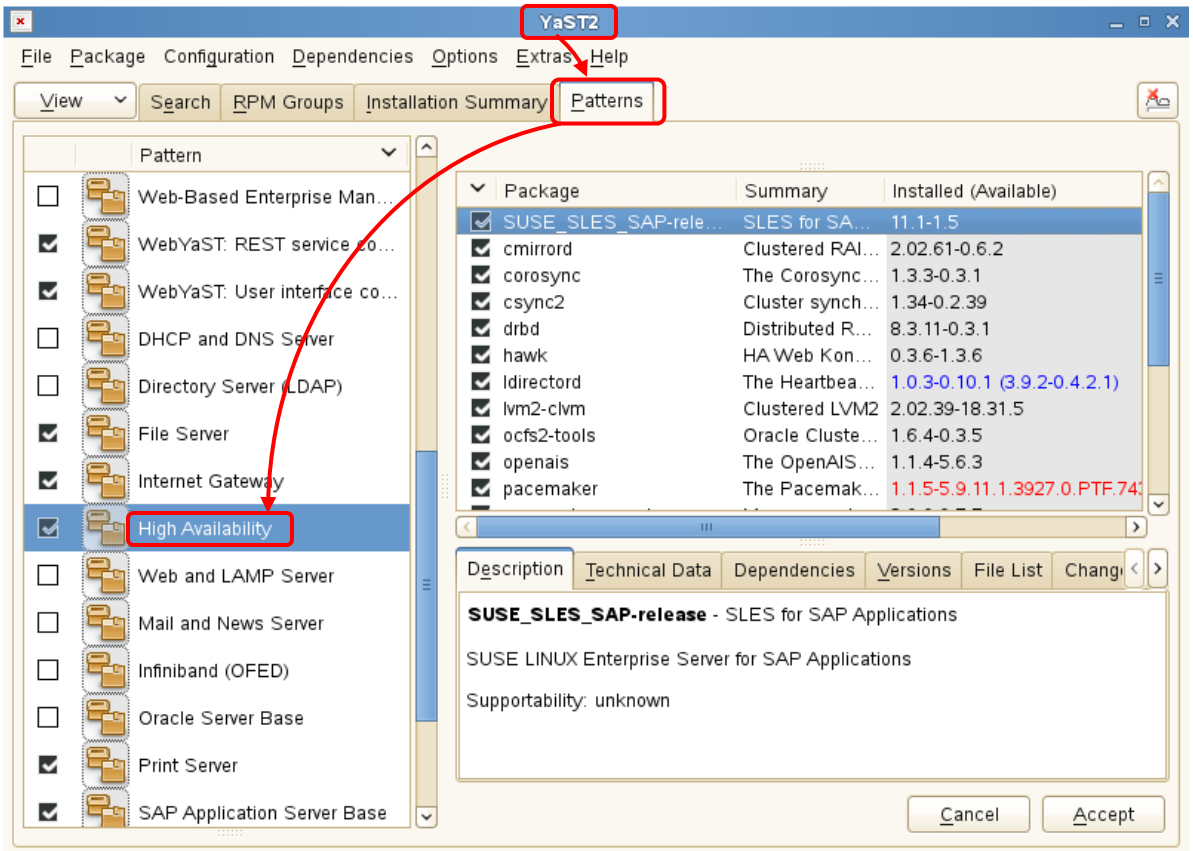

Figure 31. Installing SUSE Linux Enterprise High Availability Extension

3. In the YaST Cluster module, configure the cluster basic settings, as shown in Figure 32.

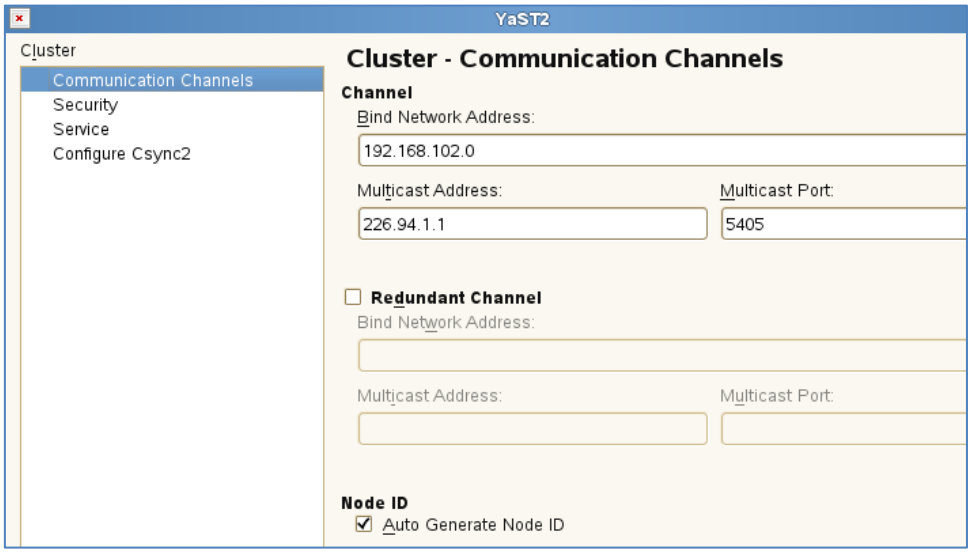

Figure 32. Configuring cluster basic settings

4. In Pacemaker GUI, configure the global cluster settings, as shown in Figure 33.

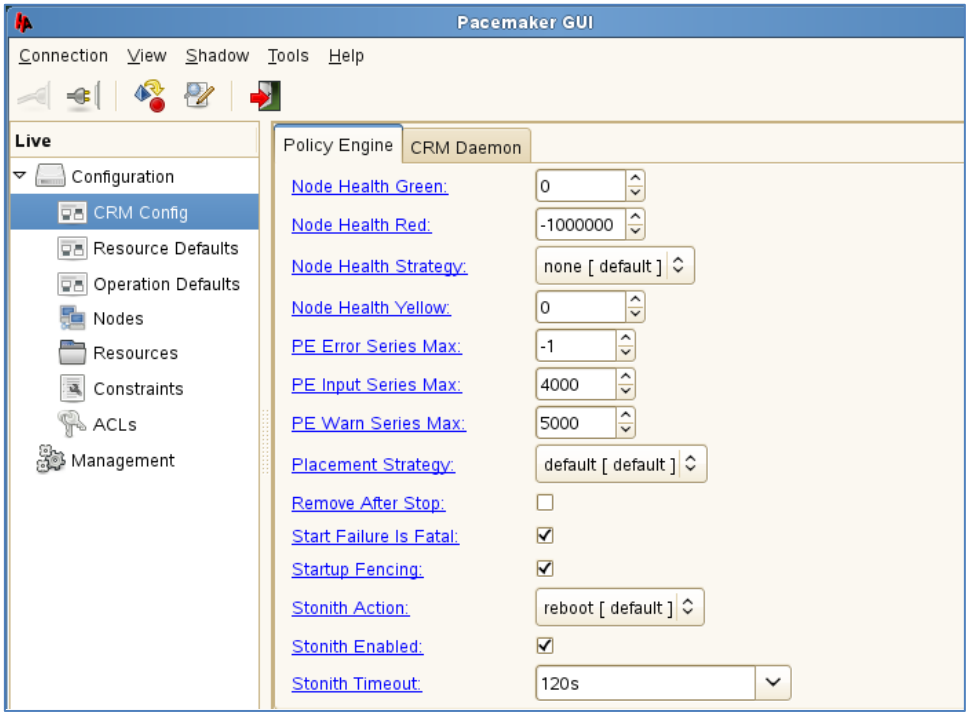

Figure 33. Configuring global cluster settings

5. In Pacemaker GUI, open the **Resources** category and configure IPaddr2, master/slave, and SAPInstance resources, as shown in Figure 34.

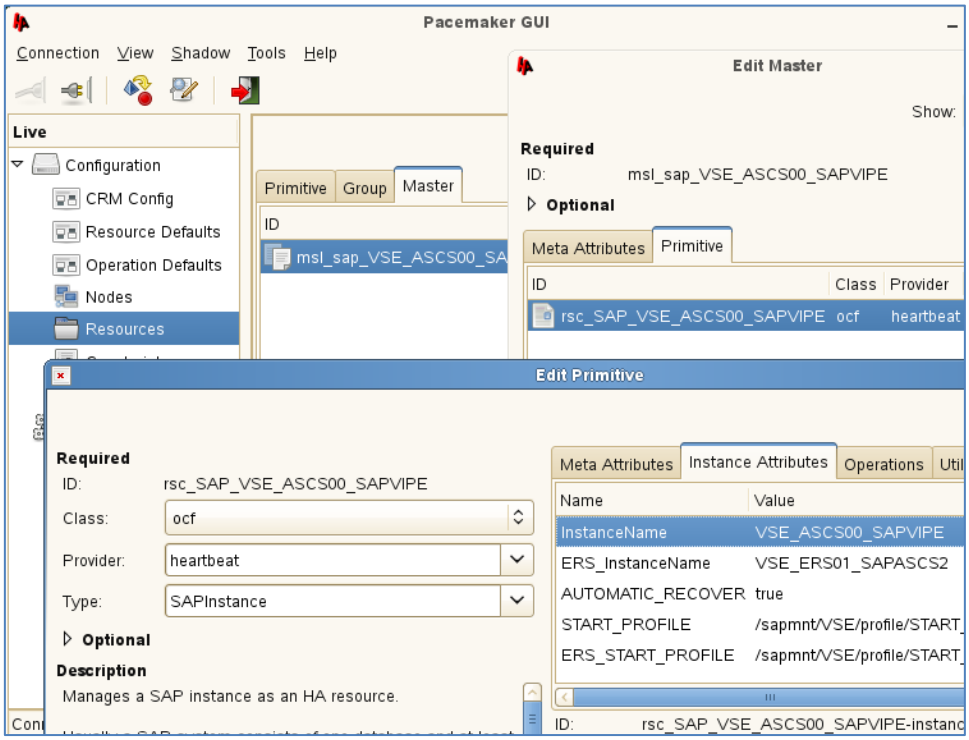

Figure 34. Configuring resources

6. In Pacemaker GUI, configure the dependencies of the resources, as shown in Figure 35.

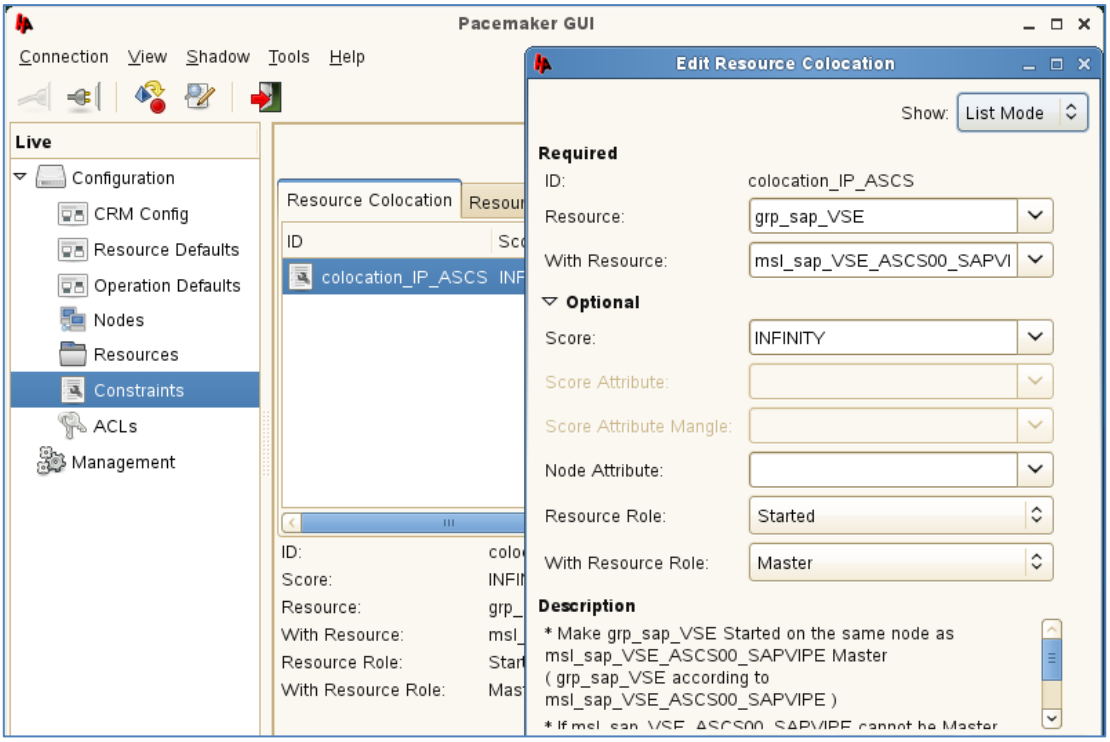

Figure 35. Configuring resource dependencies

7. In Pacemaker GUI, start the cluster and check that the cluster and all resource agents are operating normally, as shown in Figure 36.

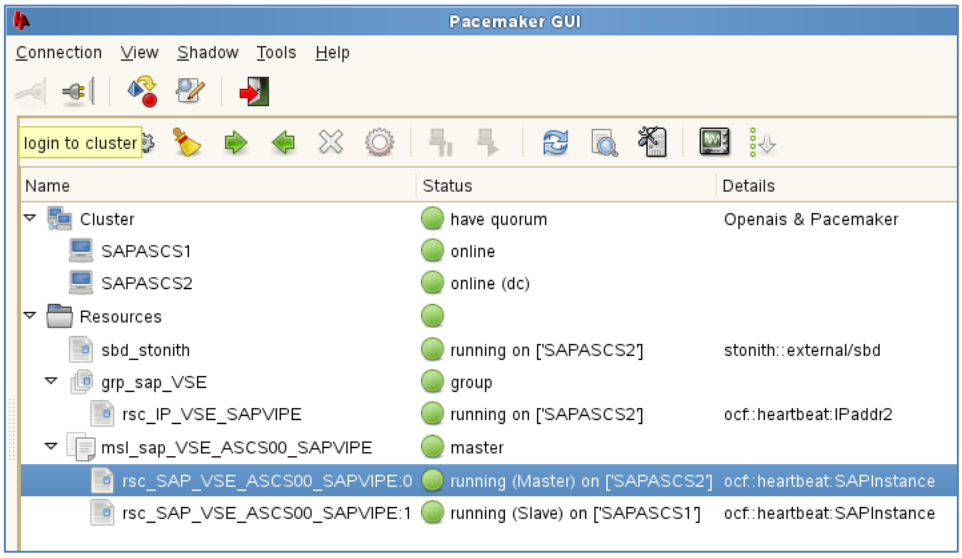

Figure 36. Checking the cluster status

Key design considerations

STONITH device configuration

SBD (STONITH block device) and STONITH (Shoot The Other Node In The Head) enable fencing (isolating nodes) in a cluster via shared storage. This solution uses a partition of a virtual disk (VMDK) as an SBD STONITH device.[10] Therefore, both cluster nodes need simultaneous access to this virtual disk. The virtual disk is stored in the same datastore as the SAP virtual machines. This is provisioned and protected by VPLEX and is available on both sites.

By default, VMFS prevents multiple virtual machines from accessing and writing to the same VMDK. However, you can enable sharing by configuring the multi-writer option[11], as shown in Figure 37.

[10] SBD is essential for handling split-brain scenarios in the cluster. A single SBD device is configured for this solution. This single SBD device configuration is for testing purposes only; for production configuration, see *Running SAP NetWeaver on SUSE Linux Enterprise Server with High Availability – Simple Stack*.

[11] For detailed information, see *VMware Knowledge Base article 1034165: Disabling simultaneous write protection provided by VMFS using the multi-writer flag*.

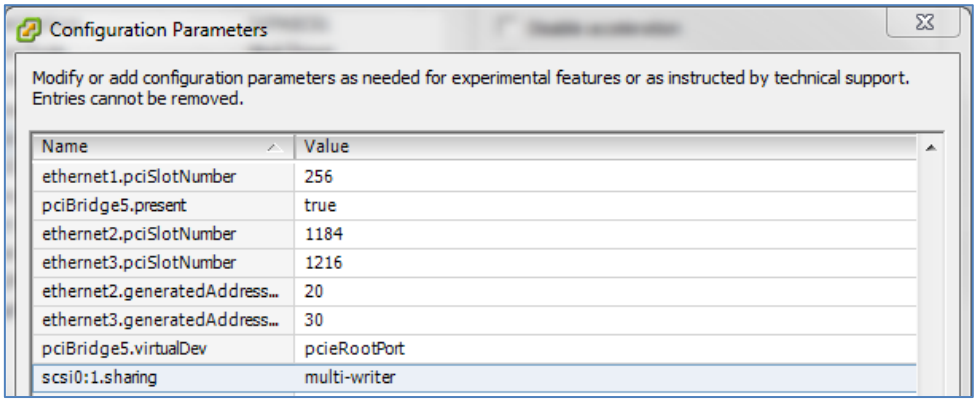

Figure 37. Multi-writer option

Master/slave configuration

The SAPInstance resource agent controls the ASCS instance as well as the appropriate ERS instance. It is configured as master/slave resource that extends the roles of the resource from **started** and **stopped** to **master** and **slave**. A promoted master instance starts the SAP ASCS instance. The demoted slave instance starts the ERS instance. The master/slave mode ensures that an ASCS instance is never started on the same node as the ERS.

Figure 38 shows the configuration of the SAPInstance resource agent.

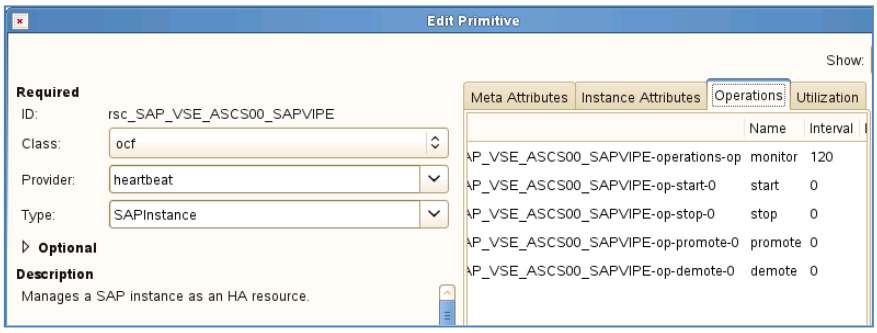

Figure 38. SAPInstance resource agent configuration

Figure 39 shows the configuration of the master/slave resource agent.

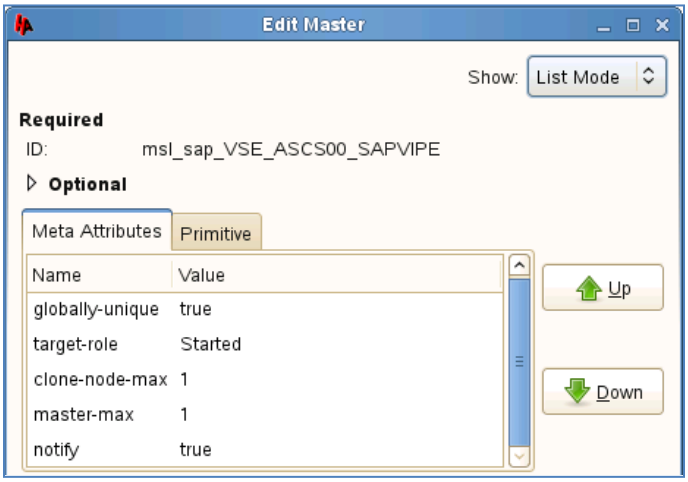

Figure 39. Master/slave resource agent configuration

Resource constraints

The ASCS instance and its virtual IP are bound together using appropriate order and colocation constraints. Figure 40 shows the configuration of the Resource Colocation and Resource Order constraints.

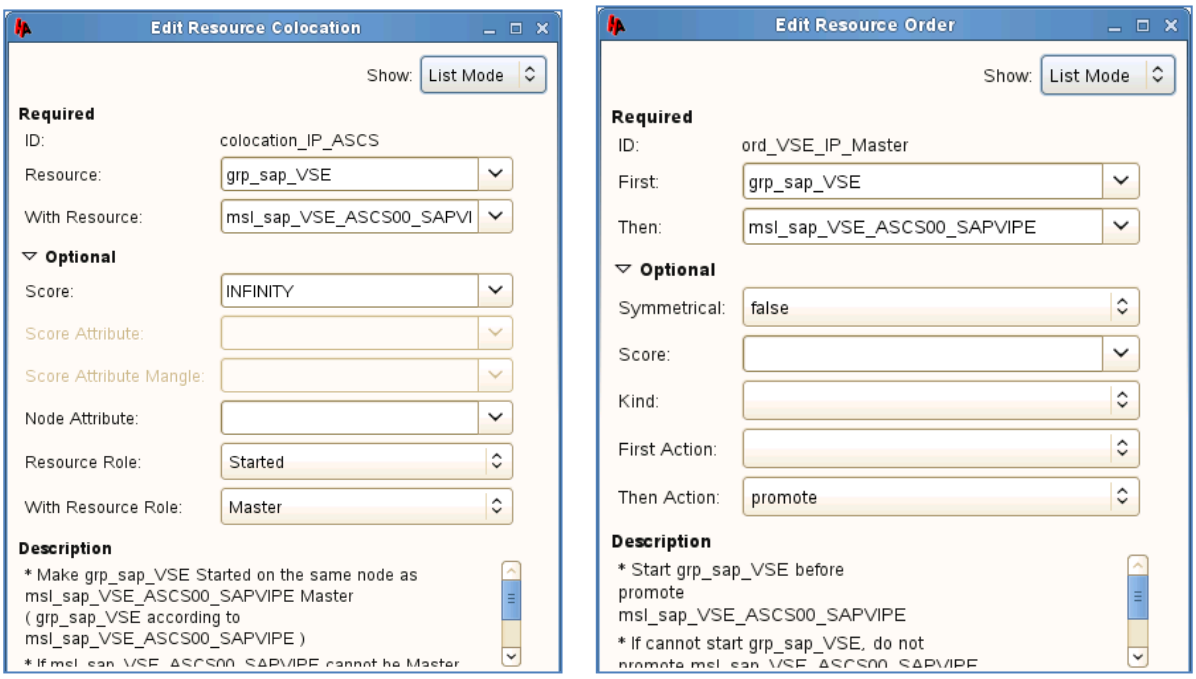

Figure 40. Resource Colocation and Resource Order constraint configuration

Corosync token parameter configuration

In the Corosync configuration file—corosync.conf—the **token** timeout specifies the time (in milliseconds) after which a token loss is declared if a token is not received. This timeout corresponds to the time spent detecting the failure of a processor in the current configuration. For this solution, the value of this parameter is set to 10,000 ms in order to cope with the switchover of the underlying layers without unnecessary cluster service failover.

Polling concept

SUSE Linux Enterprise High Availability Extension can continuously monitor the status of SAP processes on each cluster node and make the correct decisions to promote/demote the ASCS instance and ERS instance respectively.

There is no need to implement SAP polling concept. Ensure that this feature is NOT enabled in the ERS instance profile. For a sample ERS instance profile, see ERS sample instance profile on page 76.

Oracle database architecture

Introduction

Overview

This section describes the grid and database that underlies the SAP applications in the solution. At each data center, the database originated as a physical Oracle Database 11*g* single instance. To eliminate the database server as a single point of failure, the single instance database was migrated to a four-node physical Oracle RAC 11*g* cluster with the Oracle database residing on ASM.

The solution uses these Oracle components and options:

- Oracle Database 11*g* Release 2 Enterprise Edition
- Oracle Automatic Storage Management (ASM) and Oracle ASM Cluster File System (ACFS)
- Oracle Clusterware
- Oracle Real Applications Clusters (RAC) 11*g* on Extended Distance Clusters

Oracle Database 11*g* R2

Oracle Database 11*g* 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 and Oracle ACFS

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 11*g* R2, Oracle ASM has also been extended to include support for OCR and voting files to be placed within ASM disk groups.

Oracle ACFS, a feature within ASM in Oracle Database *11g,* extends ASM functionality to act as a general-purpose cluster file system. Oracle database binaries can reside on ACFS, as can supporting files such as trace and alert logs, and non-Oracle application files such as SAP ERP. Non-Oracle servers can access ACFS volumes using industry-standard NAS protocols such as NFS and Common Internet File System (CIFS).

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 11*g* R2, 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 11*g*

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.
- Third-party applications placed under Oracle Clusterware control can load balance and fail over to the Oracle RAC nodes on the surviving site—for example, NFS or Apache httpd.
- 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)[12].

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 configuration[13].

Oracle ACFS configuration

This solution uses four ACFS volumes mounted across the Oracle RAC cluster, as shown in Table 8. Three of these volumes, SAPMNT, USRSAPTRANS, and ASCS00, were then exported as NFS shares to the SAP servers, using a virtual IP address and a highly available NFS resource under control of Oracle Clusterware.

---

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

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

Table 8. Oracle ACFS volumes and mount points

ACFS volume	Size (GB)	Mount point	Description
SAP_O_HOME	16	/oracle/VSE/112	ORACLE_HOME for database VSE – shared on all Oracle RAC nodes
SAPMNT	16	/sapmnt/VSE	SAP global directory, which stores kernels, profiles etc. – shared on all SAP virtual machines
USRSAPTRANS	16	/usr/sap/trans	SAP transport directory, which stores the transport files – shared on all SAP Dialog Instance virtual machines
ASCS00	16	/usr/sap/VSE/ASCS00	SAP ASCS instance directory, which stores the instance-related files – shared on SUSE Linux Enterprise High Availability Extension cluster nodes

Oracle Extended RAC on VPLEX Metro

Figure 41 provides a logical representation of the deployment of Oracle Extended RAC on VPLEX Metro for the solution.

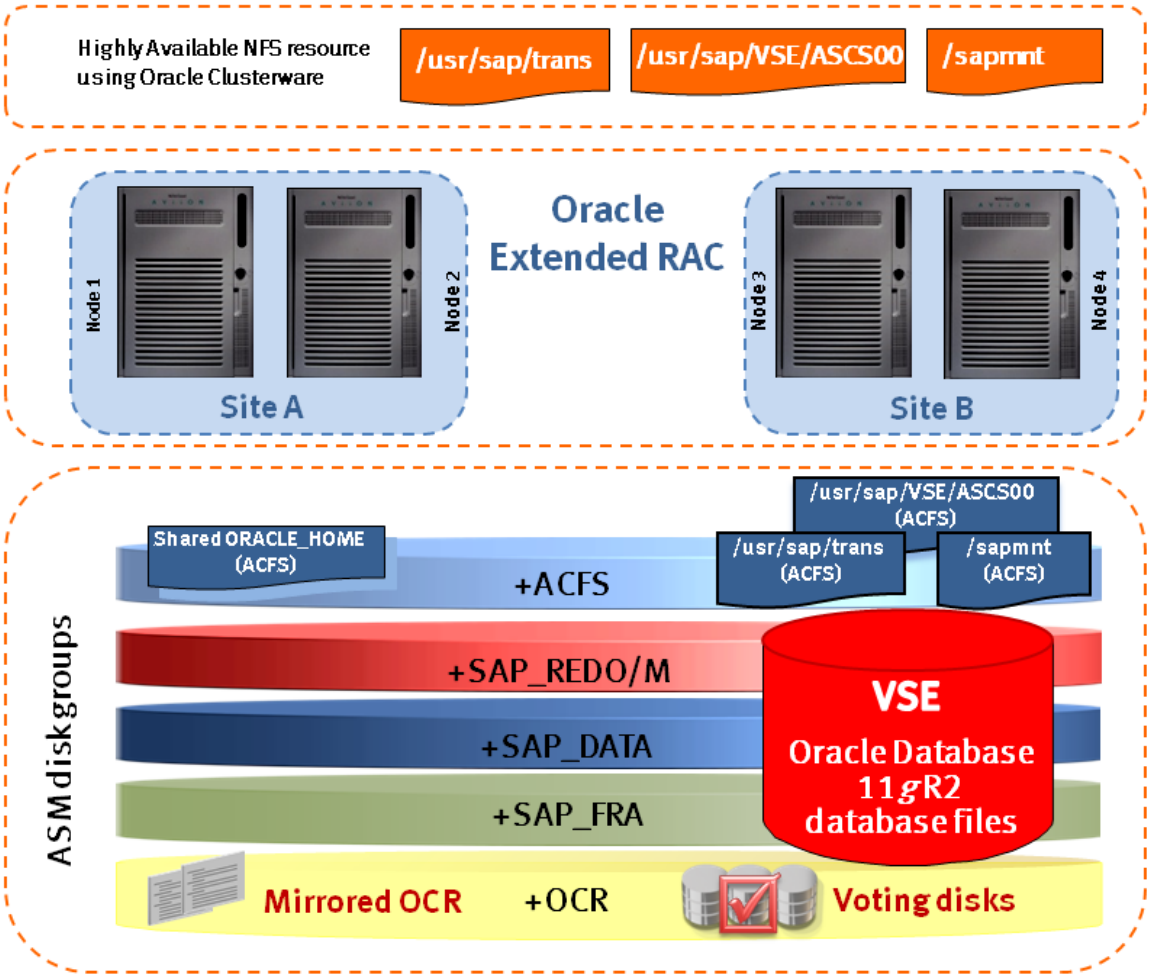

Figure 41. Oracle Extended RAC over EMC VPLEX Metro

Oracle ASM disk group configuration

The storage ASM disk groups were configured to reflect the existing single-instance Oracle database layout. Table 9 shows the ASM disk group layout and configuration.

Table 9. Oracle ASM disk group size and configuration

ASM disk group*	No of disks	Disk group size (GB)	Redundancy
OCR	5	40	Normal
EA_SAP_ACFS	4	64	External
EA_SAP_DATA	16	2,048	External
EA_SAP_REDO	4	64	External
EA_SAP_REDOM	4	64	External
EA_SAP_FRA	4	256	External

* The EA_SAP_ prefix is used to uniquely identify the ASM disk groups related to the SAP application in Extended Oracle RAC.

Oracle database migration process

For the solution, the following two-step process was used to migrate the original single instance database to an Oracle RAC cluster on ASM (see Figure 42):

1. Migrate the database from file system to ASM, using the Oracle Recovery Manager (RMAN) "duplicate from active database" method.
2. Convert the database to Oracle RAC, using the Oracle rconfig utility.

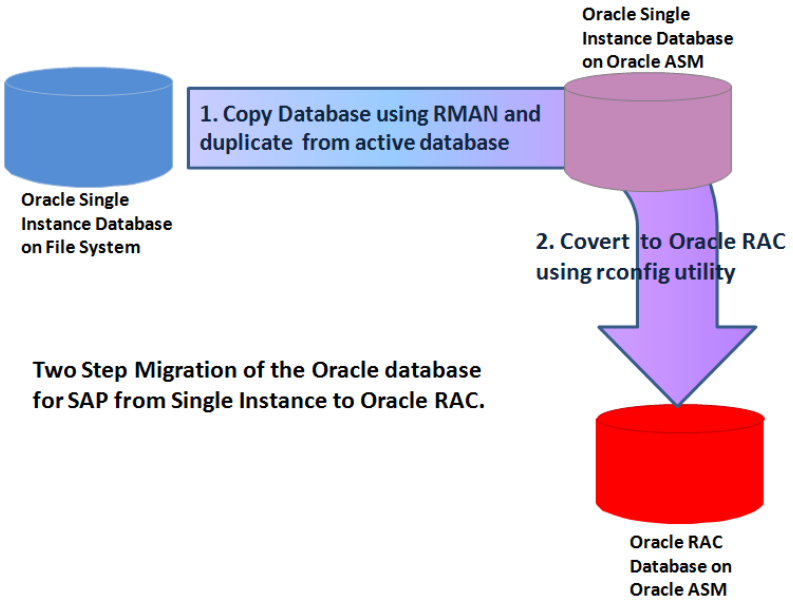

Figure 42. Two-step migration process for Oracle database

This two-step process follows the guidelines in these Oracle white papers:

- *Moving your SAP Database to Oracle Automatic Storage Management 11g Release 2: A Best Practices Guide*
- *Configuration of SAP NetWeaver for Oracle Grid Infrastructure 11.2.0.2 and Oracle Real Application Clusters 11g Release 2: A Best Practices Guide*

Preparing the source and target systems for duplicating the database

Table 10 outlines how to prepare the source and target systems before duplicating the database.

Table 10. Steps for preparing the source and target systems

Preparation—source system	
1	Ensure that key environment variables are set: ORACLE_SID, ORACLE_BASE and ORACLE_HOME.
2	Ensure that a tnsnames.ora entry is configured for the source and target/auxiliary databases for use during duplication.
3	Ensure that the Oracle password file is configured.
4	Ensure that the compatible parameter is set to 11.2.0.2.0 or later.
Preparation—target system	
1	Ensure that Oracle Clusterware is available on the local Oracle RAC node and that the ASM instance is accessible.
2	Ensure that key environment variables are set: ORACLE_SID, ORACLE_BASE, and ORACLE_HOME.
3	Ensure that a tnsnames.ora entry is configured for the source and target/auxiliary databases for use during duplication.
4	Ensure that the Oracle password file is configured.
5	Create an spfile on the target system for the duplication process. For this solution, the following parameters were amended from the default settings: *.db_domain='sse.ea.emc.com' *.db_name='VSE' *.db_create_file_dest='+EA_SAP_DATA' *.db_create_online_log_dest_1='+EA_SAP_REDO' *.db_create_online_log_dest_2='+EA_SAP_REDOM' *.db_recovery_file_dest='+ES_SAP_FRA' *.db_recovery_file_dest_size=53445918720 *.log_archive_format='VSEARC%t_%s_%r.dbf' *.control_files='+EA_SAP_DATA/vse/cntrlVSE1.ctl', '+EA_SAP_REDO/vse/cntrlVSE2.ctl', '+EA_SAP_REDOM/vse/cntrlVSE3.ctl' *.log_file_name_convert= '/oracle/VSE/origlogA/log_g11m1.dbf','+EA_SAP_REDO', '/oracle/VSE/mirrlogA/log_g11m2.dbf','+EA_SAP_REDOM', '/oracle/VSE/origlogB/log_g12m1.dbf','+EA_SAP_REDO', '/oracle/VSE/mirrlogB/log_g12m2.dbf','+EA_SAP_REDOM', '/oracle/VSE/origlogA/log_g13m1.dbf','+EA_SAP_REDO', '/oracle/VSE/mirrlogA/log_g13m2.dbf','+EA_SAP_REDOM', '/oracle/VSE/origlogB/log_g14m1.dbf','+EA_SAP_REDO', '/oracle/VSE/mirrlogB/log_g14m2.dbf','+EA_SAP_REDOM' **Note:** archive logs are written to the FRA by default

Preparation—target system	

6	Start up the target instance in nomount mode:
	`SQL> connect sys/XXXXXXXXX@DUPVSE as SYSDBA`
	`Connected to an idle instance.`
	`SQL> startup nomount`
	`ORA-32004: obsolete or deprecated parameter(s) specified for RDBMS instance`
	`Oracle instance started`
	`Total System Global Area 10689474560 bytes`
	`Fixed Size 2237776 bytes`
	`Variable Size 1644169904 bytes`
	`Database Buffers 8992587776 bytes`
	`Redo Buffers 50479104 bytes`
	`SQL>`

Migrating the database from file system to ASM

Migrating the database to ASM involves using RMAN to create a duplicate instance under ASM on the target system:

1. Start RMAN and connect both the source (target in RMAN) and target (auxiliary in RMAN) databases as sys.
2. Run the RMAN commands shown in Figure 43.

```
connect target sys/xxxxxxx@ORGVSE
connect auxiliary sys/xxxxxxxx@DUPVSE1
run {
ALLOCATE CHANNEL t1 DEVICE TYPE disk;
ALLOCATE CHANNEL t2 DEVICE TYPE disk;
ALLOCATE CHANNEL t3 DEVICE TYPE disk;
ALLOCATE CHANNEL t4 DEVICE TYPE disk;
ALLOCATE CHANNEL t5 DEVICE TYPE disk;
ALLOCATE CHANNEL t6 DEVICE TYPE disk;
ALLOCATE CHANNEL t7 DEVICE TYPE disk;

ALLOCATE AUXILIARY CHANNEL a1 DEVICE TYPE disk;
duplicate target database
to VSE
from active database
nofilenamecheck;
}
```

Figure 43. Sample RMAN duplicate script

In the solution environment, it took approximately 18 minutes to produce a duplicate of the live single instance 500 GB database using this method, as shown in Figure 44.

```
Recovery Manager: Release 11.2.0.3.0 - Production on Fri Mar 2
09:39:33 2012
...
..
.
contents of Memory Script:
{
   Alter clone database open resetlogs;
}
executing Memory Script

database opened
Finished Duplicate Db at 02-MAR-2012 09:57:37
```

Figure 44. Extract from log file of RMAN script

Validating the migration

When the migration is complete, it is important to validate the placement of data files, online redo logs, and control files, and to ensure that no corruption occurred during duplication. Figure 45 shows the RMAN script used to validate the database migration for the solution.

```
run {
ALLOCATE CHANNEL t1 DEVICE TYPE disk;
ALLOCATE CHANNEL t2 DEVICE TYPE disk;
ALLOCATE CHANNEL t3 DEVICE TYPE disk;
ALLOCATE CHANNEL t4 DEVICE TYPE disk;
ALLOCATE CHANNEL t5 DEVICE TYPE disk;
ALLOCATE CHANNEL t6 DEVICE TYPE disk;
ALLOCATE CHANNEL t7 DEVICE TYPE disk;
ALLOCATE CHANNEL t8 DEVICE TYPE disk;
ALLOCATE CHANNEL t9 DEVICE TYPE disk;
ALLOCATE CHANNEL t10 DEVICE TYPE disk;
validate database;}
```

Figure 45. Sample RMAN database validation script

The processing time for this validation script was approximately five minutes. Figure 46 shows the output of the RMAN validation script for one of the cloned data files in the VSE database.

```
File Status Marked Corrupt Empty Blocks Blocks Examined High SCN
---- ------ -------------- ------------ --------------- ----------
18   OK     0              97439        1280000         21966205
  File Name: +EA_SAP_DATA/vse/datafile/psapsr3.259.776893807
  Block Type Blocks Failing Blocks Processed
  ---------- -------------- ----------------
  Data       0              786620
  Index      0              347303
  Other      0              48638
```

Figure 46. Output from RMAN validate database command

Post duplication tasks

When duplication is complete and validated, a new spfile must be created on the target system as part of the post duplication process. Figure 47 shows the creation of the spfile for the solution—a single parameter points from the pfile to the spfile.

```
SQL> alter system reset log_file_name_convert;

System altered.

SQL> create pfile='/home/oracle/initVSE.ora_gen' from memory;

File created.

SQL> REM check and change Oracle instance parameters as required

SQL> !vi /home/oracle/initVSE.ora_gen

SQL> create SPFILE='+EA_SAP_DATA/VSE/spfileVSE.ora' from
pfile='/home/oracle/initVSE.ora_gen';

File created.
```

Figure 47. Recreating the spfile

Converting the single instance database to Oracle RAC

With the duplicate database, spfile, and pfile created, and the database not yet registered with Oracle Clusterware, the single instance database was started with sqlplus, ready for rconfig to convert it to Oracle RAC:

1. Create an rconfig instruction file for the conversion.

   For the solution, the ConvertToRAC_AdminManaged.xml template (located in the $ORACLE_HOME/assistants/rconfig/sampleXMLs directory) was used to create this file. Table 11 lists the required parameter values.

   Table 11. rconfig instruction file parameter values

Parameter	Value
Convert verify	"YES"
SourceDBHome	/oracle/VSE/112
TargetDBHome	/oracle/VSE/112
SourceDBInfo SID	“VSE”
User	Sys
Password	xxxxxxxxx
Role	Sysdba
Node name	‹n:NodeList› ‹n:Node name="sse-ea-erac-n01"/› ‹n:Node name="sse-ea-erac-n02"/› ‹n:Node name="sse-ea-erac-n03"/› ‹n:Node name="sse-ea-erac-n04"/› ‹/n:NodeList›
InstancePrefix	VSE00*
SharedStorage type	“ASM”

   * This prefix meets SAP requirements for Oracle instance naming.

2. Run rconfig as the Oracle user.

The processing time to convert the database and deploy the four instances across the cluster was 11 minutes. Figure 48 shows the rconfig output.

```
oracle@sse-ea-erac-n01:~> rconfig ./VSE.xml
Converting Database "VSE.sse.ea.emc.com" to Cluster Database. Target
Oracle Home: /oracle/VSE/112. Database Role: PRIMARY.
Setting Data Files and Control Files
Adding Database Instances
Adding Redo Logs
Enabling threads for all Database Instances
Setting TEMP tablespace
Adding UNDO tablespaces
Adding Trace files
Setting Fast Recovery Area
Updating Oratab
Creating Password file(s)
Configuring Listeners
Configuring related CRS resources
Starting Cluster Database
<?xml version="1.0" ?>
<RConfig version="1.1" >
<ConvertToRAC>
    <Convert>
      <Response>
        <Result code="0" >
          Operation Succeeded
        </Result>
      </Response>
      <ReturnValue type="object">
<Oracle_Home>
         /oracle/VSE/112
       </Oracle_Home>
       <Database type="ADMIN_MANAGED"  >
         <InstanceList>
           <Instance SID="VSE001" Node="sse-ea-erac-n01"  >
           </Instance>
           <Instance SID="VSE002" Node="sse-ea-erac-n02"  >
           </Instance>
           <Instance SID="VSE004" Node="sse-ea-erac-n03"  >
           </Instance>
           <Instance SID="VSE003" Node="sse-ea-erac-n04"  >
           </Instance>
         </InstanceList>
       </Database>     </ReturnValue>
    </Convert>
  </ConvertToRAC></RConfig>
```

Figure 48. Output from rconfig showing the new instances created

Standardizing Oracle RAC for SAP

After the database was converted, SAP requirements for the Oracle database were met by making the changes shown in Table 12.

Table 12. Matching SAP requirements for Oracle database

Description	Instance name	Changes applied
Online redo log group	VSE001 VSE002 VSE003 VSE004	Redo log groups 11 – 14 Redo log groups 21 – 24 Redo log groups 31 – 34 Redo log groups 41 – 44
Undo tablespace naming	VSE001 VSE002 VSE003 VSE004	PSAPUNDO PSAPUNDO_002 PSAPUNDO_003 PSAPUNDO_004
Listener.ora	Add the following line to the listener.ora on each node, where VSE00x is the instance name for that node: `SID_LIST_LISTENER = (SID_LIST=(SID_DESC=(SID_NAME=VSE00x) (ORACLE_HOME=/oracle/VSE/112)))`	

Connecting to Oracle RAC from SAP

To enable SAP to connect to the newly created RAC database, tnsnames.ora on each of the SAP virtual machines (SAPDI1 and SAPDI2) was amended to use the new database, as shown in Figure 49. The SAP services were then restarted.

```
VSE.WORLD=
  (DESCRIPTION =
    (LOAD_BALANCE = OFF)
    (FAILOVER = ON)
    (ADDRESS_LIST =
        (ADDRESS =
          (PROTOCOL = TCP)
          (HOST = sse-ea-erac-scan-c01.sse.ea.emc.com)
          (PORT = 1521)
        )
    )
    (CONNECT_DATA =
       (SERVICE_NAME = VSE.sse.ea.emc.com)
          (FAILOVER_MODE =
          (TYPE = SELECT)
          (METHOD = BASIC))
    )
  )
```

Figure 49. Sample tnsnames.ora file entry for the Oracle RAC database

Transparent Application Failover (TAF) is a client-side feature that allows clients to reconnect to surviving instances if a database instance fails. TAF can be configured using either a client-side specified connect string or server-side service attributes.

In the solution, database service VSE.sse.ea.emc.com was configured for TAF on Oracle RAC. It was also configured on the client side to enable SAP to use TAF. TAF was set to establish connections at failover time and to enable users with open cursors to continue fetching on them after failure of select operations.

Brocade network infrastructure

Introduction

Overview

This section describes the IP and SAN networks deployed for the solution in the two data centers, and the Layer 2 extension between the data centers. 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 VDX 6720

The Brocade VDX 6720 Data Center Switch is a high-performance, ultra-low latency, wire-speed 10 GbE fixed port switch. It is specifically designed to improve network utilization, maximize application availability, increase scalability, and dramatically simplify network architecture in virtualized data centers. With a rich set of Layer 2 features, the Brocade VDX 6720 is an ideal platform for traditional Top-of-Rack (ToR) switch deployments.

By delivering Brocade VCS Fabric technology, the Brocade VDX 6720 enables organizations to build data center Ethernet fabrics—revolutionizing the design of Layer 2 networks and providing an intelligent foundation for cloud-optimized data centers.

Brocade MLX Series

Brocade MLX Series routers are designed to enable cloud-optimized networks by providing industry-leading 100 GbE, 10 GbE, and 1 GbE wire-speed density; rich IPv4, IPv6, Multi-VRF, Multiprotocol Label Switching (MPLS), and carrier Ethernet capabilities; and advanced Layer 2 switching.

By leveraging the Brocade MLX Series, mission-critical data centers can support more traffic, achieve greater virtualization, and provide high-value cloud-based services using less infrastructure—thereby simplifying operations and reducing costs. Moreover, the Brocade MLX Series can reduce complexity in large campus networks by collapsing core and aggregation layers, as well as providing connectivity between sites using MPLS/VPLS. All of the Brocade MLX Series routers help reduce power and cooling costs with the lowest power consumption and heat dissipation in their class.

Designed for non-stop networking, the Brocade MLX Series features Multi-Chassis Trunking (MCT), which provides more than 30 TB/s of dual-chassis bandwidth, full active/active routing links, and uninterrupted traffic flow in the event of node failover. Organizations can achieve high resiliency through fully redundant switch fabrics, management modules, power supplies, and cooling systems. To further ensure network and application availability, the Brocade IronWare operating system features hitless management failover and software upgrades.

Brocade DCX 8510 Backbone

Networks need to evolve in order to support the growing demands of highly virtualized environments and private cloud architectures. Today, Fibre Channel (FC) is the de facto standard for storage networking in the data center. The introduction of 16 Gb/s Fibre Channel extends the life of this robust, reliable, and high-performance technology. This enables organizations to continue leveraging their existing IT investments as they solve their most difficult business challenges.

Brocade DCX 8510 Backbones are the industry's most powerful 16 Gb/s Fibre Channel switching infrastructure, and provide the most reliable, scalable, high-performance foundation for private cloud storage and highly virtualized environments. They are designed to increase business agility while providing non-stop access to information and reducing infrastructure and administrative costs.

The 16 Gb FC capability of the Brocade DCX 8510 offers significant benefits for data center to data center SAN Metro connectivity:

- 16 Gb provides the maximum throughput and lowest latency FC for deployments utilizing Fibre connections between data centers.
- Optional 10 Gb FC line speed for optimal line utilization if a DWDM network is deployed between sites. This feature requires a license.
- Optional frame-level Inter-Switch Link (ISL) Trunking that enables high utilization compared to standard DPS trunking. This feature requires a license.
- Optional compression for the ISLs between the data centers. This provides added bandwidth for deployments where the number of site-to-site connections are limited.
- Optional data in-flight encryption for the ISLs between the data centers for deployments requiring very high levels of data security.
- Buffer credit loss detection and recovery.
- Automatic Forward Error Correction (FEC), which proactively corrects up to 11 bit errors per 2112-bit FC frame.
- Diagnostic mode for the ISL ports between data centers can be used on any (offline) ISL port and offers the following features:
  - Electrical and optical loopback tests
  - Link saturation testing
  - Link distance measurement accuracy within 5 m when used with 8 Gb SFP+ and 50 m when used with 10 GbE SFP+.

IP network configuration

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 the solution, a separate VLAN—VLAN 10—is used for the interconnect. VLAN 20 handled 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 50 shows the IP network infrastructure.

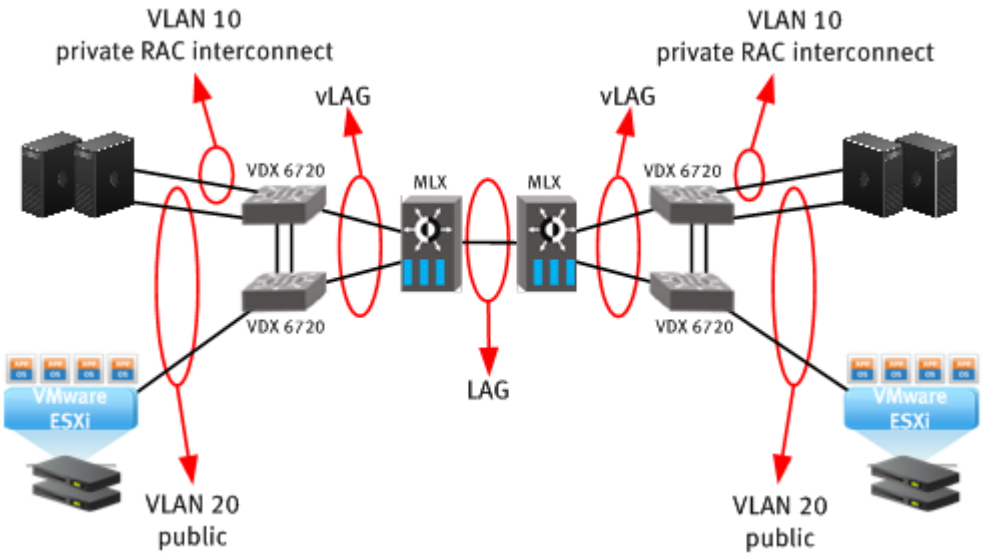

Figure 50. The solution IP networks

SAN network configuration

The SAN in each data center is built with Brocade DCX 8510 Backbones, as shown in Figure 51. 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 uses multiple FC connections between the Brocade DCX 8510 Backbones. These are used in active/active mode with failover.

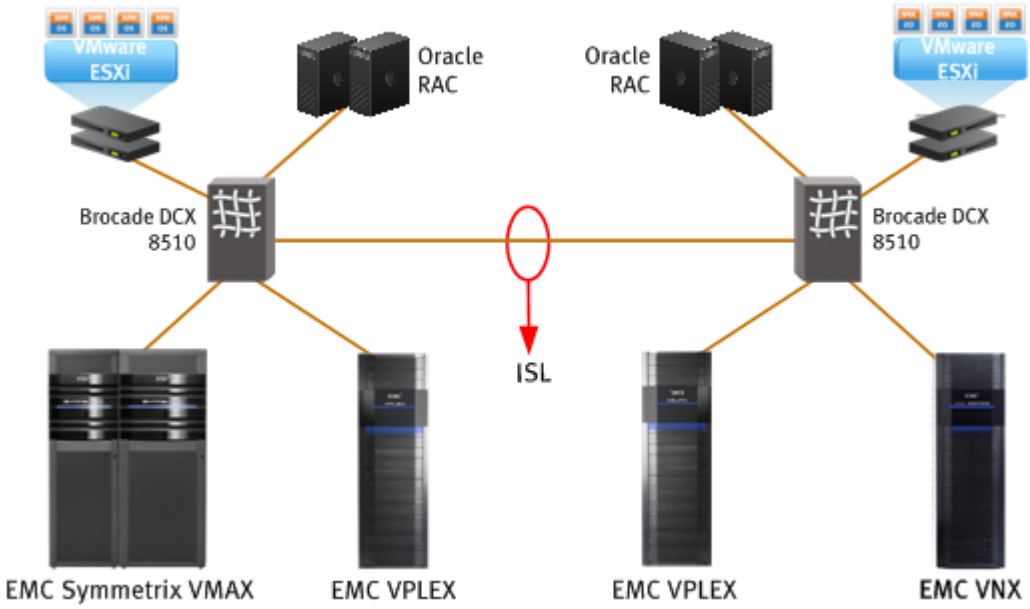

Figure 51. The solution SAN networks

EMC storage infrastructure

Introduction

Overview

This section describes the storage infrastructure for the solution:

- A Symmetrix VMAX array provides the storage platform at Site A
- An EMC VNX5700 array provides the storage platform at Site B

The two storage arrays are deployed with a matching LUN configuration.

EMC Symmetrix VMAX

EMC Symmetrix VMAX is a high-end storage array based on Intel Xeon processors and optimized for the virtual data center. Intel Stop and Scream detects poison packets in PCIe, and enables enhanced error isolation in a multiblade, highly-available environment. This results in shorter downtime, faster problem diagnosis, and simplified repair process, enabling the IT manager to optimize the virtual data center.

Built on the strategy of simple, intelligent, modular storage, the Symmetrix VMAX incorporates a highly scalable Virtual Matrix Architecture that enables it to grow seamlessly and cost-effectively from an entry-level configuration into the world's largest storage system. The VMAX supports Flash, FC, and SATA drives within a single array, and an extensive range of RAID types. EMC Fully Automated Storage Tiering for Virtual Pools (FAST VP) automates tiered storage strategies.

The EMC Enginuity operating environment provides the intelligence that controls all components in a VMAX array.

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. VNX supports Flash, SAS, and NL-SAS drives within a single array and an extensive range of RAID types. FAST VP provides automated storage tiering across all drive types.

The VNX series utilizes the Intel Xeon 5600 series processors, which help make it 2-3 times faster overall than its predecessor. The VNX quad-core processor supports 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 FAST on the VNX, with optimized performance and capacity, without tradeoffs, in a fully automated fashion.

The VNX Operating Environment (VNX OE) allows Microsoft Windows and Linux/UNIX clients to share files in multiprotocol NFS and CIFS environments. At the same time, it supports Internet SCSI (iSCSI), FC, and Fibre Channel over Ethernet (FCoE) access for high-bandwidth and latency-sensitive block applications.

Symmetrix VMAX configuration

Storage layout

For the solution, VPLEX Metro, Oracle Extended RAC, and SAP volumes are laid out using Virtual Provisioning. This configuration places the Oracle data files and log files in separate thin pools and allows each to use distinct RAID protection. The data files reside in a RAID 5 protected pool and the redo logs in a RAID 1 protected pool.

Figure 52 is a logical representation of how the storage layout corresponds to the Oracle ASM disk groups.

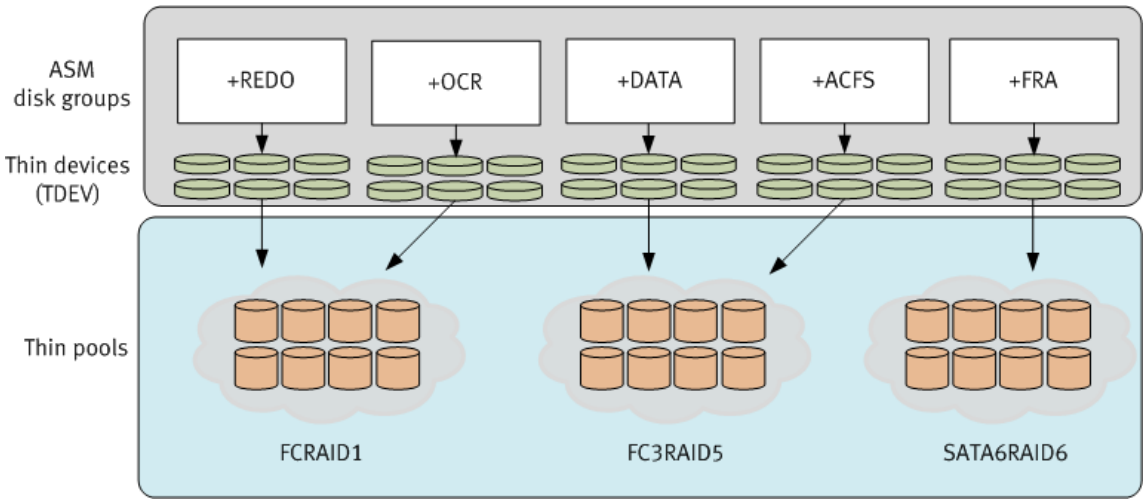

Figure 52. Storage groups and ASM disk groups

Storage was not pre-allocated to any devices, except for the Oracle REDO log devices. EMC recommends that these devices are fully pre-allocated on creation, using persistent allocation. This ensures that their storage is available up front and, if a zero space reclaim is run on the pool at any stage, their pre-allocated capacity is not returned to the pool's free space.

Device tables

Table 13 shows the size and number of devices configured for each ASM disk group.

Table 13. Device sizes

Storage group	Number of devices	Device size (GB)
OCR	5	8
FRA	4	16
REDO	8	16
DATA	16	128
ACFS	4	16

Table 14 shows the size and number of devices configured for VPLEX Metro.

Table 14. Size and number of devices configured for VPLEX Metro

VPLEX Device	Number of devices	Device size (GB)
VPLEX metadata	2	80
VPLEX log volume	2	20
VPLEX metadata backup	2	80

VNX5700 configuration

For the solution, VPLEX Metro, Oracle Extended RAC, and SAP volumes on the VNX5700 array at Site B were laid out using traditional RAID groups and LUNs, whereas these volumes were laid out on the VMAX at Site A using Virtual Provisioning. Similar EMC best practices apply to both provisioning methods, and the same ASM disk groups were created on the VNX and VMAX.

On the VNX, this configuration places the Oracle data files and log files in separate RAID groups and allows each to use distinct RAID protection. The data files reside in a RAID 5 protected RAID group and the redo logs in a RAID 10 protected RAID group. The FRA disk group resides on NL-SAS drives with RAID 6 protection.

The LUNs created on the VNX match the number and size of the thin devices created on the VMAX, as shown in Figure 52, Table 13, and Table 14. Figure 53 shows the layout of the RAID groups on the VNX.

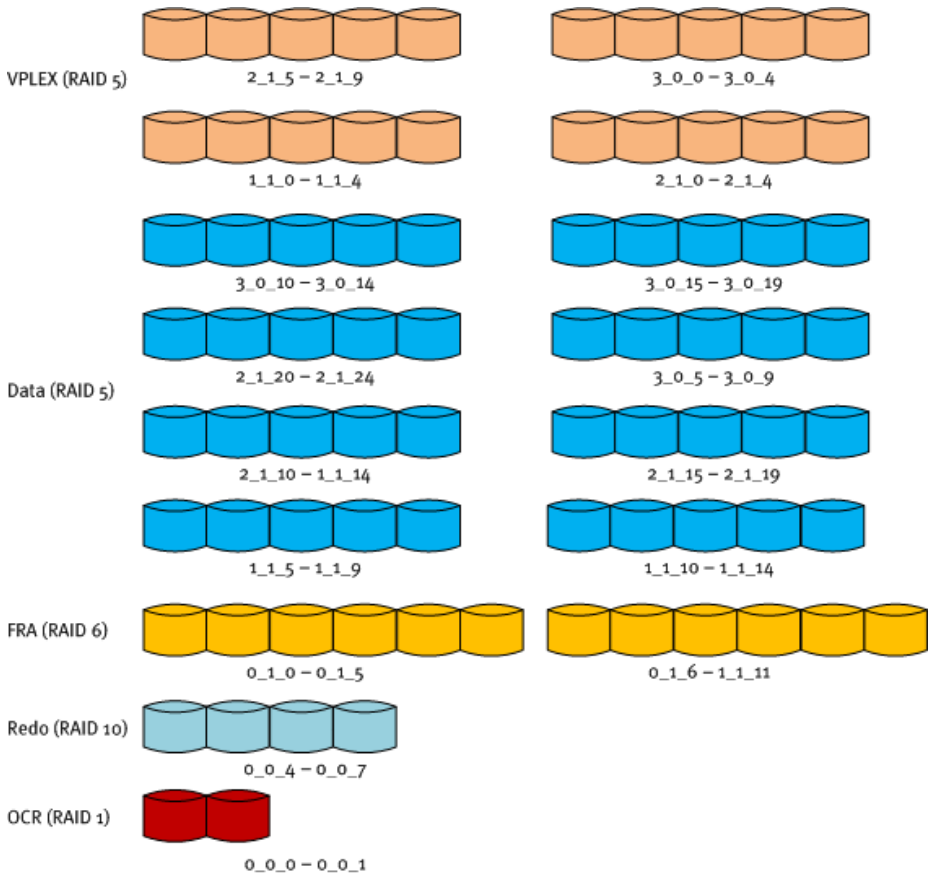

Figure 53. VNX RAID group layout

High availability and business continuity – testing and validation

Introduction

The EMC validation team initially installed and validated the environment without any high-availability or business continuity protection schemes. We then transformed the environment to the mission-critical business continuity solution described in this white paper. We carried out the following tests to validate the solution and demonstrate the elimination of all single points of failure from the environment:

- SAP enqueue service process failure
- SAP ASCS instance virtual machine failure
- Oracle RAC node failure
- Site failure
- VPLEX cluster isolation

SAP enqueue service process failure

Test scenario

This test scenario validates that, if the enqueue service process fails, the SUSE Linux Enterprise High Availability Extension cluster promotes the SAP ERS instance to a fully functional ASCS instance and takes over the lock table without end user interruption.

To test this failure scenario, we terminated the enqueue process on the active ASCS node by running the kill command:

```
kill -9 <process id>
```

System behavior

The system responds to the enqueue service process failure as follows:

1. The SAPInstance resource agent detects and reports the failure, as shown in Figure 54.

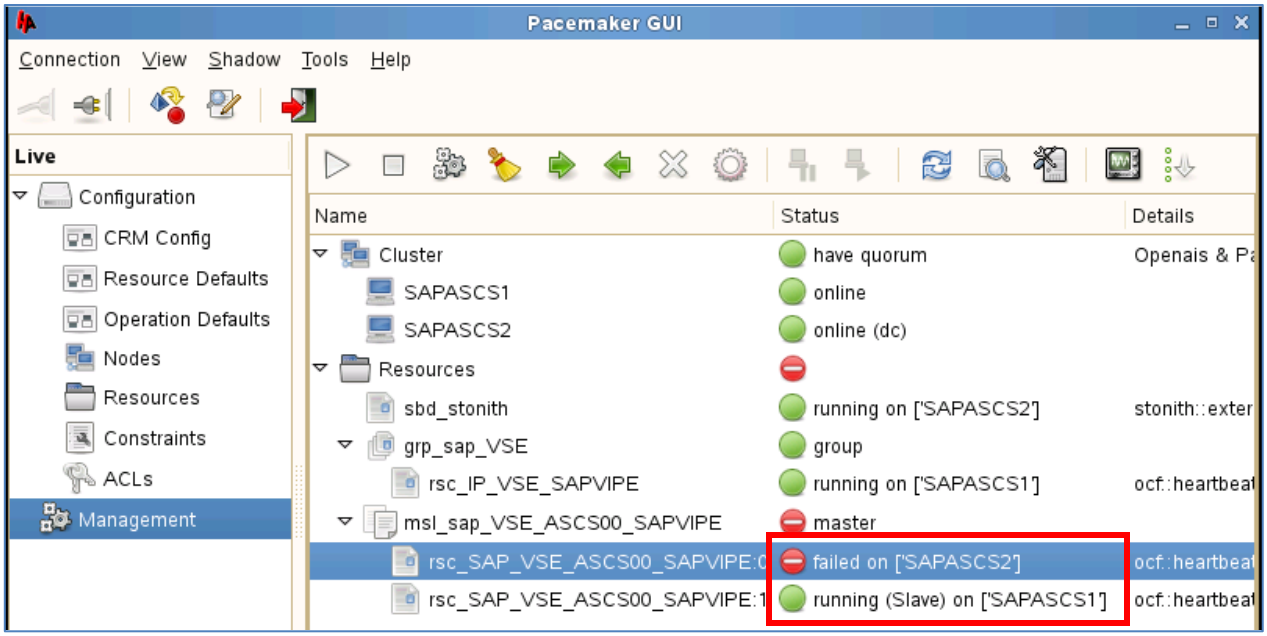

Figure 54. SAPInstance resource agent detects and reports failure

2. The master/slave resource agent promotes the previous slave node (SAPASCS1) to the master node, which hosts the ASCS services, and starts the ERS as a slave on the other node (SAPASCS2) when it rejoins the cluster (see Figure 55).

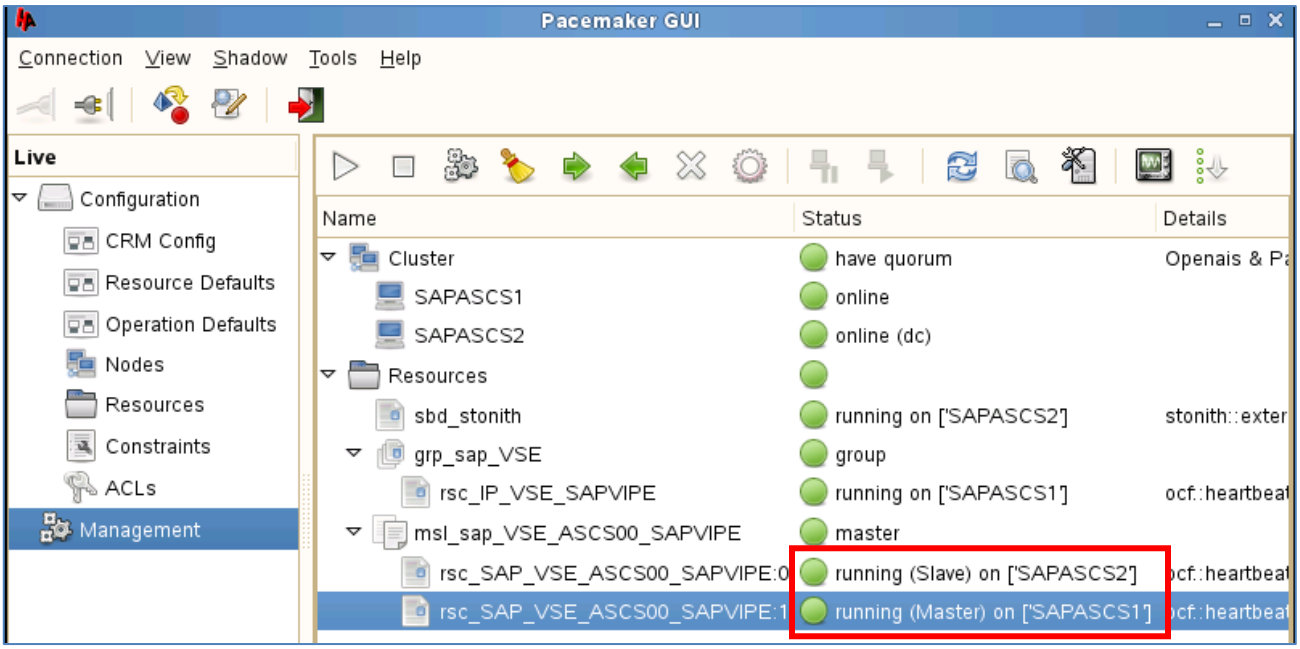

Figure 55. Master/slave resource agent switches the master and slave nodes

3. The replicated lock table is restored, as shown in Figure 56.

```
[Thr 140687136999168] profile    /sapmnt/VSE/profile/VSE_ASCS00_SAPVIPE
[Thr 140687136999168] hostname   SAPASCS2
[Thr 140687136999168] ShadowTable:attach: ShmCreate(,SHM_ATTACH,) -> 0x7ff44169a
000
[Thr 140687136999168] EnRepClass::getReplicaData: found old replication table wi
th the following data:
[Thr 140687136999168]   Line size:744,   Line count: 3603,   Failover Count: 1
[Thr 140687136999168]   EnqId: 1334757348/5794, last stamp: 1/334758434/32000
[Thr 140687136999168]   Byte order tags: int:10079666   char:Z
[Thr 140687136999168] Enqueue: EnqMemStartupAction Utc=1334758495
[Thr 140687136999168] Enqueue Info: replication enabled
[Thr 140687136999168] Enqueue Info: enque/replication_dll not set
[Thr 140687136999168] Enqueue checkpointing: start restoring entries. Utc=133475
8495
[Thr 140687136999168] ShadowTable:destroy: ShmCleanup( SHM_ENQ_REP_SHADOW_TBL)
[Thr 140687136999168] enque/backup_file disabled in enserver environment
```

Figure 56. Replicated lock table restored

Result

The end user does not notice the enqueue process failure, unless an enqueue operation is running. In this case, the end user experiences a longer transaction response time during the switchover. New users can log into the system immediately after the message server switchover. No administrative intervention is required.

SAP ASCS instance virtual machine failure

Test scenario

This test scenario validates that, in the event of an unexpected ESXi server outage (which is equivalent to a virtual machine failure), the High Availability Extension cluster promotes the SAP ERS instance to a fully functional ASCS instance and takes over the lock table, without end-user interruption.

To test this failure scenario, we powered off (via DRAC) the ESXi server that hosts the SAP ASCS instance virtual machine. We then rebooted the server without entering maintenance mode.

System behavior

The system responds to the virtual machine failure as follows:

1. SAPASCS2 becomes unavailable from vSphere Client (see Figure 57).

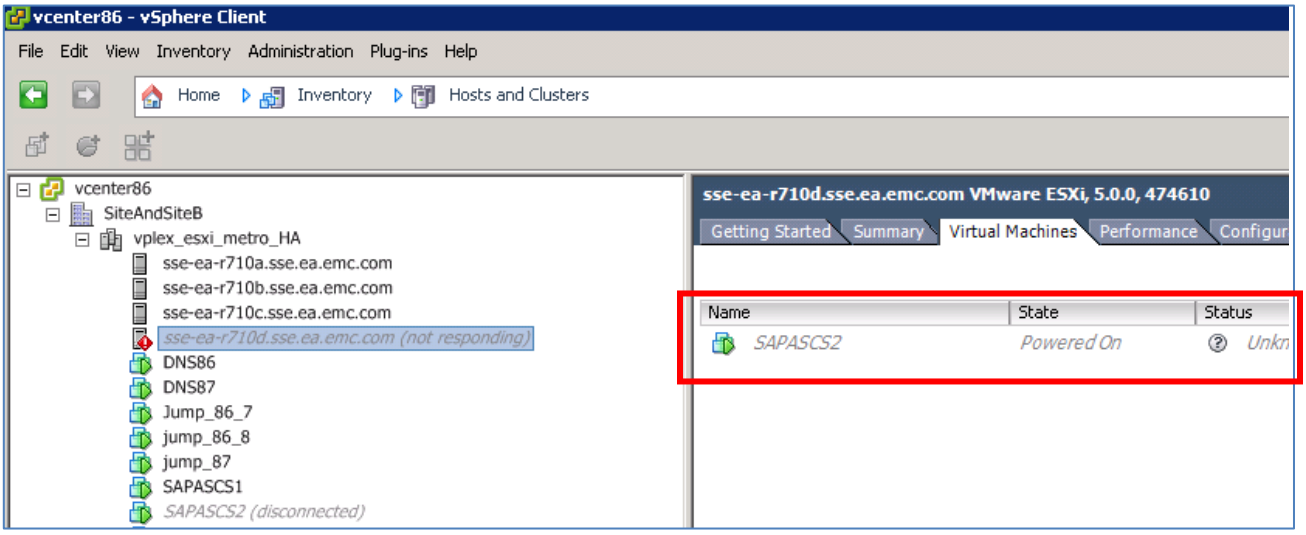

Figure 57. Virtual machine fails

2. The SAPInstance resource agent detects and reports the failure (see Figure 58).

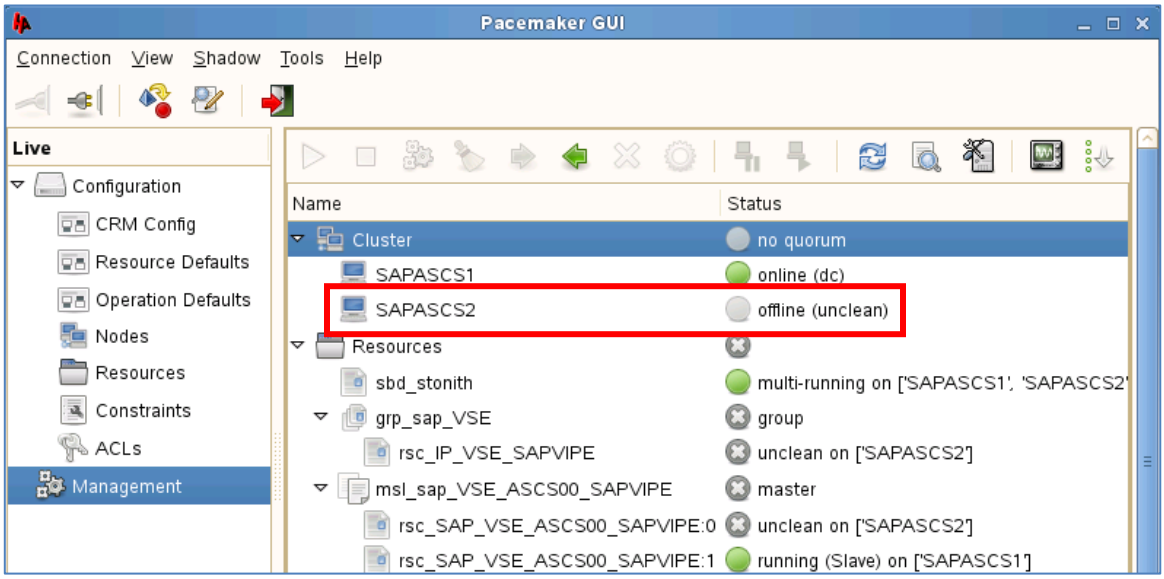

Figure 58. SAPInstance resource agent detects and reports failure

3. VMHA restarts the failed virtual machine (SAPASCS2) on the surviving ESXi host (see Figure 59).

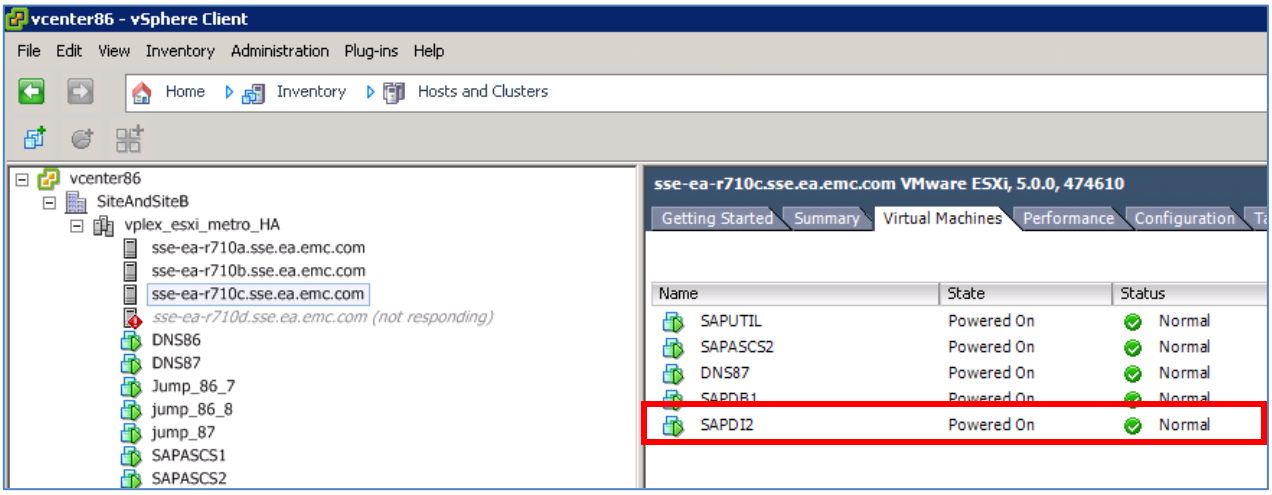

Figure 59. VMHA restarts the failed virtual machine

4. The master/slave resource agent promotes the previous slave node (SAPASCS1) to the master node, which hosts the ASCS services, and starts the ERS as a slave on the other node (SAPASCS2) when it rejoins the cluster (see Figure 60).

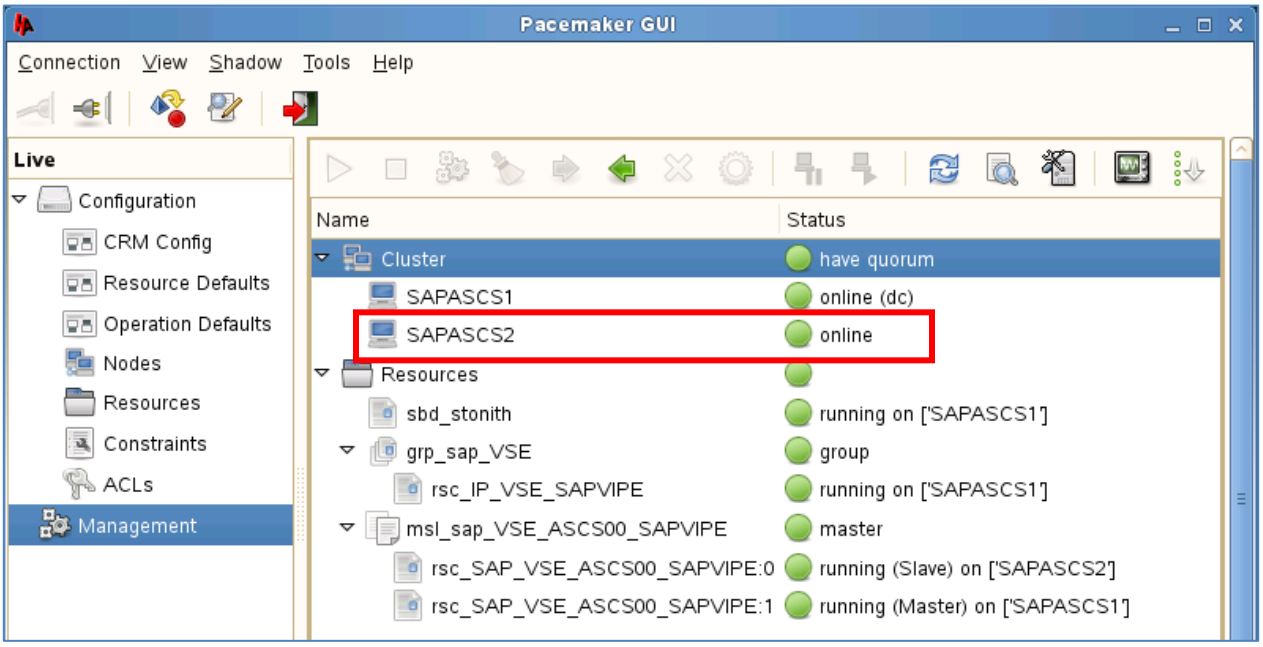

Figure 60. Master/slave resource agent switches the master and slave nodes

5. The replicated lock table is restored (see Figure 61).

```
[Thr 140101975725824] ShadowTable:attach: ShmCreate(,SHM_ATTACH,) -> 0x7f6c03153000
[Thr 140101975725824] EnRepClass::getReplicaData: found old replication table with the
following data:
[Thr 140101975725824]   Line size:744,   Line count: 3603,   Failover Count: 3
[Thr 140101975725824]   EnqId: 1334904364/24586, last stamp: 1/334913011/31000
[Thr 140101975725824]   Byte order tags: int:10079666   char:Z
[Thr 140101975725824] Enqueue: EnqMemStartupAction Utc=1334913077
[Thr 140101975725824] Enqueue Info: replication enabled
[Thr 140101975725824] Enqueue Info: enque/replication_dll not set
[Thr 140101975725824] Enqueue checkpointing: start restoring entries. Utc=1334913077
[Thr 140101975725824] ShadowTable:destroy: ShmCleanup( SHM_ENQ_REP_SHADOW_TBL)
[Thr 140101975725824] enque/backup_file disabled in enserver environment
```

Figure 61. Replicated lock table restored

Result

The end user does not notice the enqueue process failure, unless an enqueue operation is running. In this case, the end user experiences a longer transaction response time during the switchover. New users can log into the system immediately after the message server switchover. No administrative intervention is required.

Oracle RAC node failure

Test scenario

This test scenario validates that, in the event of an unexpected RAC node failure, the SAP instances automatically connect to other RAC nodes. End users can continue their transactions without interruption, unless uncommitted transactions (at the database level) are being executed on the failed RAC node.

To test this failure scenario, we rebooted the server to cause an Oracle node failure.

System behavior

The system responds to the RAC node failure as follows:

1. The RAC node goes offline and instance VSE003 is unavailable, as shown in Figure 62.

```
oracle@sse-ea-erac-n01:~> srvctl status database -d VSE
Instance VSE001 is running on node sse-ea-erac-n01
Instance VSE002 is running on node sse-ea-erac-n02
Instance VSE004 is running on node sse-ea-erac-n03
Instance VSE003 is not running on node sse-ea-erac-n04
```

Figure 62. RAC node goes offline

2. The SAP instance work process connects to another RAC instance, as shown in Figure 63.

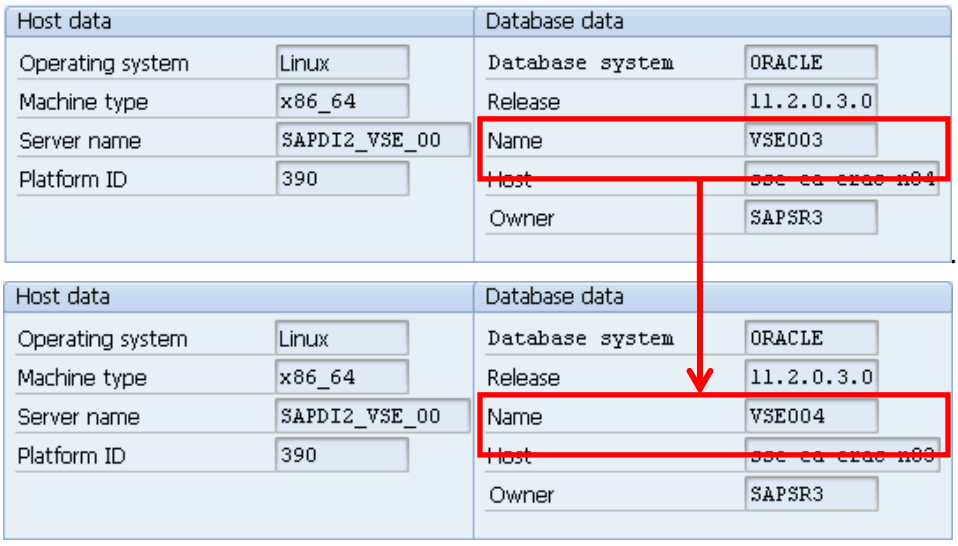

Figure 63. SAP instance connects to another RAC node

Result

The end user experiences a longer transaction response time when the dialog instance work process reconnects to another RAC node. Uncommitted transactions are rolled back at the database level to guarantee the data consistency. The end user receives a system error message (short dump) and needs to restart the transaction. No administrative intervention is required.

Site failure

Test scenario

This test scenario validates that, in the event of a complete site failure, the surviving RAC nodes preserve database operations.

To test this failure scenario, we simulated a complete failure of Site A, including VPLEX cluster, ESXi server, network, and Oracle RAC node components. The VPLEX Witness remained available on Site C. On Site B, VPLEX cluster-2 remained in communication with the VPLEX Witness.

Figure 64 shows the status of the environment before the site failure.

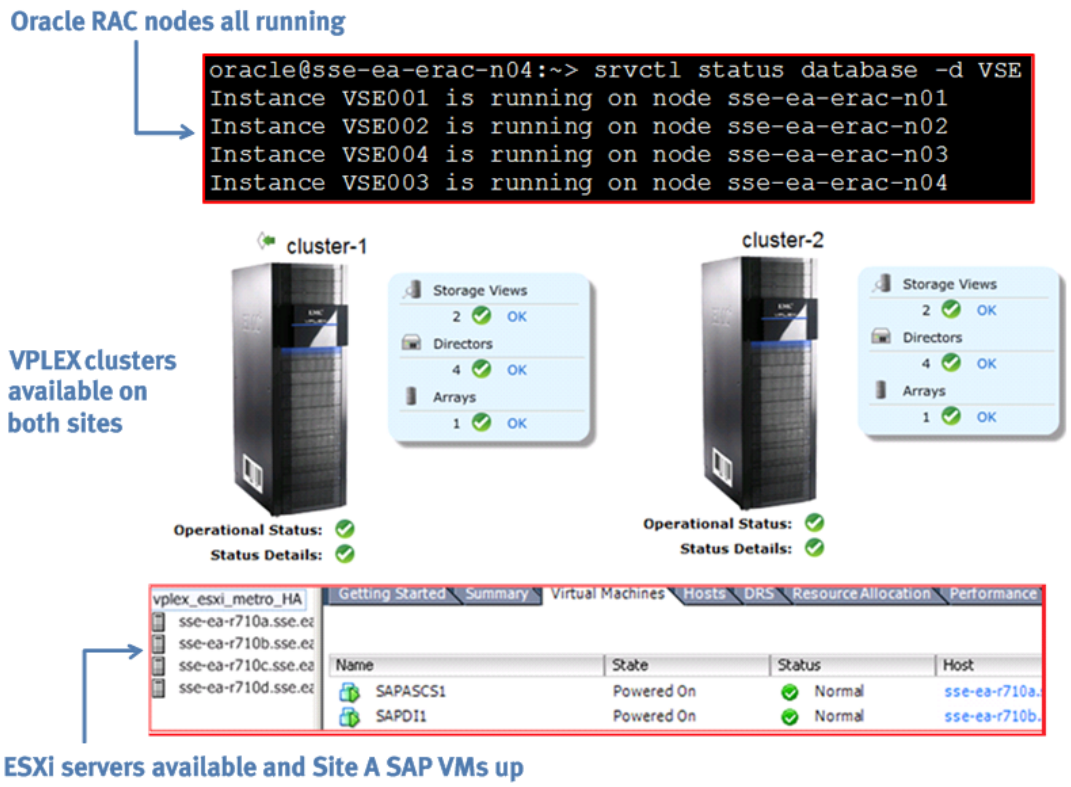

Figure 64. Status of environment prior to Site A failure

System behavior

The system responds to site failure as follows:

- When Site A fails, VPLEX Witness ensures that the consistency group's detach rule, which defines cluster-1 as the preferred cluster, is overridden and that the storage served by VPLEX cluster-2 on Site B remains available.
- RAC nodes sse-ea-erac-n03 and sse-ea-erac-n04 on Site B remain available.
- When the ESXi servers on Site A fail, VMHA restarts SAPASCS1 and SAPDI1 on Site B. SAPASCS1 is restarted on a different ESXi host to SAPASCS2, as prescribed by the defined VM-VM affinity rule.
- SUSE Linux Enterprise High Availability Extension detects the failure of cluster node SAPASCS1. Because the ERS was running on this node, the cluster takes no action except to restart the ERS when SAPASCS1 rejoins the cluster. The lock table is preserved and operational all the time.
- End users on SAPDI1 lose their sessions due to the ESXi server failure. During the restart process, new users are directed to SAPDI2. When SAPDI1 restarts on Site B, users can log into SAPDI1 again.

Figure 64 shows the status of the environment after the site failure.

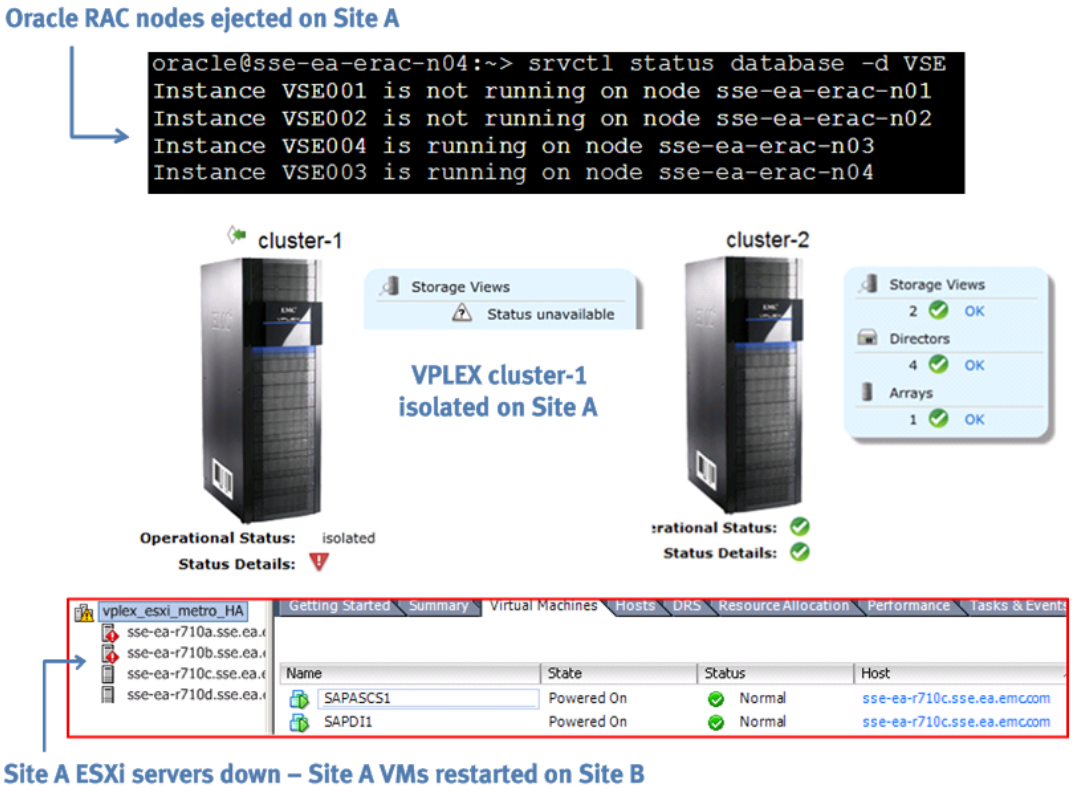

Figure 65. Environment status after Site A failure

Result

Table 15 shows the expected and observed behaviors of the system when Site A fails.

Table 15. Expected and observed behaviors

System		Status prior to test	Expected behavior	Observed behavior
Oracle RAC nodes (database VSE)	sse-ea-erac-n01 (Site A) sse-ea-erac-n02 (Site A) sse-ea-erac-n03 (Site B) sse-ea-erac-n04 (Site B)	Available Available Available Available	**Ejected** **Ejected** Available Available	**Ejected** **Ejected** Available Available
ESXi server *Virtual machine*	sse-ea-r710a (Site A) *SAPASCS1* sse-ea-r710b (Site A) *SAPDI1* sse-ea-r710c (Site B) *SAPDI2* sse-ea-r710d (Site B) *SAPASCS2*	Available *Available* Available *Available* Available *Available* Available *Available*	**Unavailable** ***VMHA restart Site B*** **Unavailable** ***VMHA restart Site B*** Available *Available* Available *Available*	**Unavailable** ***VMHA restart Site B*** **Unavailable** ***VMHA restart Site B*** Available *Available* Available *Available*

System		Status prior to test	Expected behavior	Observed behavior
VPLEX cluster	VPLEX1 – Site A – cluster-1 VPLEX2 – Site B – cluster-2	Available Available	**Unavailable** Available	**Unavailable** Available
SAP services	Enqueue Replication Server Enqueue/Message Server	Available Available	**Unavailable** **SLE HAE restart after reboot on Site B** Available	**Unavailable** **SLE HAE restart after reboot on Site B** Available

VPLEX cluster isolation

Test scenario

This test scenario validates that, in the event of isolation of a VPLEX cluster, the SAP applications and database continue operation on the surviving site without interruption.

To test this failure scenario, we simulated isolation of the preferred cluster on Site A, with both the external Management IP network and the VPLEX WAN communications network partitioned. The LAG network remains available. VPLEX Witness remains available on Site C. On Site B, VPLEX cluster-2 remains in communication with VPLEX Witness.

System behavior

The system responds to the VPLEX cluster isolation as follows:

- When the VPLEX on Site A becomes isolated, the VPLEX Witness ensures that the consistency group's detach rule, which defines cluster-1 as the preferred cluster, is overridden and that the storage served by VPLEX cluster-2 on Site B remained available.
- RAC nodes sse-ea-erac-n03 and sse-ea-erac-n04 on Site B remain available and RAC nodes sse-ea-erac-n01 and sse-ea-erac-n02 on Site A are ejected.
- The ESXi servers on Site A remain available and virtual machines SAPASCS1 and SAPDI1 remain active due to the use of VPLEX Metro HA Cross-Cluster Connect.

Figure 64 shows the status of the environment after VPLEX isolation on Site A.

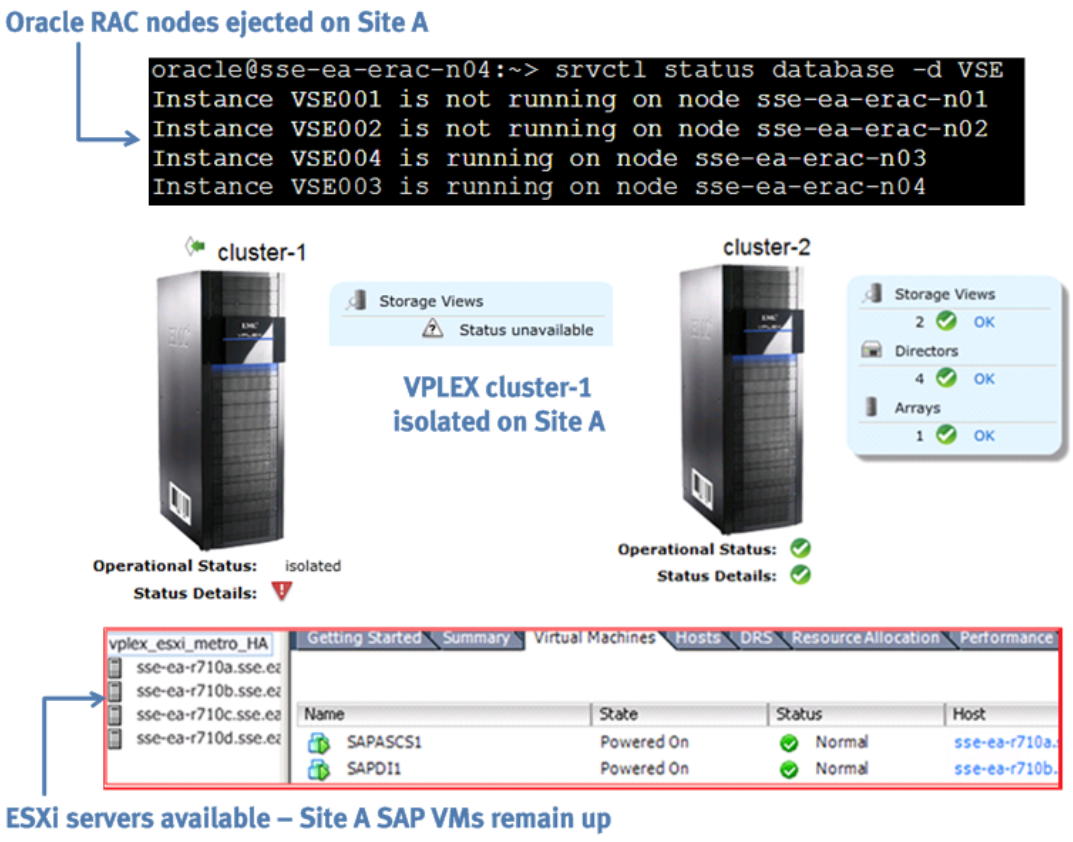

Figure 66. Environment status after isolation of VPLEX on Site A

Result

Table 16 shows the expected and observed behaviors of the system when the VPLEX at Site A is isolated.

Table 16. Expected and observed behaviors

System		Status prior to test	Expected behavior	Observed behavior
Oracle RAC nodes (database VSE)	sse-ea-erac-n01 (Site A) sse-ea-erac-n02 (Site A) sse-ea-erac-n03 (Site B) sse-ea-erac-n04 (Site B)	Available Available Available Available	**Ejected** **Ejected** Available Available	**Ejected** **Ejected** Available Available
ESXi server *Virtual machine*	sse-ea-r710a (Site A) *SAPASCS1* sse-ea-r710b (Site A) *SAPDI1* sse-ea-r710c (Site B) *SAPDI2* sse-ea-r710d (Site B) *SAPASCS2*	Available *Available* Available *Available* Available *Available* Available *Available*	Available *Available* Available *Available* Available *Available* Available *Available*	Available *Available* Available *Available* Available *Available* Available *Available*

System		Status prior to test	Expected behavior	Observed behavior
VPLEX cluster	VPLEX1 – Site A – cluster-1 VPLEX2 – Site B – cluster-2	Available Available	**Unavailable** Available	**Unavailable** Available
SAP services	Enqueue Replication Server Enqueue/Message Server	Available Available	Available Available	Available Available

Conclusion

Summary

This solution demonstrates the transformation of a traditional active/passive SAP deployment to a highly available business continuity solution with active/active data centers and always-on application availability.

The solution combines EMC, VMware, Oracle, SUSE, and Brocade high-availability components to:

- Eliminate single points of failure at all layers in the environment
- Provide active/active data centers that support near-zero RPOs and RTOs
- Enable mission-critical business continuity for SAP applications

Each single point of failure was identified and mitigated by using fault-tolerant components and high-availability clustering technologies. Resource utilization was increased by enabling active/active data access. Failure handling was fully automated to eliminate the final and often most unpredictable SPOF from the architecture—people and processes.

In addition, the use of management and monitoring tools such as the vSphere Client, EMC Virtual Storage Integrator, and the VPLEX performance tools simplifies operational management and allows monitoring and mapping of the infrastructure stack.

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

To validate the solution, the EMC validation team ran the following tests and noted the behaviors indicated:

- Simulate a SAP enqueue service process failure
  - ✓ Application continues without interruption
- Simulate a SAP ASCS instance virtual machine failure
  - ✓ Application continues without interruption
- Simulate an Oracle RAC node failure
  - ✓ Application continues without interruption

- Simulate a total site failure
  - ✓ Application continues without interruption
- Validate VPLEX Witness functionality during simulated isolation of a VPLEX cluster
  - ✓ Application continues without interruption

The testing demonstrates how VMware, SAP, SUSE, and Oracle high-availability solutions eliminate single points of failure at the local level.

It also demonstrates how VPLEX Metro, combined with SUSE Linux Enterprise High Availability Extension, Oracle Extended RAC, and Brocade networking, extends this high availability to 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 Witness and Cross-Cluster Connect provide an even higher level of resilience.

Together, these technologies enable transformation of a traditional active/passive data center deployment to a mission-critical business continuity solution with active/active data centers, 24/7 application availability, no single points of failure, and near-zero RTOs and RPOs.

References

EMC

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

- *EMC VPLEX Metro Witness Technology and High Availability*
- *Using VMware vSphere with EMC VPLEX Best Practices Planning*
- *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:

- *Moving your SAP Database to Oracle Automatic Storage Management 11g Release 2: A Best Practices Guide*
- *SAP with Oracle Real Application Clusters 11g Release 2 and Oracle Automatic Storage Management 11g Release 2: Advanced Configurations & Techniques*
- *Configuration of SAP NetWeaver for Oracle Grid Infrastructure 11.2.0.2 and Oracle Real Application Clusters 11g Release 2: A Best Practices Guide*
- *Oracle Real Application Clusters (RAC) on Extended Distance Clusters*
- *Oracle Database Upgrade Guide Upgrade to Oracle Database 11g Release 2 (11.2): UNIX For Oracle Patch Set Release 11.2.0.2 and 11.2.0.3*

VMware

For additional information, see the following VMware documents:

- *VMware vSphere Networking ESXi 5.0*
- *VMware vSphere Availability ESXi 5.0*
- *VMware Knowledge Base article 1026692: Using VPLEX Metro with VMware HA*
- *VMware Knowledge Base article 1034165: Disabling simultaneous write protection provided by VMFS using the multi-writer flag*
- *SAP Solutions on VMware vSphere: High Availability*
- *SAP Solutions on VMware: Best Practices Guide*

SUSE

For additional information, see the following SUSE documents:

- *SUSE Linux Enterprise High Availability Extension – High Availability Guide*
- *Running SAP NetWeaver on SUSE Linux Enterprise Server with High Availability – Simple Stack*
- *SAP Applications Made High Available on SUSE Linux Enterprise Server 10*
- *Protection of Business-Critical Applications in SUSE Linux Enterprise Environments Virtualized with VMware vSphere 4 and SAP NetWeaver as an Example*

SAP

For additional information, see the following SAP documents:

- *SAP Note 1552925 – Linux High Availability Cluster Solutions*
- *SAP Note 1431800 – Oracle 11.2.0 Central Technical Note*
- *SAP Note 105047 – Support for Oracle Functions in the SAP Environment*
- *SAP Note 1550133 – Oracle Automatic Storage Management (ASM)*
- *SAP Note 527843 – Oracle RAC Support in the SAP Environment*
- *SAP Note 989963 – Linux: VMware Timing Problem*
- *SAP Note 1122388 – Linux: VMware vSphere Configuration Guidelines*
- *SAP Note 1310037 – SUSE Linux Enterprise Server 11: Installation notes*
- *SAP Installation Guide for SAP ERP 6.0 – EHP 4 Ready ABAP on Linux: Oracle - Based on SAP NetWeaver 7.0 including Enhancement Package 1*
- *SAP Enqueue Replication Server Setup help portal*

Appendix – Sample configurations

CRM sample configuration

```
node SAPASCS1 \
        attributes standby="off"
node SAPASCS2 \
        attributes standby="off"
primitive rsc_IP_VSE_SAPVIPE ocf:heartbeat:IPaddr2 \
        operations $id="rsc_IP_VSE_SAPVIPE-operations" \
        op monitor interval="10s" timeout="20s" on_fail="restart" \
        params ip="xxx.xxx.xxx.xxx" \
        meta is-managed="true"
primitive rsc_SAP_VSE_ASCS00_SAPVIPE ocf:heartbeat:SAPInstance \
        operations $id="rsc_SAP_VSE_ASCS00_SAPVIPE-operations" \
        op monitor interval="120" enabled="true" role="Master" timeout="60" start_delay="5" \
        op start interval="0" timeout="180" \
        op stop interval="0" timeout="240" \
        op promote interval="0" role="Master" timeout="320" start_delay="0" \
        op demote interval="0" role="Slave" timeout="320" start_delay="0" \
        params InstanceName="VSE_ASCS00_SAPVIPE"
ERS_InstanceName="VSE_ERS01_SAPASCS2" AUTOMATIC_RECOVER="true"
START_PROFILE="/sapmnt/VSE/profile/START_ASCS00_SAPVIPE"
ERS_START_PROFILE="/sapmnt/VSE/profile/START_ERS01_SAPASCS2" \
        meta target-role="Started"
primitive sbd_stonith stonith:external/sbd \
        meta target-role="started" \
        op monitor interval="15" timeout="15" start-delay="15" \
        params sbd_device="/dev/sdb1"
group grp_sap_VSE rsc_IP_VSE_SAPVIPE \
        meta is-managed="true" target-role="started"
ms msl_sap_VSE_ASCS00_SAPVIPE rsc_SAP_VSE_ASCS00_SAPVIPE \
        meta globally-unique="true" target-role="Started" clone-node-max="1" master-max="1"
notify="true"
colocation colocation_IP_ASCS inf: grp_sap_VSE:Started
msl_sap_VSE_ASCS00_SAPVIPE:Master
order ord_VSE_IP_Master : grp_sap_VSE msl_sap_VSE_ASCS00_SAPVIPE:promote
symmetrical=false
property $id="cib-bootstrap-options" \
        dc-version="1.1.5-5bd2b9154d7d9f86d7f56fe0a74072a5a6590c60" \
        cluster-infrastructure="openais" \
        expected-quorum-votes="2" \
        last-lrm-refresh="1329421965" \
        default-resource-stickiness="1000" \
        no-quorum-policy="ignore" \
        stonith-timeout="120s"
```

ASCS sample instance profile

```
SAPSYSTEMNAME = VSE
SAPSYSTEM = 00
INSTANCE_NAME = ASCS00
DIR_CT_RUN = $(DIR_EXE_ROOT)/run
DIR_EXECUTABLE = $(DIR_INSTANCE)/exe
SAPLOCALHOST = SAPVIPE
#-----------------------------------------------------------------------
# SAP Message Server parameters are set in the DEFAULT.PFL
#-----------------------------------------------------------------------
ms/standalone = 1
```

```
ms/server_port_0 = PROT=HTTP,PORT=81$$
#---------------------------------------------------------------------
# SAP Enqueue Server
#---------------------------------------------------------------------
enque/table_size = 4096
rdisp/enqname = $(rdisp/myname)
enque/snapshot_pck_ids = 100
ipc/shm_psize_34 = 0
enque/server/replication = true
enque/server/max_requests = 1000
enque/enrep/stop_timeout_s = 0
enque/enrep/stop_retries = 0
```

ERS sample instance profile

```
SAPSYSTEM = 01
SAPSYSTEMNAME = VSE
INSTANCE_NAME = ERS01
#---------------------------------------------------------------------
# Special settings for this manually set up instance
#---------------------------------------------------------------------
DIR_EXECUTABLE = $(DIR_INSTANCE)/exe
DIR_CT_RUN = /usr/sap/VSE/SYS/exe/run
#---------------------------------------------------------------------
# Settings for enqueue monitoring tools (enqt, ensmon)
#---------------------------------------------------------------------
enque/process_location = REMOTESA
rdisp/enqname = $(rdisp/myname)
#---------------------------------------------------------------------
# standalone enqueue details from ASCS instance
#---------------------------------------------------------------------
ASCSID = 00
ASCSHOST = SAPVIPE
enque/serverinst = $(ASCSID)
enque/serverhost = $(ASCSHOST)
#---------------------------------------------------------------------
# HA polling
#---------------------------------------------------------------------
#enque/enrep/hafunc_implementation = script
#enque/enrep/poll_interval = 10000
#enque/enrep/hafunc_init =
#enque/enrep/hafunc_check = $(DIR_EXECUTABLE)/enqtest.sh
```

ERS sample START profile

```
SAPSYSTEMNAME = VSE
SAPSYSTEM = 01
INSTANCE_NAME = ERS01
#---------------------------------------------------------------------
# Special Settings for this manually set up instance
#---------------------------------------------------------------------
ASCSID = 00
DIR_CT_RUN = /usr/sap/VSE/SYS/exe/run
DIR_EXECUTABLE = $(DIR_INSTANCE)/exe
_PF = $(DIR_PROFILE)/VSE_ERS01_SAPASCS2
SETENV_00 = LD_LIBRARY_PATH=$(DIR_EXECUTABLE)
SETENV_01 = PATH=$(DIR_INSTANCE)/exe:%(PATH)
#---------------------------------------------------------------------
# Copy SAP Executables
```

```
#-----------------------------------------------------------------------
_CPARG0 = list:$(DIR_EXECUTABLE)/ers.lst
Execute_00 = immediate $(DIR_EXECUTABLE)/sapcpe$(FT_EXE) $(_CPARG0) pf=$(_PF)
#-----------------------------------------------------------------------
# Start enqueue replication server
#-----------------------------------------------------------------------
_ER = er.sap$(SAPSYSTEMNAME)_$(INSTANCE_NAME)
Execute_01 = immediate rm -f $(_ER)
Execute_02 = local ln -s -f $(DIR_EXECUTABLE)/enrepserver $(_ER)
Restart_Program_00 = local $(_ER) pf=$(_PF) NR=$(ASCSID)
```

DI sample instance profile

```
SAPSYSTEMNAME = VSE
SAPSYSTEM = 00
INSTANCE_NAME = D00
DIR_CT_RUN = $(DIR_EXE_ROOT)/run
DIR_EXECUTABLE = $(DIR_INSTANCE)/exe
exe/saposcol = $(DIR_CT_RUN)/saposcol
rdisp/wp_no_dia = 10
rdisp/wp_no_btc = 3
exe/icmbnd = $(DIR_CT_RUN)/icmbnd
icm/server_port_0 = PROT=HTTP,PORT=80$$
SAPFQDN = sse.ea.emc.com
SAPLOCALHOSTFULL = $(SAPLOCALHOST).$(SAPFQDN)
ipc/shm_psize_10 = 136000000
ipc/shm_psize_40 = 112000000
rdisp/wp_no_vb = 1
rdisp/wp_no_vb2 = 1
rdisp/wp_no_spo = 1
enque/process_location = REMOTESA
enque/serverhost = SAPVIPE
enque/serverinst = 00
enque/deque_wait_answer = TRUE
enque/con_timeout = 2000
enque/con_retries = 60
```