Dell EMC PowerStore: VMware vSphere Best Practices

Abstract

This document provides best practices for integrating VMware® vSphere® hosts with Dell EMC™ PowerStore™.

April 2021

H18116.1

Revisions

Date	Description
April 2020	Initial release: PowerStoreOS 1.0
April 2021	Updates for PowerStoreOS 2.0

Acknowledgments

Author: Darin Schmitz

DELLTechnologies

Table of contents

Executive summary

This document provides recommendations, tips, and other helpful guidelines for integrating external VMware® vSphere® hosts with the Dell EMC™ PowerStore™ platform.

Audience

This document is intended for IT administrators, storage architects, partners, and Dell Technologies™ employees. This audience also includes any individuals who may evaluate, acquire, manage, operate, or design a Dell EMC networked storage environment using PowerStore systems.

1 Introduction

Dell EMC PowerStore is a robust and flexible storage option that is ideal for use with VMware vSphere.

1.1 PowerStore overview

PowerStore achieves new levels of operational simplicity and agility. It uses a container-based microservices architecture, advanced storage technologies, and integrated machine learning to unlock the power of your data. PowerStore is a versatile platform with a performance-centric design that delivers multidimensional scale, always-on data reduction, and support for next-generation media.

PowerStore brings the simplicity of public cloud to on-premises infrastructure, streamlining operations with an integrated machine-learning engine and seamless automation. It also offers predictive analytics to easily monitor, analyze, and troubleshoot the environment. PowerStore is highly adaptable, providing the flexibility to host specialized workloads directly on the appliance and modernize infrastructure without disruption. It also offers investment protection through flexible payment solutions and data-in-place upgrades.

1.2 VMware vSphere overview

VMware vSphere is the industry-leading virtualization platform and a core building block to the software-defined data center (SDDC). VMware vSphere is primarily composed of vCenter® for management and ESXi™ hosts that provide the hypervisor for compute and memory virtualization.

1.3 Prerequisite reading

Before implementing the best practices in this document, it is recommended to review the *Dell EMC PowerStore Host Configuration Guide* and the other resources available at the PowerStore Info Hub.

1.4 Terminology

The following terms are used with PowerStore:

Appliance: The solution containing a base enclosure and attached expansion enclosures. The size of an appliance could include only the base enclosure or the base enclosure plus expansion enclosures.

Base enclosure: The enclosure containing both nodes (node A and node B) and the NVMe drive slots.

Cluster: Multiple PowerStore appliances in a single grouping.

Expansion enclosure: Enclosures that can be attached to a base enclosure to provide additional SAS-based drive slots.

Node: The component within the base enclosure that contains processors and memory. Each appliance consists of two nodes.

PowerStore Manager: The web-based user interface (UI) for storage management.

NVMe over Fibre Channel (NVMe-FC): Allows hosts to access storage systems across a network fabric with the NVMe protocol using Fibre Channel as the underlying transport.

2   Host configuration

While most settings for stand-alone ESXi hosts that are connected to PowerStore appliances can remain at the default values, some changes are required for PowerStore stability, performance, and efficiency. The recommended changes and instructions about how to set them are specified in the document *Dell EMC PowerStore Host Configuration Guide* on the [PowerStore Info Hub](). While administrators can use this section for high-level explanations and reasoning behind the recommendations, administrators should see the Host Configuration Guide for the current settings.

Caution: These recommended settings are for external ESXi hosts only, and do not apply to the ESXi instances running within PowerStore X model appliances.

Note: Dell EMC Virtual Storage Integrator (VSI) allows administrators to easily configure the ESXi host best-practice settings with PowerStore. See section 4.8 for more details.

2.1   Queue depth

There are multiple locations within ESXi and the guest operating systems to modify queue depth. While increasing the queue depth in an application, vSCSI device, or ESXi driver module can potentially increase performance, modifying or increasing queue depths can potentially overwhelm the array. For details about queue-depth settings, see the document *Dell EMC PowerStore Host Configuration Guide*.

2.2   Timeouts

Setting disk timeouts is an important factor for applications to survive both unexpected and expected node outages, such as failures or rebooting for updates. While the default SCSI timeout in most applications and operating systems is 30 seconds, storage and application vendors typically recommend increasing these timeouts to 60 seconds or more to help ensure uptime. Two of the main locations to change the timeouts are at the ESXi host level and at the virtual-machine-guest-OS level. For details on setting timeouts for ESXi, see the *Dell EMC PowerStore Host Configuration Guide*.

2.2.1   ESXi host timeouts

The timeout values set at the ESXi-host-driver level help ensure that the hosts and virtual machines can survive a storage node failover event. For details on setting timeouts for ESXi, see the *Dell EMC PowerStore Host Configuration Guide*.

2.2.2   Guest operating system timeouts

If VMware Tools is installed into a guest operating system, it automatically sets the timeout values. However, if the guest operating system does not have VMware Tools installed, the administrator can set these values manually. While [VMware documentation]() has examples on setting the disk timeouts in Microsoft® Windows® guest operating systems, consult the knowledge bases from operating system vendors to obtain specific guest settings.

2.3 Multipathing

With the vSphere Pluggable Storage Architecture (PSA), the storage protocol will determine which Multipathing Plugin (MPP) is assigned to volumes mapped from the PowerStore array. With SCSI-based protocols such as Fibre Channel and iSCSI, the Native Multipathing Plug-in (NMP) is used, whereas with NVMe-oF, the VMware High Performance Plug-in (HPP) will be used.

2.3.1 Native Multipathing Plug-in (NMP)

While SCSI-based volumes using Fibre Channel and iSCSI are automatically assigned the NMP, the ESXi Storage Array Type Plug-in (SATP) module and its corresponding path selection policy (PSP) may need claim rules to be configured to use Round Robin (RR) with PowerStore appliances. Applying the settings in the *Dell EMC PowerStore Host Configuration Guide* ensures that all volumes presented to the host use Round Robin as the default pathing policy.

Also, the recommended esxcli command sets the IOPS path-change condition to one I/O per path. While the default setting within the RR PSP sends 1,000 IOPS down each path before switching to the next path, this recommended setting instructs ESXi to send one command down each path. This results in better utilization of each path's bandwidth, which is useful for applications that send large I/O block sizes to the array.

2.3.2 High Performance Plug-in (HPP)

For NVMe-oF targets, the HPP replaces the NMP. The HPP will claim NVMe devices and is designed to improve storage performance for modern high-speed interfaces.
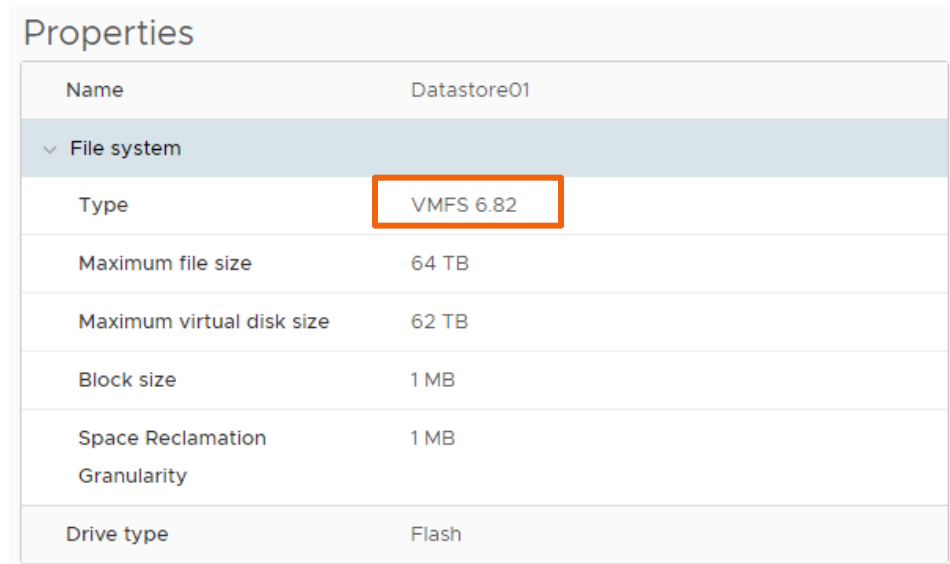
In addition, the HPP has multiple Path Selection Schemes (PSS) available to determine which physical paths are used for I/O requests. By default, the HPP uses the Load Balance – Round Robin (LB-RR) Path Selection Scheme recommended by the *Dell EMC PowerStore Host Configuration Guide*, so no additional changes or claim rules are required.

For more information about NVMe-oF and the High Performance Plug-in, see the following resources on the VMware website:

- VMware NVMe Concepts
- VMware High Performance Plug-In and Path Selection Schemes
- Requirements and Limitations of VMware NVMe Storage

2.4   Operating system disk formats

While most versions of VMFS are backwards-compatible, it is a best practice to verify and use the latest version of VMFS recommended by VMware. Typically, new VMFS versions are bundled with an ESXi upgrade. As a migration path, VMware vCenter allows administrators to use VMware vSphere® Storage vMotion® to migrate virtual machines to new VMFS datastores formatted with the latest version.

Properties	
Name	Datastore01
⌄ File system	
Type	VMFS 6.82
Maximum file size	64 TB
Maximum virtual disk size	62 TB
Block size	1 MB
Space Reclamation Granularity	1 MB
Drive type	Flash

Figure 1   Example of datastore properties showing the VMFS version

Note: PowerStore X model appliances use VMware vSphere Virtual Volumes™ (vVols) storage containers for storage of virtual machines that run on internal PowerStore X model nodes. While it is not best practice, block volumes can only be presented to the internal ESXi hosts within PowerStore X model using the REST API.

3 Sizing and performance optimization

There are several best practices for provisioning storage from a PowerStore appliance to an external vSphere cluster. The size of VMFS datastores and the number of virtual machines that are placed on each can affect the overall performance of the volume and array.

Virtual machines running within PowerStore X model internal nodes use internal vVol storage by default. However, since the PowerStore X model can simultaneously run VMs on internal nodes and serve storage to external ESXi hosts, the VMFS datastore sizing and virtual-machine placement strategies that are detailed in this section are only applicable when using PowerStore as storage to external ESXi hosts.

3.1 Volume and VMFS datastore sizing

When a volume is created on PowerStore T models, the best practice is to create a volume no larger than needed and use a single VMFS partition on that volume.

While the maximum datastore size can be up to 64 TB, it is recommended to begin with a small datastore capacity and increase it as needed. Right-sizing datastores prevents accidentally placing too many virtual machines on the datastore, thus decreasing the probability of resource contention. Since datastore and VMDK sizes can be easily increased if a virtual machine needs additional capacity, it is not necessary to create datastores larger than required. For optimal performance, best practice is to increase the number of datastores rather than increase their size.

If a standard for the environment has not already been established, the recommended starting size for a VMFS datastore volume is 1 TB (see Figure 2).

Figure 2   PowerStore volume creation wizard

3.1.1 Increasing the size of VMFS datastores

Increasing the size of VMFS datastores can be completed within PowerStore Manager by modifying a volume's properties and increasing the size. After rescanning the storage adapters on the ESXi hosts, increase the VMFS partition size. Open the wizard, right-click the datastore, and select **Increase Datastore Capacity**. The best practice is to extend datastores using contiguous space within a single volume, and to avoid spanning volumes due to recovery complexity.

Note: The Dell EMC VSI plug-in can automate the process of increasing the size of datastores with only a few clicks.
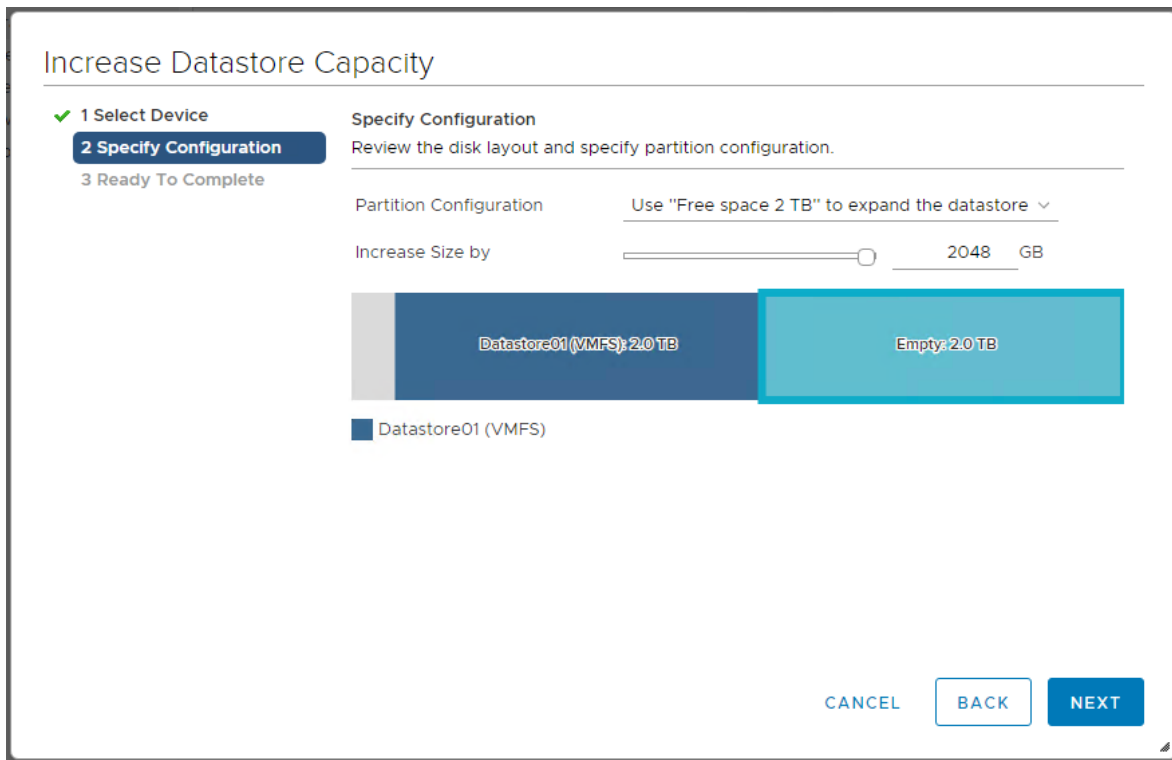

Figure 3   Increasing the VMFS datastore size

3.2 Performance optimizations

While ESXi storage performance tuning is a complex topic, this section describes a few simple methods to proactively optimize performance.

Note: The Dell EMC VSI plug-in allows administrators to quickly set host best practices for optimal operation and performance.

3.2.1 Virtual machines per VMFS datastore

While the recommended number of virtual machines per VMFS datastore is subjective, many factors determine the optimum number of VMs that can be placed on each. Although most administrators only consider capacity, the number of concurrent I/Os being sent to the disk device is one of the most important factors in the overall performance. The ESXi host has many mechanisms to ensure fairness between virtual machines competing for datastore resources. However, the easiest way to control performance is by regulating how many virtual machines are placed on each datastore. The best way to determine if a datastore

has too many virtual machines is by monitoring disk latency with either esxtop or PowerStore Manager. If the concurrent virtual machine I/O patterns are sending too much traffic to the datastore, the disk queues fill, and higher latency is generated.

Note: The vVol datastores within PowerStore X model arrays do not have the same technical architecture as VMFS datastores. The virtual machine placement strategies that are described in this section are not necessary.

3.2.2 Modifying VMFS queue depth

To regulate and ensure fairness of I/O being sent from VMs to each datastore, ESXi has an internal mechanism to control how many I/Os each virtual machine can send to the datastore at a time. This is called Disk.SchedNumReqOutstanding (DSNRO). Although DSNRO for each datastore can be tuned using esxcli, the best practice is to not modify this setting unless operating in a test environment or directed by support personnel.

3.2.3 Multiple-VM and single-VM-per-volume strategies

Although there are small performance benefits to placing a single virtual machine on a VMFS datastore, placing multiple virtual machines on each VMFS datastore is common practice. In most cases, using vVols achieves the same or better performance benefits over placing a single VM on a datastore.

The disadvantages to placing a single VM on its own datastore are that it reduces consolidation ratios and increases the management overhead of maintaining numerous items.

3.3   Virtual machine affinity with Model X appliances

When using PowerStore Model X appliances in a multi-appliance cluster, it is important to remember that while the cluster may have multiple ESXi hosts, each virtual machine's virtual disks are bound to an individual appliance. This means that virtual machine storage performance may experience higher latency if it is vMotioned to other appliances in the cluster. Best practice is to always keep the virtual machine running within the two hosts of the appliance that own the volumes. If the virtual machine needs to be relocated to another appliance in the cluster, the virtual disks can non-disruptively be moved using the migration wizard in the PowerStore user interface. To prevent the VMware Distributed Resource Scheduler (DRS) from automatically migrating a virtual machine to non-optimal nodes in the cluster, the built-in VM/Host Rules within vCenter can be assigned to any virtual machine (see Figure 4).

Figure 4   Virtual machine affinity rules

3.4   Partition alignment

Due to the PowerStore architecture, manual partition alignment is not necessary.

3.5   Guest vSCSI adapter selection

When creating a new virtual machine, vSphere automatically suggests the disk controller option based on the operating system selected (see Figure 5). The Host Configuration Guide recommends using the VMware Paravirtual SCSI controller for optimal performance. More information about the Paravirtual adapter, including its benefits and limitations, can be found in VMware documentation.
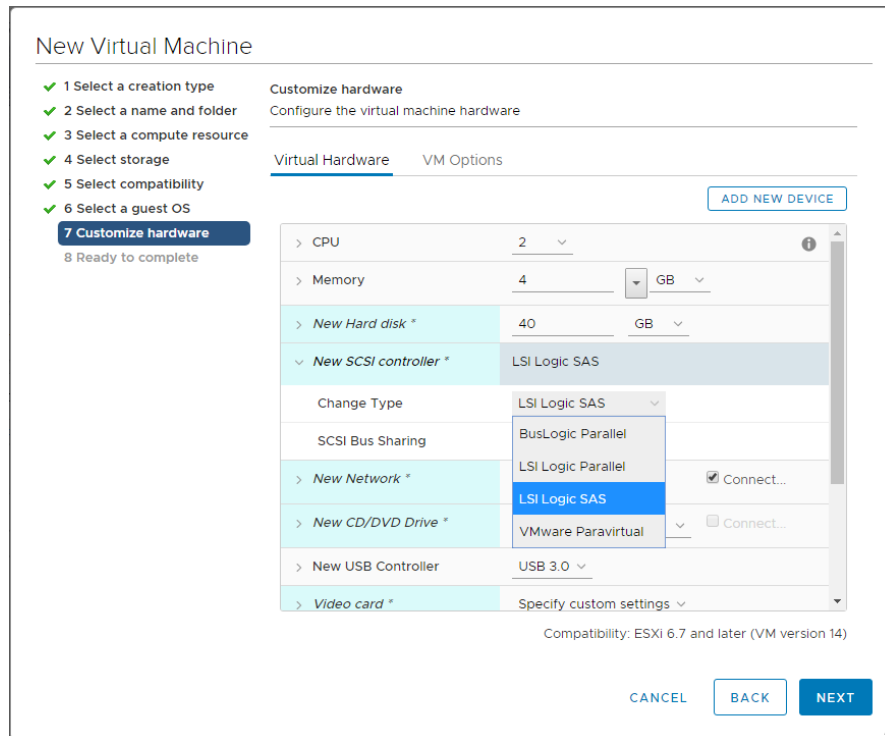

Figure 5   Selecting the virtual SCSI controller

3.6   Array offload technologies

VMware can offload storage operations to the array to increase efficiency and performance. This is performed by vStorage APIs for Array Integration (VAAI), a feature that contains four main primitives:

- Block zeroing: This primitive is primarily used for the ESXi host to instruct the storage to zero out eagerzeroedthick VMDKs.
- Full copy: Instead of the ESXi host performing the work of reading and writing blocks of data, this primitive allows the host to instruct the array to copy data which saves SAN bandwidth. This operation is typically used when cloning VMs.
- Hardware accelerated locking: This primitive replaces SCSI-2 reservations to increase VMFS scalability with changing metadata on VMFS datastores. With SCSI-2 reservations, the entire volume needed to be locked, and all other hosts in the cluster would need to wait while that ESXi host changed metadata. The hardware accelerated locking primitive allows a host to lock only the metadata on disk it needs, thus not hampering I/O from other hosts while the operation is performed.
- Dead space reclamation: This primitive uses the SCSI UNMAP command to release blocks back to the array that are no longer in use. For example, after deleting a VM, the ESXi host issues a series of commands to the PowerStore array to indicate that it is no longer using certain blocks within a volume. This capacity is returned to the pool so that it can be reused.

4 Management and monitoring

This section describes PowerStore features used to manage and monitor storage.

4.1 Mapping or unmapping practices

After a volume is created, mapping specifies the hosts that the PowerStore array presents storage to.

4.1.1 Cluster mappings

For ESXi hosts in a cluster, it is recommended to use host groups to uniformly present storage to all the initiators for reduced management complexity (see Figure 6 and Figure 7). This allows a volume or set of volumes to be mapped to multiple hosts at the same time and maintain the same logical unit number (LUN) across all hosts.
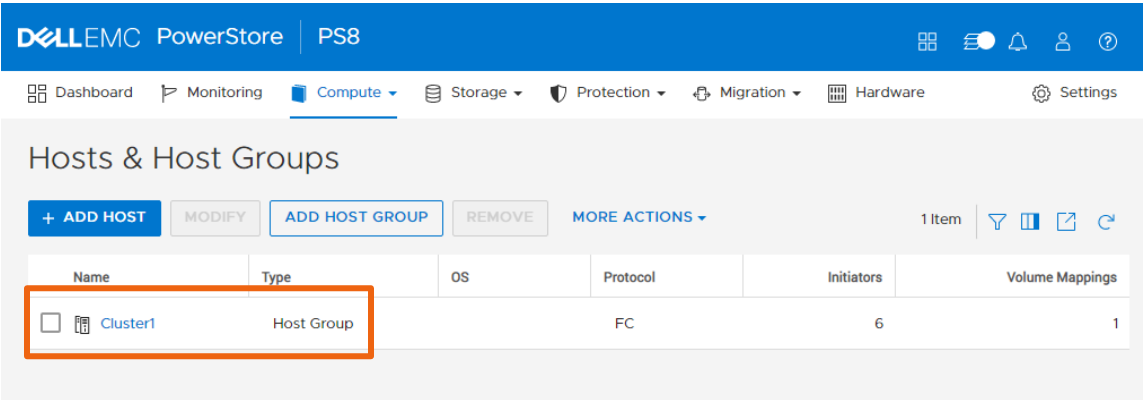

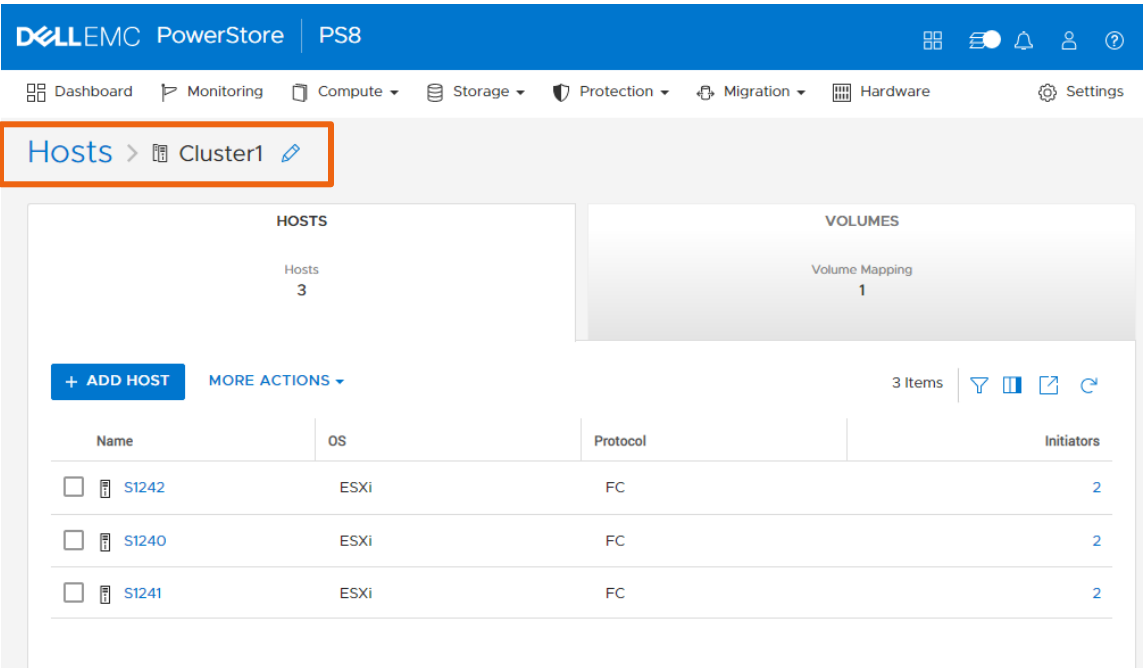Figure 6   Example of vSphere cluster containing multiple ESXi hosts

Figure 7   Host group details for the vSphere cluster showing three ESXi hosts

4.1.2 Properly unmapping volumes from ESXi hosts

If a VMFS datastore is no longer needed, the best practice is to unmount the datastore from the vCenter (Figure 8) and detach the disk devices from **each** host in the cluster (Figure 9), before unmapping and deleting the volume from the PowerStore Manager. This gracefully removes a datastore and can prevent an all paths down (APD) state from occurring.
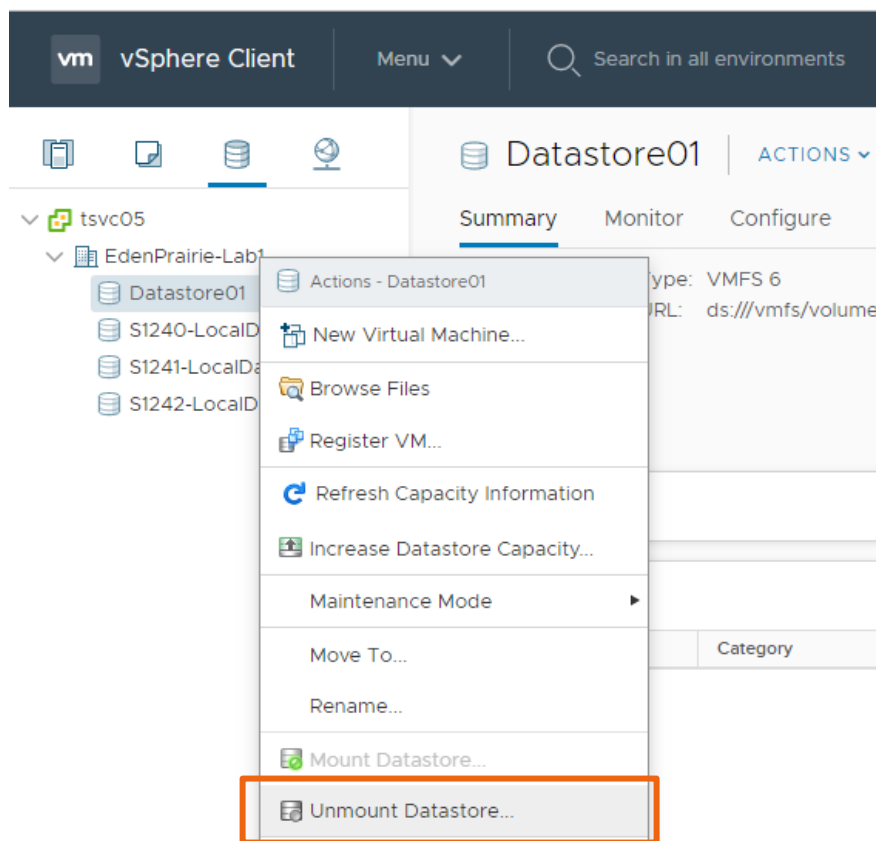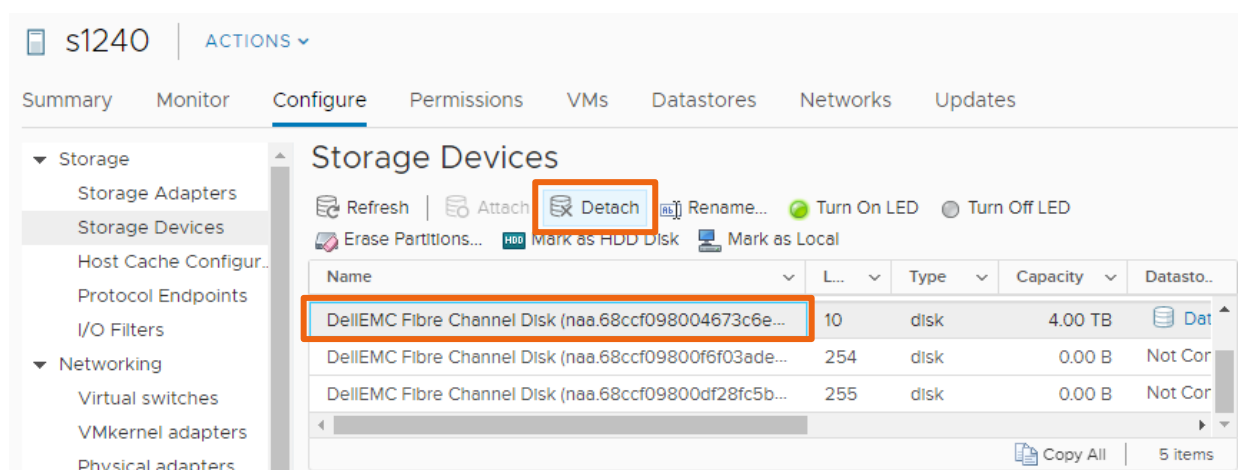

Figure 8   Unmount the datastore

Figure 9   Detach the disk device from a host

4.2 Thin clones

Thin clones make block-based copies of a volume or volume group and can also be created from a snapshot. Because the thin clone volume shares data blocks with the parent, the capacity usage of the child volume is mainly the delta changes from after it was created. Thin clones are advantageous in a vSphere environment because a VMFS datastore full of virtual machines can be duplicated for testing purposes, all while consuming less storage. For example, if a vSphere administrator needs to clone a multi-terabyte database server for a developer to run tests, the VM can be isolated and tested, and only consumes blocks that changed.

Within the PowerStore architecture, thin clones have several advantages for storage administrators:

- The thin clone can have a different data protection policy from the parent volume.
- The parent volume can be deleted, and the thin clones become their own resource.
- VMs can be cloned for testing monthly patches or development.

4.3 Data encryption

Data at rest encryption (D@RE) is enabled by default on the PowerStore array. No configuration steps are necessary to protect the drives.

4.4 Space reclamation

The VAAI dead space reclamation primitive is integrated into the array through the SCSI protocol. Depending on the version of ESXi the host is running, the primitive can automatically reclaim space.

VMFS-6 and ESXi versions that support automatic UNMAP: The best practice is to keep or reduce the reclamation rate to the **Low** or **100 MB/s**, which is the default setting (Figure 10).
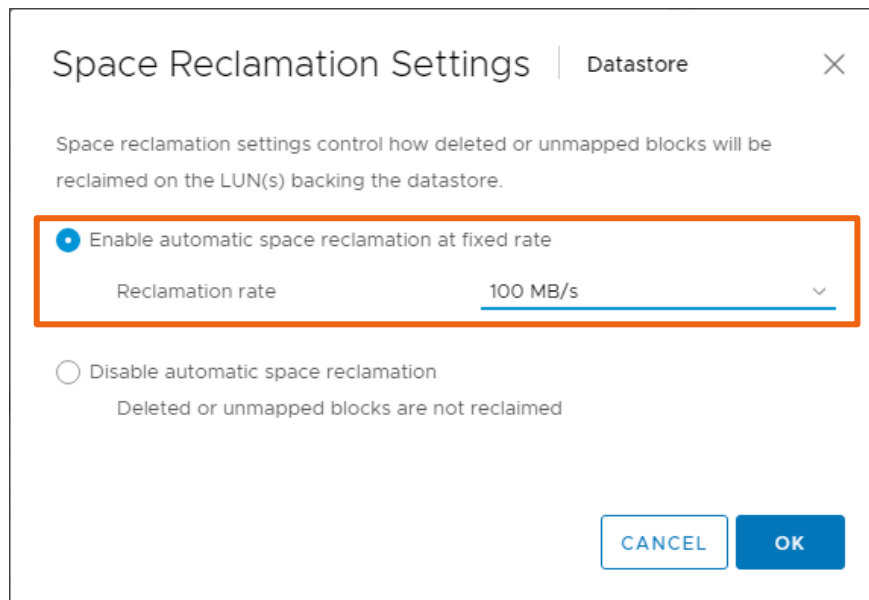

Figure 10   Setting the space reclamation rate

VMFS-5 and ESXi versions that do not support automatic UNMAP: The best practice is to set the reclamation rate to **200**.

```
esxcli storage vmfs unmap --volume-label=volume_label --reclaim-unit=200
```

Note: In certain older versions of ESXi, the dead space reclamation primitive must be invoked manually. Consult the VMware Knowledge Base for more information about which versions require additional steps.

4.5 VASA

VMware vStorage APIs for Storage Awareness (VASA) is a feature that allows vSphere hosts to gain insight into the storage types backing the datastores and gives vSphere the ability to manage storage. For example, the VASA 3.0 provider that is embedded into PowerStore allows it to manage vVols.

Note the following points about the PowerStore VASA provider:
- The provider is automatically registered on PowerStore X models.
- The provider can be manually/automatically registered on PowerStore T models.

    – Registering the VASA provider is required for vCenter integration with PowerStore.

- Before the vCenter server connection is created, the VASA provider details must be entered on the vCenter Storage Providers screen within vCenter (Figure 11).

    a. In PowerStore Manager, create an account that is named **vmadmin** and assign the VMAdmin role.
    b. In vCenter, add the storage provider specifying a name, vmadmin credentials, and the URL in the format: **https://<*management IP address*>:8443/version.xml**
    c. Do **not** select **Use storage provider certificate**.
    d. For more information, follow the instructions in PowerStore Manager help by searching for **vCenter Server Connection**.
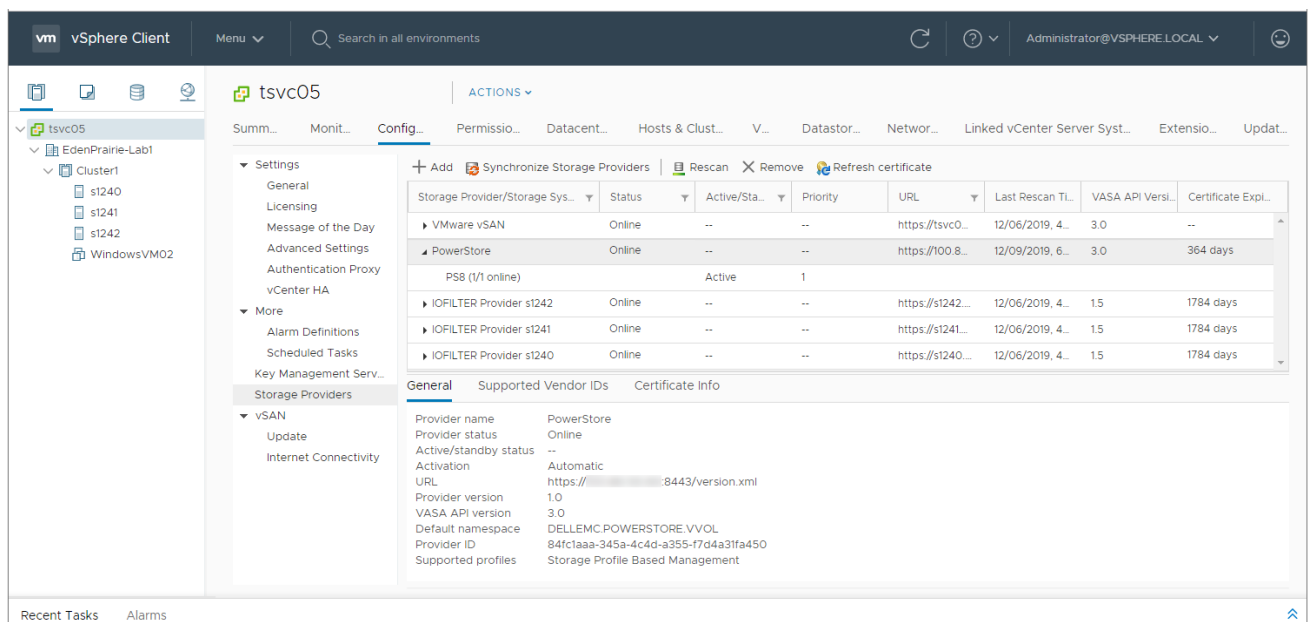

Figure 11   PowerStore VASA provider registration

4.6 Virtual Volumes

VMware vSphere Virtual Volumes (vVols) are used as the primary storage mechanism for virtual machines running on PowerStore X model appliances internal nodes and can be used by external ESXi hosts from either PowerStore X or T models. vVols is a new storage methodology that runs on top of existing storage protocols such as Fibre Channel and iSCSI, that allows administrators more granular control over virtual machines regarding performance, snapshots, and monitoring.

One of the key features of vVols is that it allows administrators to use storage policy-based management (SPBM) for their environment. This provides the ability to align application needs with the appropriate storage resources in an automated manner.

Since vCenter is required for binding and unbinding vVols from the protocol endpoints during power-on, power-off, and other operations such as vMotions, vCenter should be treated as a tier 1 application.

Caution: The VMware vCenter virtual appliance must never be migrated to a vVol datastore or storage container. Since vCenter is required for bindings to power on vVol-based virtual machines, this action may prevent powering on vCenter after the VM is shut down or has experienced an unplanned outage.

Note: As of VMware vSphere7.0 U1, vVols are not currently supported with NVMe-FC.

4.7 Scripting and automation

The PowerStore platform has a REST API to automate management tasks. Find more information at www.dell.com/support.

4.8 Dell EMC Virtual Storage Integrator

Another tool for storage management is the Dell EMC Virtual Storage Integrator (VSI) which is the vSphere web client plug-in for PowerStore and other Dell EMC storage products. The VSI appliance is downloaded as an OVA and is installed into the virtual infrastructure where it is given an IP address and added to vCenter as part of the installation process. This plug-in allows administrators to create datastores, expand datastores, apply ESXi host best-practice settings, perform capacity monitoring, and more.
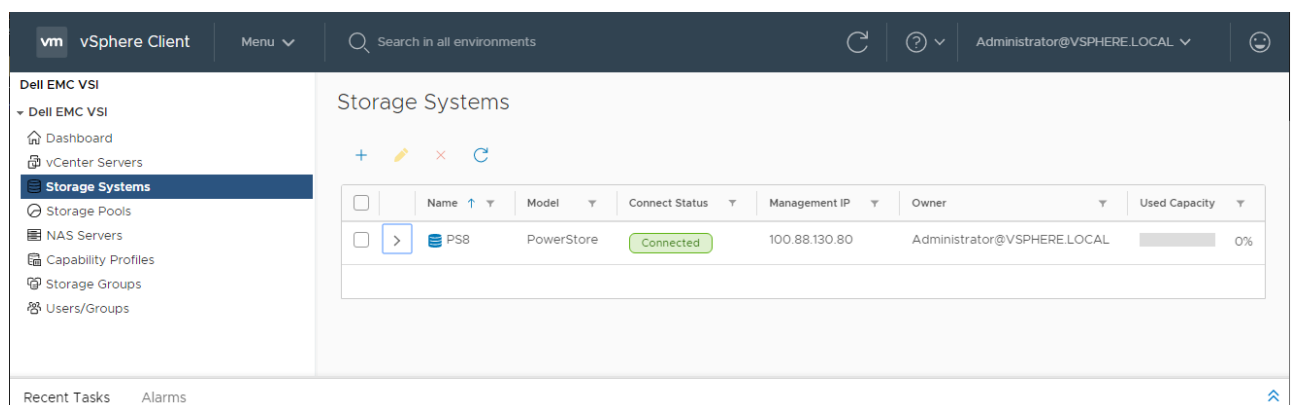

Figure 12   Dell EMC Virtual Storage Integrator vSphere web client plug-in

4.9 vRealize Orchestrator

Dell EMC PowerStore also supports VMware vRealize® Orchestrator™ (vRO) as a plug-in. vRO enables administrators to automate many common workflows with PowerStore appliances. Download the vRO plug-in for PowerStore at Dell Support.
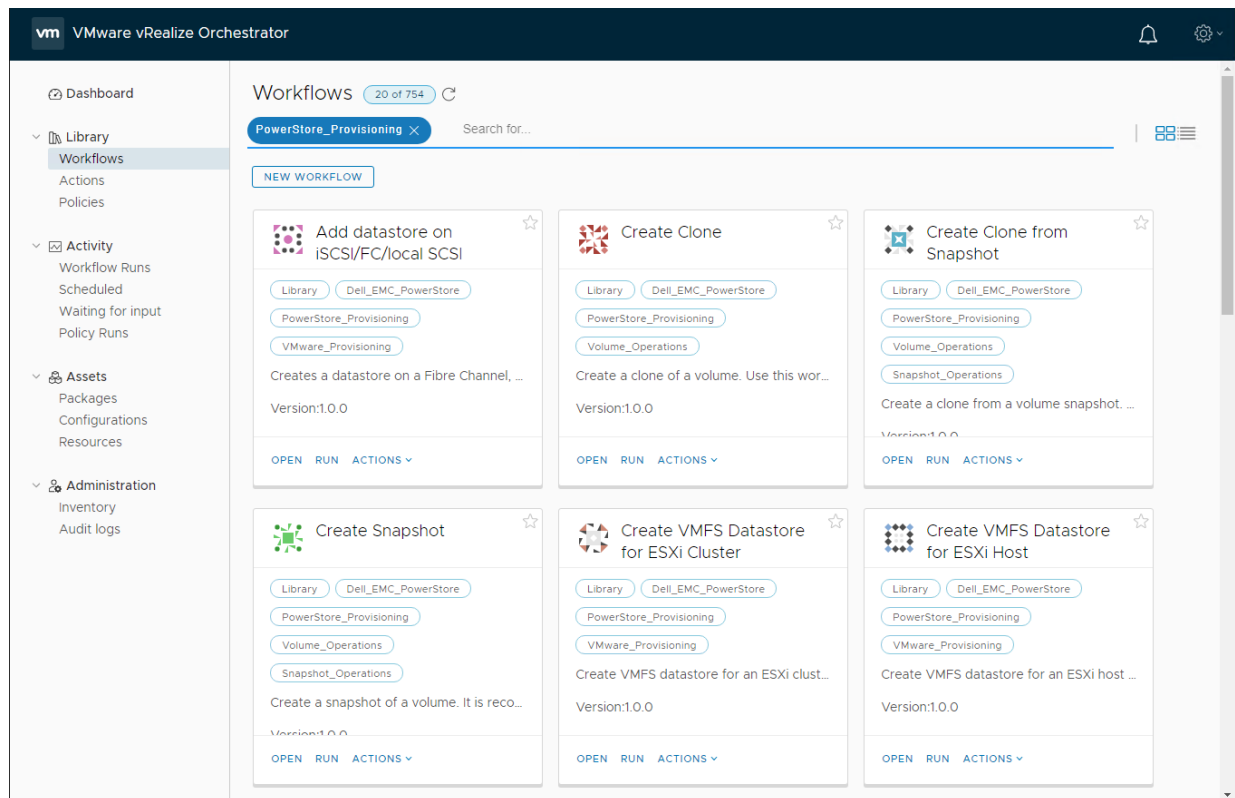

Figure 13   Dell EMC PowerStore vRealize Orchestrator plug-in example

5   Data protection and disaster recovery

PowerStore has integrated snapshot and replication capabilities to protect data, and it is policy driven for ease of administration.

5.1   Snapshots and recoveries

To automate and simplify protecting data, PowerStore uses protection policies. These are a set of snapshot and replication rules that are applied to a volume or group of volumes. Protection policies help protect data, set retention policies, and help guarantee recovery point objectives (RPOs) for an organization.

In addition, protection policies can be applied to individual volumes or to volume groups. When a protection policy is applied to a volume group, it allows multiple volumes to have snapshots taken, to be replicated, or to be recovered, simultaneously. This allows protecting complex applications that are interdependent and span across multiple volumes.

vVol snapshots can be taken from either the PowerStore Manager or the vCenter client, but they are inherently managed by vCenter. When virtual-machine snapshots are taken from the vCenter client, the best practice is to disable the option for virtual machine memory which may increase snapshot time significantly.

5.2   Snapshots and options for application backup and restore

Using array-based snapshots is an effective way to protect virtual-machine data and establish an RPO. In the PowerStore architecture, the snapshot schedule is created using protection policies. Each protection policy can define snapshot rules to establish a schedule and retention, and replication rules to specify a destination array and RPO.
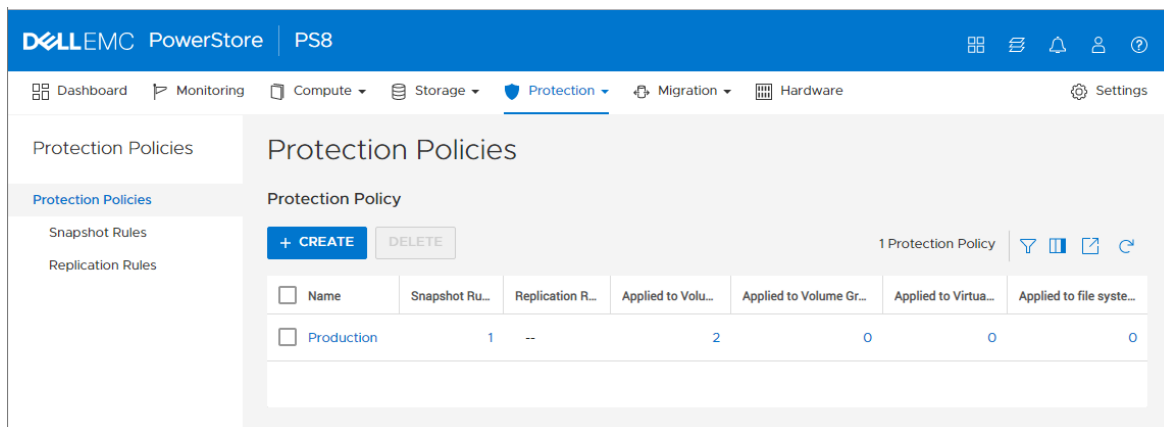

Figure 14   Protection policies screen in PowerStore Manager

PowerStore has three data-recovery mechanisms that behave differently depending on the usage scenario (see Figure 15).

- Snapshot: These represent a specific point in time for data stored on a volume that can be used for recoveries to refresh, restore, or create thin clones.
- Thin clone: This takes an existing snapshot from a parent volume and creates a child volume from that point in time.
- Refresh: This allows snapshot data to replace existing data in the volume. The existing data is removed, and snapshot data from the new source is copied to it in-place. A parent volume can refresh a child, and a child can refresh a parent.
- Restore: The restore operation replaces the contents of a parent storage resource with data from an associated snapshot. Restoring resets the data in the parent storage resource to the point in time the snapshot was taken.

Caution: Use of the refresh and restore operations on active virtual-machine volumes may cause unexpected results and behaviors. All host access to the volume must be ceased before attempting these operations.
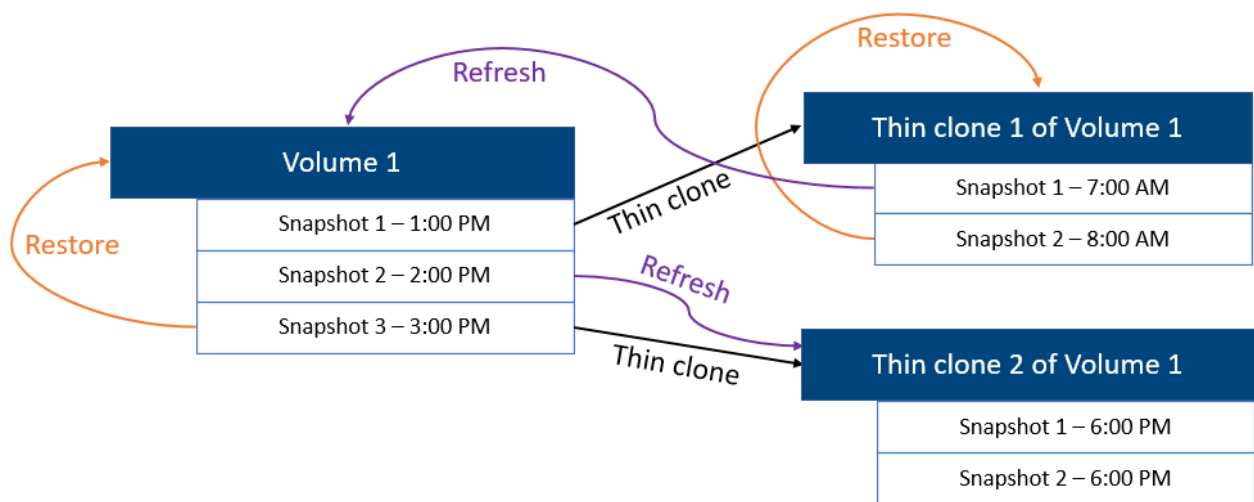

Figure 15   PowerStore snapshot and recovery

If a virtual machine residing on a VMFS datastore needs recovery, the best practice is to create a thin clone from a snapshot. The high-level steps are as follows:

1. From PowerStore Manager, create a thin clone from a snapshot and present it to the vSphere cluster.
2. From the vSphere client, rescan the storage, add a new datastore, select the newly created volume, and assign a new signature (see Figure 16).
3. Register the VM from the snap-xxxxxxxx-*originaldatastorename* datastore.
4. Use Storage vMotion to migrate the virtual machine back to the original datastore, if desired.
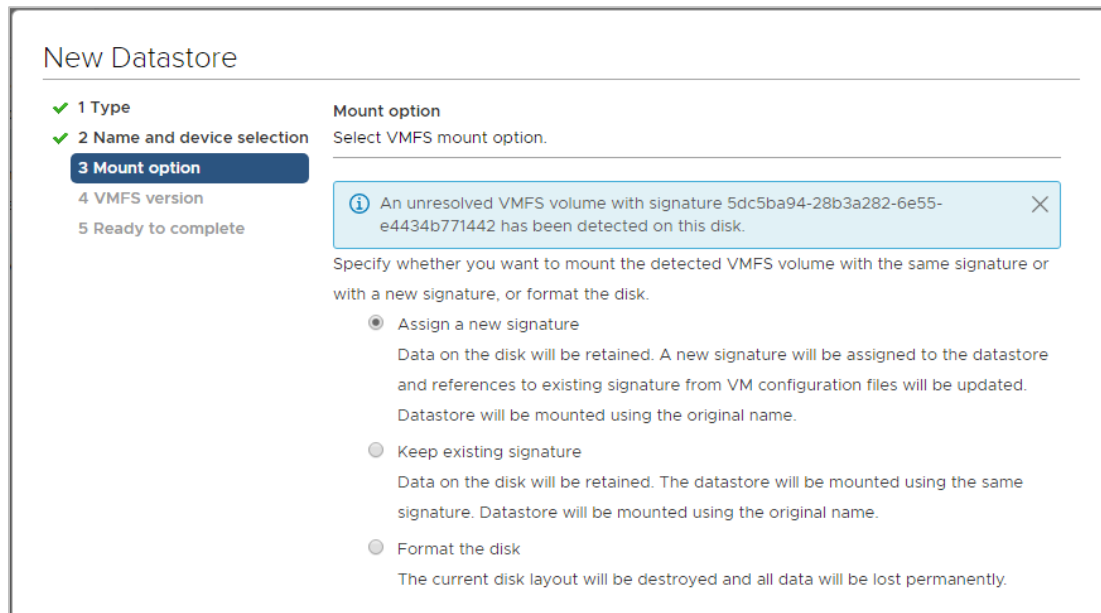

Figure 16   New datastore wizard > assign a new signature

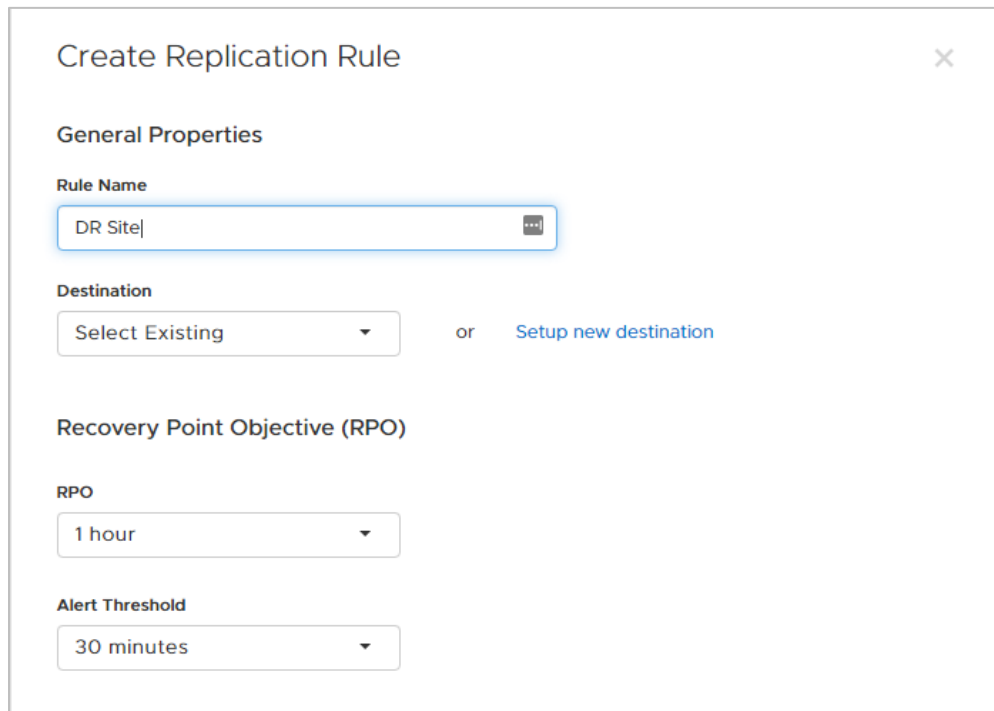## 5.3   Crash consistent and application consistent snapshots

When taking array-based snapshots of virtual machines, it is important to remember that snapshots taken without application coordination are considered **crash consistent**. Crash consistency is the storage term for data that has a snapshot taken in-progress without application awareness. While most modern applications can recover from crash consistent data, their recovery can yield varying levels of success. For example, when recovering a Microsoft Windows virtual machine, as the operating system boots, it responds as if it has encountered an unexpected power-loss event and can potentially check the disk (chkdsk) on startup.

Application consistent snapshots are supported by products such as Dell EMC AppSync™. This enables coordination between the array and the application to help assure that the data is quiesced, the caches are flushed, and the data is preserved in a known good state. Application consistent snapshots such as these offer a higher probability of recovery success.

Note: When taking managed snapshots such as with vVols from PowerStore Manager, virtual machine memory is not included. When performing snapshots from vCenter, it is recommended not to include virtual machine memory.

5.4 Replication and remote recovery

PowerStore offers asynchronous replication to transfer data to one or more secondary arrays. When the secondary array is stored at a different location, this can help to protect virtual machine data from localized geographical disasters. Replication RPOs and options are set within protection policies (Figure 17).

Figure 17   Creating a replication rule

A   Technical support and resources

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

Storage technical documents and videos provide expertise that helps to ensure customer success on Dell EMC storage platforms.

The PowerStore Info Hub provides detailed documentation on how to install, configure, and manage Dell EMC PowerStore systems.

A.1   Related resources

- VMware Documentation
- VMware Hardware Compatibility List (HCL)