Implementing Dell PowerFlex SRA with VMware Site Recovery Manager

August 2023

H19490

White Paper

Abstract

This white paper provides guidance for implementing the Dell PowerFlex Storage Replication Adapter (SRA) with VMware Site Recovery Manager (SRM) using asynchronous array replication.

Dell PowerFlex Engineering

PowerFlex Engineering Validated

DELL Technologies

Copyright

Contents

Introduction

Solution overview

VMware vCenter Site Recovery Manager (SRM) leverages storage array-based replication such as PowerFlex replication to protect virtual machines in VMware vSphere environments. The interaction between VMware SRM and storage array replication is governed through a well-defined set of specifications. These VMware-defined specifications are implemented by the storage array vendor as a lightweight application referred to as the storage replication adapter (SRA).

The PowerFlex SRA enables VMware SRM to interact with the PowerFlex storage environment. It allows VMware SRM to automate storage-based disaster restart operations on PowerFlex. The PowerFlex SRA supports a single mode of replication, asynchronous, on both PowerFlex 3.x and 4.x. This document provides implementation procedures and best practices for automated disaster recovery of VMware environments using PowerFlex, asynchronous replication, the PowerFlex SRA, and SRM.

Note: Examples provided in this guide cover methods for performing various VMware vSphere activities using PowerFlex systems and Dell 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.

Audience

This document is intended for use by storage administrators, system administrators and VMware vSphere administrators.

Readers of this document are expected to be familiar with the following topics:

- Dell PowerFlex system operation.
- Dell software including the following: PowerFlex Presentation Server (3.x), PowerFlex Gateway (3.x), and PowerFlex Manager (4.x).
- VMware vSphere products and their operation, particularly VMware Site Recovery Manager.

Terminology

The following table provides definitions for some of the terms that are used in this document.

Table 1. Terminology

Term	Definition
ITOM	IT Operational Management
LCM	Life Cycle Management
MDM	Metadata Manager
RCG	Replication Consistency Group
RPO	Recovery Point Objective

Term	Definition
RTO	Recovery Time Objective
SDC	Storage Data Client
SDR	Storage Data Replicator
SRA	Storage Replication Adapter
SRM	Site Recovery Manager (VMware)
VM	Virtual Machine

We value your feedback

Dell Technologies and the authors of this document welcome your feedback on the solution and the solution documentation. Contact the Dell Technologies Solutions team by email.

Authors: Drew Tonnesen, PowerFlex Engineering

Contributors: Storage technical marketing team

PowerFlex family overview

PowerFlex software defined infrastructure enables broad consolidation across the datacenter, encompassing almost any type of workload and architecture. The software defined architecture offers automation and programmability of the compete infrastructure and provides scalability, performance, and resiliency to enable effortless adherence to stringent workload SLAs. The PowerFlex family provides a foundation that combines compute as well as high performance storage resources in a managed unified fabric. PowerFlex comes in flexible deployment options (rack, appliance, or custom nodes and in the public cloud) that enable independent (two-layer), HCI (single-layer), or mixed architectures. PowerFlex is ideal for high performance applications and databases, building an agile private/hybrid cloud, or consolidating resources in heterogeneous environments.

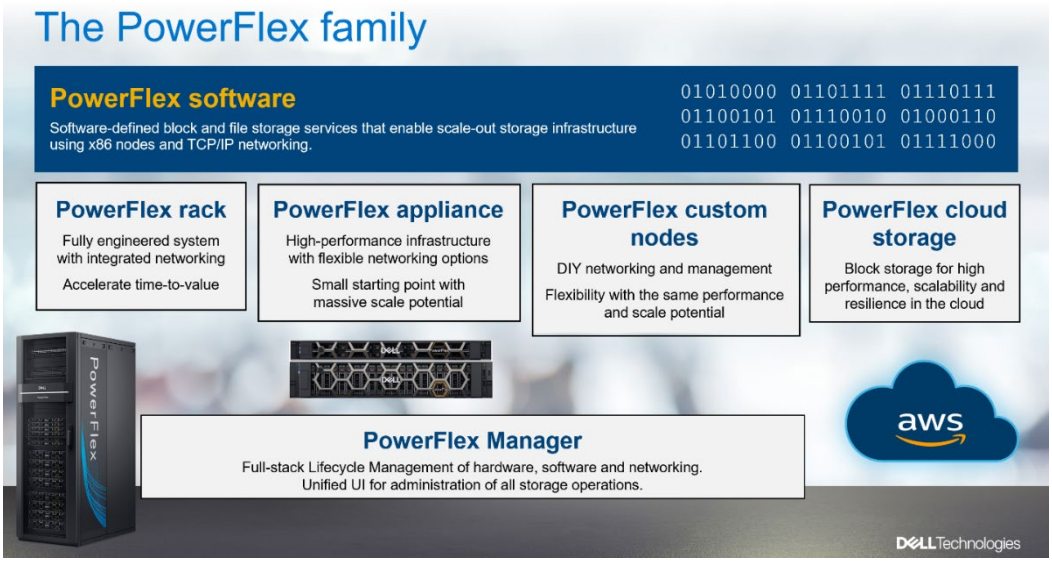

Figure 1. PowerFlex family

PowerFlex Software Components

Software is the key differentiation and the “secret sauce” in the PowerFlex offering. PowerFlex software components not only provide software defined storage services, but also help simplify infrastructure management and orchestration. This enables comprehensive IT Operational Management (ITOM) and Life Cycle Management (LCM) capabilities that span compute as well as storage infrastructure, from BIOS and Firmware to nodes, software, and networking.

PowerFlex

PowerFlex is the software foundation of PowerFlex software defined infrastructure. It is a scale-out block and file storage service designed to deliver flexibility, elasticity, and simplicity with predictable high performance and resiliency at scale.

PowerFlex Manager

PowerFlex Manager is the software component in PowerFlex family that enables ITOM automation and life cycle management capabilities for PowerFlex systems.

The PowerFlex platform is available in multiple consumption options to help customers meet their project and data center requirements. PowerFlex appliance and PowerFlex rack provide customers comprehensive IT Operations Management (ITOM) and life cycle management (LCM) of the entire infrastructure stack in addition to sophisticated high-performance, scalable, resilient storage services. PowerFlex appliance and PowerFlex rack are the two most common consumption options. PowerFlex is also available on PowerFlex custom nodes without the ITOM and LCM capabilities. While the full PowerFlex Manager functions are not available for PowerFlex on AWS, the PowerFlex element manager can be leveraged with for this deployment option.

In PowerFlex 4.0, the unified PowerFlex Manager does the job of 3 separate tools used in previous releases – PowerFlex Manager, the core PowerFlex UI, and the PowerFlex gateway / IM. By building a next-generation UI on top of Kubernetes and embracing a modern development framework, the latest release of PowerFlex Manager continues to improve the ease of management.

PowerFlex File Services

PowerFlex File Controllers (sometimes casually referred to as File Nodes) are the physical nodes that enable the presentation of PowerFlex software defined File Services. They host the NAS Servers, which in turn host the tenant namespaces and file systems, mapping PowerFlex volumes to the file systems presented by the NAS Servers. All major protocols are supported: NFS, SMB/CIFS, FTP, NDMP, etc. Moreover, the NAS Servers support multiprotocol access to the shared file systems.

PowerFlex CSI/CSM

The most important component outside of PowerFlex that enables a flexible consumption model for Kubernetes is the PowerFlex CSI plug-in; developed as part of the Dell Technologies Kubernetes strategy. After loading the CSI for PowerFlex into a single distribution (or multiple Kubernetes distributions) one can simply consume from that one single underlying PowerFlex storage resource. When the Kubernetes deployments start to run low on PowerFlex storage resources, one is able to add a PowerFlex storage node, bolstering the pools capacity, or adding to the systems performance, as each PowerFlex storage node acts as a storage controller in Kubernetes.

Customers running Kubernetes clusters on PowerFlex benefit by using Dell Technologies Container Storage Modules (CSM). These modules provide enterprise storage capabilities to Kubernetes for cloud-native stateful applications. These modules reduce management complexity, so that developers can independently consume enterprise storage with ease and automate daily operations. The CSM extend storage functionality and capabilities beyond what can be done using the CSI driver alone, and at present there are modules that assist developers with replication, observability, authorization, application mobility and resiliency.

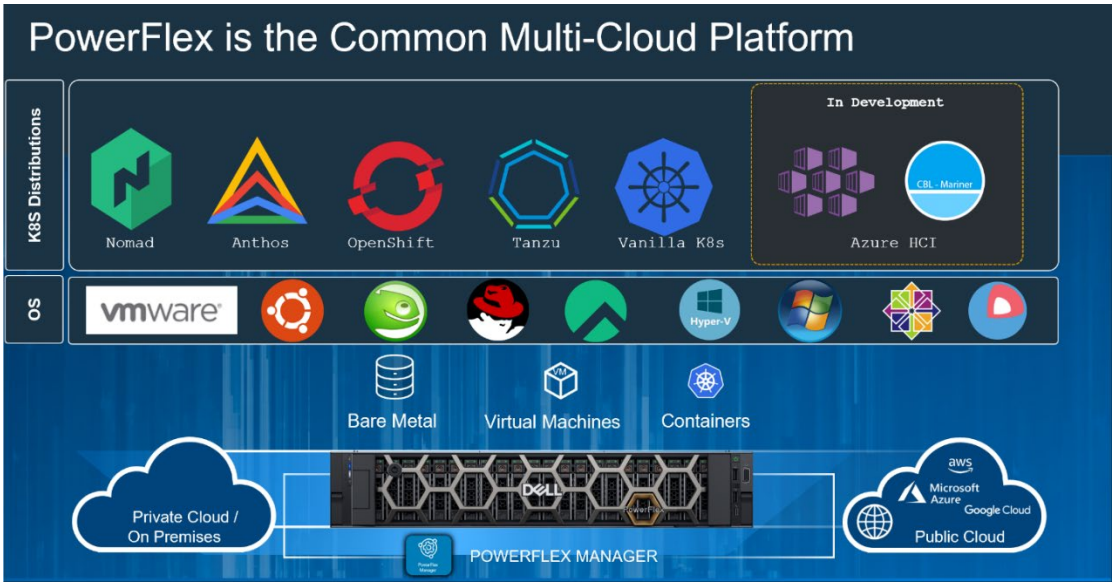

Figure 2. PowerFlex common multi-cloud

The CSI is a bridge between the PowerFlex System and single or multiple Kubernetes distributions. It is a storage broker which dynamically provisions volumes from PowerFlex by way of the PowerFlex API gateway. Once the volume is available on PowerFlex, it is immediately mapped to the requesting pod(s). Should a pod be destroyed or re-scheduled, the CSI plugin ensures that the volumes are re-mapped upon that pod's re-scheduling.

PowerFlex deployment architectures

PowerFlex software defined storage offers flexibility of deployment architecture to help best meet the specific deployment and architectural requirements. PowerFlex can be deployed in a two-layer (Independent), single-layer (Hyperconverged/HCI), or a mixture of the two architectures (Mixed).

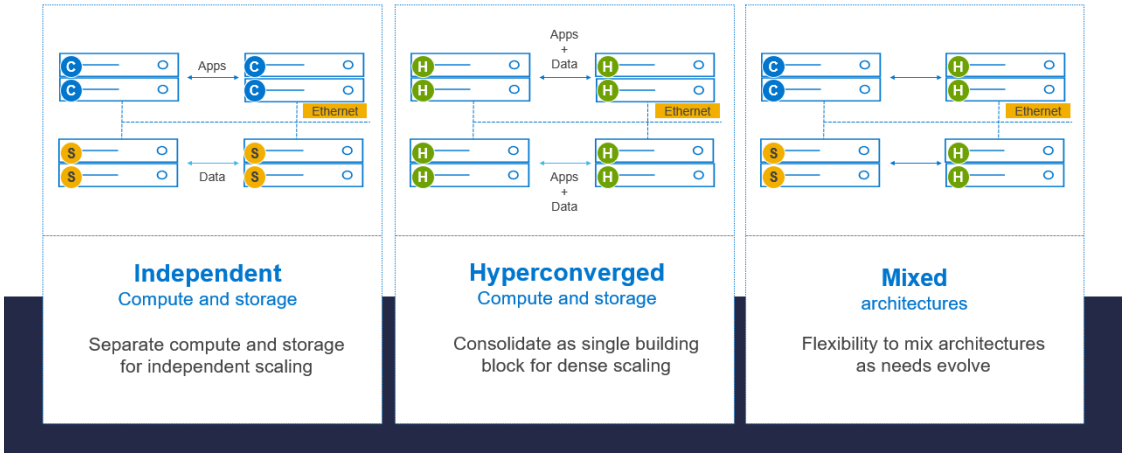

Figure 3. PowerFlex architectures

Independent architecture

In an independent architecture, or two-layer architecture, some nodes provide storage capacity and host datasets while other, separate, and independent nodes, host applications and workloads. PowerFlex manager provides LCM and IOTM for the entire infrastructure, including nodes that provide storage and nodes that host applications. Compute and storage resources can be scaled independently by adding nodes to the cluster. This separation of compute and storage resources can also help minimize software licensing costs in certain situations. And this architecture can be ideal for hosting high-performance high-value databases and application workloads.

Hyperconverged architecture

In an HCI architecture, each node in the cluster contributes storage resources as well as hosts applications and workloads. This architecture allows you to scale your infrastructure uniformly and with a pre-defined building block that adds both storage and compute resources. PowerFlex Manager provides ITOM and LCM capabilities for the entire infrastructure. This architecture is appropriate for data center and workload consolidation.

Mixed architecture

In a mixed architecture we have a combination of both the HCI and Independent architectures. As shown in the figure above there would be some storage only nodes, compute only nodes as well as HCI nodes as part of the same PowerFlex cluster. And using PowerFlex storage-only nodes, storage can be accessed and consumed by several applications and workloads that are hosted outside of the PowerFlex cluster. PowerFlex Manager provides LCM and ITOM for the storage infrastructure. This is a desirable architecture when working with an existing compute infrastructure and adding high-performance software defined storage. This can also be a starting point for a two-layer Server SAN deployment as external workloads are migrated to PowerFlex.

PowerFlex consumption options

PowerFlex rack

PowerFlex rack is a software defined storage platform that delivers flexibility, elasticity, and simplicity with predictable performance and resiliency at scale by combining compute as well as high performance storage resources in a managed unified network. This rack-scale engineered system, with integrated networking, enables customers to achieve the scalability and management requirements of a modern data center.

PowerFlex appliance

PowerFlex appliance is a PowerEdge server which has been configured to be a node in a software defined storage deployment. This offering allows customers the flexibility and savings to bring their own compatible networking.

PowerFlex Custom Nodes

PowerFlex Custom Nodes are validated server building blocks configured for use with PowerFlex. They are available with thousands of configuration options and are available for customers who prefer to build their own environments. Please consult the PowerFlex data sheet for more details on the available options.

PowerFlex on AWS

PowerFlex software can also be deployed in the public cloud and is supported as a new deployment option on AWS. It is available in the Amazon Marketplace (PowerFlex cloud storage). Offering the same performance, linear scalability and high resilience in the cloud as what our customers have experienced with on-prem deployments. In addition to larger volume sizes, higher performance, predictable scalability and high performance with low latency you can also get higher resilience when PowerFlex Fault sets are automatically distributed across multiple AWS Availability Zones. One can easily migrate data from an on-prem PowerFlex deployment to a PowerFlex on AWS, or from public cloud back to on-prem, by using the PowerFlex native replication technology.

Solution components

PowerFlex SRA

The Dell PowerFlex Storage Replication Adapter (SRA) for VMware Site Recovery Manager (SRM) allows SRM to implement disaster recovery for PowerFlex clusters. Using the PowerFlex replication engine, PowerFlex Storage Replication Adapter (SRA) supports all SRM functions including planned migration, failover, reprotect, and test failover.

PowerFlex offers the SRA for the Photon OS-based SRM appliances only. There is no SRA for the deprecated SRM Windows platform. The VMware website hosts the SRA and is available and downloadable under various versions of SRM. The SRA is titled as seen in Figure 4.

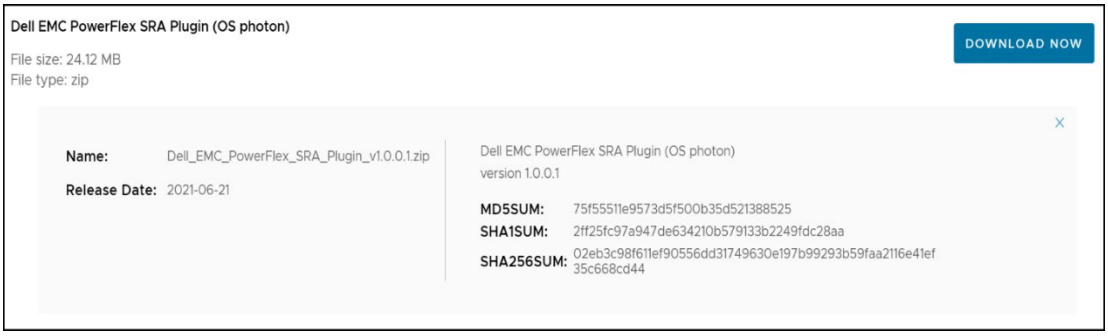

Figure 4. PowerFlex SRA download

The following table, Table 2, presents the PowerFlex SRA Compatibility Matrix for PowerFlex and VMware SRM. To determine the supported versions of ESXi and vCenter with the chosen SRM version, see the Site Recovery Manager Compatibility Matrices.

Table 2. Supported versions for PowerFlex SRA 1.0.0.1

PowerFlex 3.5.1 or higher or 4.x	VMware SRM[1][2]
PowerFlex Rack PowerFlex Appliance PowerFlex Custom Node	8.2 8.3 8.4[3] 8.5[3] 8.6[3] 8.7[3]

[1] For complete and detailed VMware supported versions with SRM, including ESXi and vCenter, see the VMware vCenter Site Recovery Manager Compatibility Matrix available on the VMware support website.

[2] The SRA supports VMware Site Recovery Manager patched releases beyond the initially supported major release. For example, if an updated version of 8.6 became available (8.6.1), it would be supported. Dell will not do additional qualification, however, so it will not appear in VMware's compatibility matrix.

[3] Granted by VMware through equivalency.

Note: PowerFlex does not support ESXi 8, though it is possible to use vCenter 8 with ESXi 7 if desired.

Note: PowerFlex does not have a VASA Provider 3.0 and therefore cannot support Virtual Volume (vVol) replication with SRM.

PowerFlex replication

Replication ensures the data protection of designated volumes within a PowerFlex environment. It maintains a remote copy of one or more volumes, configured between two PowerFlex systems designated as peer systems. PowerFlex supports asynchronous replication with an RPO as low as 15 seconds. The peer systems are connected using a LAN or WAN and are physically separated for protection purposes.

Replication is defined within the scope of a protection domain. All volumes that participate in replication are contained in Replication Consistency Groups (RCGs). An RCG is an entity that includes one or more sets of consistent volume pairs. A volume from a single protection domain on the source, is replicated to a remote volume from a single protection domain on the target. This creates a consistent pair of volumes. Replication journal capacity from storage pools in the same protection domain is shared among RCGs.

The PowerFlex Storage Data Replicator (SDR) software handles replication activities and manages the I/O of replicated volumes. The SDR is deployed on the same server as the Storage Data Server (SDS). Only I/O from replicated volumes flows through the SDR. See the Dell PowerFlex: Introduction to Replication white paper for more details on the replication technology along with deployment and configuration details and design considerations for replicating PowerFlex clusters.

VMware vSphere and SRM

Compatible versions of SRM, vCenter Server, and ESXi hosts are required for this solution. The installation and configuration of these components are not covered in this whitepaper with the exception of portions directly related to storage and replication. For more information about SRM installation and configuration see VMware's documentation.

Environment and system requirements

The following sections cover the requirements for implementing PowerFlex replication with VMware SRM. Where necessary, the differences between PowerFlex 3.5+ and 4.x are noted.

Note: For ease of reference, the PowerFlex platform versions will simply be referred to as PowerFlex 3 and PowerFlex 4.

Prerequisites

Before the Dell PowerFlex SRA can be installed, the following prerequisites must be met. This holds true for both PowerFlex versions 3 and 4:

- VMware vCenter and Site Recovery Manager (SRM) must be installed and configured at each site according to the VMware documentation. The two SRM sites must be in a paired state.
- Two PowerFlex clusters with replication configured between them.
  - Any ESXi host that is part of either SRM site must have the SDC package installed on it, such as sdc-3.6.500.106-esx7.x.zip.
  - Though not required for initial installation, Dell recommends creating at least one Replication Consistency Group (RCG) with a device pair configured between the two clusters. The source volume should be presented to one vCenter while the peer or target volume is presented Read Only to the other vCenter. This process is covered later in the paper.
- HTTPS TCP port 443 must be opened between the SRM server and PowerFlex at each site.

Solution architecture

This section describes PowerFlex software-defined storage replication architecture. Both PowerFlex and SRM support bi-directional setups, meaning a PowerFlex site can simultaneously be a protection and recovery site. The architecture is presented as logical, meaning that the type and version of the PowerFlex system, such as appliance, rack, 3, 4, etc., is not specifically noted.

Logical architectures

Figure 5 represents an environment where two equal PowerFlex systems are configured, but there is only one active protected site and one idle recovery site. Such an architecture enables the business to continue to run at full capacity at the recovery site in the event of a failure. In other words, the recovery site has all the resources required to run the production workload as it is a mirror image.

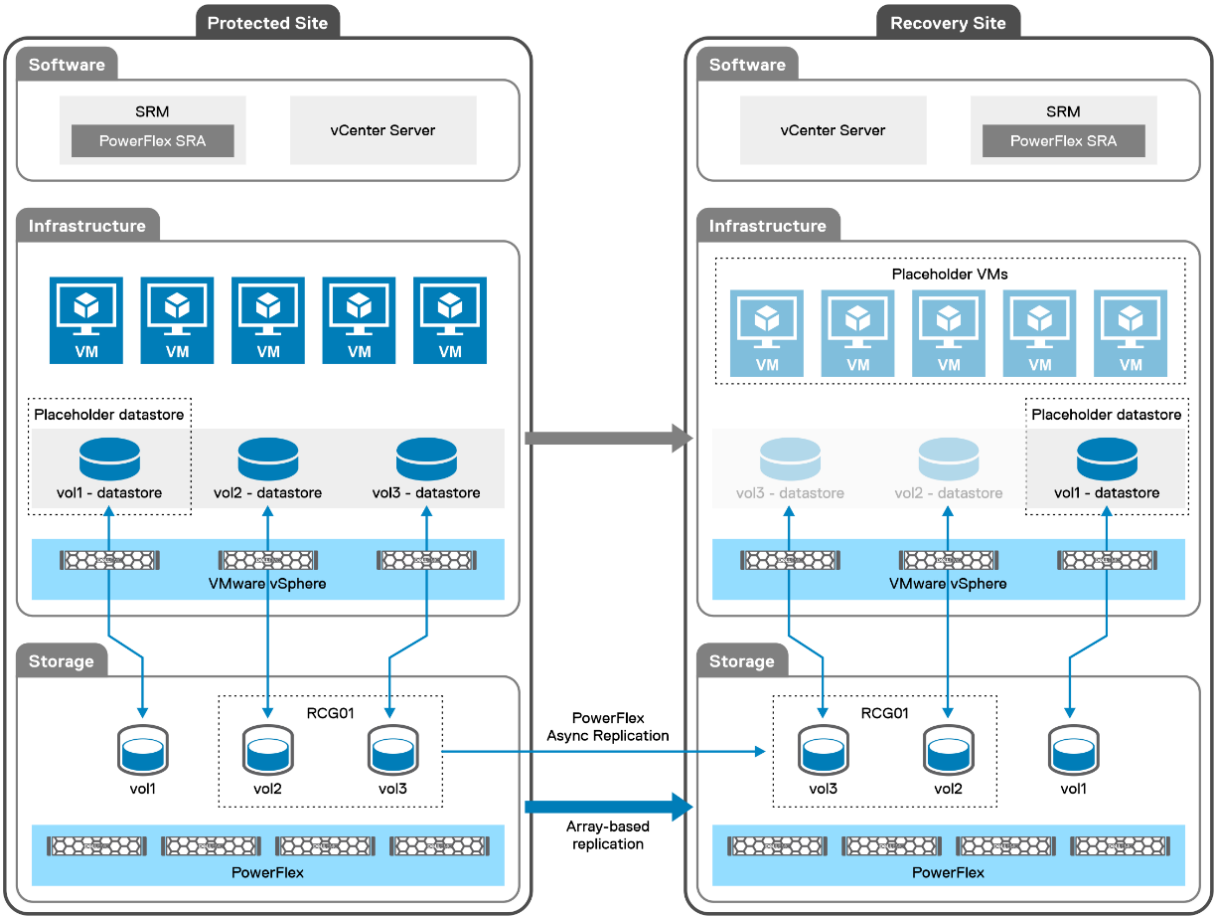

Figure 5. Single directional architecture with PowerFlex replication and SRM

As leaving hardware idle is not an option for many customers, Figure 6, displays the same two PowerFlex systems, but this time each site is actively servicing VMs and has a role as a protection and recovery site. While this configuration is able to utilize all the hardware at both sites, rather than have a system remain idle, it is important to remember that in the event of failure, a single site must run the load from both environments. As such, it is imperative that the business understands the SLAs or service level agreements (if any), and whether they permit a reduction of performance in the event of a failure. It is critical to size the environments appropriately, based on the SLAs, to account for the extra load during a failure.

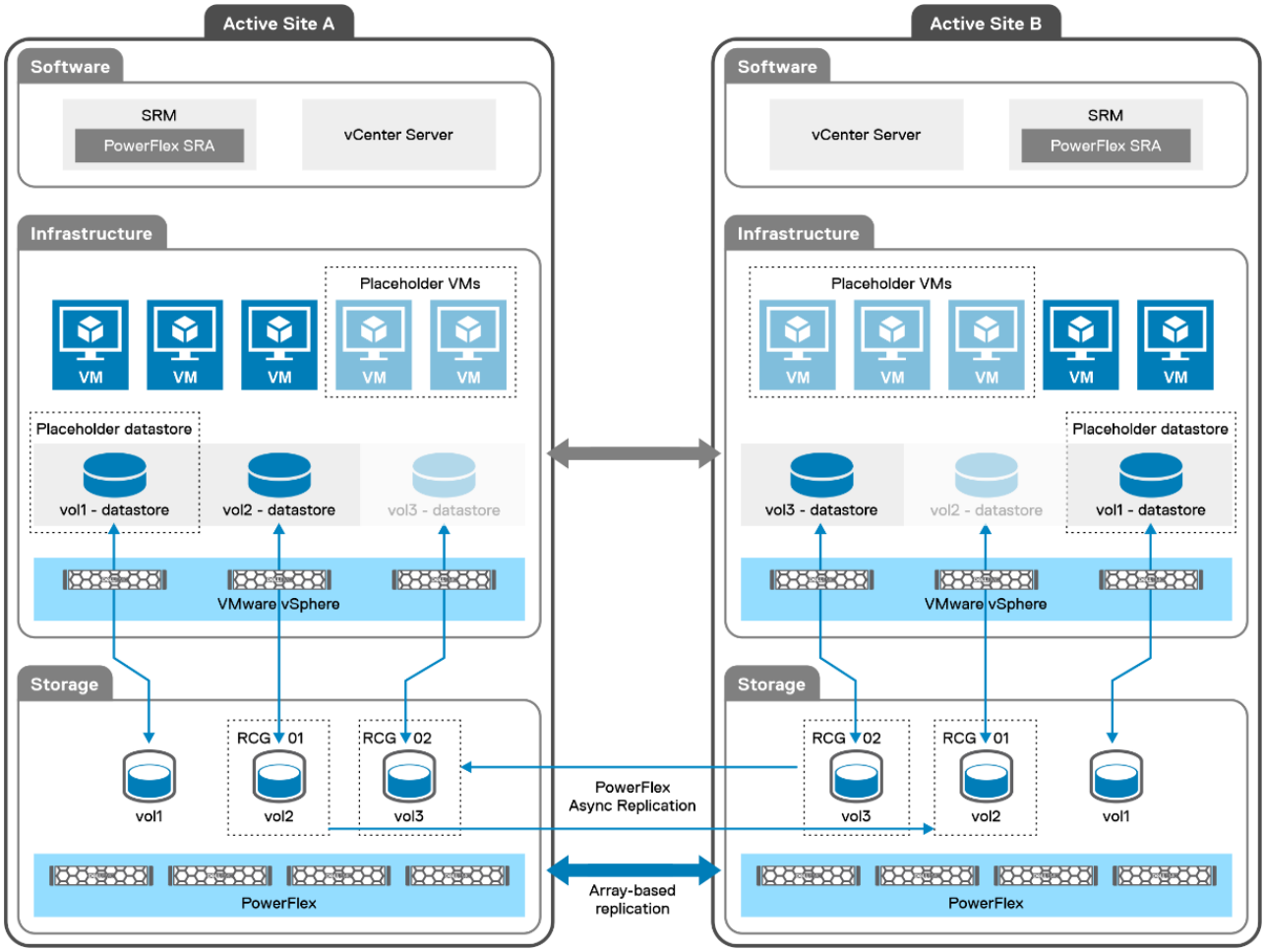

Figure 6. Bi-directional architecture with PowerFlex replication and SRM

Peered systems

The setup of PowerFlex replication is essentially the same for versions 3 and 4. It involves the peering of two systems by sharing root certificates and using the IP addresses of the SDRs. Once paired, each site will show the other as Connected as seen in Figure 7.

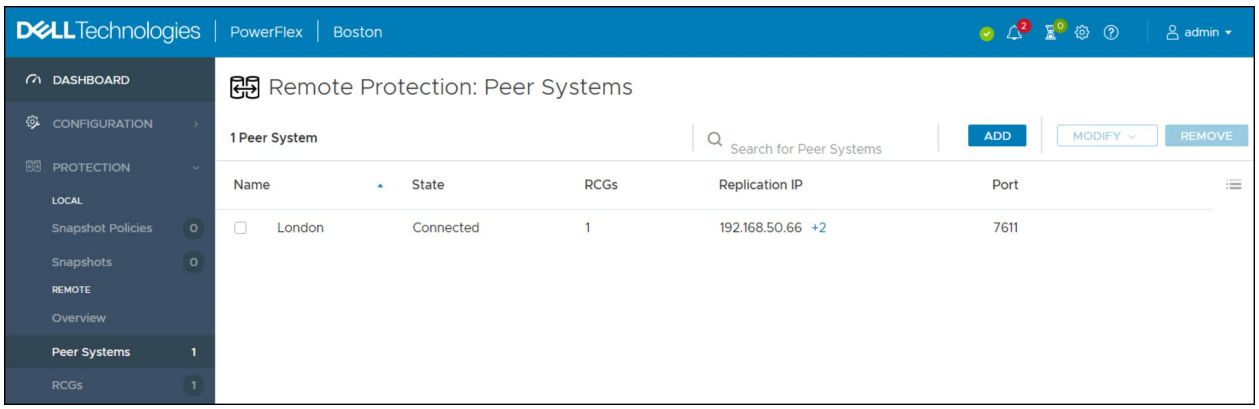

Figure 7. Peered PowerFlex systems

Replication Consistency Group

The steps to add an RCG in PowerFlex vary between versions 3.5+ and 4.x so each will be covered separately.

Note: This example, which uses a single source volume, assumes the user mapped it to the ESXi hosts at the protection site and created a datastore on it. Replication is being added after that step is complete. This is not an uncommon practice, though it is also acceptable to present volumes after replication is active.

PowerFlex 3

Begin by logging into the PowerFlex GUI as a user with administrator privilege.

1. From the PowerFlex GUI in Figure 8, select **Protection** > **RCGs > ADD**.

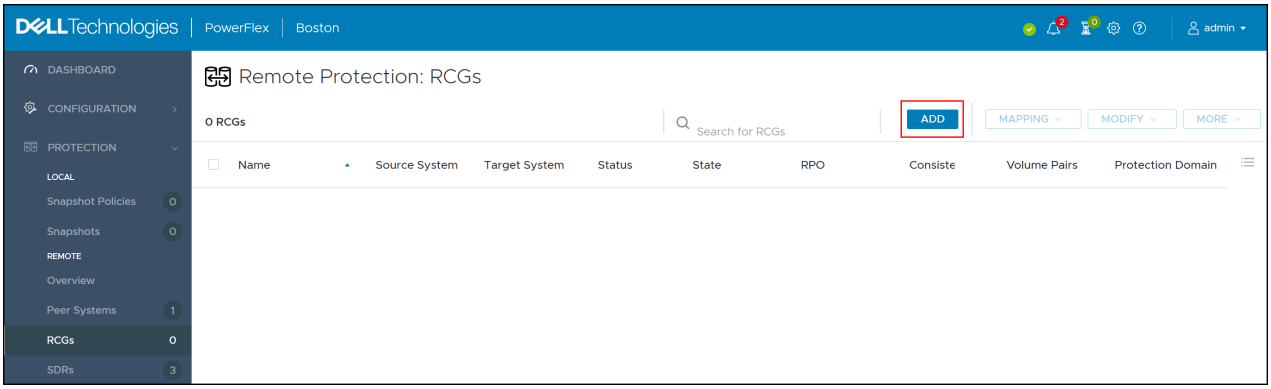

Figure 8. Add RCG – Step 1

2. In the next screen in Figure 9, provide the **RCG Name,** enter the **RPO** value in seconds or minutes (default is 60 seconds), select the **Source Protection Domain**, **Target System**, **Target Protection Domain**, and click **NEXT**.

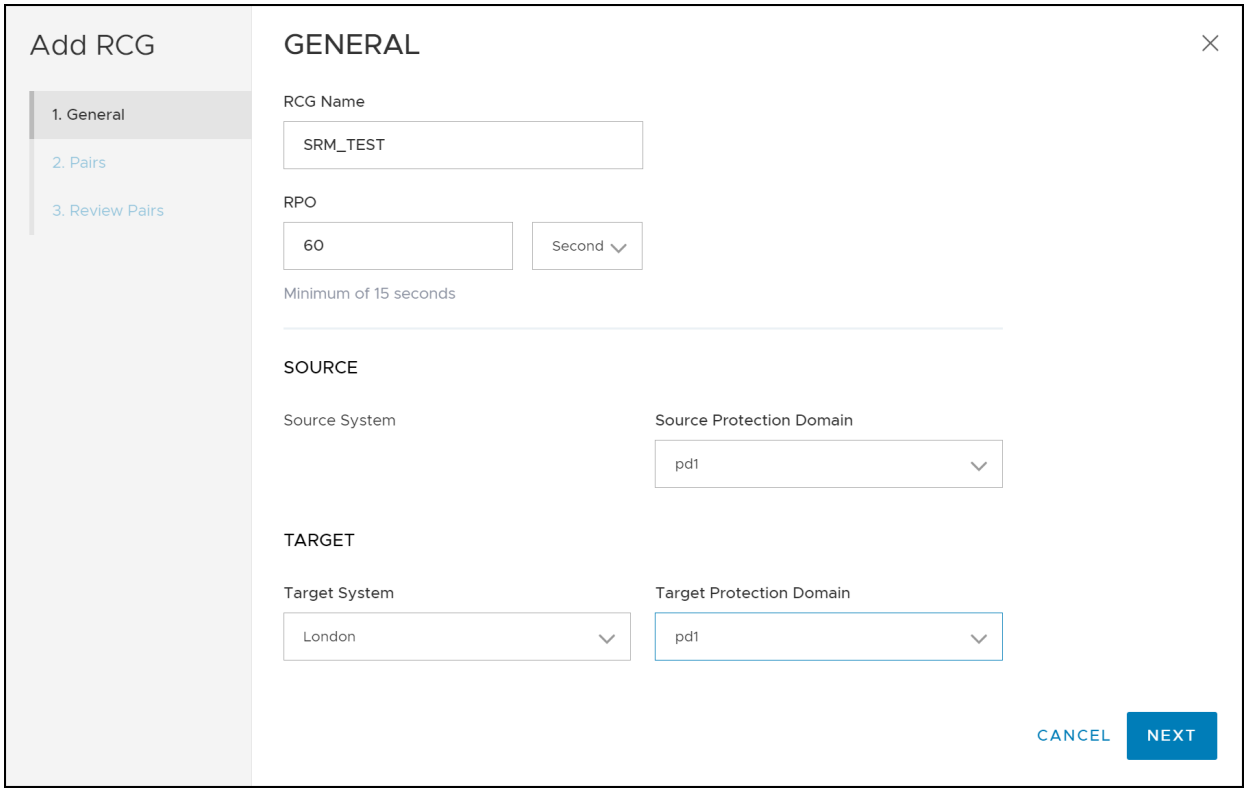

Figure 9. Add RCG – Step 2

3. Select the volume from the **Source** column and then select a volume of the same size from the **Target** column as shown in Figure 10.

   Note that the target volume will not appear in the screen until the source volume is highlighted. Once selected hit **ADD PAIR** and **NEXT.**

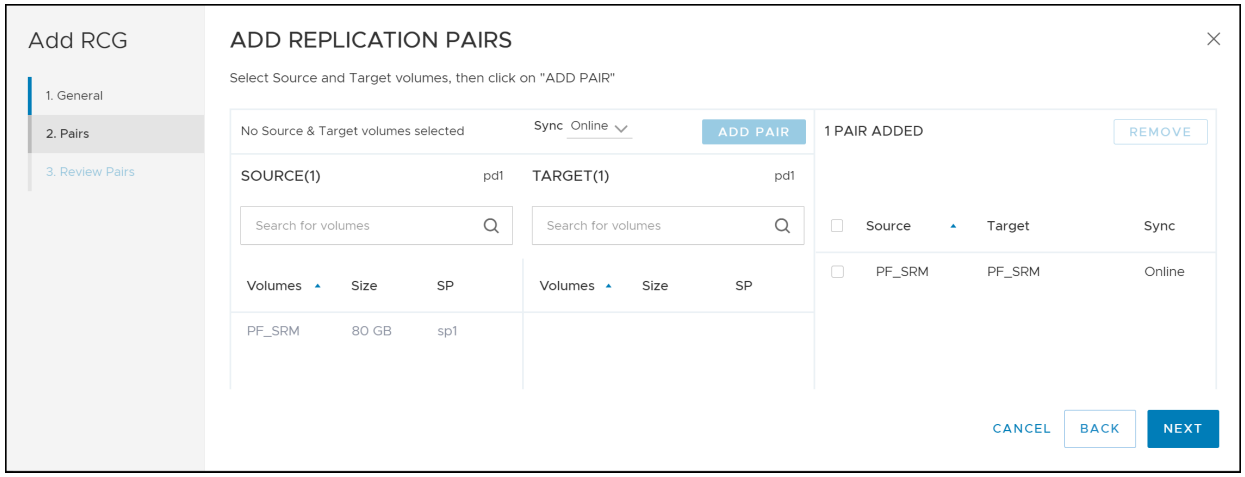

Figure 10. Add RCG – Step 3

Note: A volume of equal size must already be present at the target site. PowerFlex 3 cannot auto provision the volume.

4. Review the details shown in Figure 11 to ensure the correct source and target volume pairs are selected and click **ADD AND ACTIVATE**.

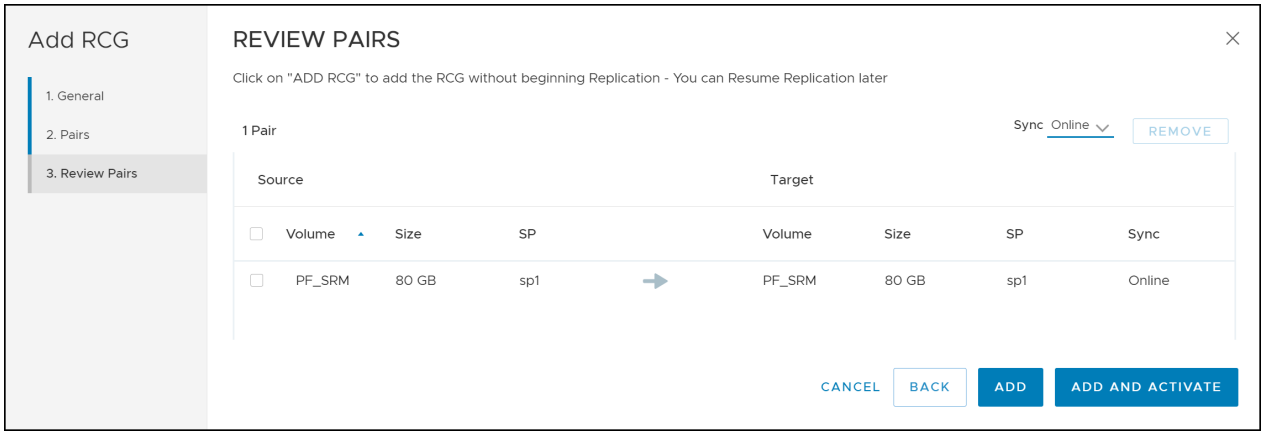

Figure 11. Add RCG – Step 4

The volumes will begin to synchronize. When the **Consistency** column moves from **Partially Consistent** (as seen in Figure 12) to **Consistent**, synchronization is complete. Since this is asynchronous replication, the remote volume is write-disabled.

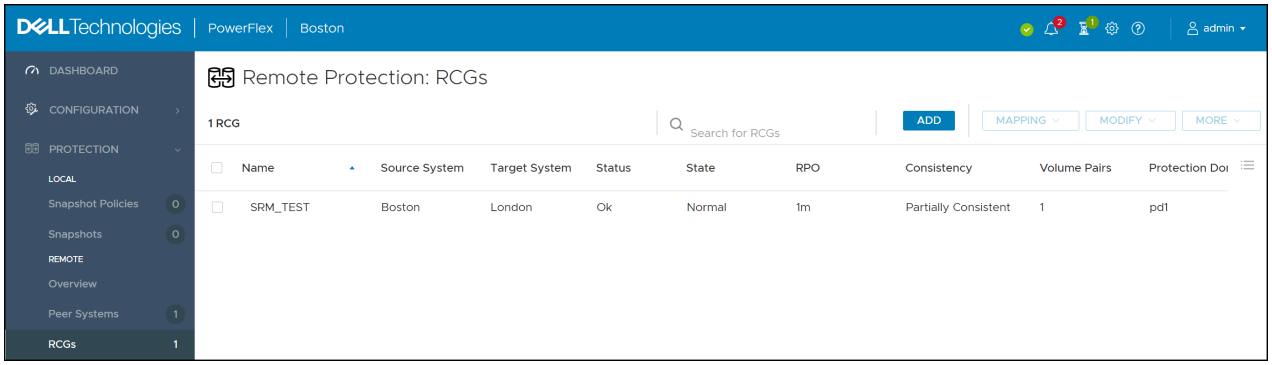

Figure 12. RCG entering consistency

Volume mapping

After creating the RCG, present the remote volume or volumes to the ESXi hosts at the recovery site. The remote volume is mapped as a read-only device. When a test failover or actual failover is required, PowerFlex will change the access of the device to read/write automatically.

The mapping of the replicated remote volume is only possible in the RCG screen, not the Volumes screen. Any attempt to map in the Volumes screen will result in the message in Figure 13.

Note: Although Dell recommends using read-only for access, using a setting of no_access will also work.

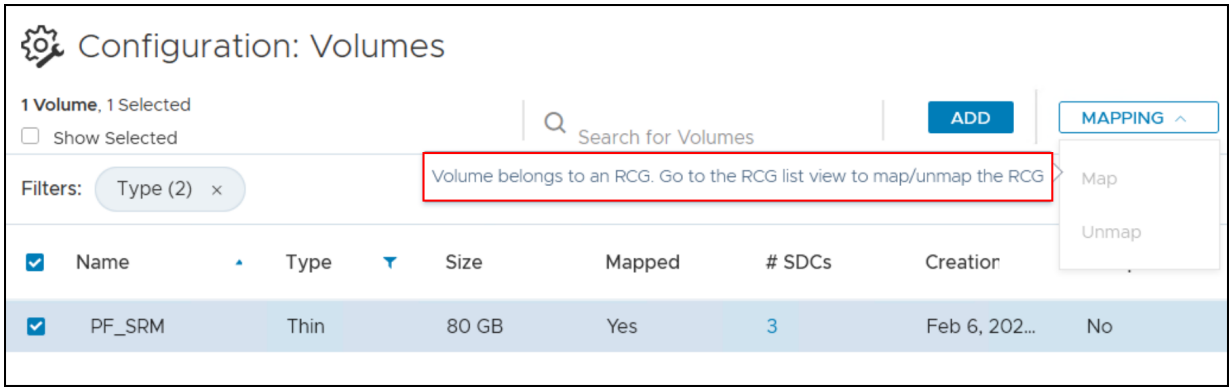

Figure 13. Attempted mapping of a volume in an RCG

Instead, navigate to the **RCGs** menu on the left-hand side, check the box next to the appropriate RCG in the screen on the right and select **Map** from the **MAPPING** drop-down menu as shown in Figure 14.

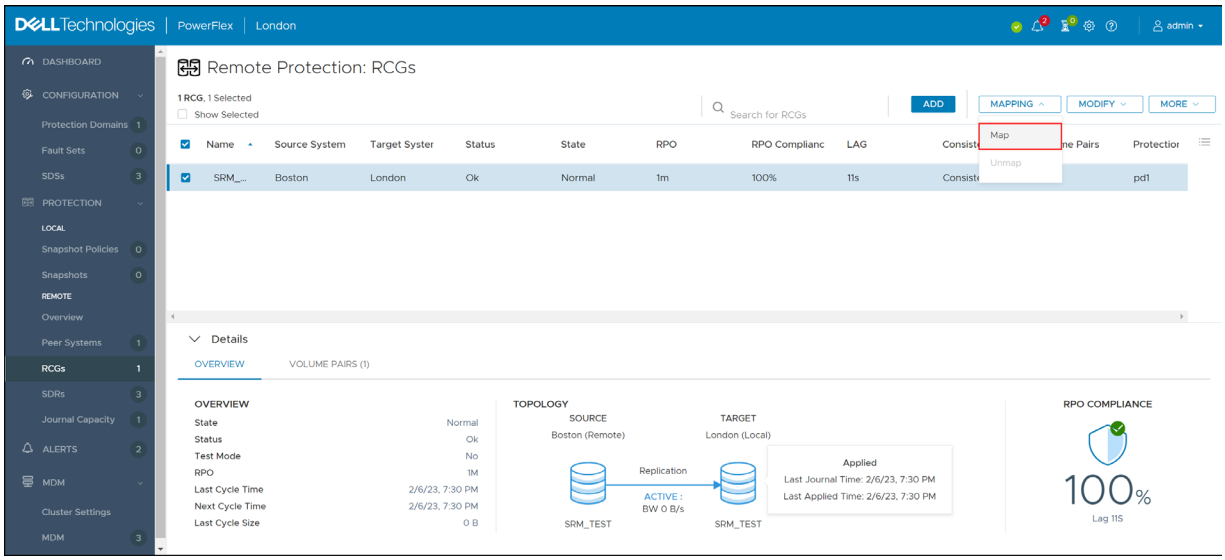

Figure 14. Map recovery volumes read-only to recovery site

Next, select the checkbox next to the ESXi hosts in the recovery site, and click **MAP**. Ensure the **Access Mode** is set to **Read Only** (default). Then click **APPLY**. This is shown in Figure 15.

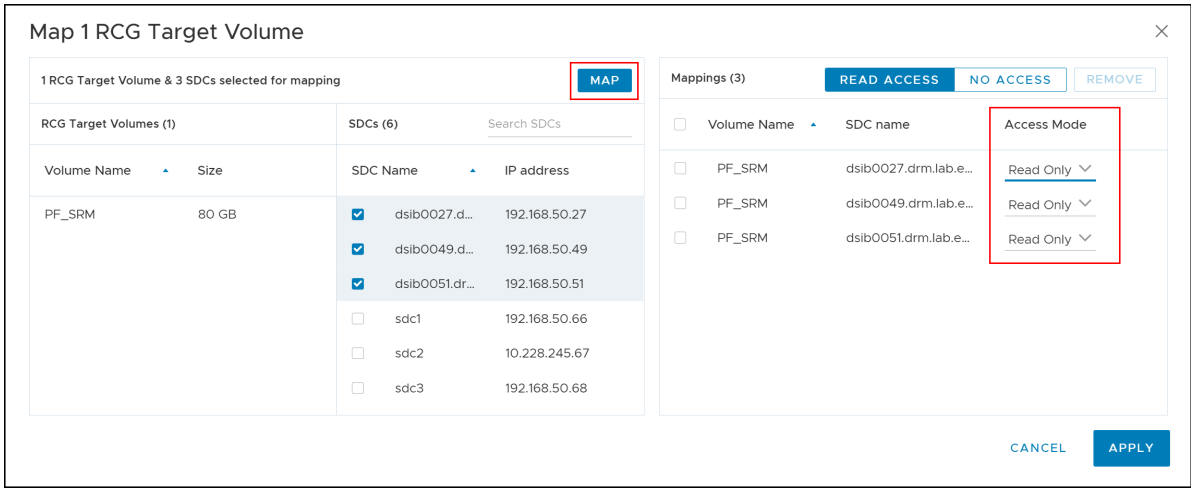

Figure 15. Mapping volumes as read-only to recovery site

Note that the device must appear in the ESXi host under the storage adapter for PowerFlex, as shown in Figure 16, so that is available for mounting when running either test failover or failover. ESXi will recognize the device regardless of the **Access Mode**.

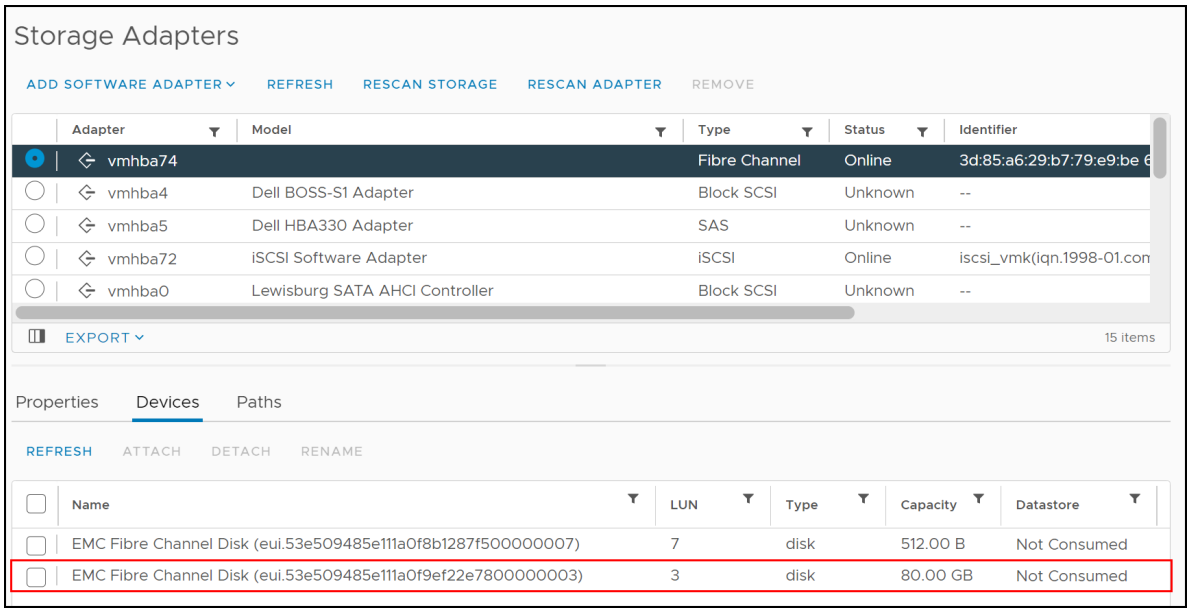

Figure 16. RCG read-only volume at recovery site

PowerFlex 4

Log into the PowerFlex Manager GUI as a user with administrator privilege.

1. From the PowerFlex Manager GUI, select the drop-down menu **Protection** > **RCGs**. Select **+Add RCG** in the right-hand corner as seen in Figure 17.

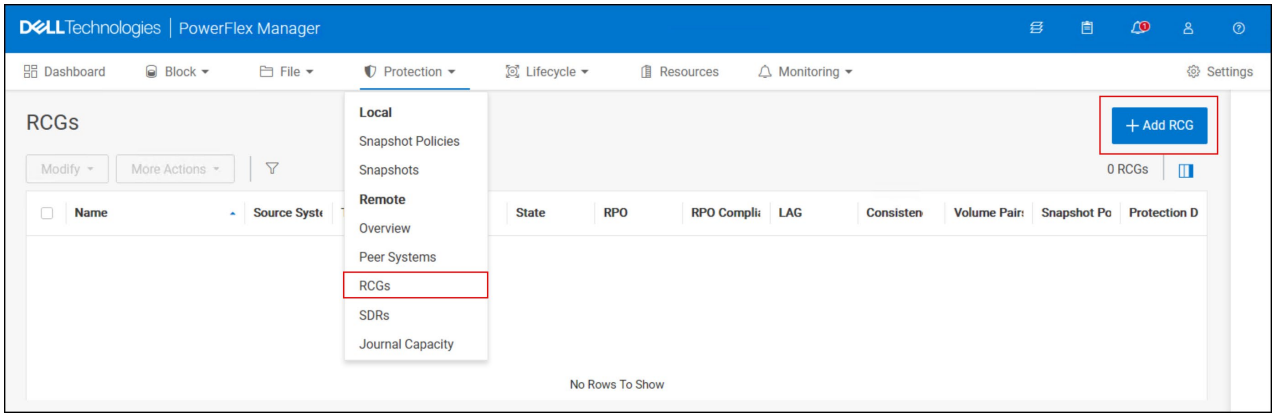

Figure 17. Add RCG – Step 1

2. As shown in Figure 18 below, enter the **RCG Name,** the **RPO** value in seconds or minutes (default is 60 seconds), select the **Source Protection Domain**, **Target System**, **Target Protection Domain**, and then click **Next**.

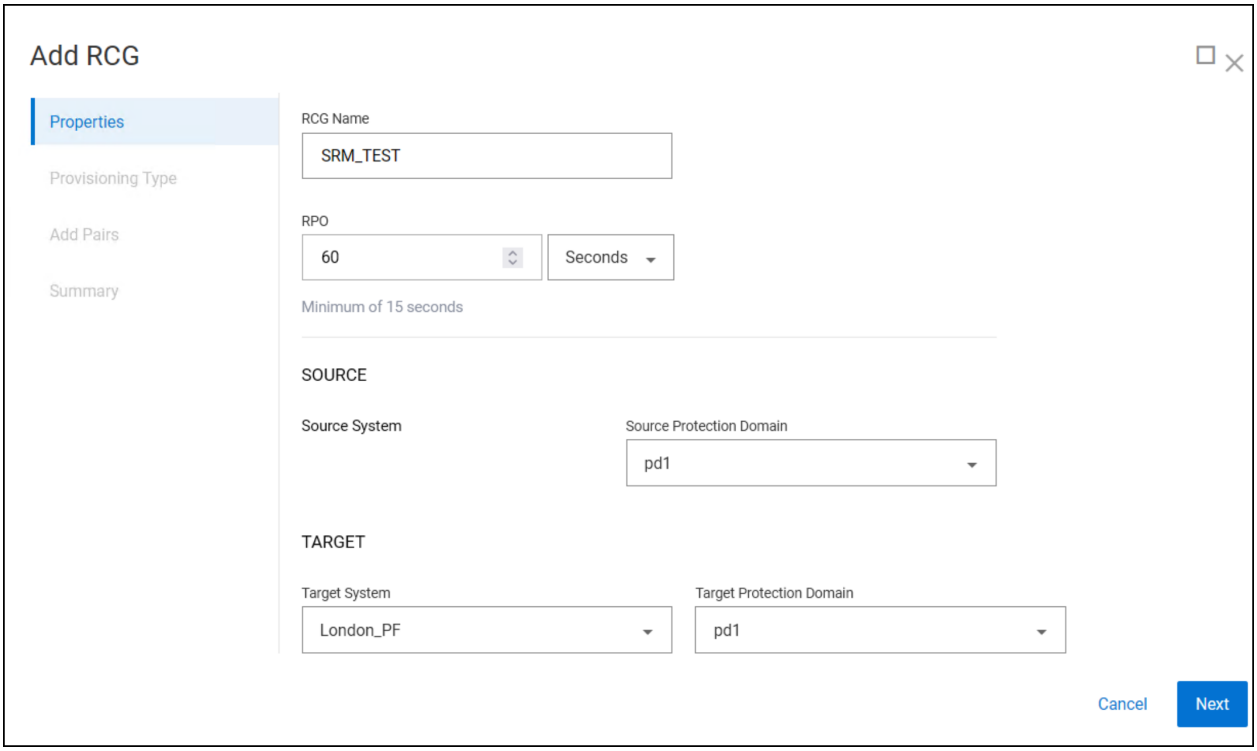

Figure 18. Add RCG – Step 2

3. In the next step, select the type of provisioning. Unlike PowerFlex 3, PowerFlex 4 offers the traditional **Manual Provisioning** option and also **Auto Provisioning.** This is shown in Figure 19.

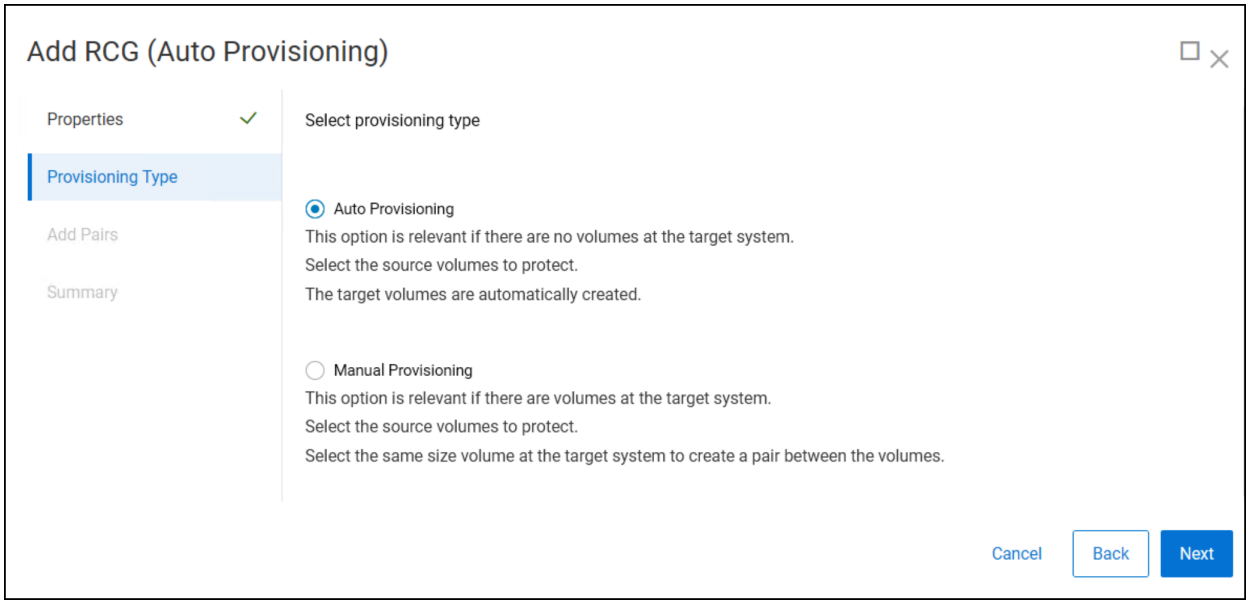

Figure 19. Add RCG – Step 3

4. If using Auto Provisioning as shown in Figure 20, choose the **Type** of volume, Thin or Thick, the **Storage Pool** and select **Add Pair** and **Next**.

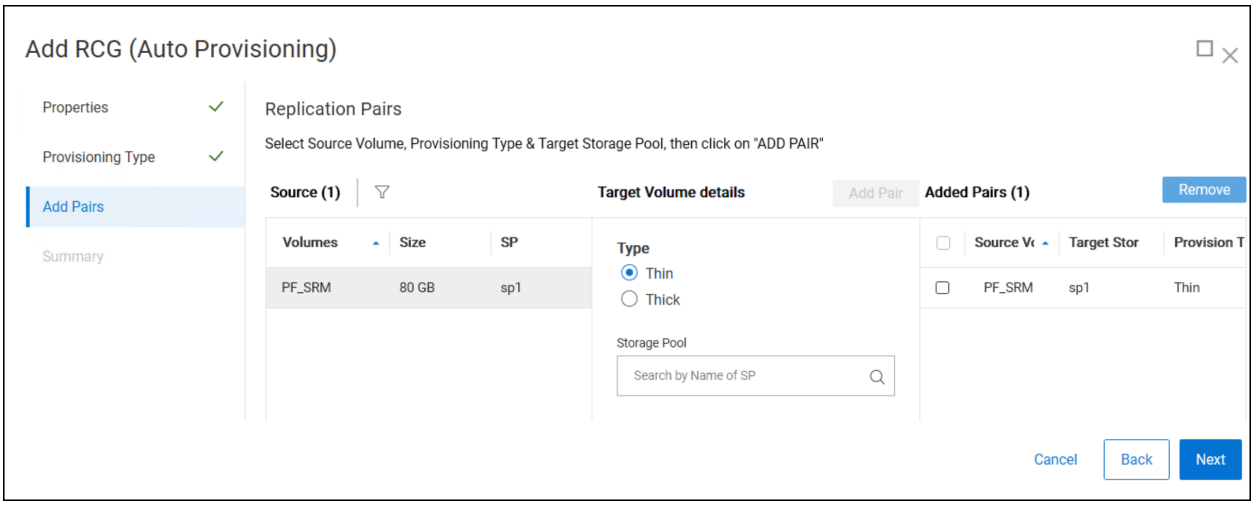

Figure 20. Add RCG – Step 4a

If instead **Manual Provisioning** is chosen as shown in Figure 21, select the volume from the **Source** column and then select a volume of the same size from the **Target** column and **Add Pair** then **Next**. This is similar to PowerFlex 3.

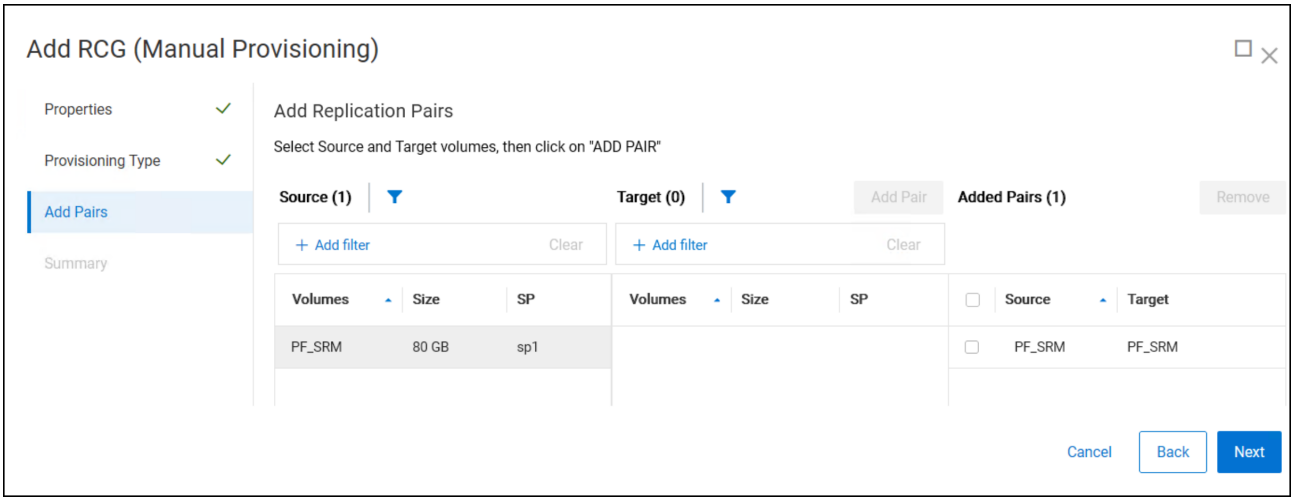

Figure 21. Add RCG – Step 4b

5. In the final step, for **Auto Provisioning** shown in Figure 22, select **Add and Activate**.

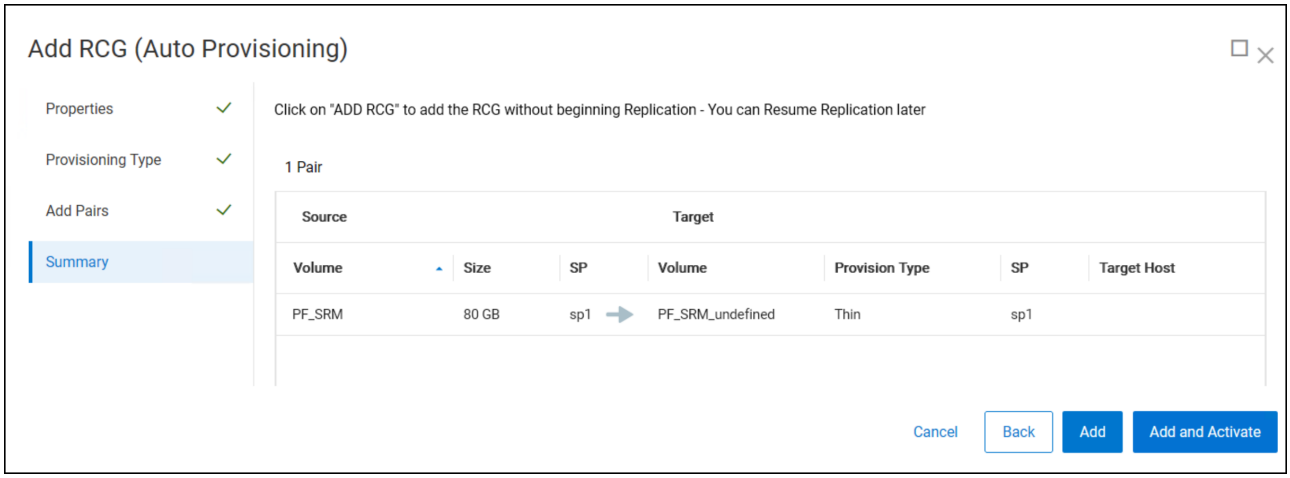

Figure 22. Add RCG – Step 5a

Similarly, for **Manual Provisioning** shown in Figure 23, select **Add and Activate**.

Figure 23. Add RCG – Step 5b

The volumes will begin to synchronize. When the **Consistency** column shows **Consistent** as seen in Figure 24, synchronization is complete. Since this is asynchronous replication, the remote volume is write-disabled.

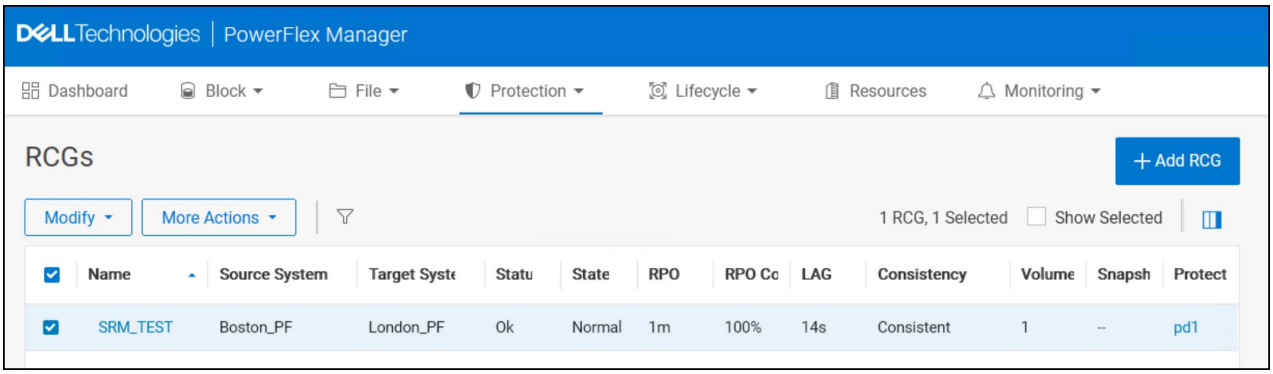

Figure 24. RCG entering consistency

Volume mapping

After creating the RCG, present the remote volume to the ESXi hosts at the recovery site. The remote volume is mapped as a read-only device. When a test failover or actual failover is required, PowerFlex will change the access of the device.

Because of a change in the PowerFlex Manager GUI in PowerFlex 4, the mapping of the replicated remote volume is not possible in the RCG screen as there is no mapping menu. The volume must be mapped through the command line interface (CLI). To do this, login to the **scli** interface as the **admin** user. Query for the existing replication pairs by running **scli –query_all_replication_pairs**, then map the volume read-only to the ESXi host(s) with **scli –map_volume_to_host**. Note that if there are multiple ESXi hosts, the flag **allow_multiple_hosts** is required. An example is shown in Figure 25.

```
[root@dsib2186 ~]#  scli --login --management_system_ip 10.228.246.183  --username admin --pas
sword U
Successfully generated login certificate at /root/.scli/cli_cert.p12
Logged in. User role is SuperUser. System ID is b9288d7890d2360f
[root@dsib2186 ~]# scli --query_all_replication_pairs

CG Id: 69febf3400000000 CG Name: SRM_TEST
Peer System Name: Boston_PF Direction: REMOTE_TO_LOCAL
    Local ID: 35d8730c00000000 Remote ID: 74bff8b900000000 Copy Type: ONLINE_COPY Initial Copy
 State: DONE Name: (null)
    Local Volume ID 06927bea00000003 Remote Volume ID: 285fbcd300000000 Local Volume Name: PF_
SRM Remote Volume Name: PF_SRM
    Initial Copy progress: 100% Local Activity State: REPLICATION_ENABLED
query_all_replication_pairs returned 1 Replication Pairs.
[root@dsib2186 ~]# scli --map_volume_to_host --volume_name PF_SRM --sdc_name dsib0078.drm.lab.
emc.com --access_mode read_only --allow_multi_map
Successfully mapped volume PF_SRM to host dsib0078.drm.lab.emc.com
[root@dsib2186 ~]#
```

Figure 25. Map volume read-only to the recovery site ESXi hosts

Once the volume is mapped, the PowerFlex environment setup is ready for SRM integration.

Additional pair volume mapping in PowerFlex 4

Although it is not possible to map a volume at the target site during the initial RCG creation, the feature is available for additional pairs. Take the following steps to add and map an additional volume pair to an existing RCG.

1. Highlight the RCG and select the **Modify** drop-down menu. Choose **Add Pair**. Note the menu, as shown in Figure 26, is relocated so as not to obscure the RCG name.

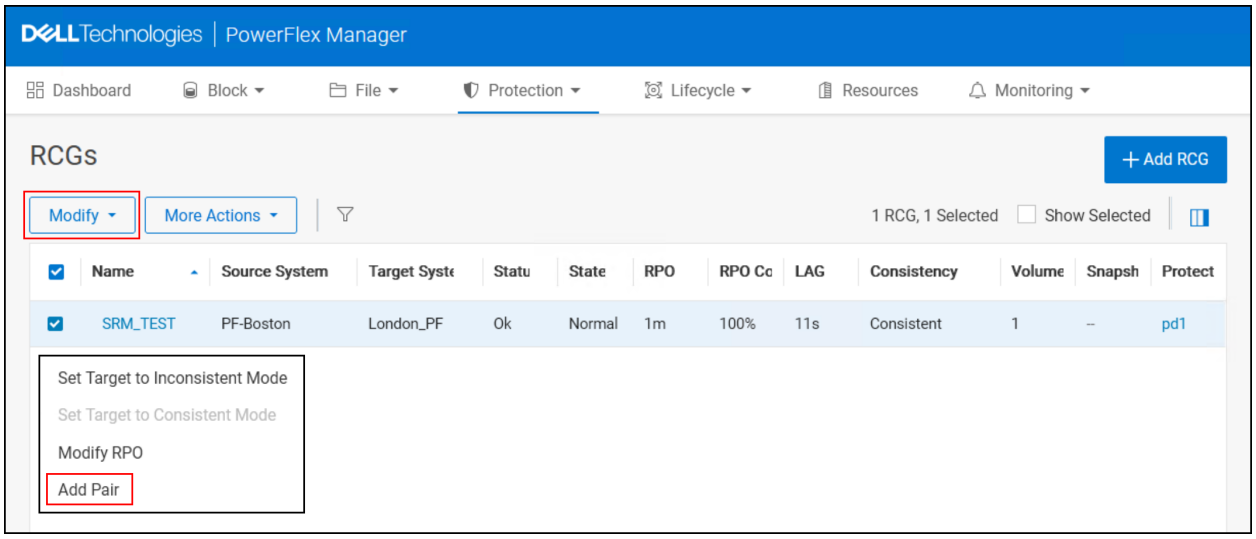

Figure 26. Adding pairs to an existing RCG

2. As previously detailed, begin by selecting the **Provisioning Type** (not shown), then add the pair. Either provisioning type permits mapping, though in this

example Auto Provisioning was selected. After adding the pair, the user is presented with a new **Mapping** screen. From here, the user can now select which hosts to map the volume to, in a similar manner as PowerFlex 3.

In the screen shown in Figure 27, select the volume, the host, and then choose **Map** before clicking **Next**.

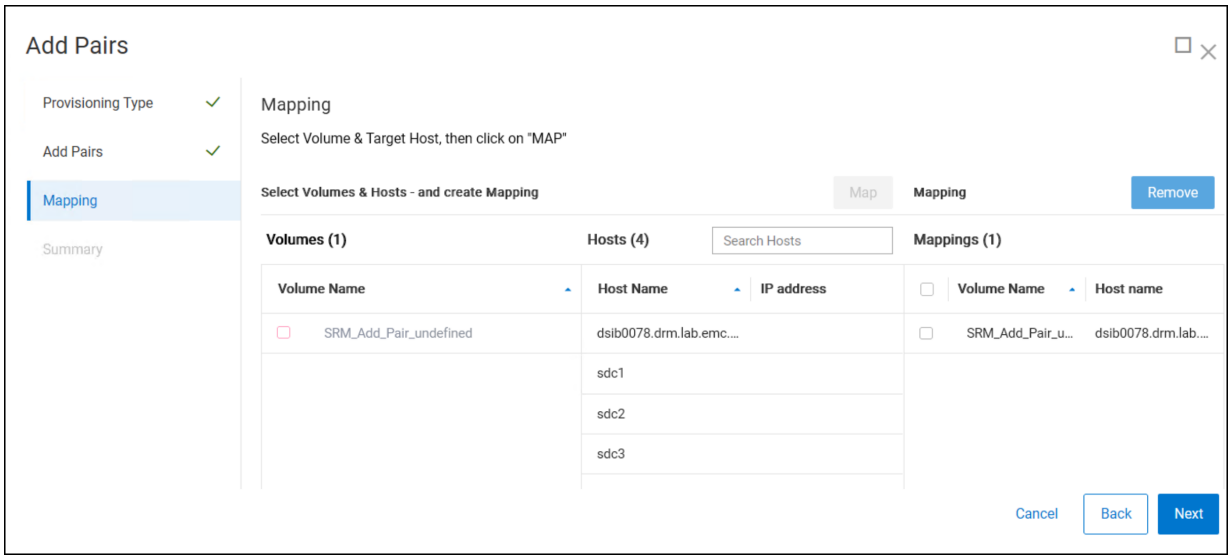

Figure 27. Mapping additional pairs

3. Review the summary and select **Add Pairs** to complete the process as seen in Figure 28.

Figure 28. Summary of add pairs

The pair is now added to the RCG and mapped to the target host, making it available for use with SRM.

Note: It is possible for the user to create an RCG using **scli** that contains no pairs, and then add all pairs using the wizard which will permit mapping. This avoids having to map the first volume using **scli** though CLI is necessary using either method because the GUI does not permit the creation of an empty RCG.

Dell PowerFlex SRA Installation

This section describes the implementation of the Dell PowerFlex SRA.

Before installing the Dell EMC PowerFlex SRA, ensure that VMware SRM is installed and configured on each vCenter and that they are paired to each other. For more information about installing Site Recovery Manager, see the VMware vCenter Site Recovery Manager documentation.

Installation steps

1. Download the Dell PowerFlex SRA from VMware as covered in the PowerFlex SRA section.
2. Verify that there is an RCG with at least one volume pair on the PowerFlex clusters. There should also be a datastore for each device and at least one VM on each datastore. This will be necessary to create protection groups and recovery plans within VMware SRM.

Note: Dell recommends using a custom System name for the PowerFlex cluster rather than the default, generated one.

3. Extract the Dell EMC PowerFlex SRA compressed tar file from the downloaded zip:

   Dell_EMC_PowerFlex_SRA_1.0.0.1.tar.gz

4. Navigate to the protection site SRM Appliance Management UI at: **https://<SRM FQDN>:5480**. The **Welcome to VMware Appliance Management** dialog box appears as shown in Figure 29.

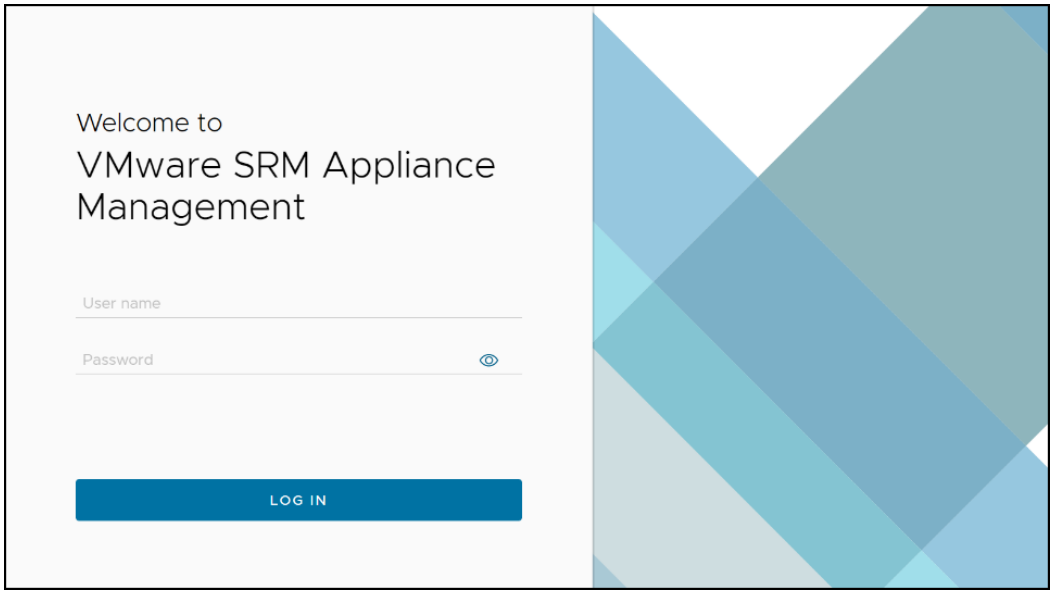

Figure 29. SRM Appliance Management screen

5. Enter the credentials (default administrator user is 'admin'), and select **LOG IN**.
6. In the SRM Appliance Management dialog box, navigate to Storage Replication Adapters in the left-hand panel. Next, select **NEW ADAPTER** in the right-hand panel as shown in Figure 30.

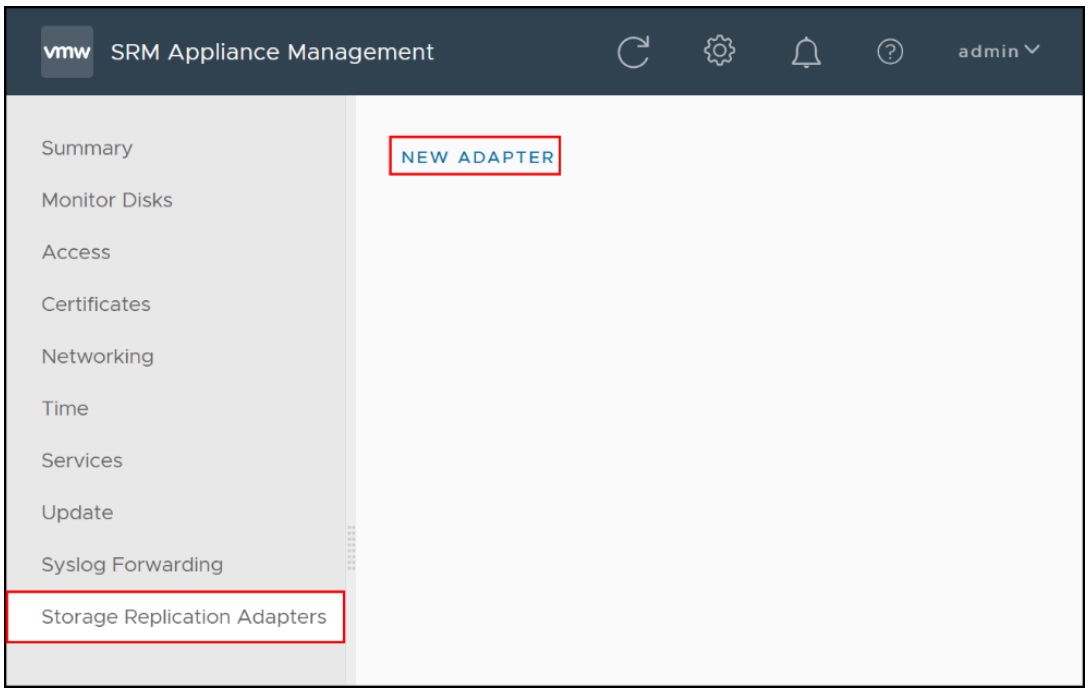

Figure 30. SRM Storage Replication Adapters – New Adapter

7. In the New Adapter dialog box shown in Figure 31, select **UPLOAD** and navigate to the file **Dell_EMC_PowerFlex_SRA_1.0.0.1.tar.gz**. The file will immediately begin uploading. Figure 32 shows the progression.

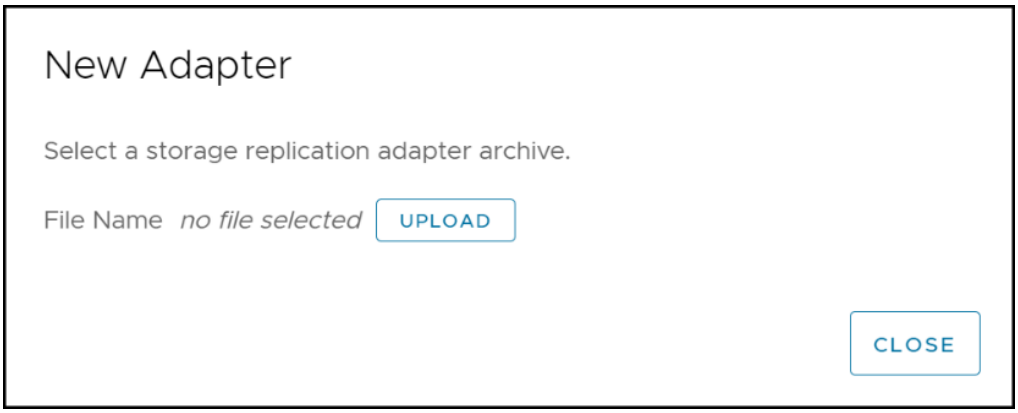

Figure 31. Upload new adapter

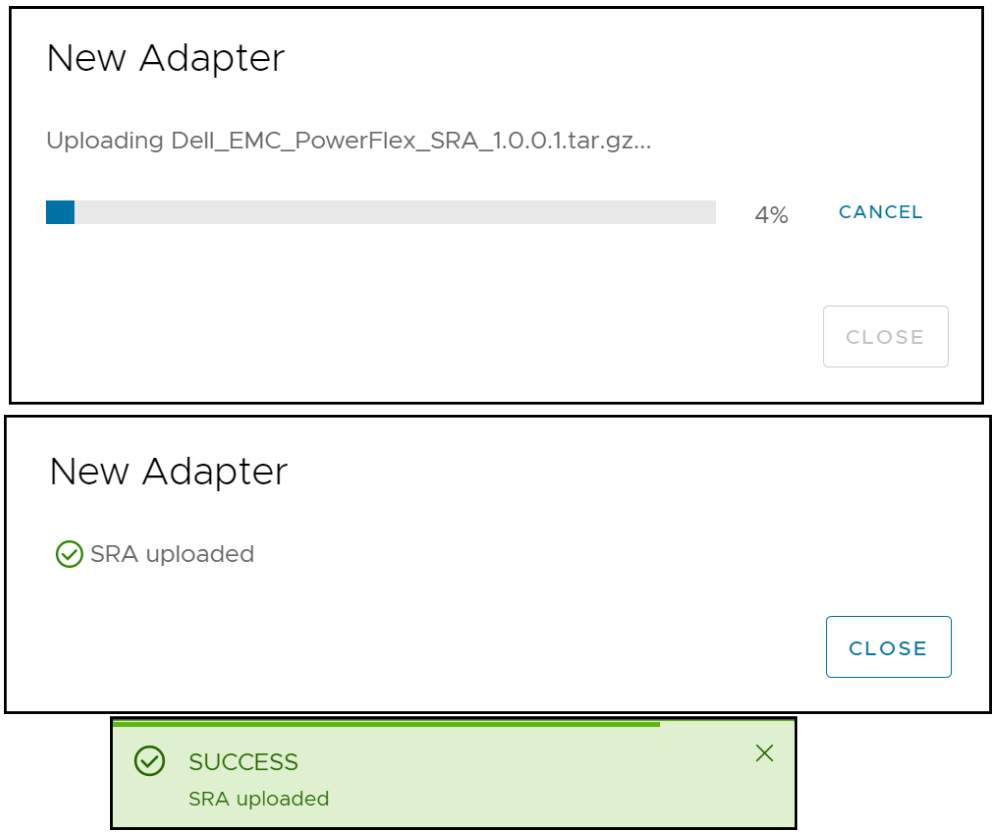

Figure 32. Successful upload

8. Repeat steps 1-7 for the recovery site.

Configuring certificates for Photon OS

The following procedure describes configuring Dell EMC PowerFlex SRA certificates for Photon OS.

Before adding the PowerFlex clusters in SRM, an additional configuration step is required. The root certificate of each PowerFlex cluster must be uploaded to its respective SRA. This certificate can be the same one used when replication was configured between the two PowerFlex clusters, or it can be extracted again. The following steps walk through the process. It must be completed for each SRA.

1. Obtain the **root** certificate, in PEM format, of the PowerFlex cluster. If there is no generated certificate, it can be exported from the primary MDM instance using CLI. Note that the name of the certificate file is user-defined. The login syntax for PowerFlex 4 requires the extra variable, **management_system_ip.**

```
scli --login --username admin --password <password>
[management_system_ip <IP of PowerFlex GUI]

scli --extract_root_ca --certificate_file /tmp/root_ca.pem
```

2. After obtaining the certificates for each cluster, log back into either SRM Appliance Management instance and navigate to **Storage Replication Adapters** again. It does not matter which site, protection or recovery, is modified first.

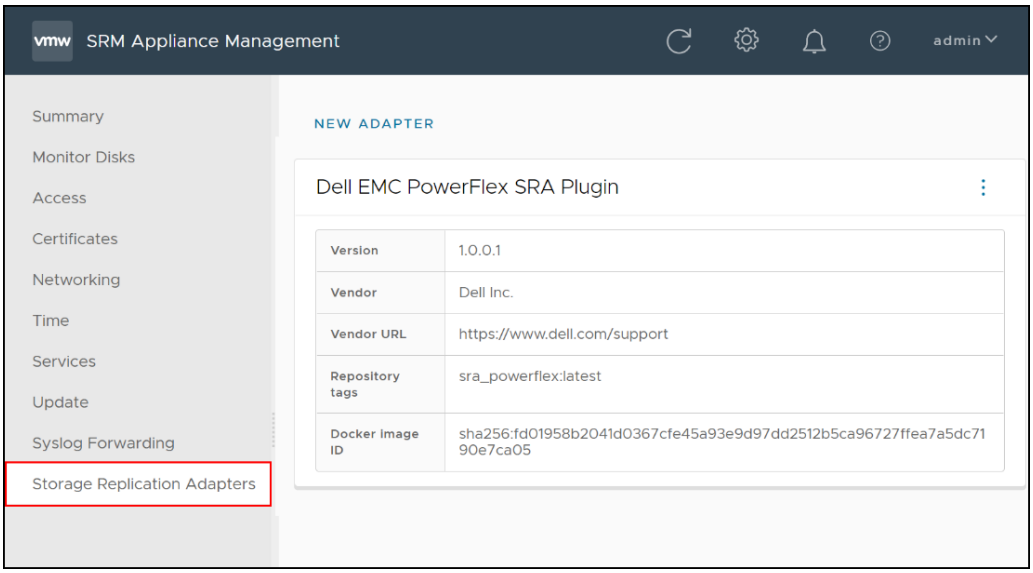

Figure 33. SRA screen in SRM Appliance Management

3. Select the three dots in the corner of the Dell EMC PowerFlex SRA Plugin box as shown in Figure 34 and choose the menu **Download configuration archive**. This will initiate a download of the SRA configuration files in the form of the file **sra_powerflex~latest.tar.gz**.

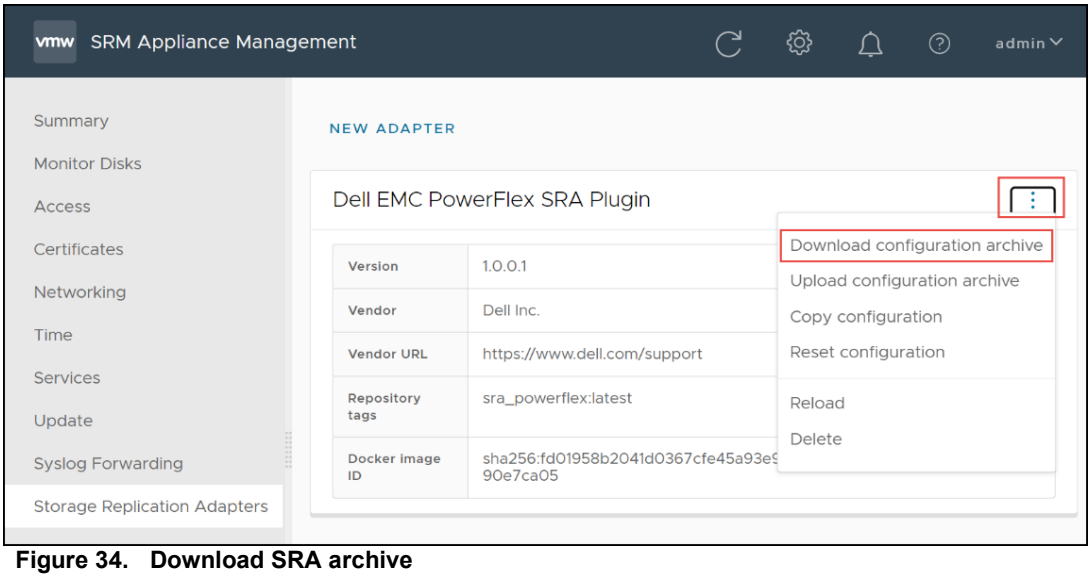

Figure 34. Download SRA archive

4. Using 7-Zip (or another unzip utility) as in this example, extract the archive zipped file and the tar file within it. There will be two files: **config.ini** and **sra-configuration-version.txt** as seen in Figure 35.

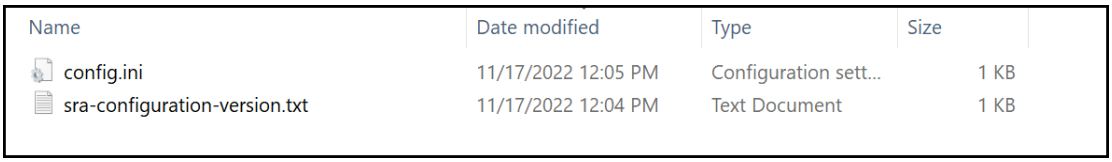

Figure 35. Contents of archive

5. Copy the root certificate of the appropriate PowerFlex cluster into the same directory. In this example shown in Figure 36, the certificate is **root_ca.pem**.

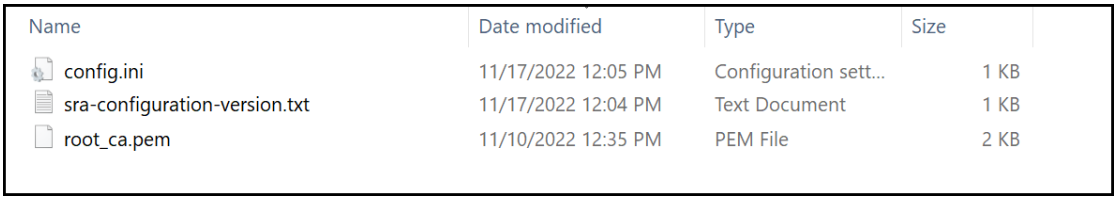

Figure 36. Copy over the root certificate

6. Next, as seen in Figure 37, edit the **config.ini** file with Notepad or a simple text editor. Do not use Microsoft Word to avoid adding extra characters. Once open, look for the variable **ca_path =** under the **[ssl]** section.

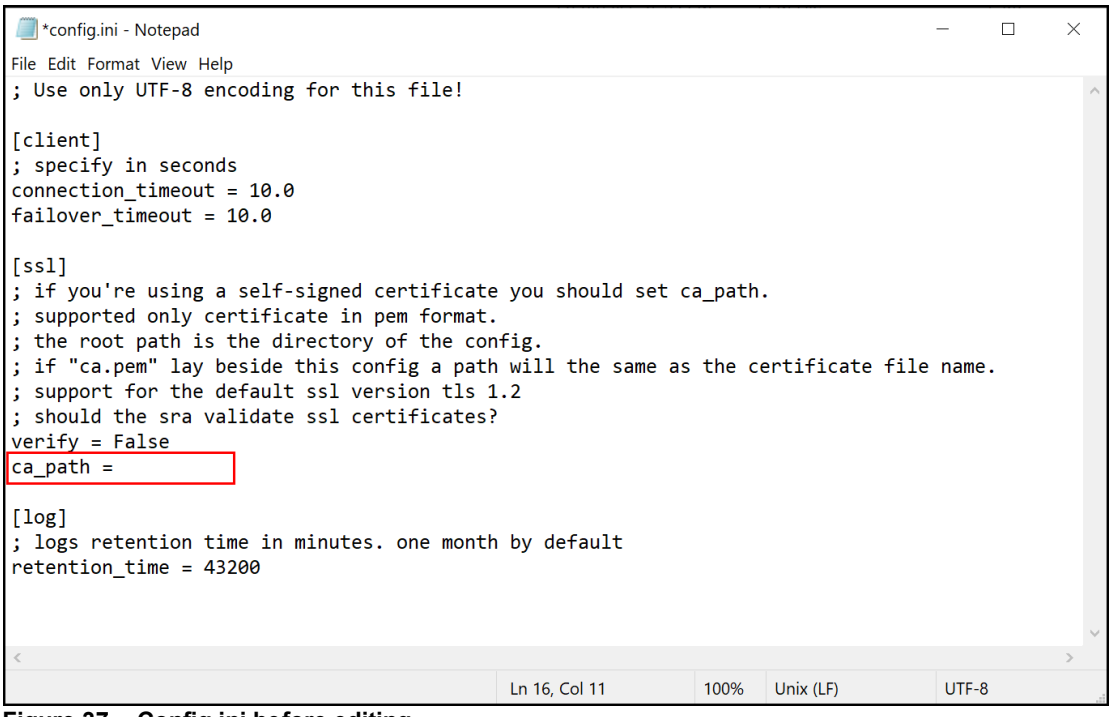

Figure 37. Config.ini before editing

7. Edit the config.ini by typing in the name of the root certificate then save the file as shown in Figure 38. The name must simply match the name given to the certificate. If desired, alter the variable **verify** to True.

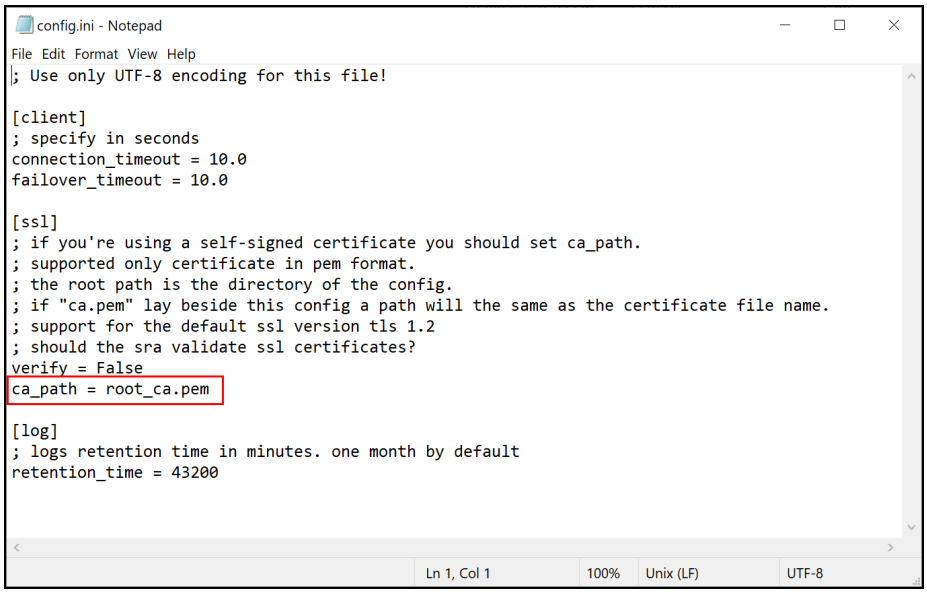

Figure 38. Config.ini after editing

8. The three files can now be re-packed. The following walks through the steps using 7-Zip.

   a. Right-click the files and select 7-Zip -> **Add to archive…** as seen in Figure 39.

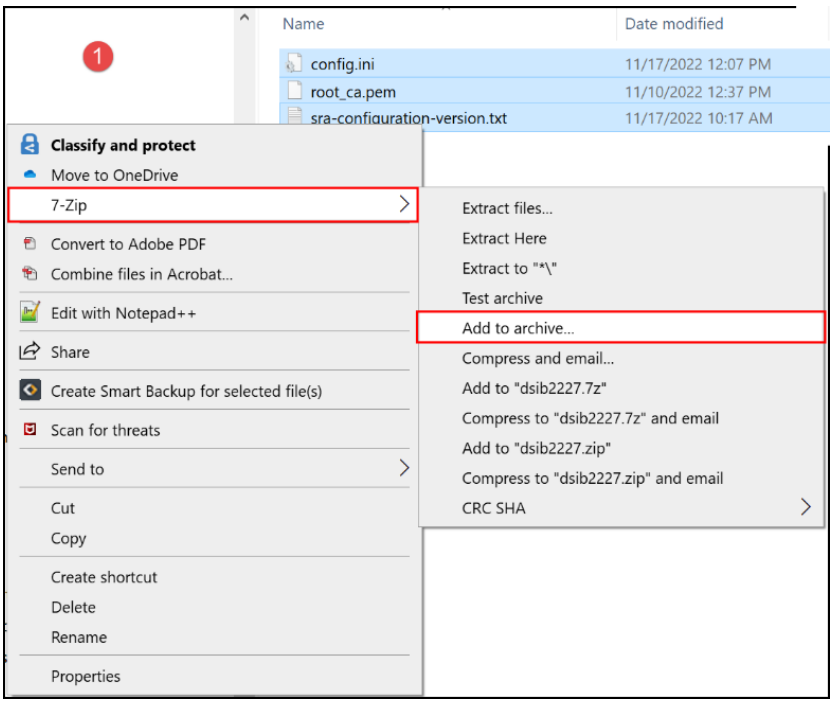

Figure 39. Create compressed archive - Step 1

b. Provide a name for the file with .tar suffix and set the **Archive format** to **tar** as shown in Figure 40.

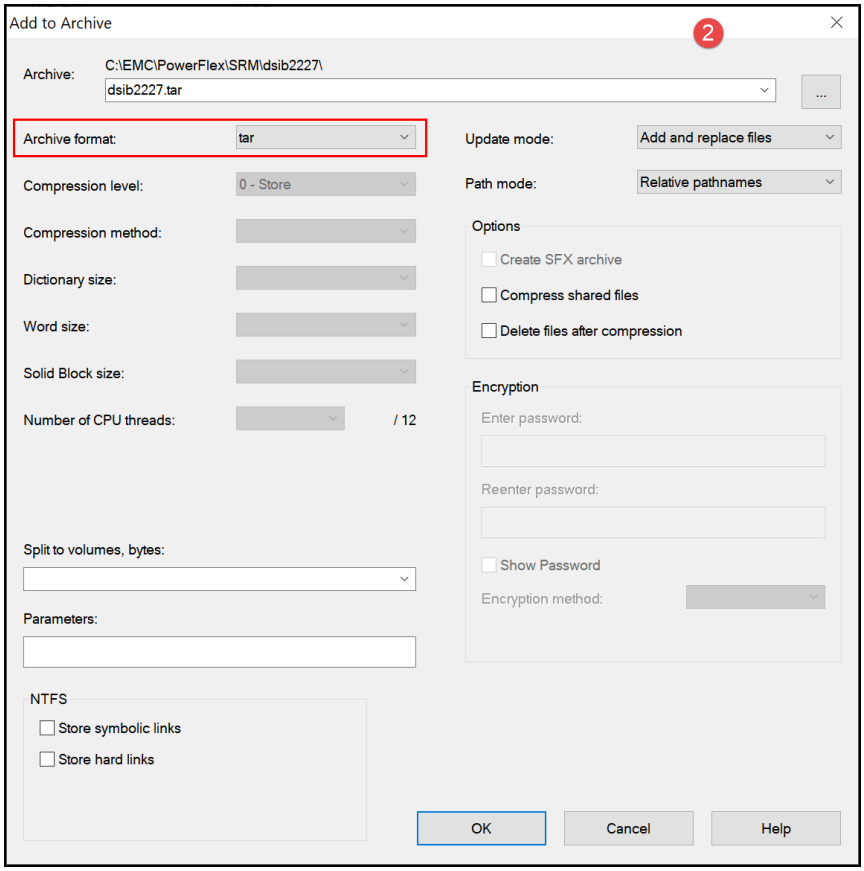

Figure 40. Create a compressed archive - Step 2

c. Next, as seen in Figure 41, right-click on the newly created tar file, in this example dsib2227.tar, and again choose **Add to archive…**

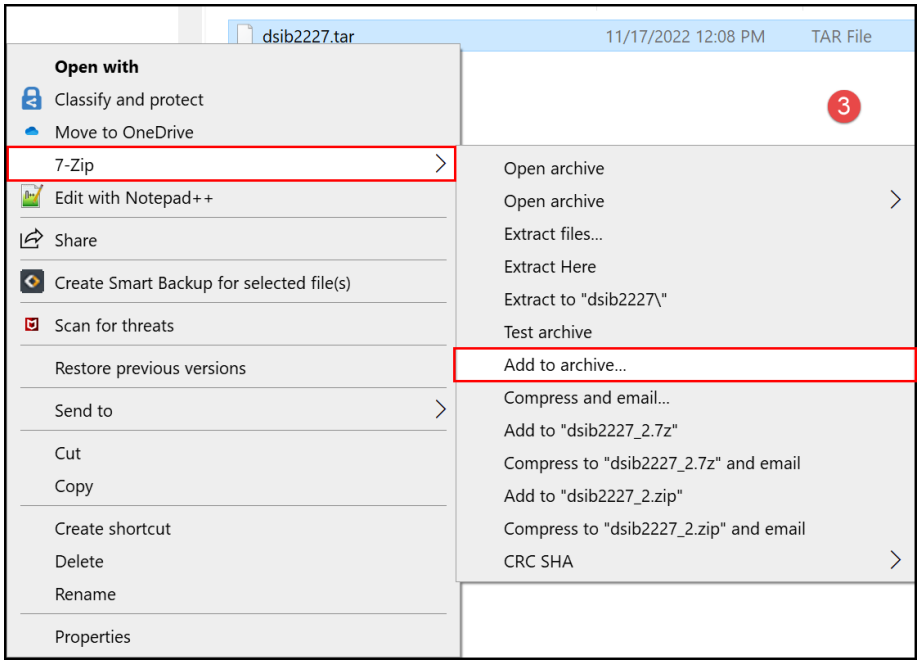

Figure 41. Create a compressed archive - Step 3

d. Provide a name for the file with .tar.gz suffix and set the **Archive format** to **gzip** as shown in Figure 42.

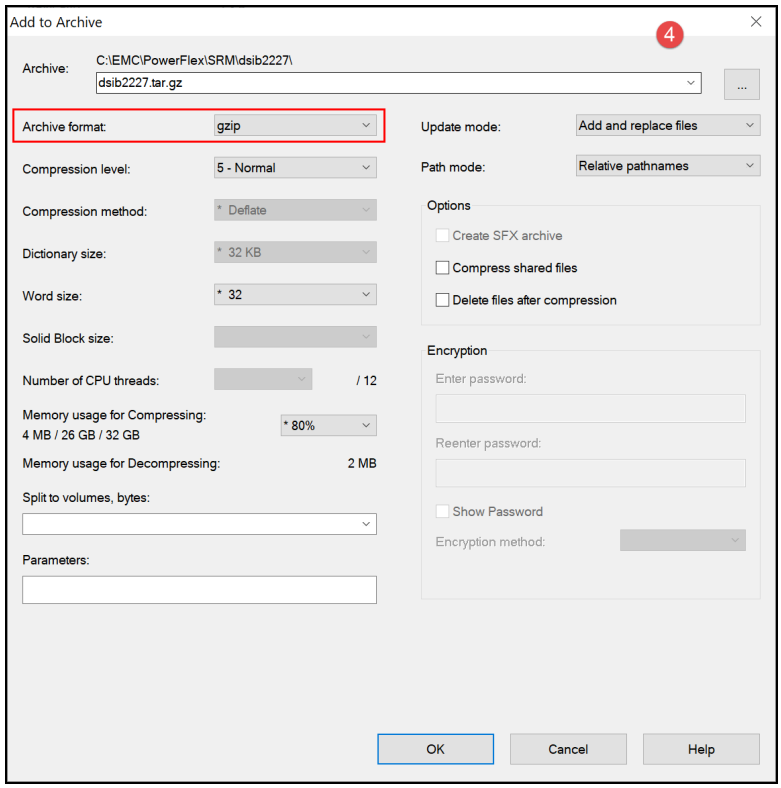

Figure 42. Create a compressed archive - Step 4

9. Now upload the file to the SRA. Using the same process as downloading, select the three buttons and choose **Upload configuration archive** as seen in Figure 43.

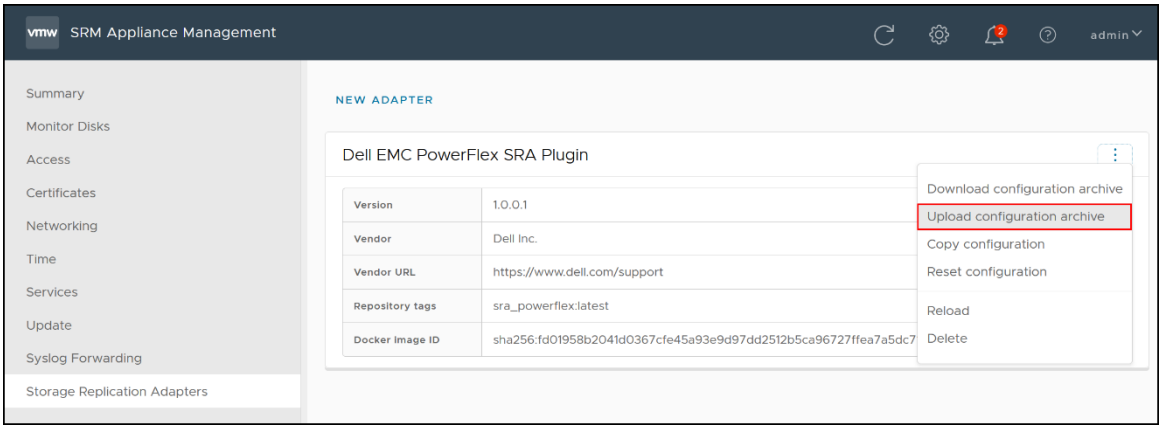

Figure 43. Upload configuration archive menu option

10. Provide the file and select **UPLOAD**. Another green box will appear indicating success as shown in Figure 44.

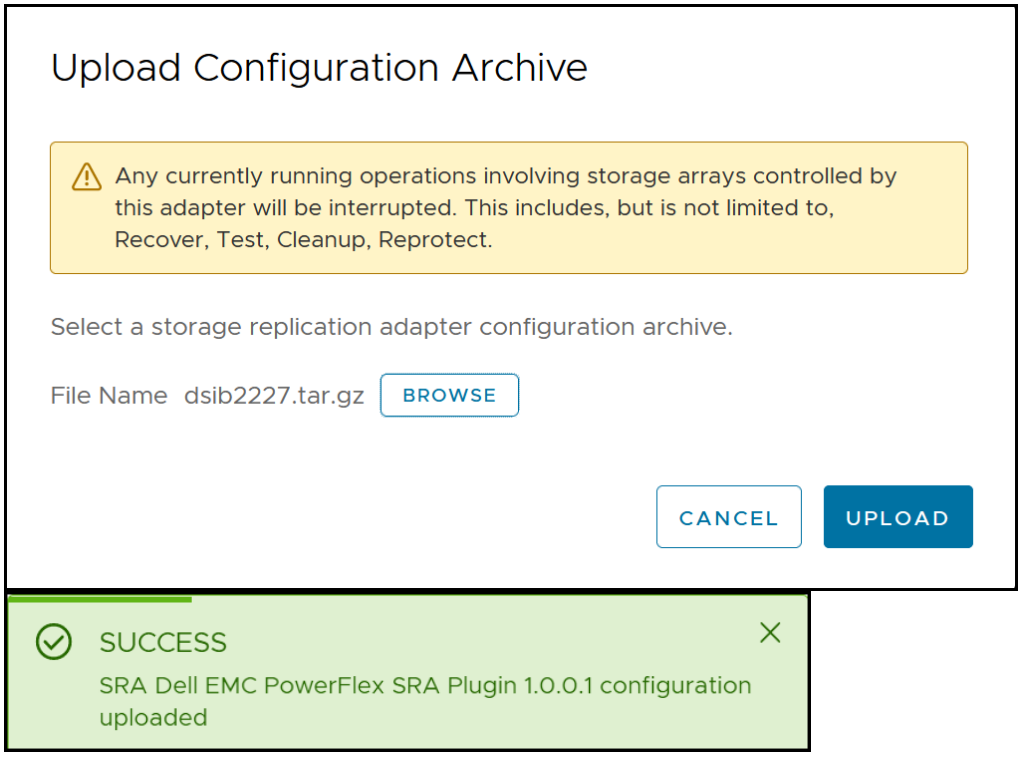

Figure 44. Upload configuration archive

For advanced configurations, see the Dell EMC PowerFlex Storage Replication Adapter (SRA) Plugin for VMware Site Recovery Manager Release Notes.

Configuring the PowerFlex SRA

The configuration of the Dell PowerFlex Storage Replication Adapter is performed through the Array Managers wizard in VMware SRM. The wizard should be invoked only after the connection between the VMware SRM servers at the protected site and recovery site has been established and the PowerFlex replication setup is complete.

Array Manager configuration

Before protection groups or recovery plans can be configured, an array manager must be configured for each instance of the PowerFlex SRA to provide connectivity information for device discovery. These array managers are referred to as local and remote. Note that because SRM supports bi-directional replication, these terms are entirely dependent on where the source volume resides.

As there is a single PowerFlex SRA for both version 3 and 4, the wizard will only be covered once. Any differences between the platforms will be included in the detail. The following steps assume the user is familiar with the SRM interface.

1. Navigate to **Site Pair -> Configure -> Array Pairs** in SRM. Select **ADD** in the right-hand panel as shown in Figure 45.

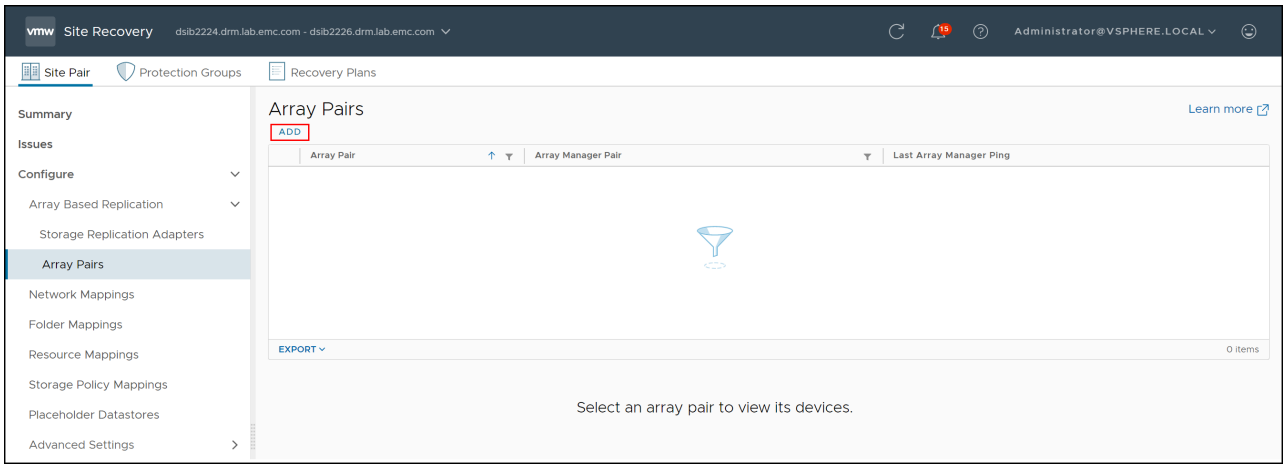

Figure 45. Add array pair – Step 1

2. In the second step shown in Figure 46, select the PowerFlex SRA and then click **NEXT.**

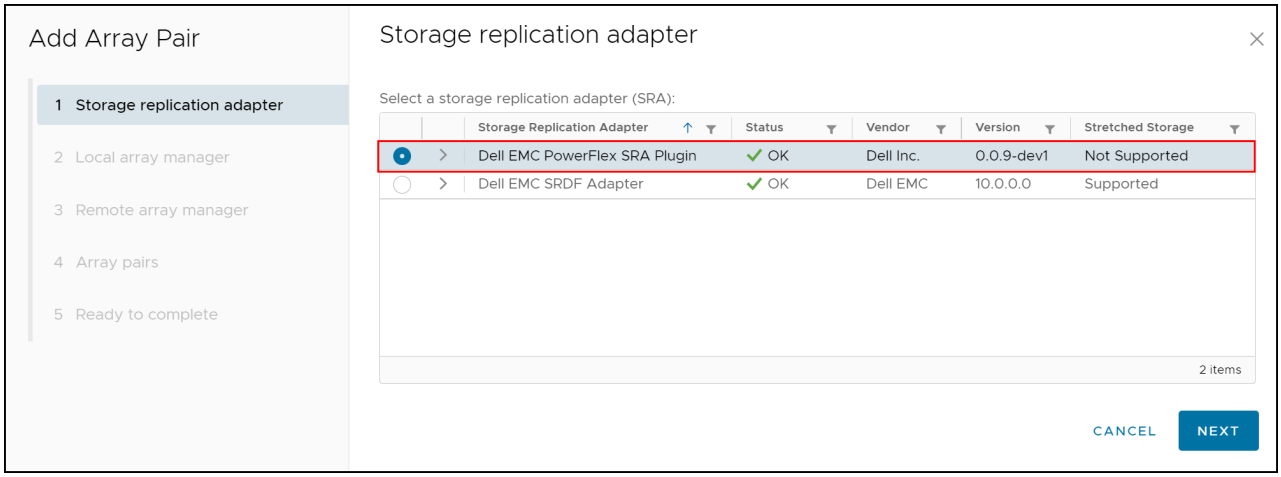

Figure 46. Add array pair – Step 2

3. Next, add the local and remote managers. Provide a name for the local array manager (source PowerFlex), being sure to use an identifying name, enter in the IP of either the PowerFlex Gateway for PowerFlex 3 or the PowerFlex GUI for PowerFlex 4, and the username admin and its password. Select **NEXT** and repeat this process for the remote array manager (target PowerFlex). This is shown in Figure 47.

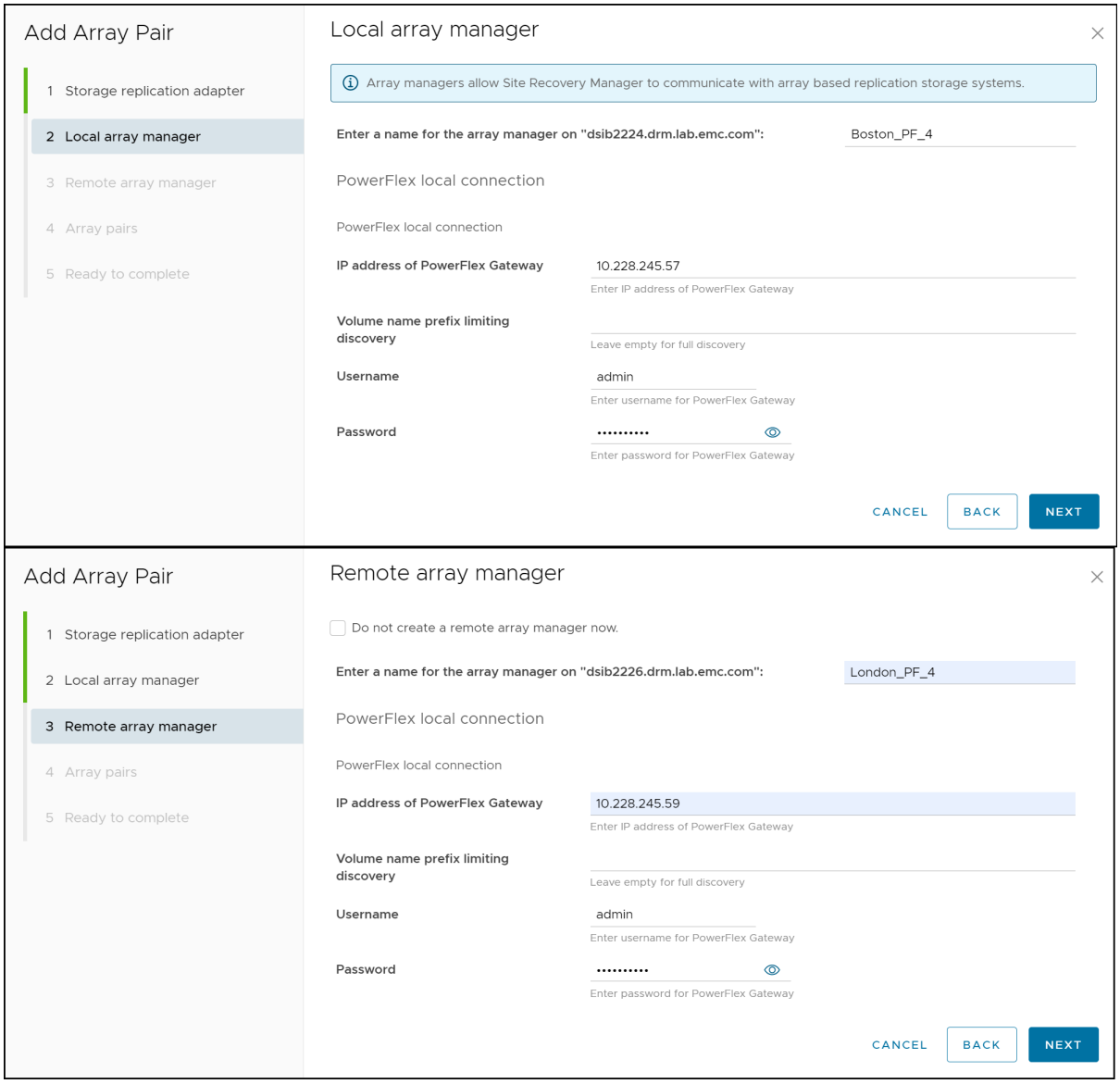

Figure 47. Add array pair – Step 3

4. SRM will discover any peered PowerFlex clusters and display them, as seen in Figure 48. The wizard automatically checks the box for the appropriate pair of arrays (one or more). Once a pair has been successfully enabled for one array manager, that pair will be enabled on the other respective array manager. It is important to note that these array pairs must be enabled (default) in order to discover devices from them.

Note: PowerFlex 3 and 4 array pairs can exist in the same SRM environment as seen in Figure 50.

Figure 48. Add array pair – Step 4

5. Review the final summary and select **FINISH**.

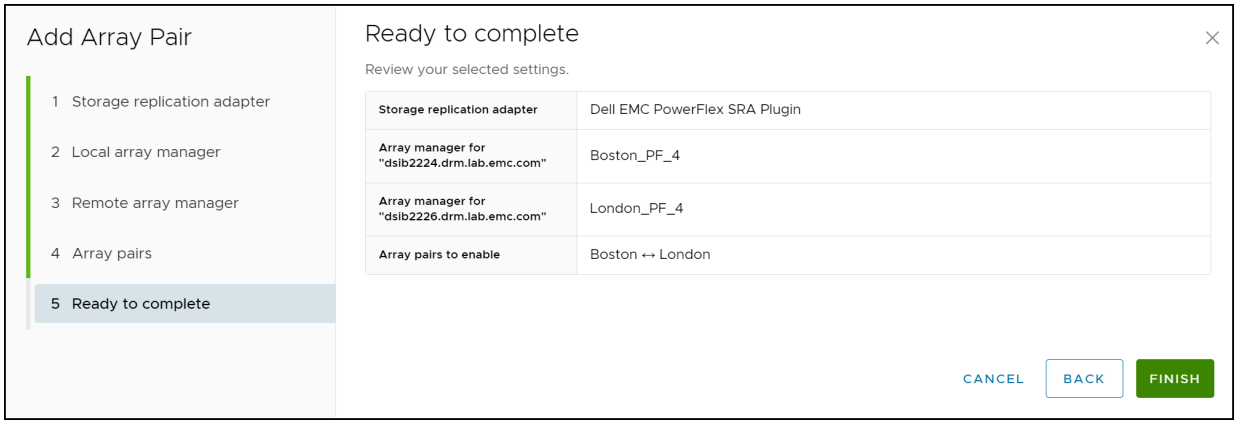

Figure 49. Add array pair – Step 5

6. If no array pairs are displayed, check that the PowerFlex clusters are properly peered. After the pairs are enabled, SRM automatically discovers any devices on the arrays. An example is shown in Figure 50 where both a PowerFlex 3 and 4 are configured.

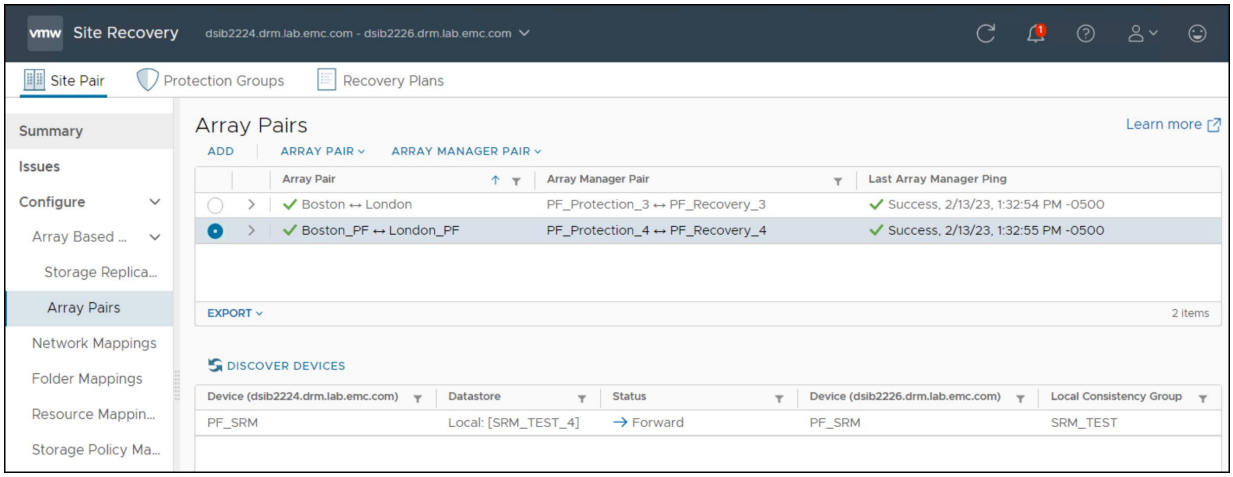

Figure 50. Device discovery – Step 6

When device discovery is complete, the user can create protection groups and recovery plans.

Configuring VMware SRM protection groups and recovery plans

Prior to creating protection groups and recovery plans in SRM, VMware requires that the user create resource mappings between the protection and recovery sites.

Configure mappings

Mappings are used to specify how SRM maps virtual machine resources on the protected site to resources on the recovery site. Site-wide mappings can be configured to map objects in the vCenter Server inventory on the protected site to corresponding objects in the vCenter Server inventory on the recovery site. Through the wizard, SRM can create the reverse objects automatically for the user. The following resources are needed:

- Network Mappings
- Folder Mappings
- Resource Mappings
- Storage Policy Mappings (advanced optional configuration)

When the user runs a recovery plan, VMware has to know in what cluster or folder the VMs should be placed and on what network they should run. Mappings need not have identical names, nor must the resources be equal on both sides. For example, two folders on the protection site can map to the same folder on the recovery site. The next section covers, by example, how one of these mappings is done.

Resource mapping

1. Navigate to **Site Pair -> Configure -> Resource Mappings**. Select **NEW** in the left-hand corner as shown in Figure 51, noting which site is highlighted above, protection or recovery.

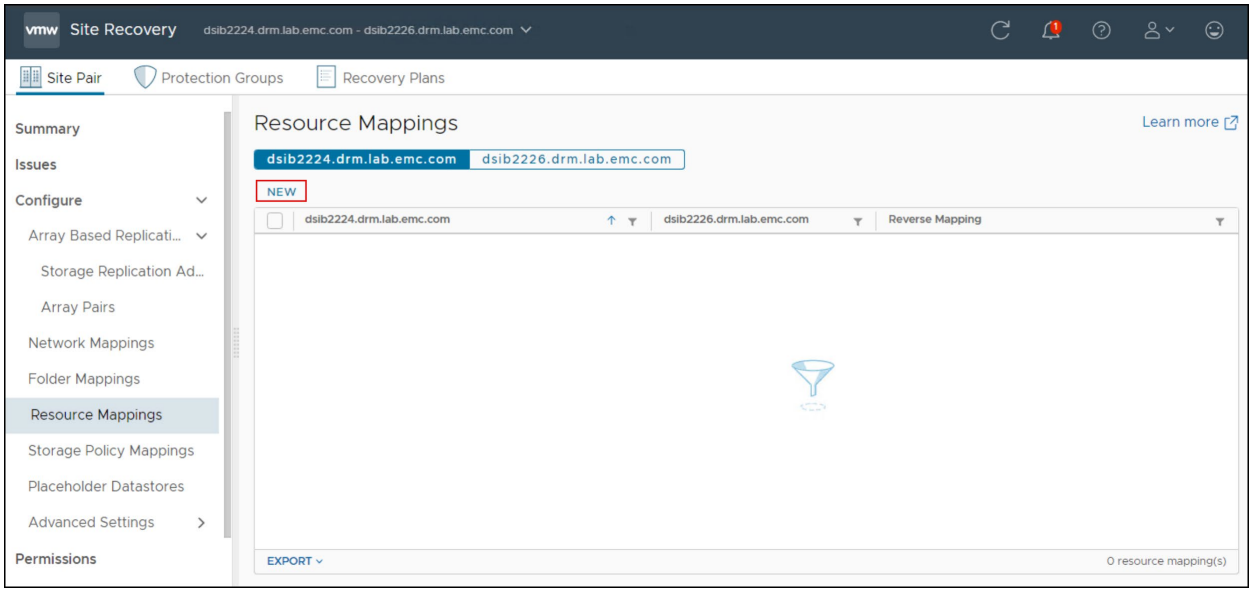

Figure 51. Resource mappings wizard – Step 1

2. In the next step, shown in Figure 52, use the checkbox in the left-hand panel to choose a resource, and then similarly on the right-hand panel use the radio button to select the matching resource. In this example the cluster **Boston_Cluster** is chosen on the protection site and **London_Cluster** is selected on the recovery site.

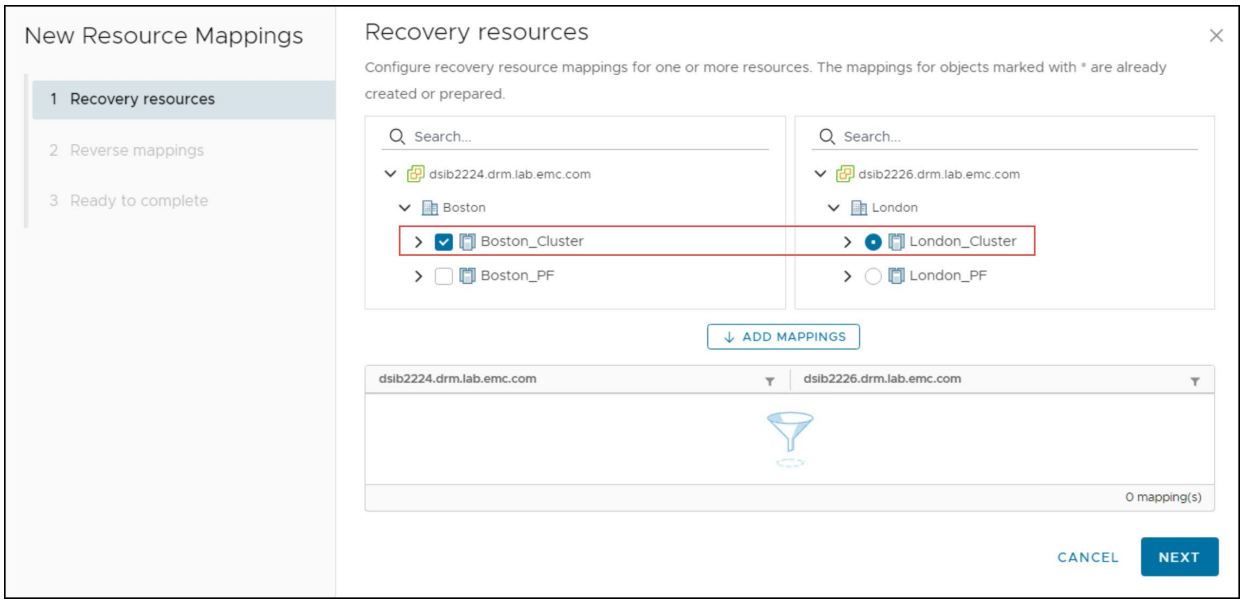

Figure 52. Resource mappings wizard – Step 2

3. Click on **ADD MAPPINGS** to place the mapping in the bottom panel. Repeat the process for the **Boston_PF** resource and the **London_PF** resource. Figure 53 shows the end result of the step. These new mappings ensure that any VM in the **Boston_Cluster** will be placed in the **London_Cluster** when a recovery plan is run. The same holds true for the other mapping. Click **NEXT** to continue.

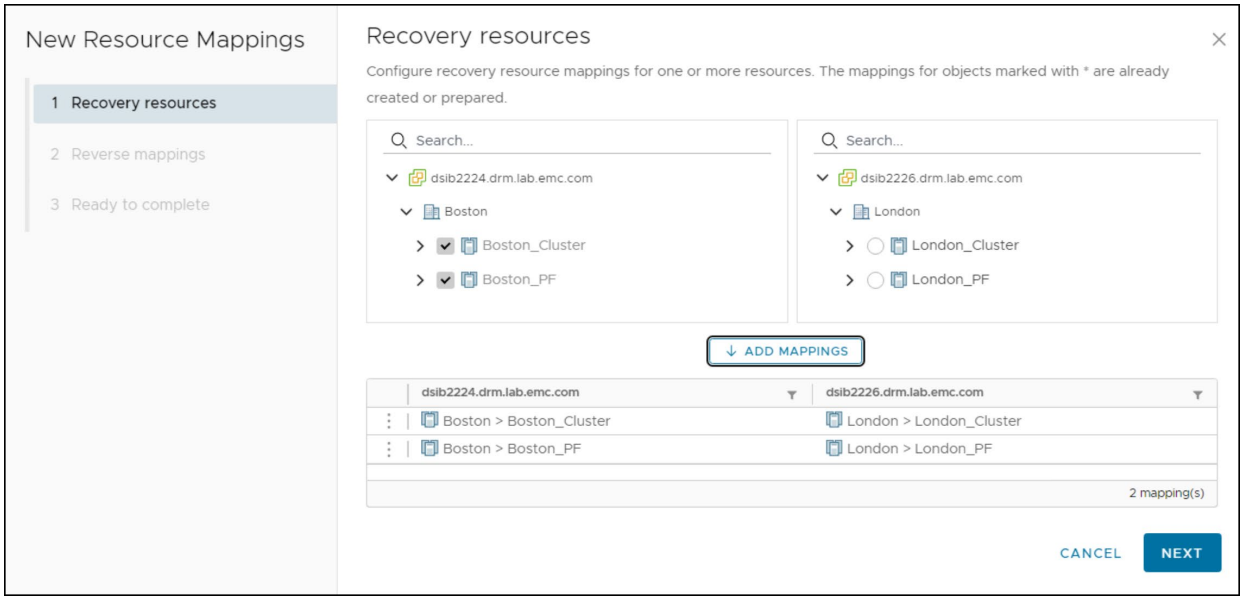

Figure 53. Resource mappings wizard – Step 3

4. Once the mappings are set, VMware offers the ability to automatically create reverse mappings. In other words, it alleviates the need to repeat the mapping process on the recovery site. If this is desired, check the boxes shown in Figure 54 to create these mappings. Note that reverse mappings are mirror images of the original mappings. So, in this case if a recovery plan is run in the opposite direction, any VMs that were in the **London_Cluster** will be placed in the **Boston_Cluster**. If there is a need to have a different mapping, do not check the boxes and instead repeat this mapping process on the recovery site.

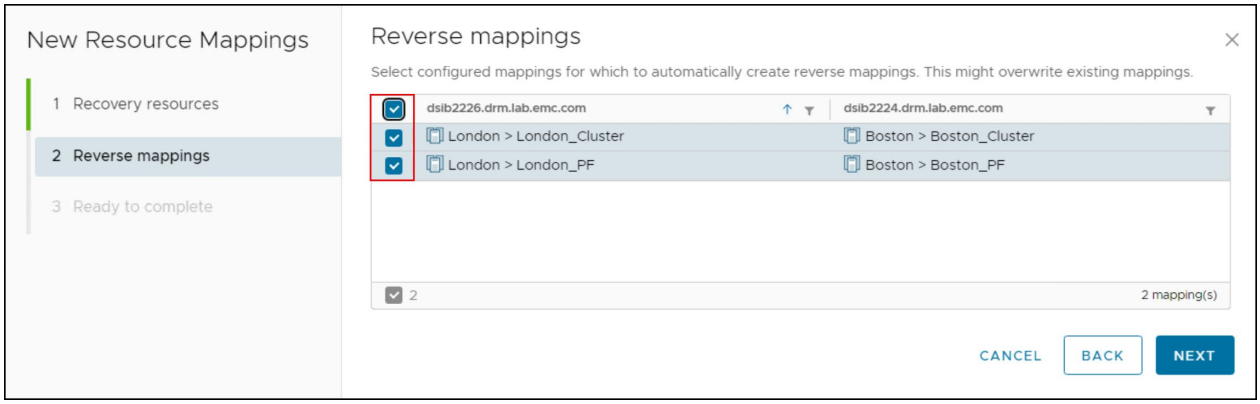

Figure 54. Resource mappings wizard – Step 4

5. The final mappings should be reviewed in the summary shown in Figure 55 before selecting **FINISH**. VMware always provides the ability to go **BACK** and fix mistakes in the wizard.

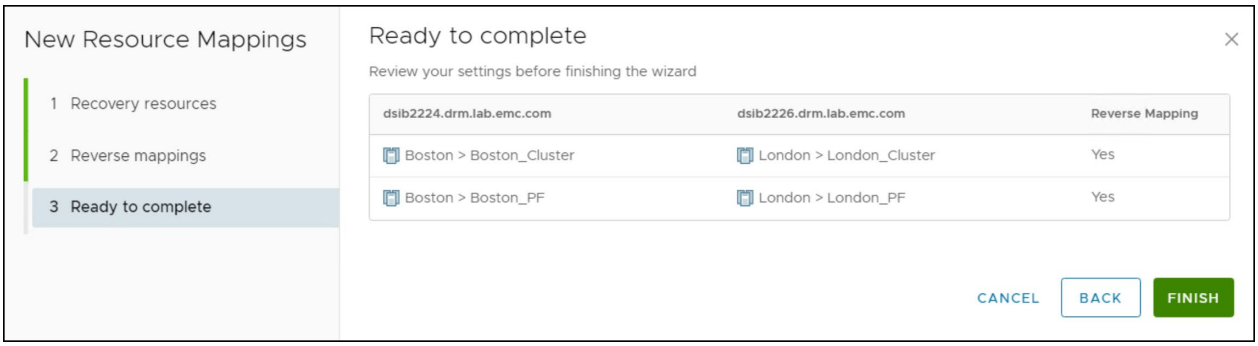

Figure 55. Resource mappings wizard – Step 5

6. When complete, SRM displays the resource mappings for the protection site. Note the column **Reverse Mapping** shown in Figure 56 indicates if the recovery site also has the mappings.

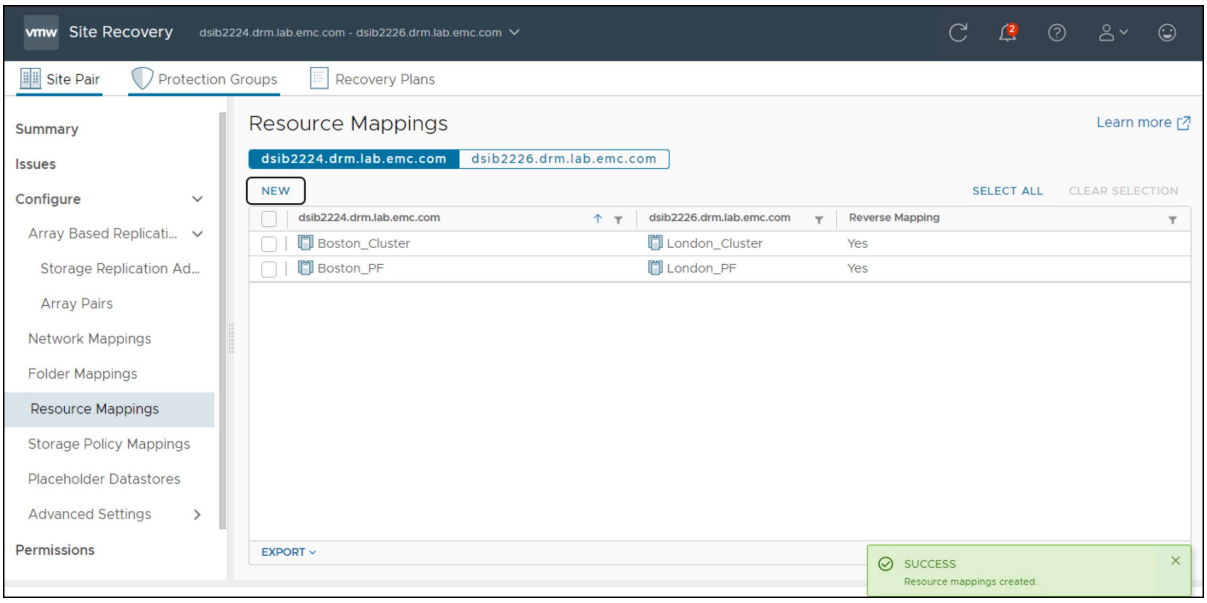

Figure 56. Resource mappings – Step 6

Placeholder datastores

In addition to the mappings, SRM requires assigning a placeholder datastore at each site. The datastore holds the configuration files for the placeholder, or “shadow”, virtual machines at the recovery site. VMware recommends using a non-replicated datastore as these placeholder VMs can be recreated by SRM if they are lost or accidentally removed. Typically, only one placeholder datastore per site is required, though more may be configured if desired. Dell recommends configuring two for redundancy. The placeholder datastore need only be large enough to hold the configuration files for all the recoverable

virtual machines. It is not a significant amount of storage, though it is perfectly acceptable to use large, existing datastores that have VMs in them. Also note that the placeholder datastores do not have to be backed by PowerFlex volumes, they only need to be available to the SRM ESXi hosts.

Protection groups

Protection groups are object that enable SRM to protect virtual machines. The groups can be independent or organized into folders, though the group names should be unique across the folders.

For array-based replication, there are two types of protection groups:

1. Datastore groups
2. Virtual Volumes

As previously noted, PowerFlex does not support Virtual Volumes with SRM, so datastore groups must be used.

A datastore group is a logical grouping of devices and is the smallest unit of storage that can be failed over or tested independently with PowerFlex. There are several rules that control how SRM calculates datastore groups:

- If a replicated device is used by a datastore containing virtual machines, all devices used by that datastore are combined.
- If a device is a part of a consistency group, all devices in the consistency group are combined.
- If a virtual machine spans multiple datastores, all devices belonging to all such datastores are combined.

Protection groups can include one or more datastore groups and are one type of building block of recovery plans. Protection groups include the datastores groups which are to be failed over simultaneously. As such, failover is absolute—either the entire protection group fails over or none of it.

Recovery plan

A recovery plan provides the automation of the protection group. It is comprised of a set of steps that SRM follows for the VMs in the protection group. These steps include the order in which SRM powers on and powers off virtual machines, the network IPs that those VMs utilize, and the running of any customized scripts the user added. During the running of a recovery plan, there are steps specific to the execution of storage activities. At these steps is when SRM communicates with the PowerFlex SRA to perform actions and then awaits a response from the array.

A recovery plan includes one or more protection groups. It is possible to have multiple recovery plans with the same protection group but remember that a VM may only be in a single protection group. If a protection group is in multiple recovery plans, no two recovery plans with that protection group can be run concurrently. Once one of these recovery plans begins, the other recovery plans change the state of the protection group to **Protection Group In Use** and they are disabled from execution.

Steps to create a protection group and recovery plan

1. Protection groups can be created by selecting the **NEW PROTECTION GROUP** hyperlink in the Protection Groups section highlighted in **Error! Reference source not found.**. This process can be initiated from either SRM server, though it may be necessary to change the **Direction** in the wizard.

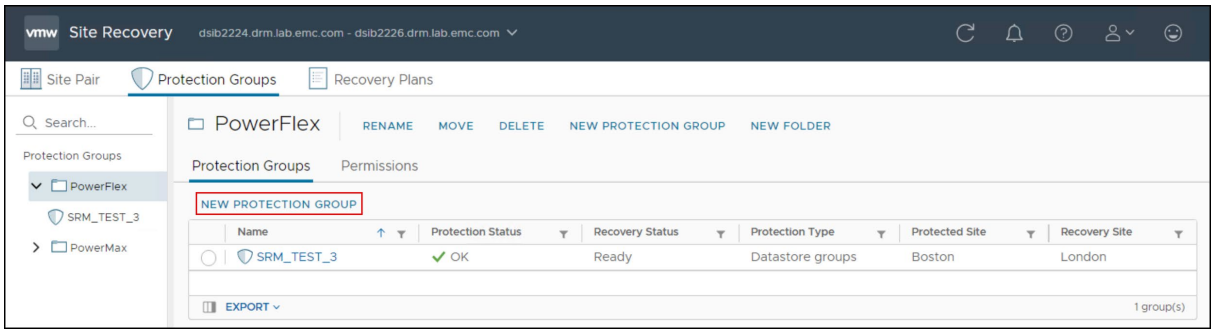

Figure 57. New Protection Group wizard – Step 1

2. The **Create Protection Group** wizard will appear. Begin by providing a name, description (optional), and the location. In the example shown in Figure 58, folders are used to distinguish between array types.

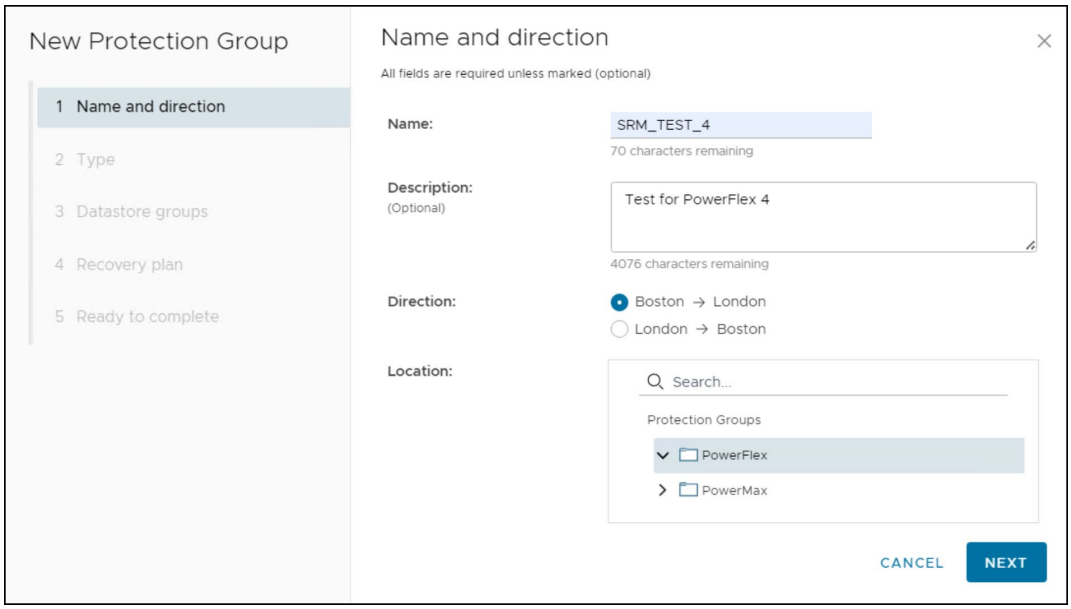

Figure 58. New Protection Group wizard – Step 2

3. In step 3, select the **Type** as **Datastore groups** (default) and choose an array pair. This screen emphasizes the need to name the array pairs properly so they can be distinguished from one another. The array pair shown in Figure 59 is a PowerFlex 4 based on naming.

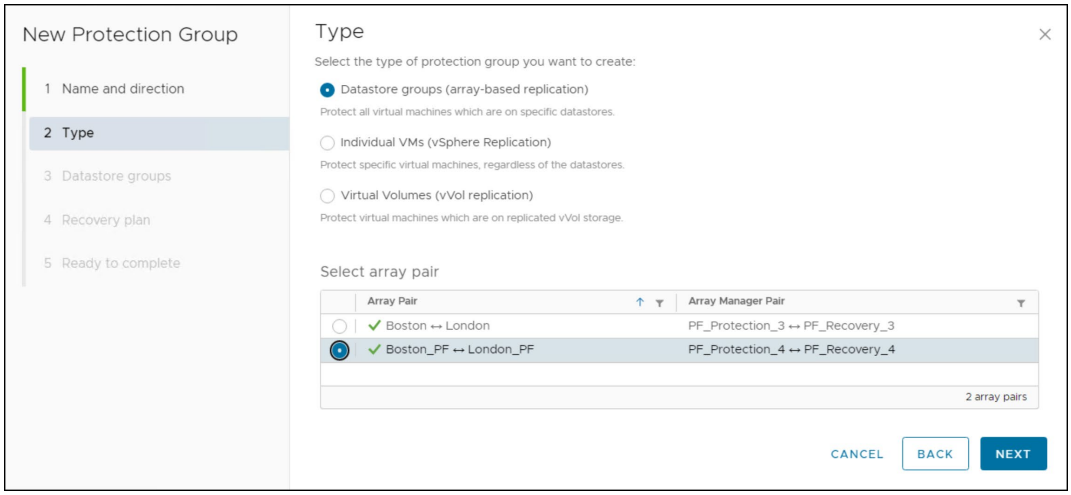

Figure 59. New Protection Group wizard – Step 3

4. The available datastore groups for protection are display as shown in Figure 60. As each datastore group can only be protected by one group, SRM will not display those already in a group.

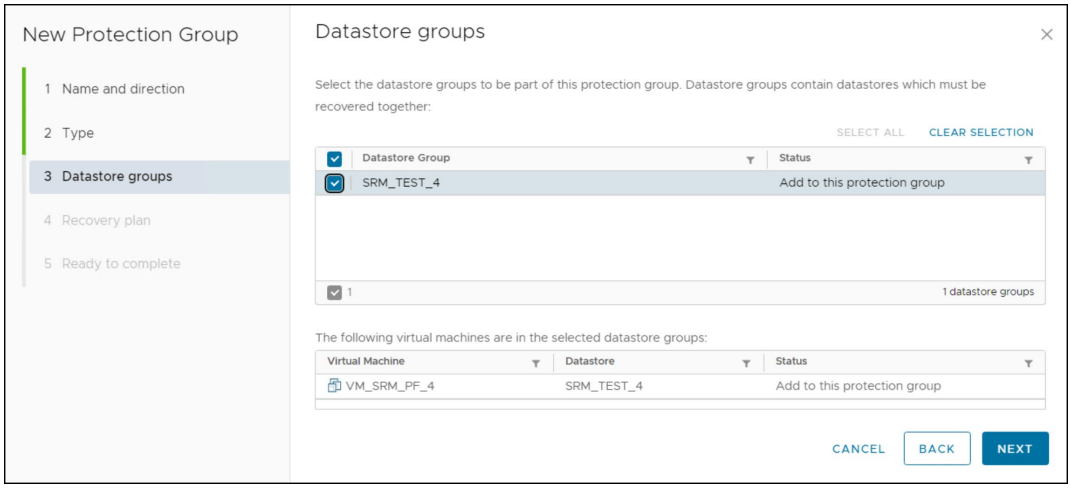

Figure 60. New Protection Group wizard – Step 4

5. The user can add the protection group to an existing recovery plan, create a new one, or choose to skip it. In this step shown in Figure 61, a new recovery plan is named.

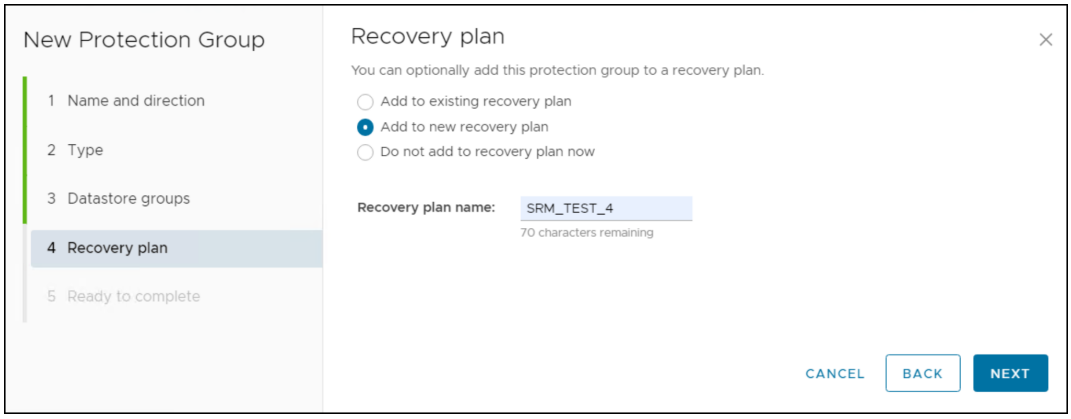

Figure 61. New Protection Group wizard – Step 5

6. With the recovery plan named, review the summary shown in Figure 62 and select **FINISH**.

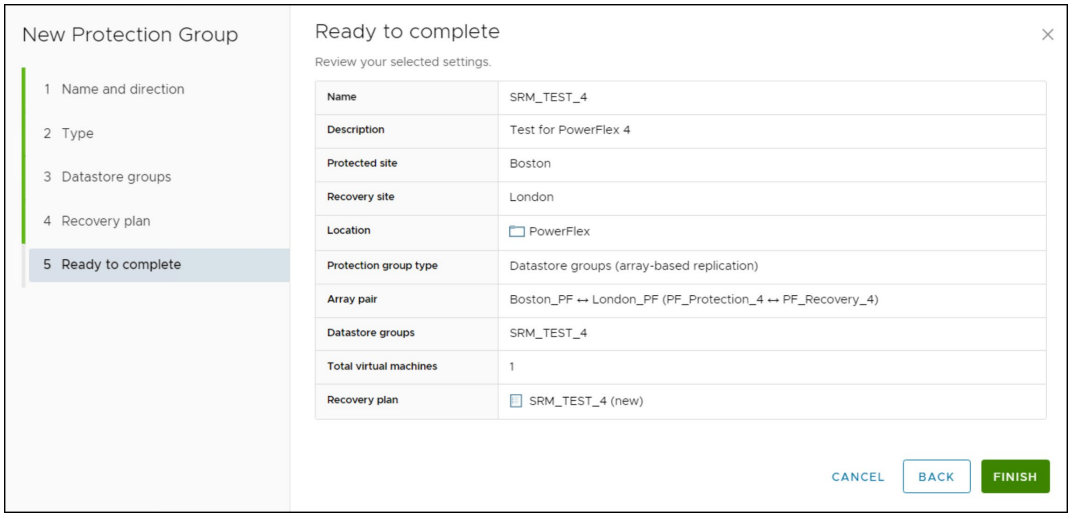

Figure 62. New Protection Group wizard – Step 6

It should be noted that while the process of protection group and recovery plan creation can be initiated from either SRM server, datastore groups will only appear for the SRM server that has access to the source devices in that group. Therefore, if no datastore groups appear in the create protection group wizard, it is likely that the incorrect SRM server was previously selected in the wizard.

Advanced settings

Under normal operations, most environments do not require additional changes to use the PowerFlex SRA with VMware SRM. Occasionally due to environment size or network latency, adjustments might be needed. The following explains some of these adjustments.

Logging

The PowerFlex SRA does not offer any additional logging parameters, but one can adjust SRM logging for any storage provider such as the PowerFlex SRA that is installed in the appliance. The following section describes these changes.

Modifying logging levels with SRM

With VMware SRM the logging level is altered from the GUI and does not require a restart of the SRM process.

To alter the logging level of all SRAs in the environment, in SRM navigate to **Site Pair -> Configure -> Advanced Settings -> Log Manager** as shown on the left-hand side in Figure 63 . Highlight the **logManager.StorageProvider** option and select **EDIT** button in the top right-hand corner. Modify the level as required on each SRM site.

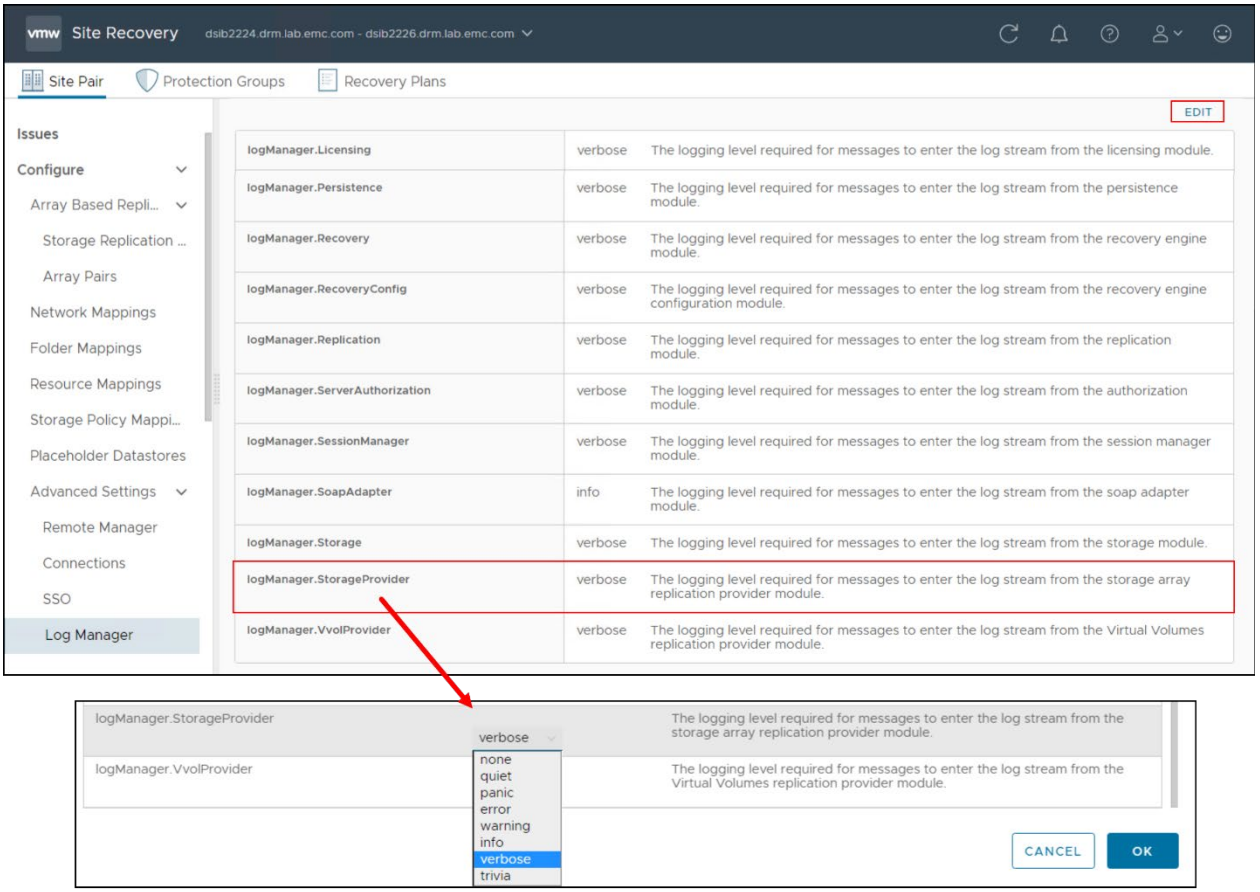

Figure 63. Storage provider logging in SRM

The logs for SRM are located on the appliance in the following directory: `/opt/vmware/support/logs/srm`.

Export logs

In general, rather than retrieving logs from the appliance, it is easier to generate a support bundle from within SRM that includes all the necessary log files for debugging. Figure 64

shows the location in the SRM Appliance Management screen for log generation. The bundle can be downloaded locally when complete.

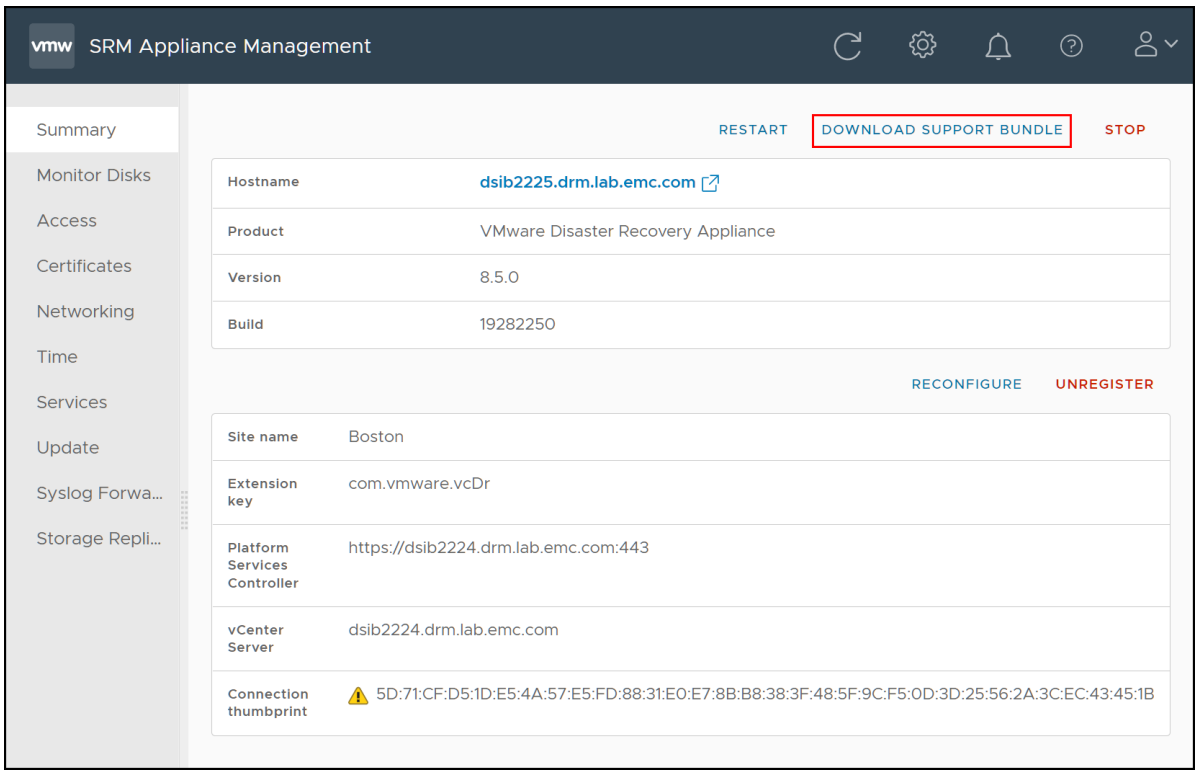

Figure 64. Export logs from SRM

Configuring advanced VMware SRM options

VMware SRM offers a wide variety of advanced options and settings that can be altered by the SRM administrator if required, or if requested by Dell or VMware support. The majority of these options are well beyond the scope of this paper, but this describes a few key settings involved in the storage configuration and behavior of SRM.

If these settings are changed from the default it is recommended to make the change on both the protected and recovery SRM servers to maintain configuration consistency between the two environments. Changes of an option on one SRM server does not propagate to the other SRM server the user must ensure it is done. Below are descriptions and recommendations for a list of six of the most common storage related advanced SRM settings.

1. **storage.commandTimeout**—This option controls the amount of time SRM allows the SRA to run an operation. In small to medium environments the default of 300 seconds is sufficient but in larger or more complex environments, or in situations where the SRA must wait for data to synchronize, increasing this setting may be advisable. It is recommended to not adjust this setting unless it is deemed necessary. This option affects both recovery and test recovery operations.

2. **storage.storagePingInterval**—This option controls the time interval between SRM-initiated automatic SRA rediscovery of the storage arrays and replication information. The default setting is to rediscover every 24 hours (configured in seconds in SRM). In general, it is not recommended to change this setting as it is important that SRM and the SRA update regularly to make sure that the environment is in a valid state.
3. **storageProvider.fixRecoveredDatastoreNames**—During recovery or test recovery, a prefix is added to datastore names when the ESXi servers re-signature and mount recovered datastores. This option is true or false. When enabled, SRM will remove the snap-xxxxxx prefixes that are automatically added to recovered datastore names. SRM will remove these prefixes after mounting the volume but before registering any virtual machines. By default, this setting is disabled and datastores will keep the prefixes unless manually removed. For customers that run scripts based on datastore name, Dell recommends enabling the setting to avoid errors. This option affects both recovery and test recovery operations.
4. **storageProvider.hostRescanDelaySec**—This option controls the amount of time between when the PowerFlex SRA returns success to SRM after a failover and when SRM initiates a storage rescan on the ESXi host(s). If the storage devices are not fully available yet, ESXi does not detect them and SRM does not find the replicated devices when it performs rescans. This would prevent the mounting of datastores and registering of the virtual machines. If an increase to this delay from the default of zero is necessary, a value of between 20 to 180 seconds is reasonable. Testing will generally surface the need to adjust this setting. This option affects both recovery and test recovery operations.
5. **storageProvider.autoResignatureMode**—This option controls the process of resignaturing during recovery of datastores. This option affects both recovery and test recovery operations. There are three different modes for this option:
   - **Disable**—Will use a selective resignature API to query and resignature only the snapshot VMFS volumes relevant to SRM. This is the default and recommended mode. SRM checks before every HBA rescan (during every test and failover operation) to make sure that the LVM.EnableResignature option on all ESXi hosts is disabled, and if it is not, SRM disables it and leaves it as such. This mode is selected by setting the option value to 0.
   - **Enable**—Turns on the LVM/EnableResignature flag in the respective ESXi hosts to enable automatic resignature of ALL unresolved VMFS volumes. SRM checks before every HBA rescan (during every test and failover operation) to make sure that the LVM.EnableResignature option on all ESXi hosts is enabled, and if it is not, SRM enables it and leaves it as such. No snapshot is left out regardless of its relationship to SRM. Dell strongly discourages the use of this mode for the reason noted in the below warning. This mode is selected by setting the option value to 1.
   - **Ignore**—Will not change the LVM/EnableResignature flag on the respective ESXi hosts. Whatever the user has pre-configured in ESXi for LVM/EnableResignature, SRM will use it. SRM checks if volumes are auto

resignatured, if not it will query and selectively resignature them. This mode is selected by setting the option value to 2.

Note: Setting the LVM.enableResignature flag on ESXi hosts is a host-wide operation and, if set, all snapshot LUNs that can be resignatured are resignatured during the subsequent host rescan. If snapshot volumes unrelated to SRM are currently forcefully mounted to ESXi hosts on the recovery site, these LUNs are resignatured as part of a host rescan. Accordingly, all the virtual machines on these volumes will become inaccessible and will require re-registration. To prevent outages, ensure that no forcefully mounted snapshot LUNs are visible to ESXi hosts on the recovery site or set auto resignature mode to disabled. For more information see VMware KB article 2010051.

6. **storageProvider.autoDetachLUNsWithDuplicateVolume**—When multiple unmounted, unresolved snapshot volumes are found, SRM will use its API to detach these unwanted and unused snapshot LUNs. SRM will then resignature the volume(s) that are to be recovered and then re-attach the previously detached LUNs. This setting is only applicable when auto resignature is either set to disable or ignore. The default for this option is enabled.

   This option is targeted for environments that have multi-extent datastores with multiple copies of them presented to the recovery environment. Resignaturing in these situations can be complex and the functionality introduced by enabling this option eliminates the possibility of errors in regard to the resignaturing process. Dell's recommendation is to leave this option enabled. It is important to note that SRM will only detach detected snapshots of the devices to be recovered in the recovery plan.

   Unmounted, unresolved devices unrelated to the recovery plan (not in the recovery plan or snapshots of devices in it) will not be detached/re-attached. This option affects both recovery and test recovery operations.

Figure 65 below shows the advanced settings options in SRM. Note the settings are per site.

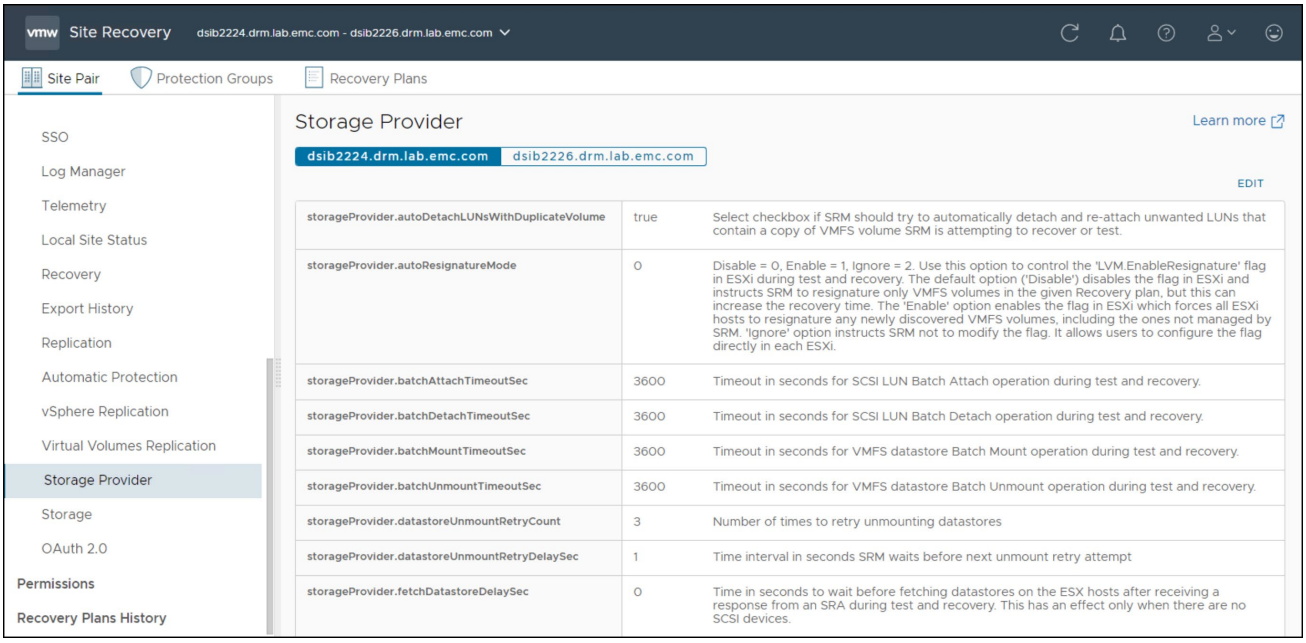

Figure 65. SRM Advanced Settings in the vSphere Client

Test Failover

SRM provides users the ability to test their recovery plans in an isolated environment. Testing recovery plans allows customers to test their disaster recovery ability without impacting the production applications and replication. The testing of the recovery plans is achieved by leveraging local replication technology provided by the storage arrays, in this case the Test Failover functionality that is inherent in PowerFlex replication.

The PowerFlex takes an on-demand snapshot during a Test Failover and point the existing mapped volume at the target to that snapshot, enabling read/write access. Any changes made to those snapshot volumes will be discarded when the test is complete.

This chapter describes the following topics related to test failover:

- Test failover SRM workflow
- Configuring test failover

Test failover workflow in VMware SRM

Test failover is a two-part process:

1. **Test**—Initiates test failover, mounts snapshot devices and powers on (if configured) virtual machines according to the selected recovery plan.
2. **Cleanup**—Powers off and removes the test environment from recovery site, unmounts and deletes and/or removes devices (if configured) and returns status to the original pre-test state.

This section covers the general workflow of the test failover operation.

Requirements

Before a test failover is run the following requirements must be met:

- A recovery plan must be configured and associated with one or more protection groups.
- Inventory mappings should be configured so that the virtual machines can be properly recovered.
- A placeholder datastore where the configuration files for the recovered VMs are stored – shadow VMs.
- A PowerFlex RCG exists that contains at least one device pair in a "consistent" state. The target volume(s) must be presented to the recovery site ESXi hosts

Once all required configurations are complete, the user can perform a test failover.

Test

1. From the **Recovery Plans** screen shown in Figure 66, highlight a recovery plan in the left-hand panel and select the **TEST** link highlighted in the red box.

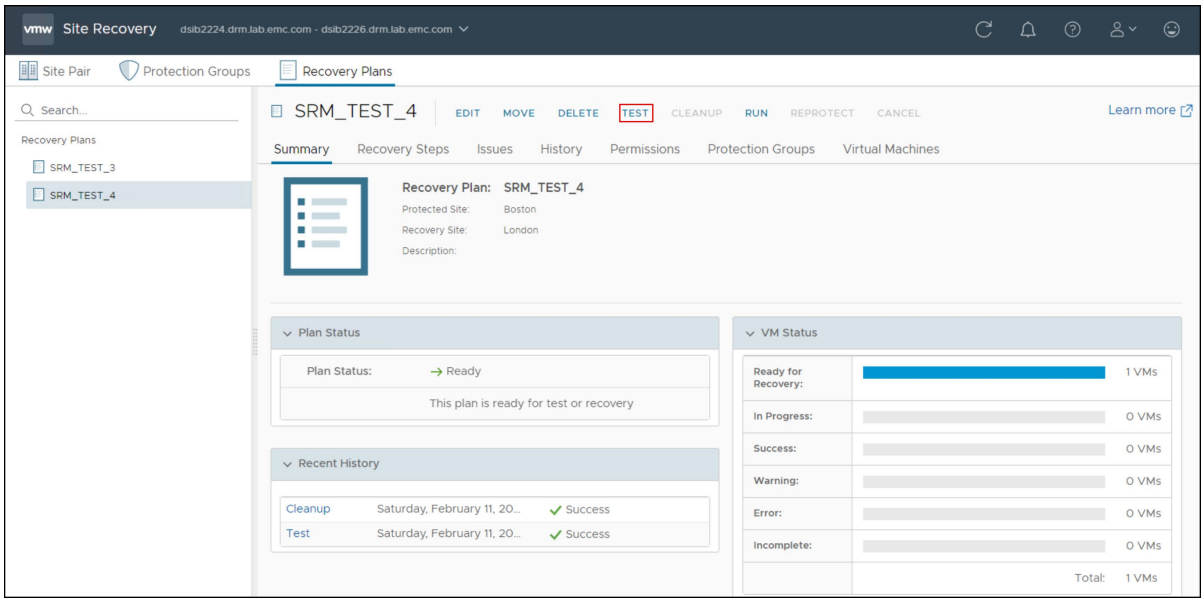

Figure 66. Initiating a recovery plan test failover operation in SRM – Step 1

2. The first screen requires the user to confirm the test operation. As part of this confirmation, there is a checkbox offered which is named **Replicate Recent changes to recovery site**. This option, which enables or disables the SyncOnce operation, is present for all storage vendors. Dell recommends leaving the box checked as show in Figure 67.

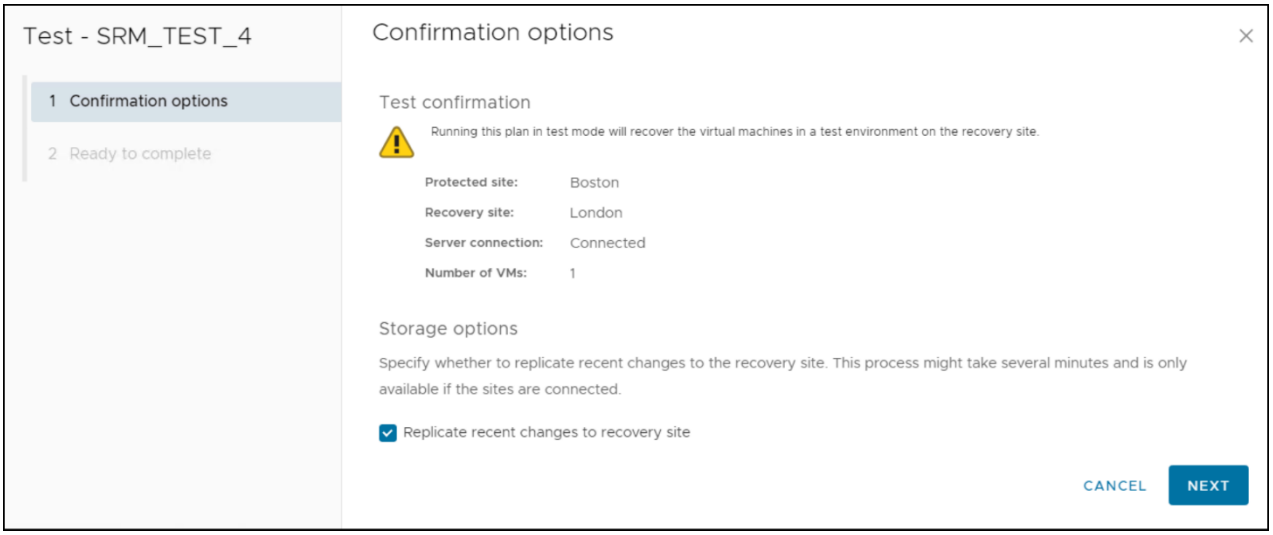

Figure 67. Initiating a recovery plan test failover operation in SRM – Step 2

3. In the final dialog shown in Figure 68, complete the wizard by selecting **FINISH**.

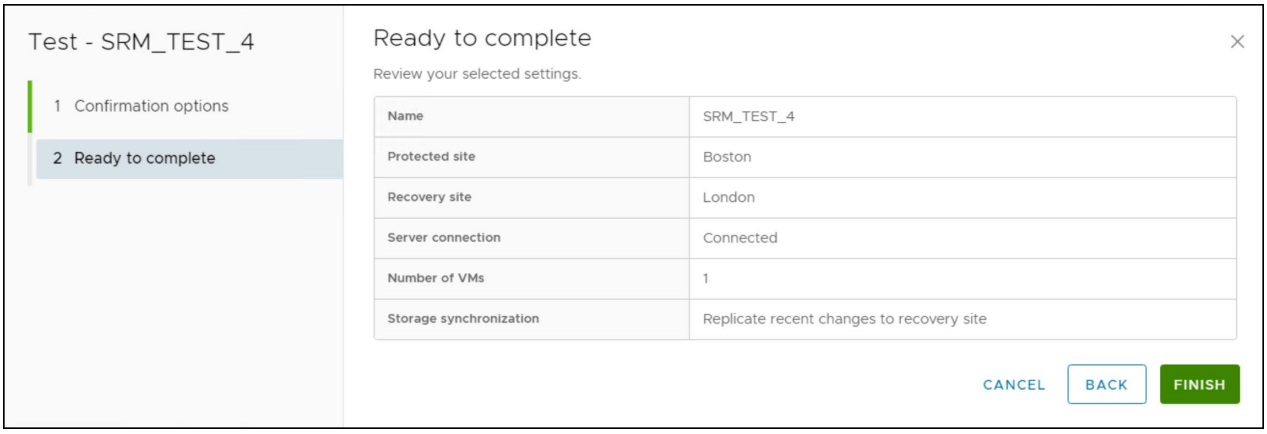

Figure 68. Initiating a recovery plan test failover operation in SRM – Step 3

At a high level, a recovery plan test involves the following:

- Creation of a bubble IP network for the virtual machines to run in, so that the production networks are not interfered with. If a special network is not created beforehand by the user, SRM creates one automatically. Be aware that the problem with the SRM-created networks is that they do not allow virtual machines on different hosts to communicate as these networks are internal to each host and do not have external up-links. If inter-host communication is required, the user should create insulated networks across ESXi hosts and configure it for use in the recovery plan.
- A temporary snapshot is created and the target device is redirected to it (pointer). The target device is then made read/write to the ESXi hosts.
- The device is resignatured and the datastore mounted at the recovery site.

- Shadow VMs are unregistered and are replaced with the copied VMs.
- If the plan requests suspension of local virtual machines at the recovery site, they are suspended during the test.
- VMs are powered-on or left off according to the recovery plan.
- Finally, it is important to remember that no operations are disrupted at the protected site.

Once the user has confirmed the test failover operation can proceed, the recovery plan will be initiated in test mode.

A completed test recovery can be seen in Figure 69. The test environment will remain operational until a **CLEANUP** operation is run.

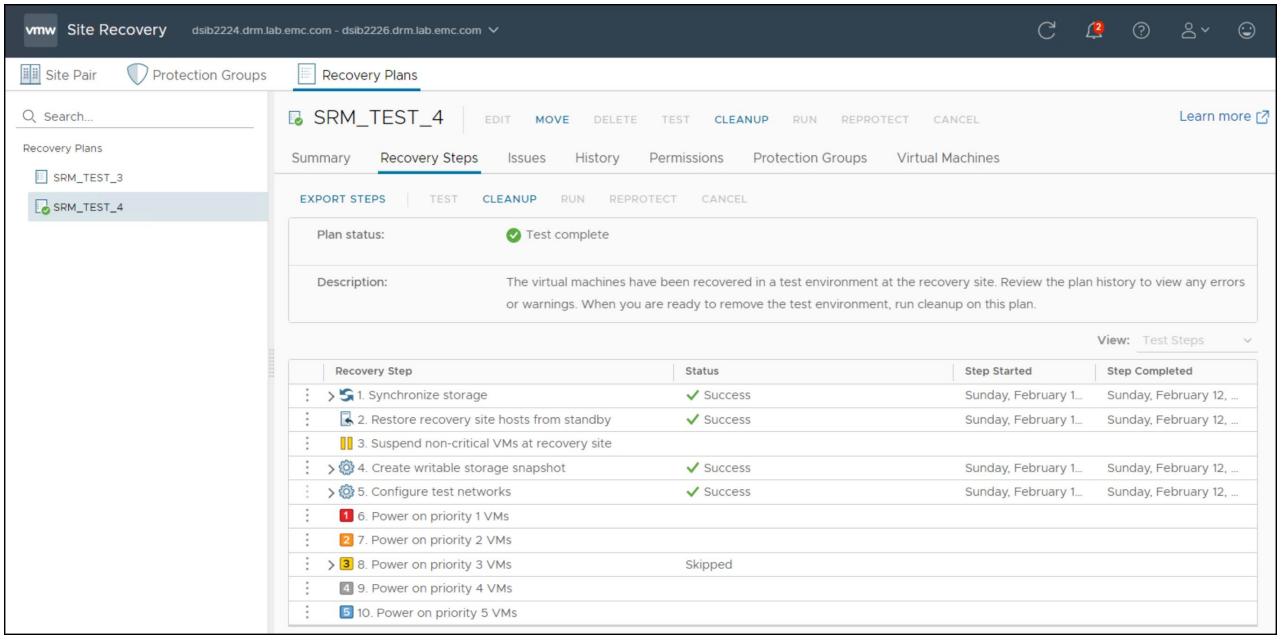

Figure 69. Completed test recovery plan in SRM

PowerFlex snapshot

The temporary snapshot, created by PowerFlex for the test failover, is shown in Figure 70. This snapshot is mounted to the target volume by a pointer redirection and used in the test. Note the **Access Mode** is set to **Read and Write** for testing purposes.

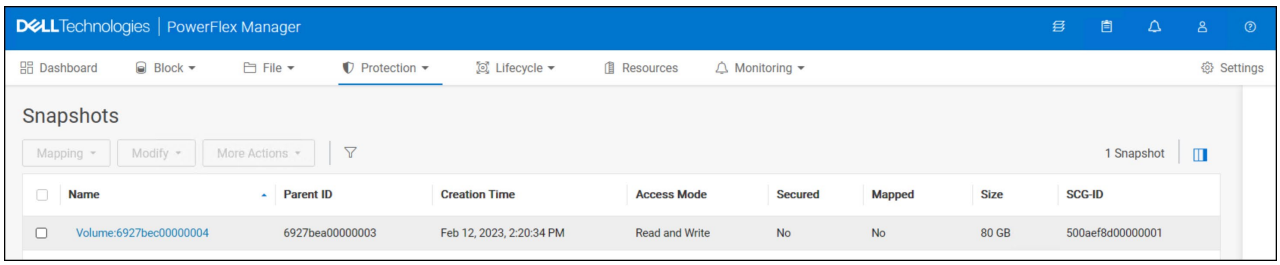

Figure 70. PowerFlex snapshot for SRM test failover

The RCG status, by necessity, is now in a **State** of **Failover Test** as shown in Figure 71. It will remain in this state until a **CLEANUP** operation is run.

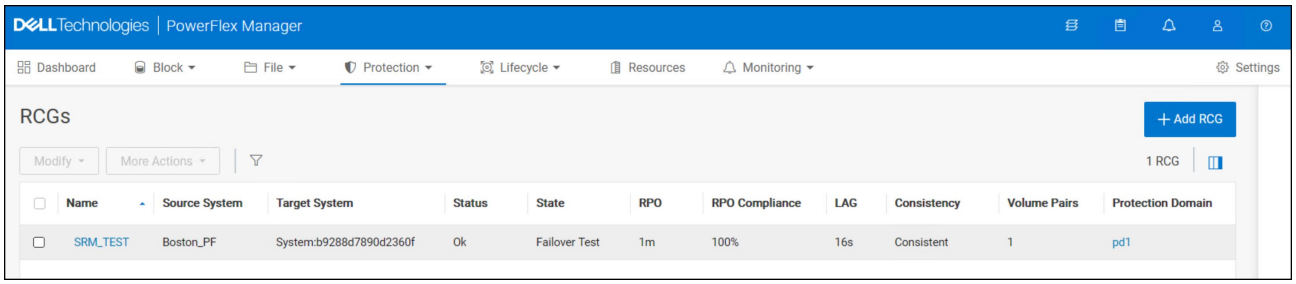

Figure 71. PowerFlex Failover Test

Note: It is important to remember that during a test failover, replication is paused between the source and target, though the source journals continue to receive updates. But if a failure were to befall the protection site during a test, the most recent data would be when the test started. Dell recommends, therefore, that if extensive, time-consuming testing is required, a manual snapshot is used with independent volumes, rather than using the test failover functionality.

Test Failover in PowerFlex GUI

Because SRM is using the Test Failover functionality in PowerFlex, it is important to understand the interaction between the two softwares. The following section describes this interaction.

PowerFlex offers the ability to run a Test Failover from within the GUI interface, as seen in Figure 72, as well as CLI.

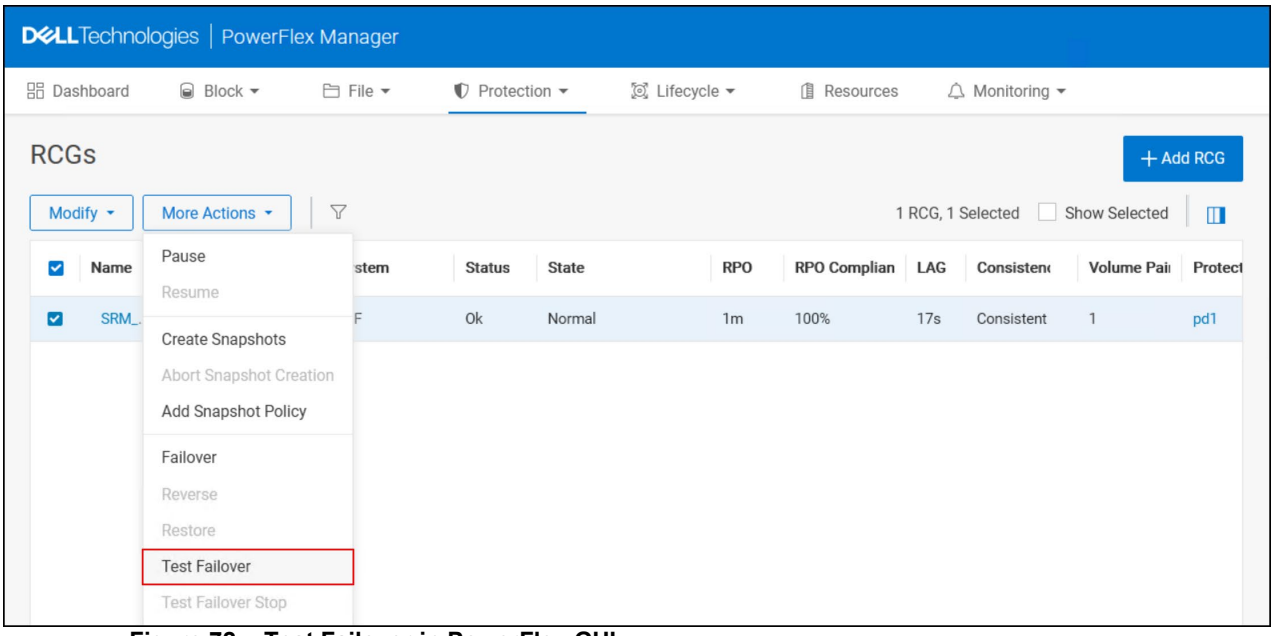

Figure 72. Test Failover in PowerFlex GUI

If Test Failover is started from the PowerFlex interface, instead of SRM, VMware has no knowledge of the activity. SRM jobs may fail to run properly due to an unknown PowerFlex state. Therefore, it is imperative not to mix PowerFlex and SRM activities. If SRM is being used, the PowerFlex GUI should not handle any test or actual failover. From a manual perspective, the two softwares are mutual exclusive.

Cleanup

Once a recovery plan has been tested, the test environment can be discarded and reset through the use of the cleanup operation offered by SRM. The cleanup operation automatically reverts all changes incurred by the recovery plan test and allows for subsequent failover operations.

The cleanup operation performs the following operations:

- Power off and unregister test virtual machines
- Unmount and detach snapshot VMFS volumes or RDMs
- Replace recovered virtual machines with original placeholders (shadow VMs), preserving their identity and configuration information.
- PowerFlex storage snapshots that were used by the recovered virtual machines during the test are removed, and thus any changes made during the test itself are discarded.

Before resetting the environment after a test failover, ensure that the recovery plan worked as desired. Verify the success of any custom scripts, application functionality, networking, etc. Once all facets of the test have been verified by the involved stakeholders, a cleanup operation can be run.

Note: After a test failover has been run, an actual failover or another test failover cannot be run until a cleanup operation has occurred. It is advisable to run a cleanup operation as soon as the test environment is no longer needed to allow for any subsequent operations to be run without delay. In addition, the replication updates on the source journals cannot be sent over until the cleanup completes.

As shown in Figure 73, a cleanup can only be run against a recovery plan if the recovery plan status is in **Test complete**. Otherwise, the **CLEANUP** button is grayed out. Furthermore, even if a test failover was not entirely successful, a cleanup operation will still need to be run before another test failover can be attempted.

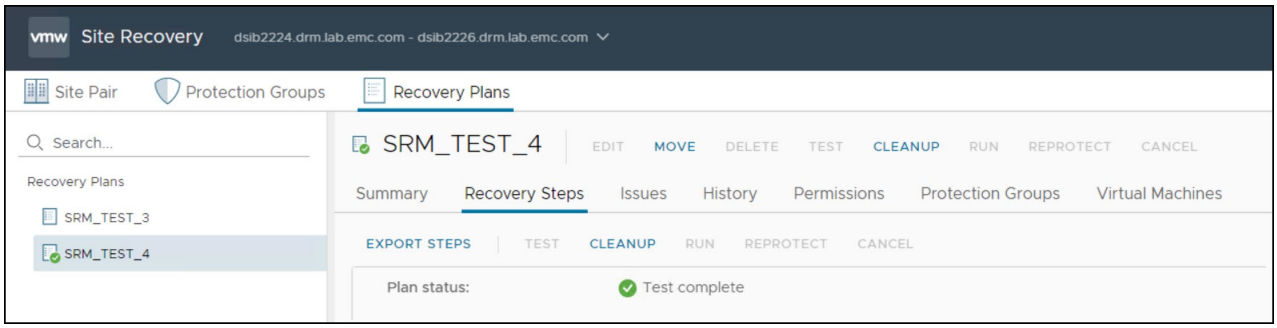

Figure 73. Recovery plan in "Test complete" status after successful test failover

The **Test complete** status will be assigned to the Recovery Plan regardless of the level of success reached by the test failover. For example, as seen in Figure 74, the test failed, but the plan status is still **Test complete**.

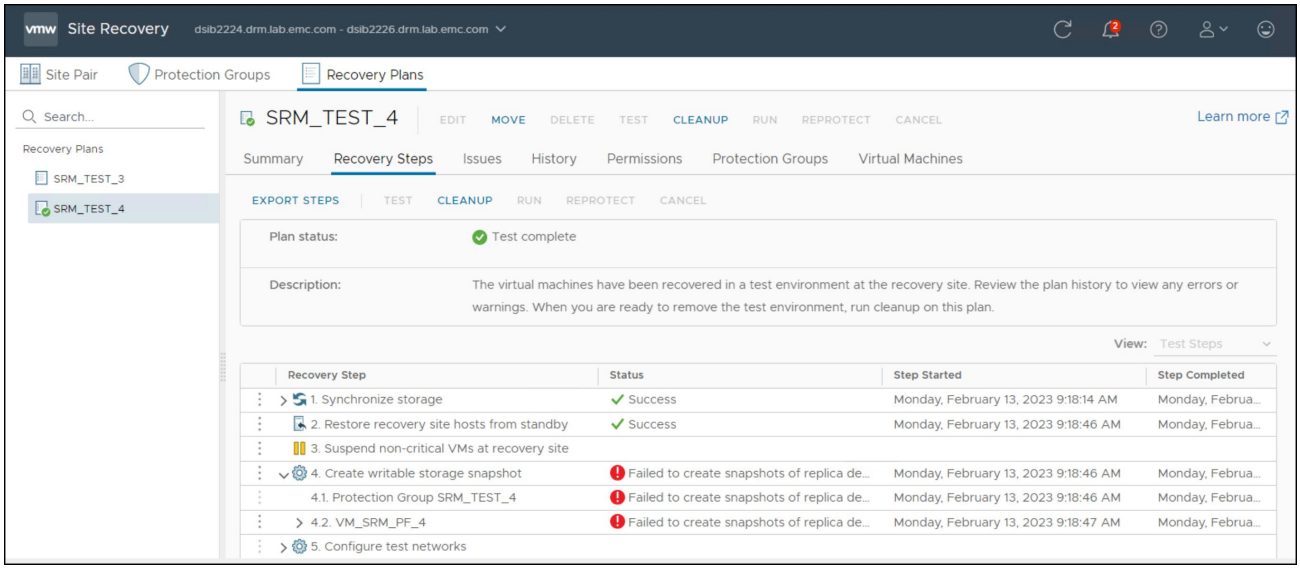

Figure 74. Recovery plan with failure, but test complete

The cleanup process is initiated, in a similar fashion to the test failover process, by clicking on the **CLEANUP** link after selecting the appropriate Recovery Plan. This can be seen in Figure 75.

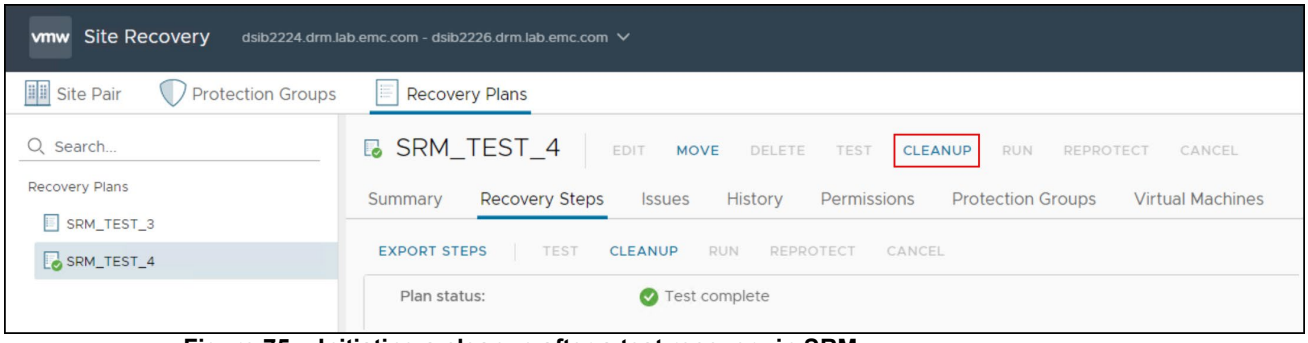

Figure 75. Initiating a cleanup after a test recovery in SRM

The **CLEANUP** link launches a similar set of windows that the original test operation brought up to confirm the reset activities that it will run. The first attempt at running this cleanup after a particular failover offers no configurable parameters and simply displays details for confirmation. This set of screens are shown in Figure 76.

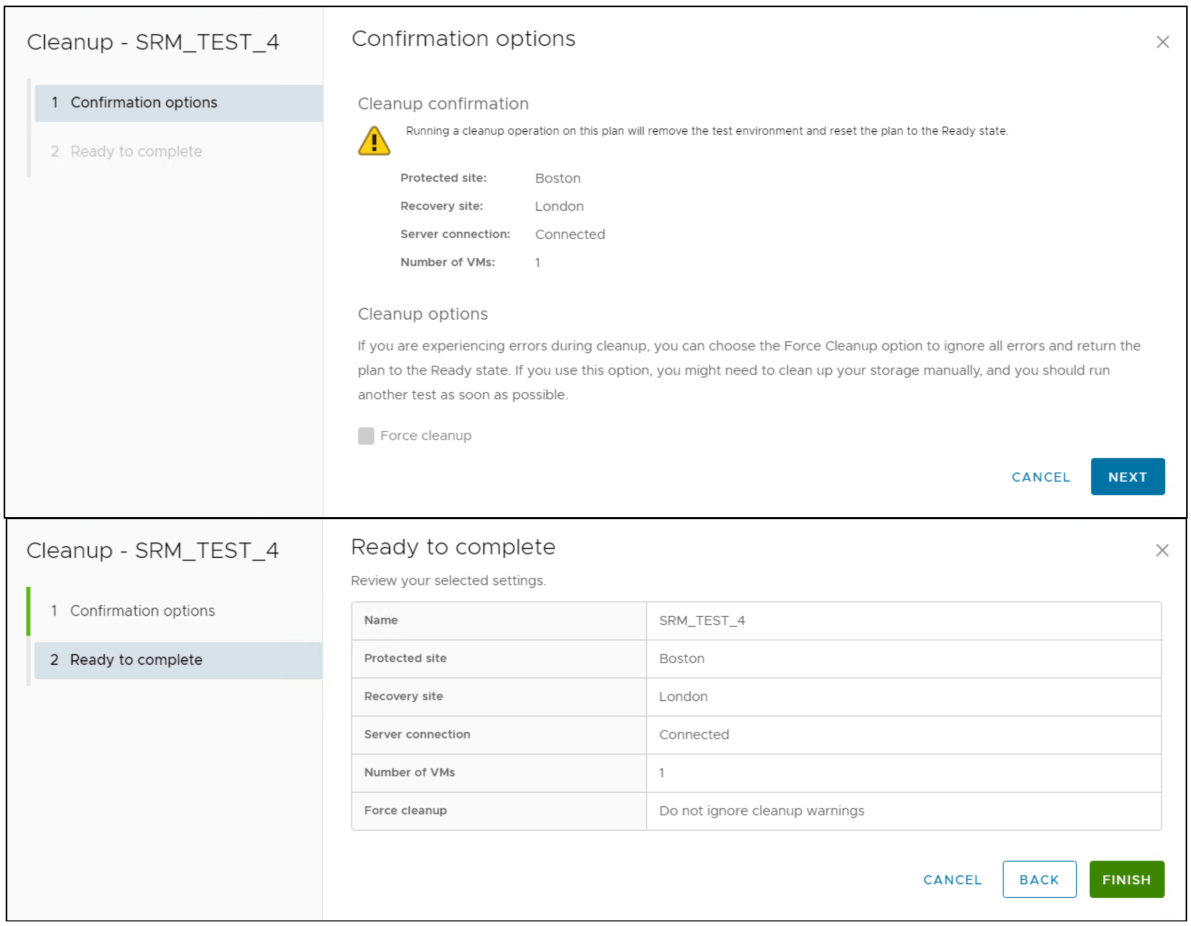

Figure 76. Cleanup operation confirmation wizard in SRM

Figure 77 shows the steps taken by the cleanup process itself.

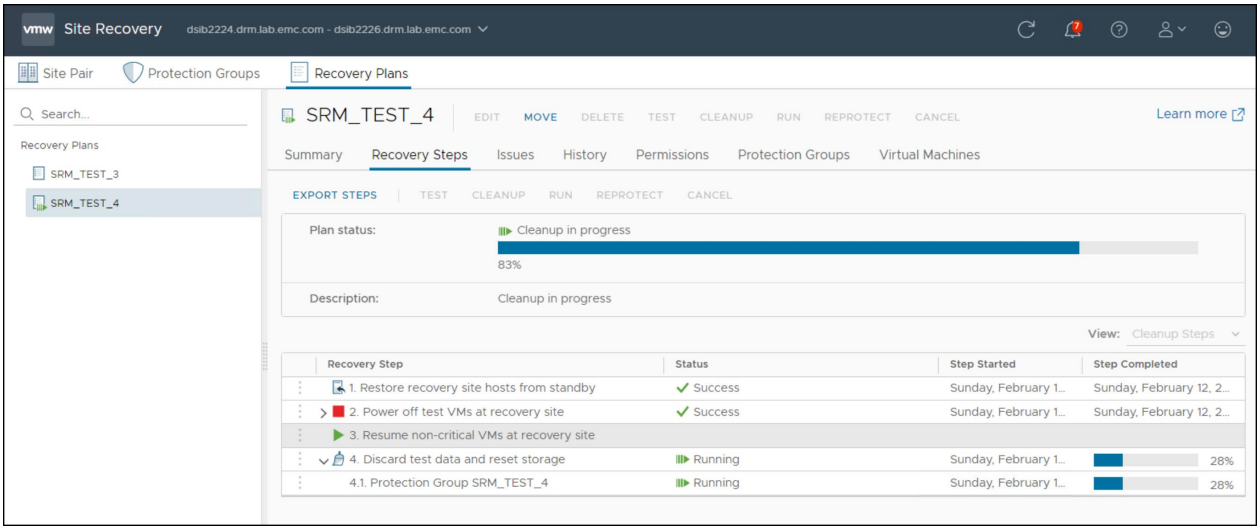

Figure 77. Cleanup operation steps in SRM

Due to a variety of reasons, the first attempt at a cleanup operation may fail. Typical causes include:

- Test failover did not complete entirely successfully
  - Storage snapshot creation failure
  - Virtual machine inventory mappings incorrect
- Environment change after the test failover but before cleanup
  - Manual change to storage outside of SRA
  - Significant protection group change
  - VMware environment failure
  - Manual change to VMware environment outside of SRM

Note: Errors reported in the SRM interface can often be very generic. Review the PowerFlex logs on the recovery site if the error indicates a failure is related to storage operations.

In cases such as these, the first cleanup operation, which does not permit the use of force, will fail. This is due to the fact that on the first run of the cleanup operation does not tolerate any failures with any step of the cleanup process. Therefore, if the cleanup process encounters an error, it will immediately fail as shown in Figure 78.

Note that the **Force cleanup** is grayed out during the first run.

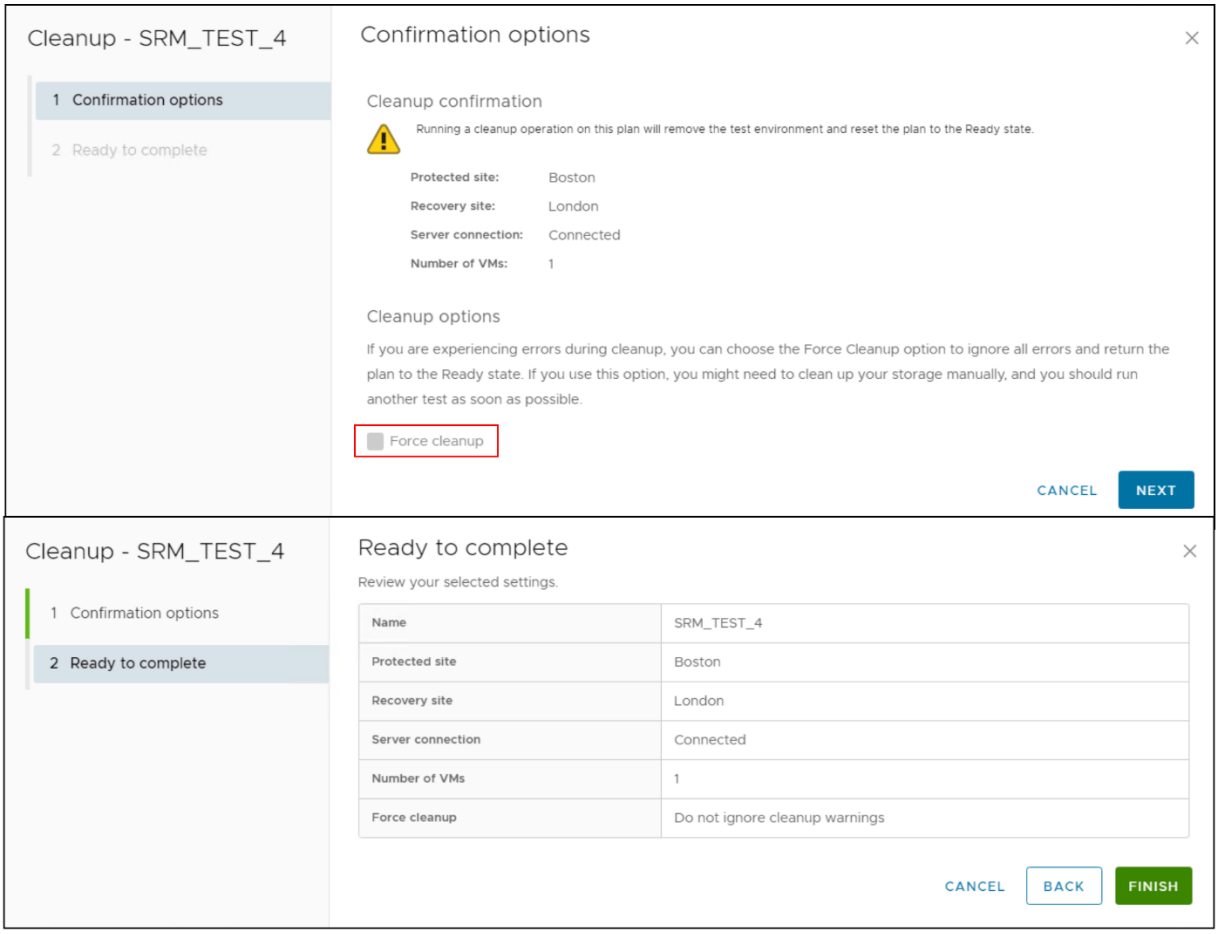

Figure 78. Initial SRM cleanup failure

Once the cleanup process has failed for the first time, the ability to force the cleanup becomes available. The cleanup confirmation wizard, when run subsequent to a failure, will now offer a checkbox to force the cleanup as seen in Figure 79. This will alter the behavior of the cleanup process to ride through any error encountered. Any operation it can complete successfully will be completed and, unlike before, any operation that encounters an error will be skipped.

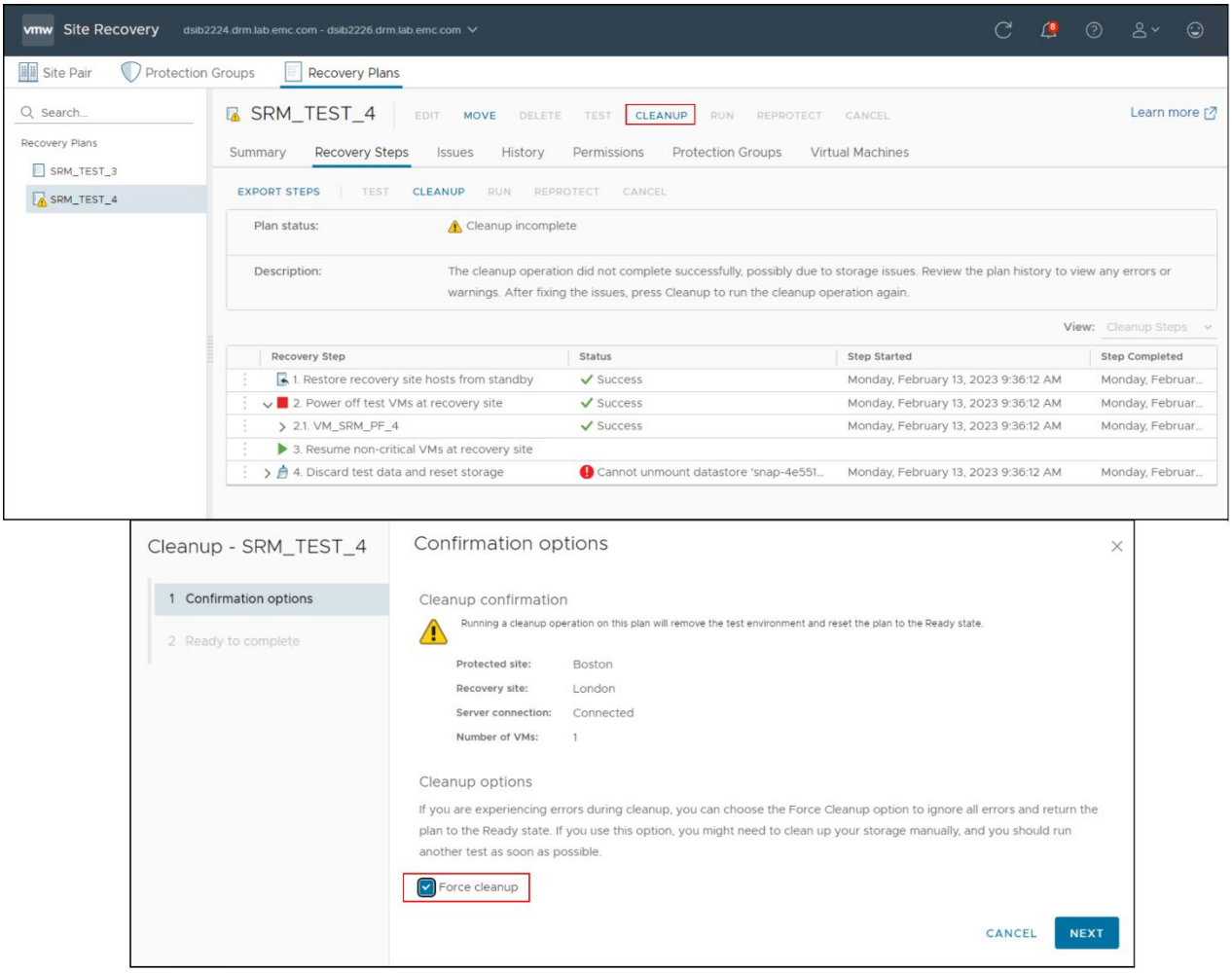

Figure 79. Executing a cleanup operation with the force option in SRM

In general, it is not advisable to resort to the force cleanup unless an actual failover operation needs to be run immediately and the time to troubleshoot any issues encountered in the cleanup cannot be afforded. Otherwise, before using the force option, attempt to resolve any issues first and then retry a non-forced cleanup again. If a force cleanup is used in haste, it may require additional manual intervention afterwards. PowerFlex and/or SRM may not be capable of making themselves ready for another test failover or failover without user intervention.

When a force cleanup is run, users should review the logs to identify the exact errors encountered. If necessary, resolve these issues and attempt to start another test failover as soon as possible to verify the environment is functioning correctly. A force cleanup is common if the PowerFlex GUI is being used in conjunction with SRM for Test Failover. As previously noted, this should be avoided.

Recovery

SRM recovery plans, whether they are for migration or failover, run as part of a planned workflow that attempts to properly shut down the protection site and synchronize data with the recovery site prior to the migration or failover. This ensures that systems are properly quiesced and that all data changes have been completely replicated prior to starting the virtual machines at the recovery site. If, however, an error is encountered during the recovery plan execution, planned migration will stop the workflow, providing an opportunity to fix the problem that caused the error before attempting to continue. If failover was used, SRM continues past the errors.

All recovery plans in SRM include an initial attempt to synchronize data between the protection and recovery sites, even during a disaster recovery scenario. During a disaster recovery event, an initial attempt is made to shut down the protection group's virtual machines and establish a final synchronization between the sites. This is designed to ensure that virtual machines are static and quiescent before running the recovery plan which minimizes data loss where possible during a disaster. If the protected site is no longer available, the recovery plan will continue to run and will run to completion even if errors are encountered. This reduces the possibility of data loss while still enabling disaster recovery to continue, balancing the requirement for virtual machine consistency with the ability to achieve aggressive recovery-point/time objectives.

Figure 80 shows the recovery plan execution wizard.

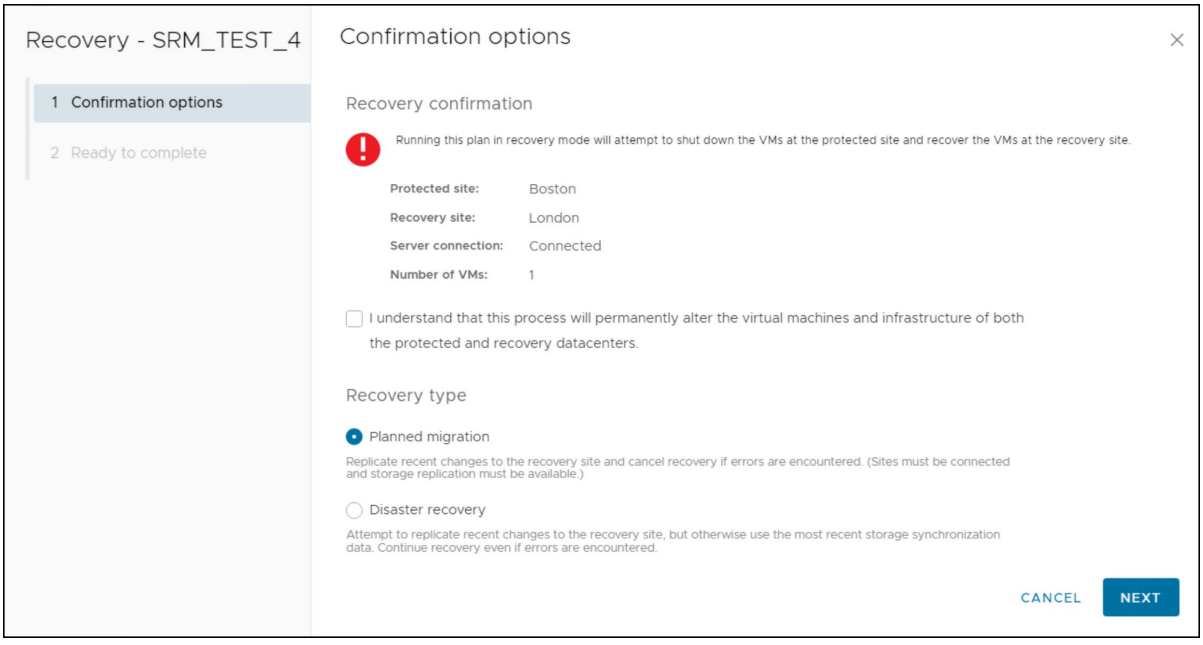

Figure 80. Recovery plan wizard

Planned migration

The recovery option **Planned migration** assures a graceful migration of virtual machines from a local vCenter to a remote vCenter. Any errors that the recovery plan encounters will immediately fail the operation and require the user to remediate these errors and

restart the migration process. Therefore, a planned migration assumes the following things (among other minor details):

- The protected and recovery VMware environments are up and running (including ESXi hosts, vCenter, virtual machines, SRM server etc.) without issues.
- The storage environment is stable and configured properly. This includes the array(s), the fabric (SAN) and the PowerFlex Gateway or PowerFlex Manager GUI interface.
- No network connectivity issues.

Before executing a recovery plan failover, Dell recommends testing the recovery plan first, preferably multiple times. Information on configuring and running a recovery test is available in the section Test Failover.

The first step is to ensure that each PowerFlex volume pair in the RCG is in a proper state. The pairs should be in a **Consistent** state as shown for PowerFlex 3 and 4 in Figure 81.

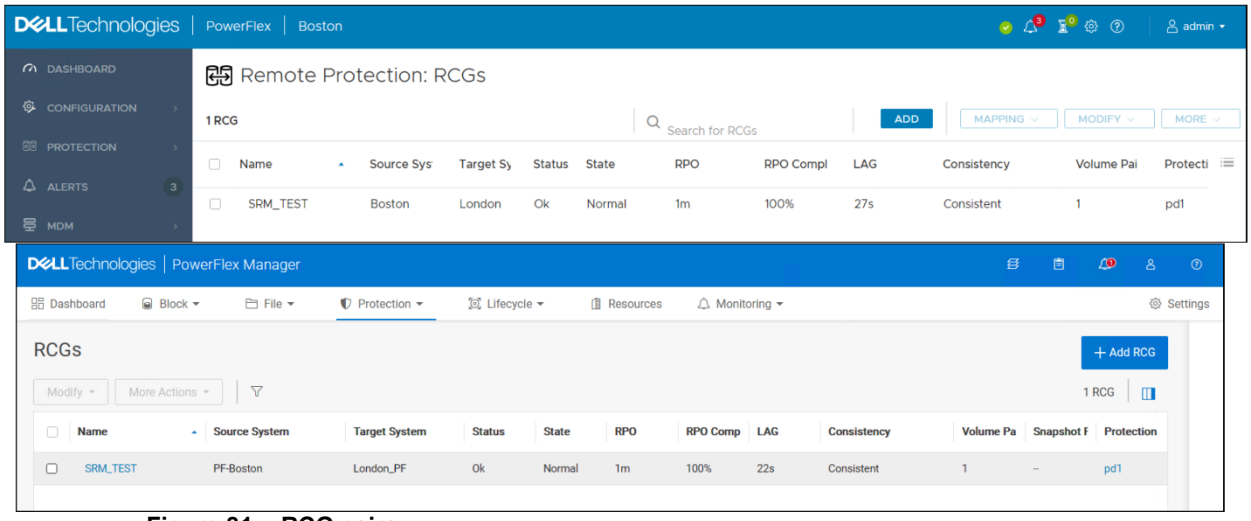

Figure 81. RCG pairs

Similarly, the **Status** in SRM should show a blue directional arrow with **Forward** as shown in Figure 82. This indicates that the pair is valid for planned migration. If the **Status** column shows a different message, such as green arrows with **Failover in Progress**, either manual intervention is required or the disaster recovery option needs to be selected instead of planned migration.

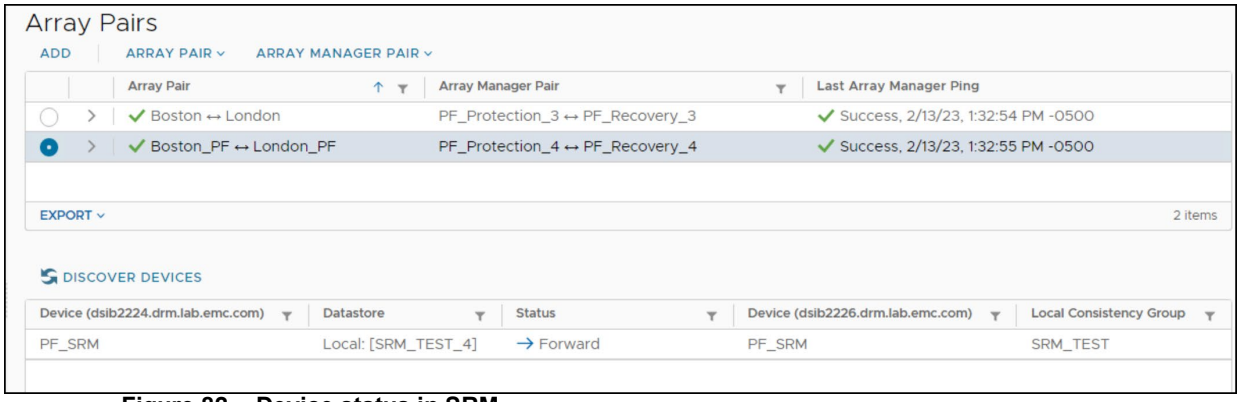

Figure 82. Device status in SRM

While the **Status** column is, in general, a good indicator of PowerFlex pair states, it is inadequate to cover all pair states. Therefore, it is advisable to use the PowerFlex 3 or 4 GUI to determine the exact status.

At this point, a planned migration can be initiated by selecting the appropriate recovery plan and selecting the **RUN** link as seen in Figure 83.

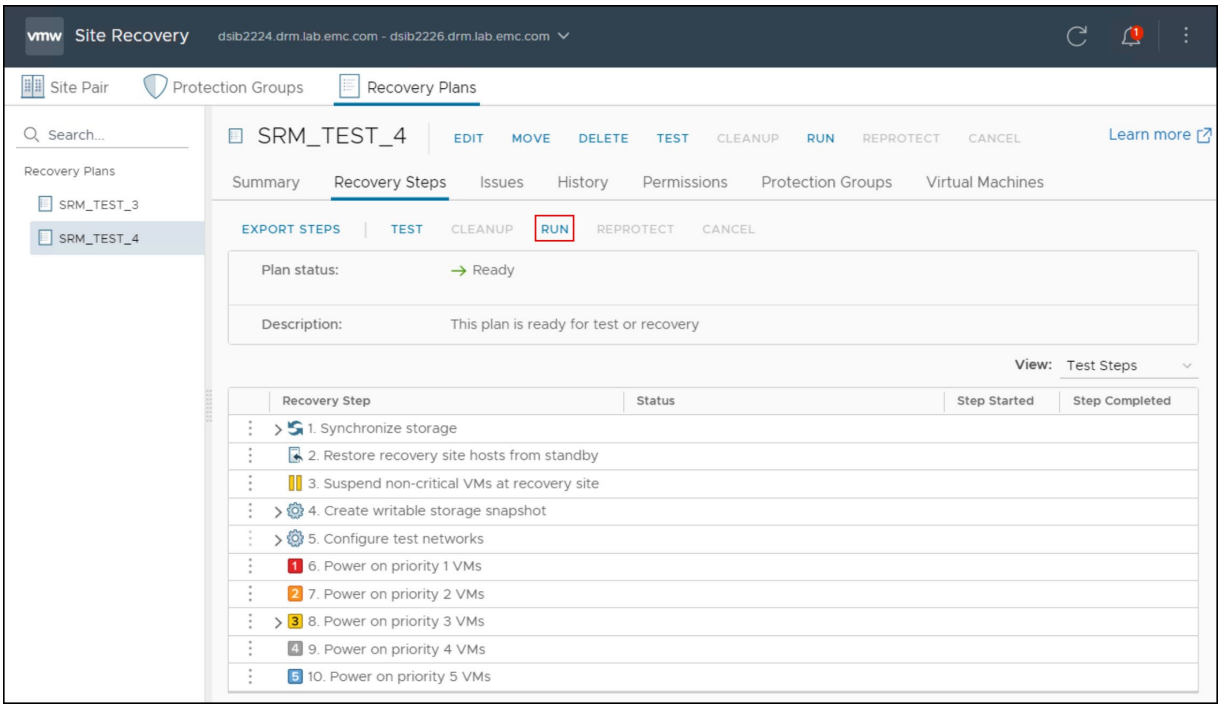

Figure 83. Initiating a planned migration with SRM

Once the Recovery link has been selected, a short confirmation wizard appears asking to confirm the initiation of the recovery operation and in which mode the recovery plan should be run.

In this example, the default **Planned migration** is chosen. The check box must be selected to proceed. This screen is shown in Figure 84.

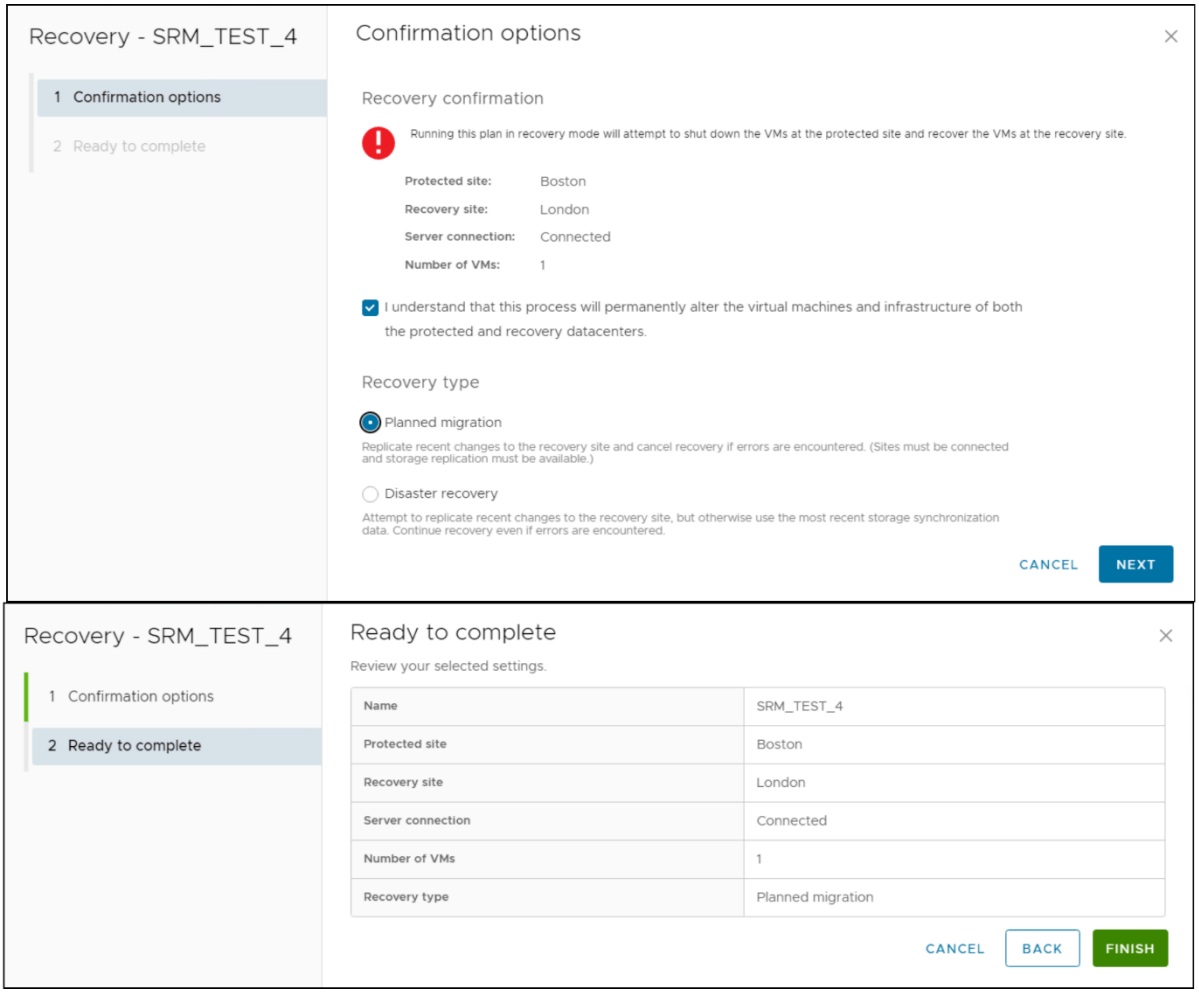

Figure 84. Recovery operation confirmation wizard in SRM

As soon as the wizard completes, the recovery operation will commence. When a recovery plan reaches the **Change recovery site storage to writable** step, as shown in Figure 85, the SRA performs a failover operation on the devices in the protection group.

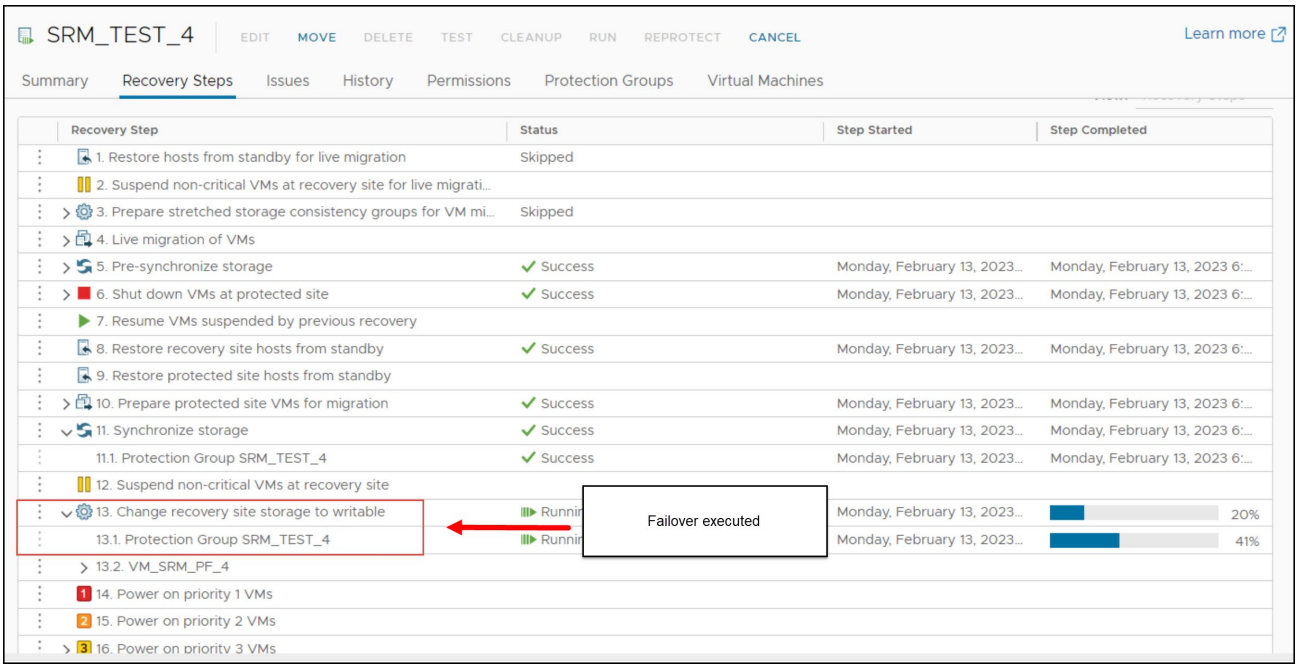

Figure 85. Steps of a recovery plan in SRM

The PowerFlex GUI will indicate that the RCG **State** is in **Failover** as shown in Figure 86.

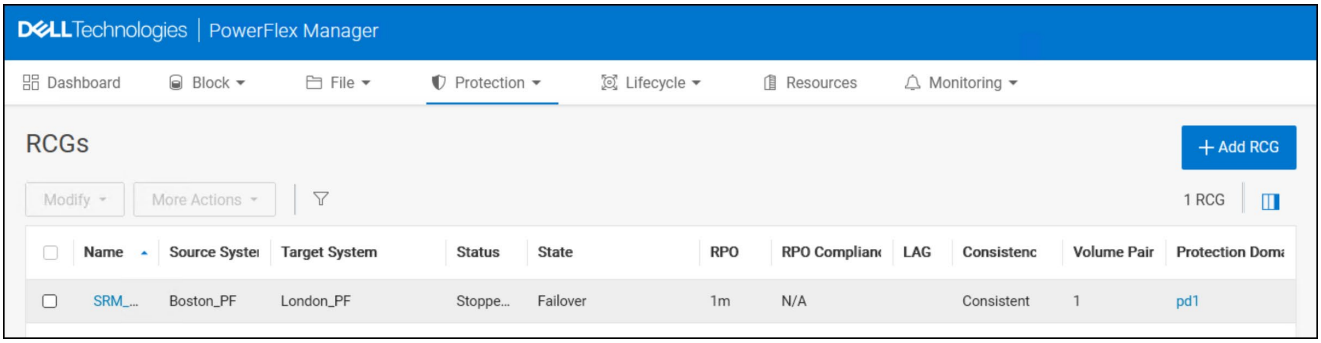

Figure 86. Failover in PowerFlex GUI

Reprotection

After a recovery plan has run, there are often cases where the environment must continue to be protected against failure to ensure its resilience and to meet objectives for disaster recovery. SRM offers “reprotection” which is an extension to recovery plan management that enables the environment at the recovery site to establish replication and protection of the environment back to the original protected site. This behavior allows users to recover the environment quickly and easily back to the original site if necessary.

It is important to note that a completely unassisted reprotection by SRM may not always be possible depending on the circumstances and results of the preceding recovery operation. Recovery plans run in planned migration mode are the likeliest candidates for a subsequent successful automated reprotection by SRM. Exceptions to this occur if certain failures or changes have occurred between the time the recovery plan was run and the reprotection operation was initiated. Those situations may cause the reprotection to fail. Similarly, if a recovery plan was run in disaster recovery mode, any persisting failures may cause a partial or complete failure of a reprotection of a recovery plan.

These different situations are described in the following sections.

Reprotect after Planned Migration

The scenario that will lead to a successful reprotection is one after a planned migration. In the case of a planned migration there are no failures in either the storage or compute environment that preceded the recovery operation. Therefore, reversing recovery plans/protections groups as well as swapping and establishing replication in the reverse direction is possible.

If failed-over virtual machines will eventually need to be returned to the original site or if they require PowerFlex replication protection, it is recommended to run a reprotect operation as soon as possible after a migration.

Reprotect is only available after a recovery operation has occurred, which is indicated by the recovery plan being in the **Recovery complete** state. Later versions of SRM will warn the user that **Reprotect needed** as shown in Figure 87.

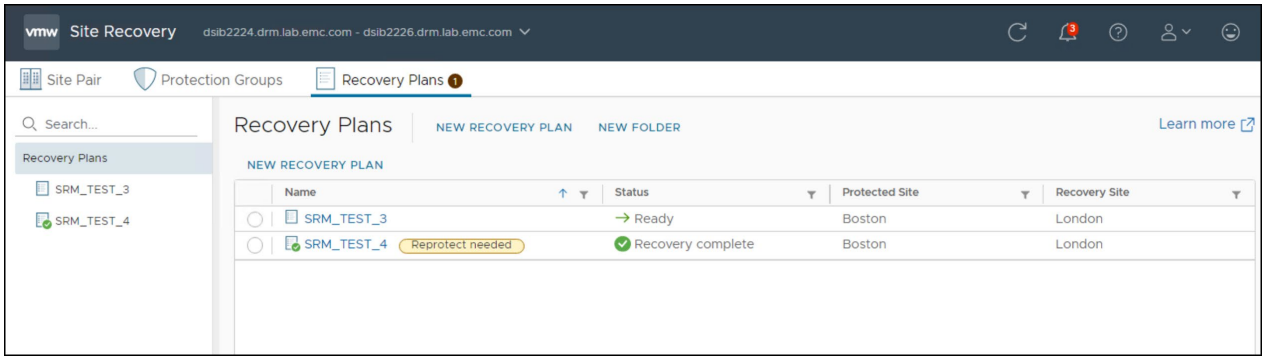

Figure 87. SRM warns user to run reprotect

A reprotect is run by selecting the appropriate recovery plan and selecting the **REPROTECT** links as shown in Figure 88.

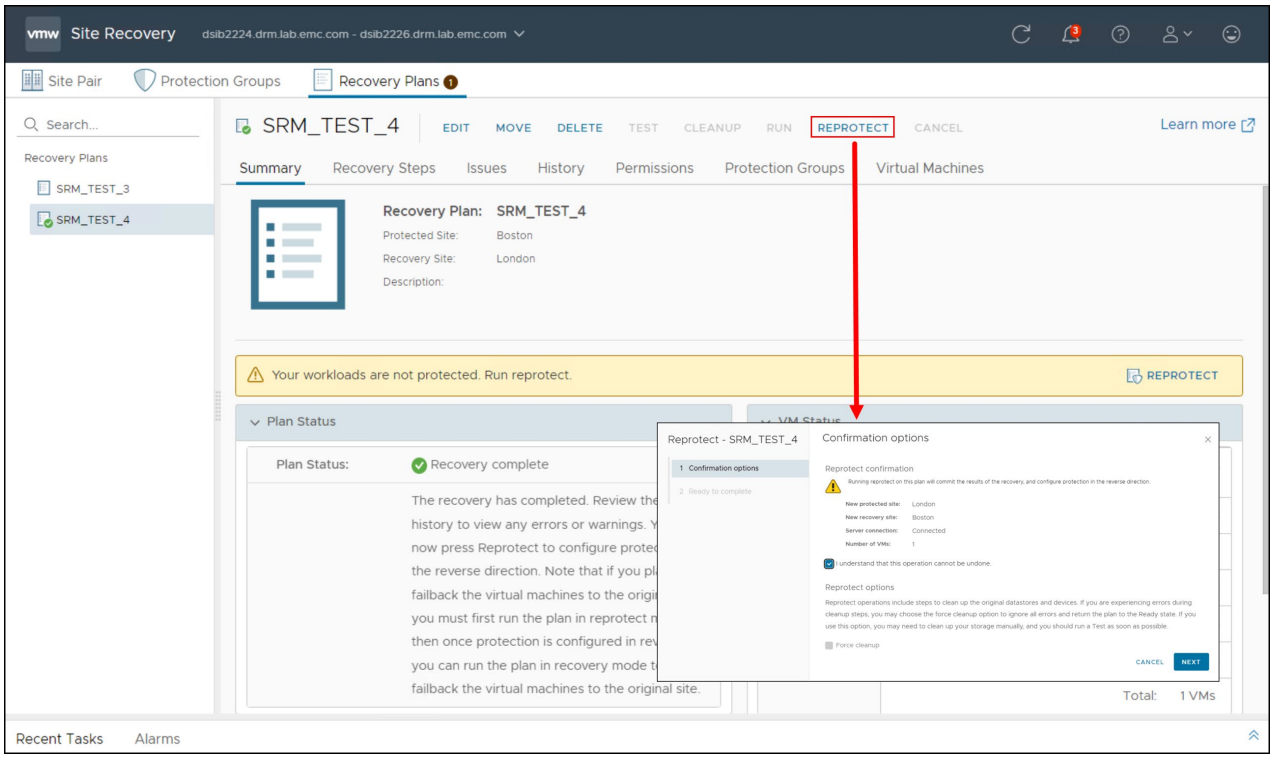

Figure 88. Executing a reprotect operation in SRM

The reprotect operation does the following things:

- **Reverses protection groups**. The protection groups are deleted on the original protection SRM server and are recreated on the original recovery SRM server. The inventory mappings are configured (assuming the user has pre-configured them in SRM on the recovery site) and the necessary shadow or placeholder VMs are created and registered on the newly designated recovery SRM server.
- **Reverses recovery plan**. The failed-over recovery plan is deleted on the original recovery SRM server and recreated with the newly reversed protection group.
- **Swaps personality of PowerFlex RCG pairs**. The PowerFlex SRA performs a swap on the target pairs which enables replication to be established back to the original site. Target becomes source and vice versa.
- **Re-establishes replication**. After the swap, the PowerFlex SRA incrementally re-establishes replication between the RCG pairs, but in the opposite direction from what it was before the failover/migration.

The PowerFlex GUI events log records the reversal of replication. An example is shown in Figure 89.

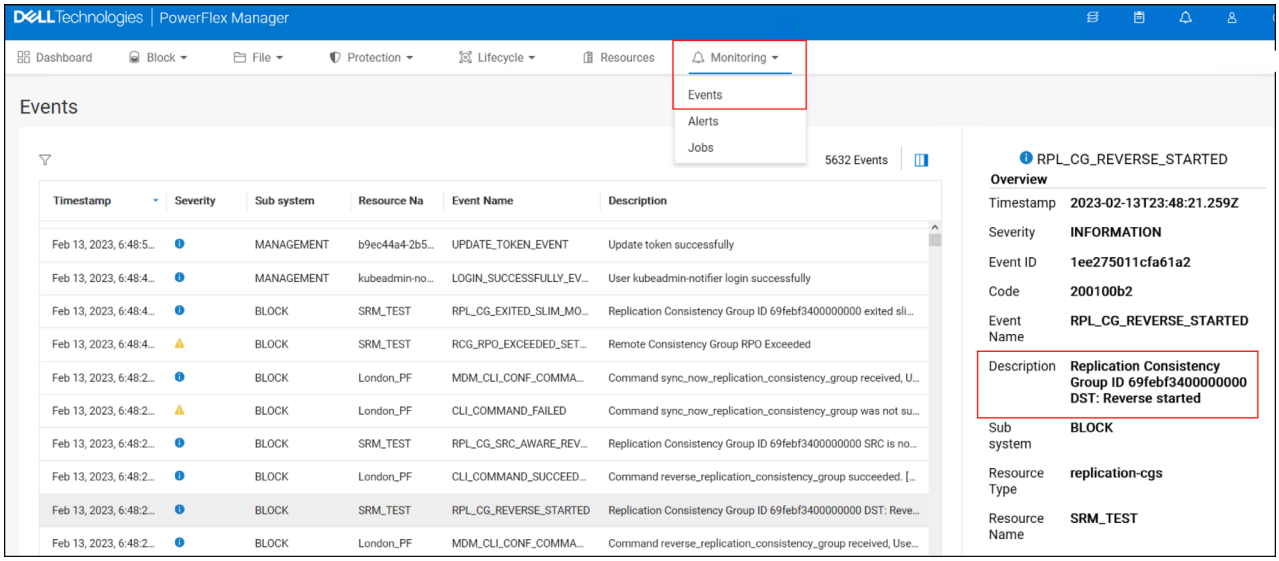

Figure 89. Reverse replication event in PowerFlex GUI

Figure 90 demonstrates the steps involved in a reprotect operation.

Note: If the command syncOnce fails during reprotect, the process will still complete successfully.

Recovery Step	Result	Step Started	Step Completed	Execution Time
1. Restore protected site hosts from standby	Success	2023-02-13 23:48:38 (UTC 0)	2023-02-13 23:48:38 (UTC 0)	00:00:01
2. Configure storage to reverse direction	Success	2023-02-13 23:48:38 (UTC 0)	2023-02-13 23:48:42 (UTC 0)	00:00:04
2.1. Protection Group SRM_TEST_4	Success	2023-02-13 23:48:38 (UTC 0)	2023-02-13 23:48:42 (UTC 0)	00:00:04
Consistency Group "SRM_TEST": Success				
2.1.1. Configure array-based storage	Success	2023-02-13 23:48:38 (UTC 0)	2023-02-13 23:48:42 (UTC 0)	00:00:04
3. Configure protection to reverse direction	Success	2023-02-13 23:48:42 (UTC 0)	2023-02-13 23:48:44 (UTC 0)	00:00:01
3.1. Protection Group SRM_TEST_4	Success	2023-02-13 23:48:42 (UTC 0)	2023-02-13 23:48:44 (UTC 0)	00:00:01
3.1.1. Configure protection	Success	2023-02-13 23:48:42 (UTC 0)	2023-02-13 23:48:44 (UTC 0)	00:00:01
3.1.2. Configure VMs protection	Success	2023-02-13 23:48:44 (UTC 0)	2023-02-13 23:48:44 (UTC 0)	00:00:00
3.1.2.1. VM_SRM_PF_4	Success	2023-02-13 23:48:44 (UTC 0)	2023-02-13 23:48:44 (UTC 0)	00:00:00
4. Clean up storage	Success	2023-02-13 23:48:44 (UTC 0)	2023-02-13 23:48:45 (UTC 0)	00:00:02
4.1. Protection Group SRM_TEST_4	Success	2023-02-13 23:48:44 (UTC 0)	2023-02-13 23:48:45 (UTC 0)	00:00:02
5. Synchronize storage	Success	2023-02-13 23:48:45 (UTC 0)	2023-02-13 23:48:46 (UTC 0)	00:00:01
5.1. Protection Group SRM_TEST_4	Success	2023-02-13 23:48:45 (UTC 0)	2023-02-13 23:48:46 (UTC 0)	00:00:01

Figure 90. Reprotect operation steps

Reprotect after a temporary failure

The previous section describes the best possible scenario for a smooth reprotection because it follows a planned migration where no errors are encountered. For recovery plans failed over in disaster recovery mode, this may not be the case.

Disaster recovery mode allows for failures ranging from the very small to a full site failure of the protection datacenter. If these failures are temporary and recoverable a fully successful reprotection may be possible once those failures have been rectified. In this case, a reprotection will behave similar to the scenario described in the previous section. If a reprotection is run before the failures are corrected or certain failures cannot be fully recovered, an incomplete reprotection operation will occur. This section describes this scenario.

For reprotect to be available, the following steps must first occur:

- A recovery must be run with all steps finishing successfully. If there were any errors during the recovery, the user needs to resolve the issues that caused the errors and then rerun the recovery.
- The original site should be available and SRM servers at both sites should be in a connected state. If the sites are disconnected, a reprotect will fail immediately as shown in Figure 91. If the original site cannot be restored (for example, if a physical catastrophe destroys the original site) automated reprotection cannot be run and manual recreation will be required if and when the original protected site is rebuilt.

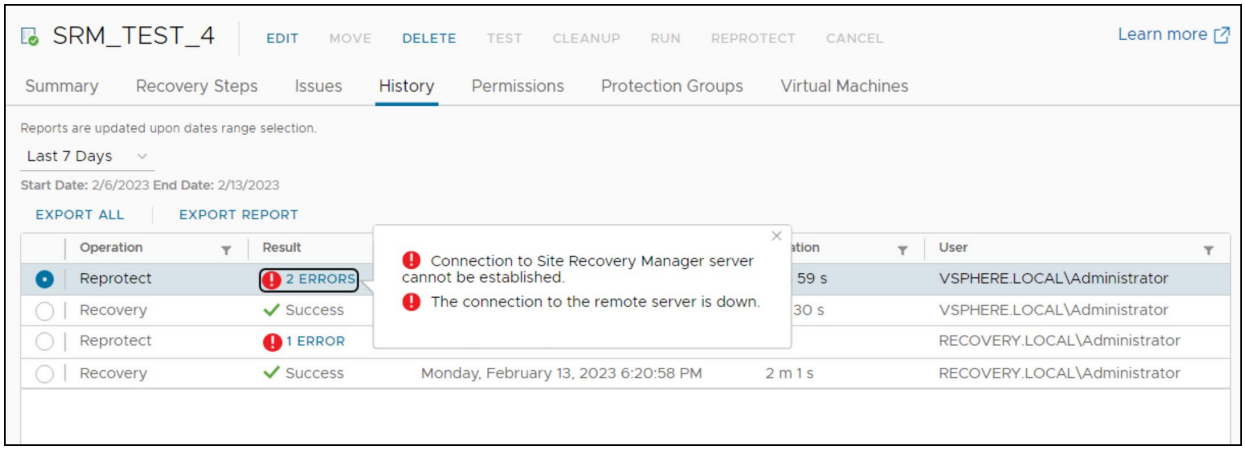

Figure 91. Reprotect fails with sites disconnected

If the protected site SRM server was disconnected during failover and is reconnected later, SRM will want to retry certain recovery operations before allowing reprotect. This typically occurs if the recovery plan was not able to connect to the protected side vCenter server and power down the virtual machines due to network connectivity issues. If network connectivity is restored after the recovery plan was failed over, SRM will detect this situation and require the recovery plan to be re-run in order to power those VMs down.

A reprotection operation will fail if it encounters any errors the first time it runs. If this is the case, the reprotect must be run a second time but with the **Force cleanup** option selected as shown in Figure 92.

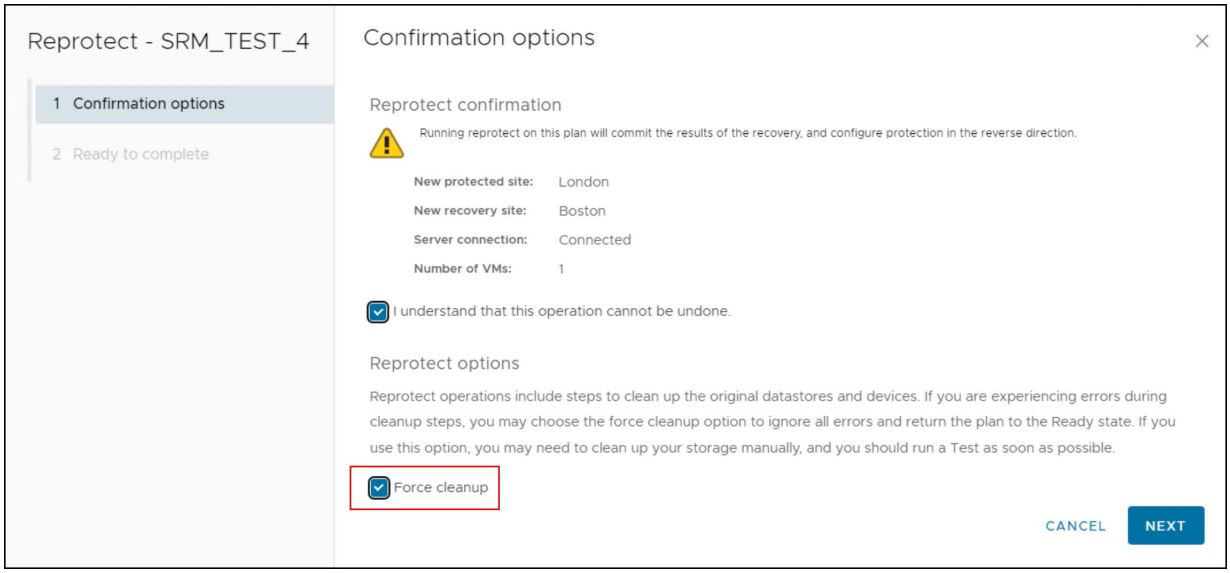

Figure 92. Forcing a reprotect operation

Once the force option is selected, any errors will be acknowledged and reported but ignored. This will allow the reprotect operation to continue even if the operation has experienced errors. It will attempt all of the typical steps and complete whichever ones are possible. Therefore, in certain situations, the PowerFlex replication may not be properly reversed even though the recovery plan and protection group(s) were. If the “Configure Storage to Reverse Direction” step fails, manual user intervention with PowerFlex GUI or CLI may be required to complete the process. The user should ensure that:

- A source/target swap has occurred by ensuring the replicated target/source devices have changed personalities
- Asynchronous replication has been re-established

In the case of a temporary storage failure or replication partition, it is likely that manual intervention will be required prior to executing a reprotect operation. In this situation the source devices may not have been unmounted.

Reprotect after a failover due to unrecoverable failure

In extreme circumstances, the storage and/or the compute environment may be rendered completely unrecoverable due to a disaster. In this scenario, reprotect will not be possible. Therefore, the process of “re-protecting” the original recovery site is no different than the original setup of the protection groups and recovery plans from scratch. An example of an unrecoverable failure would be if the protection site array was lost and then replaced, requiring new RCG pair relationships.

Failback

An automated failback workflow can be run to return the entire environment to the primary site from the secondary site. This will happen after reprotection has ensured that data replication and synchronization have been established to the original site.

Failback runs through the same workflow as shown in Figure 93 that was originally used to migrate the environment to the recovery site. It guarantees that the critical virtual machines encapsulated by the failed-over recovery plan are returned to their original environment. The workflow runs only if reprotection has successfully completed.

Failback ensures the following:

- All virtual machines that were initially migrated to the recovery site will be moved back to the primary site (assuming they still exist).
- Environments that require that disaster recovery testing be done with live environments with genuine migrations can be returned to their initial site.
- Simplified recovery processes will enable a return to standard operations after a failure.

Failback is no different in its execution or behavior than the original failover operation. Before failback can occur, valid protection groups and recovery plans must be created or re-configured through a reprotect operation or manual creation by the user.

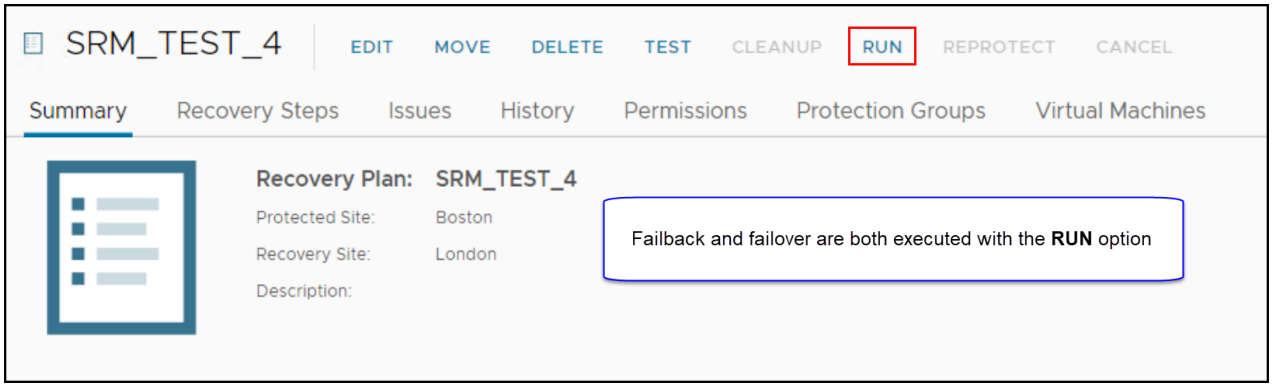

Figure 93. Executing failback in SRM

Once failback is complete, it is still necessary to run a reprotect on the primary site.

Conclusion

VMware vSphere, SRM, and PowerFlex combine to provide a highly available business platform for automated disaster recovery. This platform helps to achieve RPO as low as 15 seconds using PowerFlex native asynchronous replication and supports planned migrations for your virtual infrastructure.

References

Dell Technologies documentation

The following Dell Technologies documentation provides additional information. Access to these documents depends on your login credentials. If you do not have access to a document, contact your Dell Technologies representative.

- Introduction to Dell PowerFlex replication
- Dell PowerFlex documentation

VMware documentation

The following VMware documentation provides additional and relevant information:

- VMware Site Recovery Manager documentation

Appendix: Known Issues

The following are some known issues that a user may hit while deploying PowerFlex with VMware SRM. The intent of the appendix is to cover those most common to customers.

Journal Capacity

When creating a Replication Consistency Group (RCG) between two peered systems, there must be enough Journal Capacity available on each system. For both PowerFlex 3 and 4, each SDR must be represented by a minimum of 108 GB of Journal Capacity. A minimally configured system with only three SDRs, therefore, must have a minimum of 324 GB of Journal Capacity.

If insufficient Journal Capacity exists on either system, PowerFlex will produce the error “Could not allocate capacity for Replication Consistency Group” during Replication Consistency Group creation as shown in Figure 94 for GUI and Figure 95 for CLI. Note the error may be accompanied by additional information related to the connection handshake.

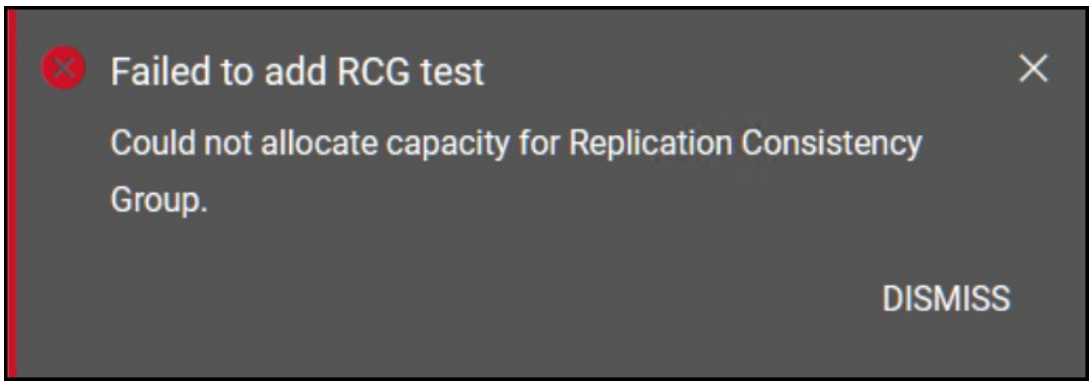

Figure 94. Insufficient Journal Capacity for RCG – GUI creation

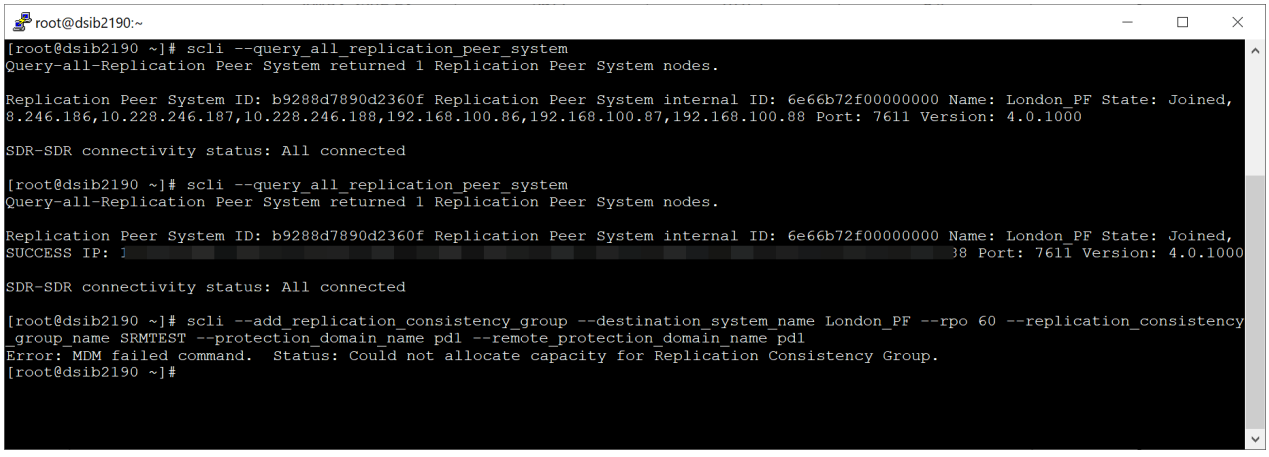

Figure 95. Insufficient Journal Capacity for RCG – CLI creation