DELL Technologies

Best Practices

Disaster Recovery for Oracle Real Application Clusters with VMware Virtual Volumes

Abstract

The paper discusses the implementation of Oracle Real Application Clusters with shared vmdks on vVol datastores and protecting that environment with VMware Site Recovery Manager.

September 2021

Dell EMC PowerMax Engineering

Revisions

Date	Description
September 2021	Initial release

Acknowledgments

Author: Drew Tonnesen, drew.tonnesen@dell.com

Examples provided in this paper cover methods for performing various VMware vSphere activities using PowerMax systems with Oracle and Dell EMC software. These examples were developed for laboratory testing and may need tailoring to suit other operational environments. Any procedures outlined in this guide should be thoroughly tested before implementing in a production environment.

Table of contents

Revisions....2
Acknowledgments....2
Table of contents....3
Executive Summary....5
1 SRDF Overview....6
2 Oracle Real Application Clusters....7
2.1 Environment....7
2.1.1 Configuration....10
3 VMware SRM setup....14
3.1 SRM restrictions....14
3.1.1 VMware....14
3.1.2 Dell EMC....14
3.2 SRM configuration for Oracle RAC....15
3.2.1 Storage policy mappings....15
3.2.2 Create a Protection Group and Recovery Plan....15
4 DRS Affinity Rule....22
4.1 VM/Host Groups....22
4.2 VM/Host Rules....26
5 VMware SRM testfailover....29
5.1 Oracle RAC testfailover without modification....29
5.2 Oracle RAC testfailover with modification....32
5.2.1 Oracle RAC networking....33
5.2.2 SRM testfailover networking....33
5.2.3 Oracle RAC – Public IP, VIP, SCAN, and GNS....38
5.2.4 Public IP and Post Power On Steps changes....41
6 VMware PowerCLI RAC testing....50
6.1 Managed vs unmanaged snapshots....50
6.1.1 Remote replication....50
6.2 Environment....51
6.2.1 Storage policies....52
6.2.2 Recovery site....55
6.2.3 Testfailover with PowerCLI....57
6.2.4 Mounting VVOL_REFRESH ASM disk group....64
6.2.5 Test cleanup....66

7 Conclusion 68
A Error/Warning messages 69
A.1.1 Testfailover 69
B Technical support and resources 70
B.1 Related resources 70
B.1.1 Dell EMC 70
B.1.2 VMware 70
B.1.3 Oracle 70

Executive Summary

VMware Virtual Volumes are an integration and management framework (referred to as the vVol framework) that moves management from the datastore to the virtual machine. vVols virtualize storage, enabling a more efficient operational model that is optimized for virtualized environments and centered on the application instead of the infrastructure. vVols map virtual disks, configuration files, clones, and snapshots, directly to virtual volume objects on a storage system, i.e., volumes. This mapping allows VMware to offload more storage operations to the PowerMax, such as VM snapshots.

vVols simplify the delivery of storage capabilities to individual applications by providing finer control of hardware resources and allowing the use of native array-based data services such as SnapVX and SRDF® at the VM level.

Enterprise-level applications like Oracle Real Application Clusters which previously required the use of raw device mappings (RDMs) to take advantage of array technologies can now be managed with greater ease on vVols, yet retain the benefits of the PowerMax like QoS and replication.

A critical component of Oracle databases is disaster recovery. Traditionally, DBAs have used RMAN and Oracle Standby/Data Guard to enable disaster recovery, but as software solutions they are limited by their networks, CPU, and memory. SRDF, the gold-standard in array replication, provides a robust solution to disaster recovery regardless of host resources. With VASA 3, VMware SRM supports vVol replication utilizing SRDF/A to protect Oracle databases from loss of data.

This paper will show how to configure a four-node Oracle RAC environment on VMware virtual volumes and replicate it with VMware SRM. Details will include how to enable an Oracle test environment to run at the recovery site using SRM testfailover with and without reconfiguration. Additionally, the final chapter will cover how to use VMware PowerCLI and a partially replicated VM to refresh an Oracle RAC test cluster at the recovery site without reconfiguring the Oracle Grid software.

This paper does not include details on the VMware Virtual Volumes implementation on the PowerMax except as needed to elucidate the solutions contained herein. Though there is some overlap with other papers by necessity, for a complete understanding of the vVol implementation on the PowerMax including the use of VMware SRM, please see the white paper Using VMware vSphere Virtual Volumes 2.0 and VASA 3.0 with Dell EMC VMAX and PowerMax.

Audience

This technical white paper is intended for VMware, Oracle, and/or PowerMax administrators responsible for deploying Oracle Real Application Clusters (RAC) on VMware vSphere 7.x, utilizing VMware Virtual Volumes, on PowerMax.

1 SRDF Overview

The Dell EMC Symmetrix® Remote Data Facility (SRDF) family of products offers a range of PowerMax-based disaster recovery, parallel processing, and data migration solutions.

SRDF disaster recovery solutions are based on active remote mirroring and dependent-write consistent copies of data maintained at one or more remote locations. A dependent write is a write operation that cannot be issued by an application until a prior, related write I/O operation completes. Dependent-write consistency is required to ensure transactional consistency when the applications are restarted at the remote location.

SRDF configurations with vVols require at least two PowerMax systems. In two-site, non-SRDF/Metro configurations, these systems are usually also known as the primary and the secondary systems. Both sites can be located in the same room, in different buildings within the same campus, or hundreds to thousands of kilometers apart.

The PowerMax enables asynchronous replication with virtual volumes through SRDF with a maximum Recovery Point Objective (RPO) of 300 seconds[1]. Replication is controlled at the VM-level and fully initiated and controlled from within VMware vCenter. VMware SRM is supported on the PowerMax beginning with vSphere 7 and VASA 3.

[1] Asynchronous SRDF is the only mode supported with vVols.

2 Oracle Real Application Clusters

Oracle Real Application Clusters, or RAC, is a feature of the Oracle database which permits clustering of multiple instances accessing a single database on shared devices. In order to enable Oracle RAC, Oracle Clusterware is required to provide the infrastructure to share the database. The Clusterware also manages network and service resources. Although Oracle Clusterware can be used with non-RAC databases, it is mostly associated with RAC. The Clusterware, when used in conjunction with Oracle Automatic Storage Management (ASM), make up the Oracle Grid Infrastructure. ASM allows the user to create and make available a pool of storage through ASM disk groups. Using the Oracle Grid Infrastructure negates the requirement for third-party clustering software such as Veritas, though it is still possible to use it.

Oracle RAC provides both scalability and high availability. Additional nodes (instances) can be added at any time to increase processing power and resiliency. Oracle also offers Extended RAC which as its name suggests extends the instances across a campus cluster. Dell EMC can enable this using SRDF/Metro on the PowerMax, though not supported for vVols.

Oracle RAC can be run on both physical and virtual hosts, including VMware. While ASM provides for sharing devices across RAC instances, when using VMware, it is necessary to make the devices accessible to all VMs in the cluster through the multi-writer flag, regardless of the type of VMware storage used, e.g., VMFS, RDMs, or vVols.

2.1 Environment

In the current use case, vVols is being used as storage. As noted, it is necessary to set the multi-writer flag for each vmdk involved with the Oracle database to enable sharing across the VMs. Oracle does not require any additional configuration such as SCSI 2 or 3 reservations to operate. Dell EMC recommends deploying the Clusterware files, OCR and voting, on their own ASM disk group using “Normal” redundancy, while using “External” redundancy for the other disk groups that will house the Oracle RAC database. The RAC environment contains four nodes, each with two non-shared vVols for the OS and Oracle software, and the remaining seventeen devices for the RAC database shared by all using the multi-writer attribute. Figure 1 below shows a local disk (Hard disk 2) and a shared disk (Hard disk 3) with multi-writer attribute set for one of the RAC nodes. Note that both disks have the same VM storage policy and Replication Group.

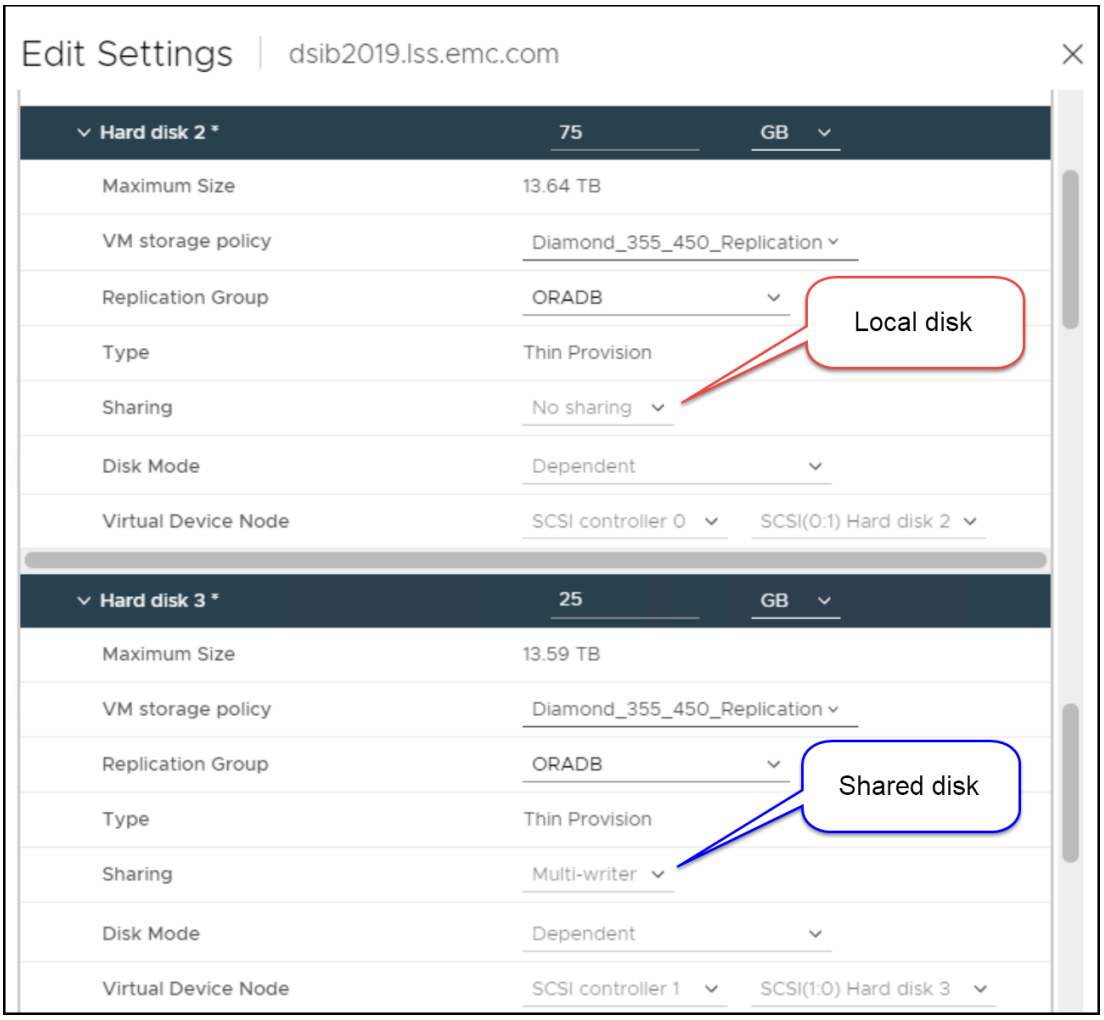

Figure 1. Local and shared disk for RAC node

The policy, shown in Figure 2, has two attributes assigned to it from the VASA Provider: Diamond service level and Asynchronous replication. Note that although each of the other three nodes share seventeen disks between them, all nodes must be assigned to the same storage policy, not just the first RAC node.

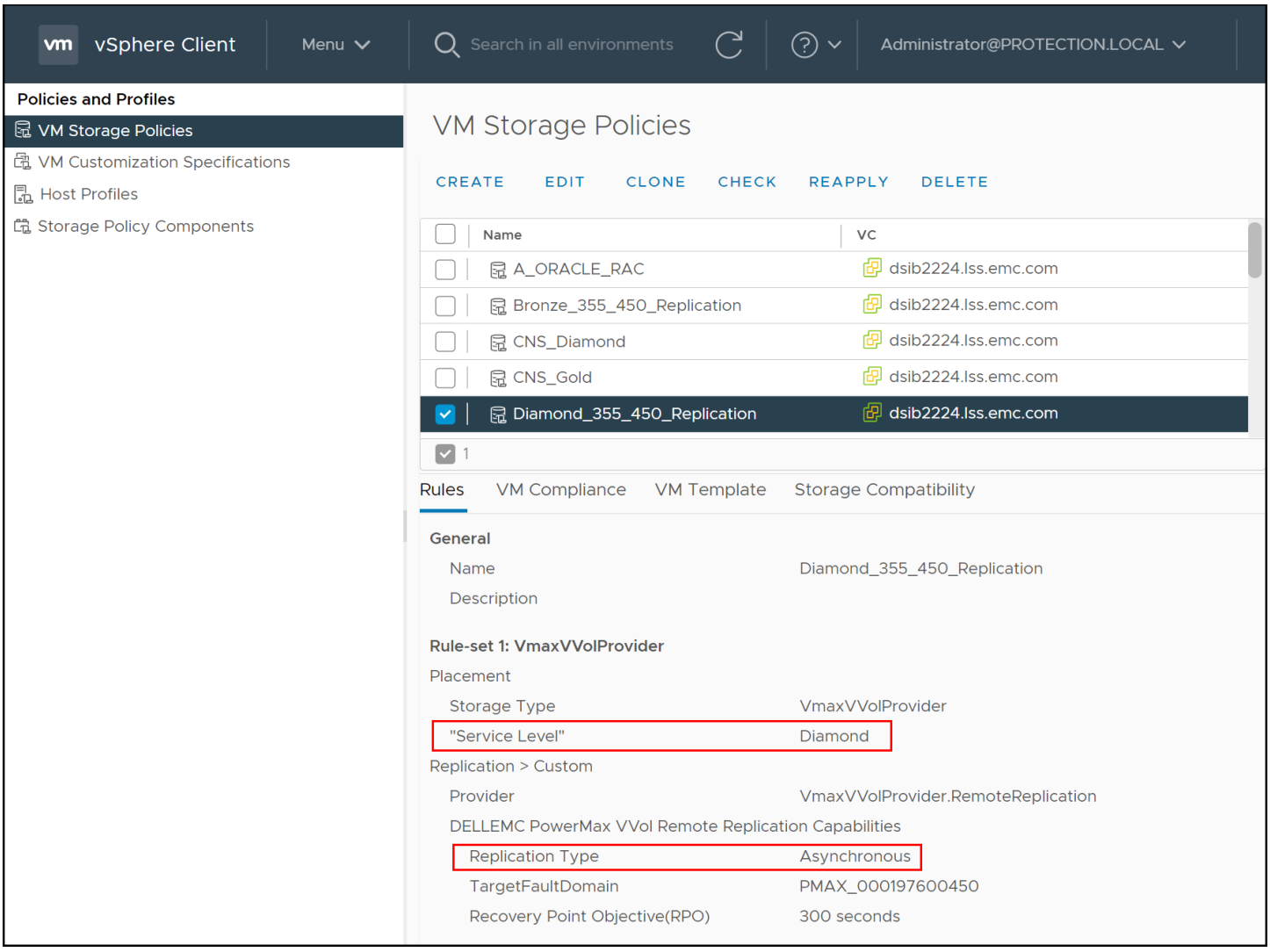

Figure 2. Storage policy for Oracle RAC nodes

Once assigned, VMware will prompt the user to select a Replication Group – in this example ORADB. This should be the same for every disk that is part of the VM. In Figure 3, the first node, dsib2019.lss.emc.com is displayed. Because all of the devices are in a consistent replication state, the VM is reported as compliant with the storage policy.

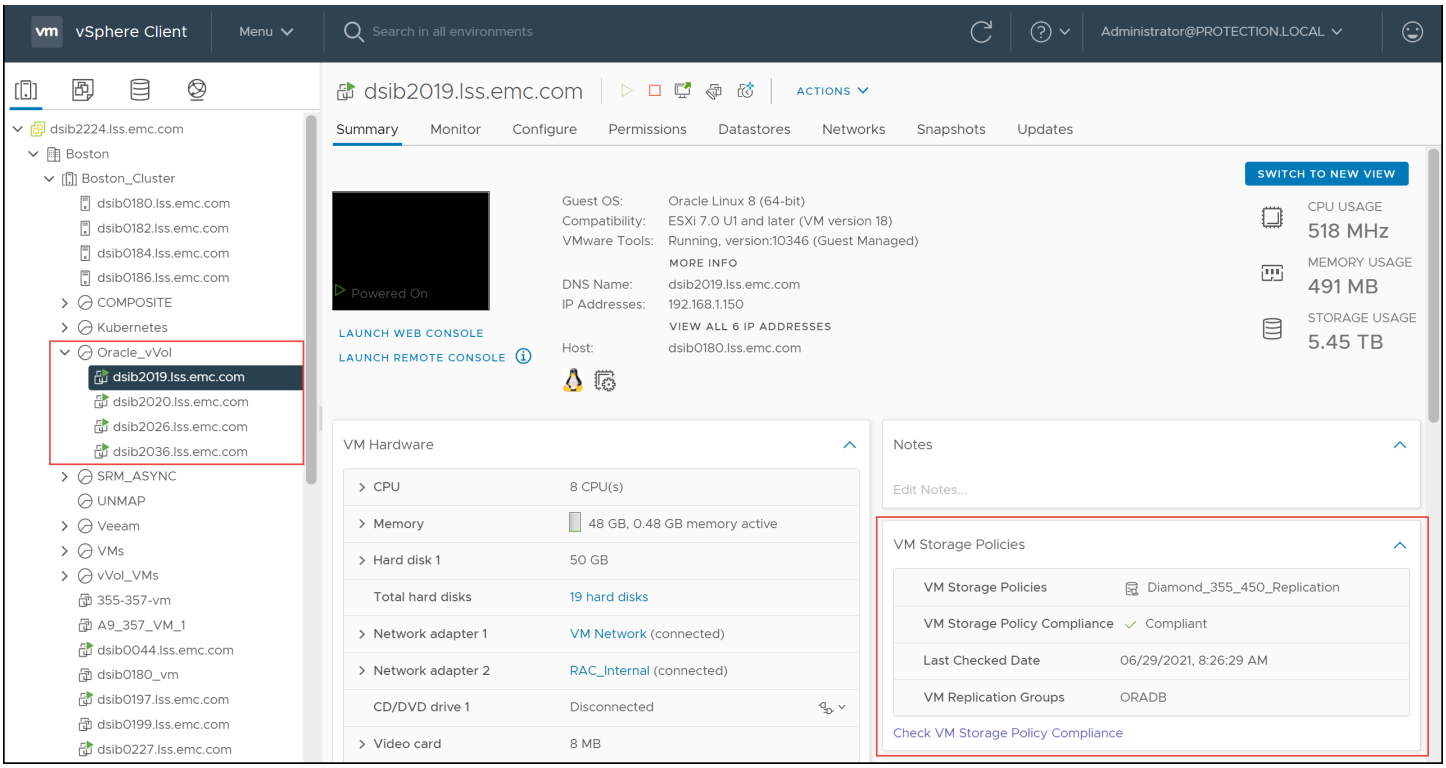

Figure 3. Oracle RAC nodes assigned the same VM Storage Policy

2.1.1 Configuration

The following steps provide the high-level process for creating the four-node Oracle RAC environment in this use case running on vVols. The procedure assumes that the basic setup of vVols with VASA 3 and replication is complete before undertaking. This includes, but is not limited to:

- Creation of storage containers on both arrays
- Setup of VASA Replication Groups between the storage containers
- Provisioning of protocol endpoints to ESXi hosts at the protection and recovery sites
- Registration of active and standby VASA Providers on each vCenter at protection and recovery sites respectively
- Creation of vVol datastores of sufficient size to support SRM testing. Dell EMC recommends that the storage resource on the recovery site should be at least double in size of the protection site to accommodate testfailover testing as the VASA Provider will create an equal number of target vVol devices as the production devices. Remember that a storage container is only logical space and does not consume storage from the SRP except when data is written. Therefore there is no penalty in oversizing it.
- Creation of storage policies
- Configuration of network for public and private interfaces

2.1.1.1 Installation overview

This installation overview is not meant to take the place of the Oracle Clusterware and Database installation documentation, rather it should provide overall guidance to the process. The list below follows the installation present in this whitepaper, and thus references objects specific to it, e.g., storage policy.

1. Configure an NTP client for use in the environment.
2. Create a single VM with Guest OS set to Oracle Linux 8 (64-bit) with two vmdks, one to hold the OS, and another to hold the Oracle database binaries. These disks are assigned the storage policy Diamond_355_450_Replication in this environment. When selecting the policy as shown in Figure 4, VMware will list the compatible datastores. After selecting the proper vVol datastore which supports replication, VMware will require the user to set a Replication Group, in this example ORADB.

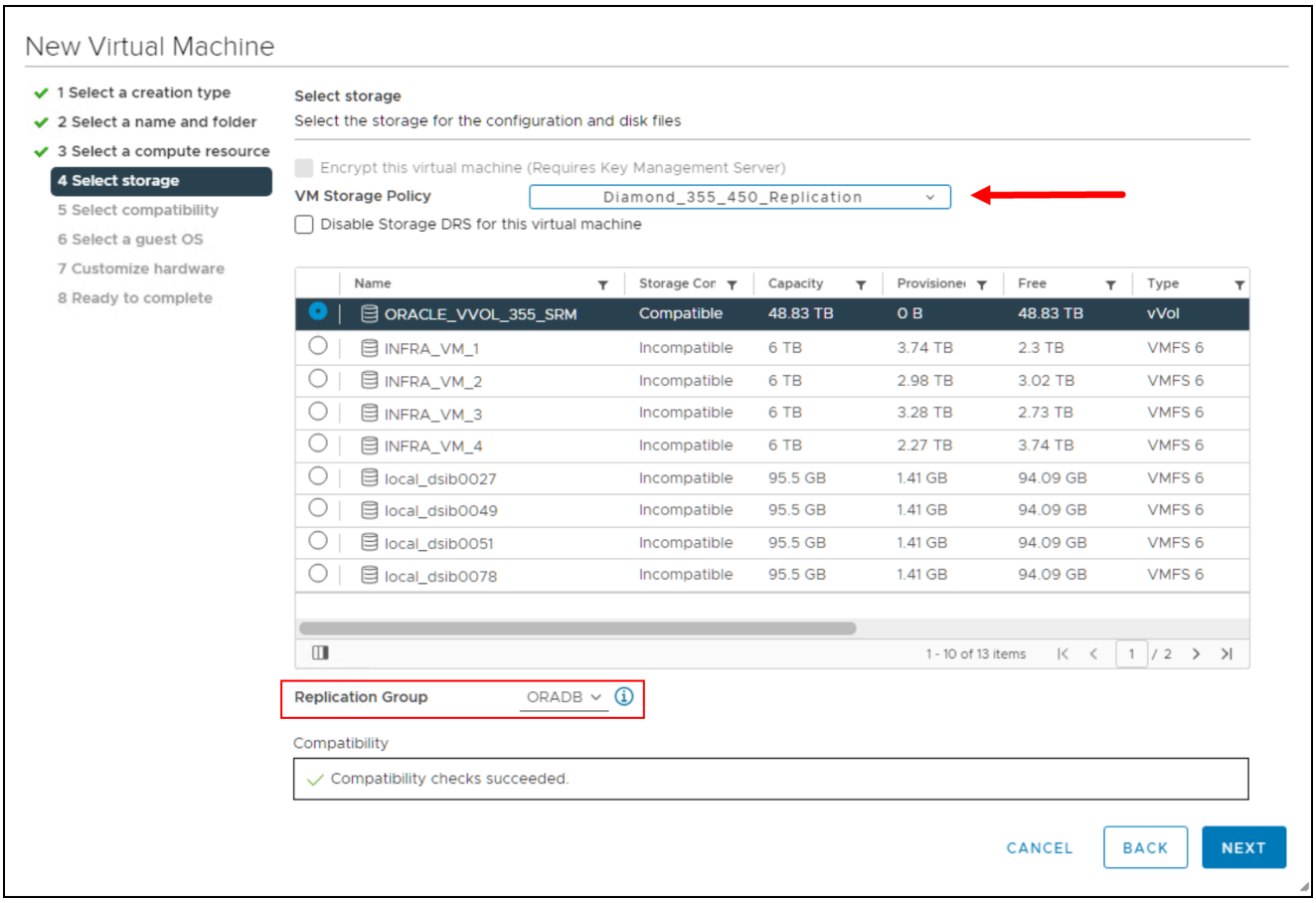

Figure 4. Selecting storage policy, compatible datastore and replication group

The VASA Replication Group can be seen in Unisphere for PowerMax in Figure 5.

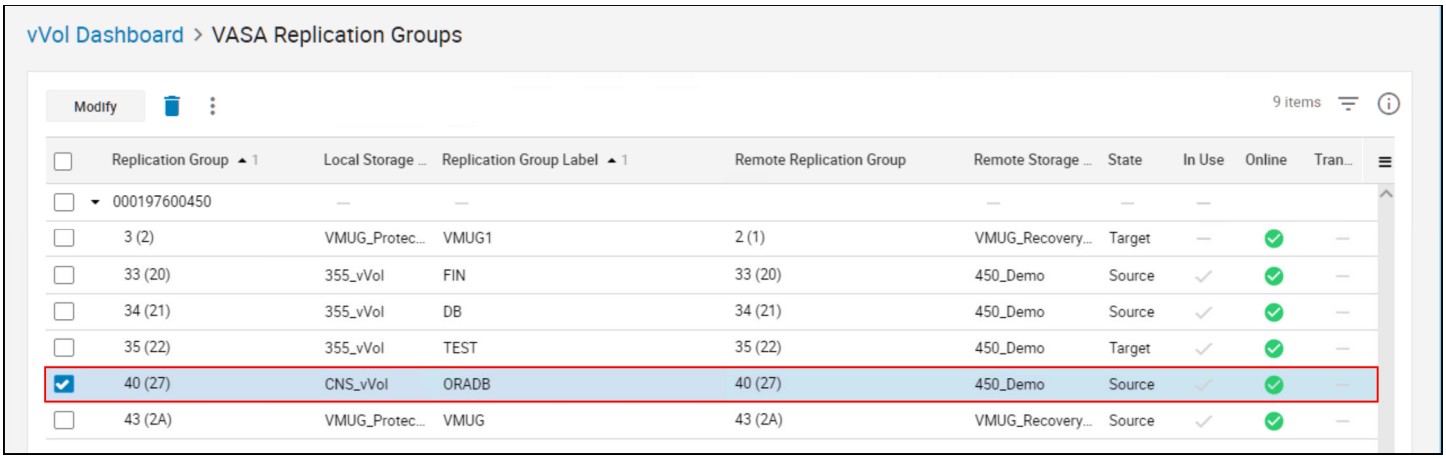

Figure 5. ORADB VASA Replication Group

3. Add a second NIC to the VM which will be used for the private network. Set both NICs to VMXNET3. These NICs are on VMware networks configured on their own physical adapter to separate traffic.
4. Install Oracle Enterprise Linux. Once complete, this VM can be used as a template for the other nodes as was done in this use case.
5. Configure NTP on the node.
6. Create vmdks for the Oracle database and CRS/voting disks being sure to assign the correct storage policy and replication group. For this environment a total of seventeen disks were created in addition to the two previously created, with three ASM disk groups: DATA, REDO and OCR. Normally all vmdks for RAC (and shared disks) need to be disk type eagerzeroedthick (EZT); however, vVols are an exception as all vVols are thin, yet they act as thick and reserve space in the datastore like EZT. The disks should be assigned to a Paravirtual SCSI controller. Be sure the controller is set to the default of “None” for SCSI Bus Sharing so that vMotion will work.
7. Use VMware KB article 1034165 to set the multi-writer flag on the vmdks, allowing more than one node to access the disk.
8. Use `fdisk` to configure a single partition on all the disks. It is important to align the partition on the disk since neither Oracle nor VMware will do this automatically. VMware ensures the VMFS is aligned but not the file systems on vmdks.
9. Change the ownership of the disks/partitions so that the ASM assistant will recognize them as available for use.
10. Install and configure Oracle Grid Infrastructure 19c. There are many prerequisites that must be completed before installation. Oracle uses the Cluster Verification Tool (CVU) to ensure the nodes are ready for installation. The Clusterware install will setup and use a single ASM disk group during installation. Set this disk group to “Normal” redundancy which will ensure multiple voting disks as Oracle automatically determines the number of files based on redundancy. This ASM disk group (e.g., OCR) need not be large as the Clusterware files are small.
11. Create the remaining ASM disk groups using the ASM GUI or ASM CLI, being sure to set redundancy to “External.”
12. Install Oracle database binaries on all nodes at once through the installer. Note that it is possible to add nodes after the initial RAC node is configured rather than doing them all at once.
13. Use DBCA to create the new database and all the RAC instances.

The RAC nodes are now ready for configuration with VMware Site Recovery Manager.

3 VMware SRM setup

VMware vCenter Site Recovery Manager (SRM) leverages storage array-based replication such as Dell EMC Symmetrix Remote Data Facility (SRDF) to protect virtual machines in VMware vSphere environments. The interaction between VMware SRM and storage array replication is governed by a well-defined set of specifications. These VMware-defined specifications are implemented by the storage array vendor as either a lightweight application referred to as the storage replication adapter (SRA) for VMFS/NFS/RDMs, or within the VASA 3.0 Provider for virtual volumes.

Dell EMC embedded VASA 3.0 Provider (EVASA) enables VMware SRM to interact with a PowerMax environment running vVols. It allows VMware SRM to automate storage-based disaster restart operations on PowerMax arrays in an SRDF configuration. Unlike the SRDF SRA, there is no configuration required outside of SRM for vVols.

The installation and general configuration of SRM will not be covered in this section. If instructions are needed, please follow the VMware documentation for installing SRM 8.4 with vSphere 7.0 as there is nothing about its configuration that requires changes for vVols. No additional software must be installed (e.g. an SRA) to use SRM with vVols, save for registering the VASA 3.0 Provider which is necessary to run vVols even without replication. Where necessary, however, specific SRM vVol information pertinent to the environment will be included below.

3.1 SRM restrictions

VMware and Dell EMC have a number of restrictions when using SRM with vVols which are important to understand prior to the configuration with Oracle RAC.

3.1.1 VMware

Some of the most important restrictions from VMware are the following. Note these apply only to SRM.

- Site Recovery Manager does not support protection of vVol virtual machines that have non-replicated virtual disks.
- Site Recovery Manager does not support placing replication groups from different fault domain pairs in the same vVol protection groups. For example, if Array A has VMs replicating to Array B in replication group 1, and other VMs replicating to Array C in replication group 2, there cannot be a single protection group with both group 1 and group 2. Each group would have to be in a different protection group, though they could be in the same recovery plan.
- Site Recovery Manager does not support the protection of virtual machines with different vVols-based disks, replicated by different storage policies or different vVol replication groups.
- vVols does not support the recovery of template virtual machines.

Be sure to review the VMware SRM documentation for more details around other limitations, particularly maximums for object like protection groups and VMs.

3.1.2 Dell EMC

The following limits apply to Dell EMC's implementation of SRM with vVols on the PowerMax:

- 250 VMs supported with SRM
  - 2000 vVols – average of 8 vVols per VM; however, the 2000 vVols may allocated as required across VMs. In this use case, for example, each RAC node has 19 devices; however as only two

are unique to each VM, the total devices for all four VMs is only 33 (one config vVol, one swap vVol, and two local vVols for each VM and seventeen shared vVols across all VMs).

- 25 VASA Replication Groups

3.2 SRM configuration for Oracle RAC

The following sections cover the salient points of the SRM configuration for the Oracle RAC environment running on vVols. As much of the setup of SRM, including mappings between the sites, is universal to all environments, it is for the most part not included.

3.2.1 Storage policy mappings

For this environment, a single storage policy mapping was configured, though the policy itself is named differently on each site as shown in Figure 6. On the protection site, the policy is named **Diamond_355_450_Replication** and the recovery site, **Diamond_450_355_Replication**.

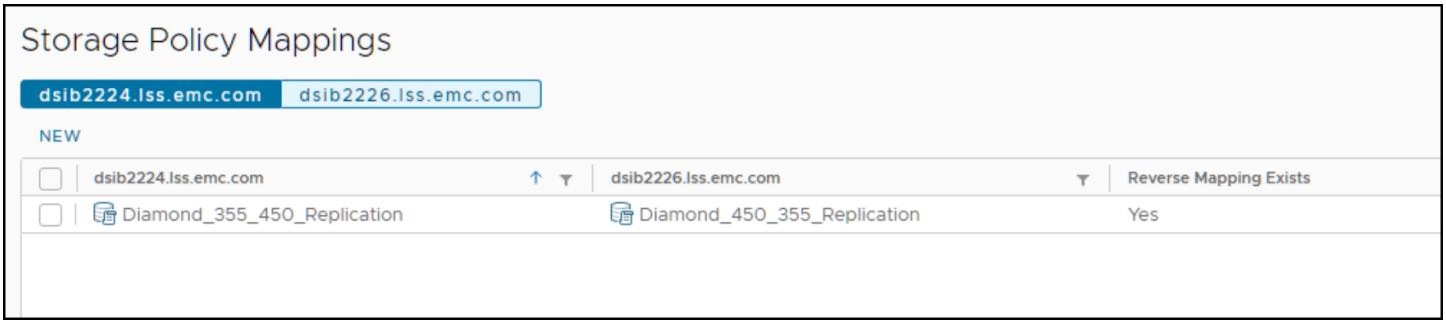

Figure 6. Storage policy mapping SRM

3.2.2 Create a Protection Group and Recovery Plan

The following is a walk-through of setting up the protection group and recovery plan for this Oracle RAC implementation. For more detail, please see the VMware documentation.

Within SRM, in step 1, select **NEW** under the Protection Groups screen to start the dialog as shown in Figure 7.

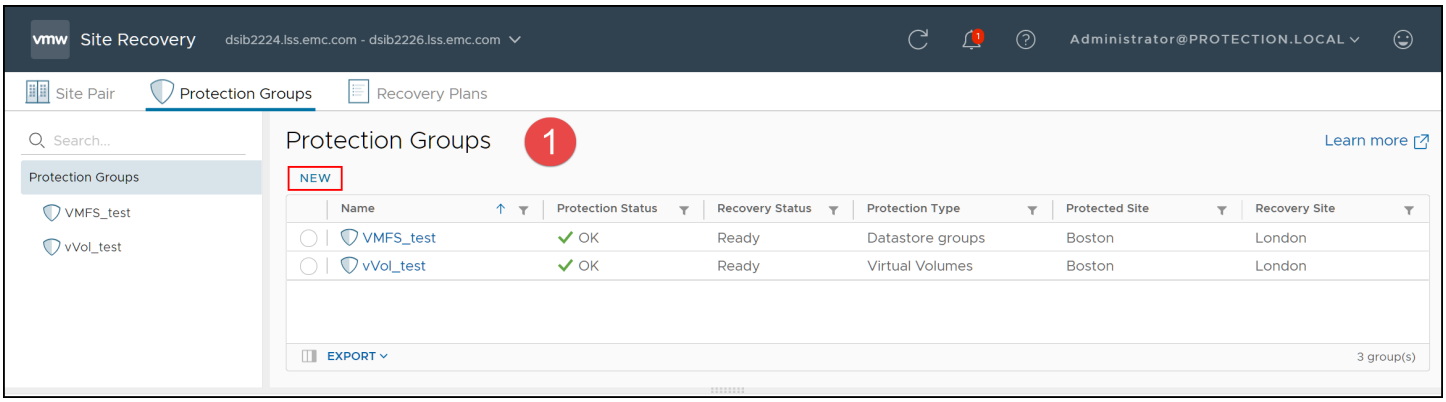

Figure 7. Create SRM Protection Group: Step 1

In step 2 shown in Figure 8, name the group, add a description if desired, and select **NEXT.** Be sure the Direction is correct.

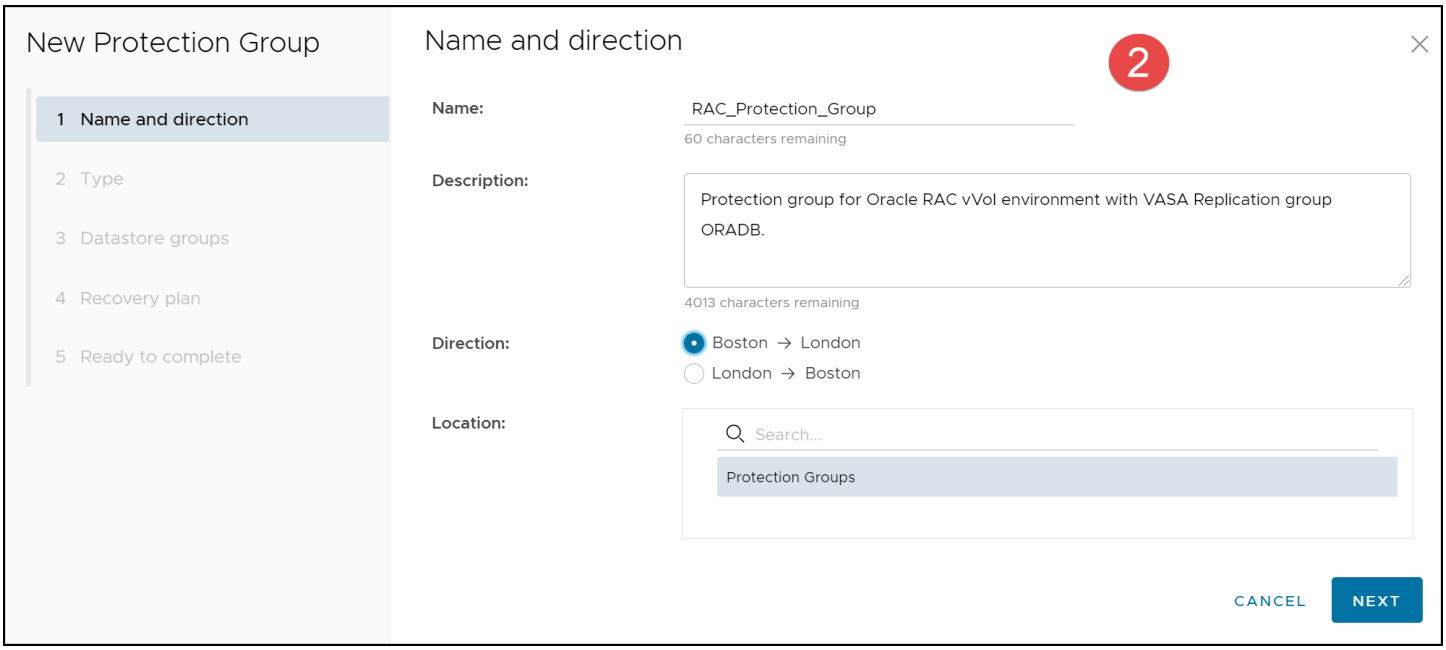

Figure 8. Create SRM Protection Group: Step 2

Next, in step 3 shown in Figure 9, select the **Virtual Volumes** radio button. Select the desired **Fault Domain** (remote array) and then **NEXT**.

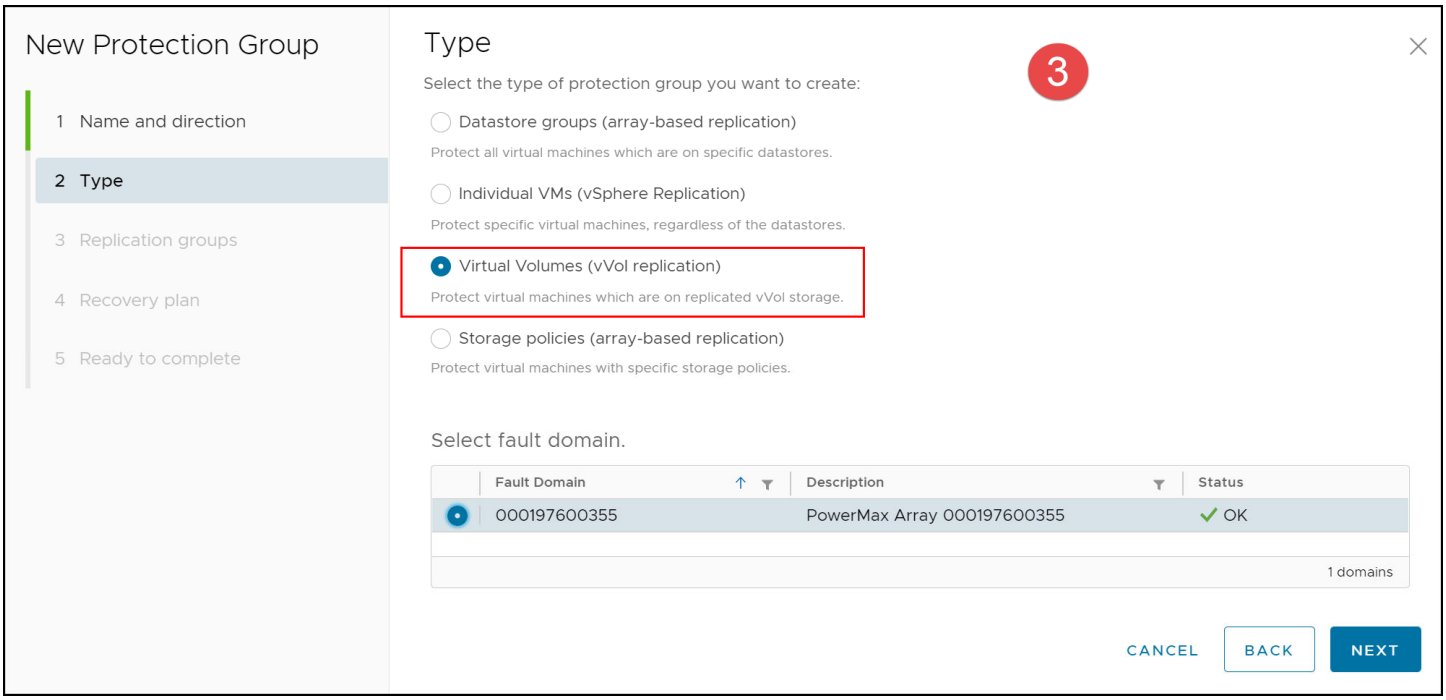

Figure 9. Create SRM Protection Group: Step 3

In step 4 shown in Figure 10, locate the VASA Replication Group being used with the Oracle RAC VMs and check the box next to it. Be sure all RAC nodes are listed and then select **NEXT**.

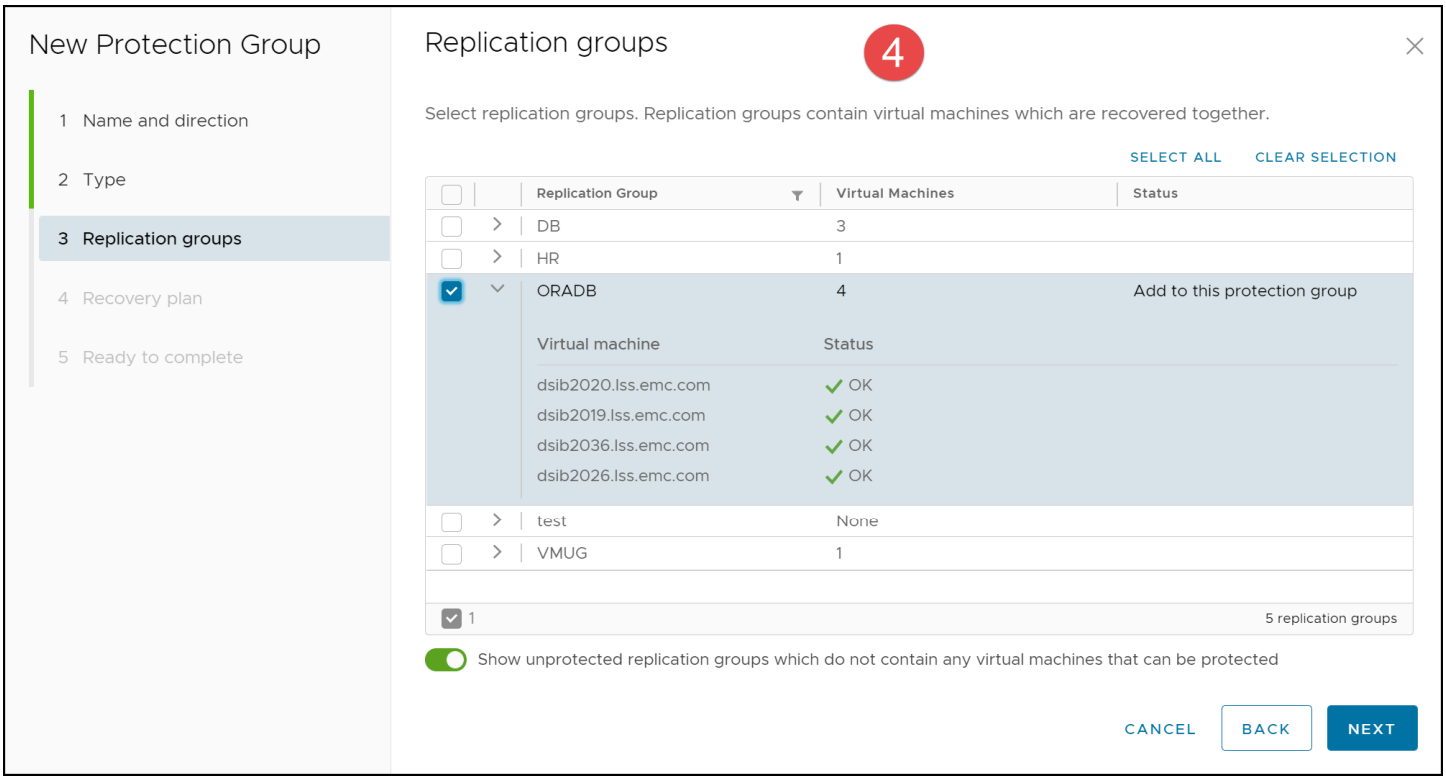

Figure 10. Create SRM Protection Group: Step 4

Finally, in step 5 shown in Figure 11, choose whether to create a new recovery plan or skip it for now and select **NEXT**. The option to skip is selected here so the entire Recovery plan wizard can be shown.

Figure 11. Create SRM Protection Group: Step 5

Review the options shown in Figure 12 and then execute **FINISH**.

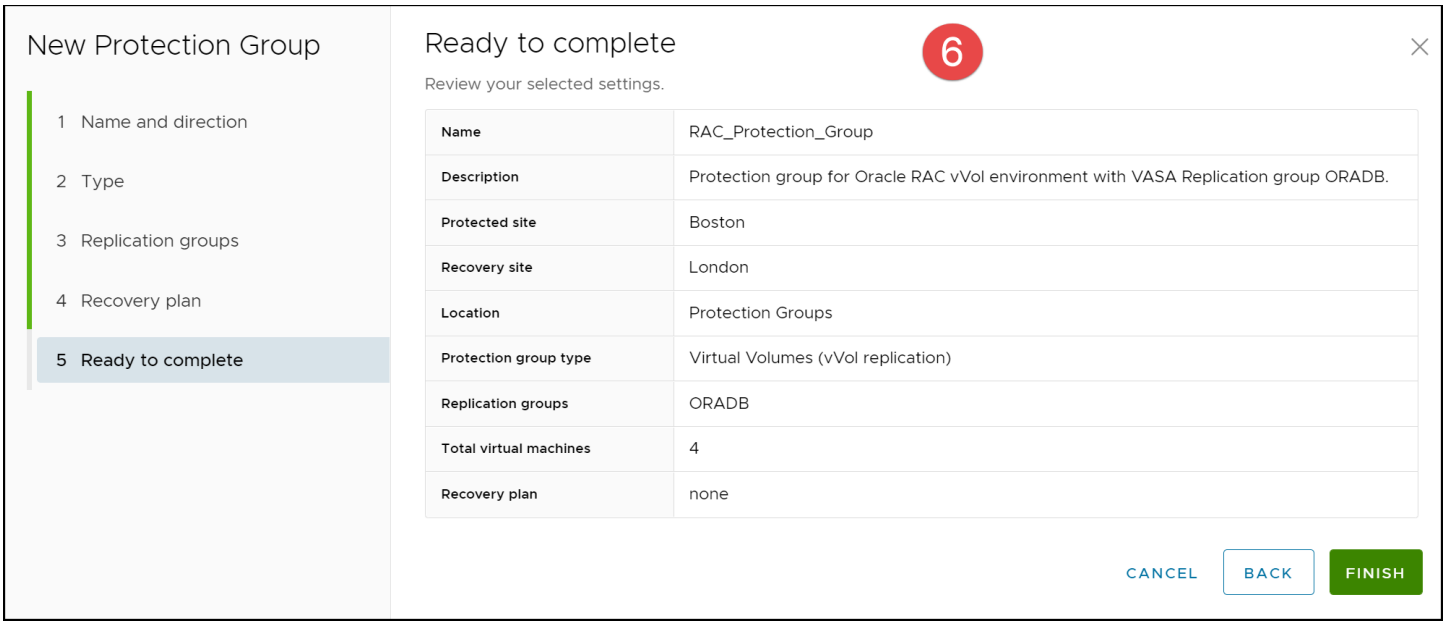

Figure 12. Create SRM Protection Group: Step 6

If the recovery plan was not created as part of the protection group wizard, do so now. From the Recovery Plans screen, select **NEW** in step 1 shown in Figure 13.

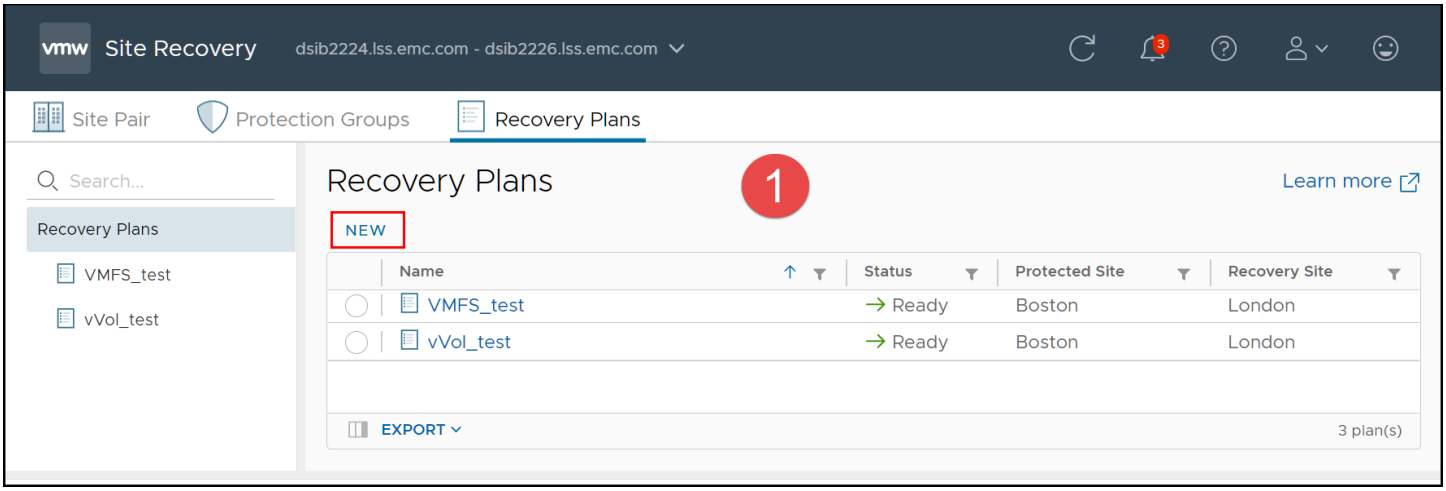

Figure 13. Create SRM Recovery Plan: Step 1

In step 2 shown in Figure 14, type in a Name and Description if desired, and then select **NEXT**. Be sure the Direction is correct.

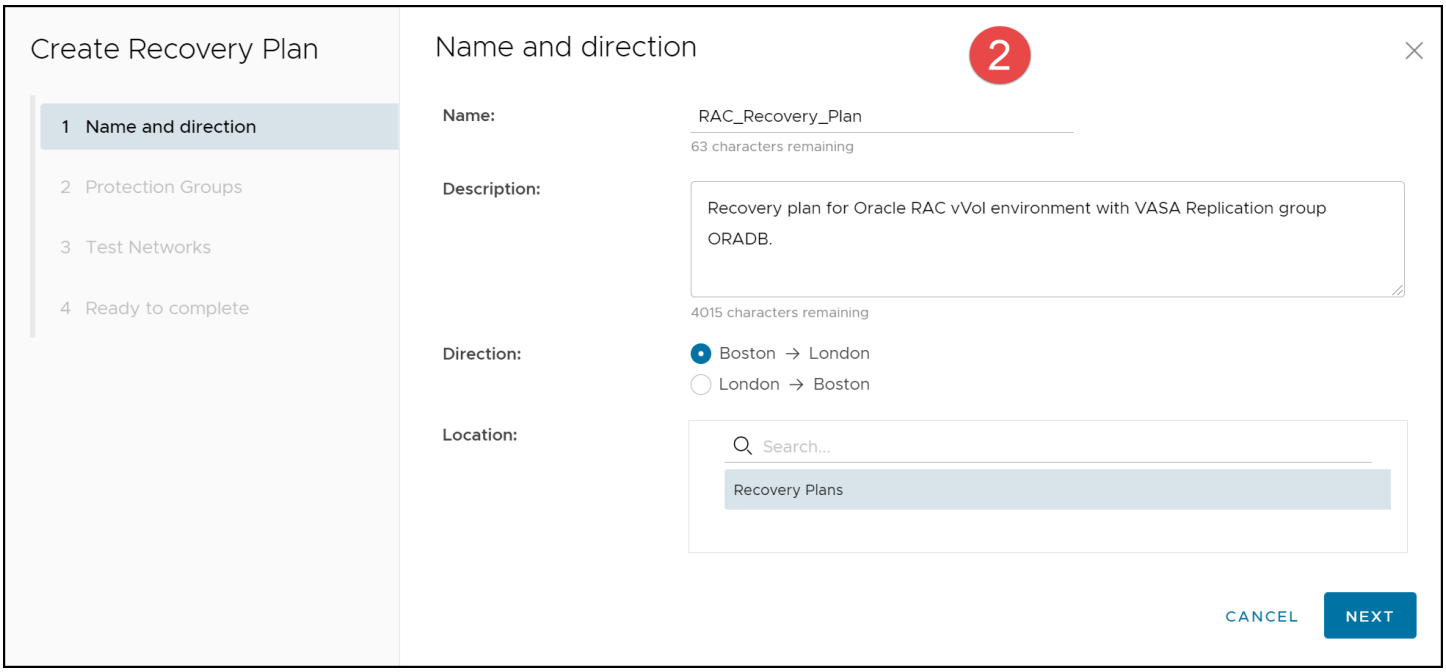

Figure 14. Create SRM Recovery Plan: Step 2

In step 3 shown in Figure 15, select the protection group previously created by checking the box. Then select **NEXT**.

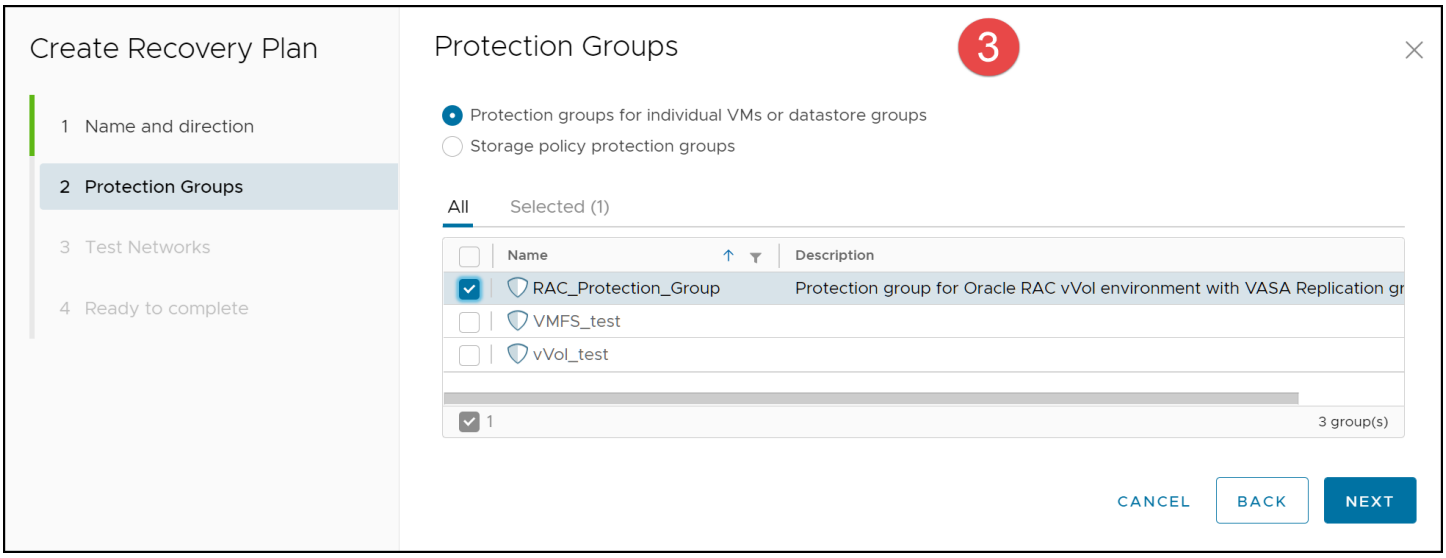

Figure 15. Create SRM Recovery Plan: Step 3

For the test networks shown in step 4 in Figure 16, keep the default site-level mapping and select **NEXT**. Network re-mapping will be addressed as part of the modifications in the section SRM testfailover networking.

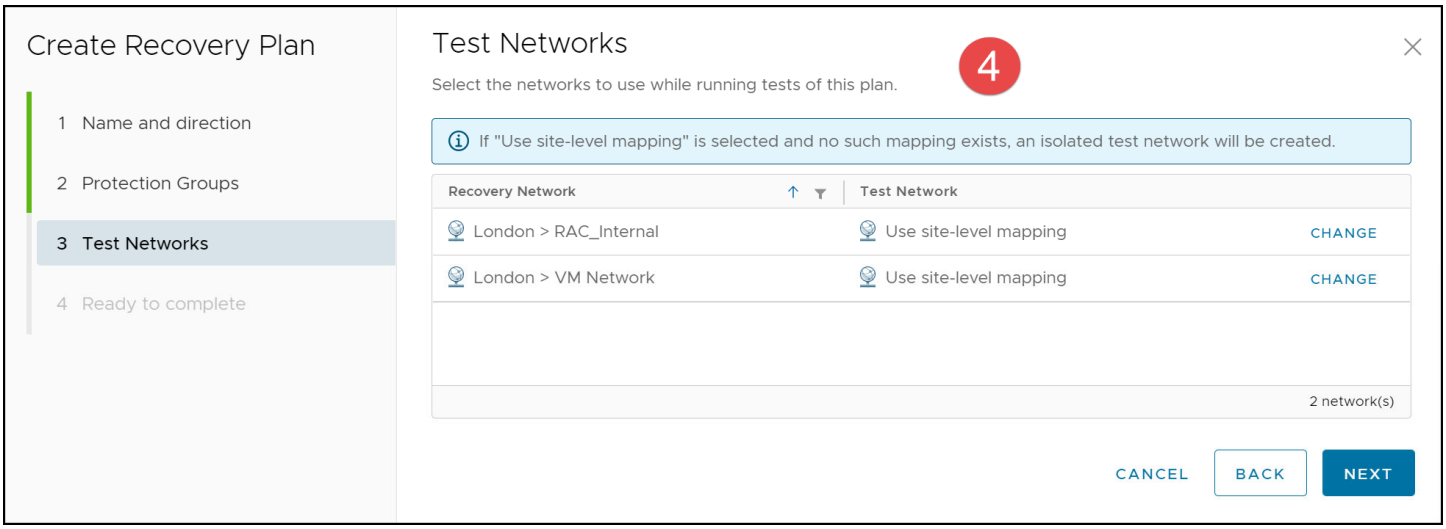

Figure 16. Create SRM Recovery Plan: Step 4

In step 5 shown in Figure 17, select **FINISH**.

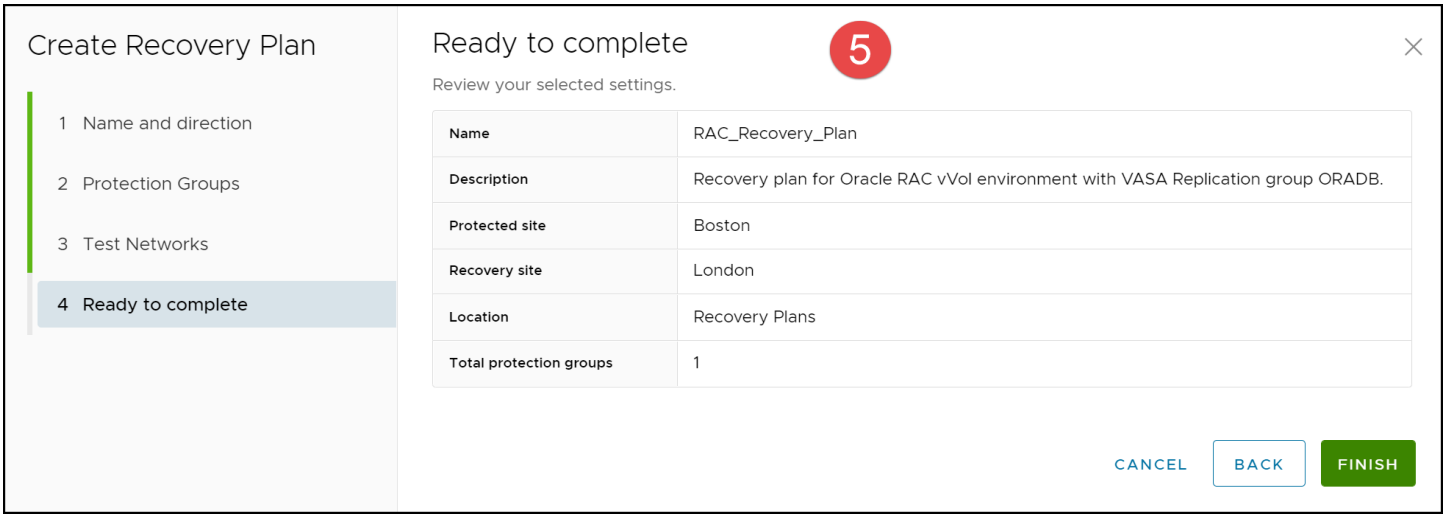

Figure 17. Create SRM Recovery Plan: Step 5

An architectural view of the completed SRM environment is shown in Figure 18.

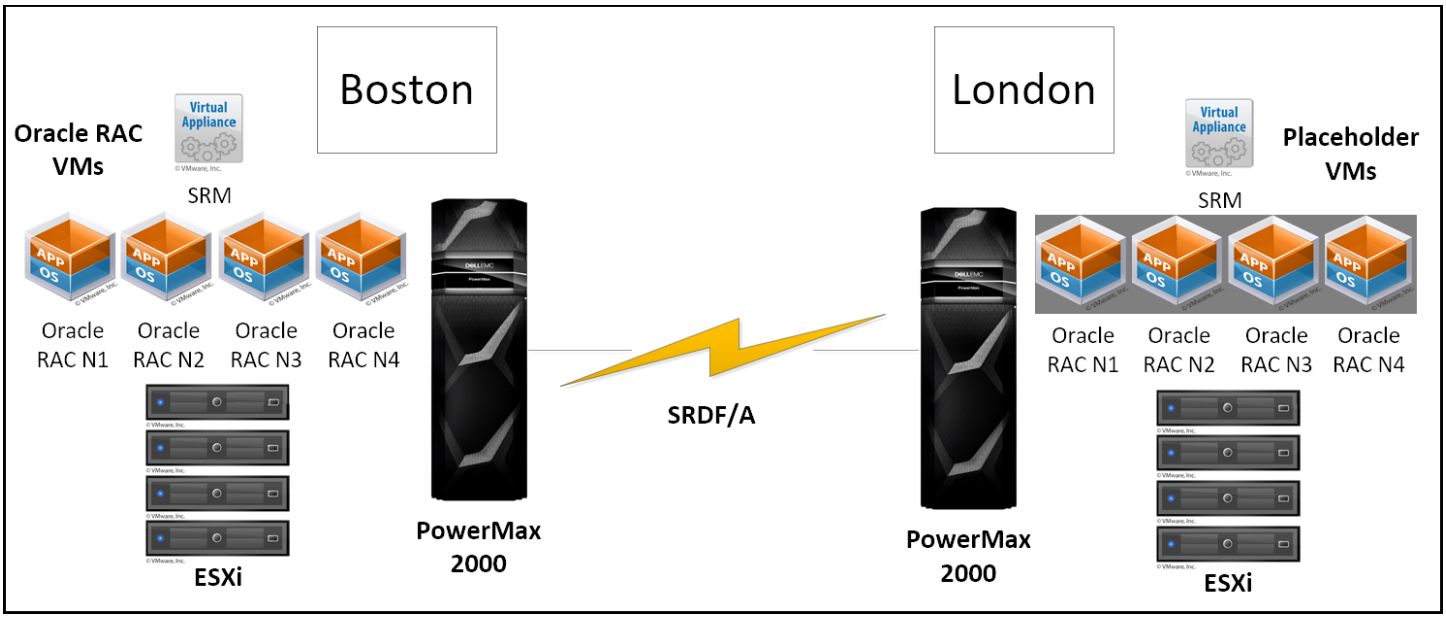

Figure 18. Oracle RAC environment showing the protection and recovery site

In the next chapter, DRS Affinity Rules are covered. These will dictate how RAC nodes are assigned to hosts at the recovery site during testfailover or failover.

4 DRS Affinity Rule

DRS Affinity Rules dictate on which ESXi hosts VMs must or should run. When using DRS Affinity Rules on the protection site in an SRM configuration, however, they do not carry over to the recovery site because there is no mapping capability between them as there is for objects like storage policies. Therefore, it is necessary to pre-configure rules on the recovery site so that when a testfailover or failover is run, the VMs will be placed on the desired hosts.

The following sections detail how to configure the DRS Affinity Rules for the recovery site described in this paper.

4.1 VM/Host Groups

First, create a host rule for each ESXi host at the recovery site. In step 1 shown in Figure 19, highlight the recovery site cluster in the left-hand panel, select the **Configure** tab on the right, then highlight the **VM/Host Groups** option under the **Configuration** section. Select **ADD…**.

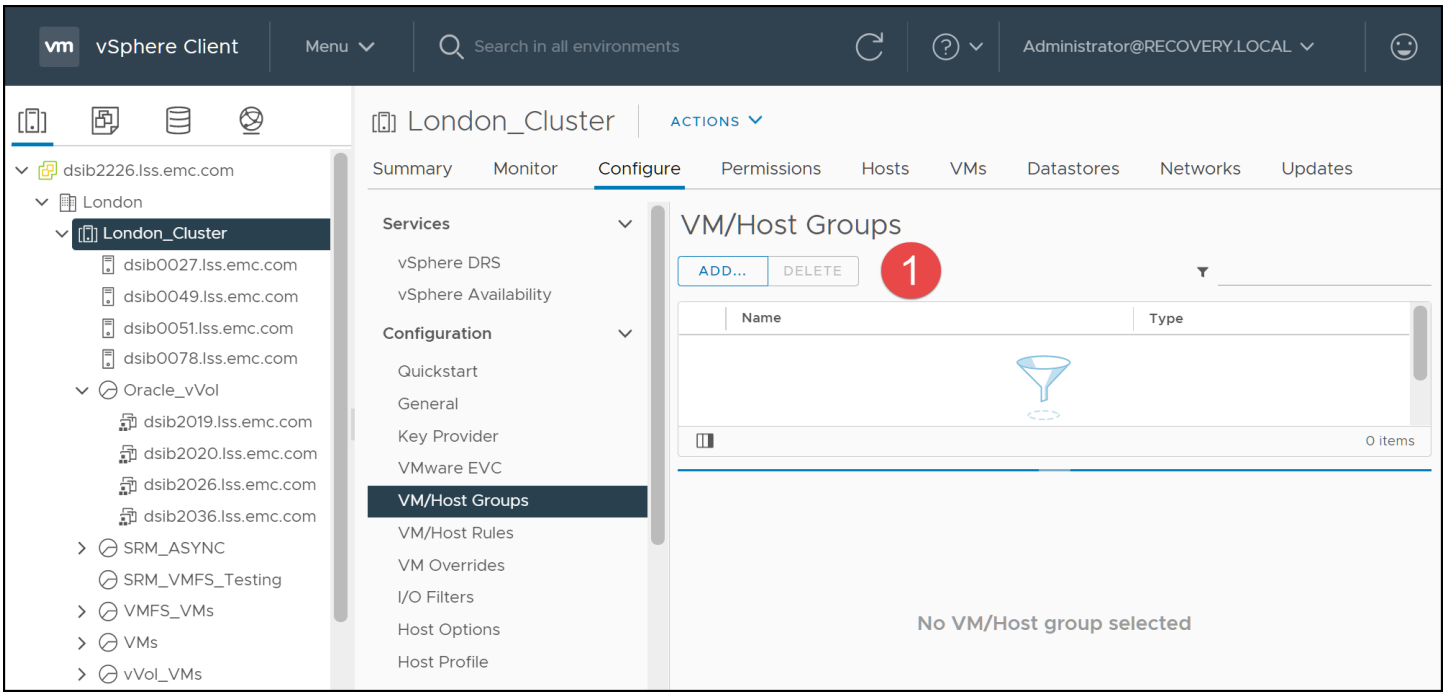

Figure 19. Host Group creation: Step 1

Next, in step 2 shown in Figure 20, type in a name for the host group, select **Host Group** from the drop-down menu next to **Type**, and select **ADD…**.

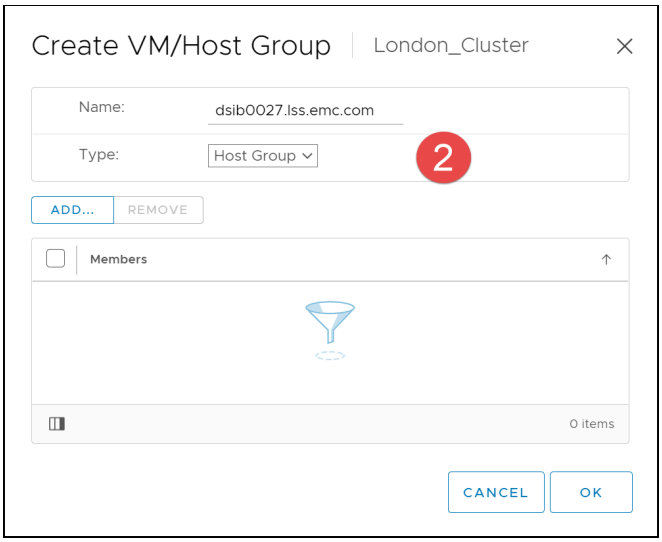

Figure 20. Host Group creation: Step 2

Using the checkbox, select the appropriate host in step 3 shown in Figure 21. Select **OK** to save.

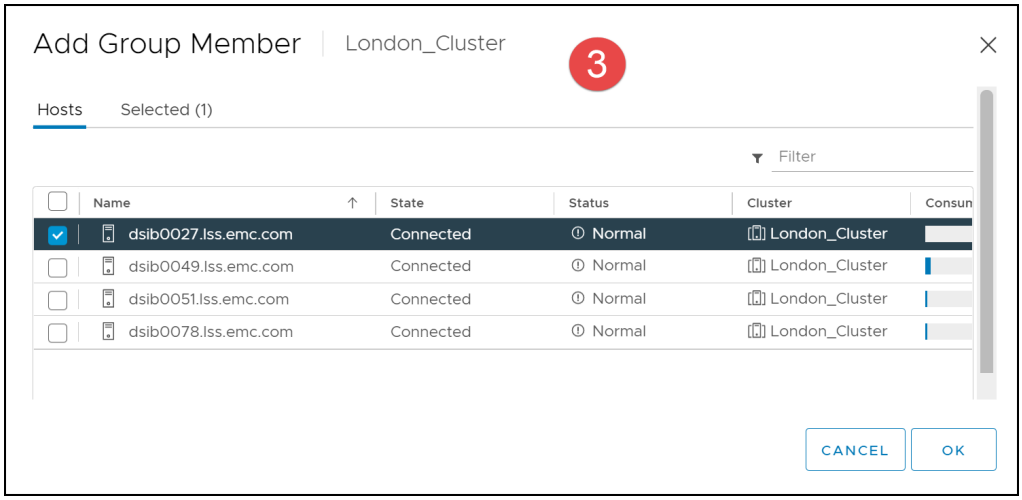

Figure 21. Host Group creation: Step 3

In the final screen in step 4 shown in Figure 22, review and then select **OK**.

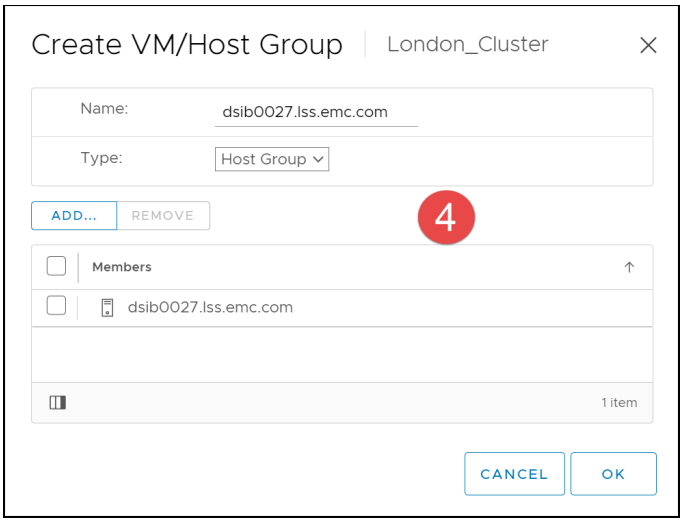

Figure 22. Host Group creation: Step 4

Repeat these four steps for all four hosts in the recovery site cluster. When complete, the final result will appear similar to what is shown in Figure 23.

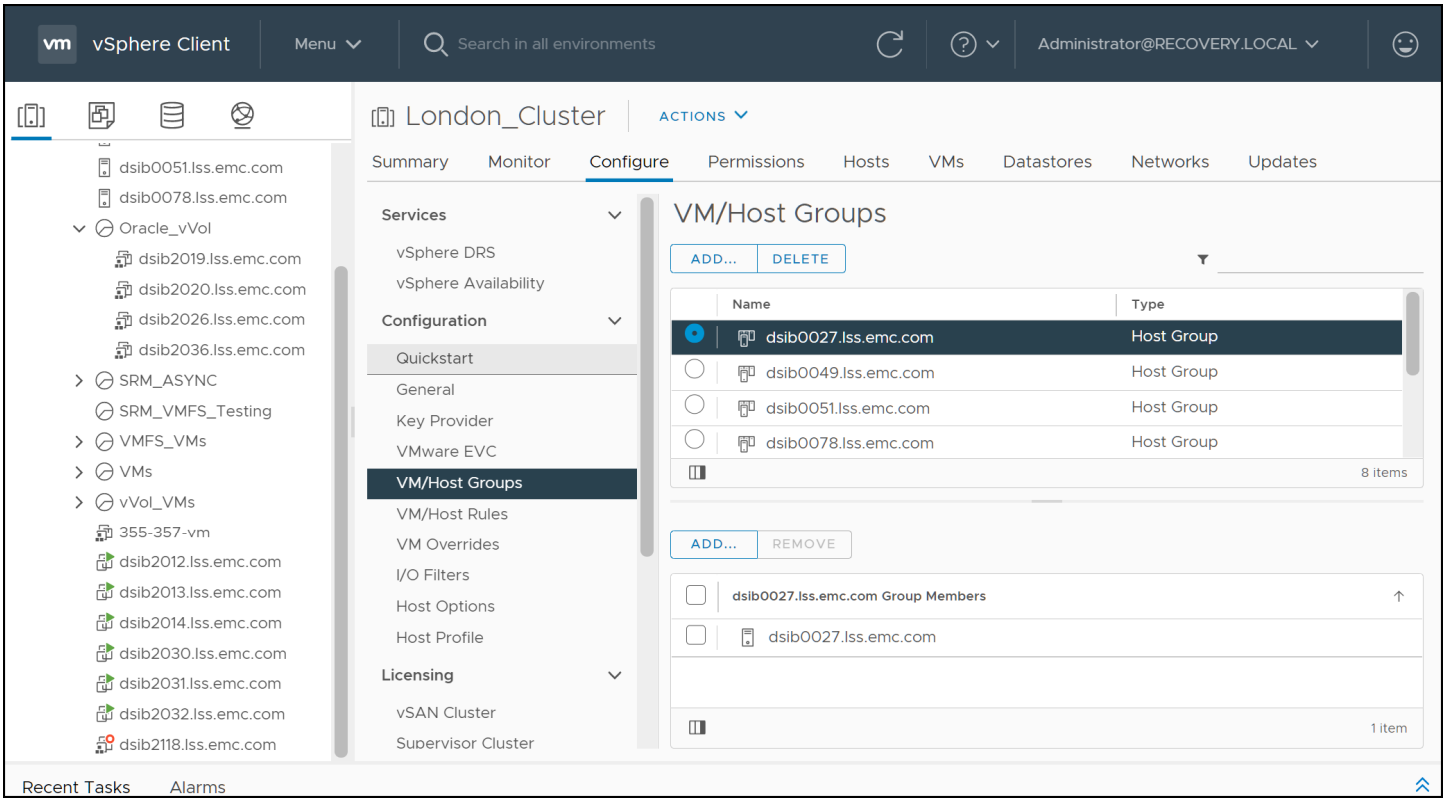

Figure 23. Host Group creation summary

Using the same wizard shown in Figure 19 for step 1, now create the VM Groups. In step 2 shown in Figure 24, provide a name for the VM Group and use the drop-down box next to **Type** to select **VM Group**. Select **ADD…** to choose the appropriate VM.

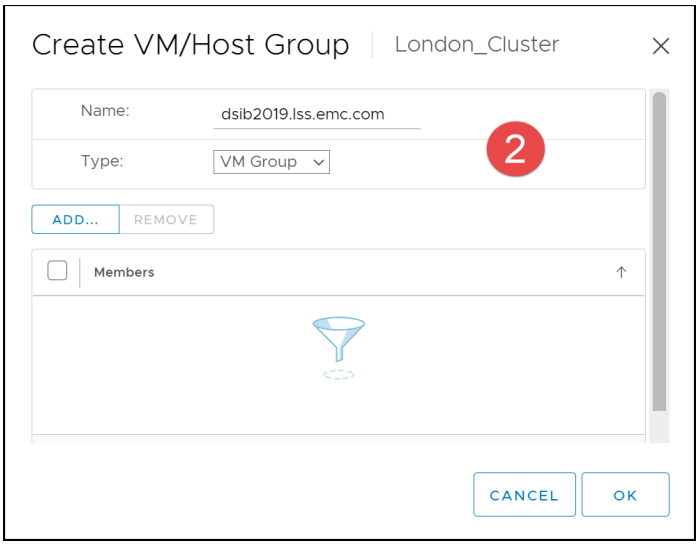

Figure 24. VM Group creation: Step 2

In step 3 shown in Figure 25, use the check box to select one of the RAC nodes and hit **OK**.

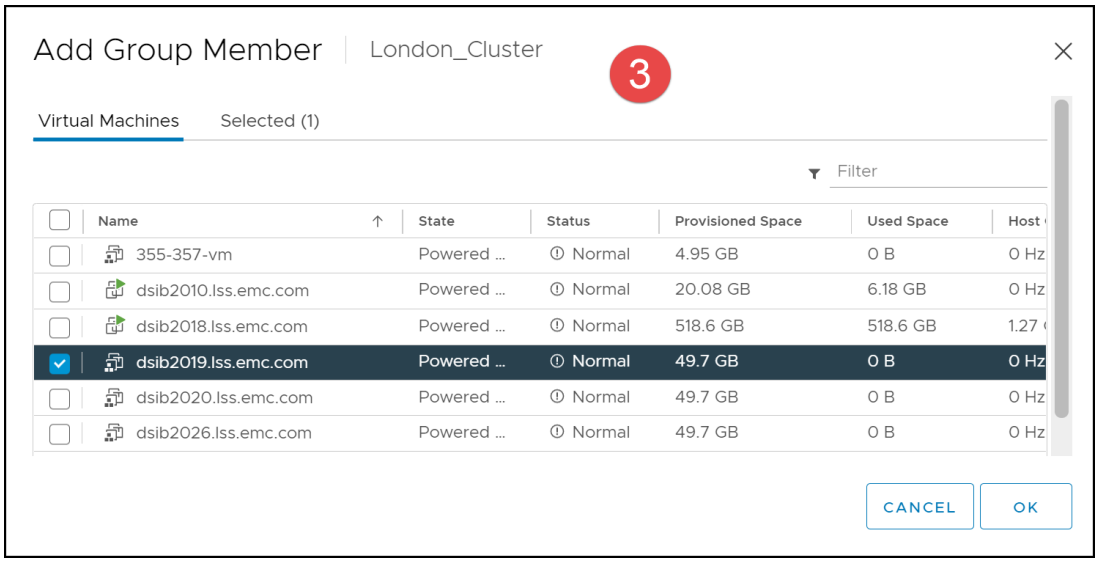

Figure 25. VM Group creation: Step 3

In step 4 shown in Figure 26, review the screen and select **OK**.

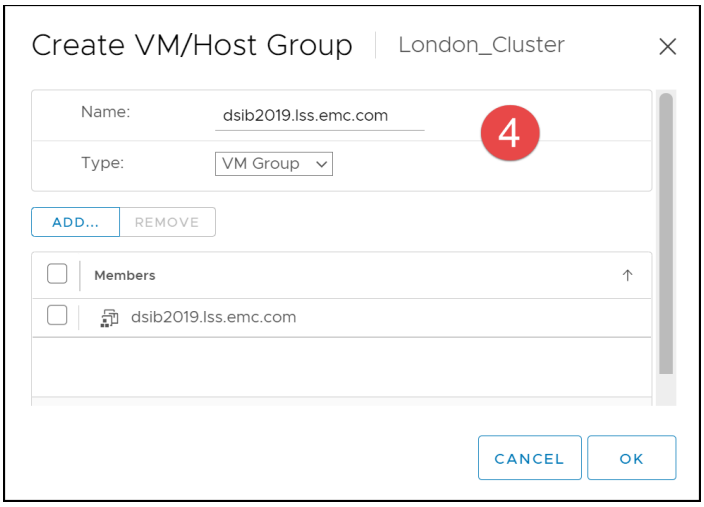

Figure 26. VM Group creation: Step 4

Repeat the VM Group creation for the other three RAC nodes. Next, setup the rules which govern how the startup of the VMs behave.

4.2 VM/Host Rules

After the VM/Host Groups are available, rules need to be configured. These will be similar to those created on the protection site. Each VM will be assigned to one of the ESXi hosts. The exact mapping of VM to host is not important so long as there is a one to one assignment as there are four VMs and four hosts. In step 1 shown in Figure 27, again navigate to the cluster on the left-hand panel, then the **Configure** tab on the right, and finally **VM/Host Rules** under **Configuration**. Select the **+ Add…** button

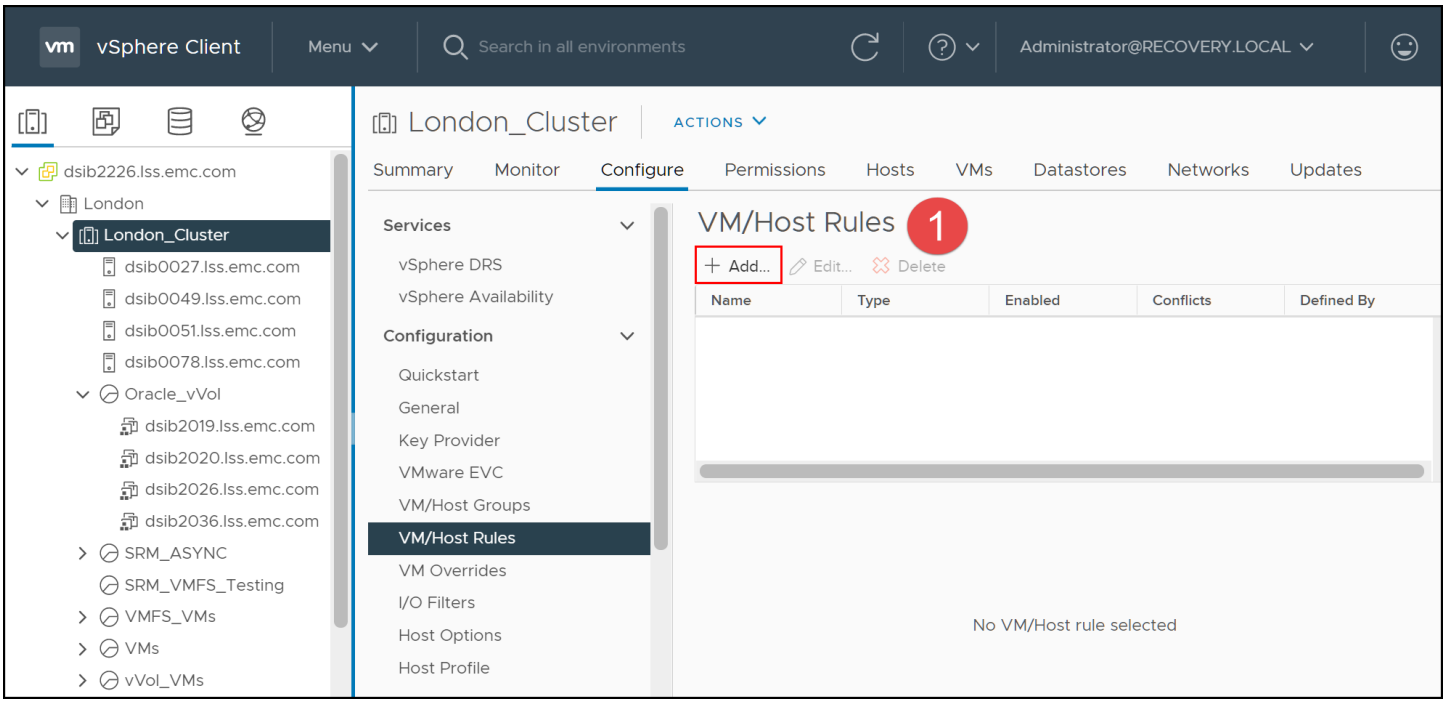

Figure 27. Host Rules creation: Step 1

Steps 2 through 6 are shown in Figure 28. In step 2 add a name. Note the **Enable rule** box is automatically checked. When creating the rule, in step 3 select the **Type Virtual Machines to Hosts** using the drop-down box which then populates the three options in the bottom panel. Selecting this type tells VMware to start a selected VM Group on a particular Host Group. Using the drop-down under **VM Group** in step 4, select the appropriate one previously created. Then, in step 6, select one of the host groups under **Host Group**. The most important option, however, is for dictating the condition under which the rule is enacted which is shown in step 5. The option, **Should run on hosts in group,** must be selected, and not **Must run on hosts in a group**. The use of *should* provides VMware the flexibility to start the VM on a different host if there is an issue with the configured one. Setting the value to *must* would prevent the VM from starting at all as it would be bound only to the chosen host. So, as shown in Figure 28, the VM dsib2019.lss.emc.com, whether in testfailover or failover, will attempt to start on host dsib0027.lss.emc.com. If the host is unavailable, VMware will select one of the other three hosts.

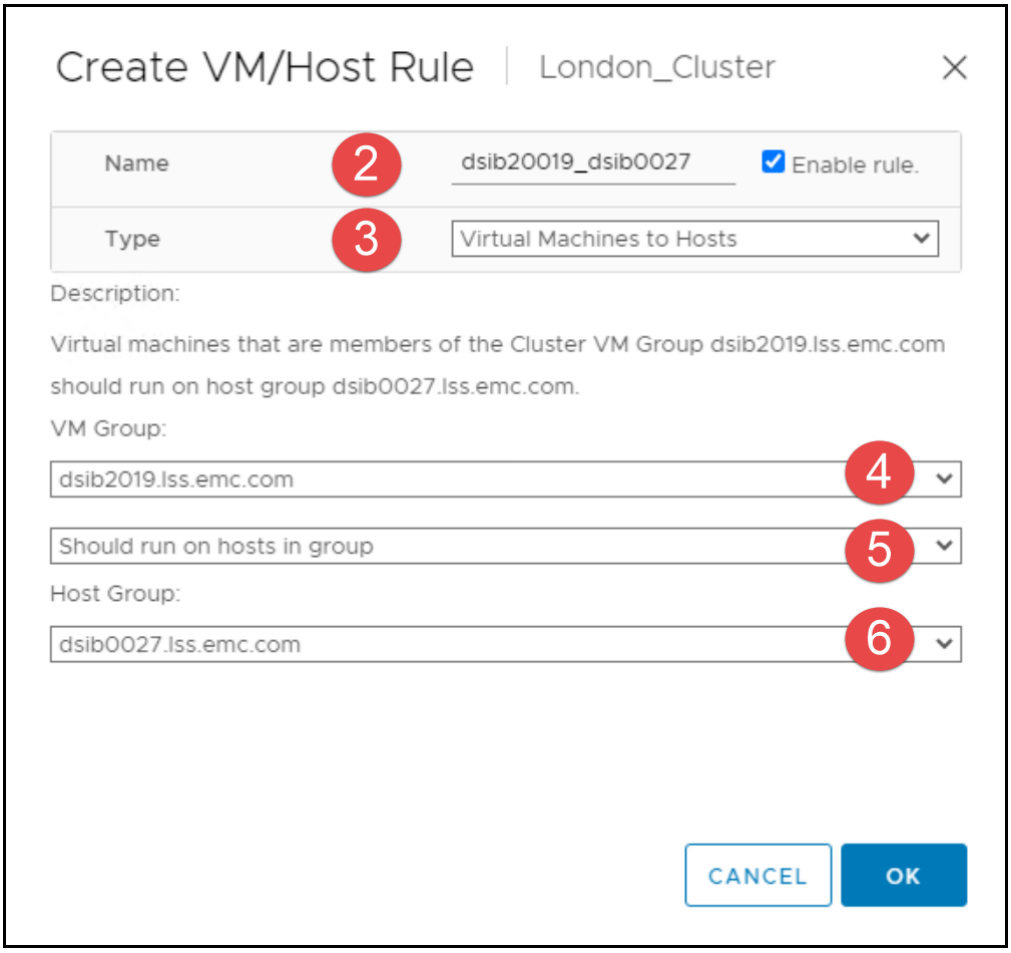

Figure 28. Host Rules creation: Steps 2 - 6

Create three additional rules for the other VMs. The new rules for each VM are shown in Figure 29.

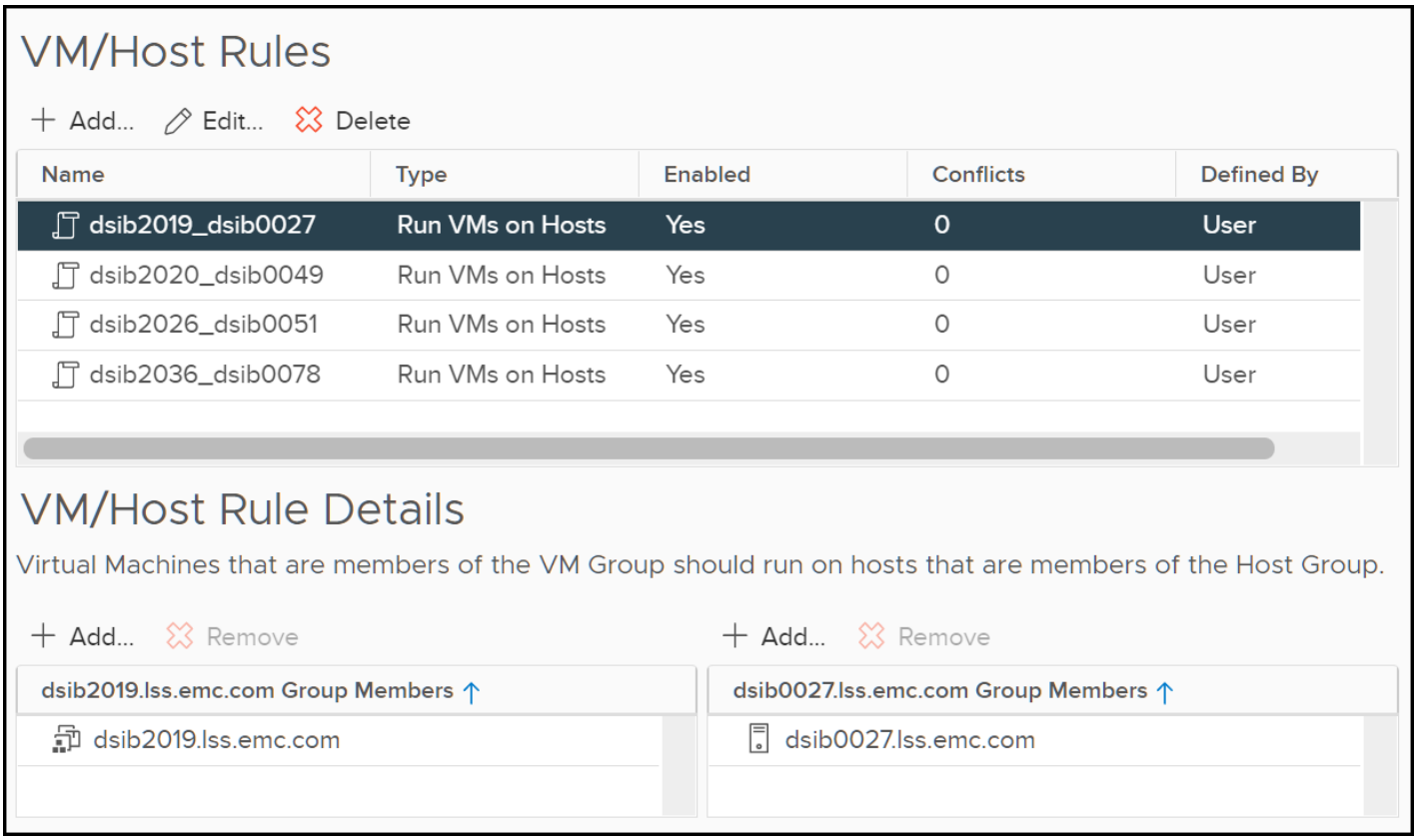

Figure 29. VM/Host rules for recovery site

5 VMware SRM testfailover

Before demonstrating the testfailover it is worthwhile to mention how vVol testfailover differs from VMFS/RDM testfailover on the PowerMax. When using a VMFS/RDM configuration, the PowerMax requires the use of the SRDF SRA which integrates the storage with SRM. SRM makes calls to the SRA and the SRA then performs storage related commands. For example, when running testfailover, the SRDF SRA will take a snapshot of the R2 devices in real time, and then create the snapshot targets and perform the link. These target devices become the test devices which avoids impacting the R2s.

When using vVols, there is no SRDF SRA. Instead, the VASA Provider takes over that role and works in conjunction with SRM. The workflow, however, changes somewhat. Rather than creating real time snapshots, the PowerMax automatically takes snapshots of all the R2 devices in a VASA Replication Group on a five-minute cycle, always retaining five copies on the recovery array. When the user runs a testfailover of a recovery plan, the latest snapshot will be used for that test. The VASA Provider will create new vVol target devices and then link the snapshot to those devices. Like VMFS/RDM, the production devices are never impacted.

5.1 Oracle RAC testfailover without modification

This section will demonstrate an SRM testfailover of the recovery plan without making any modifications to SRM or the VMs. SRM will create a bubble or test network in which to run the VMs based on the default network mapping.

5.1.1.1 Oracle RAC testfailover

The following is the execution of a testfailover after creation of the recovery plan. No modifications to the recovery plan have been made, in particular as shown previously in Figure 16, SRM will create a test network. From within SRM, navigate to the **Recovery Plans** tab and highlight the **RAC_Recovery_Plan** on the left-hand side. Then select **TEST** as shown in step 1 in Figure 30.

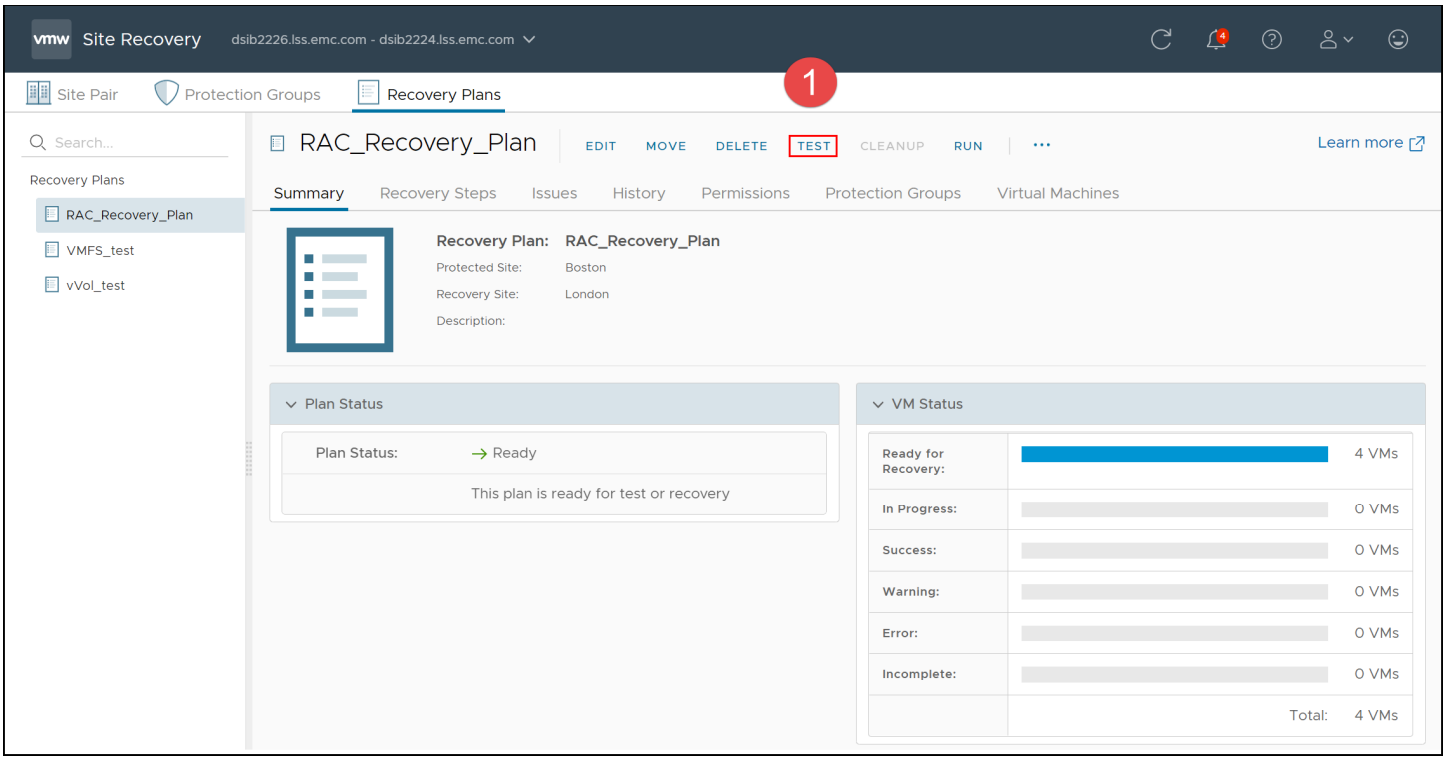

Figure 30. Testfailover: Step 1

In step 2 shown in Figure 31, the box to replicate changes will be checked, though as the message shows it has no impact on the data. Select **NEXT**.

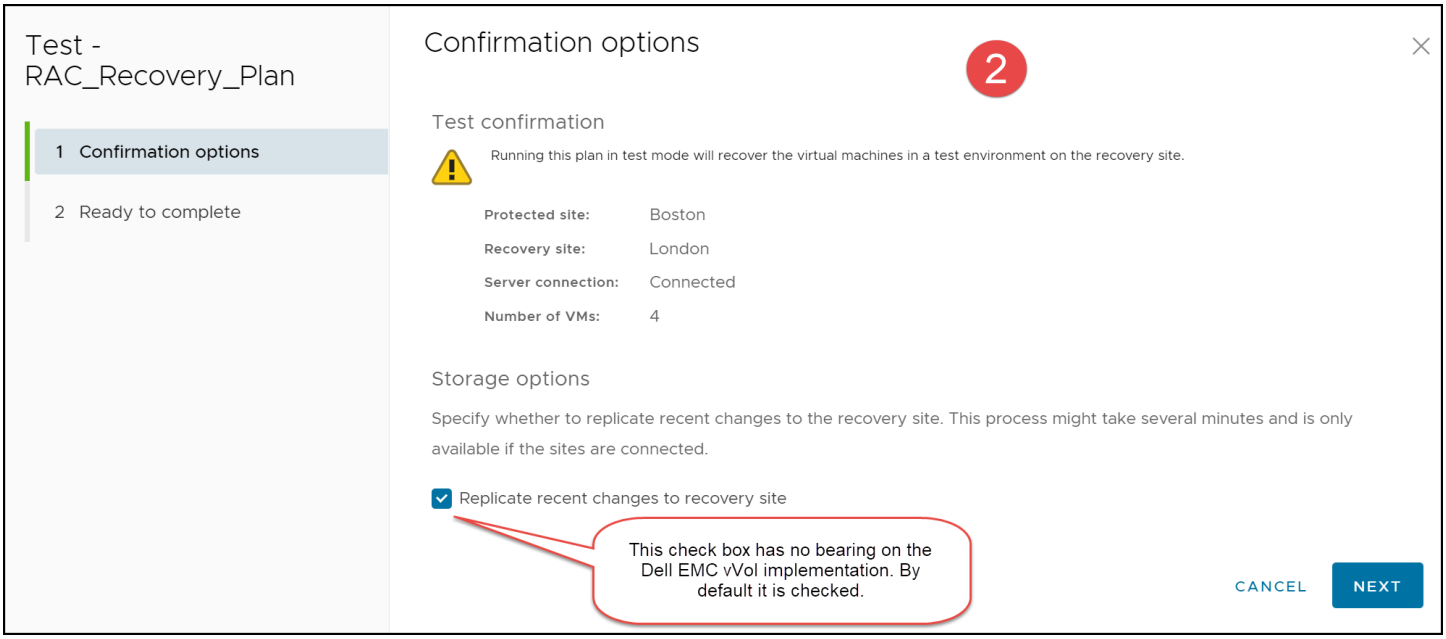

Figure 31. Testfailover: Step 2

In the final step shown in Figure 32, select **FINISH**.

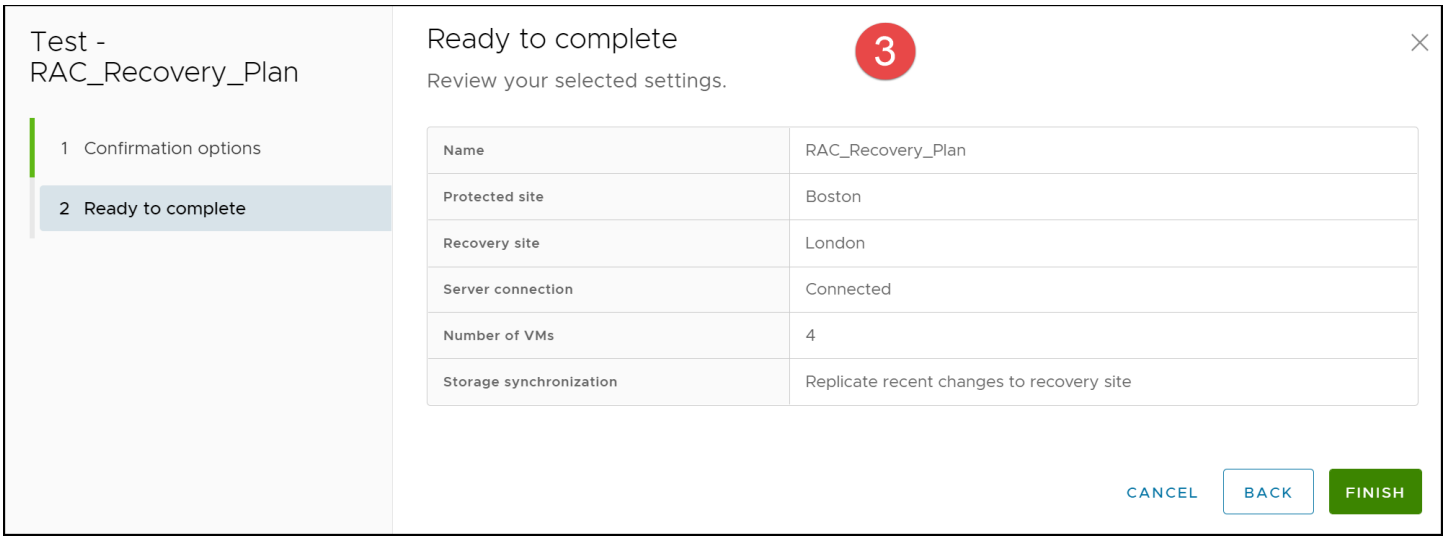

Figure 32. Testfailover: Step 3

The recovery plan finishes in under five minutes and reports a successful result shown in Figure 33.

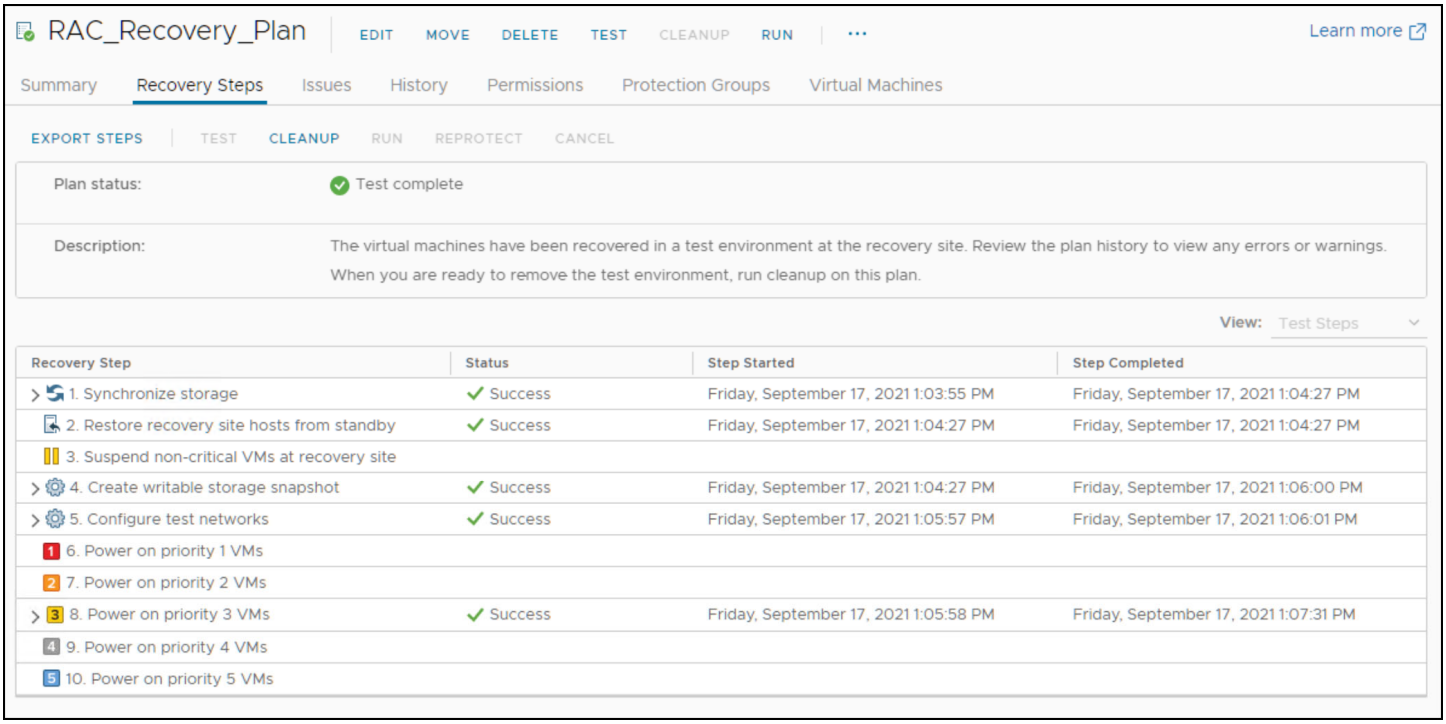

Figure 33. Testfailover result

5.1.1.2 RAC testing

With the testing complete, there is now an isolated four node RAC cluster available for testing. The problem, however, is the test network. Oracle RAC is very reliant on host names and in this test network they are no longer viable. Neither the public nor the private network functions, and thus the Clusterware cannot come up as shown in Figure 34.

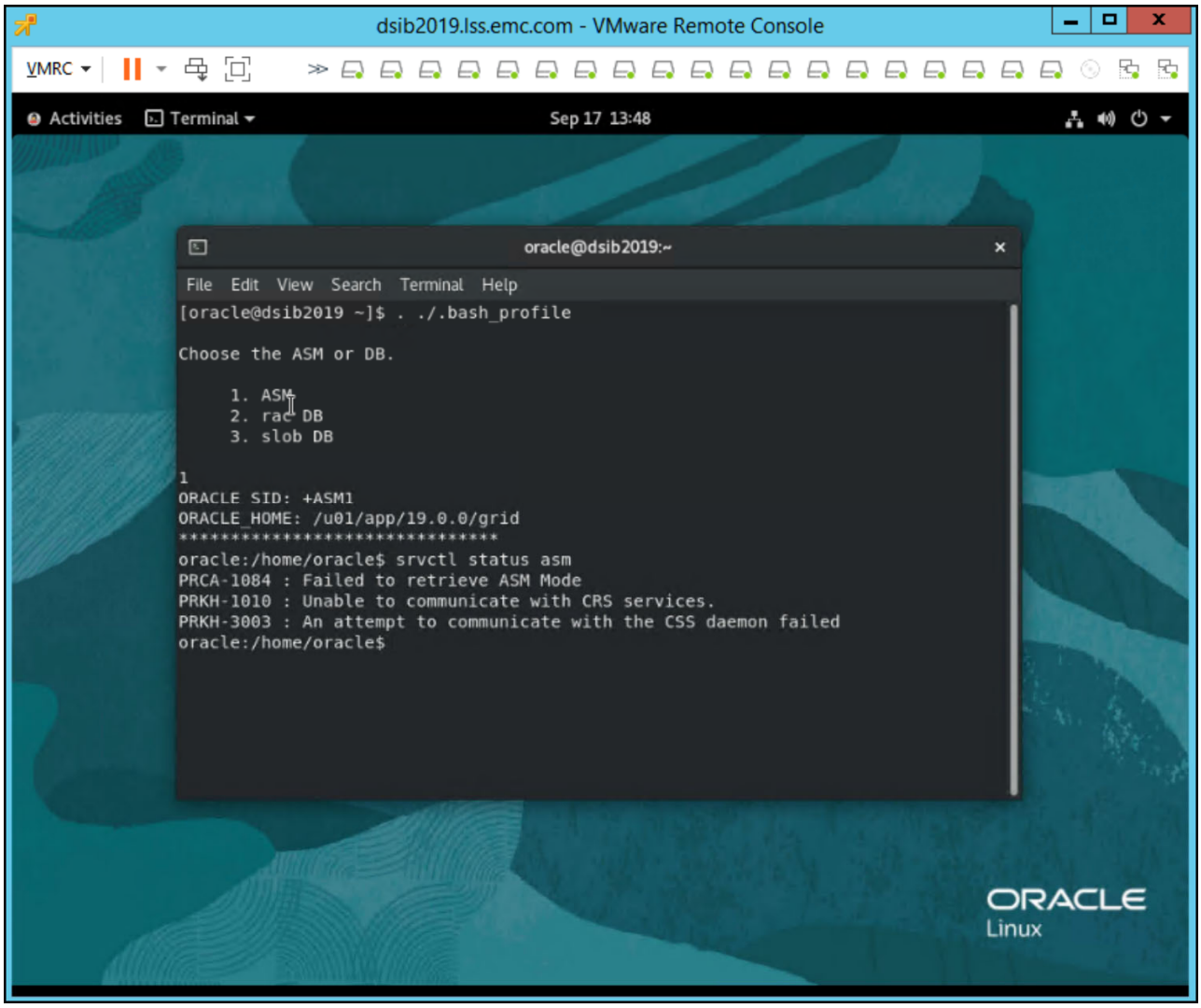

Figure 34. Inoperable Clusterware on SRM test network

Each node has the same issue and without changes to the network, the Oracle RAC database will never come up. Therefore, in order to create a viable test environment, modifications are needed to either the network, the VMs or both. The next section will detail how to use the existing production network on the recovery site by modifying some network components of the RAC nodes themselves so that both production and test nodes can be operational at the same time.

5.2 Oracle RAC testfailover with modification

The following procedure will detail how to automate the SRM post-testfailover modifications to the Oracle RAC nodes (VMs) so that a viable environment is produced with both ASM and the Oracle database operational on all four nodes. As mentioned above, the changes are made within the production network so subnets will remain the same, though IPs are altered to prevent conflict.

Generally, enterprise customers have many network options available, in particular the ability to use VLANs which avoid many of the issues the following procedure is meant to overcome. The process herein assumes a lab environment where only the existing network is available.

5.2.1 Oracle RAC networking

Oracle RAC networking can be complicated to configure and even more so to alter after being successfully implemented. This is mostly due to Oracle's reliance on the host (node) name. The host becomes part of the cluster information and cannot be altered without modifying the Clusterware underlying it. In a scenario with VMware SRM, it is impractical to consider changing the software upon failover as essentially it is a reinstallation. If customers desire using a different set of host (node) names at the recovery site, rather than SRM, it is easier to pre-configure the Clusterware on hosts at the recovery site and avoid replicating those disks. In that scenario it is only necessary to failover the database files, though using SRM is not possible as partial VM failover is not supported. A similar situation is covered later in this paper in the section VMware PowerCLI RAC testing.

5.2.2 SRM testfailover networking

As explained in the previous section, by default, SRM will auto create "test" networks which are isolated. The VMs are assigned to these temporary networks that are not connected to any physical network. Therefore, there is no concern about conflicting IPs with the production environment. While this might be useful for VMs where the network is not critical, it is not practical for an Oracle RAC environment which relies on viable IP addresses to function as previously shown. Assuming that the production nodes must remain online and that a non-conflicting network is not available (e.g. VLAN), the best option is to use the production network at the recovery site but change the IP addresses of the Oracle RAC environment. Recall that Oracle will not tolerate host name changes without reconfiguration of the Clusterware, i.e., reinstallation, so that is not an option.

There are two methods to change the SRM test network. It can be done at the site level (site-level mapping) so that every recovery plan will automatically use the same ones, or it can be done individually on the recovery plan level. Both methods are addressed below.

5.2.2.1 Site-level default test network

By default, when creating the network mappings, VMware will assign the **Test Network** to **Isolated network (auto created)** as shown in Figure 35.

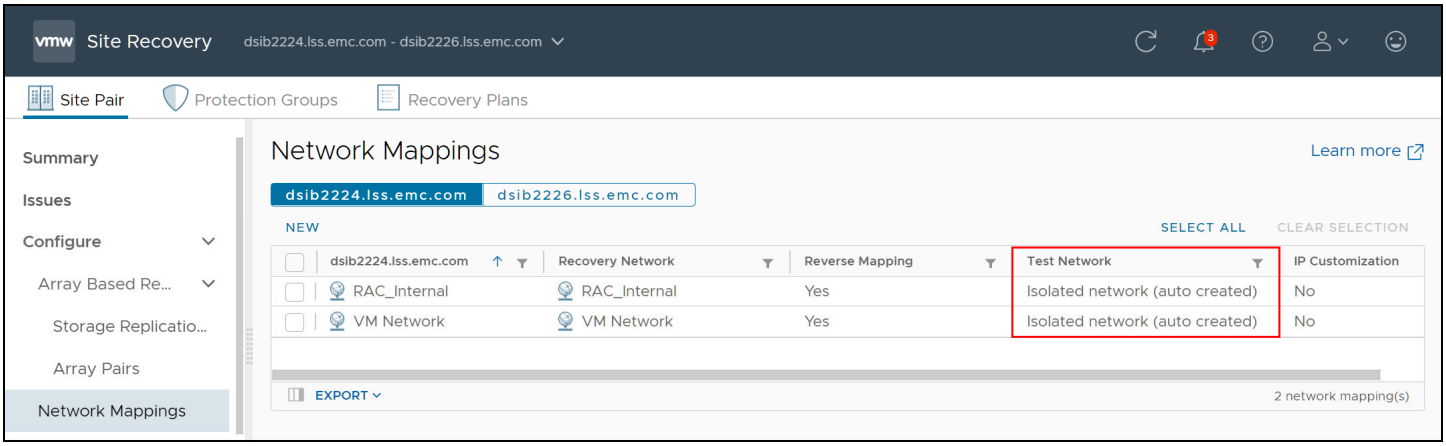

Figure 35. SRM Site-Level Test Network

To change the test network, use the check box to select one of the networks as shown in step 1 in Figure 36.

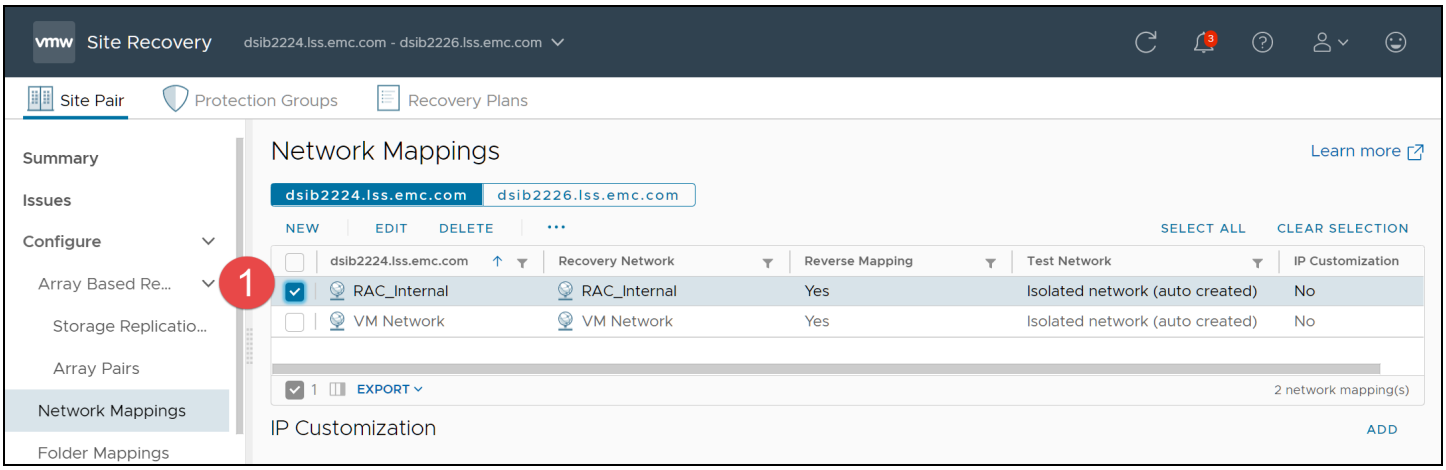

Figure 36. Assign site-level test network: Step 1

In step 2, use the three buttons and select **Edit Test Network Mapping** as shown in Figure 37.

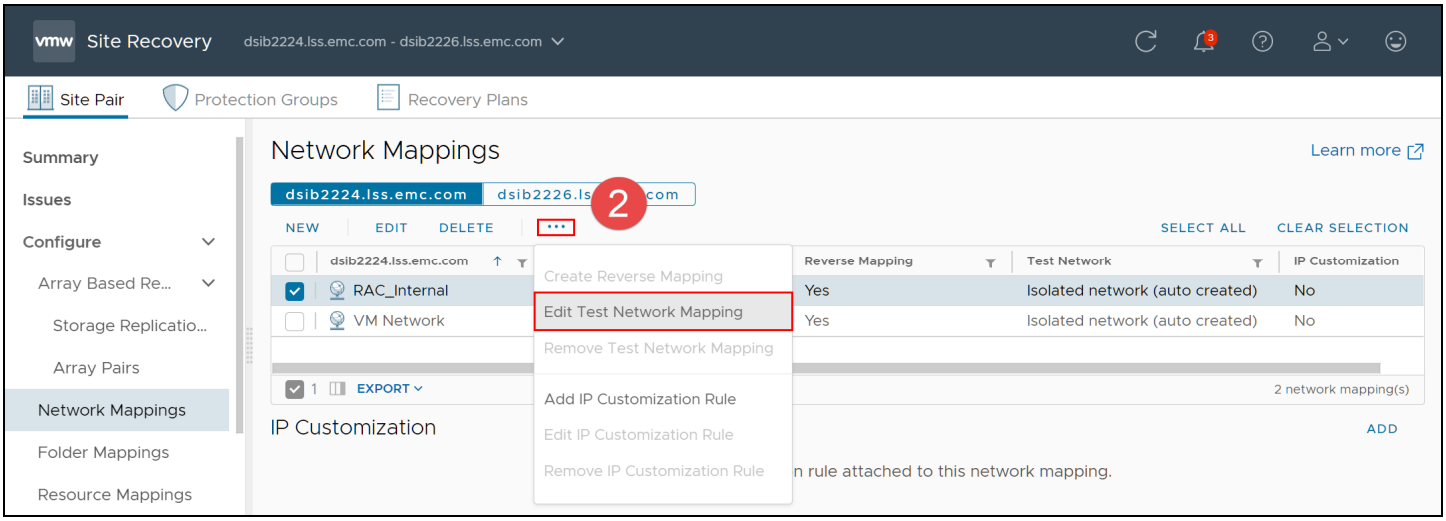

Figure 37. Assign site-level test network: Step 2

Finally in step 3 shown in Figure 38, select the radio button **Select a specific network** and choose the production network at the recovery site that matches the protection site, in this case **RAC_Internal**.

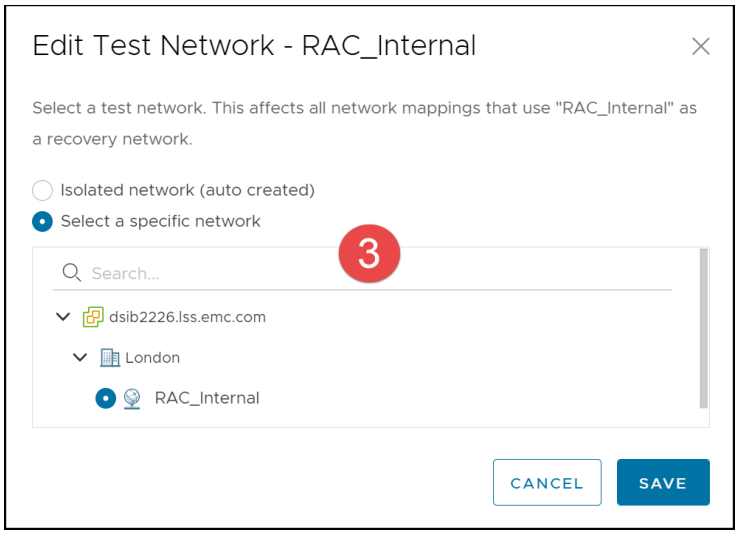

Figure 38. Assign site-level test network: Step 3

Complete the previous steps for the other network used in the Oracle VMs, **VM Network**. The final result required for failover testing is shown in Figure 39.

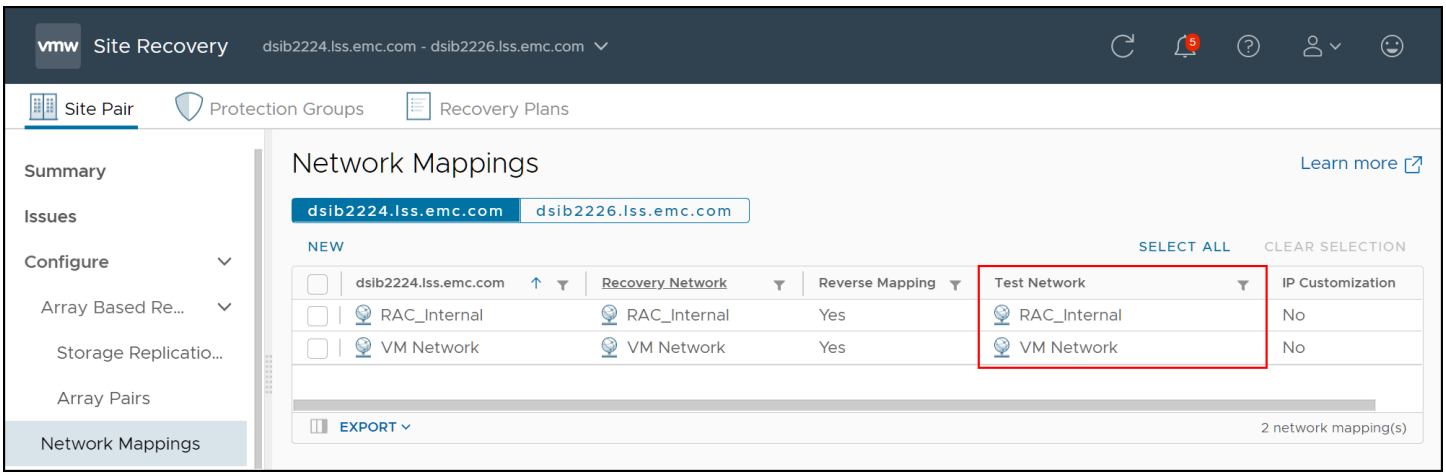

Figure 39. Complete site-level test networks

5.2.2.2 Recovery plan level test network

If multiple recovery plans are being used, and each needs a different test network, it is advisable to change the test network at the recovery plan level rather than at the site level. To do this, begin by highlighting the recovery plan in SRM in the left-hand panel and selecting **EDIT** shown in step 1 on the right-hand panel in Figure 40.

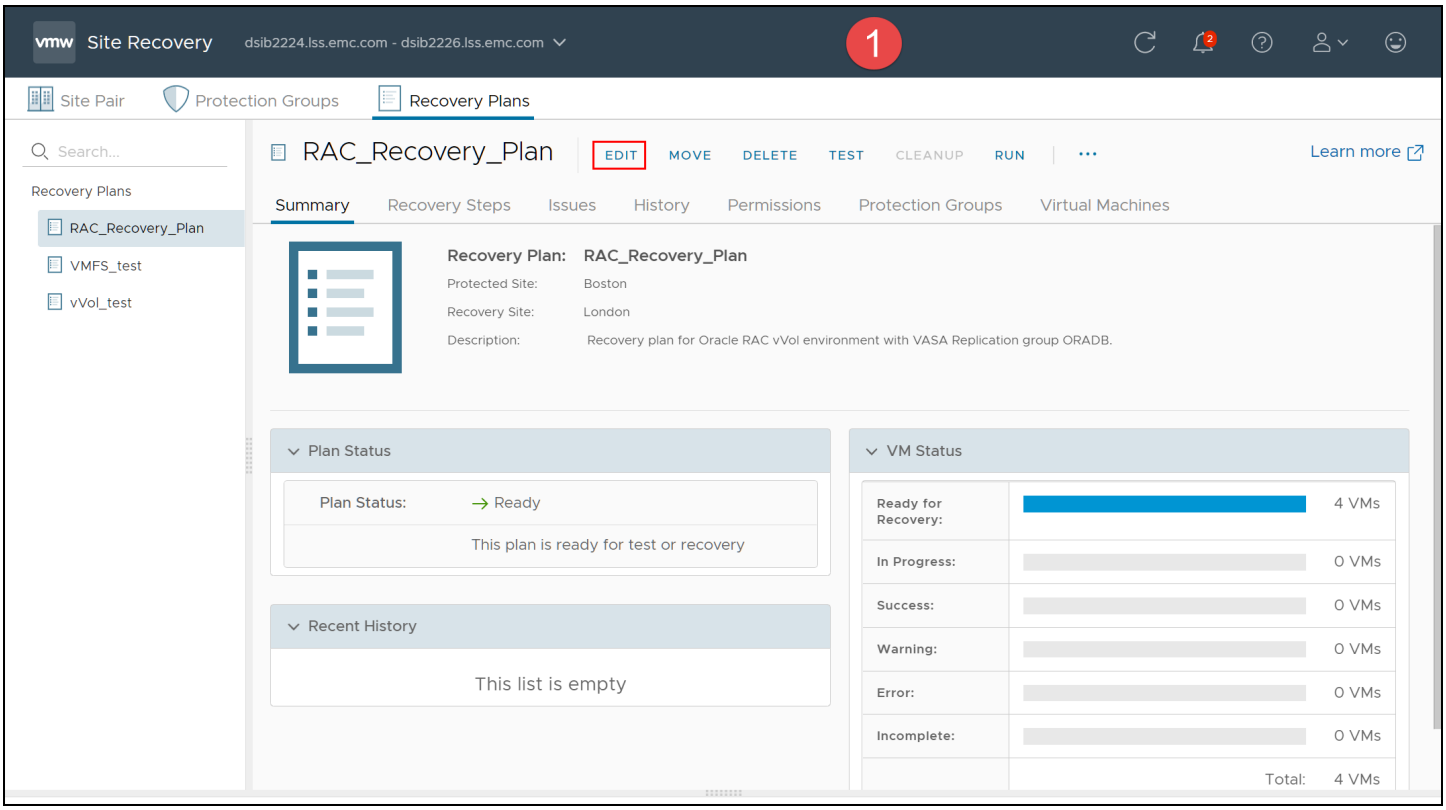

Figure 40. Assign recovery plan-level test network: Step 1

In step 2 shown in Figure 41, select **CHANGE** next to the chosen network.

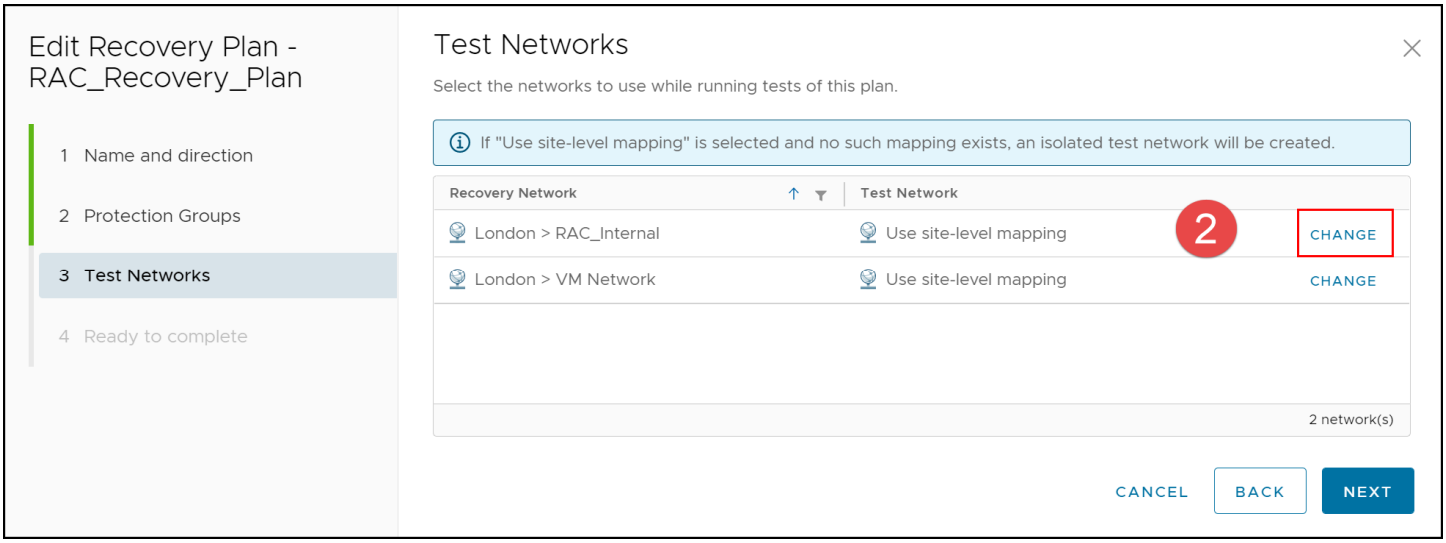

Figure 41. Assign recovery plan-level test network: Step 2

In step 3 shown in Figure 42, select the radio button **Select a specific network** and choose the production network at the recovery site that matches the protection site, in this case **RAC_Internal**.

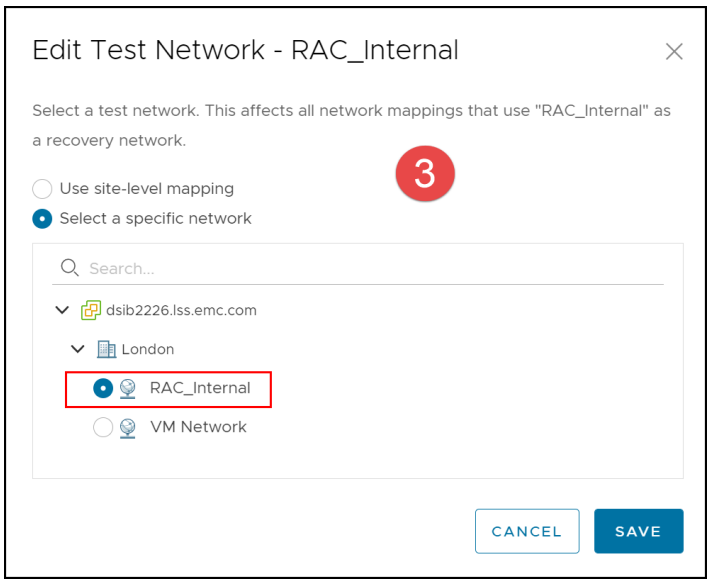

Figure 42. Assign recovery plan-level test network: Step 3

Follow steps 1-3 for the other network **VM Network** shown in step 4 in Figure 43 and then select **NEXT**.

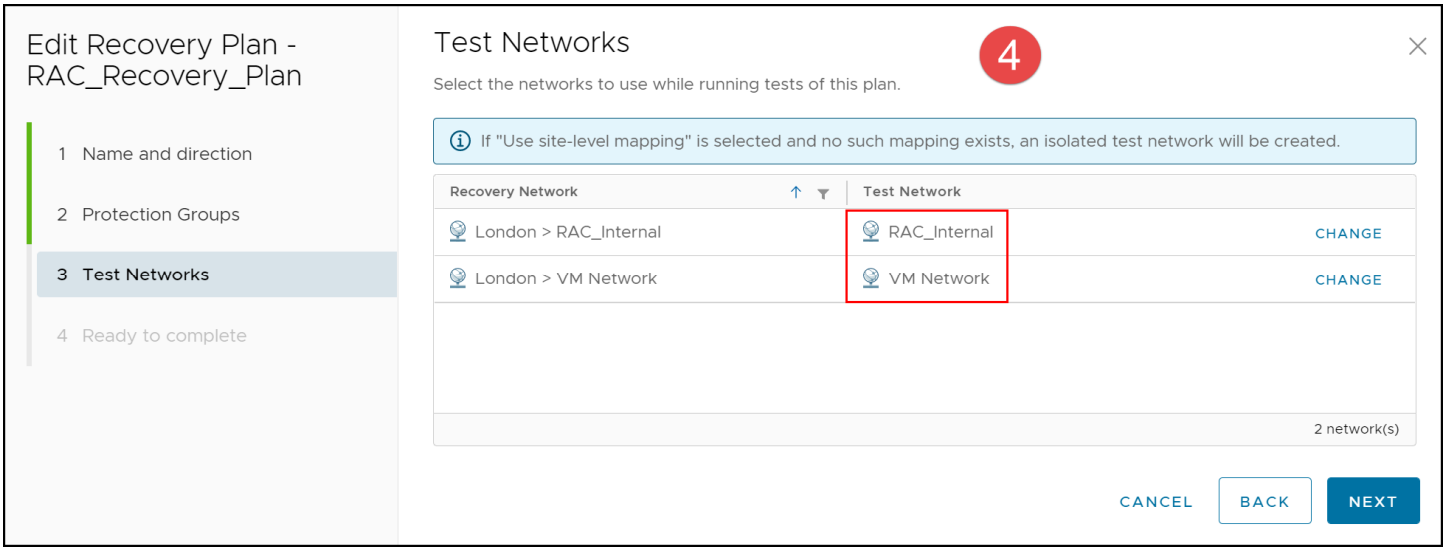

Figure 43. Assign recovery plan-level test network: Step 4

Complete the test network assignment by selecting **FINISH** as shown in Figure 44.

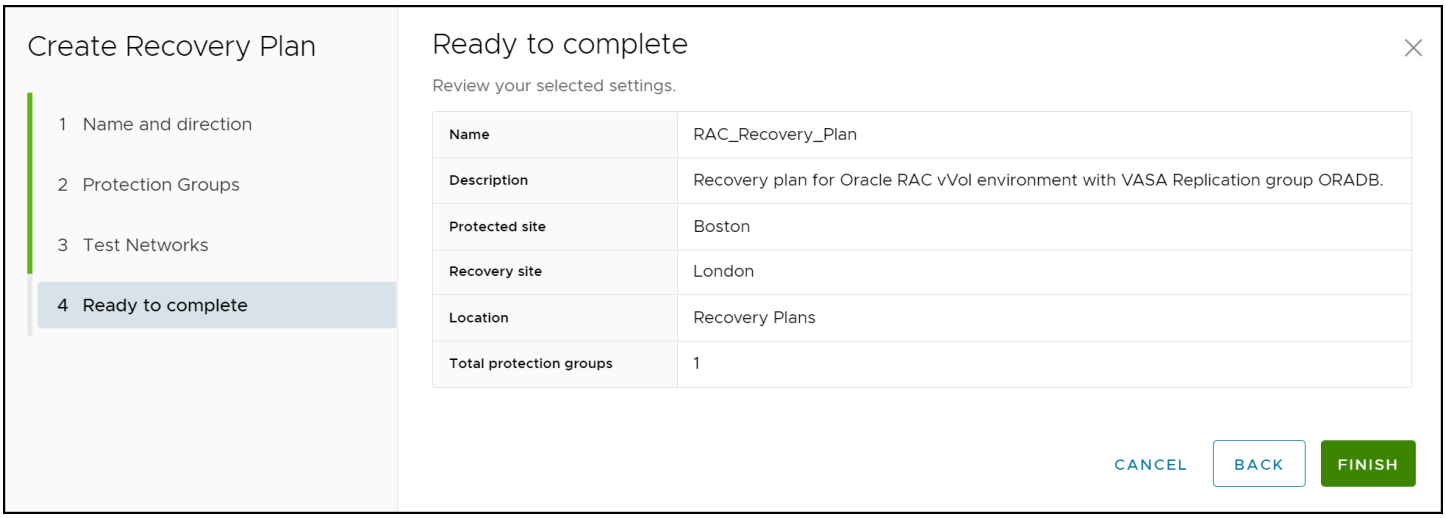

Figure 44. Complete recovery plan-level test network

Now when the recovery plan is executed, it will use the newly configured test networks without changing the site-level test networks.

5.2.3 Oracle RAC – Public IP, VIP, SCAN, and GNS

When configuring Oracle RAC 11.2 and higher, the user can choose between using Virtual Internet Protocol (VIP) addresses and Single Client Access Name (SCAN) or Grid Naming Service (GNS). GNS is designed to simplify management of the networking component as it does not require using the VIP or SCAN. Instead it uses DNS and a DHCP server to assign addresses to the RAC nodes. Using GNS with SRM is possible but requires a different process than is covered in this paper. To understand GNS implementation, please review Oracle support document 946452.1.

In a non-GNS configured Oracle RAC environment, there are four different IP addresses a user assigns to a node: Public IP, Virtual IP (VIP), SCAN, and an internal IP for RAC inter-communication. Of these, the only IP address that can be carried over to the recovery site is the internal IP since it is only accessible between the nodes on the ESXi hosts through the internal network and will not conflict with the production ESXi hosts. The other three addresses need to be altered so as not to conflict with production as in this environment they all share the same network. As noted, as long as the host or node name is not changed, Oracle will permit IP changes for the public, VIP and SCAN addresses. Since the host names will not change, the hosts file on the Oracle nodes will be used instead of DNS. It is also possible to use a different DNS if there is one configured that re-assigns the hostnames to the new IPs.

Two of the IP changes will be initiated by a Post Power On script, but the public IP will be changed directly by VMware through the SRM interface. As the hosts file is essential in this configuration, a new one will be generated that will replace the original at the test site.

5.2.3.1 Hosts file

The two hosts files are listed below. The first is from the protection site, the original *hosts* file, and the second from the recovery site, *hosts.testfailover*. The recovery site hosts file (hosts.testfailover) is stored on the production nodes so it is available on the test nodes once the testfailover is run. The script will then replace the original *hosts* with it. Note that the private IPs of the four nodes, named **<host>-priv**, remain the same between the files as explained above.

hosts

```
10.228.246.19 dsib2019.lss.emc.com dsib2019
10.228.246.20 dsib2020.lss.emc.com dsib2020
10.228.246.26 dsib2026.lss.emc.com dsib2026
10.228.246.36 dsib2036.lss.emc.com dsib2036
10.228.246.21 dsib2021.lss.emc.com dsib2021
10.228.246.22 dsib2022.lss.emc.com dsib2022
10.228.246.28 dsib2028.lss.emc.com dsib2028
10.228.246.37 dsib2037.lss.emc.com dsib2037
192.168.1.150 dsib2019-priv
192.168.1.151 dsib2020-priv
192.168.1.152 dsib2026-priv
192.168.1.153 dsib2036-priv
10.228.246.23 dsib-scan dsib-scan.lss.emc.com
10.228.246.24 dsib-scan dsib-scan.lss.emc.com
10.228.246.25 dsib-scan dsib-scan.lss.emc.com
```

hosts.testfailover

```
10.228.246.137 dsib2019.lss.emc.com dsib2019
10.228.246.138 dsib2020.lss.emc.com dsib2020
10.228.246.139 dsib2026.lss.emc.com dsib2026
10.228.246.140 dsib2036.lss.emc.com dsib2036
10.228.245.214 dsib2021.lss.emc.com dsib2021
10.228.245.215 dsib2022.lss.emc.com dsib2022
10.228.245.216 dsib2028.lss.emc.com dsib2028
10.228.245.217 dsib2037.lss.emc.com dsib2037
192.168.1.150 dsib2019-priv
192.168.1.151 dsib2020-priv
192.168.1.152 dsib2026-priv
192.168.1.153 dsib2036-priv
10.228.246.161 dsib-scan dsib-scan.lss.emc.com
10.228.246.162 dsib-scan dsib-scan.lss.emc.com
10.228.246.144 dsib-scan dsib-scan.lss.emc.com
```

5.2.3.2 Post Power On Script

The basic script (no error handling) below is designed to make all the necessary changes to the IP addresses which will enable Oracle RAC to operate in the test environment. The script has multiple steps to accomplish the IP changes after the VMs power on at the recovery site. Before SRM executes the script, it changes the public IP which is essential to the successful running of the script. There are four different scripts, as each must be customized to the individual RAC node. Each customized script should be placed on the appropriate production node so it is available after testfailover. Only the script for node dsib2019 is included below. Each script takes the following actions:

- The script begins by sleeping ten minutes because the Clusterware must be given time to start as the VIP and SCAN changes are driven by Oracle Server Control (`srvctl`) which relies on the Clusterware daemons.
- After ten minutes, the Clusterware environment (e.g. GRID_HOME) is sourced by root.
- Before the IPs can be changed, the old hosts file is then removed and replaced with the *hosts.testfailover* file which points to the new IPs.

- At this point, with the Clusterware running and the hosts pointing to new IPs, Oracle Server Control can be used to update the network. It is essential for Clusterware to be operational before the hosts file is changed, otherwise it would fail to start and the IP changes could not be made.
- Using the command `srvctl modify network`, a new IP is assigned for VIP at the appropriate NIC and network number. It's necessary to stop and start VIP for the changes to take effect.
- For the SCAN IPs, it is unnecessary to directly assign the IPs. Instead the services for SCAN and the SCAN_LISTENER are stopped and then the command `srvctl modify scan` is run which will pick up the new IPs from the changed hosts file. Note that since only the hosts file is being used here, only a single IP is going to be selected, though it won't necessarily select the first one. It is possible to setup a local DNS on the host so that the scan name is associated with all three IPs.
- Restart the SCAN and SCAN_LISTENER services.

ora_change_IPs.sh

```
#!/bin/sh
#
#    AUTHOR
#    drew.tonnesen@dell.com
#
#    NAME
#      ora_change_IPs.sh
#
#    DESCRIPTION
#      Script to assign new IPs for VIP and SCAN to Oracle RAC in SRM
testfailover environment
#
#    NOTES
#      Runs as root. Must sleep 10 minutes for Clusterware to start. Relies on
second file, hosts.testfailover, with new IPs.
#
#    MODIFIED                    (MM/DD/YYYY)
#      drew.tonnesen@dell.com    06/30/2021 - Creation
#
sleep 600
. /home/oracle/.grid_profile
rm -f /etc/hosts
cp -f /home/oracle/hosts.testfailover /etc/hosts
srvctl modify network -netnum 1 -subnet 10.228.245.214/255.255.252.0/ens192
srvctl stop vip -n dsib2019 -f
srvctl start vip -n dsib2019
srvctl stop scan_listener
srvctl stop scan
srvctl modify scan -scanname dsib-scan
srvctl start scan
srvctl start scan_listener
exit 0
```

With the script in place, the changes to the recovery plan can be added in SRM. First the new public IP will be assigned, then the script added as a Post Power On job.

5.2.4 Public IP and Post Power On Steps changes

The following will use the **Configure Recovery** wizard to change the public IP of each VM and add a Post Power On script. The public IP can be changed through customization within SRM without a script. In SRM in step 1 highlight the recovery plan in the left-hand panel, then use the checkbox to select the first RAC VM as shown in Figure 45.

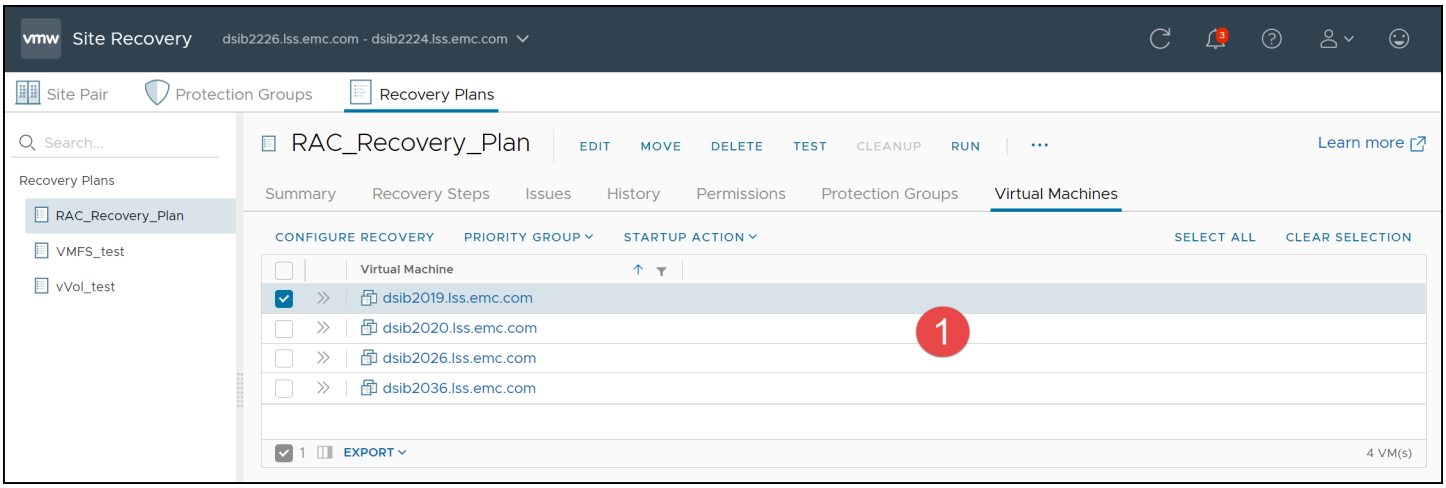

Figure 45. Configure Recovery: Step 1

Next, in step 2, right-click on the VM and select **Configure Recovery** from the **Actions** menu as shown in Figure 46.

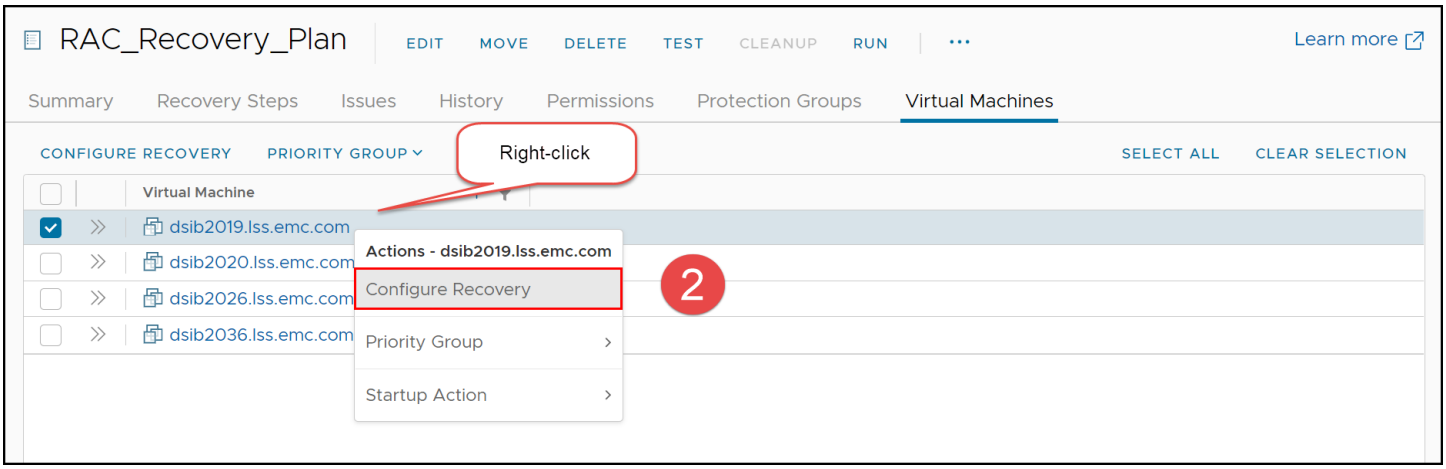

Figure 46. Configure Recovery: Step 2

Selecting this brings up the dialog where the user can customize how VMware treats the VM during failover. By default, the **Recovery Properties** tab is selected. Select the **IP Customization** tab shown in step 3 and 4 in Figure 47.

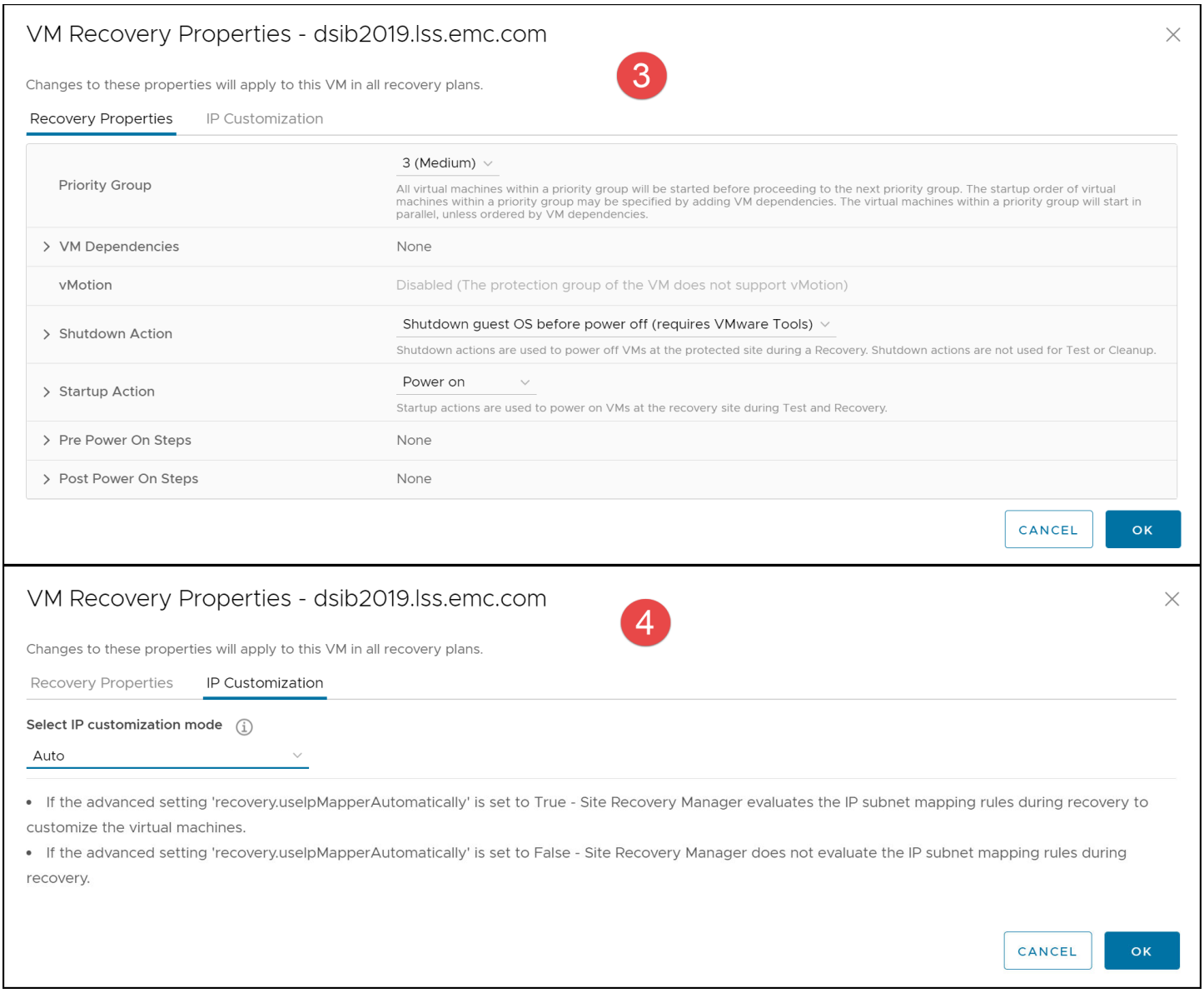

Figure 47. Configure Recovery: Steps 3-4

In step 5, under **Select IP customization mode,** use the drop-down box to select **Manual IP customization** shown in Figure 48.

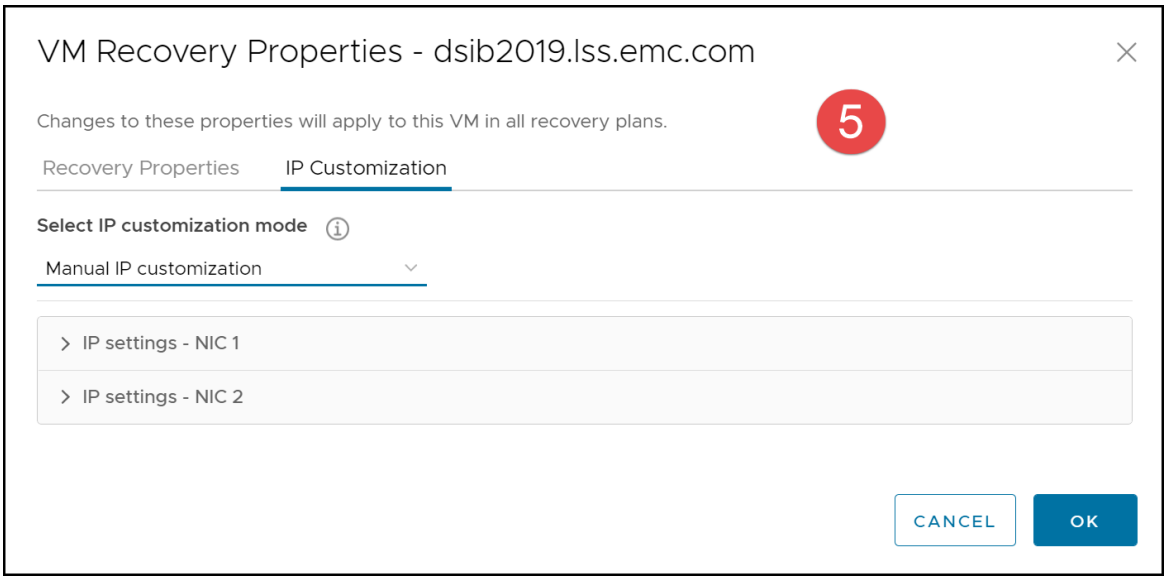

Figure 48. Configure Recovery: Step 5

In step 6 expand the first selection, **IP settings – NIC 1**, and select the **CONFIGURE** button next to the **Recovery Site**. This is the public IP NIC that needs to be updated, while NIC 2 is the private one which does not require alteration.

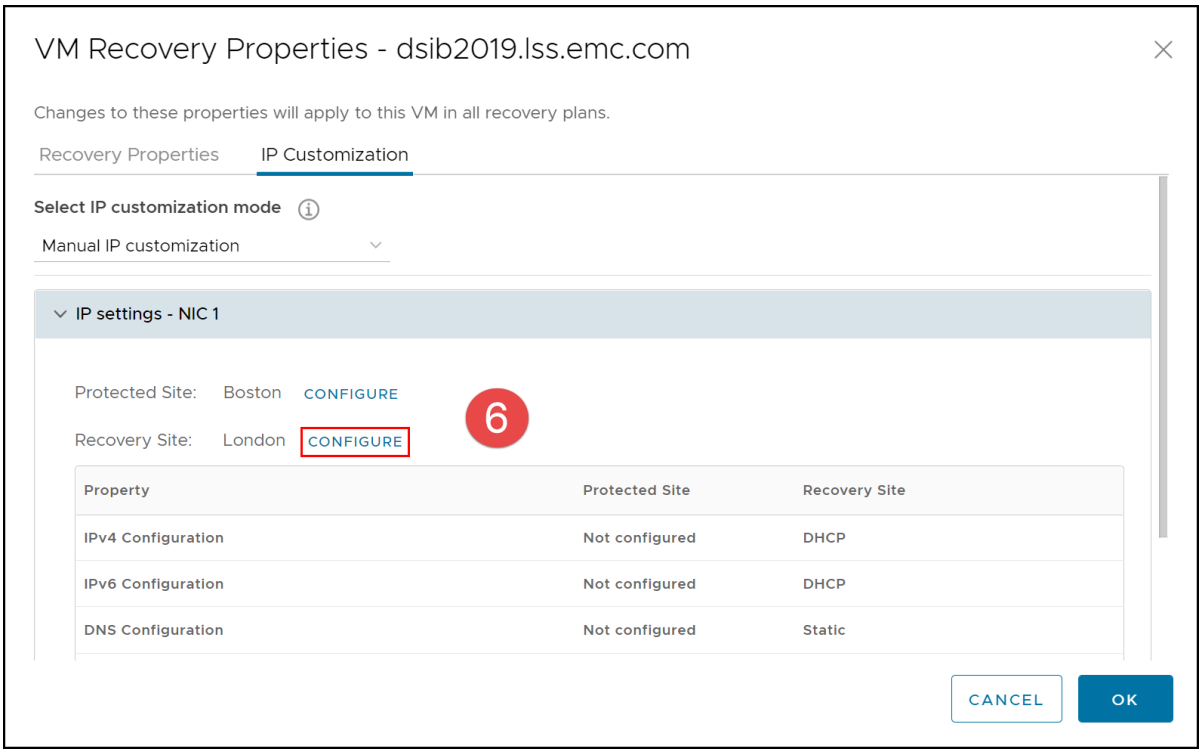

Figure 49. Configure Recovery: Step 6

In step 7 shown in Figure 50, enter the new IP address. This should be the IP that is matched to the hostname in the new hosts file. Select **OK** to save.

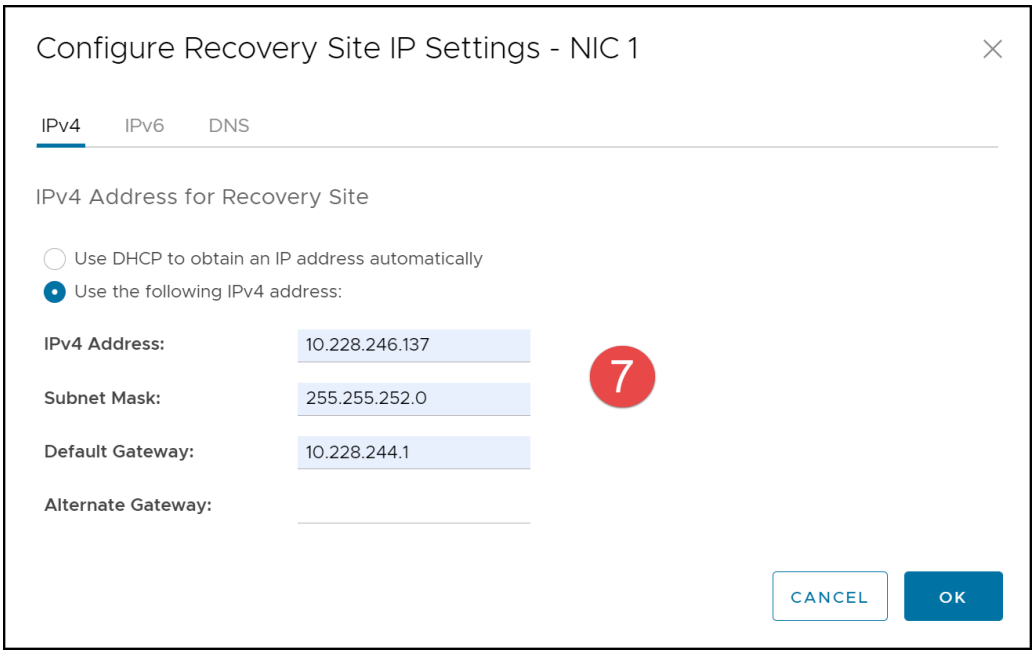

Figure 50. Configure Recovery: Step 7

Step 8 shown in Figure 51 demonstrates the final result. Do not select **OK** or the dialog will close. Rather, return to the **Recovery Properties** tab at the top.

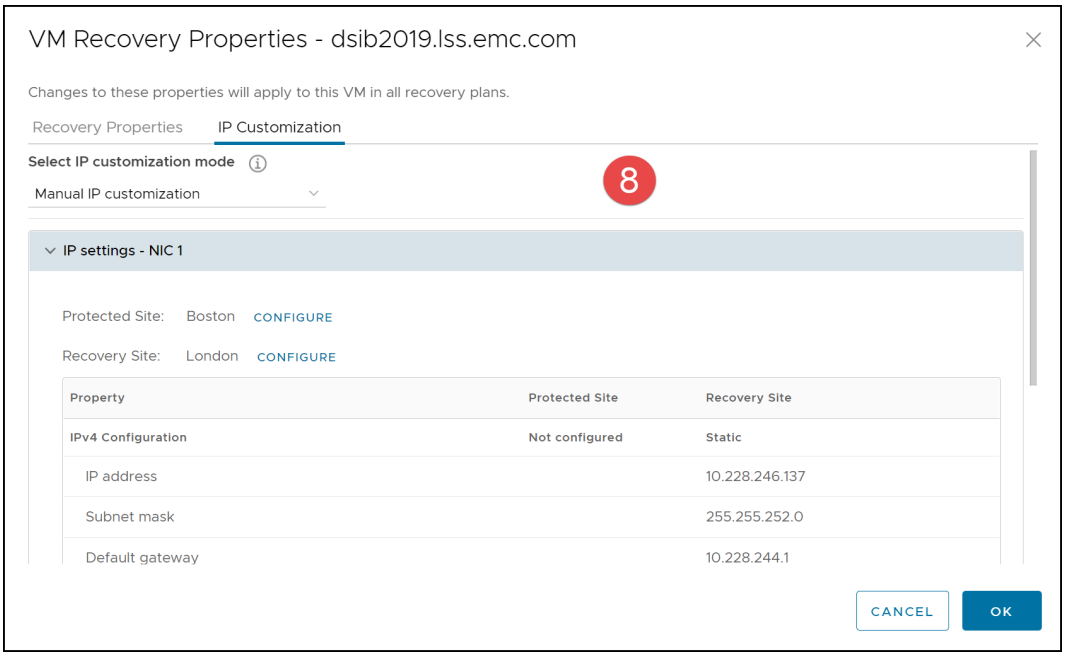

Figure 51. Configure Recovery: Step 8

In step 9 shown in Figure 52, expand the bottom option, **Post Power On Steps**.

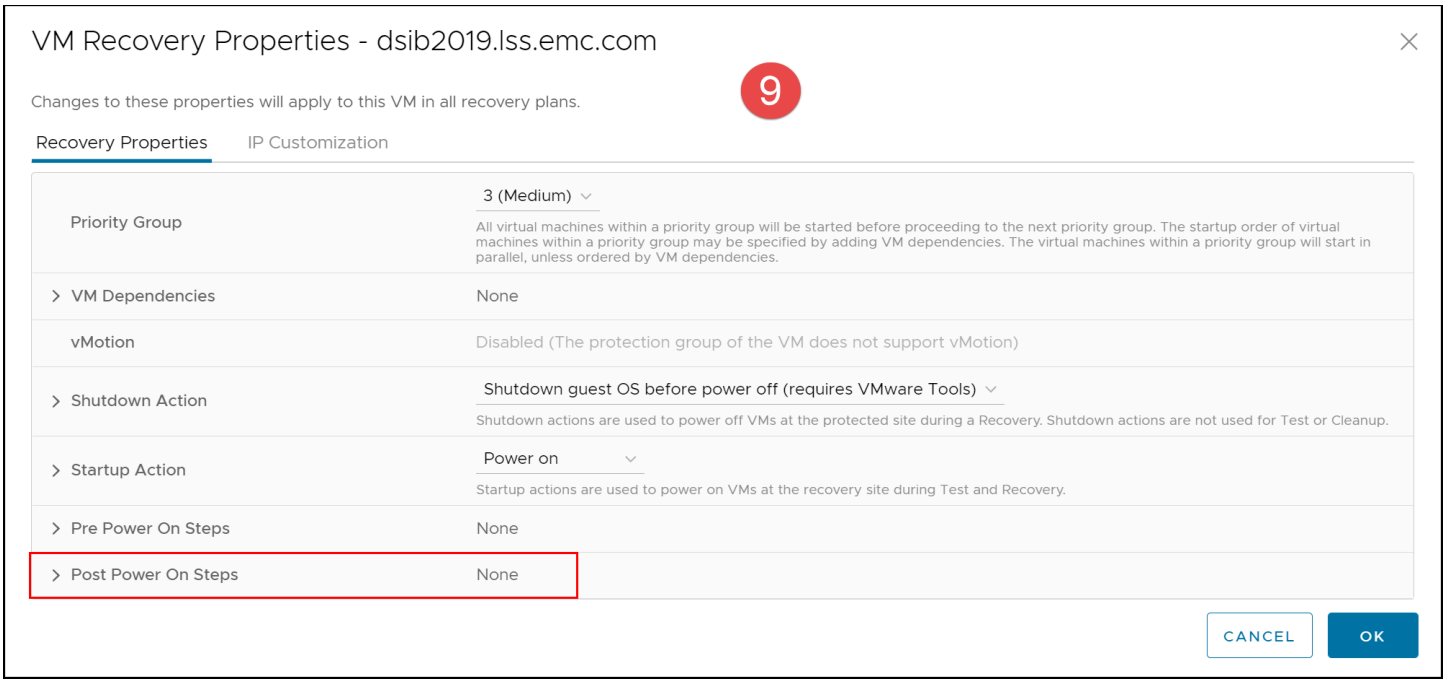

Figure 52. Configure Recovery: Step 9

Select the **+ NEW** button to add a step to the VM recovery. This is shown in Figure 53.

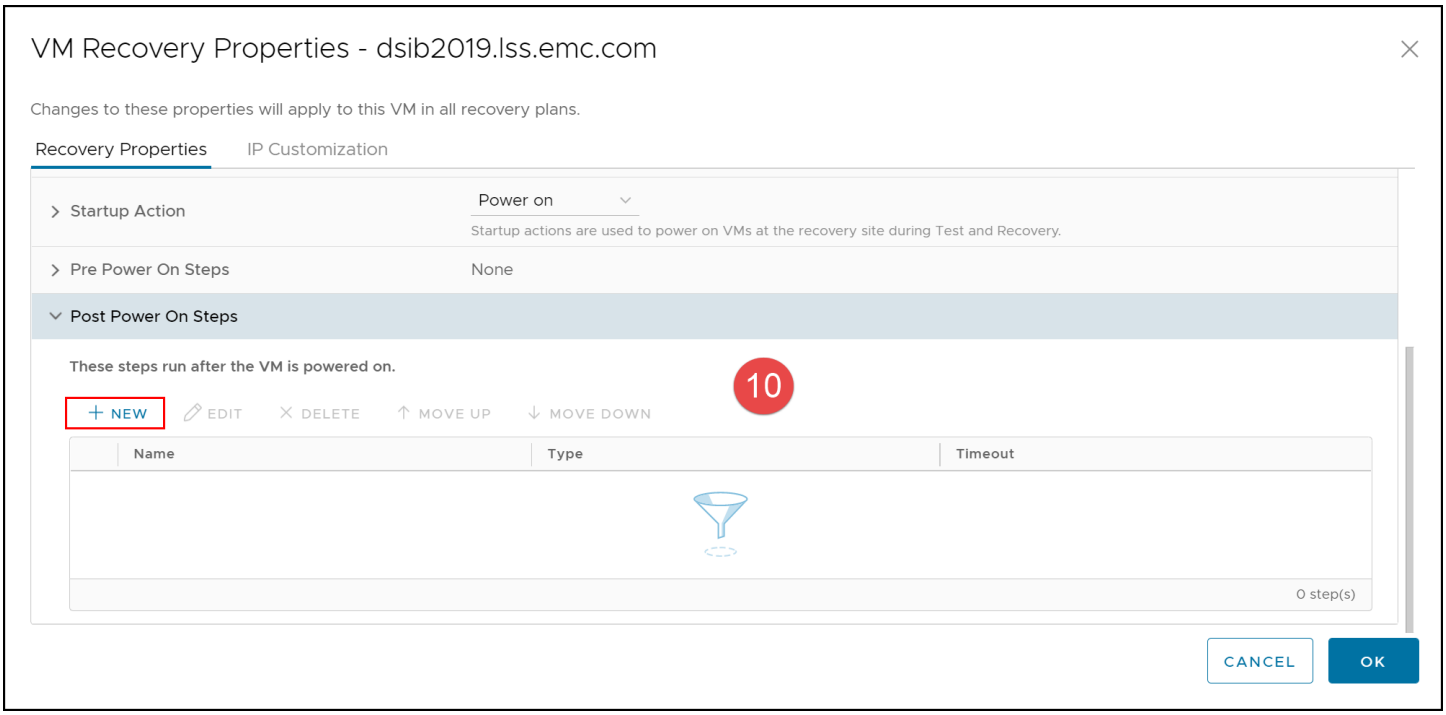

Figure 53. Configure Recovery: Step 10

There are three options available when executing a step for the post power on VM recovery shown in step 11 in Figure 54. In this example select **Command on Recovered VM** because the script is on the recovered VM and that is where it needs to be run. Provide a name to the step. In the **Content** box, add the the script. Here, a shell is included in the command and the script is redirected to a log that can be reviewed after the recovery. Most important, change the **Timeout** from the default of five minutes to fifteen minutes. Remember

the script will sleep for ten minutes so that the Clusterware can come up, which means this post-recovery step needs to account for that. Finally select **ADD** to finish.

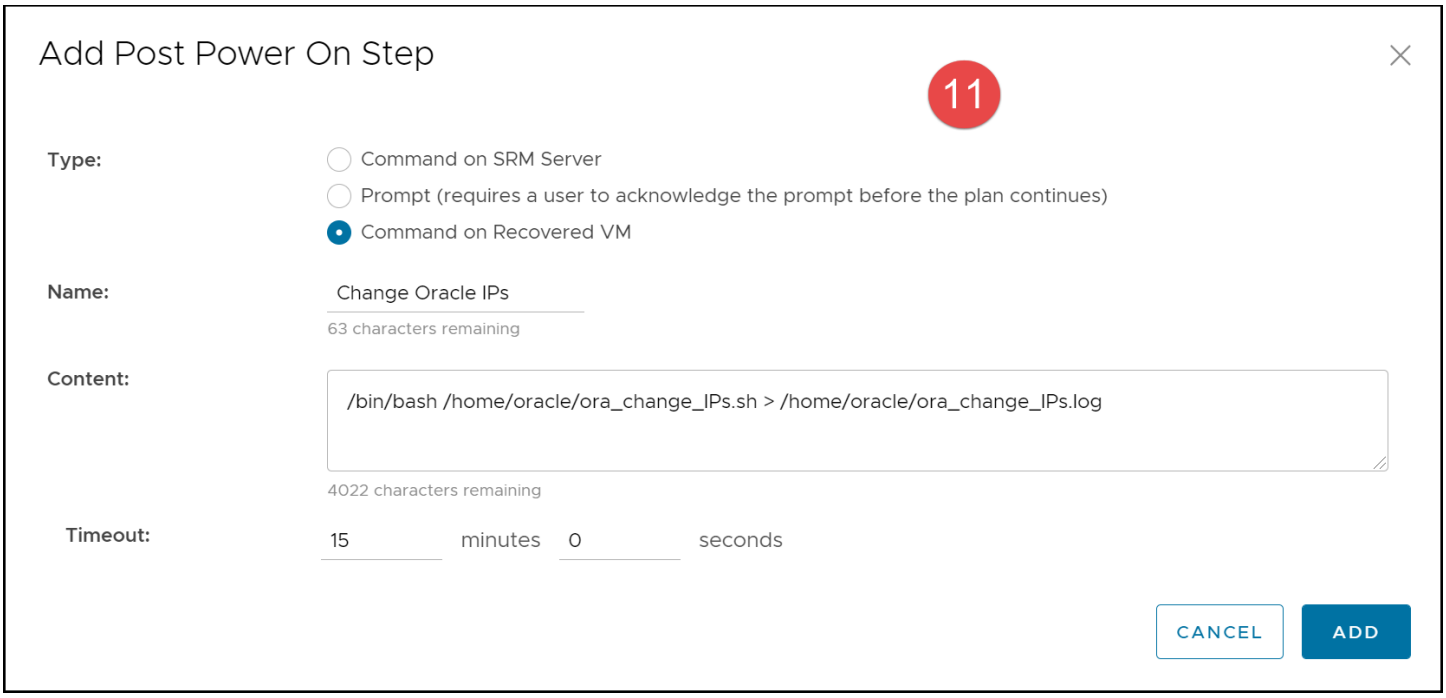

Figure 54. Configure Recovery: Step 11

Step 12 shown in Figure 55 now lists a new **Post Power On Step**. Select **OK** to save it which will also save the public IP change.

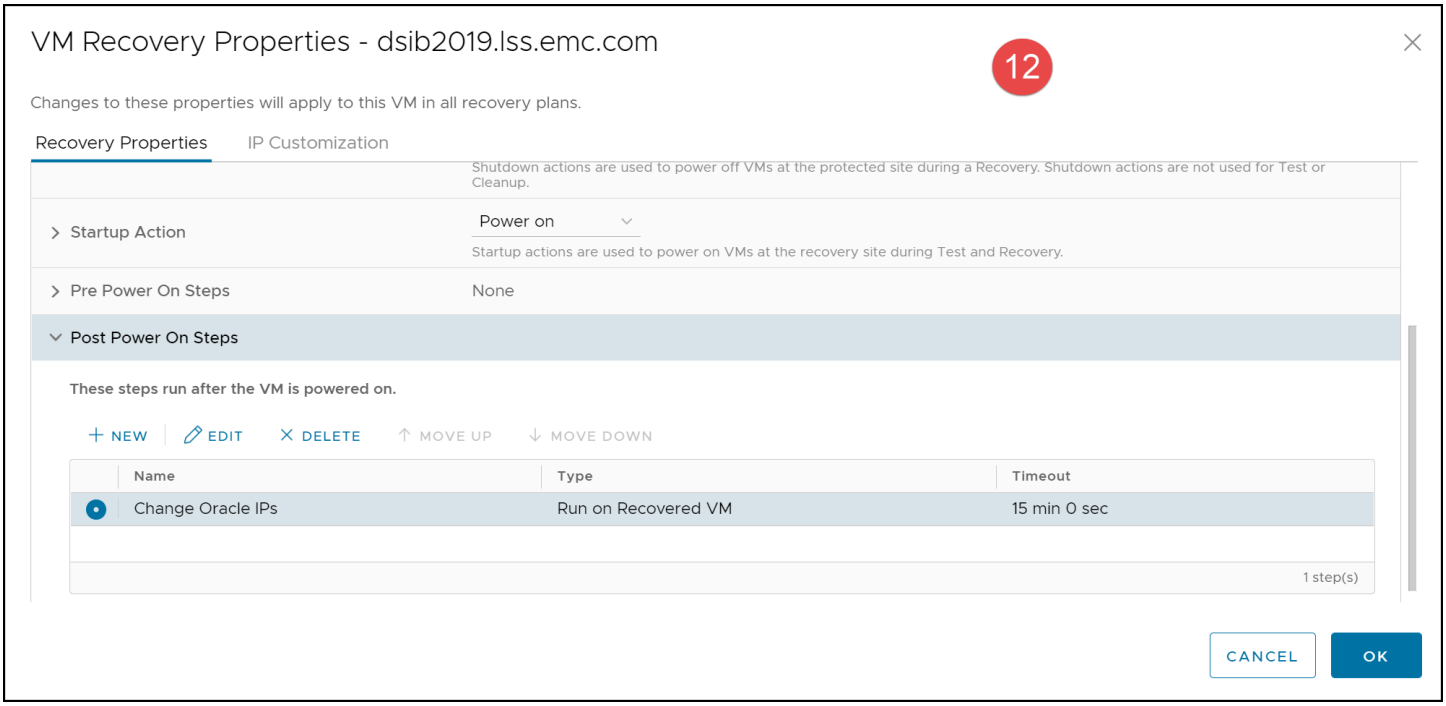

Figure 55. Configure Recovery: Step 12

Figure 56 shows the result of the modifications. These steps need to be repeated for each of the other three RAC VMs before the testfailover is run. Note that Oracle RAC nodes can start in any order so it is unnecessary to prioritize one over another.

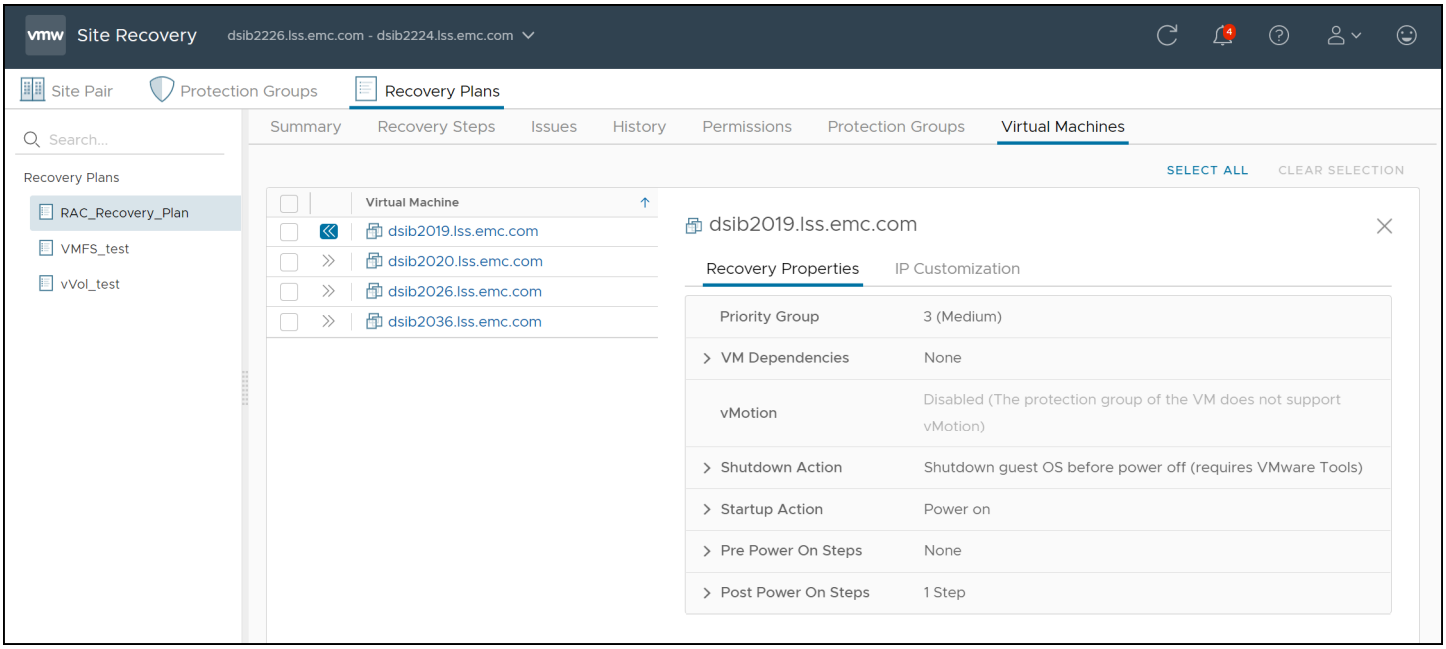

Figure 56. Configure Recovery result

5.2.4.1 Testfailover

With the changes complete, run the testfailover in SRM for the recovery plan. Because of the script, the plan will take more than fifteen minutes to complete, far longer than the three minutes when making no alterations to the plan; yet the storage actions themselves do not take any longer than without the script present.

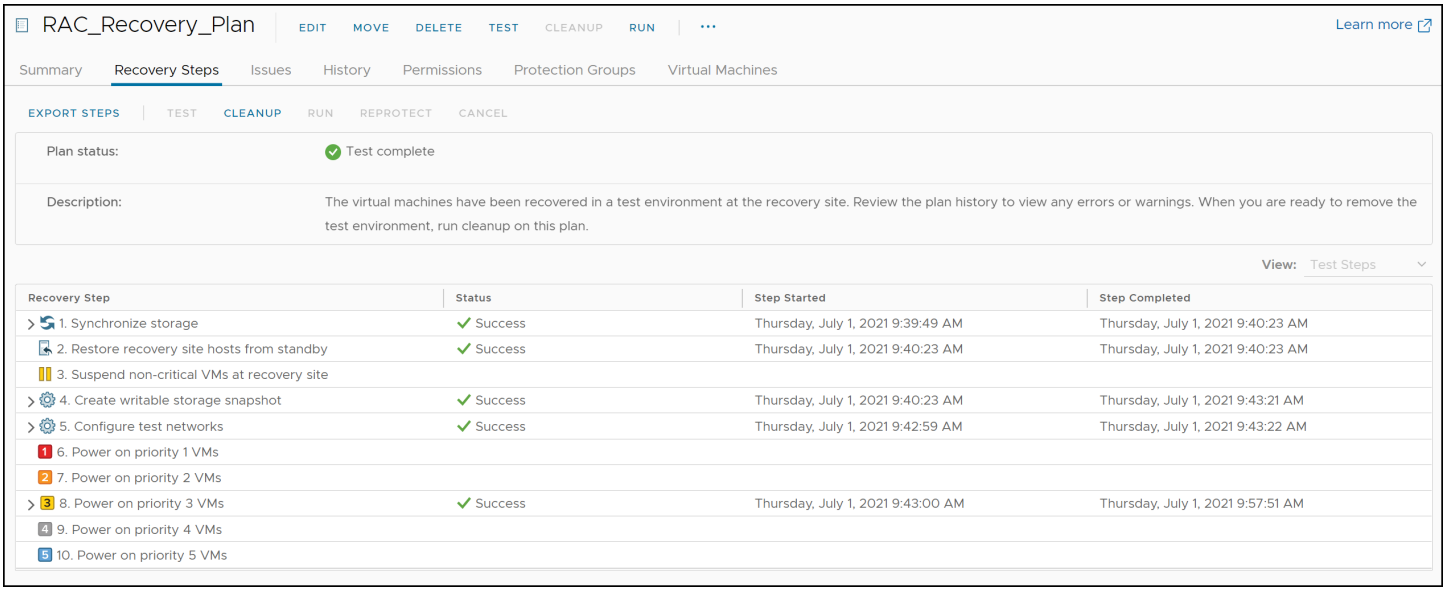

Figure 57. Testfailover with modified recovery plan

After the recovery plan completes, SSH or use VMware remote console into the box to check if the changes were made. Note that when using SSH the new public IP is used, as using the hostname will resolve to the old IP unless the accessing box has a different DNS. Once accessed, run the OS command `ifconfig` at the shell prompt to retrieve IP information. Shown in Figure 58, each new IP in the Oracle RAC environment is colored and labeled. Note that the public IP is changed by VMware during the initial steps of testfailover before the post-script is executed. The post-script changes the virtual IP (VIP) and the SCAN IPs. Recall that as the internal network is shielded from the production vCenter network ESXi hosts, it was unnecessary to change it.

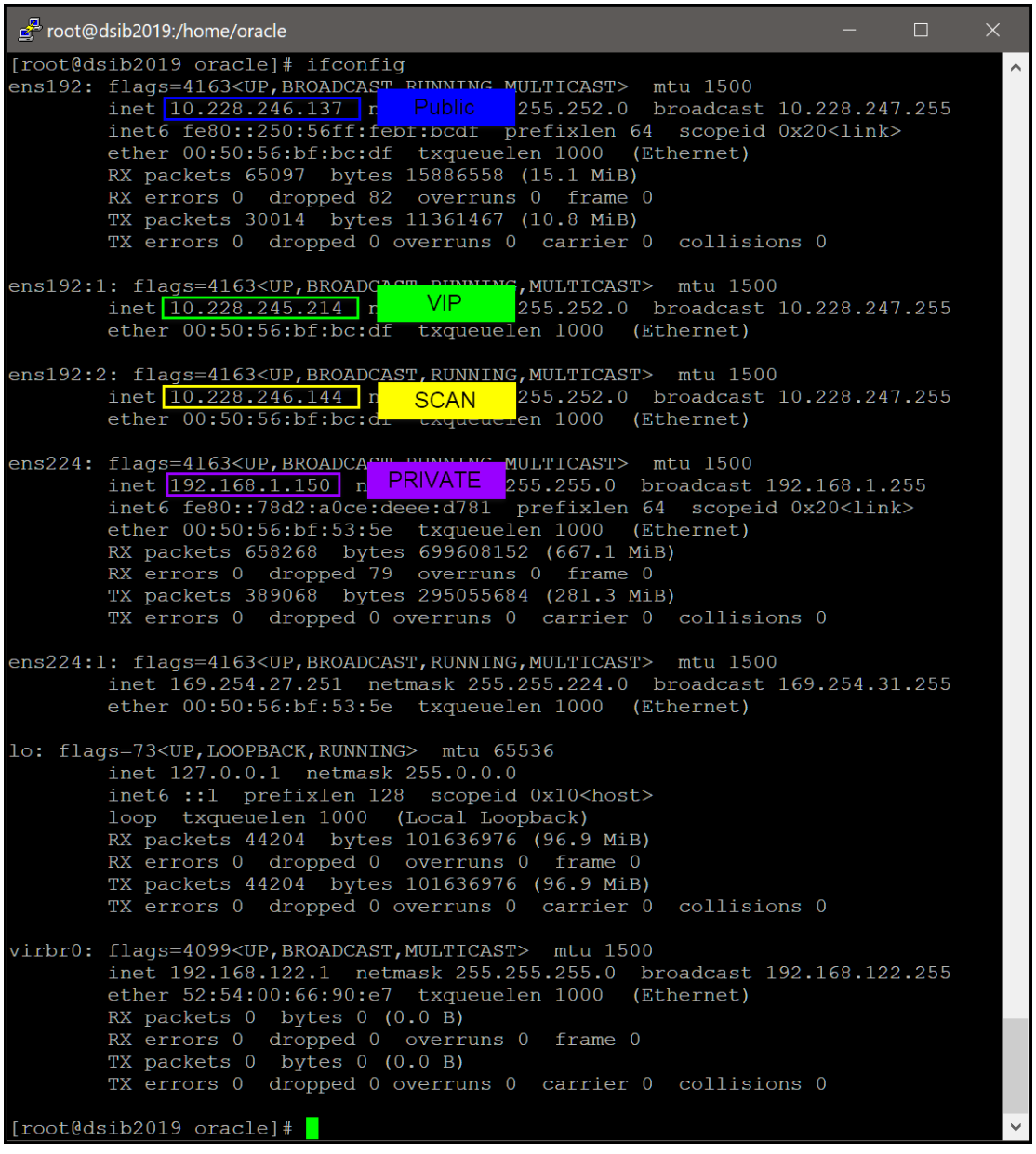

Figure 58. IP changes in recovery site VM

With the IP changes made, one can see in Figure 59, both ASM and the RAC database are fully operational on all nodes.

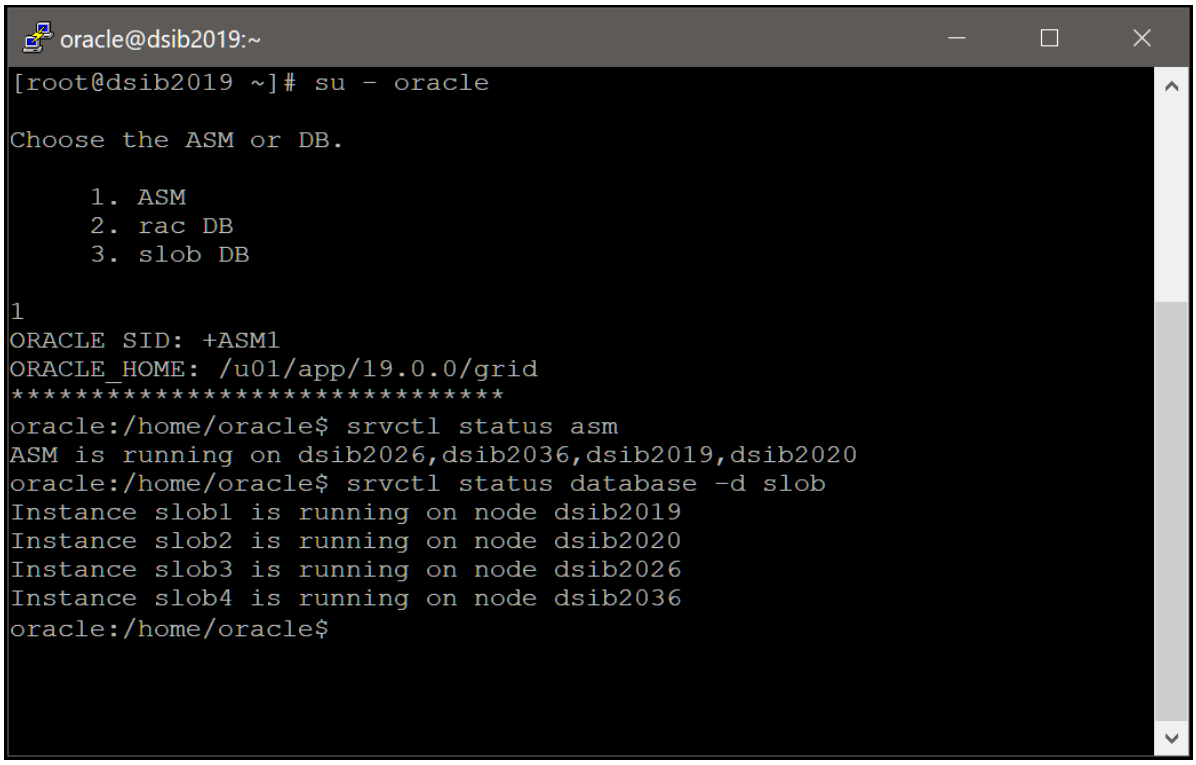

Figure 59. ASM and Oracle RAC database operational on all nodes

5.2.4.2 Testfailover and failover behavior

It's critical to remember that any modification to the recovery plan VMs such as those outlined here, are honored when running testfailover, planned migration or failover. Normally, customers use a single recovery plan for testfailover and failover, but when a recovery plan is modified in the manner discussed here, the best practice is to create a second recovery plan which has no modification. This will ensure that a planned migration or failover results in the same configuration as the production VMs.

6 VMware PowerCLI RAC testing

While VMware Site Recovery Manager (SRM) replicates and recovers at the VM level, there are use cases in which it is desirable to recover only some of the vmdks of a VM, whether in a testfailover or failover scenario. One of those particular scenarios is in order to refresh an Oracle RAC database but not refresh the software (grid and database) to avoid the issues with host names and IPs covered in the previous sections. This can be accomplished using VMware PowerCLI (PowerCLI).

6.1 Managed vs unmanaged snapshots

VMware supports both managed and unmanaged vVol snapshots. The distinction between them is whether VMware is aware of the snapshot or not. A managed snapshot, therefore, is one VMware initiates through its software (e.g. vCenter), while an unmanaged is one is created through array software. On the PowerMax the snapshot would normally be initiated through Solutions Enabler or Unisphere for PowerMax, however, the PowerMax does not support unmanaged snapshots so there is no way to create these vVol copies outside of VMware. Therefore, vVol snapshots initiated by the user on the PowerMax are done at the VM level, not the device level, typically in vCenter with the Take Snapshot functionality.[2]

6.1.1 Remote replication

Remote replication presents a special use case for snapshots in a PowerMax environment. Normally, when a testfailover or failover is executed against a VASA replication group, VMware would instruct the array through the VASA Provider to either create a snapshot off the remote site for failover testing, or in the case of failover, activate the remote devices directly. On the PowerMax, the remote devices (aka R2) are never activated directly whether for testfailover or failover. Instead, the array maintains five snapshots of the devices in a VASA replication group, five minutes apart. This achieves the published RPO of 300 seconds and offers redundancy in the event a snapshot cannot be taken. Using testfailover as the example, when VMware requests that the array present a copy of the devices in a replication group, the Dell EMC PowerMax VASA Provider will use the most recent snapshot of the five, and create linked target devices, presenting them as the requested snapshot. Note that this is true whether testfailover is being executed in VMware SRM, or manually through PowerCLI.

In some ways, this implementation of snapshotting in remote replication meets the definition of unmanaged snapshots since VMware is not initiating, nor is it aware of the five snapshots; however, since the user is unable to initiate or manage them either, they are not unmanaged. The snapshots will only be used as a result of a command issued through the VASA Provider by VMware, whether GUI or CLI. But there is a difference in how the snapshots are handled in SRM (GUI) versus PowerCLI (CLI).

6.1.1.1 VMware SRM and PowerCLI

SRM is an orchestration tool that is fully integrated with vVols and the VASA Provider. It is privy to the vCenter information including all the VMs and whether or not they are replicated. It is a complete solution since it ties VMware to the array through a set of orchestration steps. But as previously noted, SRM works with the entire VM, not just part of it. The only way to work with snapshots for a subset of replicated vVols in a VM is with PowerCLI. By necessity in such a case, therefore, a VM might have multiple storage policies associated with it. This, in and of itself, would invalidate it for use with SRM. It is important to remember, however, that the user still has no control over the snapshots. Using PowerCLI for testfailover or failover will

[2] Utilities such as PowerCLI also offer this capability.

still access one of the five snapshots on the array, not create new ones. The biggest difference between SRM and PowerCLI, therefore, is that while SRM is able to recover the VMs from the snapshot as part of the orchestration, VMware knows nothing about the snapshot when manipulated through PowerCLI. After testfailover or failover, therefore, the user must tell VMware about it. This is accomplished by “importing” the snapshot into the VASA database. Note that this is the same step that is required of an unmanaged snapshot, though technically these are still managed.

6.2 Environment

The following sections will walk through the process of refreshing the Oracle database in a test environment using PowerCLI. Though this is not an SRM configuration, the source site will be referred to as the production or protection site, while the target site will be called disaster recovery or simply recovery site.

In this example, the production database consists of two Oracle RAC nodes using ASM for storage in a similar configuration as the other database in this paper. There are only two ASM disk groups in this setup, DATA and VVOL_REFRESH. The vvol database is solely located in the VVOL_REFRESH disk group seen in Figure 60. It is comprised of three vmdks/vVols as shown in Figure 61. The DATA disk group is used for other objects and purposely will remain mounted through this process to demonstrate the online nature of the refresh.

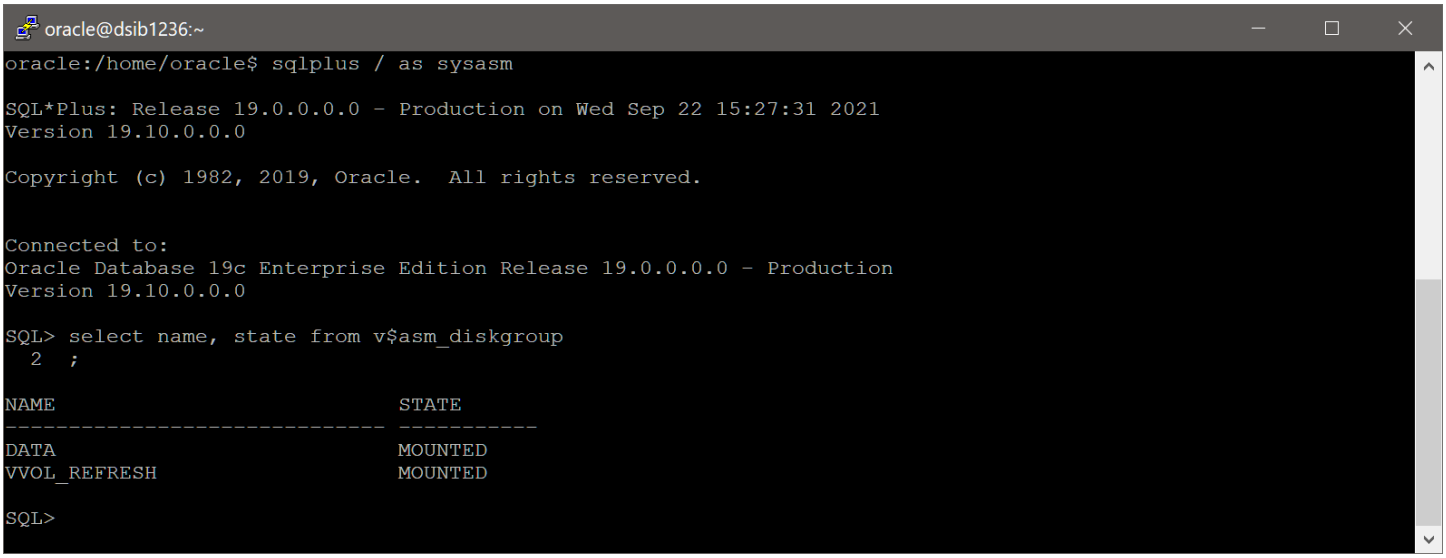

Figure 60. ASM disk groups

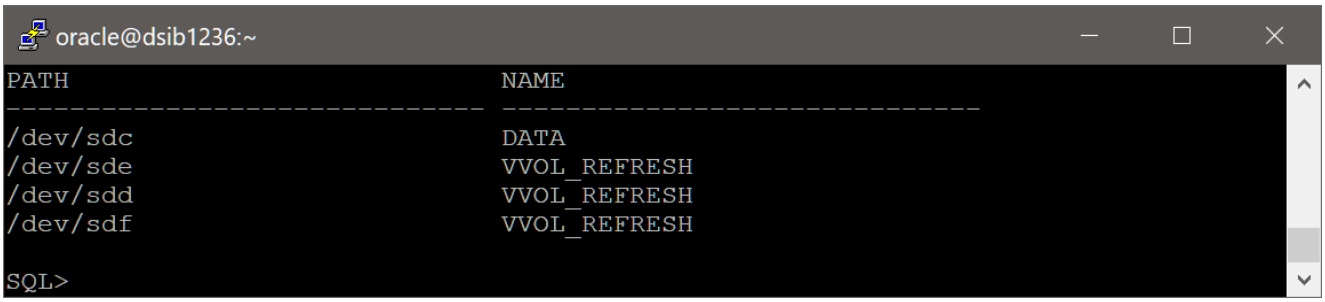

Figure 61. VVOL_REFRESH ASM disks

The production vvol database is comprised of two RAC instances running on nodes dsib1249 and dsib1236 as shown in Figure 62.

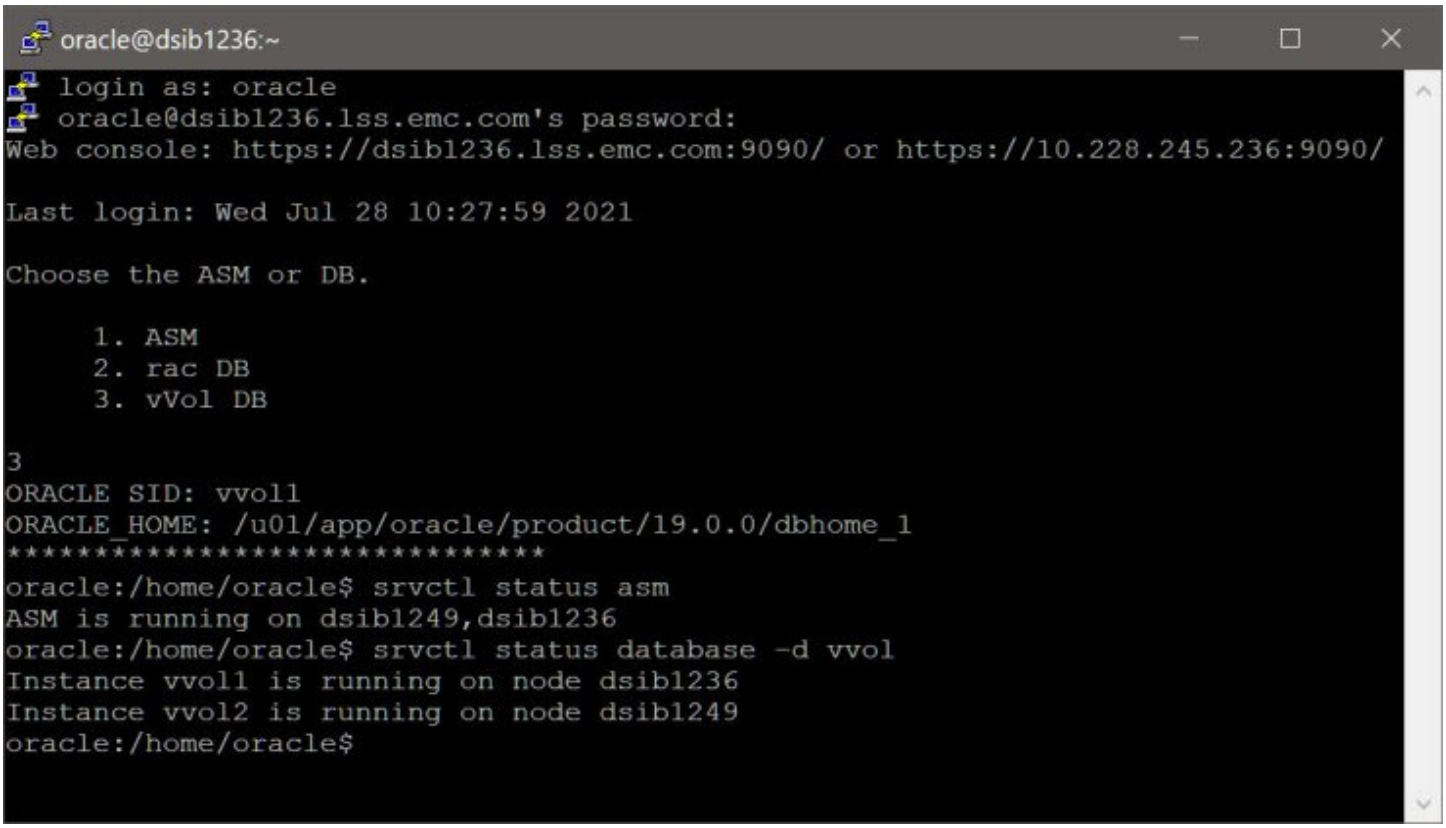

Figure 62. ASM and RAC instances on protection site

6.2.1 Storage policies

For each RAC node there are two different policies because only part of the VM requires replication. The first storage policy is named ORA-vVol. It is associated with a service level of Diamond and has no replication associated with it as shown in Figure 63.

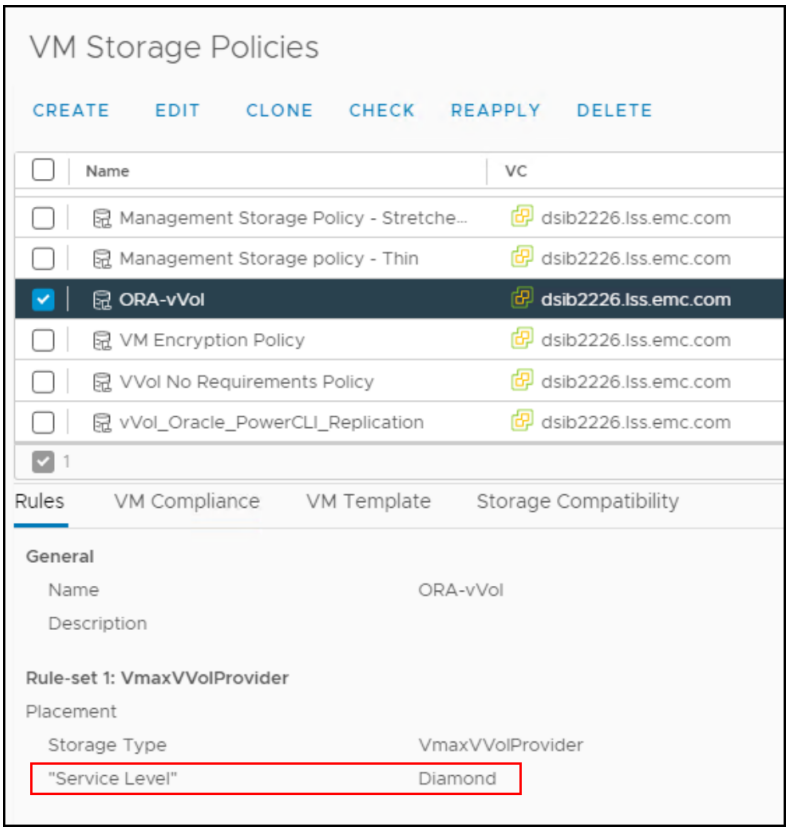

Figure 63. ORA-vVol storage policy

The second storage policy is named vVol_Oracle_PowerCLI_Replication. It also has a service level of Diamond but supports replication as shown in Figure 64.

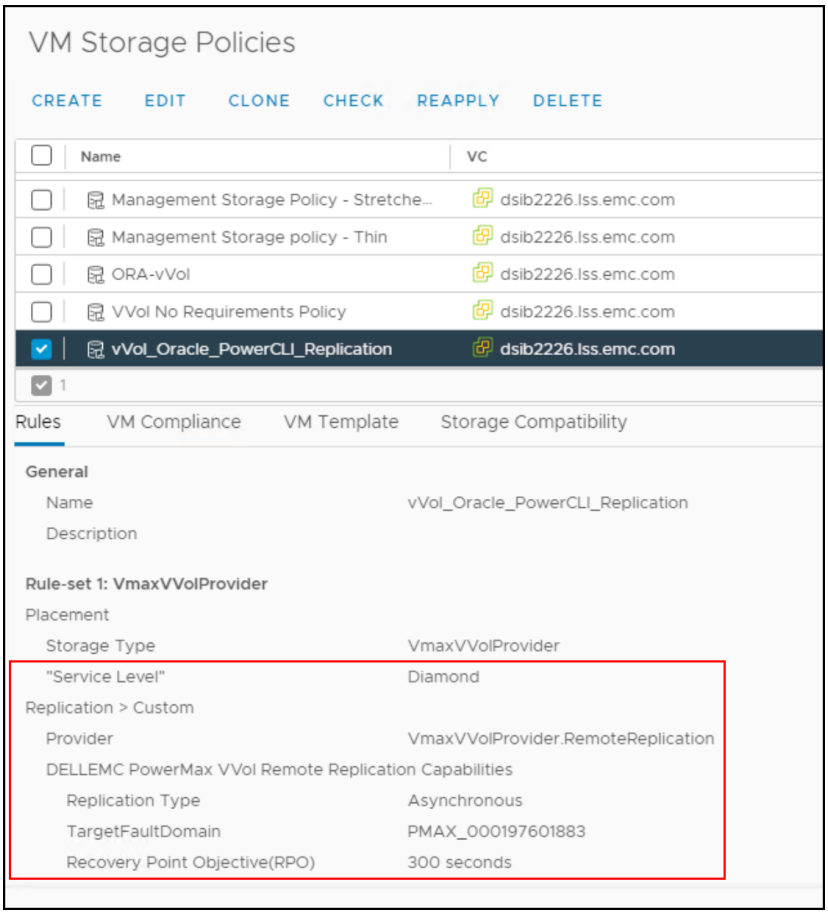

Figure 64. vVol_Oracle_PowerCLI_Replication storage policy

The three vmdks for the VVOL_REFRESH disk group are assigned the vVol_Oracle_PowerCLI_Replication storage policy. The vmdks that hold the software and the DATA disk group are assigned the ORA-vVol storage policy as they are not needed at the recovery site. Note in Figure 65 below where the DATA disk (Hard disk 3) and the VVOL_REFRESH disk (Hard disk 4) are assigned the different policies. As is required, the VVOL_REFRESH disk is also associated with a VASA replication group, ORAPCLI. Such a VM configuration is not permitted by SRM.

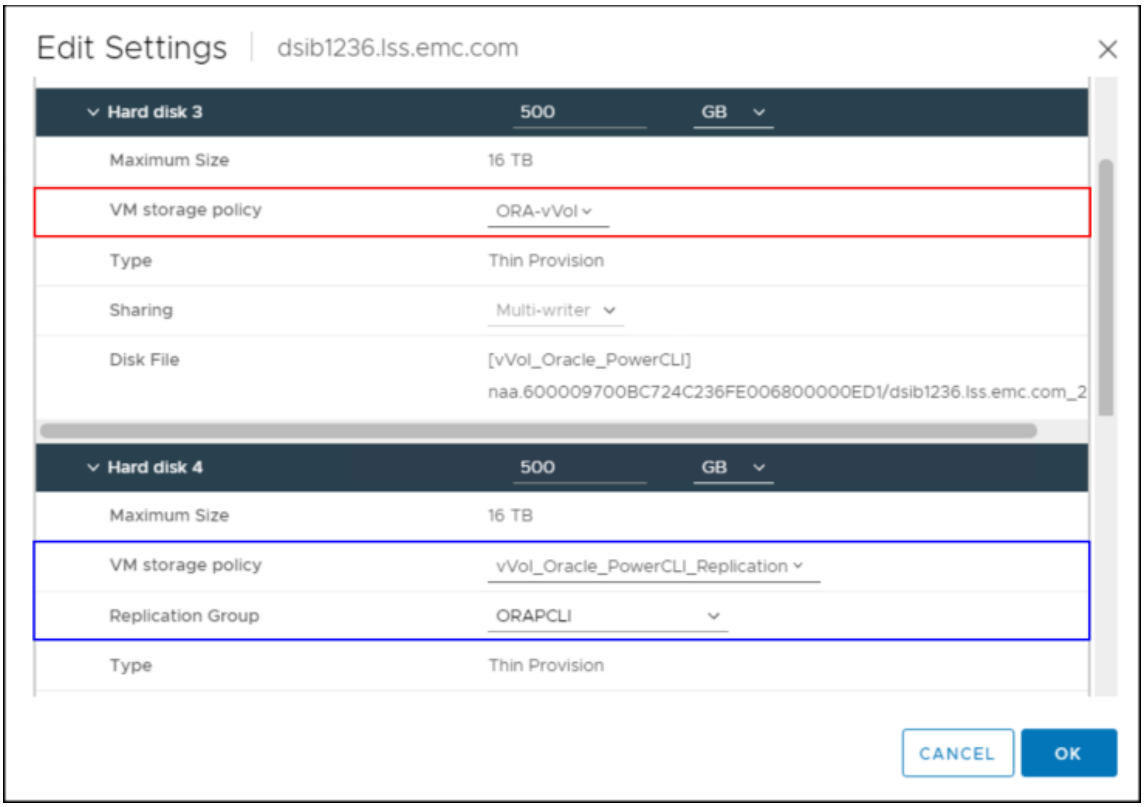

Figure 65. Oracle VM storage policies

The three disks of the VVOL_REFRESH disk group are assigned the ORAPCLI group label which is VASA replication group 37 on the array as seen in Figure 66. Since the replication is SRDF/A, all disks must be associated with the same group to guarantee consistency. But because ASM disk groups are independent, and the database is only located in the VVOL_REFRESH disk group, the DATA disk group does not have to be replicated.

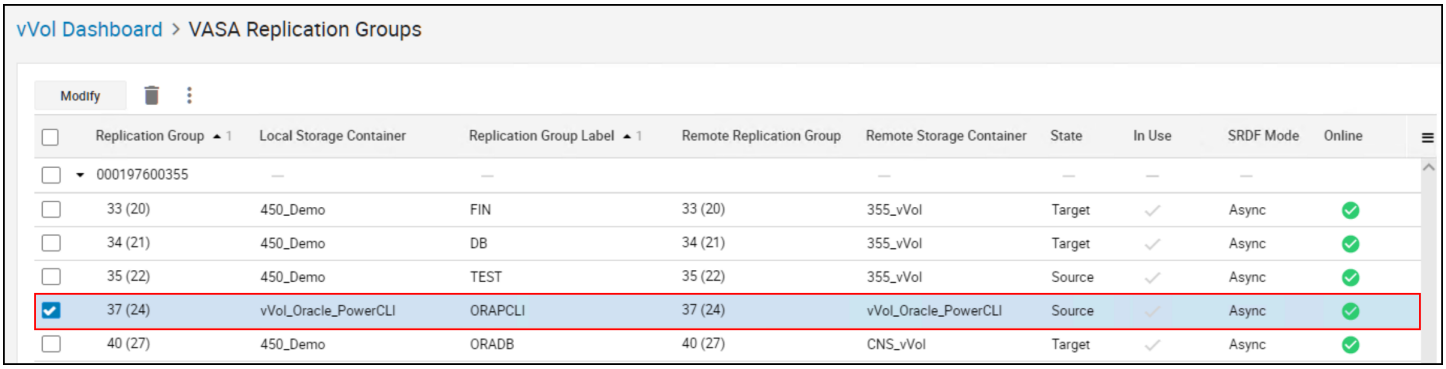

Figure 66. Unisphere VASA Replication Group

6.2.2 Recovery site

As SRM is not being used, there are no placeholder VMs configured at the recovery site. Rather, the point of using PowerCLI is so the recovery site VM can be operational, fully online. In this example, there is a single VM at the recovery site, dsib1246.lss.emc.com, with Oracle RAC configured using ASM. The VM's

configuration is similar to the production VMs. There are two disks for the software and one disk for a DATA disk group (which contains the voting and cluster information). All disks are using the same storage policy, VVol No Requirements Policy, as performance is not of primary concern on this test VM and this policy instructs the VASA Provider to use the least performance service level (e.g., Bronze) in the container. The software and DATA disk are shown in Figure 67. Note the multi-writer flag for the ASM disk would permit more test nodes if desired.

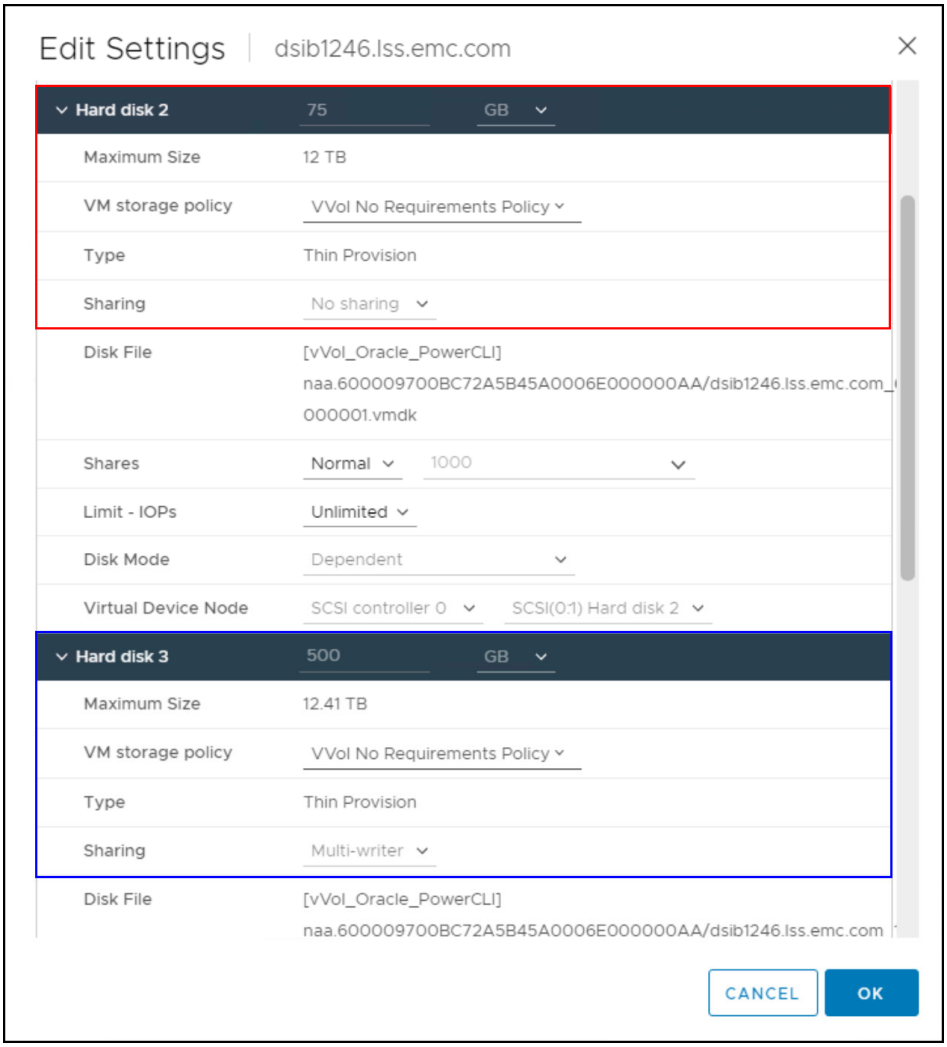

Figure 67. Recovery site VM dsib1246.lss.emc.com

The test VM currently has a single RAC node as shown in Figure 68, with ASM running, but the database is not mounted or open as it is located in the VVOL_REFRESH disk group which is not yet present. There is, however, a parameter file for the vvol instance which references the missing disk group.

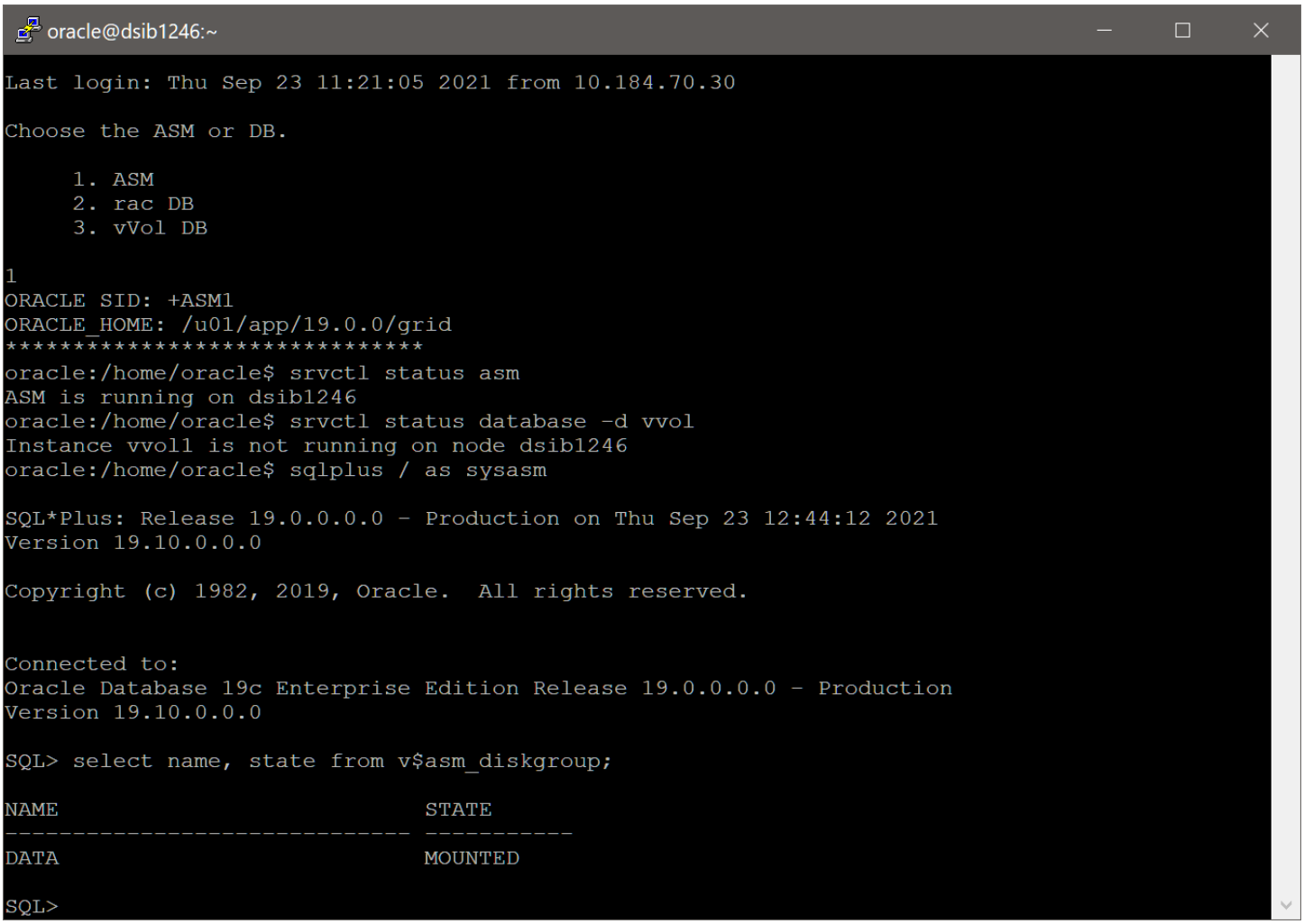

Figure 68. ASM and RAC instance on recovery site

With the environment prepared, the testfailover can be executed.

6.2.3 Testfailover with PowerCLI

The procedure covered here assumes the user is familiar with both PowerShell and VMware PowerCLI. VMware PowerCLI is a set of modules that is executed from within PowerShell. The modules can be installed within PowerCLI or downloaded as a zip and manually extracted to the proper directory. The procedure in this chapter uses PowerShell version 5.1.x and PowerCLI 12.3.x.

```
PS C:\WINDOWS\system32> $PSVersionTable
Name Value
---- -----
PSVersion 5.1.18362.1593
PSEdition Desktop
PSCompatibleVersions {1.0, 2.0, 3.0, 4.0...}
BuildVersion 10.0.18362.1593
CLRVersion 4.0.30319.42000
WSManStackVersion 3.0
PSRemotingProtocolVersion 2.3
```

```
SerializationVersion 1.1.0.1


PS C:\WINDOWS\system32> Find-Module -Name VMware.PowerCLI
Version Name Repository Description
------- ---- ---------- -----------
12.3.0.... VMware.PowerCLI PSGallery This Windows PowerShell module contains
VMware.PowerCLI
```

Begin by connecting to the target vCenter where the test VM is running as shown in Figure 69. All commands will be run against this target vCenter at the recovery site.

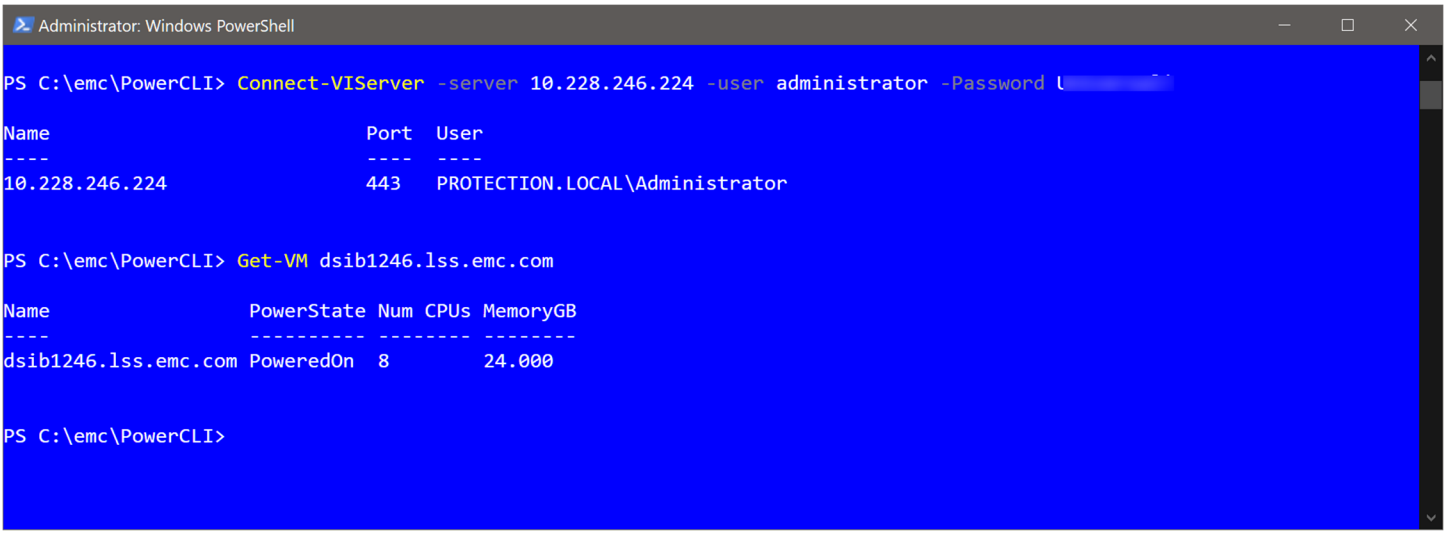

Figure 69. PowerCLI connection to target vCenter with test VM

Prior to running the testfailover, information about the snapshot target vVols is required. The next section details how to obtain that data.

6.2.3.1 WWN identification

One of the challenges of using PowerCLI instead of SRM is the identification of the snapshot target devices. When using SRM, the VASA Provider is able to communicate this information to VMware and the VMs are automatically assigned the devices. With PowerCLI, the user must obtain this information as the WWNs are needed as part of the testfailover process.

The identification processes explained here are purposely basic so that a user need not be proficient in scripting to obtain the information. More advanced scripting methods are possible and the user is encouraged to explore those options; however, be mindful that when working with vVols there are command limitations in the Solutions Enabler CLI.

The most basic way to get the information, assuming the vVol environment is static during the test (no new vVol creation), is to run the following command as shown in Figure 70 before and after the testfailover.

```
symdev list -wwn -vvol -sid xxx
```

Comparing the two results will reveal the new vVol devices.

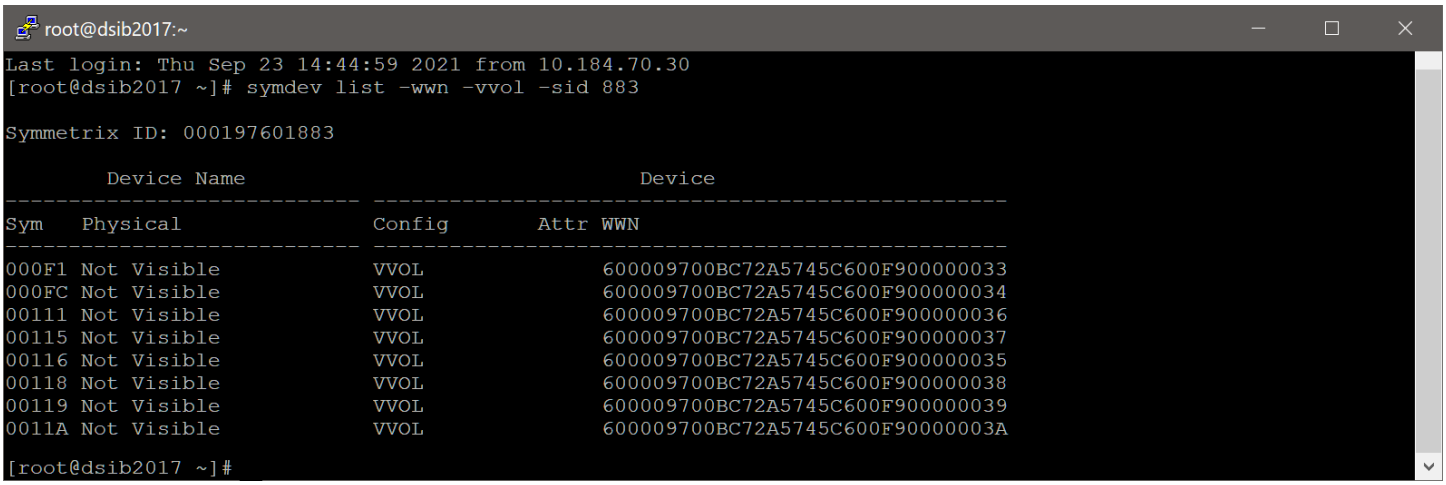

Figure 70. Listing WWN of vVol devices

A similar method is to determine which vVol devices are snapshot targets before and after testfailover. If the environment is actively creating vVols, this way is preferable as the list is likely smaller. The previous command, `symdev`, can also be used to find this information using the verbose output. A row called *Snapvx Target* will provide a Boolean output, so the user is looking for **True** as shown in Figure 71. The command is:

```
symdev list -vvol -sid xxx -v|grep "Snapvx Target"|grep "True"
```

Figure 71. Snapvx Target

Each line returned in the previous command represents a vVol snapshot target. If any results are returned, therefore, repeat the command without the filter and redirect the results to a file:

```
symdev list -vvol -sid xxx -v > verbose_output.txt
```

The output can be searched to obtain the device WWN (and device ID) as shown in Figure 72.

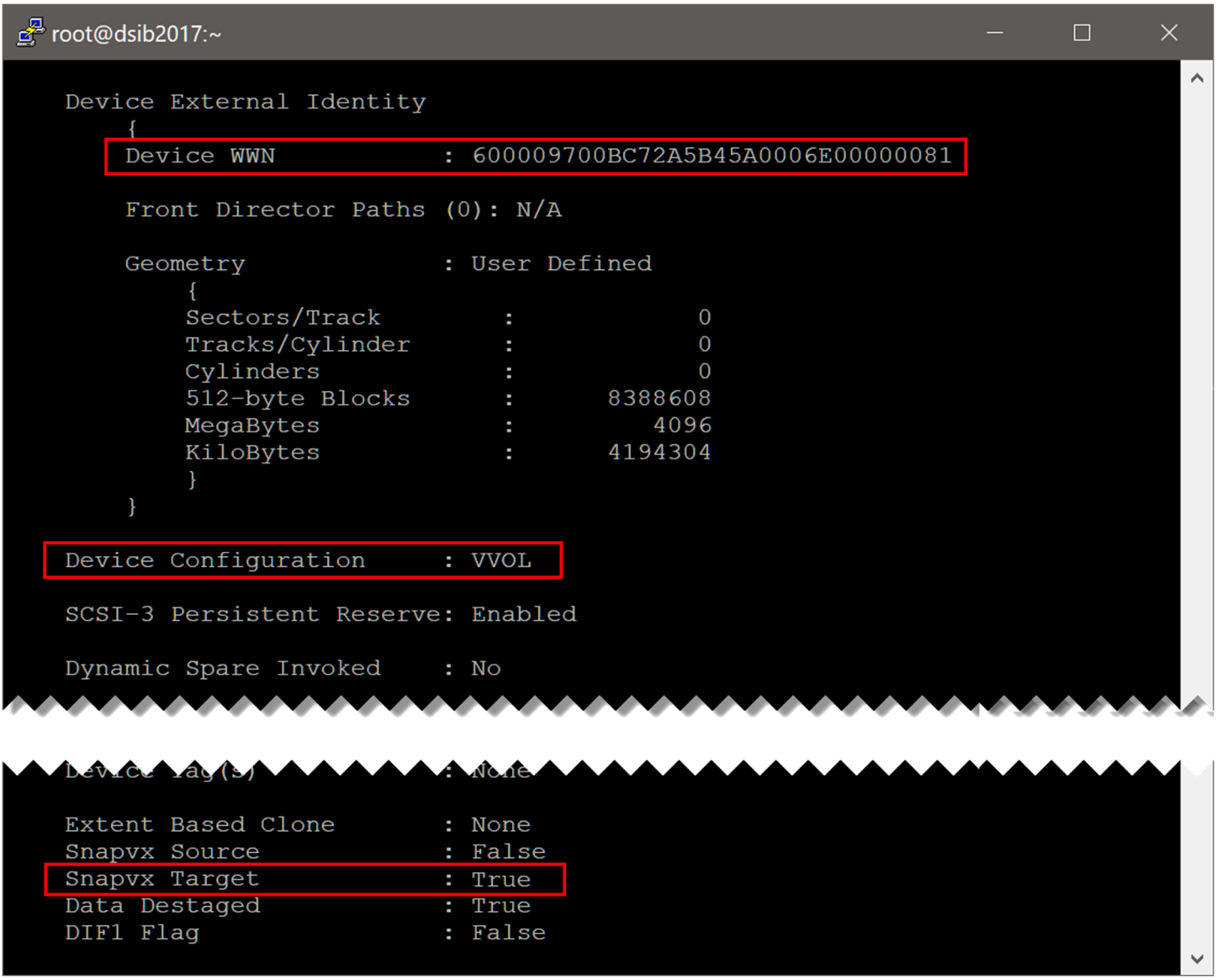

Figure 72. SnapVX target device

This procedure is run prior, and subsequent to testfailover. Record the WWNs that did not appear prior to testfailover.

6.2.3.2 Testfailover command

Two commands are required for running testfailover with PowerCLI, one to set the replication group variable, the second to run the testfailover. Figure 73 shows that the variable *$RG* is first set to ORAPCLI. Then, in the second command, the testfailover cmdlet *Start-SpbmReplicationTestFailover* is called. The testfailover takes less than a minute to complete. Note that the option **-Server** is only required in the commands if more than one vCenter has a connection in PowerShell.

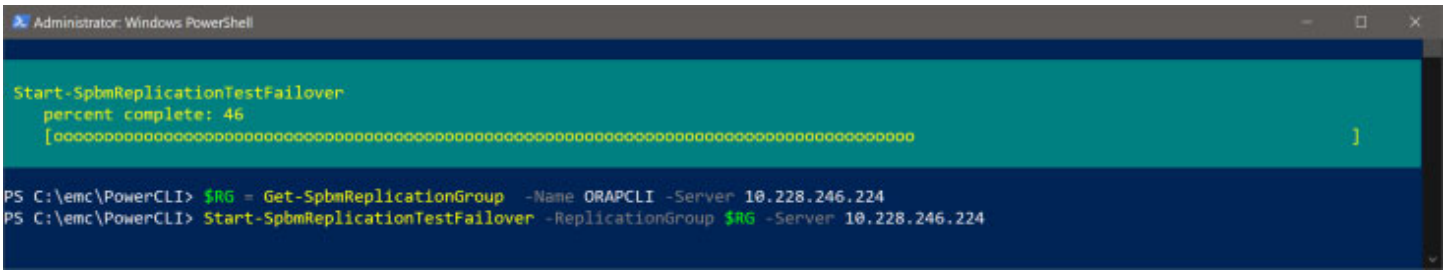

Figure 73. PowerCLI testfailover

When viewing the VASA replication group in Unisphere on the target array, the group will change from **Target** to **Intest** as seen in Figure 74, just as it does with testfailover in SRM.

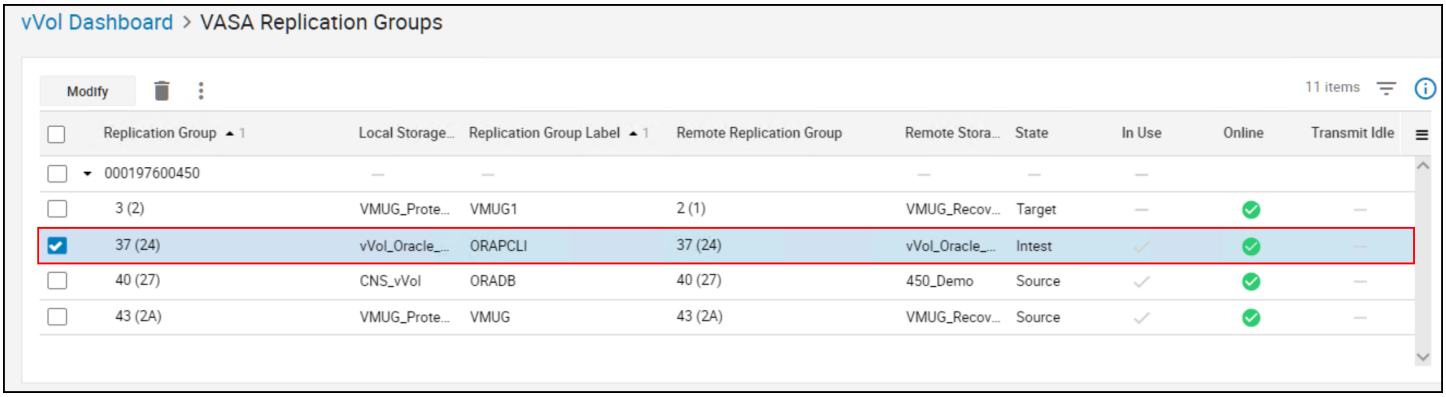

Figure 74. Intest VASA Replication Group: ASM VVOL_REFRESH disk group

When the command completes, in this example there should be three snapshot target devices created for the VVOL_REFRESH ASM disk group. Using the process outlined in the section WWN identification, the three WWNs for the linked targets are determined to be:

- 600009700BC7246338010036000011C1
- 600009700BC7246338010036000011C2
- 600009700BC7246338010036000011C3

As VMware knows nothing of these new snapshot target devices, they must now be imported into the VASA database. Be aware that until they are imported, the devices cannot be seen in the vVol datastore on the target vCenter.

6.2.3.3 Import the snapshots (linked targets)

The method used in PowerCLI to import the snapshots into VASA/VMware is called **ImportUnmanagedSnapshot** which is a part of **virtualDiskManager**. The method takes three parameters:

- The vmdk path (any file location in the datastore set by the user)
- The target vCenter datacenter name
- The WWN of the vVol. Note as with other disks presented to VMware, the prefix **naa.** is required.

Begin by setting the virtualDiskManager to the view of the target vCenter by running the following command:

```
$virtualDiskManager = Get-View (Get-View
ServiceInstance).Content.virtualDiskManager -Server 10.228.246.224
```

Next, assign the datacenter to a variable, *$dc*, so that it can be used in the import snapshot call. Note only the managed object reference (MoRef) will be necessary in the import call.

```
$dc = get-datacenter "Boston"
```

Each of the three vVols now must be assigned a variable as there are three import calls. Again, note the naa prefix.

```
$uuid1 = "naa.600009700BC7246338010036000011C1"
$uuid2 = "naa.600009700BC7246338010036000011C2"
$uuid3 = "naa.600009700BC7246338010036000011C3"
```

Finally, assign a vmdk path in the target vVol datastore to three more variables. These vmdks can be placed anywhere in the vVol datastore, so long as the path exists. It may be necessary to pre-create folders for this purpose. In this example the vmdks will go into the folder with the test VM, though the file names themselves reflect the production VM from whence they came. The naming convention for the path is: datastore in brackets, followed by folder(s) name, then vmdk name.

```
$vvolpath1 = "[vVol_Oracle_PowerCLI]
dsib1246.lss.emc.com/dsib1236.lss.emc.com_1.vmdk"
$vvolpath2 = "[vVol_Oracle_PowerCLI]
dsib1246.lss.emc.com/dsib1236.lss.emc.com_2.vmdk"
$vvolpath3 = "[vVol_Oracle_PowerCLI]
dsib1246.lss.emc.com/dsib1236.lss.emc.com_3.vmdk"
```

With the last of the variables set, run the import command as shown in Figure 75. When the user executes the command, VMware will import the vVol into the VASA DB, bind it to a Protocol Endpoint, and create the pointer file in the datastore so that it can be added it to the VM (as an existing disk). The commands complete fairly quickly and do not return a response unless there is a failure.

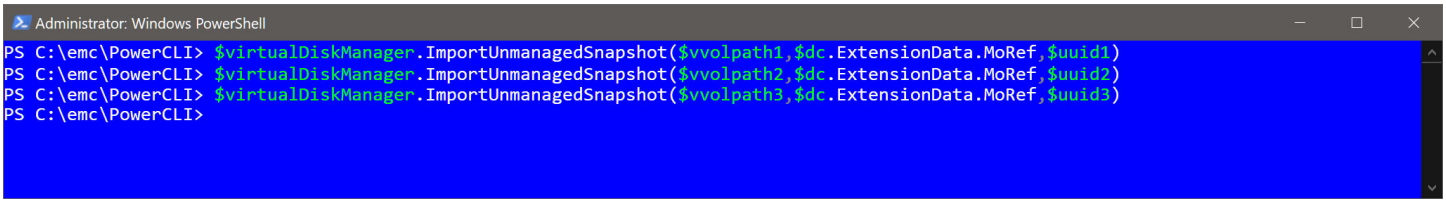

Figure 75. Import snapshots

When the commands complete, check the vVol datastore for the imported vVols at the path specified. The imported snapshots are shown in Figure 76.

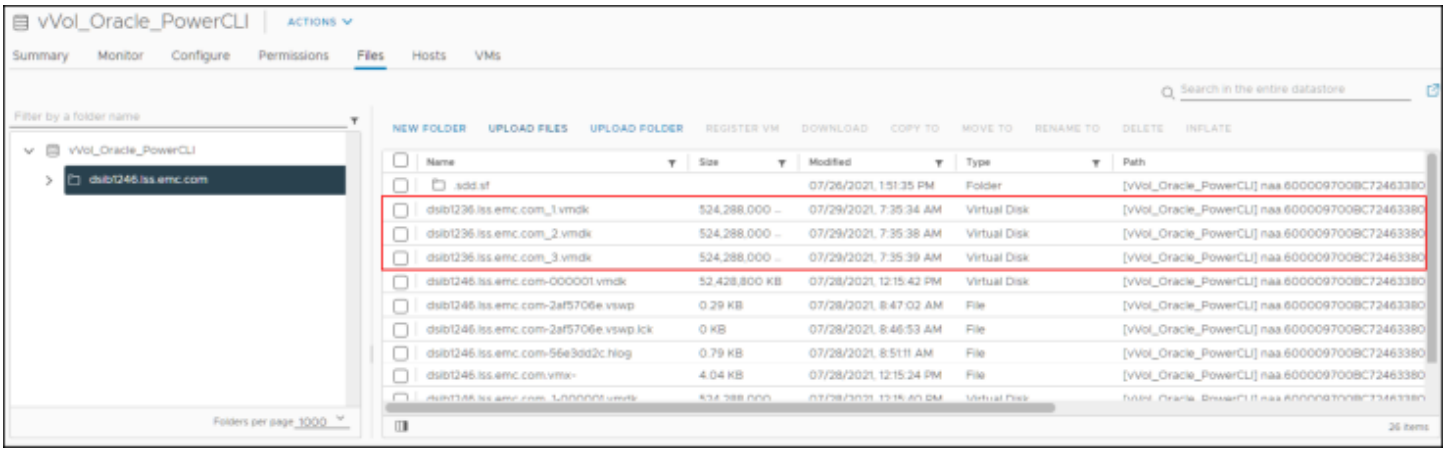

Figure 76. vVol datastore vmdks

6.2.3.4 Add to VM

With the vVols available they are like any other vmdk, add them to the test VM using the add "Existing Hard Disk" functionality. Normally this action should be available through the vCenter GUI or CLI, but the vCenter GUI will produce a VMware storage policy error that is seen in Figure 77.

Figure 77. vCenter error

The reason the vCenter does not work is because of the SPBM plugin. Basically, VMware is looking for a replication group for this vVol as it is a mandatory field in the spec but there is none, so it produces the error. VMware has chosen not to address this bug, as it is still possible to use the vSphere Client GUI or CLI to add the disks.

To import the disks through PowerCLI, issue the following commands:

```
$vm = get-vm dsib1246.lss.emc.com
$vm |new-harddisk -Diskpath $vvolpath1
$vm |new-harddisk -Diskpath $vvolpath2
$vm |new-harddisk -Diskpath $vvolpath3
```

The devices will come in with a **Datastore Default** policy and **Sharing** set to *No sharing* as shown in Figure 78. In this example the lack of multi-writer is not a problem since the RAC test environment is a single node; however, some test environments are likely to have multiple nodes and will need sharing. While it is possible to change the multi-writer flag with PowerCLI while the VM is powered on, VMware does not provide for

passing the sharing flag when adding the disk. This has been an issue for many years. There are scripts[3] available to modify the flag but they are beyond the scope of this example.

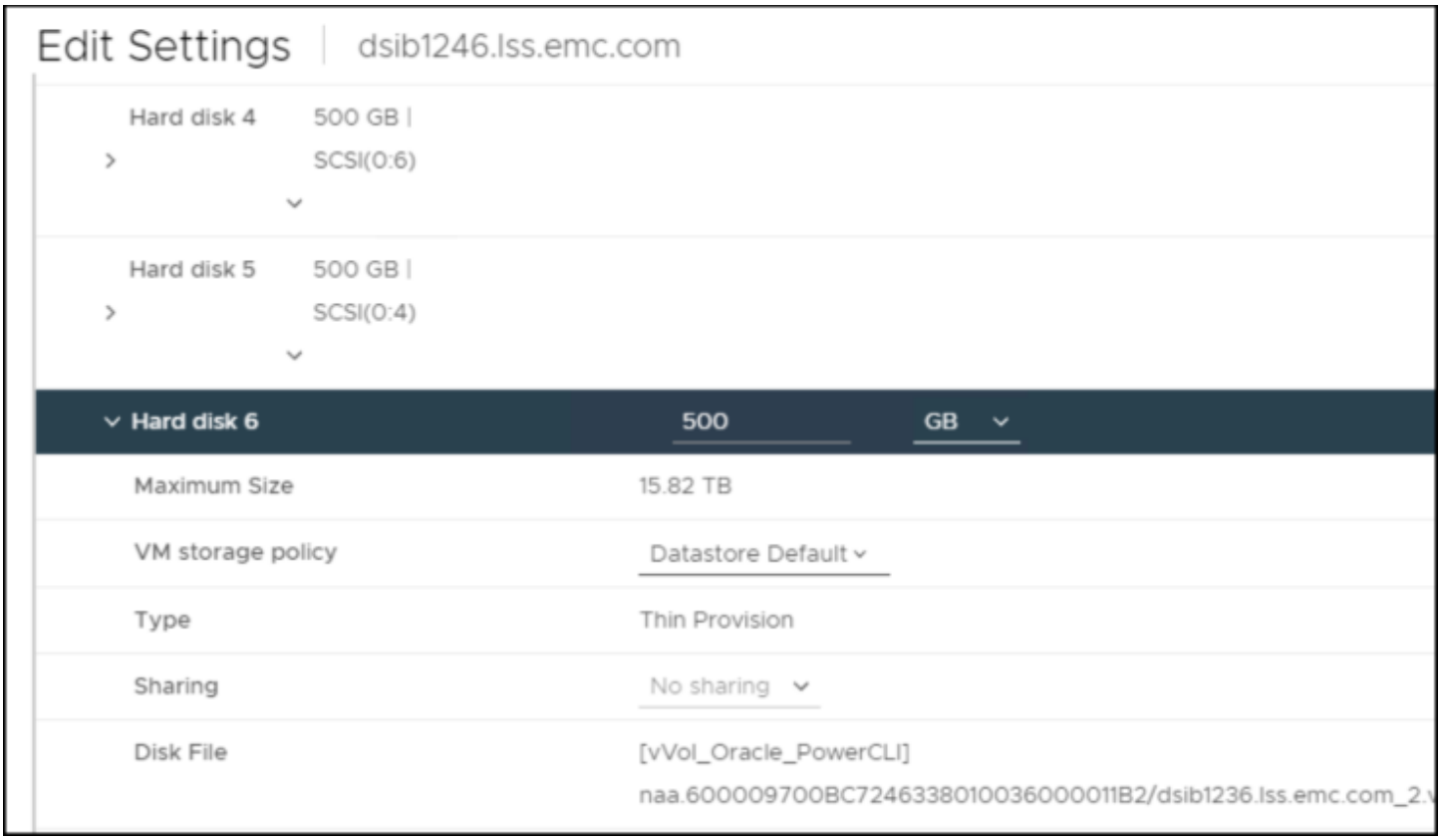

Figure 78. Adding existing snapshot vVols

6.2.4 Mounting VVOL_REFRESH ASM disk group

Once the vmdks are added to the test VM, ASM will immediately recognize the devices as having an ASM header and show the VVOL_REFRESH ASM disk group as **DISMOUNTED**. Recall that the three disks comprising the group are part of the same SRDF/A group and thus are consistent with each other. All that remains is to mount the disk group and then start the vvol database as shown in Figure 79.

[3] For information about adding the device online with multi-writer set, please see https://williamlam.com/2015/10/new-method-of-enabling-multiwriter-vmdk-flag-in-vsphere-6-0-update-1.html.

```
oracle@dsib1246:~

Choose the ASM or DB.

     1. ASM
     2. rac DB
     3. vVol DB

1
ORACLE SID: +ASM1
ORACLE_HOME: /u01/app/19.0.0/grid
******************************
oracle:/home/oracle$ sqlplus / as sysasm

SQL*Plus: Release 19.0.0.0.0 - Production on Fri Sep 24 10:34:58 2021
Version 19.10.0.0.0

Copyright (c) 1982, 2019, Oracle.  All rights reserved.


Connected to:
Oracle Database 19c Enterprise Edition Release 19.0.0.0.0 - Production
Version 19.10.0.0.0

SQL> select name, state from v$asm_diskgroup;

NAME                           STATE
------------------------------ -----------
VVOL_REFRESH                   DISMOUNTED
DATA                           MOUNTED

SQL> alter diskgroup VVOL_REFRESH mount;

Diskgroup altered.

SQL> exit
Disconnected from Oracle Database 19c Enterprise Edition Release 19.0.0.0.0 - Production
Version 19.10.0.0.0
oracle:/home/oracle$ . ./.db_vVol_profile
oracle:/home/oracle$ sqlplus / as sysdba

SQL*Plus: Release 19.0.0.0.0 - Production on Fri Sep 24 10:38:41 2021
Version 19.10.0.0.0

Copyright (c) 1982, 2019, Oracle.  All rights reserved.

Connected to an idle instance.
SQL> startup pfile=/u01/app/oracle/product/19.0.0/dbhome_1/dbs/initvvol1.ora
ORACLE instance started.

Total System Global Area 7532967056 bytes
Fixed Size                  8914064 bytes
Variable Size            1308622848 bytes
Database Buffers         6207569920 bytes
Redo Buffers                7860224 bytes
Database mounted.
Database opened.
SQL> show parameter CONTROL_FILES

NAME                                 TYPE        VALUE
------------------------------------ ----------- ------------------------------
control_files                        string      +VVOL_REFRESH/VVOL/CONTROLFILE
                                                 /current.280.1078917417, +VVOL
                                                 _REFRESH/VVOL/CONTROLFILE/curr
                                                 ent.279.1078917417

SQL>
```

Figure 79. Mounting VVOL_REFRESH ASM disk group and starting vvol database

It is now possible to conduct testing, all of this accomplished without altering any IP addresses or having to reconfigure the ASM Clusterware.

6.2.5 Test cleanup

When testing is complete, or a refresh is required, the process should be reversed before repeating. The following steps need to be completed. Note that failure to complete these initial steps before stopping testfailover will result in an error as VMware will recognize that the devices are still in use.

- Shutdown the vvol database and dismount the VVOL_REFRESH ASM disk group as shown in Figure 80.

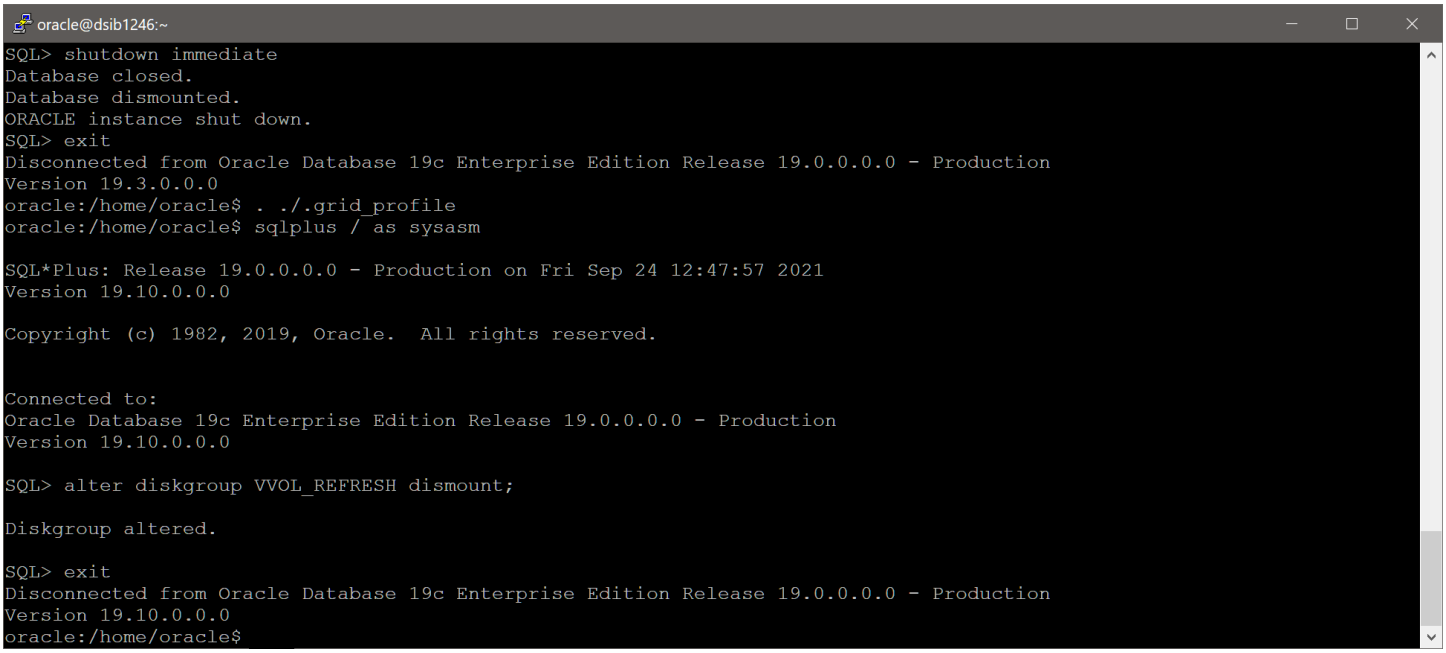

Figure 80. Shutdown RAC database and dismount VVOL_REFRESH disk group

- Remove the devices from the VM (both GUI and CLI work for this), include the checkbox to remove the vmdk pointer file if using GUI. A CLI example is shown in Figure 81.

Figure 81. Remove vVols from test VM

- Stop the testfailover as shown in Figure 82.

```
PS C:\emc\PowerCLI> Stop-SpbmReplicationTestFailover -ReplicationGroup $RG -Server 10.228.246.224

Name                           ReplicationState
----                           ----------------
ORAPCLI                        Target
```

Figure 82. End testfailover for RAC test environment

The environment is now reset and a new refresh could be executed.

7 Conclusion

This paper served to demonstrate how to configure an Oracle RAC environment on VMware Virtual Volumes and enable VMware SRM as a disaster recovery solution. VMware vVols enables Oracle RAC customers to move away from raw device mappings (RDM) yet retain the one to one mapping of vmdk to array volume that they desire. In addition, the paper included detail on creating an Oracle RAC test environment when only a subset of vmdks are replicated using PowerCLI, avoiding Oracle software reconfiguration at the test site.

A Error/Warning messages

The following sections detail some expected errors or warnings that the user may encounter.

A.1.1 Testfailover

When an SRM testfailover is executed, the test VMs no longer satisfy the storage policy requirements, e.g., replication. Therefore, the VM will show as noncompliant as shown in Figure 83.

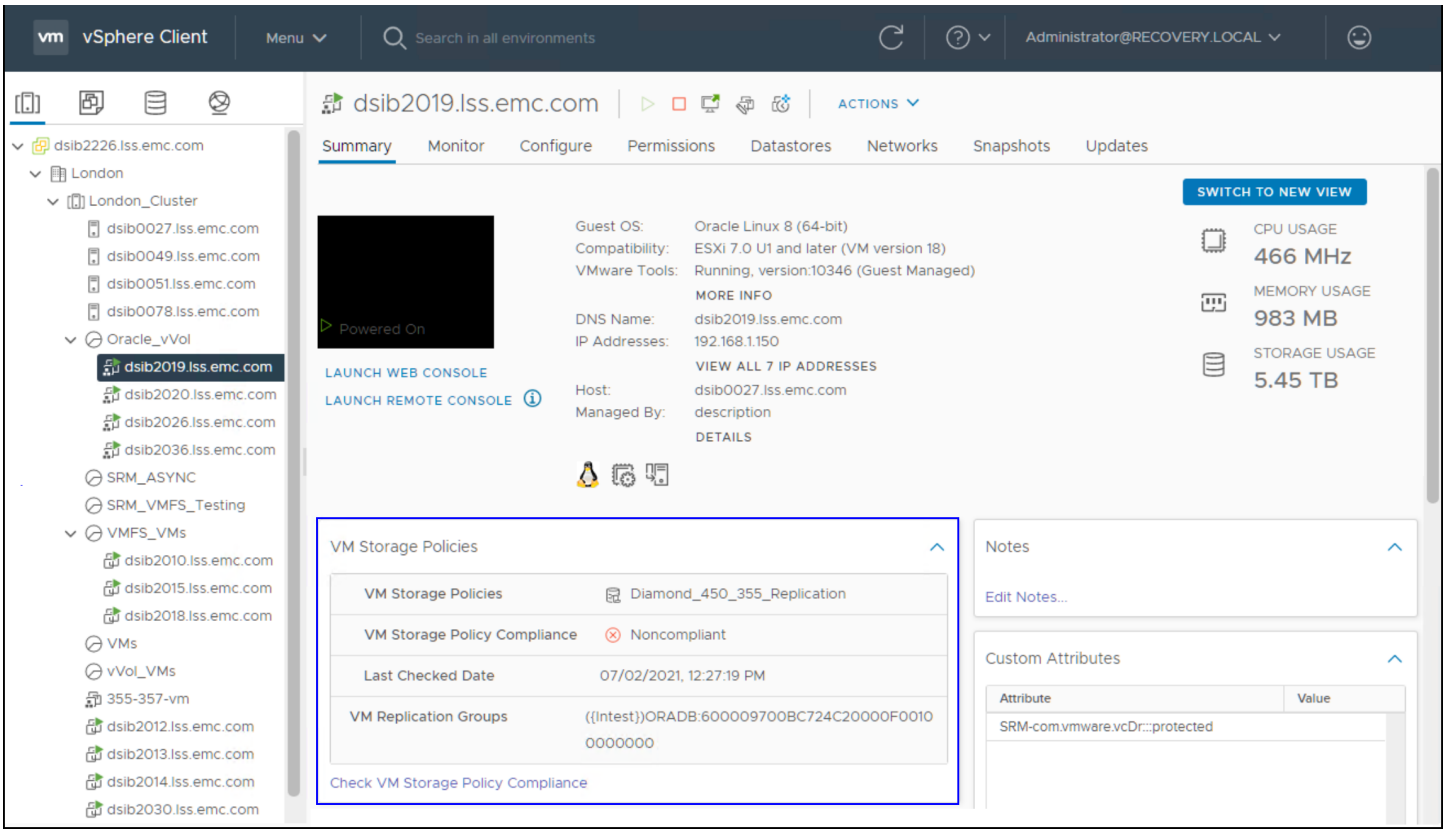

Figure 83. Noncompliant VM after testfailover in SRM

B Technical support and resources

Dell.com/support is focused on meeting customer needs with proven services and support.

B.1 Related resources

B.1.1 Dell EMC

- Using Dell EMC VMAX and PowerMax Storage in VMware vSphere Environments TechBook

  https://www.dellemc.com/resources/en-us/asset/technical-guides-support-information/products/storage-2/h2529-vmware-esx-svr-w-symmetrix-wp-ldv.pdf

- Using VMware vSphere Virtual Volumes 2.0 and VASA 3.0 with Dell EMC VMAX and PowerMax

  https://www.dellemc.com/resources/en-us/asset/white-papers/products/storage/h18344-vmware-vvol2-vasa3-powermax-wp.pdf

- Deployment Best Practices for Oracle Databases with Dell EMC PowerMax

  https://www.delltechnologies.com/en-us/collaterals/unauth/white-papers/products/storage/h17390-deployment-bp-for-oracle-with-powermax-wp.pdf

- Oracle Database Backup, Recovery, and Replications Best Practices with Dell EMC VMAX All Flash Storage

  https://www.delltechnologies.com/en-us/collaterals/unauth/white-papers/products/storage/h14232-oracle-database-backup-recovery-vmax3.pdf

B.1.2 VMware

- VMware vSphere

  https://docs.vmware.com/

B.1.3 Oracle

- Oracle database

  https://docs.oracle.com/en/database/oracle/oracle-database/index.html