Disaster Recovery of Tier 1 Applications on VMware vCenter™ Site Recovery Manager™

vmware®

VMware, Inc.
3401 Hillview Ave
Palo Alto, CA 94304
www.vmware.com

Contents

Table of Figures

1. Introduction

1.1 Overview

Tier 1 applications, also known as business critical applications (BCAs), need to have robust disaster recovery. In legacy physical server infrastructures, protecting Tier 1 applications against disasters was a monumental task due to the following reasons:

- Identical dedicated hardware was needed in the recovery site. The hardware was expensive and underutilized.
- Recovery site applications, operating systems, and hardware needed to be kept up to date.
- Any testing would require downtime of the production environment.

1.2 Purpose

This document shows how Tier 1 applications can be easily protected using VMware vCenter™ Site Recovery Manager™. Techniques to protect against disaster for common BCAs such as Microsoft Exchange, Microsoft SQL Server, SAP, and Oracle Databases are presented.

1.3 Target Audience

This document is intended for SAP, Exchange, SQL Server, Oracle, and VMware administrators responsible for managing the operations of business critical software applications on the VMware vSphere® platform.

1.4 Disaster Recovery Terminology

Business Continuity Plan (BCP) is a plan to continue operations if a place of business (office, work site, or data center) is affected by adverse physical conditions, such as a storm, fire, or crime. Such a plan typically explains how the business would recover its operations or move operations to another location.

Disaster recovery (DR) involves a set of policies and procedures to enable the recovery or continuation of vital technology infrastructure and systems following a natural or human-induced disaster. DR is a subset of business continuity and focuses on the technology systems supporting critical business functions, as opposed to business continuity.

Recovery Point Objective (RPO) is the maximum tolerable period in which data might be lost from an IT service due to a major incident.

Recovery Time Objective (RTO) is the targeted duration of time and a service level within which a business process must be restored after a disaster or other disruption to avoid unacceptable consequences.

2. Overview

2.1 vCenter Site Recovery Manager

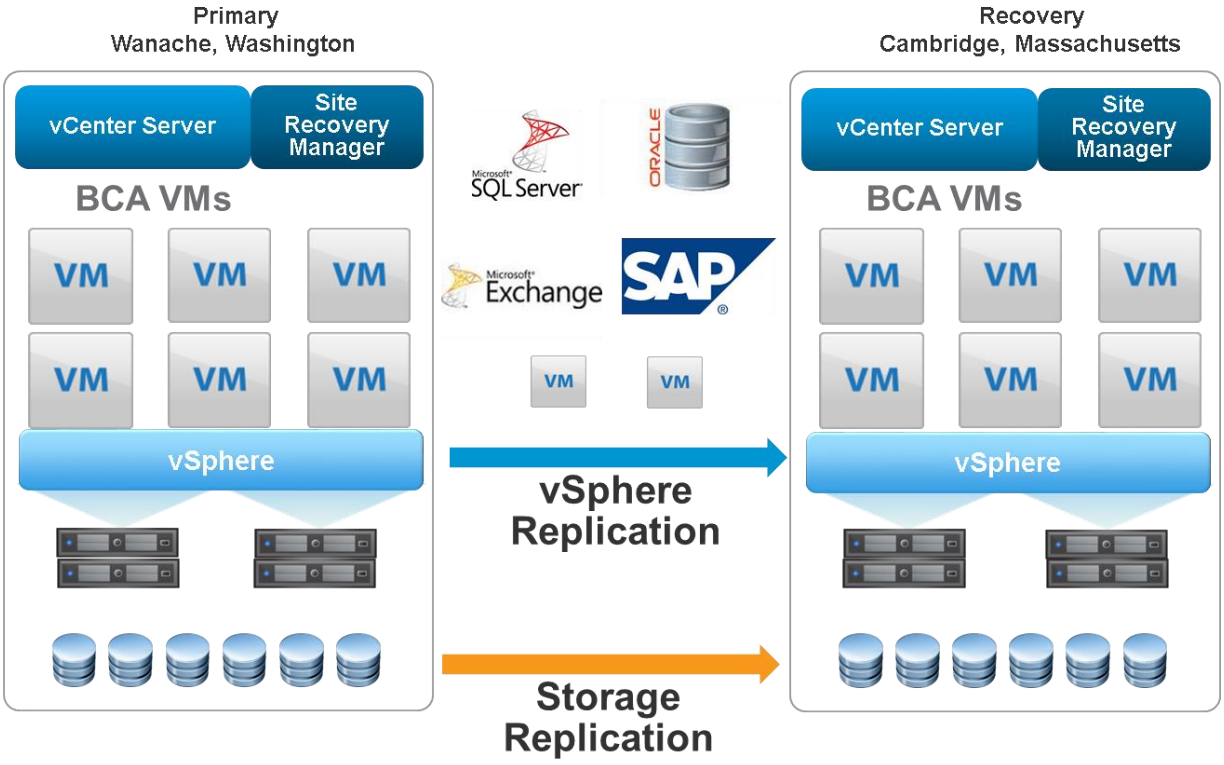
With the advent of virtualization and the concept of encapsulation of virtual machines, replicating entire business critical workloads has been greatly simplified. Virtual machines are represented as a set of files that can be replicated to the recovery site easily and re-instantiated on different hardware. vCenter Site Recovery Manager leverages the unique aspects of virtual machines combined with replication management and workflows to automate disaster recovery for BCA.

- Automated disaster recovery failover
  - Initiate recovery plan execution from the VMware vSphere Web Client with a single click of a button.
  - Halt replication and promote replicated virtual machines for fastest possible recovery.
  - Execute user-defined scripts and pauses during recovery.
- Planned migration and disaster avoidance
  - Graceful shutdown of protected virtual machines at the original site.
  - Replication synchronization of protected virtual machines prior to migration to avoid data loss.
  - Restart of protected virtual machines in an application consistent state.
- Seamless workflow automation with centralized recovery plans
  - Create and manage recovery plans directly from the vSphere Web Client.
  - Predefine the boot sequence of virtual machines for automated recovery.
  - Reconfigure IP addresses upon failover at the subnet or individual address level.

3. Overview of the Site Recovery Manager Environment

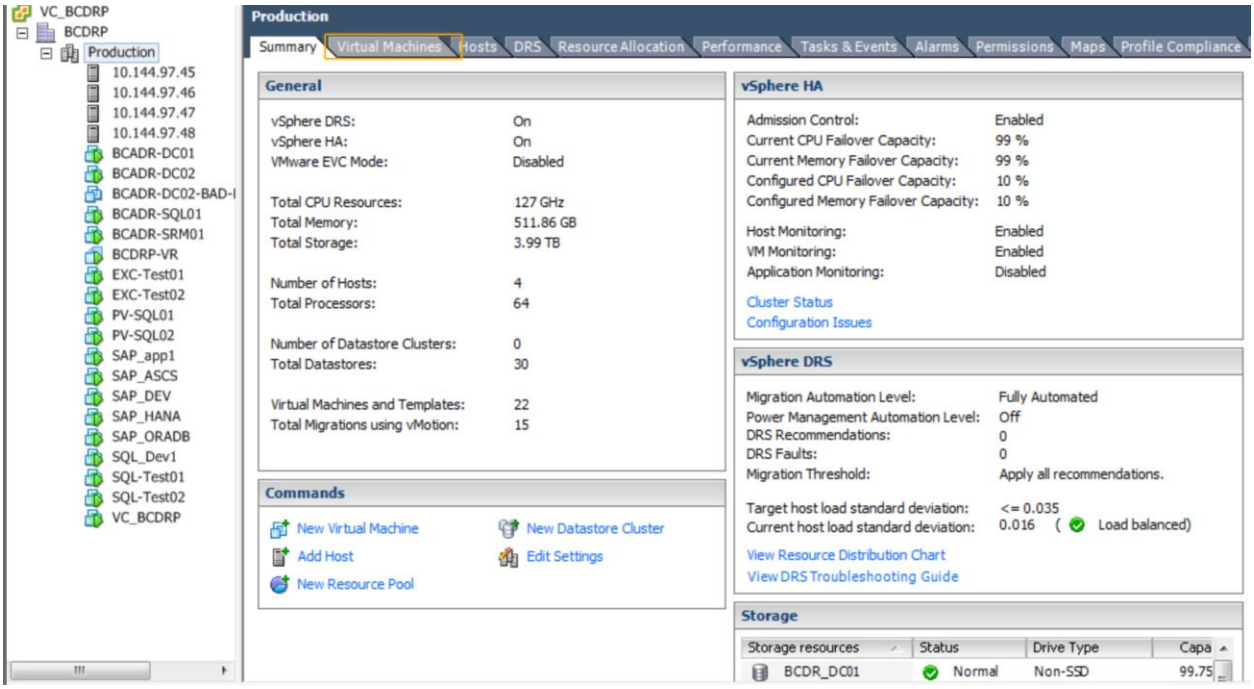
For this paper, consider a sample environment with two sites, the primary site and the recovery site. As shown in the following figure, the primary site is located in Wanache, Washington State and the recovery site is located in Cambridge, Massachusetts. Subject matter experts in SAP, Oracle, SQL, and Exchange have created instances of these applications in the primary site. Infrastructure components such as domain controllers and Site Recovery Manager servers are also deployed.

Figure 1. Primary and Secondary Sites



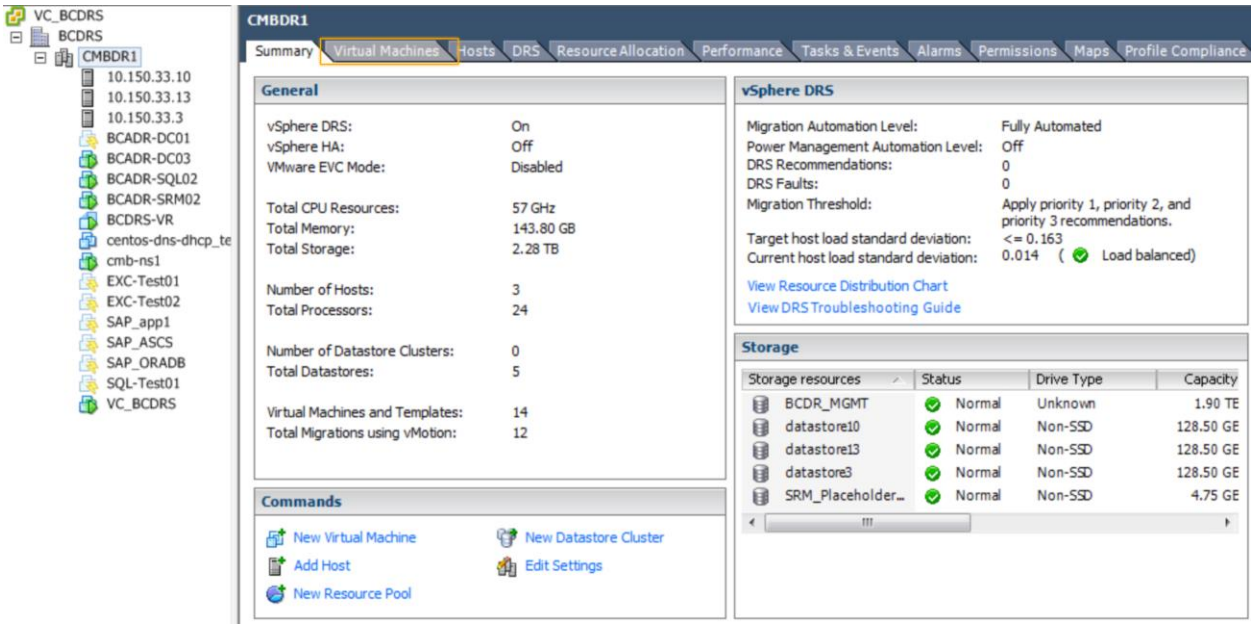
The following figure shows details about the environment and the virtual machines representing the applications.

Figure 2. Primary Site with Infrastructure and BCA Components



The recovery site has three vSphere servers and other local systems relating to infrastructure, including domain controllers, Site Recovery Manager servers, and other local applications. The remote site has placeholder virtual machines created by Site Recovery Manager for all the protected virtual machines from the primary site.

Figure 3. Recovery Site with Site Recovery Manager and Infrastructure Components



4. Replication

The backbone for disaster recovery is replication of the protected workloads from the primary to the recovery site. Replication can be synchronous or asynchronous.

Synchronous replication is used in active-active environments that have zero RPO requirements. The scope of synchronous replication is within metro areas, as there is a requirement to have latencies below 10 ms. Every write in the primary site is acknowledged only when it has been written to both sites. This solution is typically very expensive and is used by organizations that have zero RPO as a requirement. Examples of synchronous replication are EMC vPLEX and NetApp MetroClusters.

Asynchronous replication is used for the majority of disaster recovery deployments. The RPO for asynchronous replication can range from a few minutes to hours depending on the customer requirements. This replication is usually constrained by the bandwidth between the primary and recovery sites. Site Recovery Manager is typically used with asynchronous replication. The following types of replication are used in Site Recovery Manager deployments:

- *Storage replication* is array-based replication provided by the storage vendor. This requires the same type and vendor of the storage solution on both the primary and recovery sites. Examples are EMC SRDF and NetApp SnapMirror. These mature solutions have been used over the past few decades and provide granular features and robust recovery mechanisms. Storage replication with Site Recovery Manager has a storage replication adapter (SRA) that is provided by the storage vendor. Site Recovery Manager communicates with the storage array using this adapter and uses it as a proxy for replication using the array.
- *VMware vSphere Replication™* is a VMware solution that can replicate at the individual virtual machine level. The storage backing the virtual machines can be of any type, including local storage. The primary and recovery sites can have different types of storage. vSphere Replication can replicate all the virtual machine disks or a chosen subset of disks. vSphere Replication works at the VMware kernel level with any changes to storage captured and replicated. The RPO for vSphere Replication can range from 15 minutes to 24 hours based on customer requirements. It requires vSphere Replication appliances that are deployed by Site Recovery Manager. These appliances coordinate the replication of changes between the primary and the recovery site.

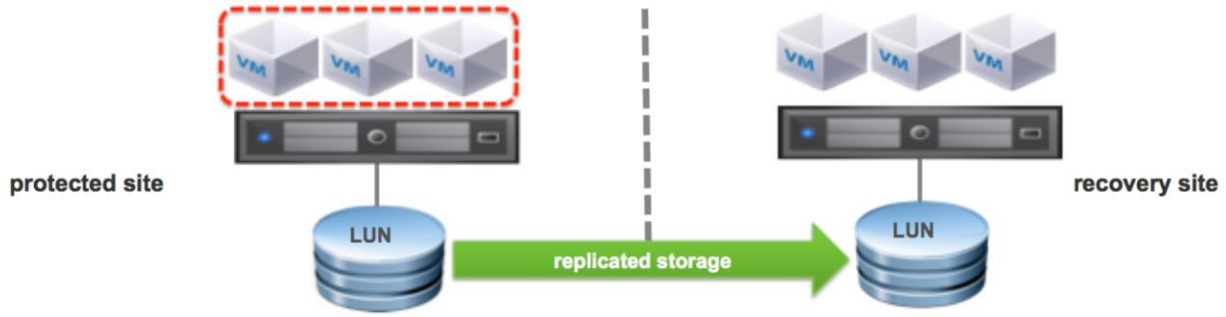
Both of these replication methodologies use crash-consistent recovery. The application has not been quiesced and hence the recovery is akin to that of a machine following a power outage. There is a small probability of data corruption for database type workloads. In this situation, application-level replication can protect against the risk of data corruption. See Section 6, Recovery Plans.

5. Protection Groups

A protection group is a group of virtual machines that fail over together to the recovery site. Protection groups contain virtual machines with data that has been replicated by array-based replication or by vSphere Replication. A protection group typically contains virtual machines that are related in some way, as in these examples:

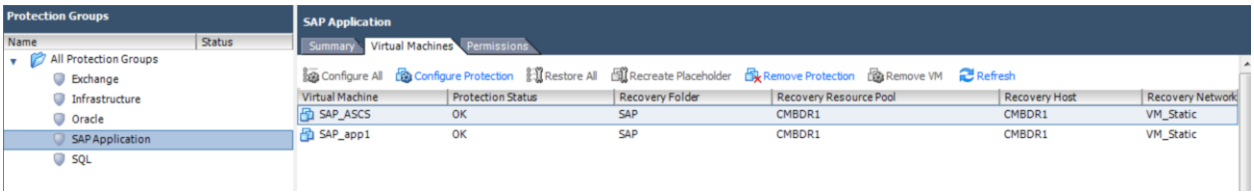
- A three-tier application (application server, database server, and web server).
- Virtual machines with virtual machine disk files that are part of the same datastore group.

Figure 4. Protection Groups



The following figure shows protection groups for the sample use case. Protection groups for Exchange, Infrastructure, Oracle and SQL were created. The details for the SAP Application protection group are shown.

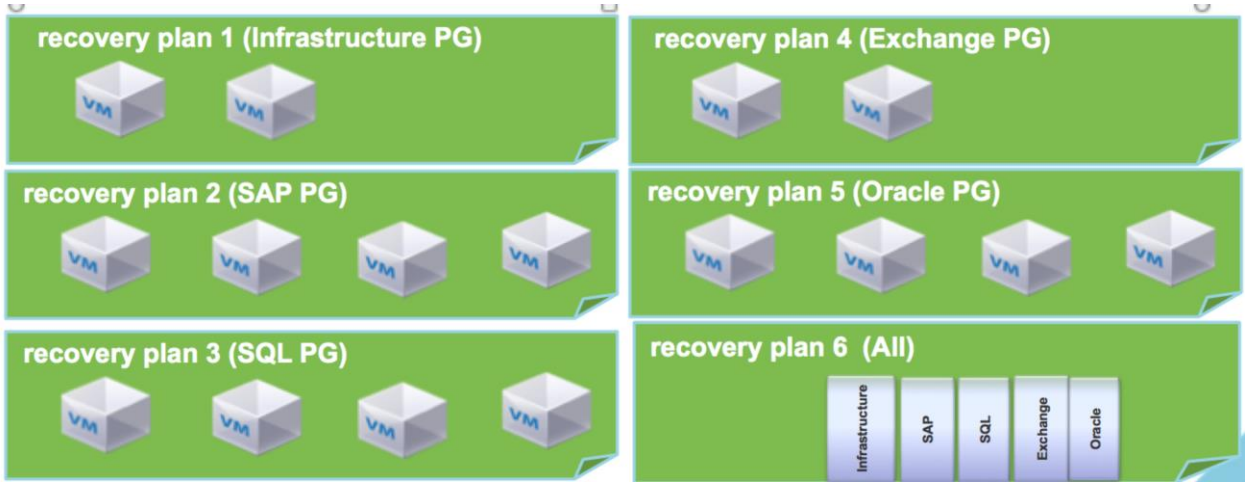
Figure 5. SAP Application Protection Group



6. Recovery Plans

Protection groups are the building blocks of recovery plans. A protection group can be included in multiple recovery plans. Each recovery plan is a sequence of steps executed to recover virtual machines in a specified sequence with specified priority.

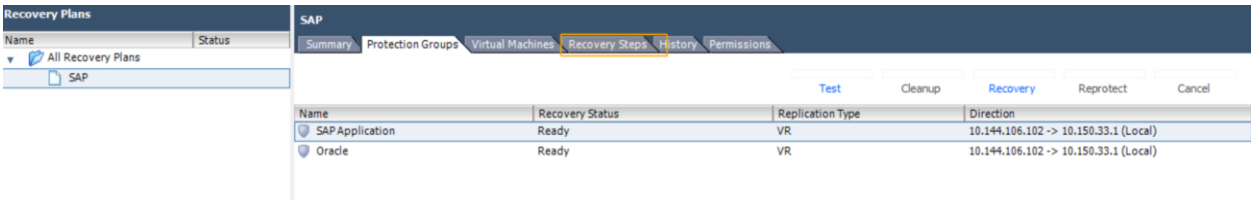
Figure 6. Recovery Plans



Each individual application can have its own recovery plan and can be recovered independently of others. Most applications have dependencies on infrastructure components such as Active Directory and DNS. The infrastructure protection group needs to be included in most application recovery plans so that the application is usable after recovery. A recovery plan for the entire site (All) can include all applications. The recovery plan allows setting of priorities for applications to create an order for the recovery process.

The following figure shows the recovery plan for the SAP application. SAP application requires the Oracle database in addition to the application servers, so two protection groups are included in the plan.

Figure 7. Recovery Plan for the SAP Application



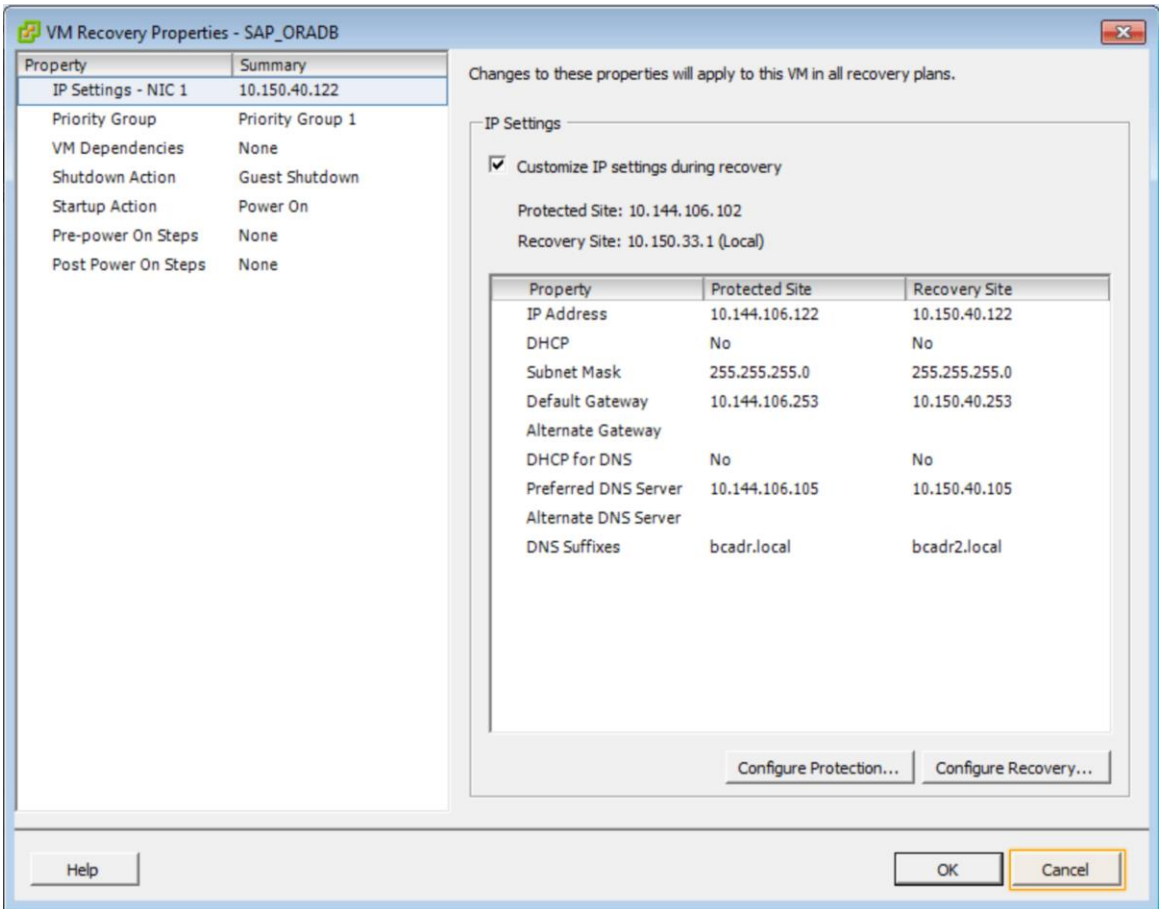
6.1.1 IP Addressing at the Recovery Site

The recovery site can use the same IP addresses as the primary site or use different IP addresses. The appropriate choice depends on the customer infrastructure:

- Some customers leverage capabilities such as stretched VLANs or VLANs that can be relocated to have the same IP address in the primary and the recovery site.
- It is common to have a completely different set of IP addresses at the recovery site for the virtual machines. Site Recovery Manager recovery plans provide the capability to automatically reassign the IP addresses for the virtual machines prior to recovery.

The following figure shows an example IP mapping for the SAP Oracle VM at the recovery site.

Figure 8. IP Customization During Recovery



7. Site Recovery Manager Use Cases

Site recovery helps automate the cumbersome process of disaster recovery planning, testing, and recovery. It provides the workflow, runbooks, and automation to make the process of recovery seamless.

Site Recovery Manager typically addresses the following use cases:

- Disaster recovery
  - Testing
  - Recovery from actual disasters
  - Re-protect
- Planned migration

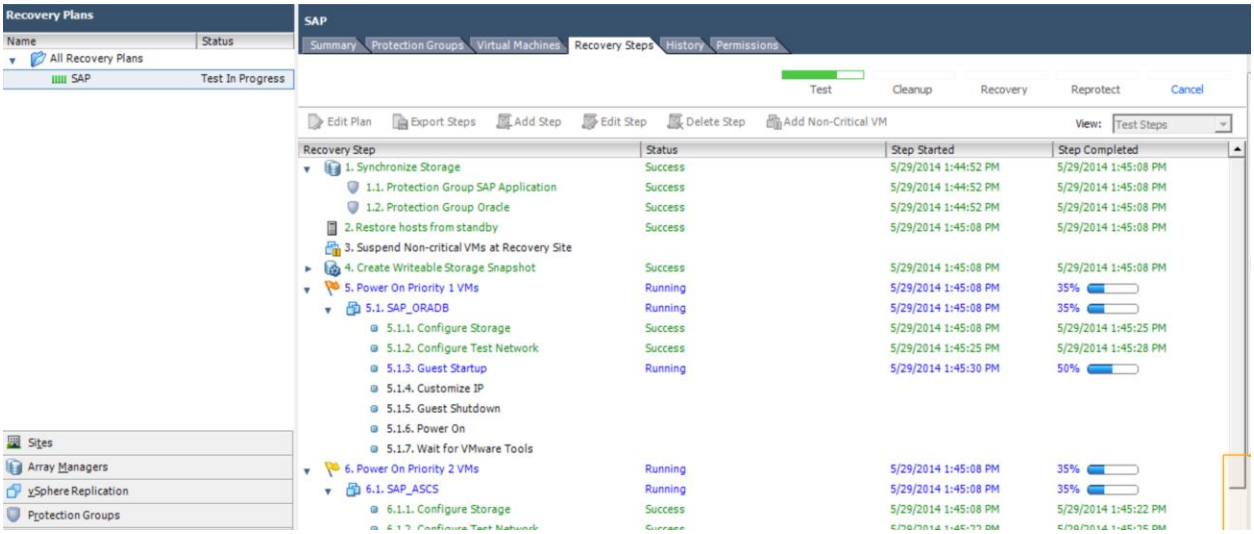
7.1 Protecting SAP Applications with Site Recovery Manager

A major advantage of Site Recovery Manager is the ability to test DR of BCAs with no impact to the running production applications. This is an important differentiator for Site Recovery Manager compared with DR for legacy physical environments. Site Recovery Manager provides the ability to test recovery in an isolated environment without any downtime and helps tune the recovery by learning from the tests. When an actual disaster happens, the RTO is optimized. The processes for testing and actual recovery are identical, except that in the disaster the primary site goes down and actual recovery occurs. In a test scenario, recovery happens in isolation with the primary site remaining up.

The environment that was set up for BCAs was tested with Site Recovery Manager. DR tests are executed by running recovery plans that had been set up for the different applications. The following figure shows the testing process for the SAP recovery plan. For a real recovery the recovery use case is run instead of the test use case.

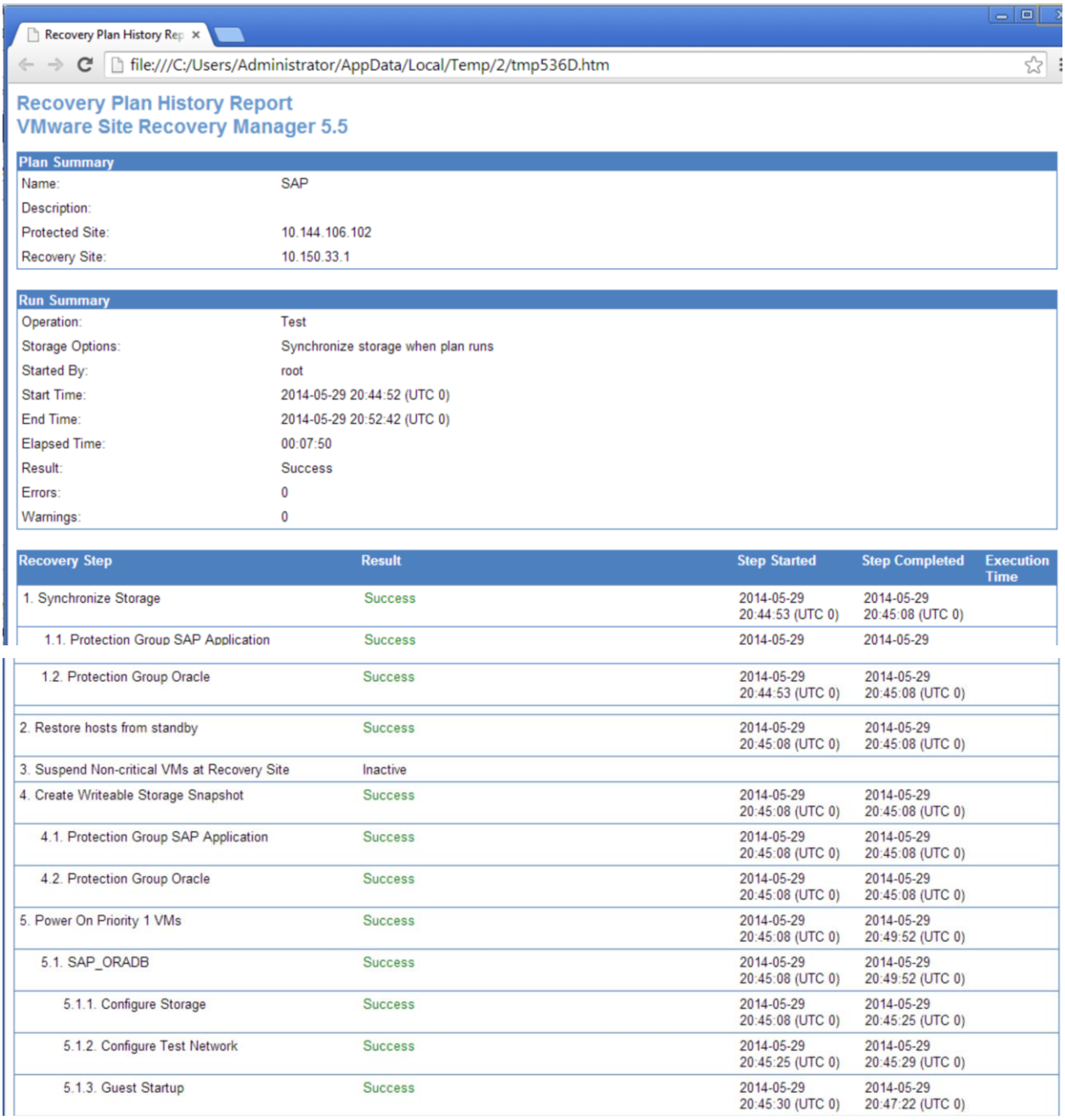
As the following figure shows, there are multiple steps in the recovery. Some steps are done in parallel, while others are done based on priority to optimize recovery time.

Figure 9. SAP Recovery with Site Recovery Manager



When the recovery is complete, history reports are available to view all the steps and their duration. These reports can help in tuning the recovery process.

Figure 10. History Report for SAP Recovery



Recovery Plan History Rep ×

file:///C:/Users/Administrator/AppData/Local/Temp/2/tmp536D.htm

Recovery Plan History Report
VMware Site Recovery Manager 5.5

Plan Summary

Name:	SAP
Description:	
Protected Site:	10.144.106.102
Recovery Site:	10.150.33.1

Run Summary

Operation:	Test
Storage Options:	Synchronize storage when plan runs
Started By:	root
Start Time:	2014-05-29 20:44:52 (UTC 0)
End Time:	2014-05-29 20:52:42 (UTC 0)
Elapsed Time:	00:07:50
Result:	Success
Errors:	0
Warnings:	0

Recovery Step	Result	Step Started	Step Completed	Execution Time
1. Synchronize Storage	Success	2014-05-29 20:44:53 (UTC 0)	2014-05-29 20:45:08 (UTC 0)	
1.1. Protection Group SAP Application	Success	2014-05-29	2014-05-29	
1.2. Protection Group Oracle	Success	2014-05-29 20:44:53 (UTC 0)	2014-05-29 20:45:08 (UTC 0)	
2. Restore hosts from standby	Success	2014-05-29 20:45:08 (UTC 0)	2014-05-29 20:45:08 (UTC 0)	
3. Suspend Non-critical VMs at Recovery Site	Inactive			
4. Create Writeable Storage Snapshot	Success	2014-05-29 20:45:08 (UTC 0)	2014-05-29 20:45:08 (UTC 0)	
4.1. Protection Group SAP Application	Success	2014-05-29 20:45:08 (UTC 0)	2014-05-29 20:45:08 (UTC 0)	
4.2. Protection Group Oracle	Success	2014-05-29 20:45:08 (UTC 0)	2014-05-29 20:45:08 (UTC 0)	
5. Power On Priority 1 VMs	Success	2014-05-29 20:45:08 (UTC 0)	2014-05-29 20:49:52 (UTC 0)	
5.1. SAP_ORADB	Success	2014-05-29 20:45:08 (UTC 0)	2014-05-29 20:49:52 (UTC 0)	
5.1.1. Configure Storage	Success	2014-05-29 20:45:08 (UTC 0)	2014-05-29 20:45:25 (UTC 0)	
5.1.2. Configure Test Network	Success	2014-05-29 20:45:25 (UTC 0)	2014-05-29 20:45:29 (UTC 0)	
5.1.3. Guest Startup	Success	2014-05-29 20:45:30 (UTC 0)	2014-05-29 20:47:22 (UTC 0)	

After testing, the Site Recovery Manager clean-up function makes it easy to clean up the test environment. The temporary storage, virtual machine, and networking resources used for the tests are cleaned up within a few minutes.

For an actual disaster, clean-up is not done. Instead, a reprotect process is initiated when the primary site infrastructure becomes available. The reprotect process reverses the direction of replication and the roles of the primary and recovery sites.

7.2 Protecting Oracle Databases Using Oracle Data Guard and Site Recovery Manager

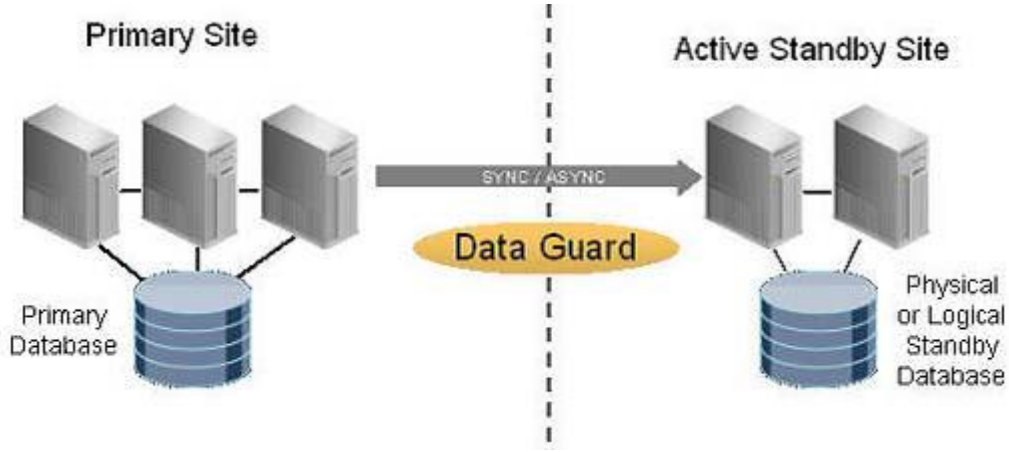
Oracle Data Guard provides the management, monitoring, and automation software infrastructure to create and maintain one or more standby databases to protect Oracle data from failures, disasters, errors, and data corruption. A *standby database* is a database copy that is in a constant state of recovery. The standby database is a near-real-time replica of the primary database, which is the standby's source. Data can be transported to the standby using either of the following methods:

- Archived redo logs can be copied to the standby site and subsequently applied to the standby database.
- Data can be transmitted through the network transport method, which transports the individual change vectors of the redo log stream. After the transport, the redo logs are applied through the continuous recovery process.

Data Guard is unique among Oracle replication solutions in supporting both synchronous (zero data loss) and asynchronous (near-zero data loss) configurations. To maintain high availability (HA) for mission-critical applications, administrators can choose either manual or automatic failover of production to a standby system if the primary system fails. A Data Guard system can be set up in one of the following modes:

- *Maximum Performance* allows the standby to lag behind the primary so it never affects the performance of the primary. It uses the transfer of individual archive logs and subsequent application during the recovery process or the asynchronous copy of the change vectors in the redo stream.
- *Maximum Availability* is the most common Data Guard usage mode because it allows for the synchronous transportation of the redo stream. Under these circumstances, no data is lost, because the primary and standby are always in sync. Only if a certain configurable timeout threshold is exceeded does the standby lag behind the primary.
- *Maximum Protection* requires the redo stream transfer to always remain synchronous. The primary cannot continue unless the redo transfer is acknowledged at the standby site. If the preconfigured lag time is exceeded, the primary database crashes to guarantee that no deviation exists between the standby's and the primary's log stream.

Figure 11. Oracle Data Guard



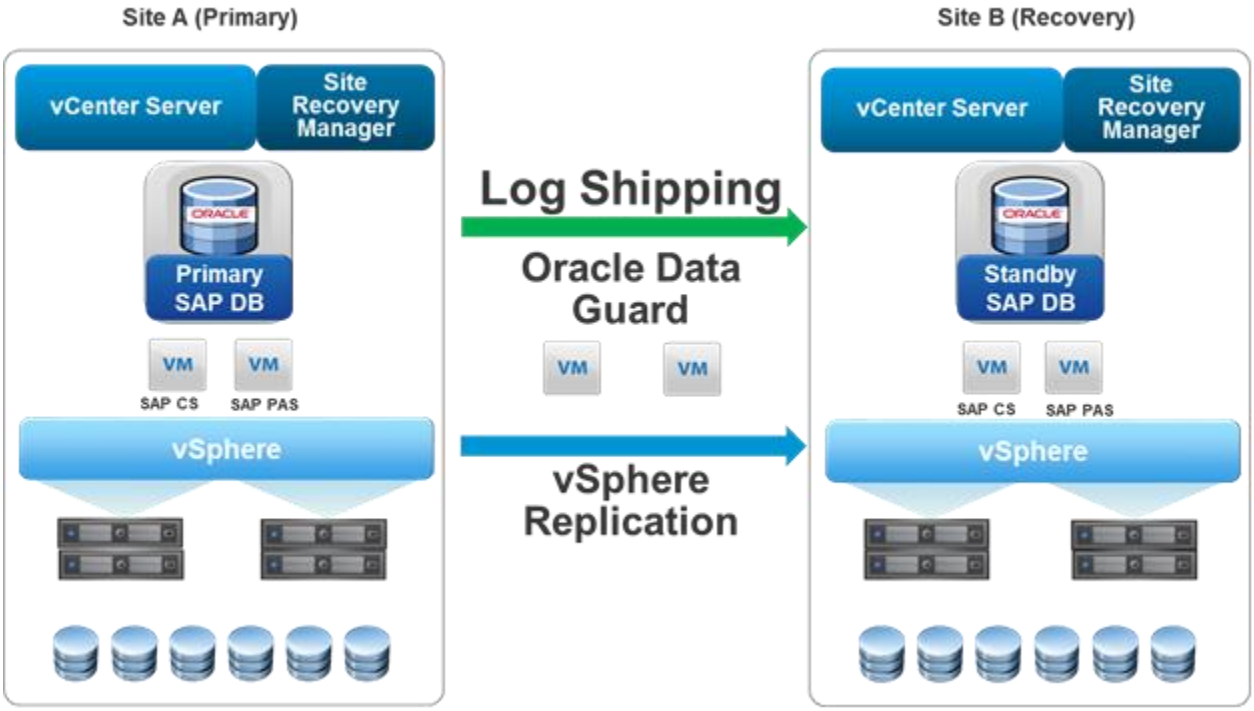
7.2.1 Combining Site Recovery Manager, vSphere Replication, and Oracle Data Guard

There are many different approaches to developing and implementing comprehensive DR architectures. Some involve storage replication, while others rely only on host-based replication. Some solutions primarily use replication provided by the application vendors, whereas others trust in the database vendor's ability to replicate the data.

Many of the tools and technologies do a similar job. The best approach is to determine the application's DR SLA requirements and then apply the appropriate tools and technologies to meet those requirements.

The solution shown in the following figure combines vSphere Replication to protect a set of SAP application VMs and Oracle Data Guard to protect Oracle database VMs. The SAP data is protected by vSphere Replication and the database data is protected by Data Guard using the Maximum Availability Data Guard mode.

Figure 12. Oracle Data Guard and vSphere Replication Solution



This solution elegantly integrates vSphere Replication with Data Guard by allowing Site Recovery Manager to remain the controlling application. When the simulated DR event occurs, Site Recovery Manager orchestrates the entire system failover. The standby database is failed over first, as an embedded callout script is used to execute the appropriate Data Guard commands to perform the switchover of the primary to the standby. After the database switchover is complete and the callout script exits, the next step in the orchestration occurs, which is the failover of the VMs that run the SAP application.

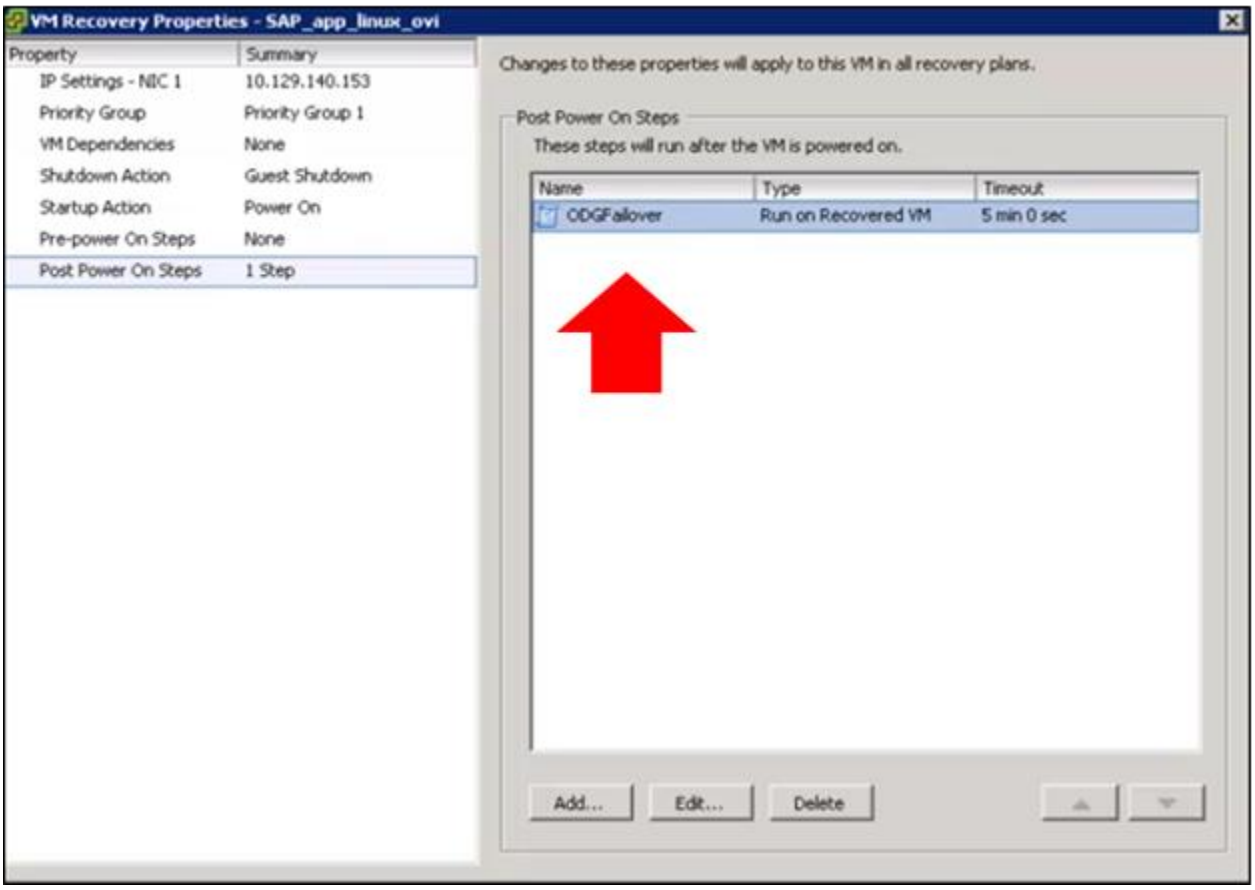
The beauty of this orchestration is that if a step within the orchestration fails, the entire workflow reverses course. Because Data Guard uses a switchover instead of a failover, even the database role reversal will smoothly return to the original state.

7.2.2 Testing Site Recovery Manager vSphere Replication

Testing of the Site Recovery Manager, vSphere Replication, and Oracle Data Guard solution follows this process:

1. Initially, the Oracle primary database is at Site A and the standby database is at Site B with the Data Guard setup.
2. The SAP application is connected to the primary database at Site A.
3. The SAP application and central services VM are replicated to Site B using vSphere Replication.
4. When the simulated DR event occurs, the database administrator switches over the Oracle standby to a primary using a Site Recovery Manager callout script from the SAP application VM or the vSphere Replication server at Site B. See the following figure.
5. Connect or resume the SAP application to the Oracle database at Site B.

Figure 13. Site Recovery Manager Call Out Script



The following script is used to perform the database switchover.

```
~ # cat odgfail.sh
#! /bin/sh
################################################################################
#
#    file name   : odgfail.sh
#    location    : /scripts
#    called from : Application VM/VR server on Site B
################################################################################
#
echo  "Job `basename $0`: started at `date`"
#
# Set up standard ORACLE environment variables
ORACLE_SID=stdby; export ORACLE_SID
ORACLE_BASE=/oracle; export ORACLE_BASE
ORACLE_HOME=/oracle/PRD/102_64; export ORACLE_HOME
PATH=/oracle/PRD/102_64/bin:.:/oracle/PRD:/usr/sap/PRD/SYS/exe/run:/usr/kerberos/bin:/
usr/local/bin:/bin:/usr/bin:/usr/X11R6/bin;export PATH
LD_LIBRARY_PATH=/usr/sap/PRD/SYS/exe/run:/oracle/client/10x_64/instantclient; export
LD_LIBRARY_PATH
#
# Failover to Standby
$ORACLE_HOME/bin/sqlplus /nolog <<EOFarch1
connect / as sysdba
--shutdown Primary database(in case of RAC, shutdown all RAC instances)
--Initiate failover to Standby Database:
ALTER DATABASE RECOVER MANAGED STANDBY DATABASE FINISH FORCE;
--Convert the physical standby database to the production role:
ALTER DATABASE COMMIT TO SWITCHOVER TO PRIMARY;
--Comment/Uncomment either of the 2 sets of commands below
--If the database was never opened read-only since the last time it was started,
--open new production database via:
ALTER DATABASE OPEN;
--If the physical standby database has been opened in read-only mode since the last
time it was started,
--shutdown standby database and restart it
--SHUTDOWN IMMEDIATE
--STARTUP pfile=initSTDBY.ora
exit
EOFarch1
echo  "Job `basename $0`: ended at `date`"
```

```
########################## end of script
~ #
```

Application needs for DR are unique to each company and the reason that the particular application is being used. Some applications do not need to be part of a DR plan, while others require zero recovery point and zero recovery time.

7.3 Protecting Microsoft Applications with Site Recovery Manager

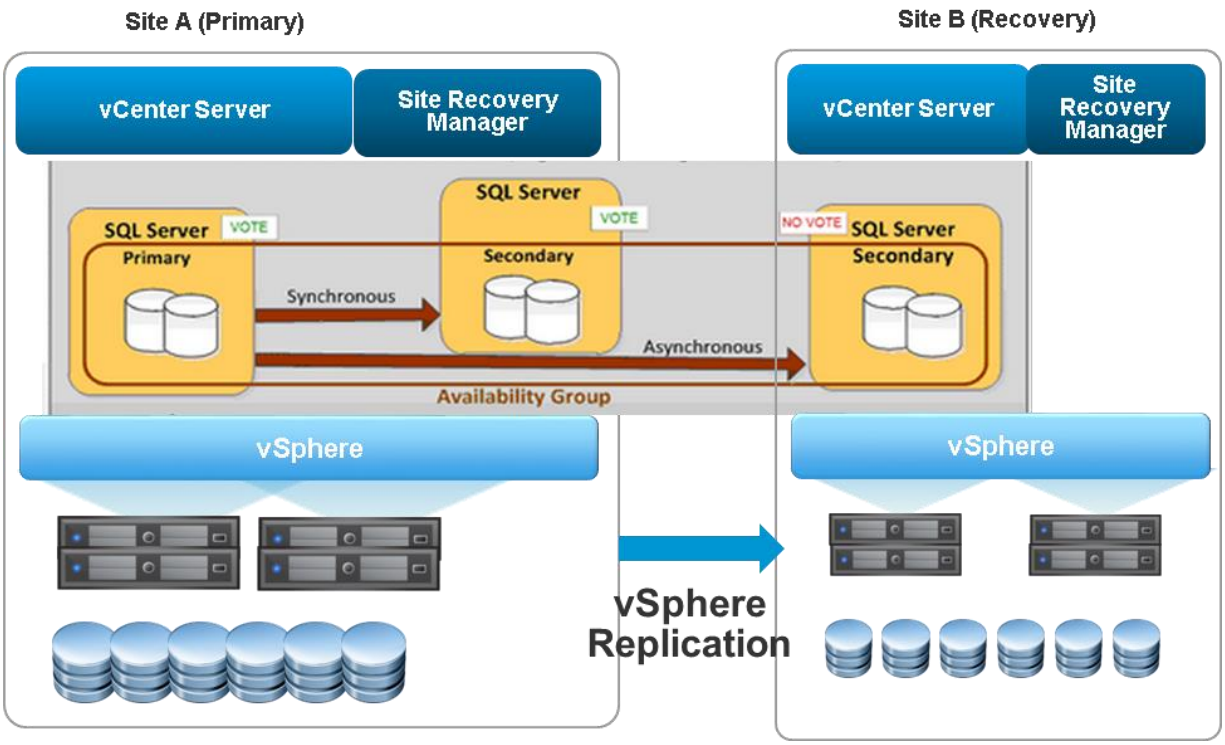
Microsoft applications have features that can be used for disaster recovery with Site Recovery Manager. Risk averse customers who require application consistent recovery can leverage application consistent copies for recovery.

7.3.1 Microsoft SQL Server

Many BCAs use Microsoft SQL server as a backend database. Business critical data that exists in these databases must be protected. SQL provides application consistent protection using the concept of Always on Availability Groups (AAG). Synchronous copies of the database are available locally and asynchronous copies are available for remote sites. The synchronous copy is for local recovery. The asynchronous AAG copy (with callout script) can be used with Site Recovery Manager for application consistent SQL database recovery.

The recovery site has a local Microsoft SQL server that mounts the asynchronous copy of the database in read/write mode in the event of a disaster. All other virtual machines that are required for the application are usually stateless and can use vSphere replication, as shown in the following figure.

Figure 14. SQL Always on Availability groups with Site Recovery Manager



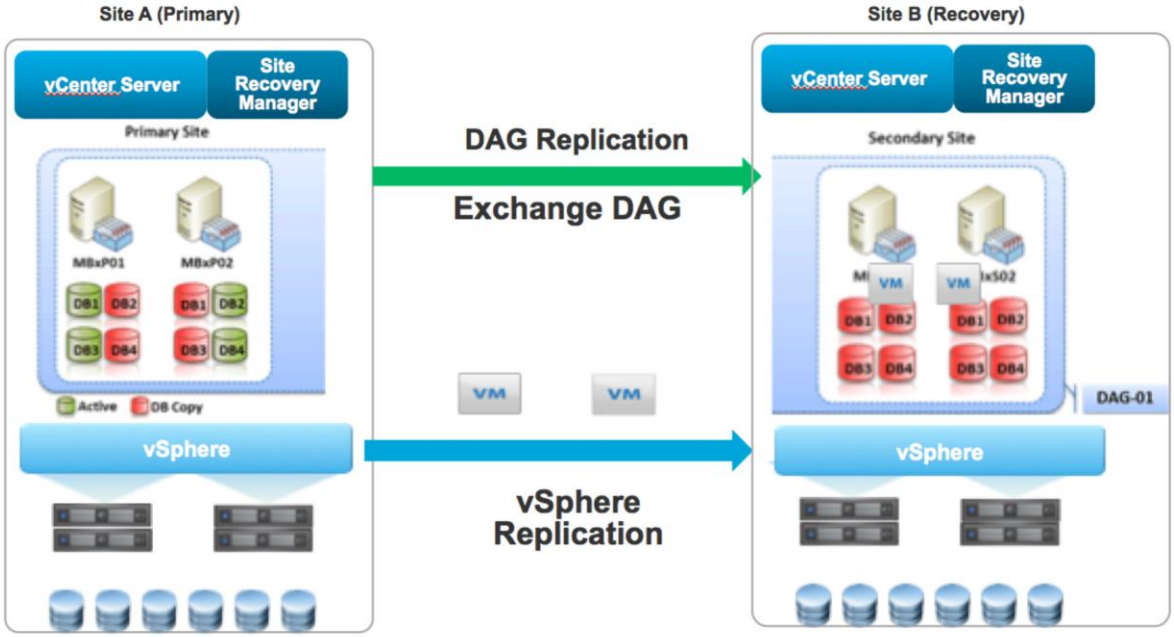
7.3.2 Microsoft Exchange Server

Microsoft Exchange provides application consistent protection using Exchange Database Availability Group (DAG). Synchronous copies of the database are available locally and asynchronous copies are

available for remote sites. The synchronous copy is for local recovery. The asynchronous DAG copy (with callout script) can be used with Site Recovery Manager for application consistent SQL database recovery.

The recovery site has a local Exchange server that mounts the asynchronous DAG copy in read/write mode in the event of a disaster. All other virtual machines that are required for Exchange are stateless and can use vSphere replication as shown in the following figure.

Figure 15. Microsoft Exchange DAG with Site Recovery Manager

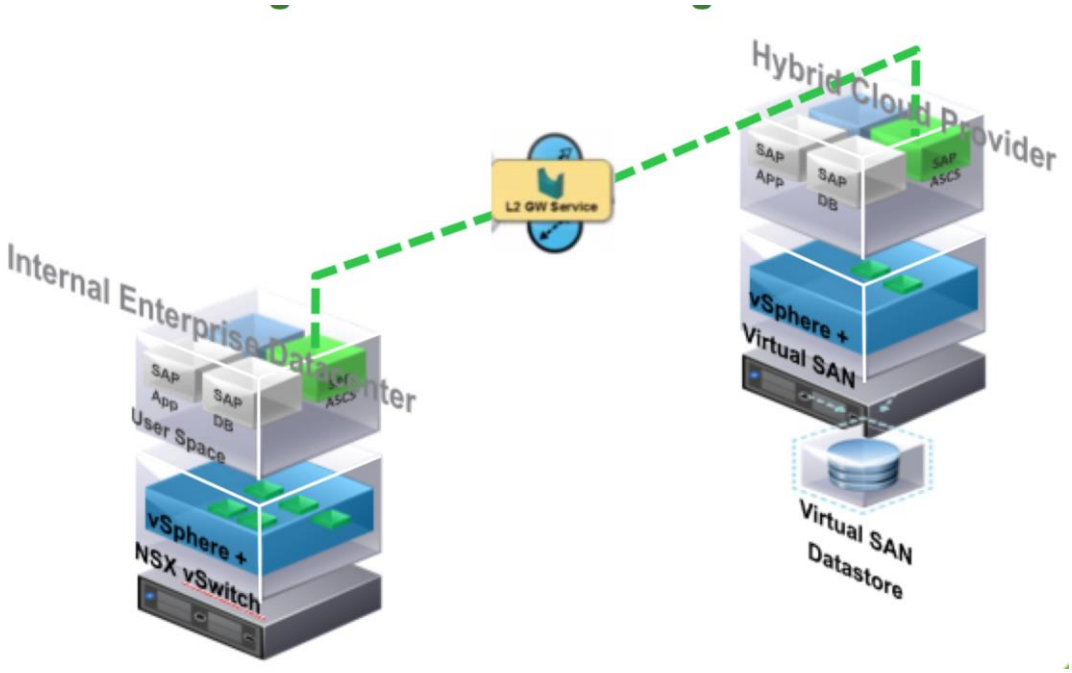


7.4 Tier 1 Application Planned Migration to Cloud

Planned migration is a relatively new use case for Site Recovery Manager. Customers can migrate their Tier 1 workloads to the recovery or cloud provider site in a planned manner. This is helpful when preparing for an upcoming threat, such as a hurricane, or for a data center migration to a different location or cloud provider.

The following figure shows the logical infrastructure for a planned migration of SAP from an enterprise data center to a cloud provider virtual data center.

Figure 16. Logical Planned Migration Infrastructure



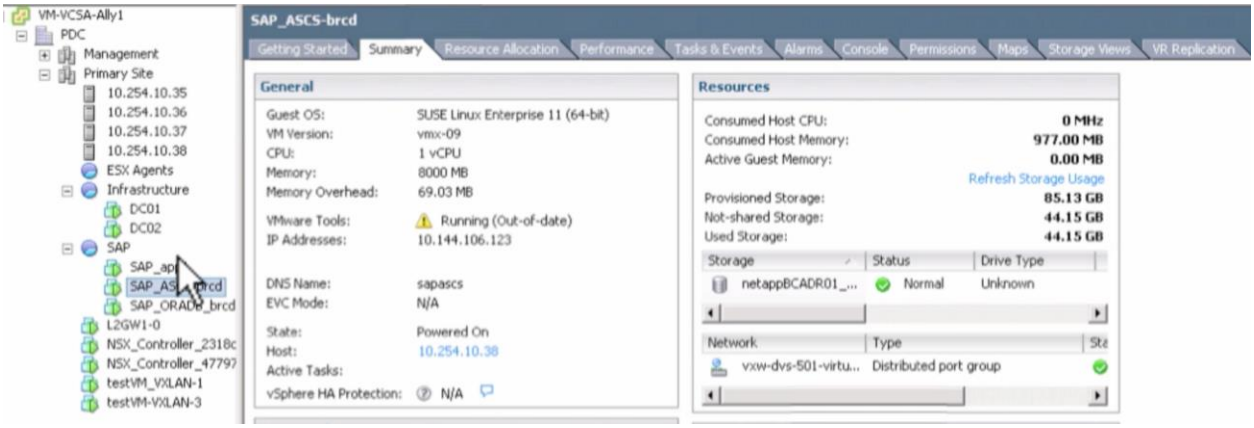
The internal enterprise data center hosts the SAP application stack that will be migrated to a cloud provider. SAP is running in a VMware software-defined data center (SDDC) with VMware NSX™ networking in the internal data center with traditional shared storage. The hybrid cloud virtual data center has vSphere for computing, VMware vCloud® Networking and Security™ for networking, and VMware Virtual SAN™ for storage.

The NSX Layer 2 gateway provides the capability to extend the virtual machine VLANs across the two data centers. This allows the solution to keep the same IP addresses for the virtual machines through the planned migration.

Site Recovery Manager servers are set up on both sides with vSphere Replication to replicate the SAP VMs to be migrated. The destination datastore for the migration is the Virtual SAN datastore in the hybrid cloud data center. The three SAP virtual machines, SAP App, DB, and ASCS, are replicated to the cloud data center using vSphere Replication.

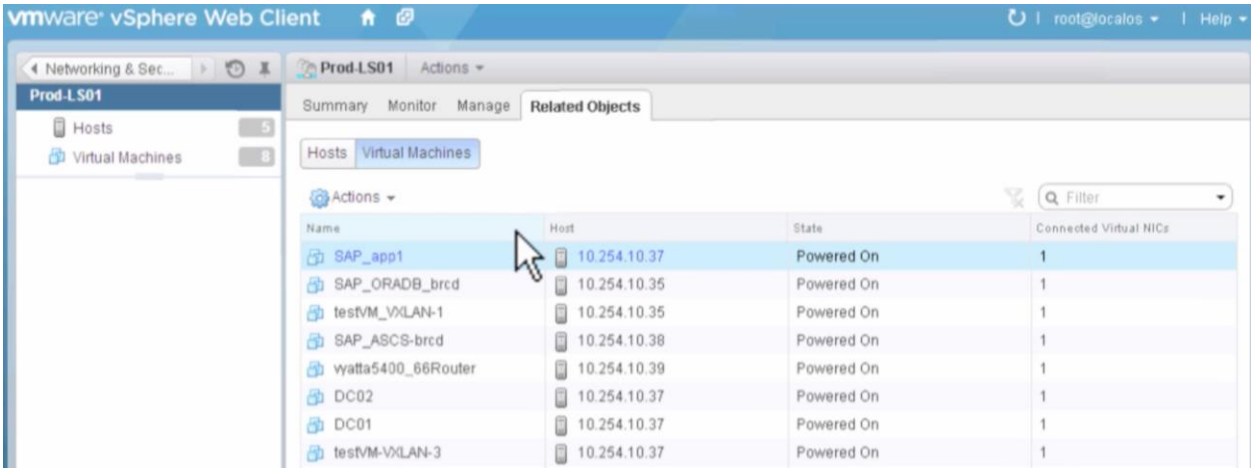
The following figure shows details for the components relating to the planned migration. The virtual machines listed under **Infrastructure** are the domain controllers that are required for access to SAP. One of the domain controllers, DC01, is also replicated to the cloud provider site.

Figure 17. Primary Data Center Components for Planned Migration



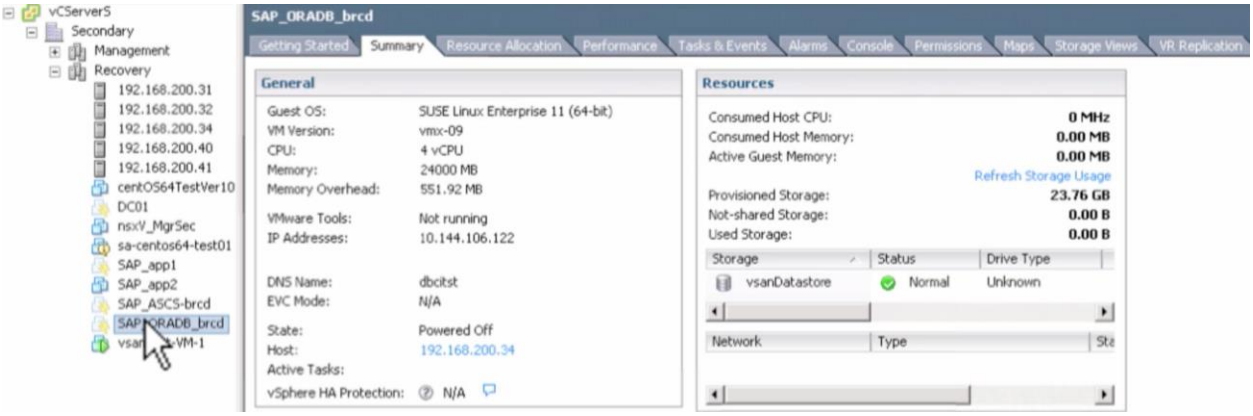
The following figure shows how the virtual machines related to SAP are connected to the NSX logical switch in the primary data center.

Figure 18. NSX Logical Switch Connected to SAP Virtual Machines



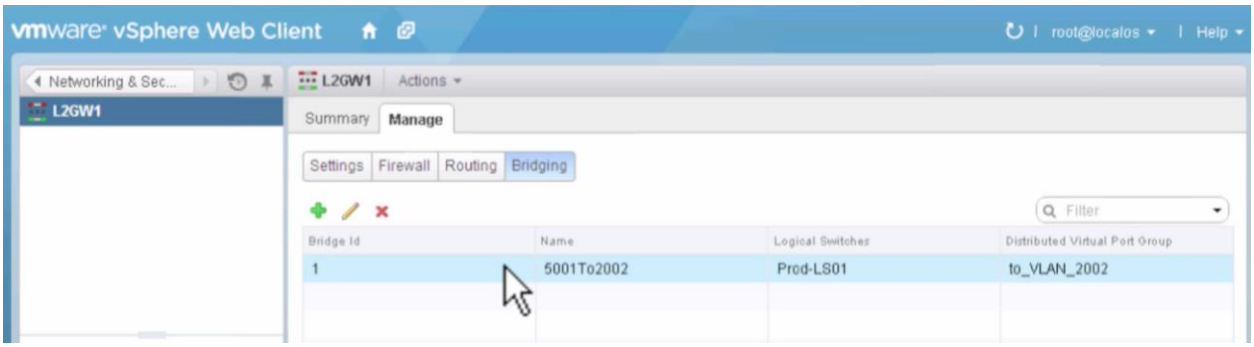
The following figure shows details for the cloud data center and its virtual components. The placeholder virtual machines for the SAP and its related infrastructure that were created by Site Recovery Manager are also shown.

Figure 19. Cloud Provider Virtual Data Center



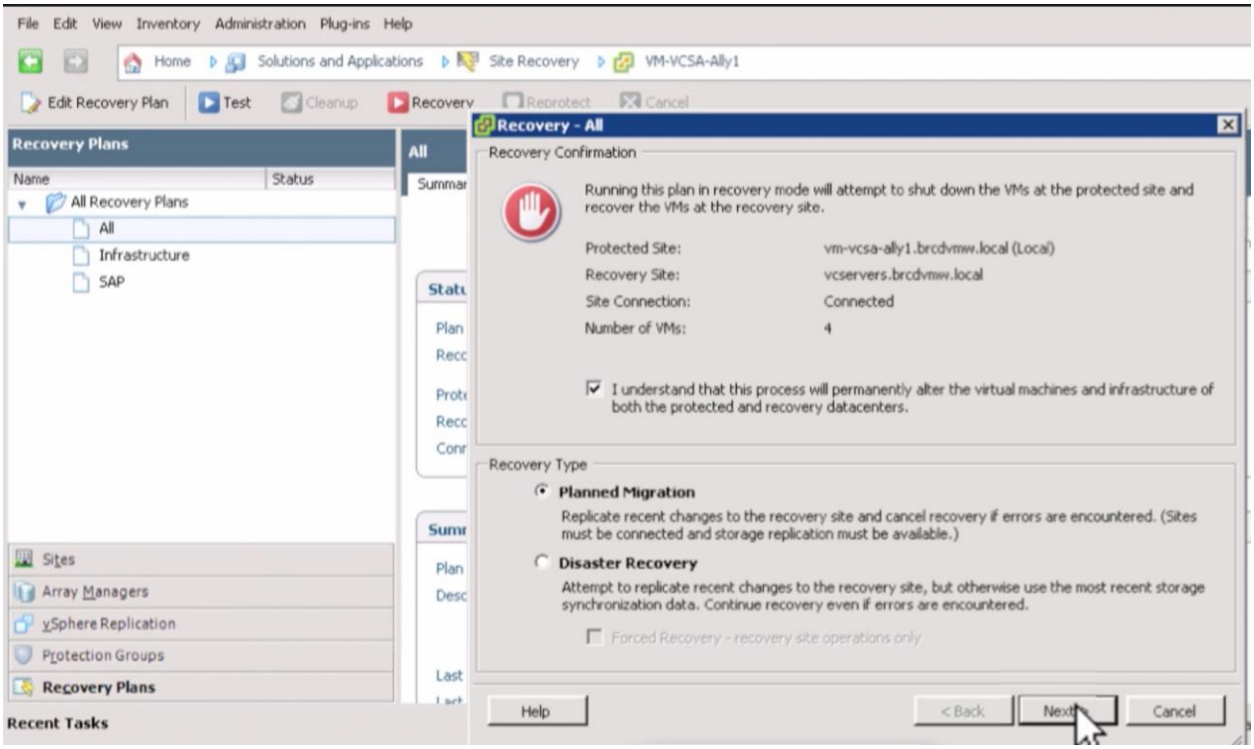
The VMware NSX Layer 2 gateway is used to extend the virtual machine VLANs across the wide area network connecting the two data centers. There is no need to change IP addresses for the virtual machines in the destination data center. The following figure shows the configuration of the NSX gateway to bridge the VLANs.

Figure 20. NSX Layer 2 Gateway Configuration



With the two sites synchronized through vSphere Replication and the network successfully extended, Site Recovery Manager can now be used to perform a planned migration to the hybrid cloud provider data center for the SAP application components.

Figure 21. Initiation of Planned Migration to Cloud Provider

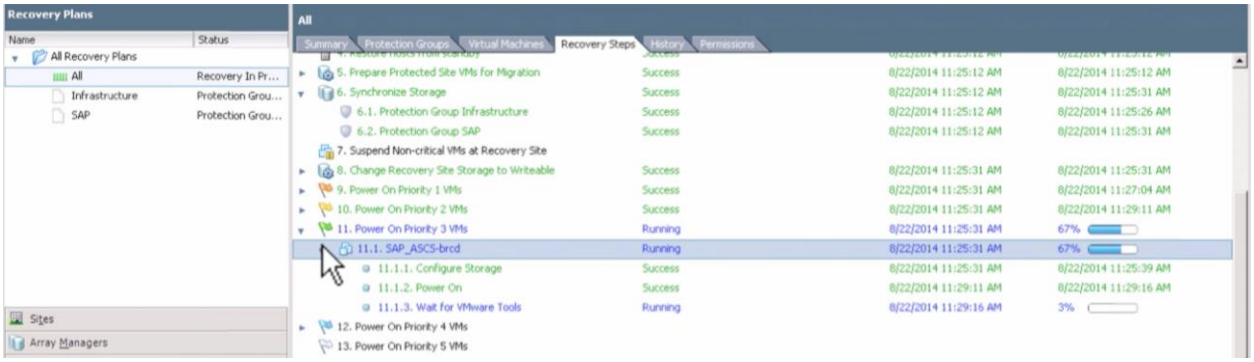


In a planned migration, the virtual machines are systematically shut down before the final synchronization process for the virtual machines. The virtual machines are brought down in reverse priority order, as shown in the following figure.

Figure 20. Shutdown of Virtual Machines During a Planned Migration

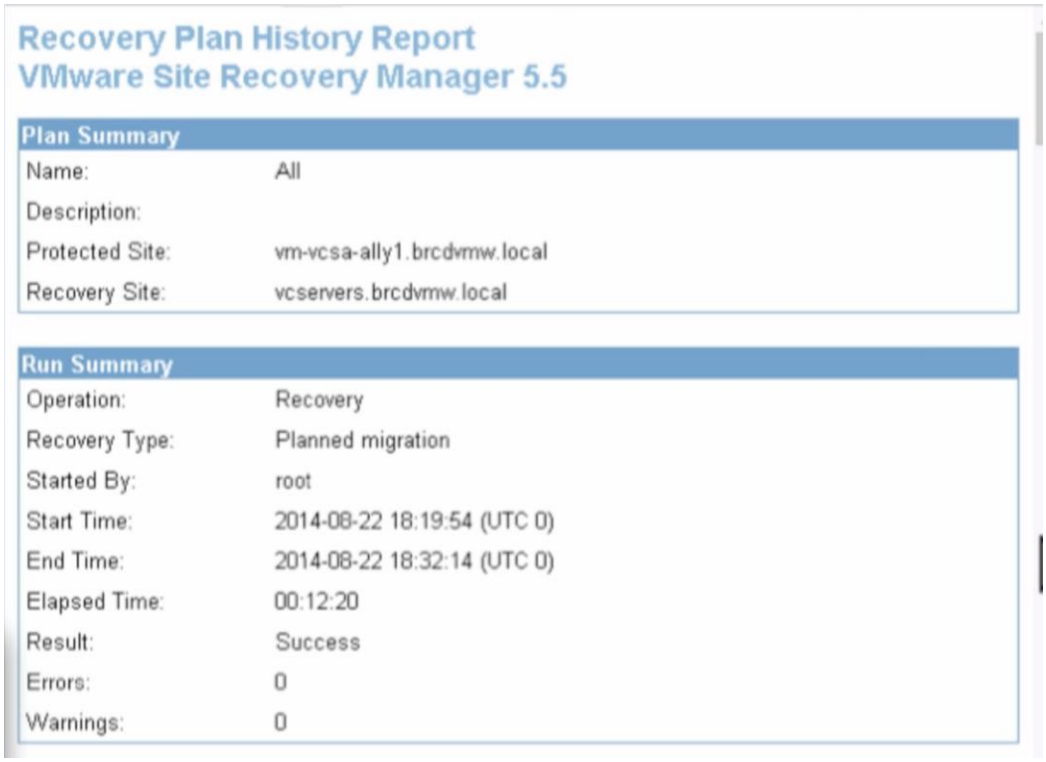
In the next phase after the final sync is complete, the virtual machines are automatically powered up in priority order. The planned migration is complete when all the virtual machines have come up and the SAP application is validated by the application owners.

Figure 21. Virtual Machines Powered On at the Cloud Provider Data Center



The following history report shows that the SAP planned migration was completed successfully in less than 15 minutes.

Figure 22. Planned Migration History Report



Recovery Plan History Report
VMware Site Recovery Manager 5.5

Plan Summary

Name:	All
Description:	
Protected Site:	vm-vcsa-ally1.brcdvmw.local
Recovery Site:	vcservers.brcdvmw.local

Run Summary

Operation:	Recovery
Recovery Type:	Planned migration
Started By:	root
Start Time:	2014-08-22 18:19:54 (UTC 0)
End Time:	2014-08-22 18:32:14 (UTC 0)
Elapsed Time:	00:12:20
Result:	Success
Errors:	0
Warnings:	0

8. Conclusion

Site Recovery Manager enhances the capability of businesses to protect business critical applications from disasters. It also provides the ability for customers to migrate applications temporarily or permanently between data centers with minimal downtime and impact to business. The offline testing, workflow, and runbook capabilities provided by Site Recovery Manager are indispensable for businesses seeking to minimize their RTO and RPO for business critical applications.

9. Acknowledgements

The following are the authors involved in creating this document

- Kannan Mani
- Mohan Potheri

Other major contributors and reviewers are:

- Vas Mitra
- Sudhir Balasubramanian
- Deji Akomolafe
- Don Sullivan
- Scott Salyer