TECHNICAL WHITE PAPER:
February 2024

VMware vSAN™ Architecture Overview & Setup

Proof of Concept (PoC) Guide

Table of contents

Introduction 4

Overview 4

What's New in vSAN 8.0 4

Hardware Selection 5

General Approach to Testing vSAN 5

Prerequisites 7

Hardware Compatibility 7

vSAN ESA or OSA 7

vSAN ESA ReadyNode™ & ReadyNode Emulated Configuration Guidance 7

vSAN OSA Hardware Guidance 9

vSAN Setup Assumptions 10

vSAN OSA All-Flash or Hybrid 10

Enabling a Single vSAN HCI Cluster 11

Using Quickstart to Enable Single vSAN HCI Cluster 11

Manually Enabling vSAN Services on a Cluster 24

Enabling vSAN Services on a VMware Cloud Foundation™ based Cluster 24

Enabling vSAN Max™ - Disaggregated Storage 26

vSAN Max Sizing Considerations 26

Disaggregated Storage for vSAN OSA (AKA: HCI Mesh) 26

Using Quickstart to Enable vSAN Max Cluster 27

Manually Enabling vSAN Max on a Cluster 32

Enabling vSAN Max/HCI Mesh Services on a VMware Cloud Foundation™ based Cluster 32

Enabling vSAN Compute Cluster 33

Overview 33

Enabling vSAN Compute Clusters in VMware Cloud Foundation 33

Enabling Two-Node and vSAN Stretched Clusters 34

Overview 34

Enabling vSAN Stretched Clusters in VMware Cloud Foundation 34

Check the vSAN Cluster Thoroughly 35

Why Is This Important? 35

Use Health Check to Verify vSAN Functionality 35

Manually Checking against the VCG 40
vSAN Basics 42
Deploy your first Virtual Machine 42
Verifying Disk Layout of a VM on vSAN 48
Configuring Fault Domains 52
Creating a Snapshot 53
Clone a Virtual Machine 58
vMotion a Virtual Machine Between Hosts 61
VMware vSphere® Storage vMotion® (Storage vMotion) VM Between Datastores 65
VM Storage Policies and vSAN 67
Create a New VM Storage Policy 68
Deploy a new VM with a New Storage Policy 72
Assign a new Storage Policy to an Existing VM 76
Modify a VM Storage Policy 79
IOPS Limits 82
Space Efficiency & Encryption Defined by Policy (vSAN ESA) 82
APPENDIX A: vSAN RDMA Configuration & Troubleshooting 84
Overview 84
Configuration Example 85
RDMA troubleshooting 88
APPENDIX B: Cleanly Removing vSAN Configuration 91
vCLS Retreat Mode 91
Remove vSAN Partitions and Clear Data 91

Introduction

Overview

This document primarily details the overall architecture and setup of vSAN Express Storage Architecture™ (ESA) cluster environments. vSAN Original Storage Architecture™ (OSA) environments are covered where they differ from vSAN ESA. The specific focus of the document includes overviews of:

- Single vSAN HCI cluster deployments (ESA and OSA)
- vSAN Max™ cluster deployments (disaggregated storage)
- vSAN Compute cluster deployments
- vSAN Two-Node and Stretched Cluster deployments.

Additionally, the document includes reviews of basic functions such as vSAN storage policies and virtual machine deployment.

Separate guides are dedicated to more detailed discussions of specific vSAN related areas:

- vSAN Proof of Concept: vSAN Management, Monitoring & Hardware Testing
- vSAN Proof of Concept: vSAN Features
  - Space efficiency features (e.g., compression, deduplication, RAID-5/RAID-6 erasure coding, and Trim/Unmap)
  - Encryption
  - File Services
- vSAN Proof of Concept: vSAN Performances Testing
- vSAN Proof of Concept: vSAN Stretched Cluster & Two-Node Overview & Testing

What's New in vSAN 8.0

vSAN 8.0 is a major milestone release that includes a new architecture option for new vSAN clusters, and there are now two distinct ways to configure a vSAN cluster. vSAN 8.0 introduced ESA, which optimizes the use of certified NVMe drives in a single tier, with a dynamic approach to data placement. vSAN 8.0 Update 2 introduces Max. A disaggregated deployment model, providing scalable, cost-effective storage solutions to empower your cloud infrastructure. This innovative approach caters to three primary use cases: driving cost optimization for infrastructure and applications, ensuring operational simplicity with unified storage, and facilitating rapid scaling for cloud-native applications. The classic OSA model, which is updated from vSAN 7, remains an option.

For a comprehensive overview of ESA, visit: https://core.vmware.com/vsan-esa.

Updates to the OSA include:

- Change in buffer device capacity, increasing from 600GB to 1.6TB.
- For HCI Mesh, the number of server and client clusters increases from 5 to 10
- Improvements in vSAN File Services operations
- vSAN Boot time optimizations

Hardware Selection

Choosing the appropriate hardware is one of the most important factors in the successful validation of vSAN. Hardware recommendations will differ between ESA and OSA deployments.

There are many variables with hardware (drivers, controller firmware versions) so be sure to choose hardware that is on the VMware Compatibility List. See the 'Prerequisites' section of this guide for more information.

General Approach to Testing vSAN

Once the appropriate hardware is selected for testing, it is useful to define the use case, goals, expected results and success metrics.

The testing lifecycle can be broken into three phases:

Day-0

The post-design phase, installation of the infrastructure components, including: the hypervisor VMware ESXi™ (ESXi); control plane VMware vCenter®(vCenter); physical network uplinks and upstream network devices; essential services, such as DNS and NTP; VLANs, etc.

Day-1

Setup and configuration of the required solution (in the case of vSAN).

Day-2

Operational aspects and monitoring. The most important aspects to validate when testing:

- Successful vSAN configuration and deployment
- VMs successfully deployed to vSAN Datastore
- Reliability: VMs and data remain available in the event of failure (host, disk, network, power)
- Serviceability: Maintenance of hosts, disk groups, disks, clusters
- Performance: vSAN and selected hardware can meet the application, as well as business needs
- Validation: vSAN features are working as expected (File Service, Deduplication and Compression, RAID-5/6, checksum, encryption)
- Day-2 Operations: Monitoring, management, troubleshooting, and upgrades

These can be grouped into three common types: resiliency testing, performance testing, and operational testing.

Resiliency Testing
Operational Testing
Performance Testing

Resiliency Testing

As with any storage solution, failures can occur on hardware components at any time due to age, temperature, firmware, etc. Such failures can occur among storage controllers, disks, nodes, and network devices among other components. As a software solution, vSAN is designed to be resilient against these failures. In this guide, we will examine how to systematically test against disk, host, network, and control plane failures.

Operational Testing

Understanding how the solution behaves during normal (or "day two") operations is important to consider as part of the evaluation. Fortunately, because vSAN is embedded within the ESXi hypervisor, many vSAN operations tasks are simply extensions of normal VMware vSphere® (vSphere) operations. Adding hosts, migrating VMs between nodes, and cluster creation are some of the many operations that are consistent between vSphere and vSAN.

Performance Testing

Before embarking on testing, it is important to set clear objectives and understand the performance requirements of the environment. Close attention to details such as workload I/O profiles, latency and hotspots is needed. In this guide, we will explore how to conduct performance tests with a consistent, methodological approach.

Prerequisites

Hardware Compatibility

Plan on testing a reasonable hardware configuration resembling a production-ready environment that suits your business requirements. Refer to the VMware vSAN Design and Sizing Guide for information on design configurations and considerations when deploying vSAN.

As vSAN is a software solution, it is critical to ensure that well supported, enterprise class hardware components are used. The VMware Compatibility Guide (or "VCG") lists components that have been tested and certified for use with vSAN. In addition, the vSAN Hardware Compatibility List (or "HCL") lists hardware compatibility specific to vSAN.

Note: The terms "VCG" and "HCL" may be used interchangeably (both within this guide, and in other documentation), but essentially pertain to hardware compatibility.

Note: If using a VMware vSAN ReadyNode™ (ReadyNode) or appliance, the factory-installed hardware is guaranteed to be compatible with vSAN. However, be mindful that BIOS updates and firmware and device driver versions may be out of date and should be checked.

vSAN ESA or OSA

If using a certified ReadyNode or vSAN appliances with specific certified devices, then vSAN ESA (which has an optimized data path and placement for NVMe devices) is an option.

The table below summarizes the minimum requirements for each architecture:

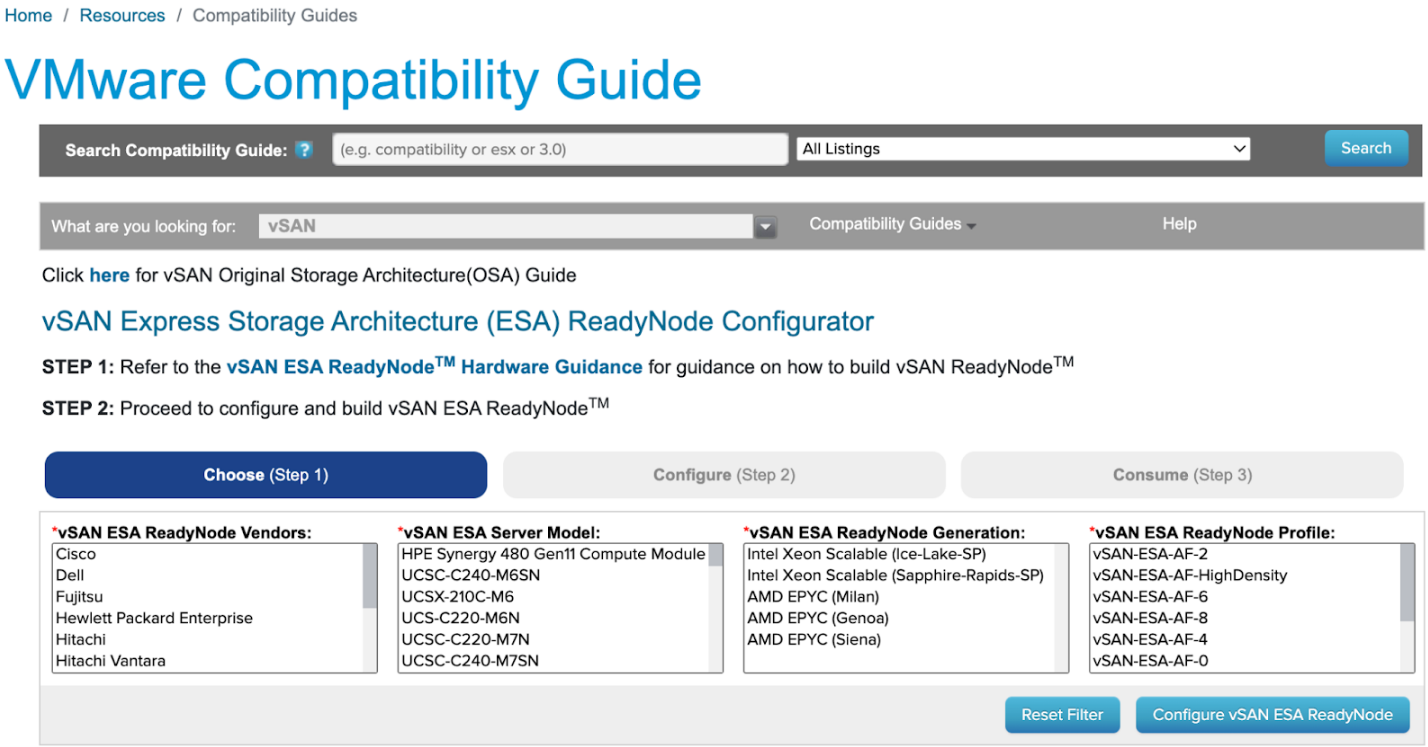
	vSAN 8.0 ESA	vSAN 8.0 OSA
Storage Device Minimums	4 devices per host	1 capacity device + 1 cache device per host
Hardware Support	ReadyNodes, Appliances, build your own with certified devices Only vSAN ESA certified NVMe devices	ReadyNodes, Appliances, build your own with certified devices Any SATA, SAS. NVMe certified device
Networking Requirements (minimum)	25Gbps	10Gbps

Note that 25Gbps networking (or faster) is required vSAN ESA. The one exception is the vSAN-ESA-AF-0 ReadyNode configuration, which allows for 10Gbps networking. It is designed for Edge, 2-node, and other small environments with relatively few VMs, and minimal workload demands. For more information, please refer to: https://core.vmware.com/blog/esa-all-platforms-and-all-workloads-esa-af-0

vSAN ESA ReadyNode™ & ReadyNode Emulated Configuration Guidance

Customers can pick a server manufacturer's ESA certified ReadyNode platform listed on the VMware Compatibility Guide (VCG) for vSAN ESA, and can build out a server configuration using hardware components as they are listed on the given ReadyNode specification, emulating a ReadyNode purchased using a single SKU. This approach can help customers who chose not to purchase a ReadyNode through an official SKU but have the same or better hardware found in the desired ReadyNode classification.

Here is the direct link to the VMware Compatibility Guide vSAN ESA page: https://www.vmware.com/resources/compatibility/search.php?deviceCategory=vsanesa



Home / Resources / Compatibility Guides

VMware Compatibility Guide

Search Compatibility Guide: (e.g. compatibility or esx or 3.0) All Listings

What are you looking for: vSAN

Click here for vSAN Original Storage Architecture(OSA) Guide

vSAN Express Storage Architecture (ESA) ReadyNode Configurator

STEP 1: Refer to the vSAN ESA ReadyNode^{TM} Hardware Guidance for guidance on how to build vSAN ReadyNode^{TM}

STEP 2: Proceed to configure and build vSAN ESA ReadyNode^{TM}

*vSAN ESA ReadyNode Vendors:	*vSAN ESA Server Model:	*vSAN ESA ReadyNode Generation:	*vSAN ESA ReadyNode Profile:
Cisco	HPE Synergy 480 Gen11 Compute Module	Intel Xeon Scalable (Ice-Lake-SP)	vSAN-ESA-AF-2
Dell	UCSC-C240-M6SN	Intel Xeon Scalable (Sapphire-Rapids-SP)	vSAN-ESA-AF-HighDensity
Fujitsu	UCSX-210C-M6	AMD EPYC (Milan)	vSAN-ESA-AF-6
Hewlett Packard Enterprise	UCS-C220-M6N	AMD EPYC (Genoa)	vSAN-ESA-AF-8
Hitachi	UCSC-C220-M7N	AMD EPYC (Siena)	vSAN-ESA-AF-4
Hitachi Vantara	UCSC-C240-M7SN		vSAN-ESA-AF-0

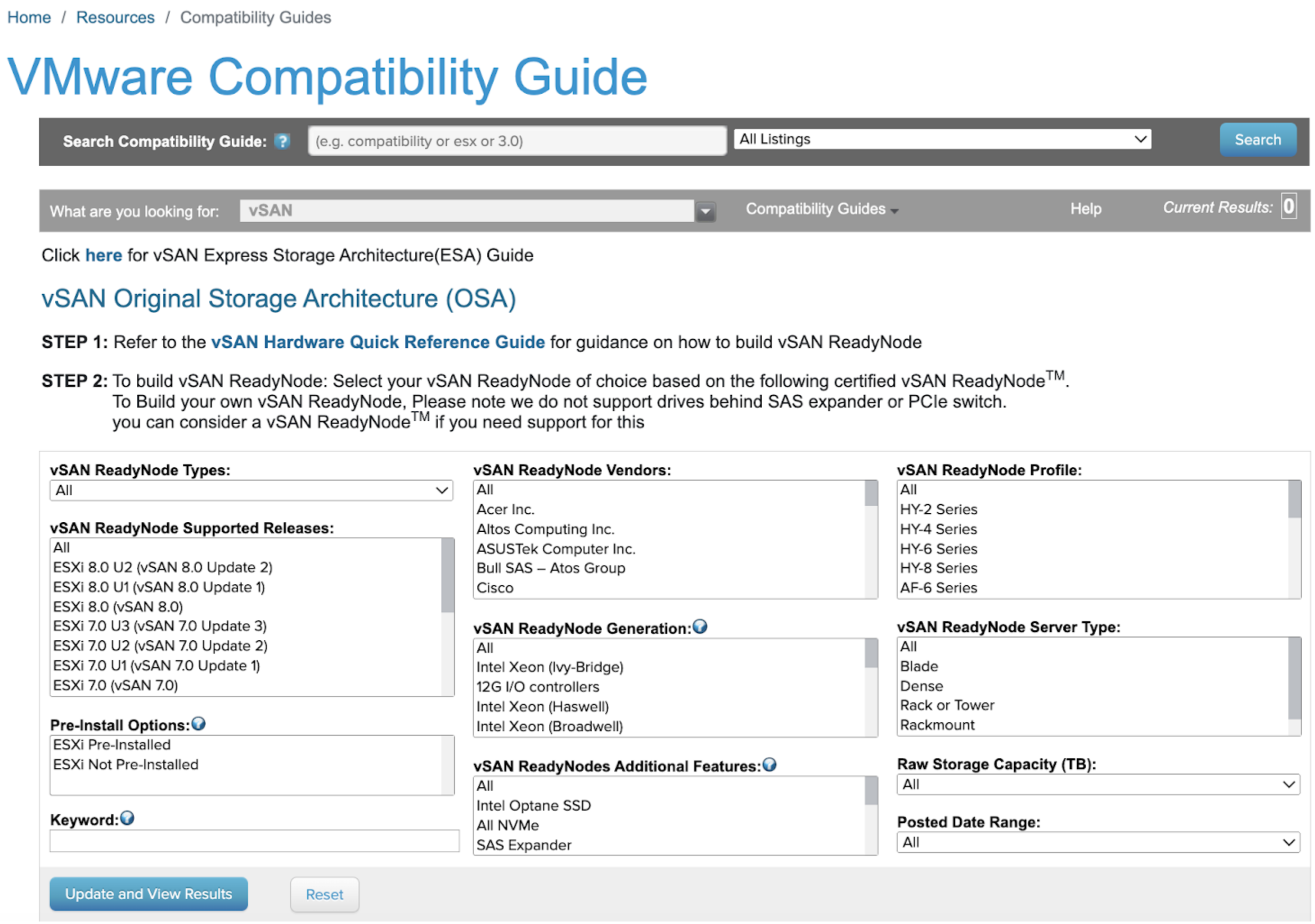
For further vSAN ESA hardware guidance, see:

- vSAN ESA ReadyNode Hardware Guidance - https://www.vmware.com/resources/compatibility/vsanesa_profile.php
- KB article on vSAN ESA architecture - https://kb.vmware.com/s/article/90343
- Support for ReadyNode Emulated Configurations in vSAN ESA Blog - https://core.vmware.com/blog/support-readynode-emulated-configurations-vsan-esa

vSAN OSA Hardware Guidance

As with ESA deployments, OSA supports both prescriptive ReadyNode and customized architectures. Below is the direct link to the VMware Compatibility Guide vSAN OSA page:

https://www.vmware.com/resources/compatibility/search.php?deviceCategory=vsanosa



Home / Resources / Compatibility Guides

VMware Compatibility Guide

Search Compatibility Guide: (e.g. compatibility or esx or 3.0) All Listings Search

What are you looking for: vSAN Compatibility Guides Help Current Results: 0

Click **here** for vSAN Express Storage Architecture(ESA) Guide

vSAN Original Storage Architecture (OSA)

STEP 1: Refer to the **vSAN Hardware Quick Reference Guide** for guidance on how to build vSAN ReadyNode

STEP 2: To build vSAN ReadyNode: Select your vSAN ReadyNode of choice based on the following certified vSAN ReadyNode^{TM}.
To Build your own vSAN ReadyNode, Please note we do not support drives behind SAS expander or PCIe switch.
you can consider a vSAN ReadyNode^{TM} if you need support for this

vSAN ReadyNode Types: All

vSAN ReadyNode Supported Releases:
All
ESXi 8.0 U2 (vSAN 8.0 Update 2)
ESXi 8.0 U1 (vSAN 8.0 Update 1)
ESXi 8.0 (vSAN 8.0)
ESXi 7.0 U3 (vSAN 7.0 Update 3)
ESXi 7.0 U2 (vSAN 7.0 Update 2)
ESXi 7.0 U1 (vSAN 7.0 Update 1)
ESXi 7.0 (vSAN 7.0)

Pre-Install Options:
ESXi Pre-Installed
ESXi Not Pre-Installed

Keyword:

vSAN ReadyNode Vendors:
All
Acer Inc.
Altos Computing Inc.
ASUSTek Computer Inc.
Bull SAS – Atos Group
Cisco

vSAN ReadyNode Generation:
All
Intel Xeon (Ivy-Bridge)
12G I/O controllers
Intel Xeon (Haswell)
Intel Xeon (Broadwell)

vSAN ReadyNodes Additional Features:
All
Intel Optane SSD
All NVMe
SAS Expander

vSAN ReadyNode Profile:
All
HY-2 Series
HY-4 Series
HY-6 Series
HY-8 Series
AF-6 Series

vSAN ReadyNode Server Type:
All
Blade
Dense
Rack or Tower
Rackmount

Raw Storage Capacity (TB): All

Posted Date Range: All

For further vSAN OSA hardware guidance, see:

- vSAN OSA Hardware Quick Reference Guide - https://www.vmware.com/resources/compatibility/vsan_profile.html?locale=
- KB article on vSAN OSA architecture - https://kb.vmware.com/s/article/52084

vSAN Setup Assumptions

This document assumes a generic vSphere deployment (one vCenter and at least four ESXi hosts). This document does not assume the test environment is deployed using the VMware Cloud Foundation™ (VCF) model. If the VCF model is used, there may be an impact to the information in this guide. Those impacts are noted and references to the proper VCF versions are provided.

The list below details the setup requirements prior to using this guide:

- Hosts with supported hardware and firmware (see above)
- All hosts with ESXi 8.0 Update 2 (build number 22380479) or newer deployed
- vCenter Server 8.0 Update 2 (build number 22617221) or newer has been deployed to manage these ESXi hosts
- DNS and NTP services are available
- The hosts should not be added to any cluster
- If using the Quickstart workflow
  - Each host should have a management network VMkernel
  - VMware vSphere® vMotion (vMotion) and vSAN networks will be configured during the workflow
- If not using the Quickstart workflow, each host should have a management and a vMotion network VMkernel port already configured
- A set of IP addresses
  - One per ESXi host will be needed for the vSAN network VMkernel ports (the recommendation is that these are all on the same VLAN and IPv4 or IPv6 network segment)
  - One per ESXI will be needed for the vMotion network VMkernel port (the recommendation is that these are all on the same VLAN and IPv4 or IPv6 network segment)
- If possible, configure internet connectivity for vCenter such that the HCL database can be updated automatically. (Internet connectivity is also a requirement to enable Customer Experience Improvement Program or CEIP)
- Optionally, for the purposes of testing Storage vMotion operations, an additional datastore type (such as NFS or VMFS) should be presented to all hosts

vSAN OSA All-Flash or Hybrid

When reviewing whether to deploy all-flash or hybrid configurations (all-flash recommend), there are several factors to consider:

- All-Flash vSAN requires at least a 10Gb Ethernet network
- Flash devices are used for both cache and capacity
- Deduplication and Compression are space-efficiency features available in all-flash configuration and not available with hybrid configuration
- Erasure Coding (RAID 5/ RAID 6) is a space efficiency feature available on all-flash configuration only
- Flash read cache reservation is not used with all-flash configurations; reads come directly from the capacity tier SSDs
- Endurance and performance classes now become important considerations for both cache and capacity layers

Enabling a Single vSAN HCI Cluster

Using Quickstart to Enable Single vSAN HCI Cluster

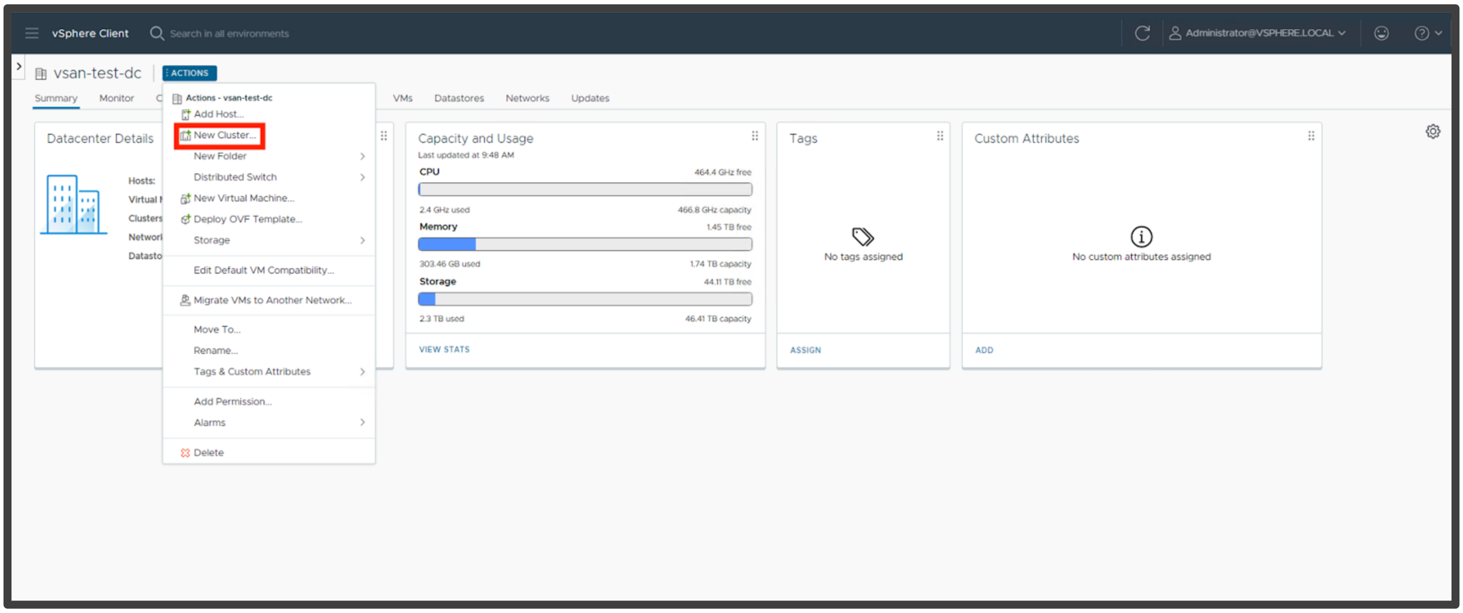
Follow this section to configure a single vSAN HCI cluster via the Quickstart process (recommended).

Note:

- If you wish to manually setup a vSAN HCI Cluster, please refer to the Manually Enabling vSAN Services on a Cluster section
- If you deployed vSAN using the bootstrap feature of vCenter deployment, you will not be able to use Quickstart to configure your environment
- vSAN Max deployments, disaggregated storage, requires a specific vSAN ESA deployment, if you plan to test vSAN Max, please skip this section a go directly to the Enabling vSAN Max - Disaggregated Storage Section
- ESA and OSA HCI deployment steps are similar
- The walkthrough below calls out where OSA and ESA steps may differ.

Initialize Cluster

Navigate to your **Datacenter** > Click **Actions** > **New Cluster**.



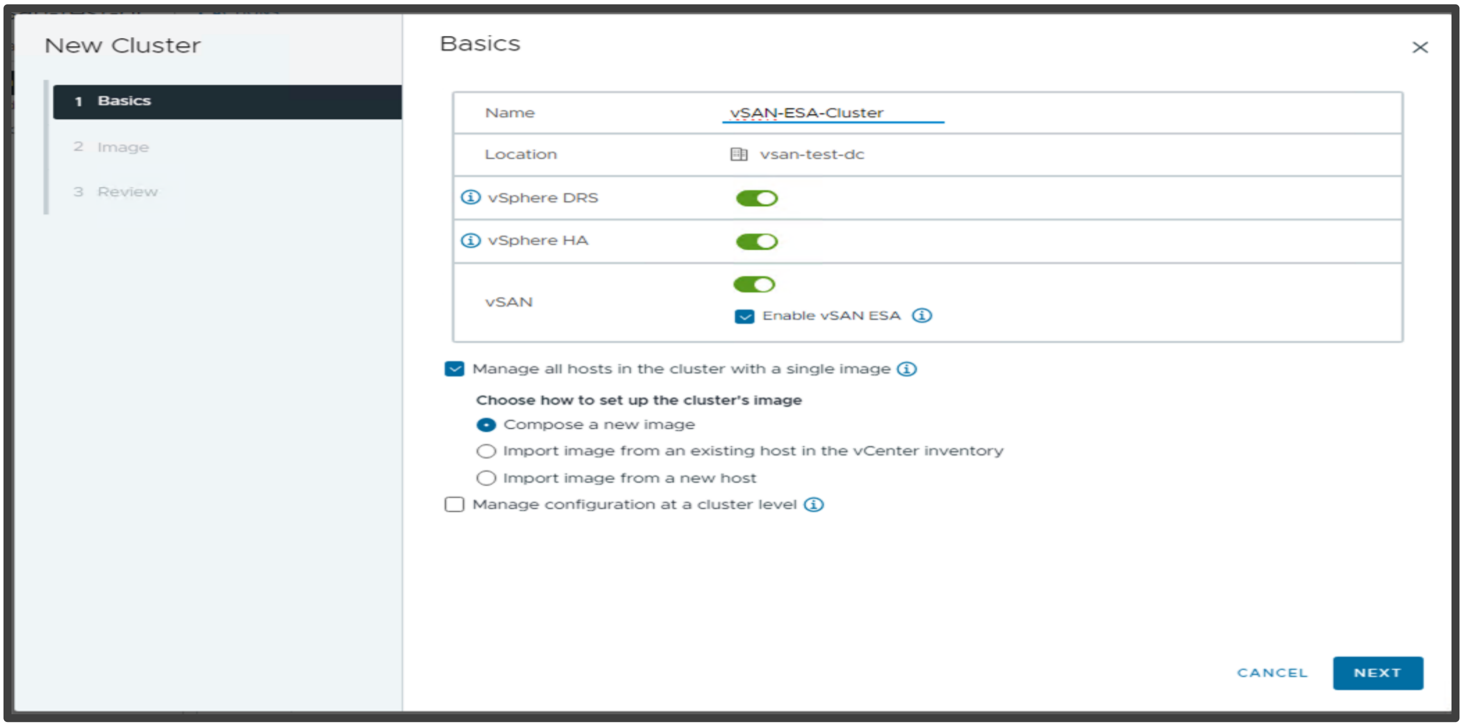
The New Cluster screen pops-up and we are presented with a dialog to enable services. Provide a name for the cluster and select vSAN from the list of services. We are also able to enable vSAN ESA (the default).

Note: Once the cluster is created with the ESA flag, it cannot be changed unless the cluster is re-created.

Be aware that:

- If targeting an OSA deployment, the “Enable vSAN ESA” checkbox must be unchecked
- For the Quickstart workflow to configure the vMotion VMkernel, vSphere DRS must be set to enabled

vSAN ESA Example:



Basics

Name	vSAN-ESA-Cluster
Location	vsan-test-dc
ⓘ vSphere DRS	●
ⓘ vSphere HA	●
vSAN	● ☑ Enable vSAN ESA ⓘ

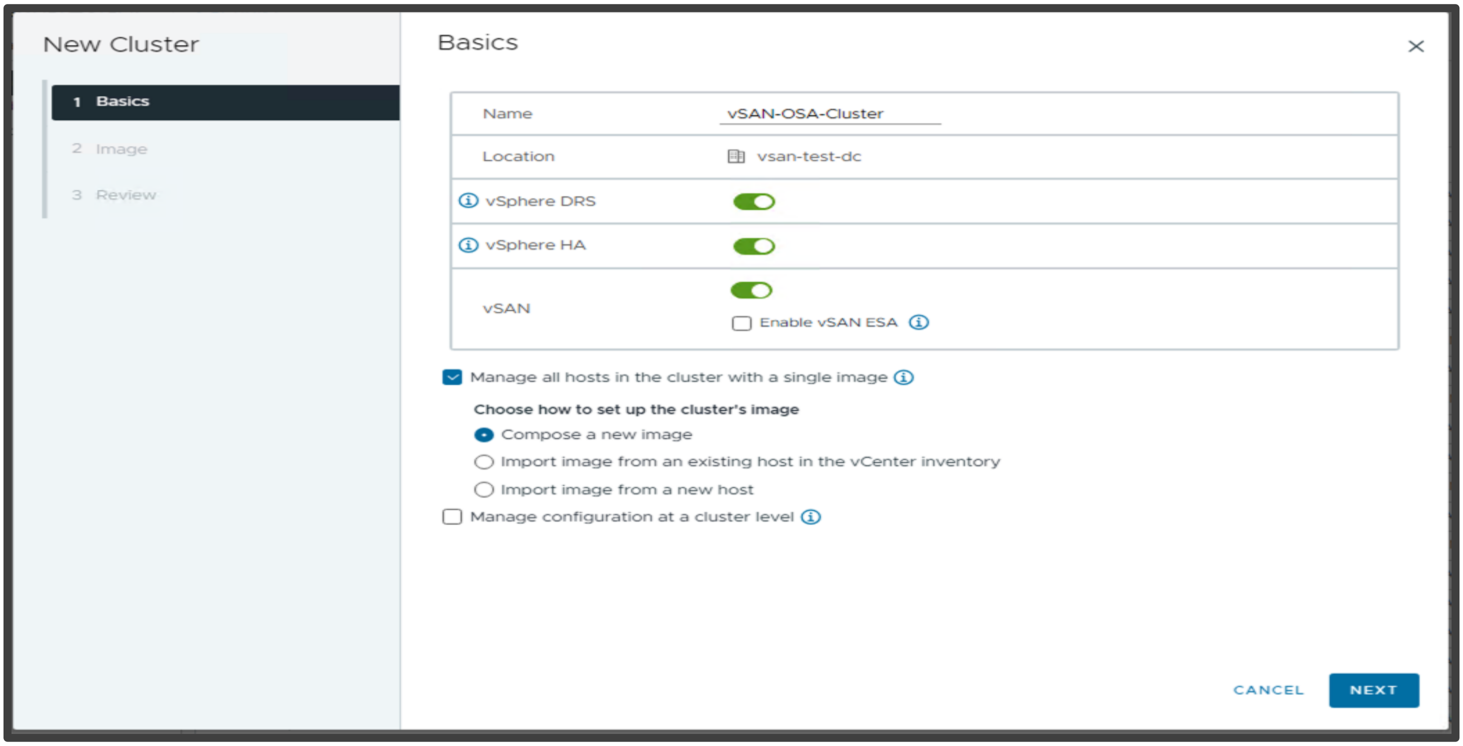
☑ Manage all hosts in the cluster with a single image ⓘ

Choose how to set up the cluster's image

- ● Compose a new image
- ○ Import image from an existing host in the vCenter inventory
- ○ Import image from a new host

☐ Manage configuration at a cluster level ⓘ

vSAN OSA Example:



Basics

Name	vSAN-OSA-Cluster
Location	vsan-test-dc
ⓘ vSphere DRS	●
ⓘ vSphere HA	●
vSAN	● ☐ Enable vSAN ESA ⓘ

☑ Manage all hosts in the cluster with a single image ⓘ

Choose how to set up the cluster's image

- ● Compose a new image
- ○ Import image from an existing host in the vCenter inventory
- ○ Import image from a new host

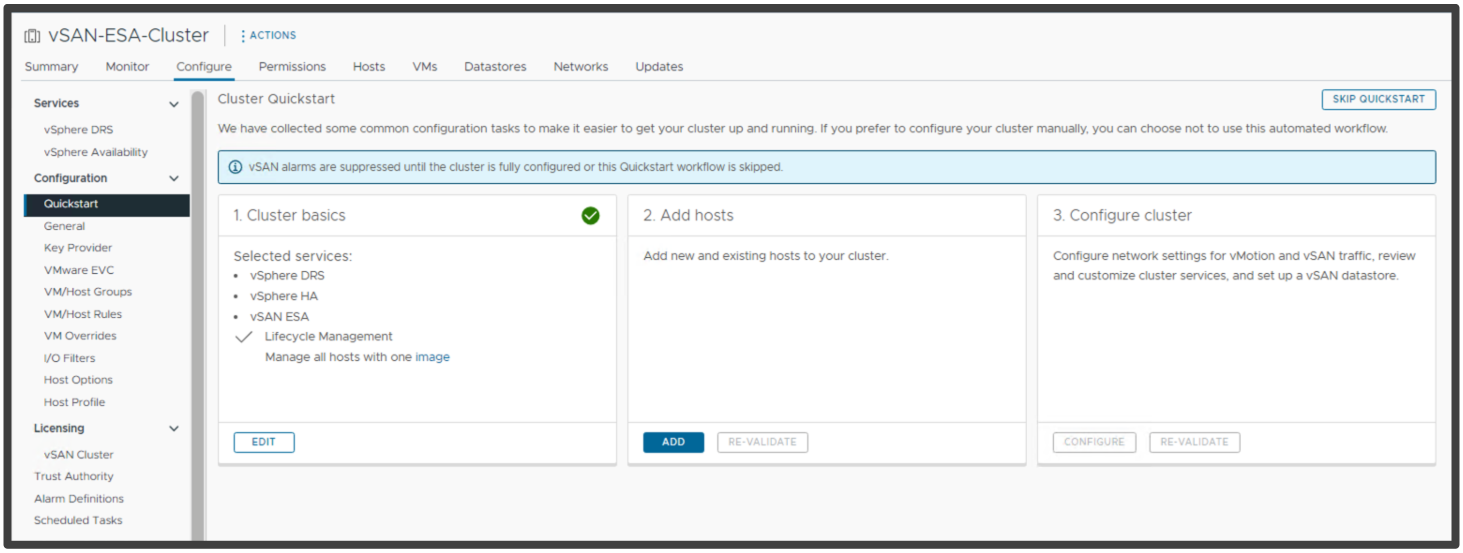
☐ Manage configuration at a cluster level ⓘ

We can also setup the cluster to use a single image (thereby enabling vLCM). For more information on vLCM, see: https://core.vmware.com/resource/introducing-vsphere-lifecycle-management-vlcm.

Quickstart - Cluster Basics

The initial cluster creation above initializes the Quickstart process. Once the cluster has been created, navigate to [vSAN Cluster] **> Configure > Quickstart**. On this screen you will be able to confirm the basic services selected previously then move to the add hosts and configuration phases.

vSAN ESA Example:



Cluster Quickstart

We have collected some common configuration tasks to make it easier to get your cluster up and running. If you prefer to configure your cluster manually, you can choose not to use this automated workflow.

ⓘ vSAN alarms are suppressed until the cluster is fully configured or this Quickstart workflow is skipped.

1. Cluster basics ✓

Selected services:
- vSphere DRS
- vSphere HA
- vSAN ESA

✓ Lifecycle Management
Manage all hosts with one image

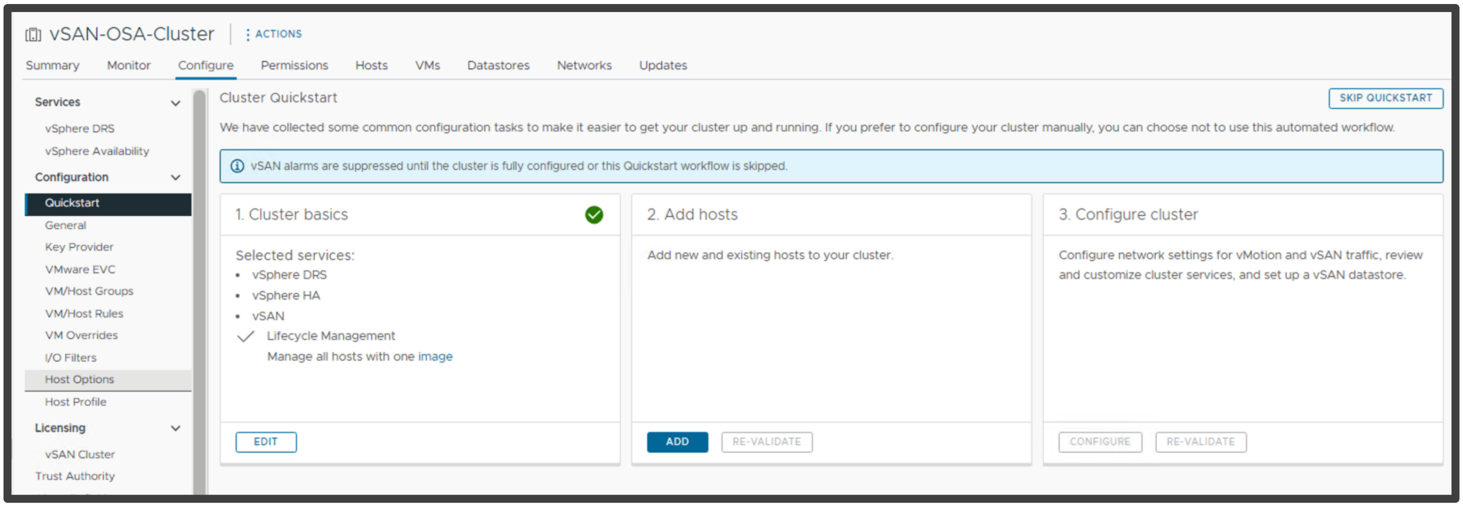
2. Add hosts

Add new and existing hosts to your cluster.

3. Configure cluster

Configure network settings for vMotion and vSAN traffic, review and customize cluster services, and set up a vSAN datastore.

vSAN OSA Example:



Cluster Quickstart

We have collected some common configuration tasks to make it easier to get your cluster up and running. If you prefer to configure your cluster manually, you can choose not to use this automated workflow.

ⓘ vSAN alarms are suppressed until the cluster is fully configured or this Quickstart workflow is skipped.

1. Cluster basics ✓

Selected services:
- vSphere DRS
- vSphere HA
- vSAN

✓ Lifecycle Management
Manage all hosts with one image

2. Add hosts

Add new and existing hosts to your cluster.

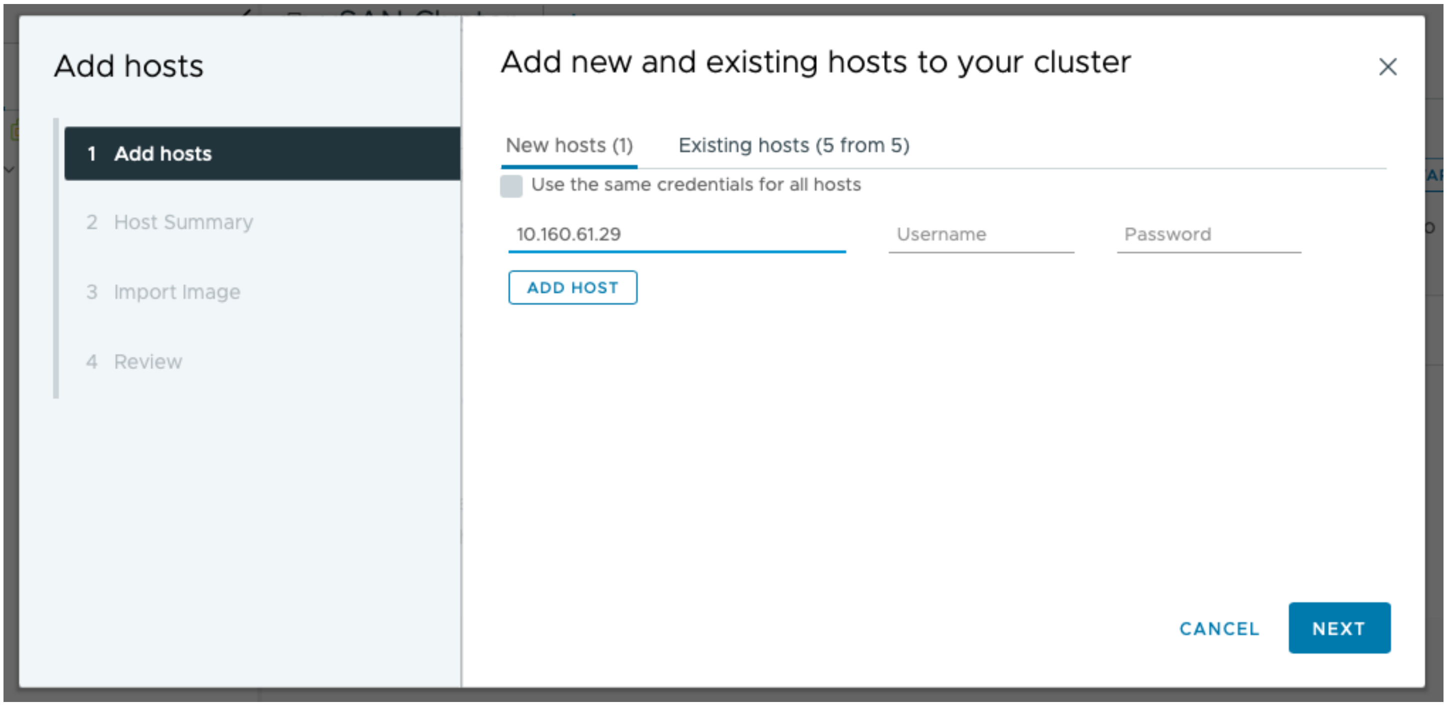
3. Configure cluster

Configure network settings for vMotion and vSAN traffic, review and customize cluster services, and set up a vSAN datastore.

Quickstart - Add Hosts

The next step is to add hosts. The process for ESA and OSA is identical, clicking on the 'Add' button on the 'Add hosts' section presents the dialog below. Multiple hosts can be added at the same time (by IP or FDQN). Additionally, if the credentials of every host are the same, tick the checkbox above the list to quickly complete the form. If the hosts are already in vCenter, they will appear in the 'Existing hosts' tab.

Note: It may be useful to leave one host out of the configuration here to later demonstrate cluster expansion.



Add hosts

1 Add hosts

2 Host Summary

3 Import Image

4 Review

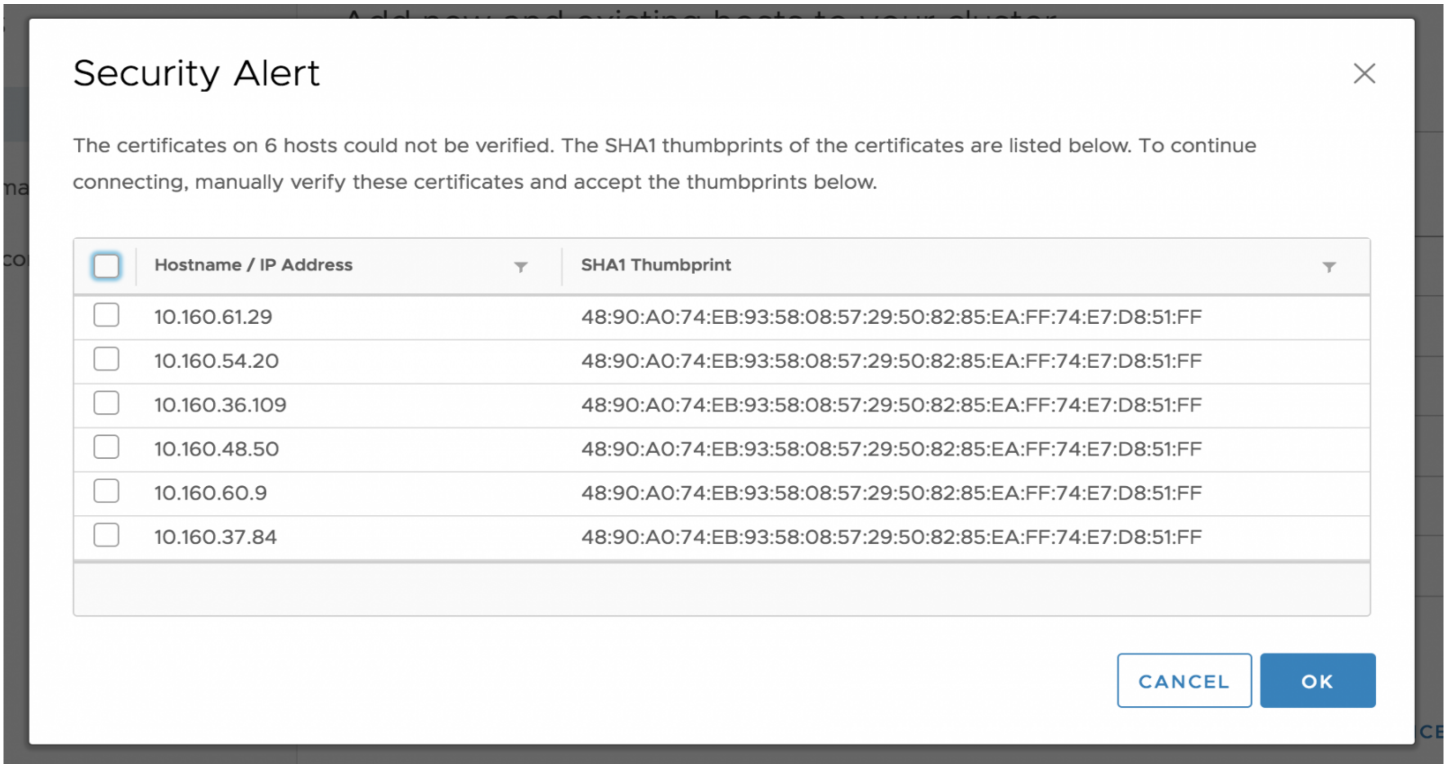
Add new and existing hosts to your cluster

New hosts (1) Existing hosts (5 from 5)

□ Use the same credentials for all hosts

10.160.61.29 | Username | Password

Once the host details have been entered, click Next. You are then presented with a dialog showing the thumbprints of the hosts. If these are as expected, tick the checkbox(es) and then click OK.

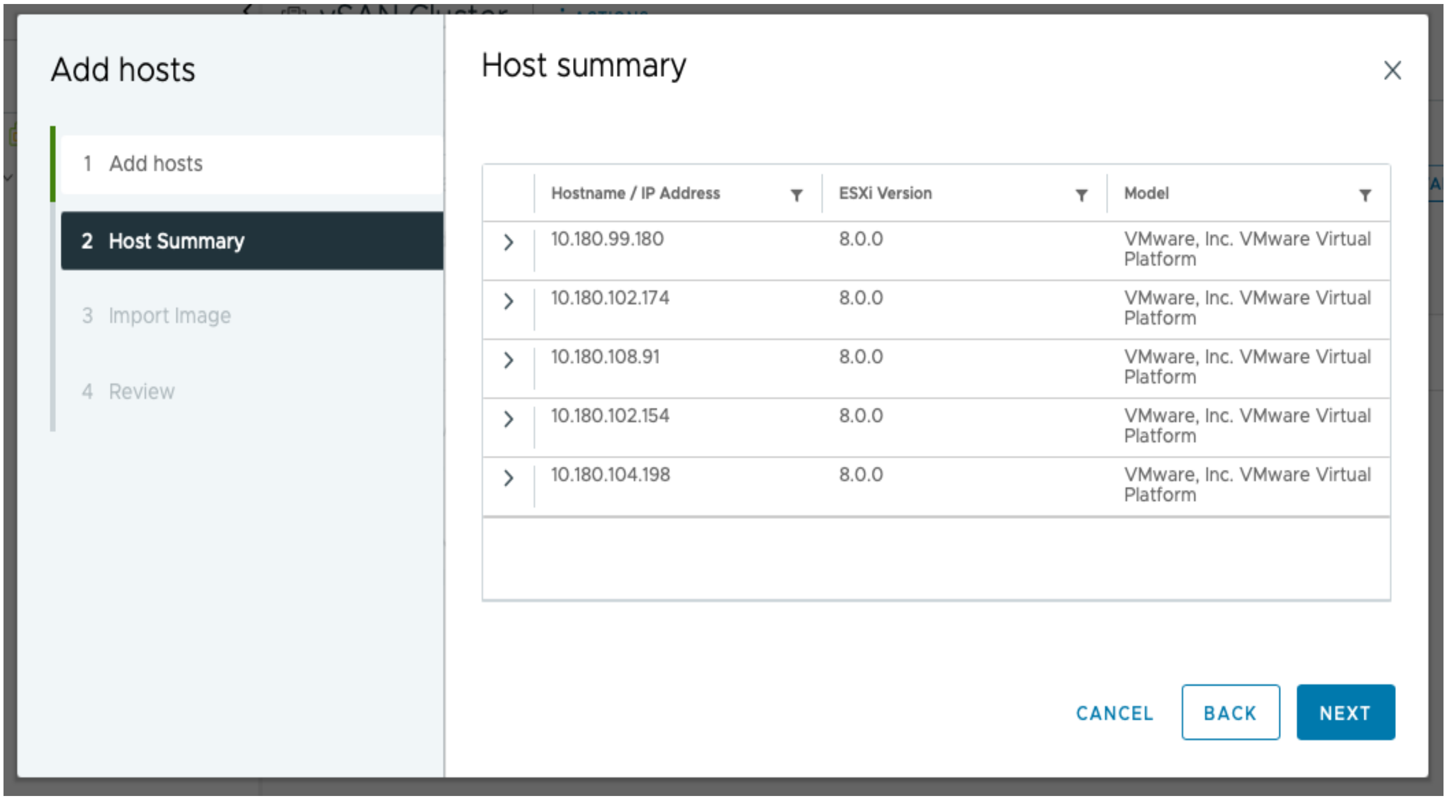


Security Alert

The certificates on 6 hosts could not be verified. The SHA1 thumbprints of the certificates are listed below. To continue connecting, manually verify these certificates and accept the thumbprints below.

□	Hostname / IP Address	SHA1 Thumbprint
□	10.160.61.29	48:90:A0:74:EB:93:58:08:57:29:50:82:85:EA:FF:74:E7:D8:51:FF
□	10.160.54.20	48:90:A0:74:EB:93:58:08:57:29:50:82:85:EA:FF:74:E7:D8:51:FF
□	10.160.36.109	48:90:A0:74:EB:93:58:08:57:29:50:82:85:EA:FF:74:E7:D8:51:FF
□	10.160.48.50	48:90:A0:74:EB:93:58:08:57:29:50:82:85:EA:FF:74:E7:D8:51:FF
□	10.160.60.9	48:90:A0:74:EB:93:58:08:57:29:50:82:85:EA:FF:74:E7:D8:51:FF
□	10.160.37.84	48:90:A0:74:EB:93:58:08:57:29:50:82:85:EA:FF:74:E7:D8:51:FF

A summary will be displayed, showing the vSphere version on each host and other details. Verify the details are correct:

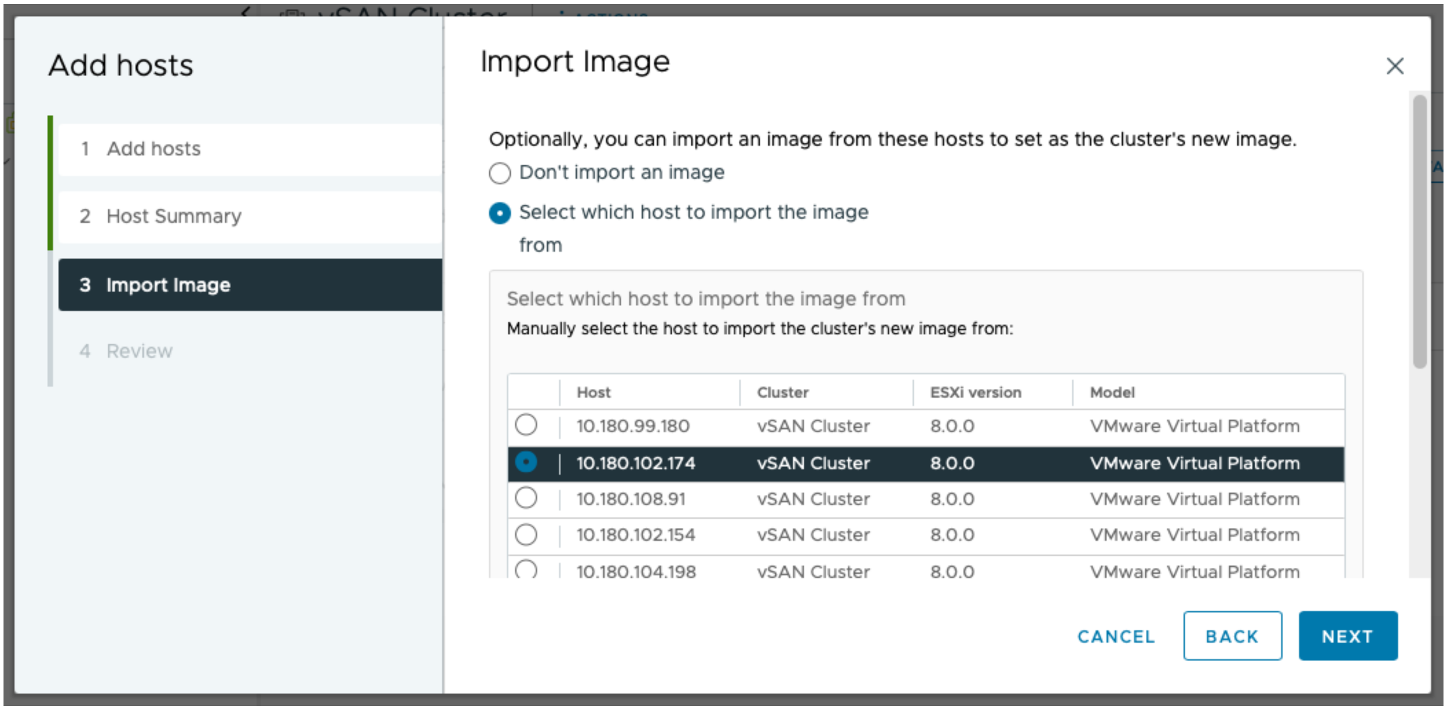


Add hosts

Host summary

	Hostname / IP Address	ESXi Version	Model
>	10.180.99.180	8.0.0	VMware, Inc. VMware Virtual Platform
>	10.180.102.174	8.0.0	VMware, Inc. VMware Virtual Platform
>	10.180.108.91	8.0.0	VMware, Inc. VMware Virtual Platform
>	10.180.102.154	8.0.0	VMware, Inc. VMware Virtual Platform
>	10.180.104.198	8.0.0	VMware, Inc. VMware Virtual Platform

On the next screen, we have the option to import the ESXi an image from a host (to set as the cluster's new image). Select an option and continue.



Add hosts

Import Image

Optionally, you can import an image from these hosts to set as the cluster's new image.

○ Don't import an image

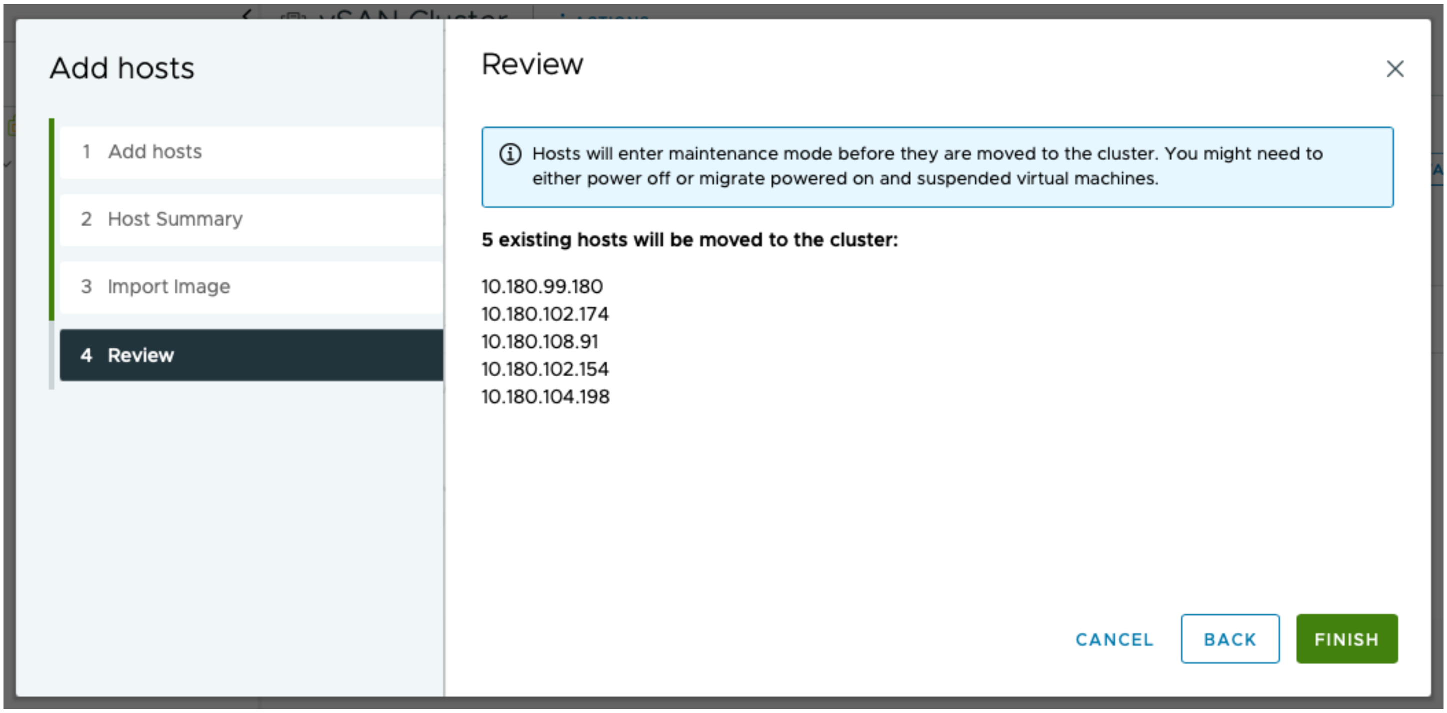
● Select which host to import the image from

Select which host to import the image from

Manually select the host to import the cluster's new image from:

	Host	Cluster	ESXi version	Model
○	10.180.99.180	vSAN Cluster	8.0.0	VMware Virtual Platform
●	10.180.102.174	vSAN Cluster	8.0.0	VMware Virtual Platform
○	10.180.108.91	vSAN Cluster	8.0.0	VMware Virtual Platform
○	10.180.102.154	vSAN Cluster	8.0.0	VMware Virtual Platform
○	10.180.104.198	vSAN Cluster	8.0.0	VMware Virtual Platform

Finally, review and click **Finish** if everything is in order.



Add hosts

1 Add hosts
2 Host Summary
3 Import Image
4 Review

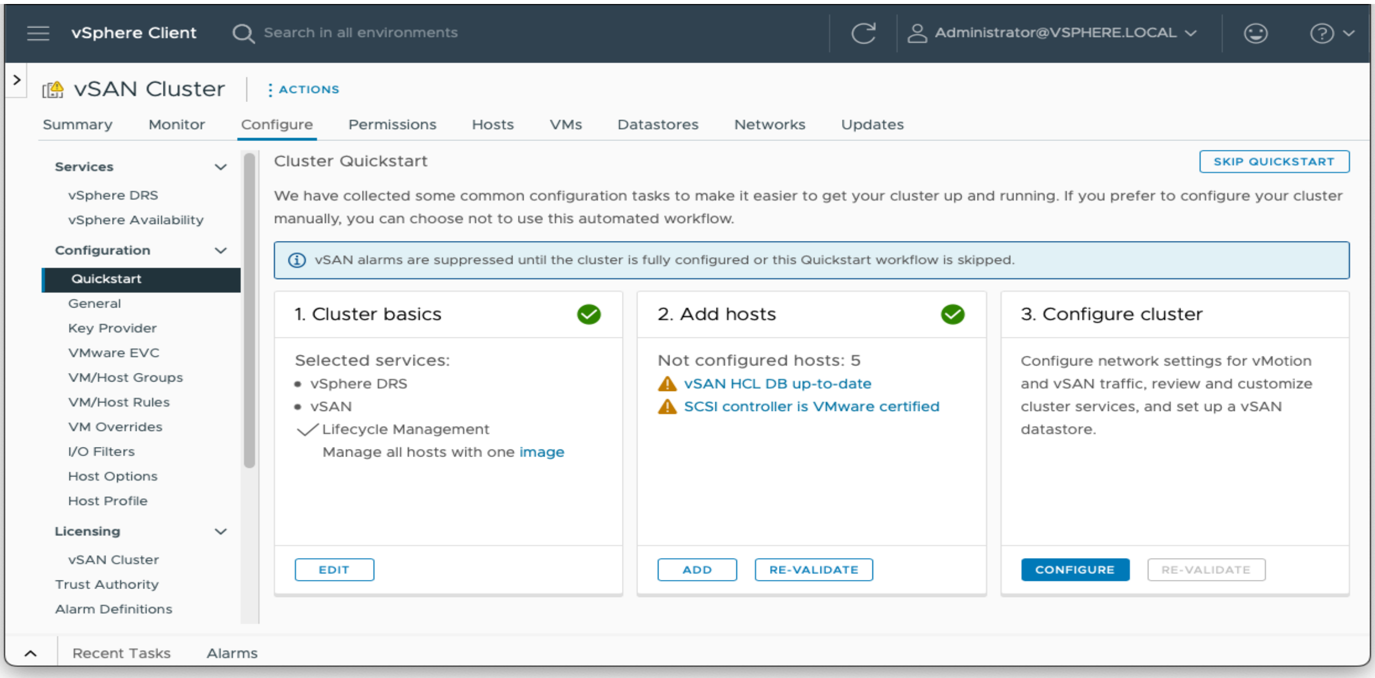
Review

ⓘ Hosts will enter maintenance mode before they are moved to the cluster. You might need to either power off or migrate powered on and suspended virtual machines.

5 existing hosts will be moved to the cluster:

10.180.99.180
10.180.102.174
10.180.108.91
10.180.102.154
10.180.104.198

After the hosts have been added, validation will be performed automatically on the cluster. Check for any errors and inconsistencies and re-validate if necessary.



Cluster Quickstart

SKIP QUICKSTART

We have collected some common configuration tasks to make it easier to get your cluster up and running. If you prefer to configure your cluster manually, you can choose not to use this automated workflow.

ⓘ vSAN alarms are suppressed until the cluster is fully configured or this Quickstart workflow is skipped.

1. Cluster basics ✓

Selected services:
- vSphere DRS
- vSAN

✓ Lifecycle Management
Manage all hosts with one image

EDIT

2. Add hosts ✓

Not configured hosts: 5
⚠ vSAN HCL DB up-to-date
⚠ SCSI controller is VMware certified

ADD RE-VALIDATE

3. Configure cluster

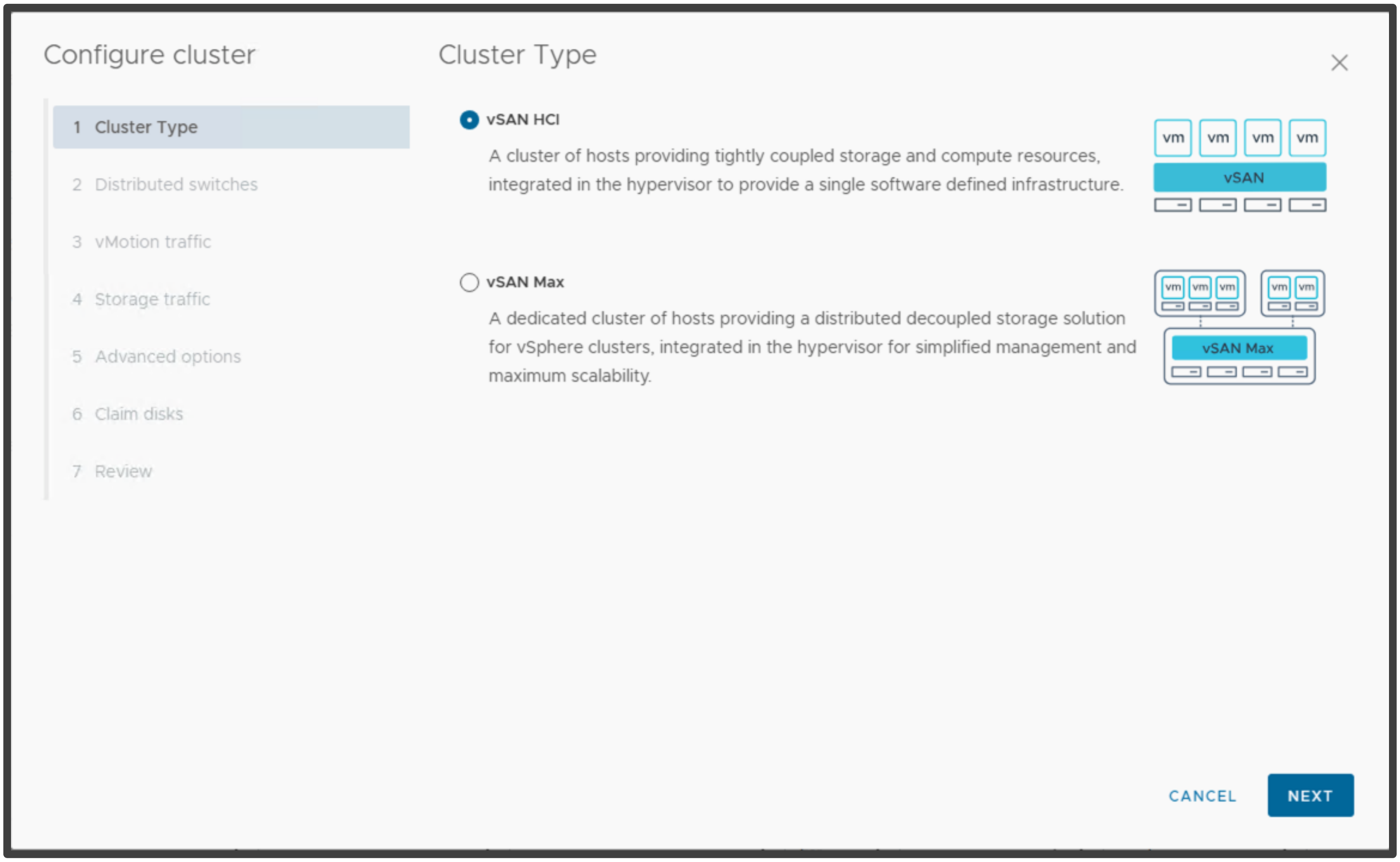
Configure network settings for vMotion and vSAN traffic, review and customize cluster services, and set up a vSAN datastore.

CONFIGURE RE-VALIDATE

Quickstart - Configure Cluster

The next step is to configure the vSAN HCI cluster. After clicking on **Configure** under Step 3: Configure Cluster, the Configure Cluster workflow will start. For ESA deployments the first step will ask whether this is a vSAN HCI or vSAN Max deployment. Ensure the **vSAN HCI** radio button is selected. OSA deployments do not have this screen. Instead, they start at the Distributed Switches step.

Step 1: Select Cluster Type:



Configure cluster

1 Cluster Type
2 Distributed switches
3 vMotion traffic
4 Storage traffic
5 Advanced options
6 Claim disks
7 Review

Cluster Type

● vSAN HCI

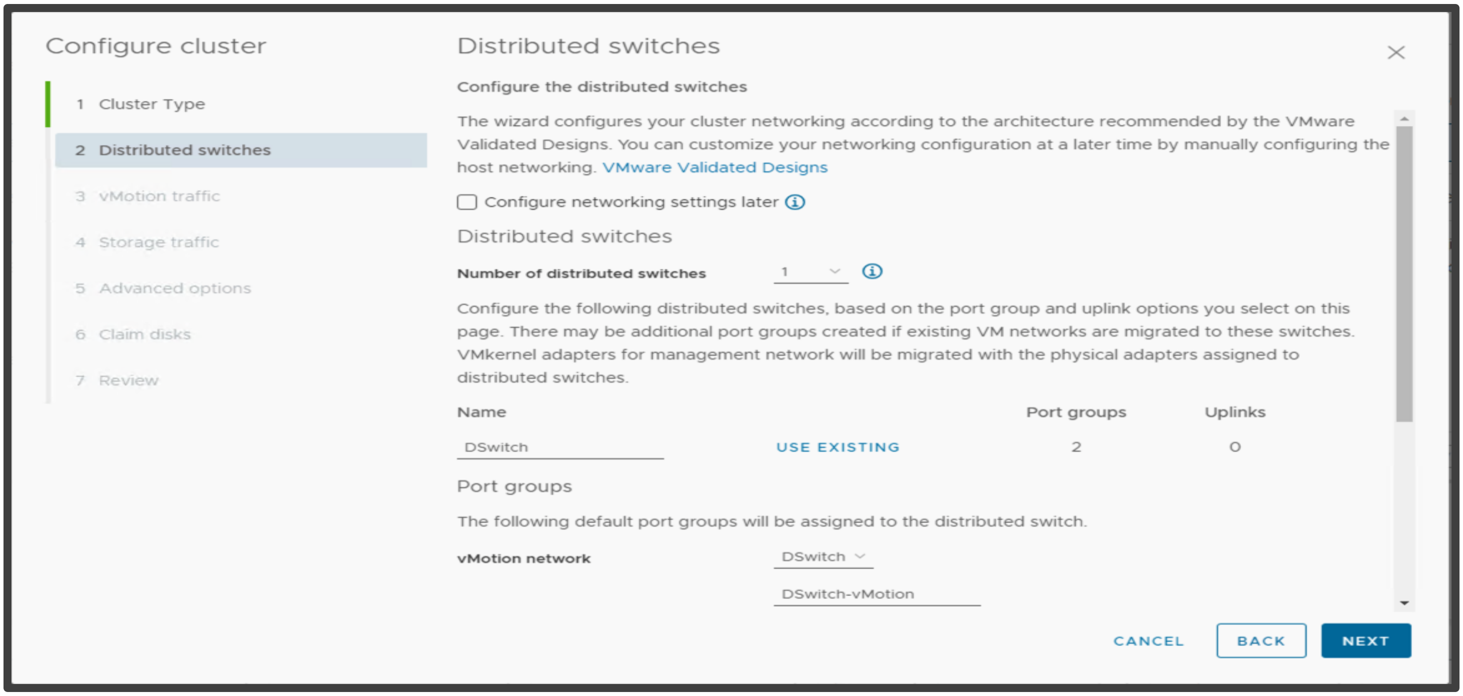
A cluster of hosts providing tightly coupled storage and compute resources, integrated in the hypervisor to provide a single software defined infrastructure.

○ vSAN Max

A dedicated cluster of hosts providing a distributed decoupled storage solution for vSphere clusters, integrated in the hypervisor for simplified management and maximum scalability.

Step 2: Configure Distributed Switches:

The next dialog allows for the configuration of the distributed switch(es) for the cluster. Leave the default 'Number of distributed switches' set to 1 and assign a name to the switch and port groups.



Distributed switches

Configure the distributed switches

The wizard configures your cluster networking according to the architecture recommended by the VMware Validated Designs. You can customize your networking configuration at a later time by manually configuring the host networking. VMware Validated Designs

□ Configure networking settings later

Distributed switches

Number of distributed switches 1

Configure the following distributed switches, based on the port group and uplink options you select on this page. There may be additional port groups created if existing VM networks are migrated to these switches. VMkernel adapters for management network will be migrated with the physical adapters assigned to distributed switches.

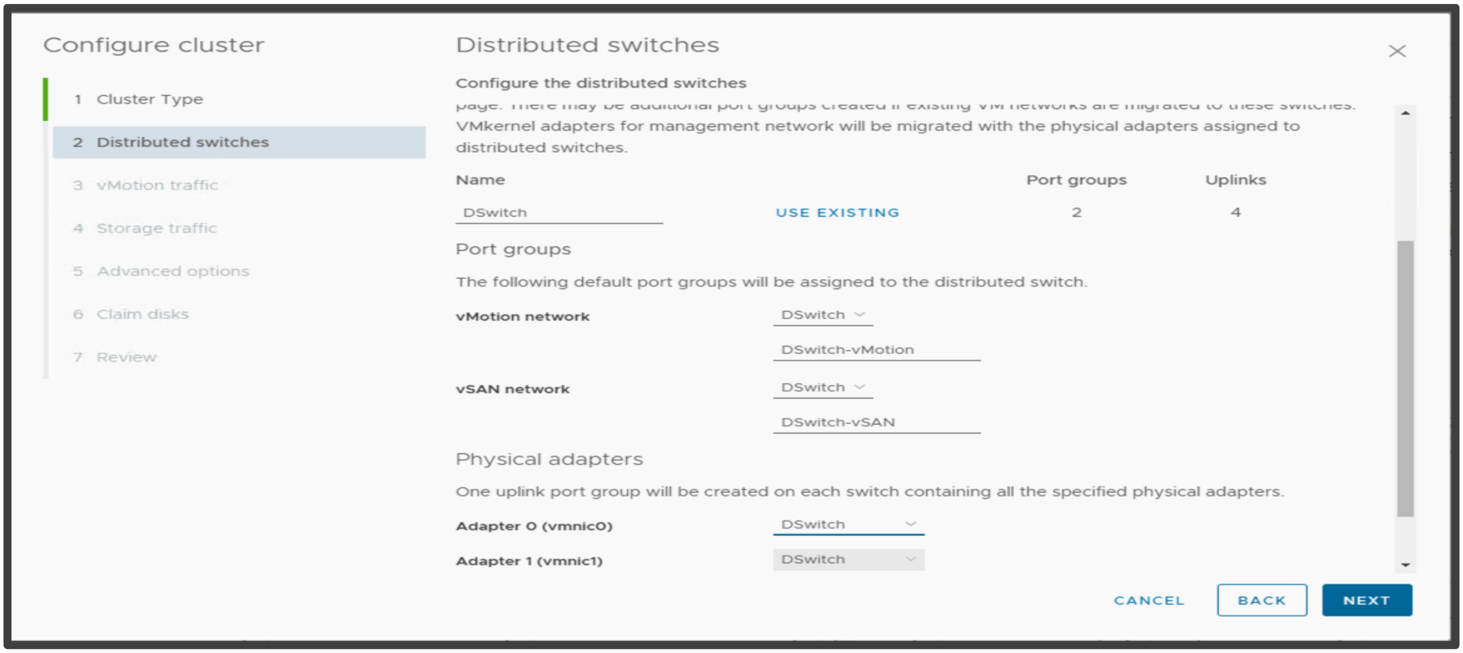
Name		Port groups	Uplinks
DSwitch	USE EXISTING	2	0

Port groups

The following default port groups will be assigned to the distributed switch.

vMotion network DSwitch
DSwitch-vMotion

Scroll down and configure the port groups and physical adapters as needed.



Distributed switches

Configure the distributed switches

page. There may be additional port groups created if existing VM networks are migrated to these switches. VMkernel adapters for management network will be migrated with the physical adapters assigned to distributed switches.

Name		Port groups	Uplinks
DSwitch	USE EXISTING	2	4

Port groups

The following default port groups will be assigned to the distributed switch.

vMotion network DSwitch
DSwitch-vMotion

vSAN network DSwitch
DSwitch-vSAN

Physical adapters

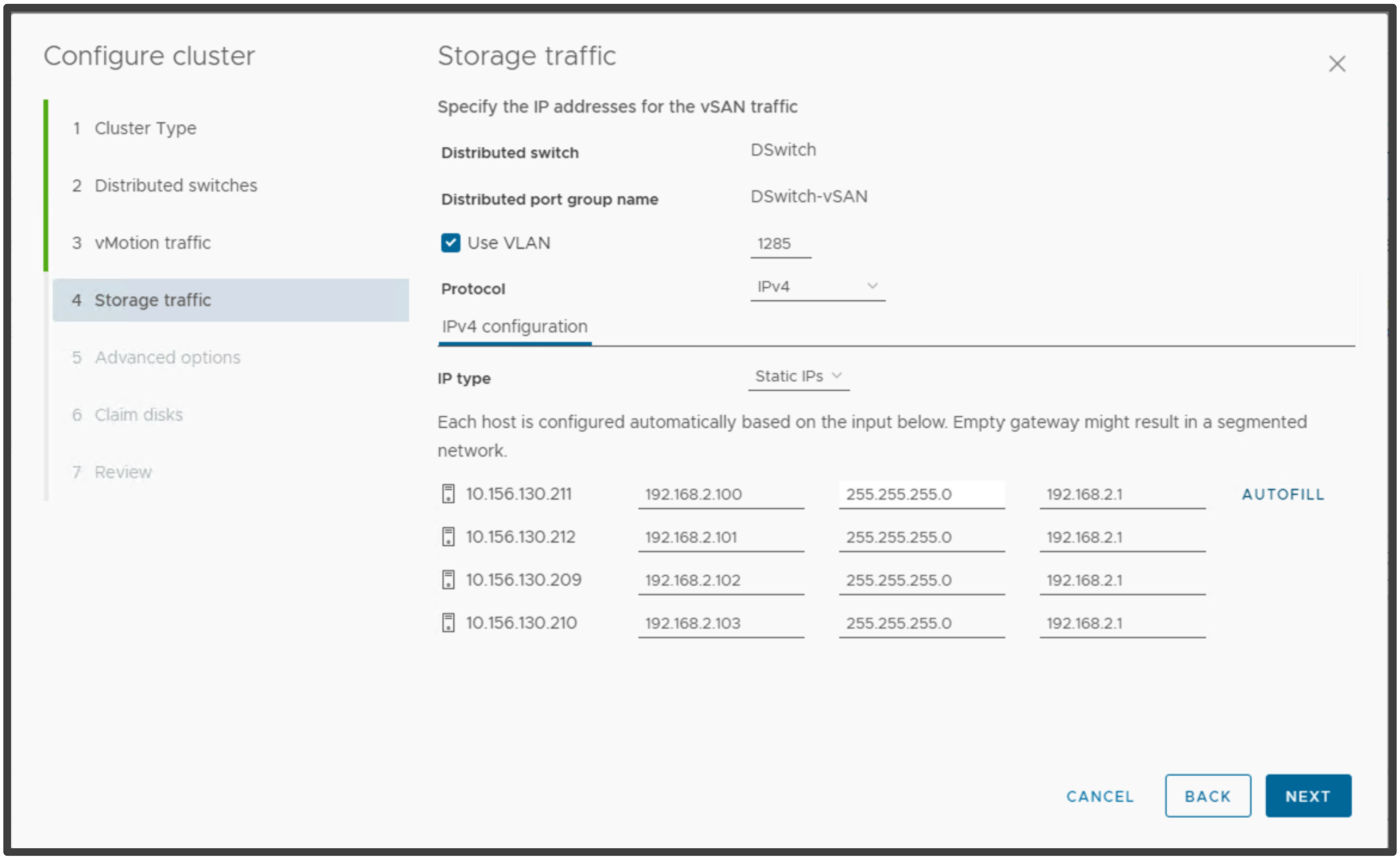
One uplink port group will be created on each switch containing all the specified physical adapters.

Adapter 0 (vmnic0) DSwitch

Adapter 1 (vmnic1) DSwitch

Steps 3 & 4: Configure vMotion and Storage Traffic:

On the next two screens, set the VLAN and IP addresses to be used for the vMotion and vSAN for each host. The 'autofill' function can be used to input consecutive IP addresses.



Storage traffic

Specify the IP addresses for the vSAN traffic

Distributed switch DSwitch

Distributed port group name DSwitch-vSAN

☑ Use VLAN 1285

Protocol IPv4

IPv4 configuration

IP type Static IPs

Each host is configured automatically based on the input below. Empty gateway might result in a segmented network.

10.156.130.211	192.168.2.100	255.255.255.0	192.168.2.1	AUTOFILL
10.156.130.212	192.168.2.101	255.255.255.0	192.168.2.1	
10.156.130.209	192.168.2.102	255.255.255.0	192.168.2.1	
10.156.130.210	192.168.2.103	255.255.255.0	192.168.2.1	

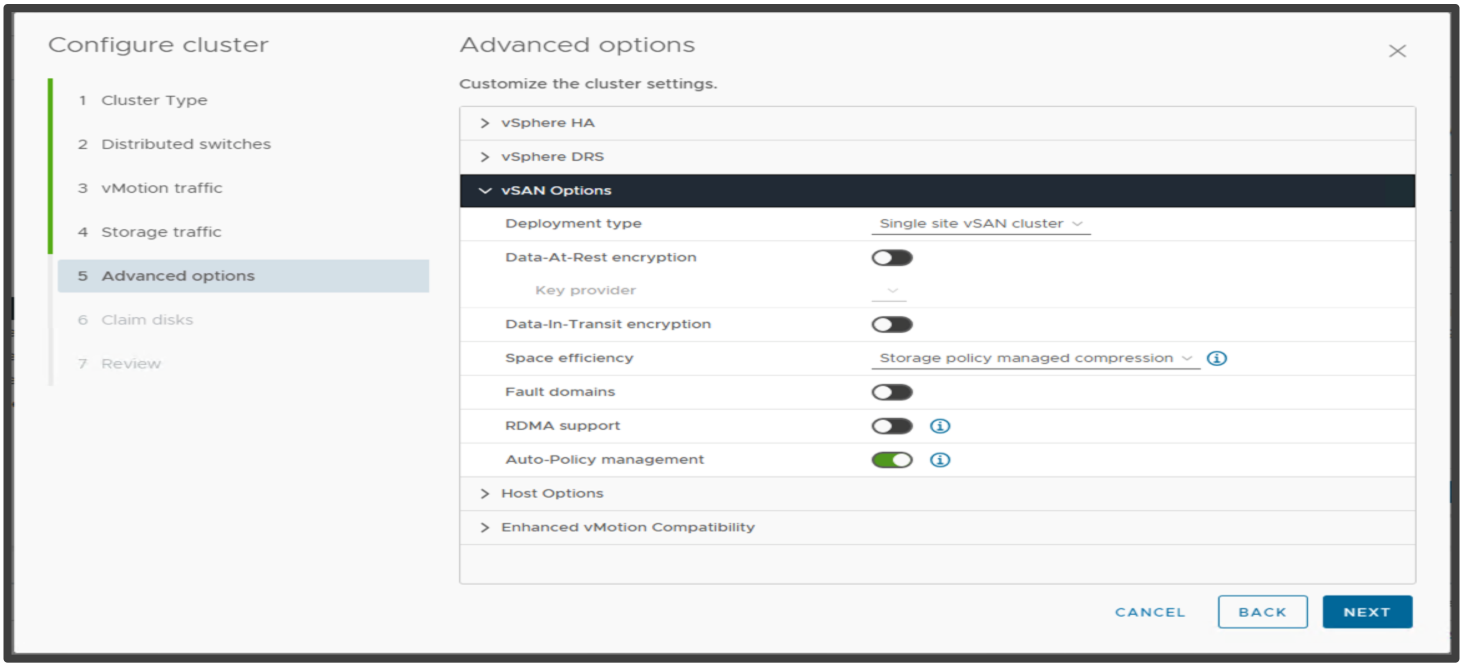
Step 5: Configure Advanced Options:

Here we select 'Single site cluster' as the deployment type. For stretched clusters, refer to the stretched cluster guide

Auto-Policy Management is a vSAN ESA specific feature that ensures resilience settings for the environment are optimally configured automatically. Enabling this feature does not preclude the creation of custom storage policies as needed post deployment. For this deployment Auto-Policy management is set to the default "enabled."

Note: OSA deployments do not support Auto-Policy Management or the RDMA support features. These options will be grayed out in an OSA deployment.

For more information please review: https://core.vmware.com/blog/auto-policy-remediation-enhancements-esa-vsan-8-u2

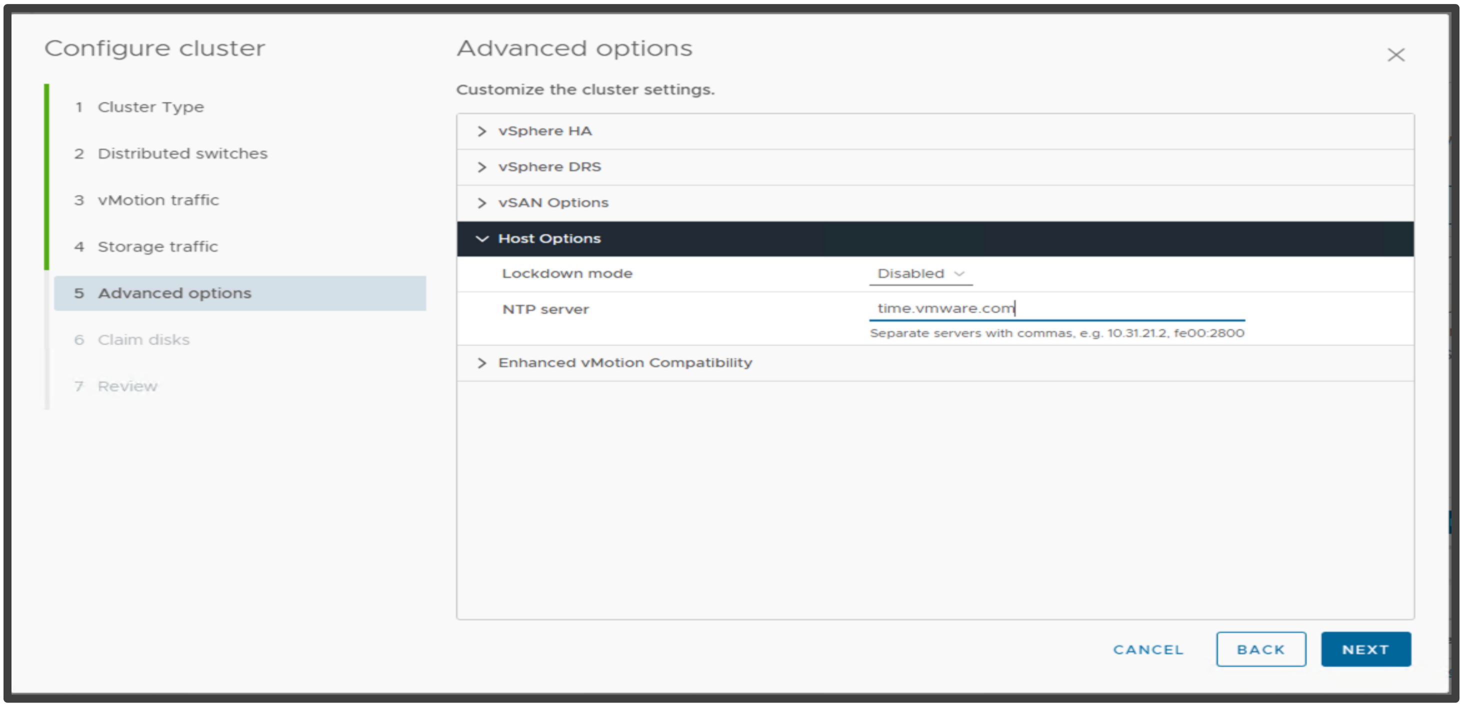


Advanced options

Customize the cluster settings.

- > vSphere HA
- > vSphere DRS
- ∨ vSAN Options
  - Deployment type: Single site vSAN cluster
  - Data-At-Rest encryption: ○ (off)
    - Key provider
  - Data-In-Transit encryption: ○ (off)
  - Space efficiency: Storage policy managed compression
  - Fault domains: ○ (off)
  - RDMA support: ○ (off)
  - Auto-Policy management: ● (on)
- > Host Options
- > Enhanced vMotion Compatibility

Ensure that an NTP server is added to the configuration (under ‘Host Options’). As vSAN is a distributed system, features may not work as expected if there is a time drift between servers (ideally, also ensure that vCenter has the same time source).



Advanced options

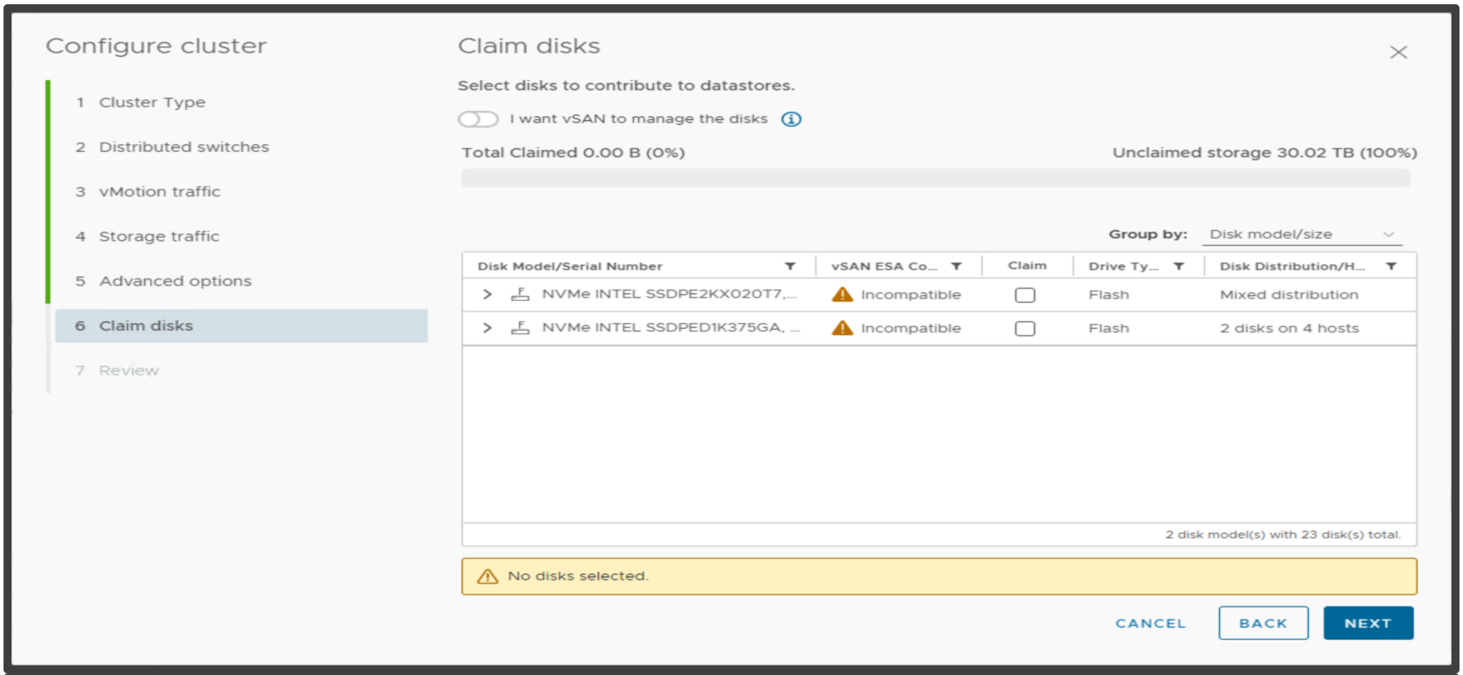
Customize the cluster settings.

- > vSphere HA
- > vSphere DRS
- > vSAN Options
- ∨ Host Options
  - Lockdown mode: Disabled
  - NTP server: time.vmware.com
    - Separate servers with commas, e.g. 10.31.21.2, fe00:2800
- > Enhanced vMotion Compatibility

Step 6: Claim Disks:

On the next screen, select the disks to use. For a vSAN OSA cluster, the system will attempt to select disks automatically as cache or capacity (on a vSAN ESA cluster, there is no tier selection). As the ESA example images indicate, the workflow checks drive compatibility and warns if an issue is detected.

vSAN ESA example:



Claim disks

Select disks to contribute to datastores.

○ I want vSAN to manage the disks

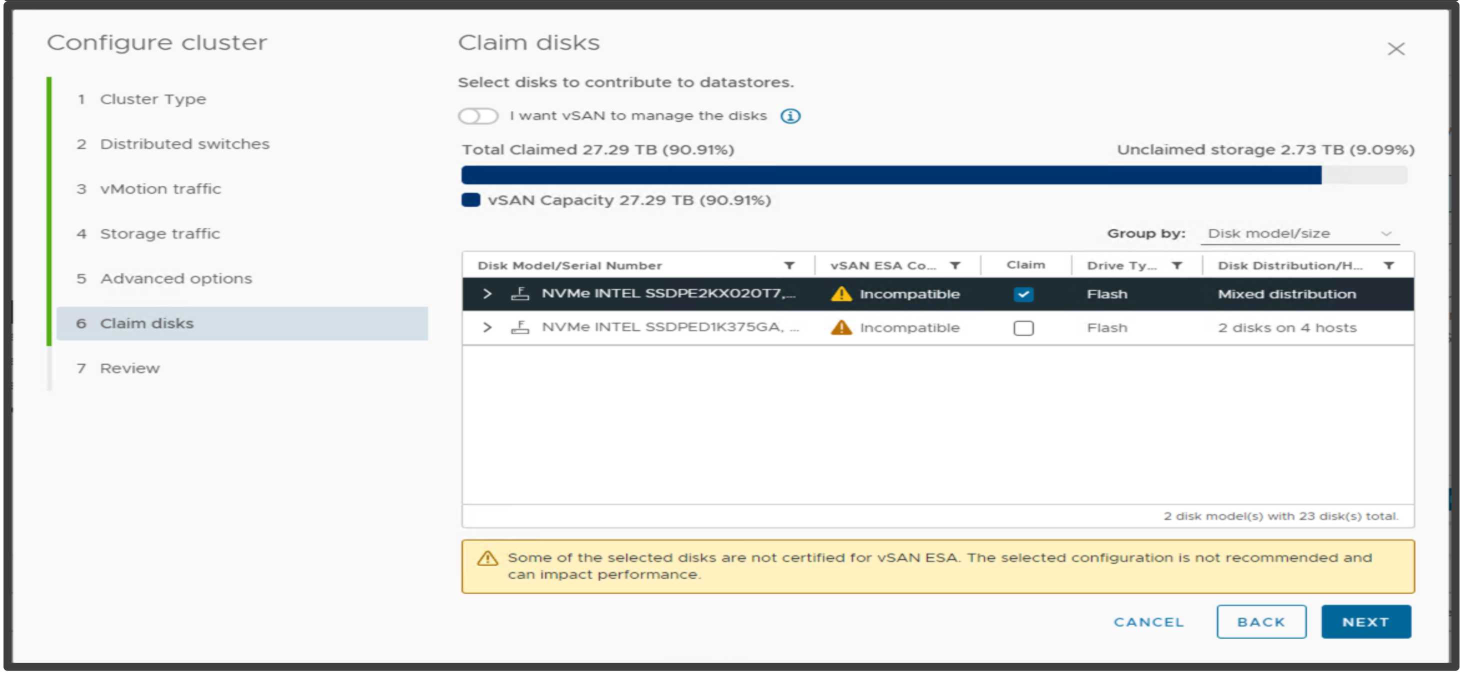
Total Claimed 0.00 B (0%) Unclaimed storage 30.02 TB (100%)

Group by: Disk model/size

Disk Model/Serial Number	vSAN ESA Co...	Claim	Drive Ty...	Disk Distribution/H...
NVMe INTEL SSDPE2KX020T7,...	⚠ Incompatible	□	Flash	Mixed distribution
NVMe INTEL SSDPED1K375GA, ...	⚠ Incompatible	□	Flash	2 disks on 4 hosts

2 disk model(s) with 23 disk(s) total.

⚠ No disks selected.



Claim disks

Select disks to contribute to datastores.

○ I want vSAN to manage the disks

Total Claimed 27.29 TB (90.91%) Unclaimed storage 2.73 TB (9.09%)

■ vSAN Capacity 27.29 TB (90.91%)

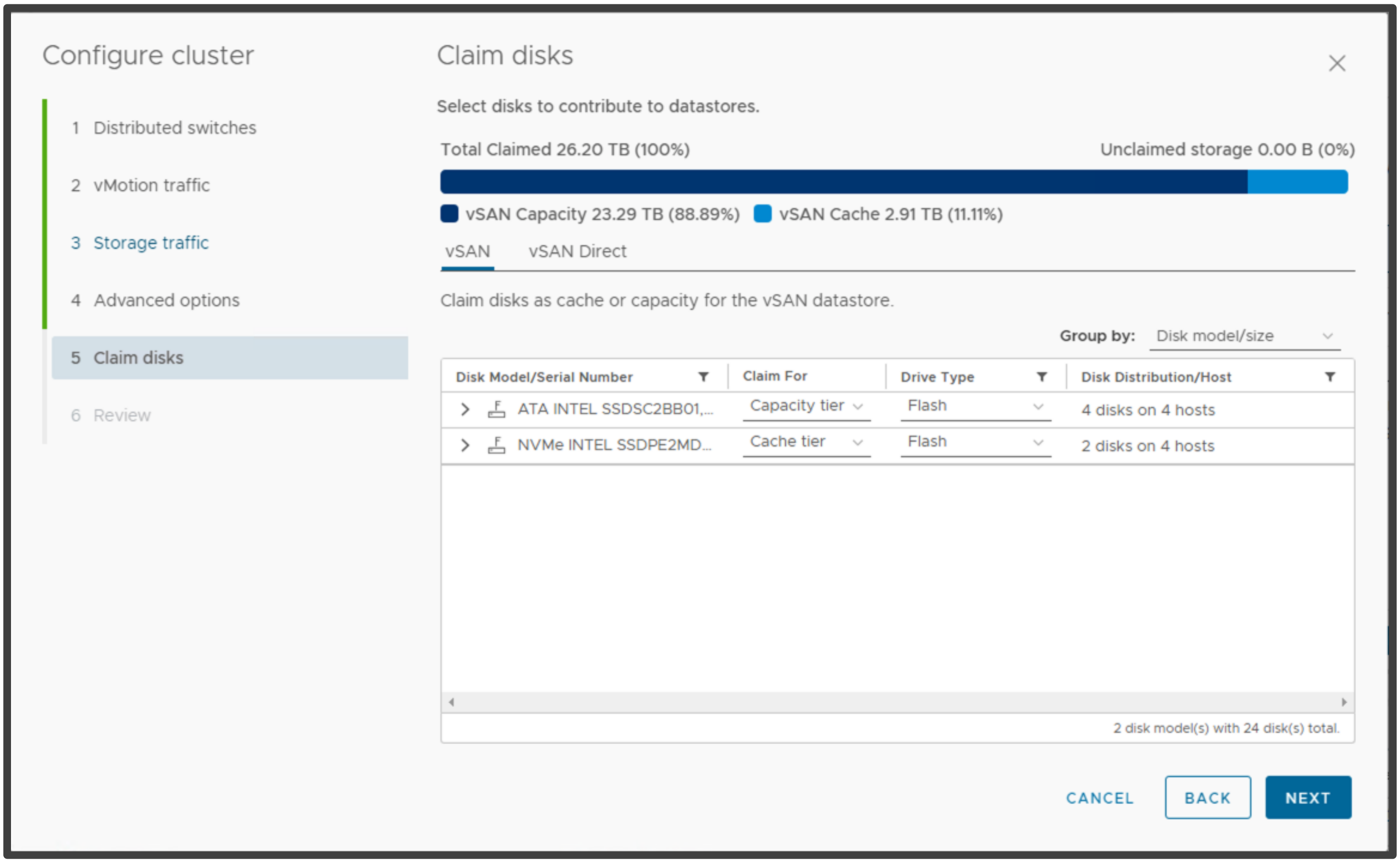
Group by: Disk model/size

Disk Model/Serial Number	vSAN ESA Co...	Claim	Drive Ty...	Disk Distribution/H...
NVMe INTEL SSDPE2KX020T7,...	⚠ Incompatible	☑	Flash	Mixed distribution
NVMe INTEL SSDPED1K375GA, ...	⚠ Incompatible	□	Flash	2 disks on 4 hosts

2 disk model(s) with 23 disk(s) total.

⚠ Some of the selected disks are not certified for vSAN ESA. The selected configuration is not recommended and can impact performance.

vSAN OSA Example:



Configure cluster

1 Distributed switches

2 vMotion traffic

3 Storage traffic

4 Advanced options

5 Claim disks

6 Review

Claim disks

Select disks to contribute to datastores.

Total Claimed 26.20 TB (100%)

Unclaimed storage 0.00 B (0%)

■ vSAN Capacity 23.29 TB (88.89%) ■ vSAN Cache 2.91 TB (11.11%)

Claim disks as cache or capacity for the vSAN datastore.

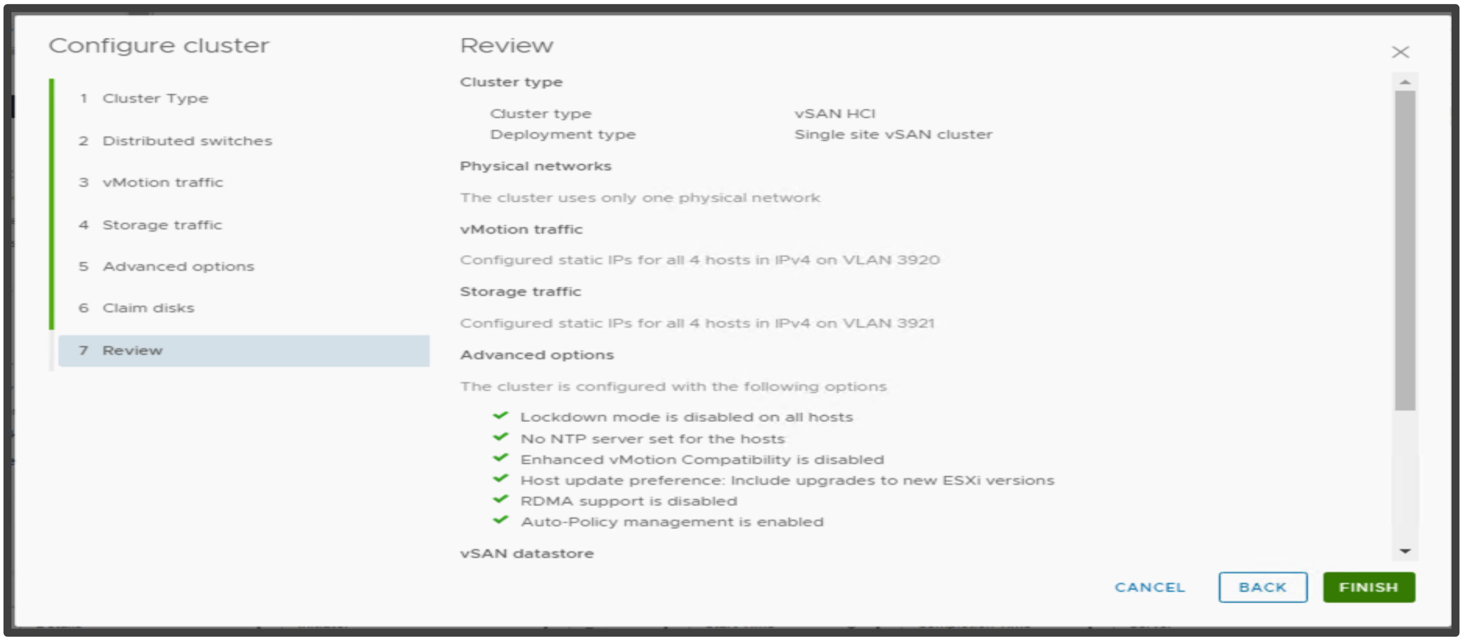
Group by: Disk model/size

Disk Model/Serial Number	Claim For	Drive Type	Disk Distribution/Host
ATA INTEL SSDSC2BB01,...	Capacity tier	Flash	4 disks on 4 hosts
NVMe INTEL SSDPE2MD...	Cache tier	Flash	2 disks on 4 hosts

2 disk model(s) with 24 disk(s) total.

Step 7: Review:

Finally, check everything is as expected and click **Finish**.



Review

Cluster type

Cluster type	vSAN HCI
Deployment type	Single site vSAN cluster

Physical networks

The cluster uses only one physical network

vMotion traffic

Configured static IPs for all 4 hosts in IPv4 on VLAN 3920

Storage traffic

Configured static IPs for all 4 hosts in IPv4 on VLAN 3921

Advanced options

The cluster is configured with the following options

- ✓ Lockdown mode is disabled on all hosts
- ✓ No NTP server set for the hosts
- ✓ Enhanced vMotion Compatibility is disabled
- ✓ Host update preference: Include upgrades to new ESXi versions
- ✓ RDMA support is disabled
- ✓ Auto-Policy management is enabled

vSAN datastore

Monitor the creation in the task view and wait for the cluster to be formed.

Using Quickstart, we have created a cluster with vSAN enabled with the correct networking in place.

Manually Enabling vSAN Services on a Cluster

Note: If Quickstart was used (as per the earlier section) then this section can be skipped.

Manual vSAN enablement is available for those that do not wish to use the Quickstart process.

For this scenario, please follow the Manually Enabling vSAN instructions on the VMware Docs page linked below:
https://docs.vmware.com/en/VMware-vSphere/8.0/vsan-planning/GUID-53571374-B3E5-4767-A372-FEB7C995AF71.html

Enabling vSAN Services on a VMware Cloud Foundation™ based Cluster

VCF includes dedicated processes to automate the deployment and configuration of core infrastructure including vSAN services. In fact, these processes are required and are the only supported methods within VCF. This applies to standing up the initial Management Domain, a subsidiary Workload Domain, or a cluster with a Domain.

For more information on enabling vSAN services within the VCF Model please review these links:

- vSAN Storage with VCF – https://docs.vmware.com/en/VMware-Cloud-Foundation/5.1/vcf-admin/GUID-766FBED1-6FE8-46D0-99C5-62355478F2CF.html
- VCF Admin Guide main page – https://docs.vmware.com/en/VMware-Cloud-Foundation/5.1/vcf-admin/GUID-D5A44DAA-866D-47C9-B1FB-BF9761F97E36.html

- Management Domain Deployment Walkthrough – https://core.vmware.com/cloud-foundation-automation#domain-operations
- Workload Domain Deployment Walkthrough -https://core.vmware.com/cloud-foundation-automation#domain-operations
- Add a Cluster to an existing Workload Domain – https://core.vmware.com/cloud-foundation-automation#cluster-operations

Enabling vSAN Max™ - Disaggregated Storage

What is vSAN Max? VMware's new disaggregated storage offering that provides Petabyte-scale centralized shared storage for your vSphere clusters.

Built on the foundation of vSAN ESA, vSAN Max is a fully distributed architecture, where access to data is NOT funneled through centralized I/O controllers but using the full power of each node (host) in the vSAN Max cluster. The aggregated resources across all hosts in a vSAN Max cluster contribute to the processing of I/O. The addition of new hosts means that it can scale capacity and performance linearly.

For more additional technical information please review:

- Main vSAN Max informational page – https://core.vmware.com/vsan-max
- vSAN Max Design & Operations Guide – https://core.vmware.com/resource/vsan-max-design-and-operational-guidance
- vSAN Max introduction blog post – https://core.vmware.com/blog/introducing-vsan-max
- vSAN Max scalability blog post – https://core.vmware.com/blog/vsan-max-and-advantage-scalability

vSAN Max Sizing Considerations

When sizing vSAN Max deployments, consider that vSAN Max clusters support:

- A maximum of 32 ESXi hosts in the cluster (24 ESXi hosts recommended)
- A maximum of 10 compute clusters mounting to a single vSAN Max cluster
- A maximum of 128 total ESXi hosts (both within the vSAN Max cluster and the vSAN Compute clusters connecting to a single vSAN Max datastore)

Note: Limiting the vSAN Max cluster size to 24 ESXi hosts will allow for up to 104 ESXI hosts from vSAN compute clusters to mount the datastore, offering a 4.3:1 ratio. A vSAN Max cluster size of 32 ESXi hosts would allow for up to 96 ESXI hosts from vSAN compute clusters to mount the datastore, offering a storage ratio of 3:1.

Disaggregated Storage for vSAN OSA (AKA: HCI Mesh)

Although the vSAN Max is explicitly a vSAN ESA function, vSAN OSA deployments still support disaggregated storage.

vSAN OSA datastores can be shared between two vSAN clusters, utilizing vSAN's native data path for cross-cluster connections. Compute Only Clusters are also supported.

Each vSAN OSA client cluster can mount a maximum of ten remote vSAN OSA datastores. A vSAN OSA server cluster can export its datastore up to a maximum of ten client clusters.

All vSAN features are supported except for Data-in-Transit encryption, Cloud Native Storage (including vSAN Direct), Stretched Clusters, and 2-Node Clusters. Additionally, HCI Mesh will not support remote provisioning of File Services Shares, iSCSI volumes, or First-Class Disks (FCDs). File Services, FCDs, and the iSCSI service can be provisioned locally on clusters participating in a mesh topology but may not be provisioned on a remote vSAN datastore.

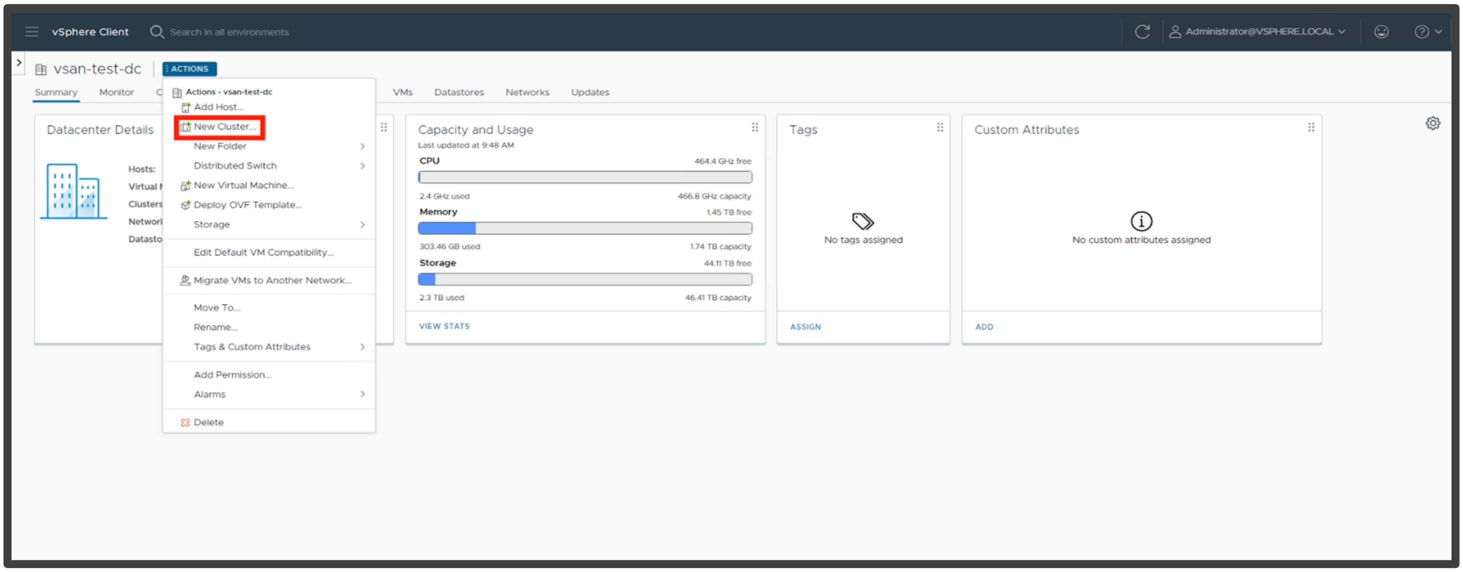
The same MTU sizing is required for both the Client and Server clusters.

Using Quickstart to Enable vSAN Max Cluster

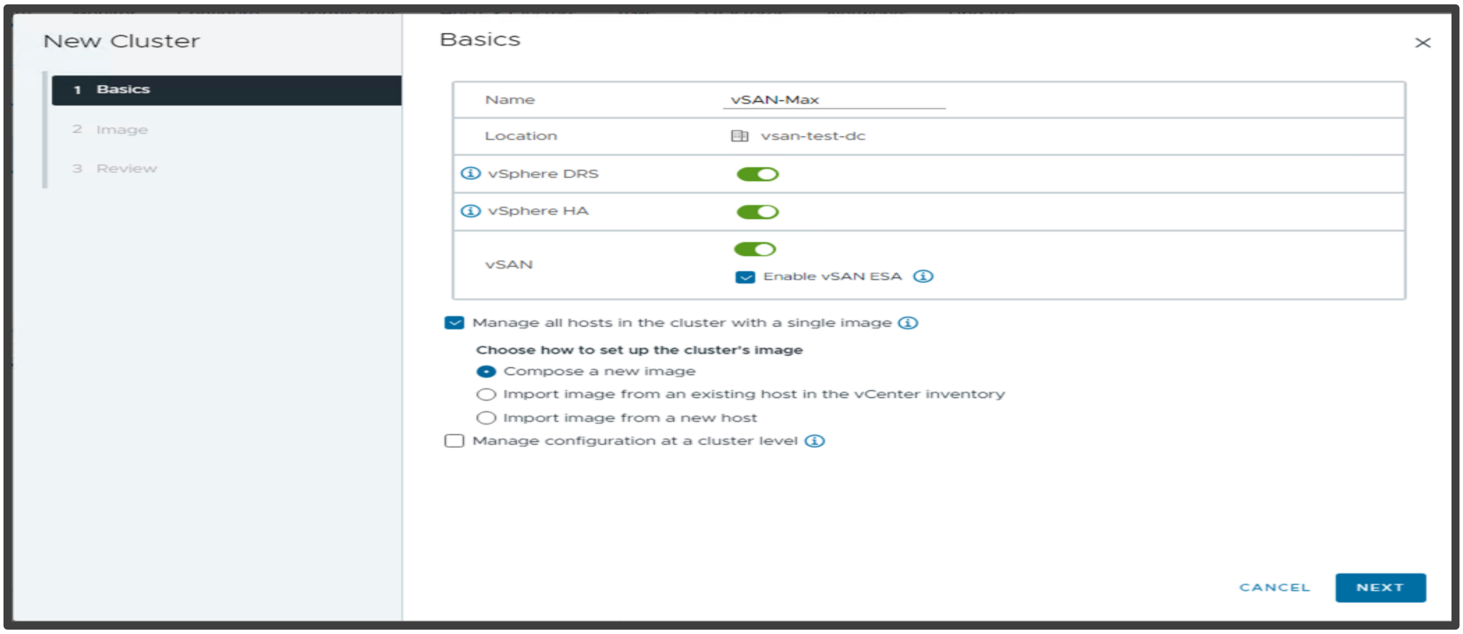
vSAN Max leverages vSAN ESA, as such the initial enablement process is very similar to the steps reviewed in the Using Quickstart to Enable Single vSAN HCI Cluster section of this document.

Initialize Cluster

Navigate to your **Datacenter** > Click **Actions** > **New Cluster**.



The New Cluster screen pops-up and we are presented with a dialog to enable services. Provide a name for the cluster and select vSAN from the list of services. Ensure that vSAN ESA (the default). For the Quickstart workflow to configure the vMotion VMkernel, vSphere DRS must be set to enabled.



Name	vSAN-Max
Location	vsan-test-dc
vSphere DRS	● on
vSphere HA	● on
vSAN	● on ☑ Enable vSAN ESA

☑ Manage all hosts in the cluster with a single image

Choose how to set up the cluster's image

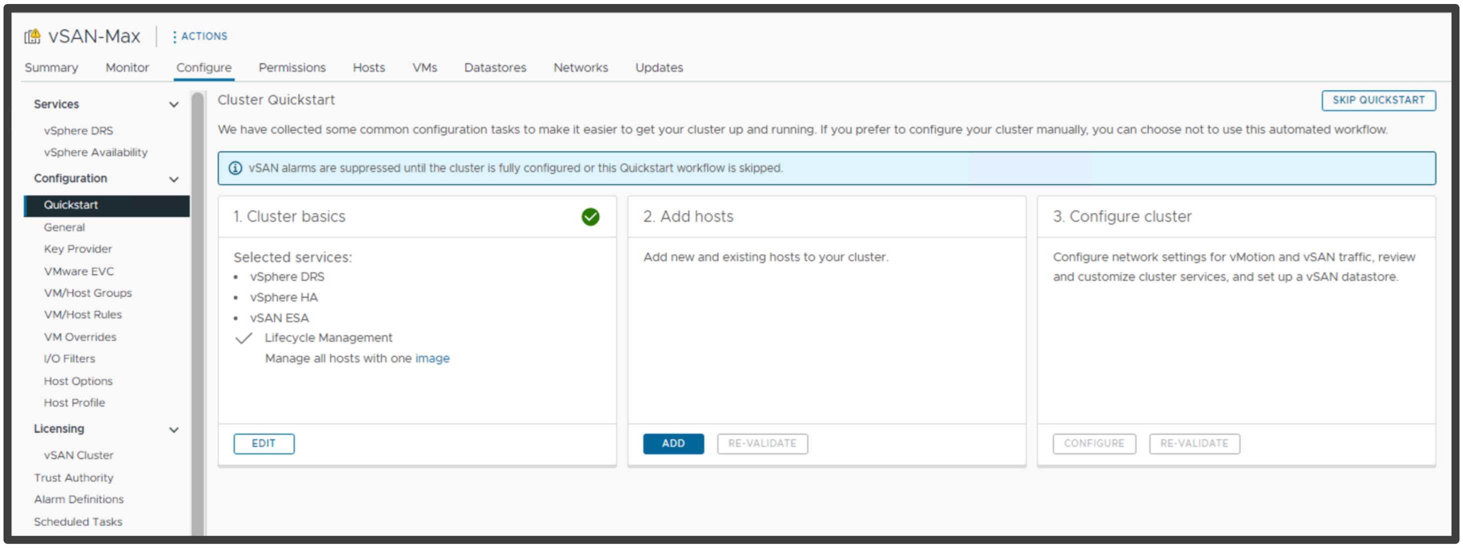
- ● Compose a new image
- ○ Import image from an existing host in the vCenter inventory
- ○ Import image from a new host

☐ Manage configuration at a cluster level

We can also setup the cluster to use a single image (thereby enabling vLCM). For more information on vLCM, see: https://core.vmware.com/resource/introducing-vsphere-lifecycle-management-vlcm.

Quickstart – Cluster Basics

The initial cluster creation above initializes the Quickstart process. Once the cluster has been created, navigate to **[vSAN Cluster] > Configure > Quickstart**. On this screen you will be able to confirm the basic services selected previously then move to the add hosts and configuration phases.



Cluster Quickstart

We have collected some common configuration tasks to make it easier to get your cluster up and running. If you prefer to configure your cluster manually, you can choose not to use this automated workflow.

vSAN alarms are suppressed until the cluster is fully configured or this Quickstart workflow is skipped.

1. Cluster basics ✓

Selected services:

- vSphere DRS
- vSphere HA
- vSAN ESA

✓ Lifecycle Management
Manage all hosts with one image

2. Add hosts

Add new and existing hosts to your cluster.

3. Configure cluster

Configure network settings for vMotion and vSAN traffic, review and customize cluster services, and set up a vSAN datastore.

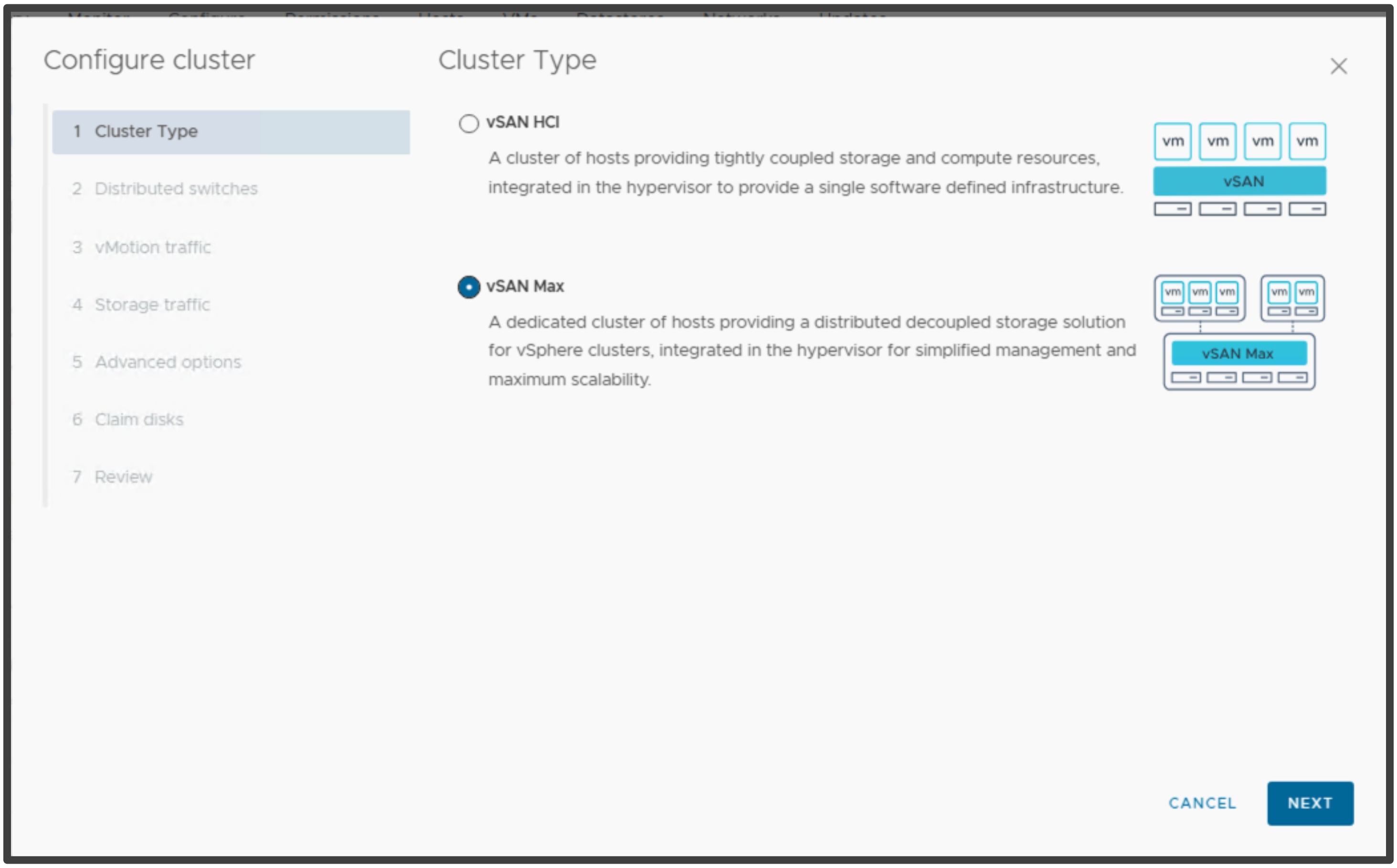
Quickstart – Add Hosts

The Adding Hosts steps for vSAN Max are identical to those for a single vSAN HCI cluster. Refer to the steps documented in the Enable a Single vSAN HCI, Quickstart – Add Hosts section of this document.

Quickstart – Configure Cluster

The next step is to configure the vSAN Max cluster. After clicking on **Configure** under Step 3: Configure Cluster, the Configure Cluster workflow will start. Ensure that **vSAN Max** is selected.

Step 1: Select Cluster Type



Configure cluster

1 Cluster Type
2 Distributed switches
3 vMotion traffic
4 Storage traffic
5 Advanced options
6 Claim disks
7 Review

Cluster Type

○ **vSAN HCI**

A cluster of hosts providing tightly coupled storage and compute resources, integrated in the hypervisor to provide a single software defined infrastructure.

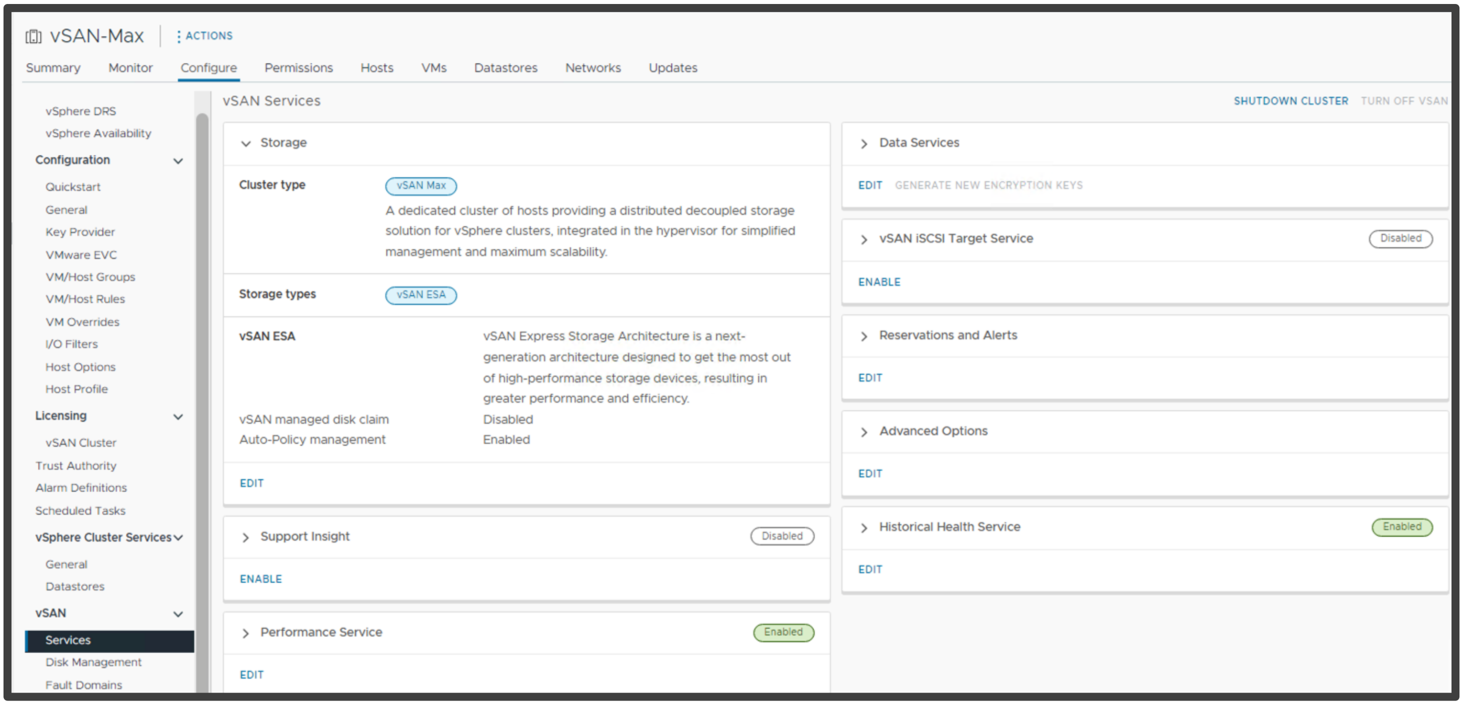
● **vSAN Max**

A dedicated cluster of hosts providing a distributed decoupled storage solution for vSphere clusters, integrated in the hypervisor for simplified management and maximum scalability.

Steps 2 -7: Configuring the Cluster

The remaining steps to configure the vSAN Max cluster are identical to those for a single vSAN HCI cluster. Refer to the steps documented in the Enable a Single vSAN HCI, Quickstart – Configure Cluster section of this document.

After the new vSAN Max cluster creation completes, navigate to **[vSAN Cluster] > Configure > vSAN > Services**. The screen will show that the vSAN Max cluster is ready to provide disaggregated storage to vSAN Computer clusters.



vSAN Services

Storage

Cluster type	vSAN Max A dedicated cluster of hosts providing a distributed decoupled storage solution for vSphere clusters, integrated in the hypervisor for simplified management and maximum scalability.
Storage types	vSAN ESA
vSAN ESA	vSAN Express Storage Architecture is a next-generation architecture designed to get the most out of high-performance storage devices, resulting in greater performance and efficiency.
vSAN managed disk claim	Disabled
Auto-Policy management	Enabled

EDIT

Support Insight — Disabled — ENABLE

Performance Service — Enabled — EDIT

Data Services — EDIT — GENERATE NEW ENCRYPTION KEYS

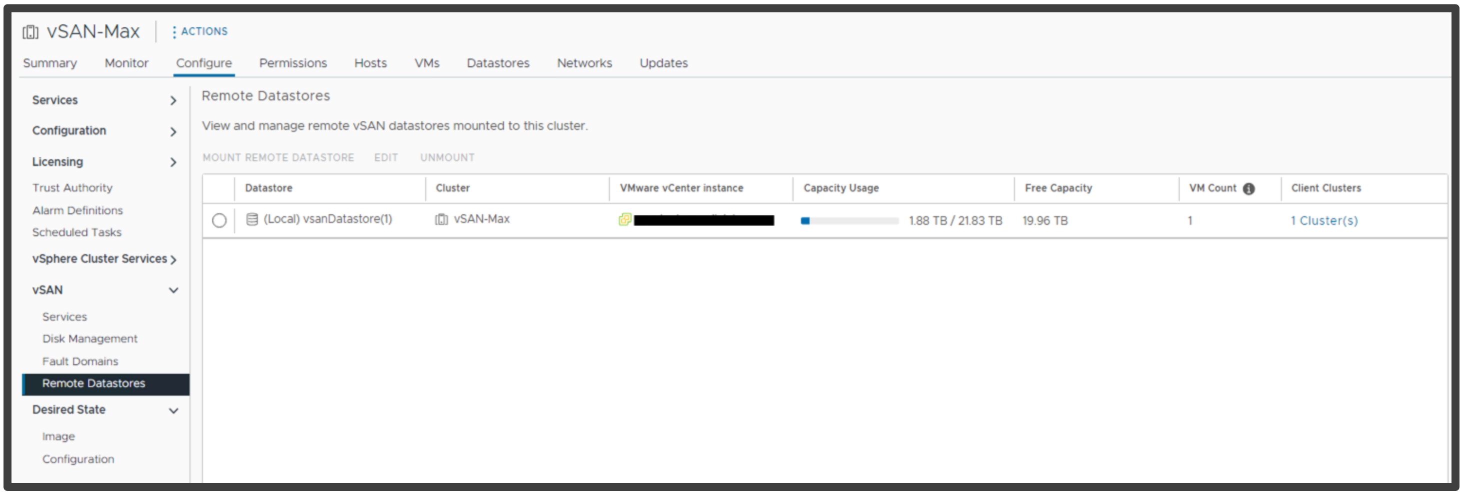
vSAN iSCSI Target Service — Disabled — ENABLE

Reservations and Alerts — EDIT

Advanced Options — EDIT

Historical Health Service — Enabled — EDIT

Now navigate to **[vSAN Cluster] > Configure > vSAN > Remote Datastore**. This screen shows the name of the remote datastore created by the vSAN Max cluster configuration workflow. The datastore name is a default name. If you wish to rename this datastore please refer to the Post-Configuration – Renaming vSAN Datastore section of this document.



Remote Datastores

View and manage remote vSAN datastores mounted to this cluster.

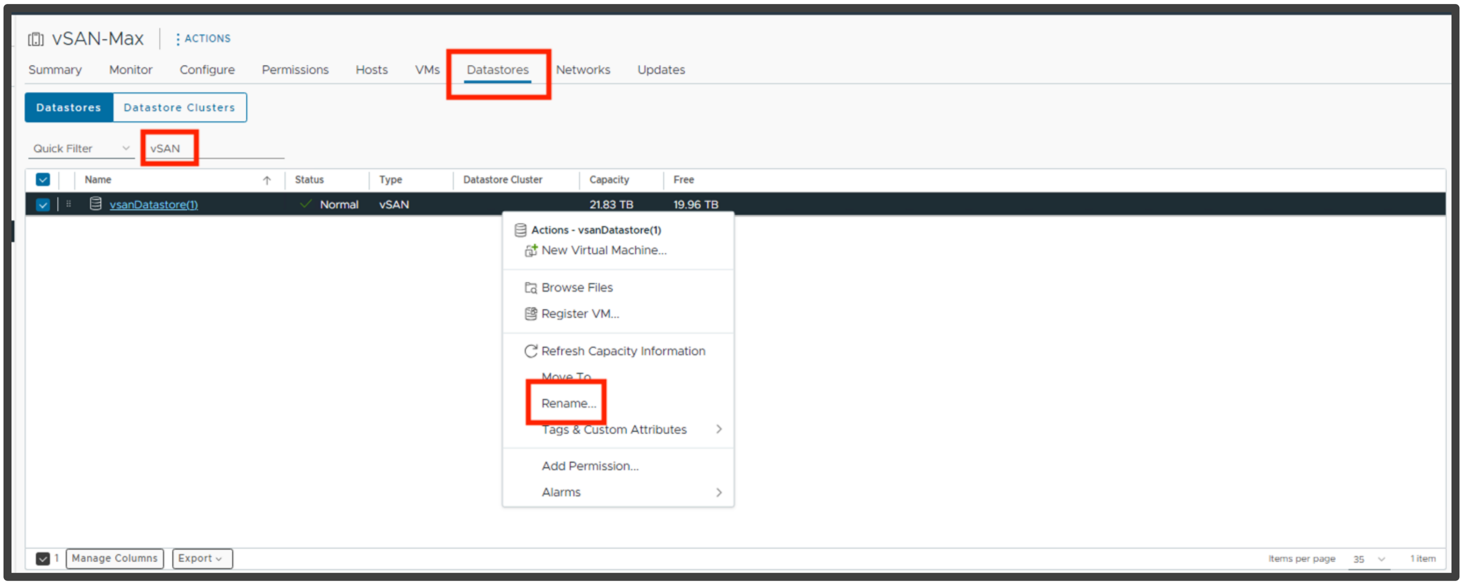
	Datastore	Cluster	VMware vCenter instance	Capacity Usage	Free Capacity	VM Count	Client Clusters
○	(Local) vsanDatastore(1)	vSAN-Max	[redacted]	1.88 TB / 21.83 TB	19.96 TB	1	1 Cluster(s)

If you do not wish to rename the datastore, you are now ready to configure a vSAN Compute cluster and mount this datastore to the Compute cluster. Go to the Enabling vSAN Compute Cluster section.

Post-Configuration – Renaming vSAN Datastore (optional)

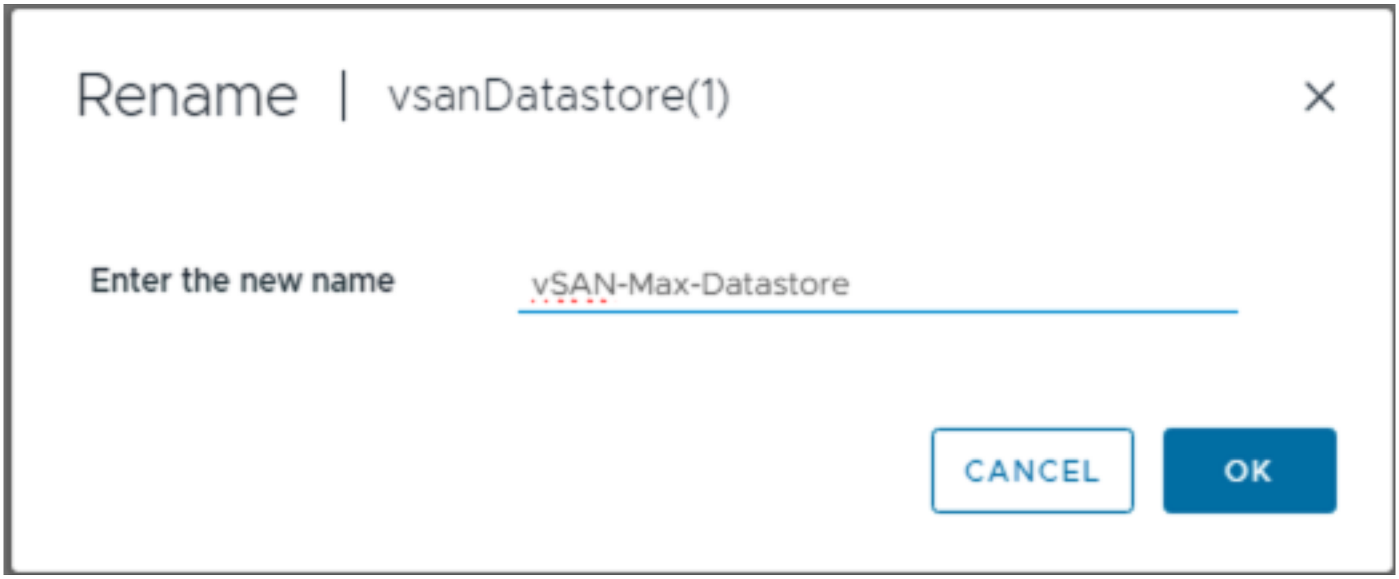
Once the vSAN Max cluster creation completes, the vSAN Max datastore is ready to be shared with vSAN Compute Clusters. The datastore will have the default name of "vsanDatastore." If the default name is not suitable for your environment, use these steps to rename the datastore as needed.

Navigate to [vSAN Cluster] > **Datastores**. Once on that screen, filter on vSAN (to make it easier to find the new datastore otherwise one may see the local datastore for each cluster host as well). Then right-click on the **vSAN datastore** and select **Rename**.



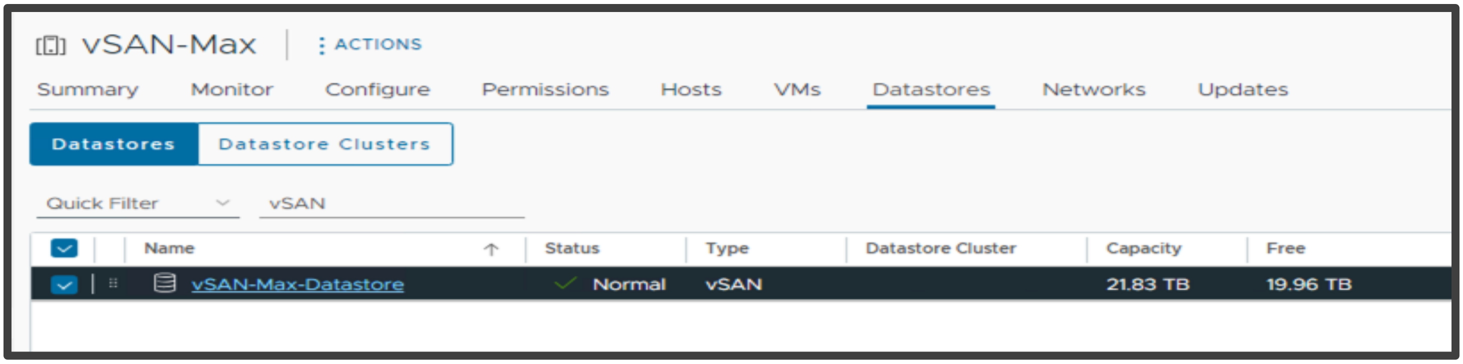
Name	Status	Type	Datastore Cluster	Capacity	Free
vsanDatastore(1)	Normal	vSAN		21.83 TB	19.96 TB

This will open the Rename workflow. In the workflow rename the datastore as needed then select **OK**



Enter the new name: vSAN-Max-Datastore

The datastore will now reflect the newly assigned name.



Name	Status	Type	Datastore Cluster	Capacity	Free
vSAN-Max-Datastore	✓ Normal	vSAN		21.83 TB	19.96 TB

You are now ready to configure a vSAN Compute cluster and mount this datastore to the Compute cluster. Go to the Enabling vSAN Compute Cluster section.

Manually Enabling vSAN Max on a Cluster

Note: If Quickstart was used (as per the earlier section) then this section can be skipped.

Manual vSAN Max enablement is available for those that do not wish to use the Quickstart process.

For this scenario, please follow the vSAN Max Cluster Services Configuration instructions in the vSAN Max Design and Operational Guidance document. Direct link to the section listed below:

https://core.vmware.com/resource/vsan-max-design-and-operational-guidance – sec32263-sub1

Enabling vSAN Max/HCI Mesh Services on a VMware Cloud Foundation™ based Cluster

VCF includes dedicated processes to automate the deployment and configuration of core infrastructure including vSAN services. In fact, these processes are required and are the only supported methods within VCF.

As of the writing of this guide, VCF 5.1 supports HCI Mesh. For more information, please review below.

HCI Mesh with VCF - https://docs.vmware.com/en/VMware-Cloud-Foundation/5.1/vcf-admin/GUID-1F86850D-E95E-40A8-AFC5-BE58D504D739.html

Enabling vSAN Compute Cluster

Overview

A vSAN Compute cluster is simply a vSphere cluster that has a thin layer of vSAN installed for the purposes of mounting the remote vSAN Max datastore.

The configuration process is solely manual. We do not have a Quickstart process for configuring vSAN Compute clusters. For this scenario, please follow the vSAN Compute Cluster instructions in the vSAN Max Design and Operational Guidance document. Direct link to the section is listed below:

https://core.vmware.com/resource/vsan-max-design-and-operational-guidance#sec32263-sub2

Prerequisites:

- The hosts in the vSAN Compute cluster will need a vSAN VMkernel configured. (routable to the vSAN network used by the target vSAN Max cluster)
- If the vSAN Compute cluster is initialized as a vSAN ESA cluster, it will only mount vSAN ESA/vSAN Max datastores
- If the vSAN Compute cluster is initialized as a vSAN OSA cluster, it will be able to mount vSAN OSA as well as vSAN ESA/Max datastores
- If direct access to the datastore is not required, one can configure vSAN File Services
- For more information on configuring vSAN File Shares refer to the “vSAN Proof of Concept: vSAN Features” guide

Enabling vSAN Compute Clusters in VMware Cloud Foundation

VCF includes dedicated processes to automate the deployment and configuration of core infrastructure including vSAN services. In fact, these processes are required and are the only supported methods within VCF.

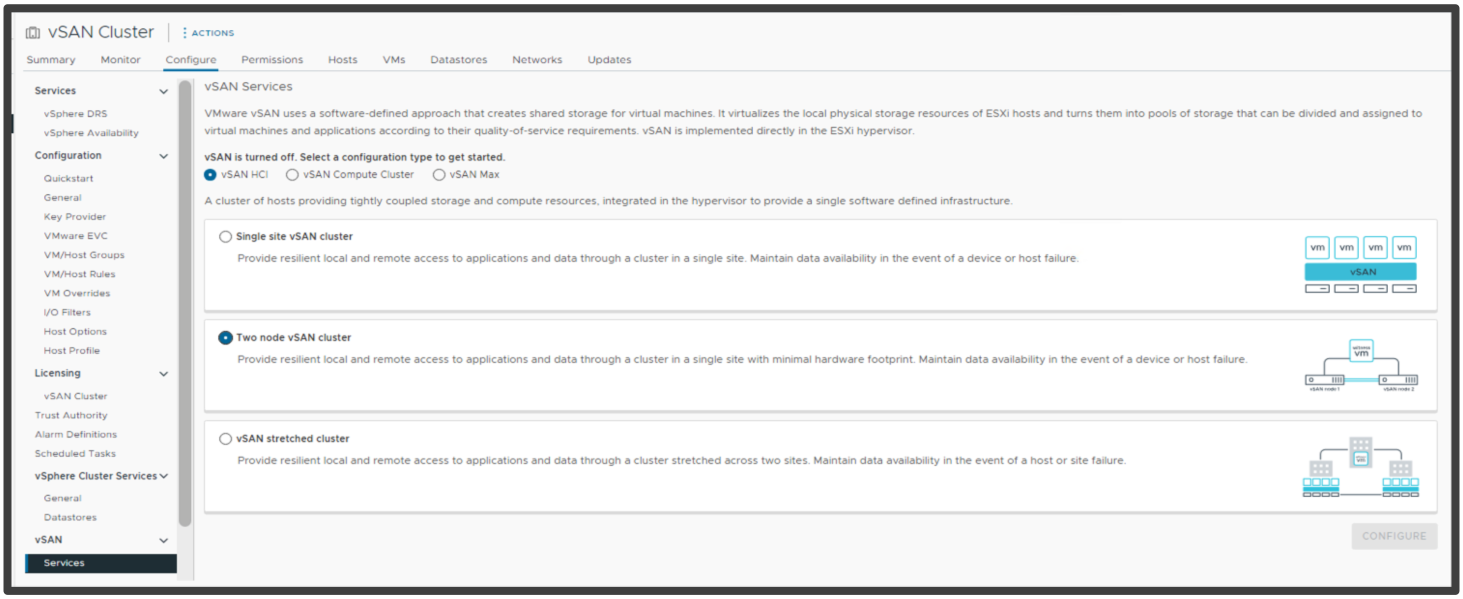
As of the writing of this guide, VCF 5.1 supports vSAN Compute clusters. For more information, please review below:

- Commissioning Hosts into VCF - https://docs.vmware.com/en/VMware-Cloud-Foundation/5.1/vcf-admin/GUID-45A77DE0-A38D-4655-85E2-BB8969C6993F.html
- HCI Mesh with VCF - https://docs.vmware.com/en/VMware-Cloud-Foundation/5.1/vcf-admin/GUID-1F86850D-E95E-40A8-AFC5-BE58D504D739.html

Enabling Two-Node and vSAN Stretched Clusters

Overview

To configure a vSAN two node cluster or stretched, select the appropriate option from the options given in **[vSAN Cluster] > Configure > Services:**



vSAN Services

VMware vSAN uses a software-defined approach that creates shared storage for virtual machines. It virtualizes the local physical storage resources of ESXi hosts and turns them into pools of storage that can be divided and assigned to virtual machines and applications according to their quality-of-service requirements. vSAN is implemented directly in the ESXi hypervisor.

vSAN is turned off. Select a configuration type to get started.

● vSAN HCI ○ vSAN Compute Cluster ○ vSAN Max

A cluster of hosts providing tightly coupled storage and compute resources, integrated in the hypervisor to provide a single software defined infrastructure.

○ **Single site vSAN cluster**

Provide resilient local and remote access to applications and data through a cluster in a single site. Maintain data availability in the event of a device or host failure.

● **Two node vSAN cluster**

Provide resilient local and remote access to applications and data through a cluster in a single site with minimal hardware footprint. Maintain data availability in the event of a device or host failure.

○ **vSAN stretched cluster**

Provide resilient local and remote access to applications and data through a cluster stretched across two sites. Maintain data availability in the event of a host or site failure.

Note: A witness appliance is required for two node and stretched vSAN clusters.

For more information on vSAN two-node and stretched cluster, see the '"vSAN Proof of Concept: vSAN Stretched Cluster and Two-Node Overview and Testing" guide.

Enabling vSAN Stretched Clusters in VMware Cloud Foundation

VCF includes dedicated processes to automate the deployment and configuration of core infrastructure including vSAN services. In fact, these processes are required and are the only supported methods within VCF.

You can stretch a vSAN cluster in a workload domain across two availability zones within a region. Both availability zones must contain an equal number of hosts to ensure failover in case any of the availability zones goes down.

The default management vSphere cluster must be stretched before a VI workload domain cluster can be stretched. This ensures that the NSX control plane and management VMs (vCenter, NSX management cluster, and SDDC Manager) remain accessible if the stretched cluster in the primary availability zone goes down.

You cannot stretch a cluster in the following conditions:

- The cluster uses vSAN ESA.
- The cluster has a vSAN remote datastore mounted on it.
- The cluster shares a vSAN Storage Policy with any other clusters.
- The cluster includes DPU-backed hosts.
- The cluster is enabled for Workload Management (vSphere with Tanzu).

Review this link for more information: https://docs.vmware.com/en/VMware-Cloud-Foundation/5.1/vcf-admin/GUID-7B4CC729-20BD-4CC9-B855-B38F02F74D40.html

Check the vSAN Cluster Thoroughly

Once the vSAN network has been created and vSAN is enabled, you should check that each ESXi host in the vSAN cluster is able to communicate to all other ESXi hosts in the cluster. The easiest way to achieve this is via the vSAN Health Check.

Why Is This Important?

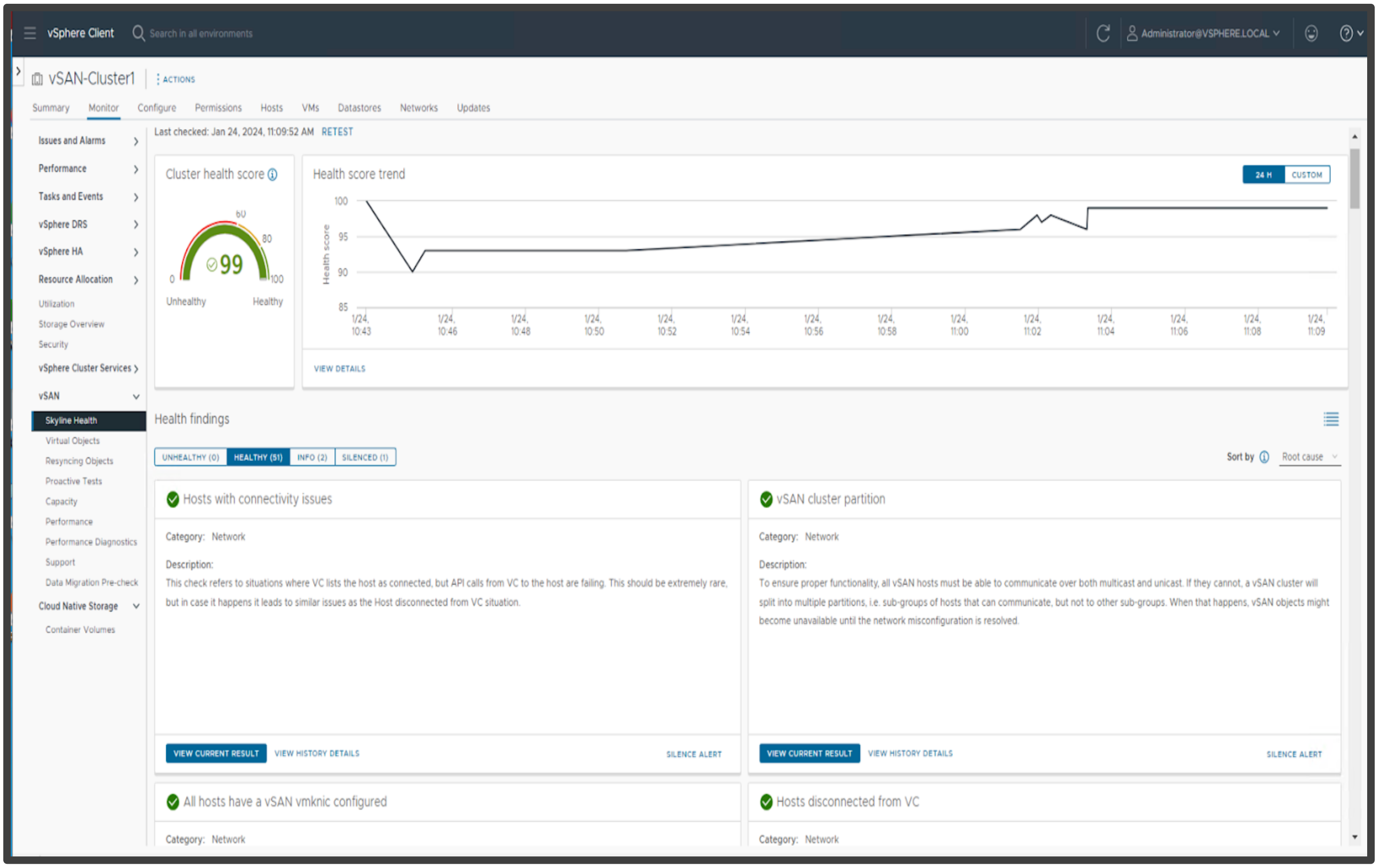
vSAN is dependent on the correct hardware and firmware combinations, as well as the network (configuration, reliability, performance, etc.). One of the most frequent causes of requesting support is either an incorrect network configuration or the network not performing as expected.

Use Health Check to Verify vSAN Functionality

Running individual commands from one host to all other hosts in the cluster can be tedious and time-consuming. vSAN has an integrated health check system. One of the first tasks to do after setting up any vSAN cluster is to perform a vSAN Health Check.

To run a vSAN Health Check, navigate to **[vSAN Cluster] > Monitor > vSAN > Skyline Health** and click the **RETEST** button.

Note: The Skyline Health interface provides the option to see the health findings in both a tile and list views. In this case, the screenshots below, show a “Cluster Health Score” of 99% with no Unhealthy findings in both tile and list views.



Last checked: Jan 24, 2024, 11:09:52 AM

✓ Hosts with connectivity issues

Category: Network

Description:
This check refers to situations where VC lists the host as connected, but API calls from VC to the host are failing. This should be extremely rare, but in case it happens it leads to similar issues as the Host disconnected from VC situation.

✓ vSAN cluster partition

Category: Network

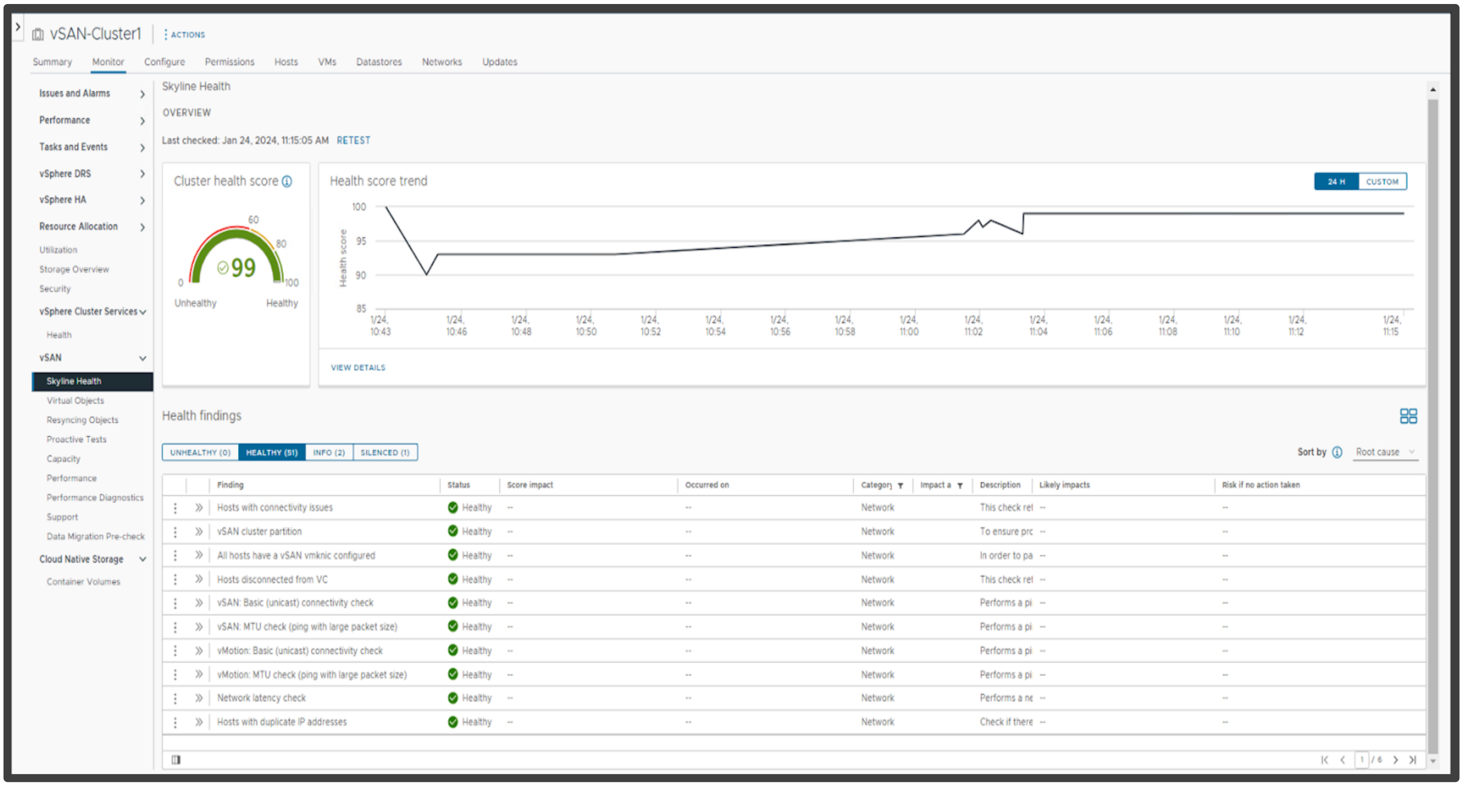
Description:
To ensure proper functionality, all vSAN hosts must be able to communicate over both multicast and unicast. If they cannot, a vSAN cluster will split into multiple partitions, i.e. sub-groups of hosts that can communicate, but not to other sub-groups. When that happens, vSAN objects might become unavailable until the network misconfiguration is resolved.

✓ All hosts have a vSAN vmknic configured

Category: Network

✓ Hosts disconnected from VC

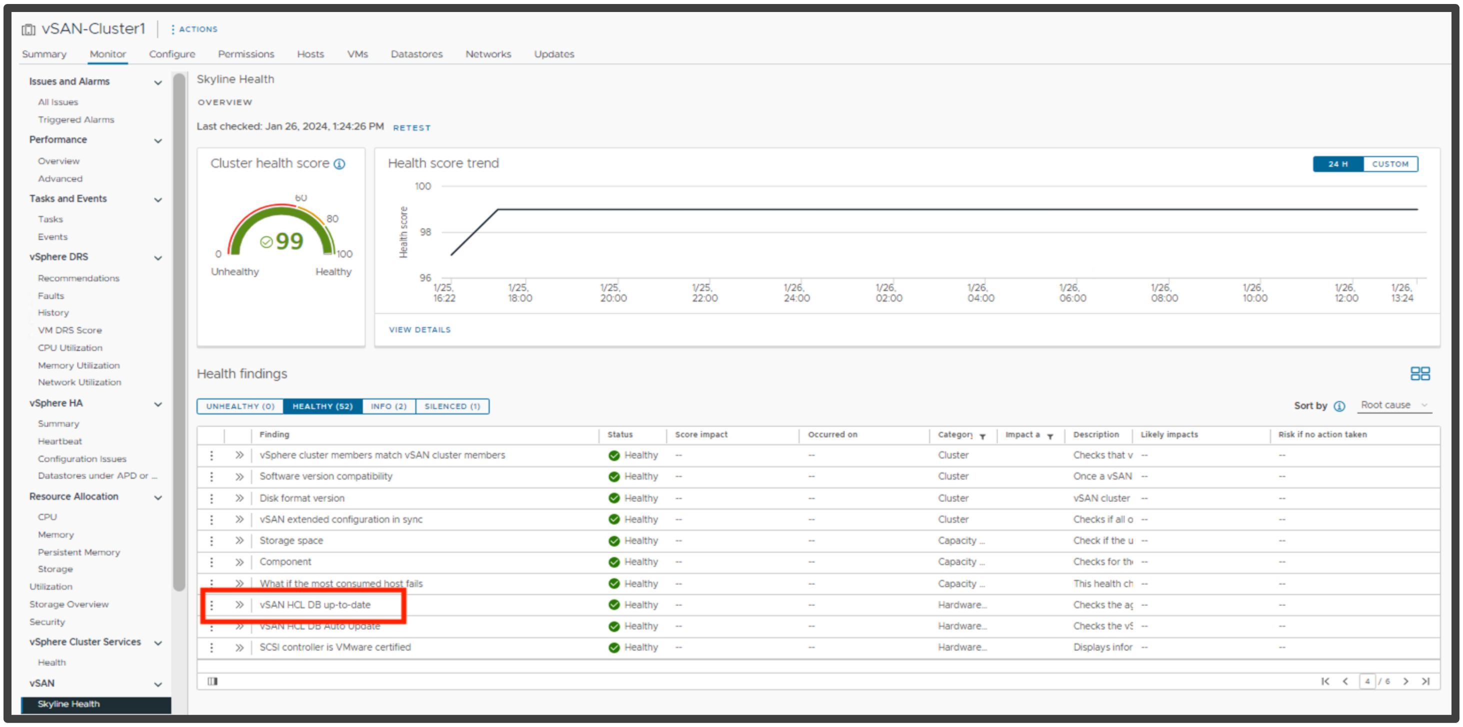
Category: Network



Last checked: Jan 24, 2024, 11:15:05 AM

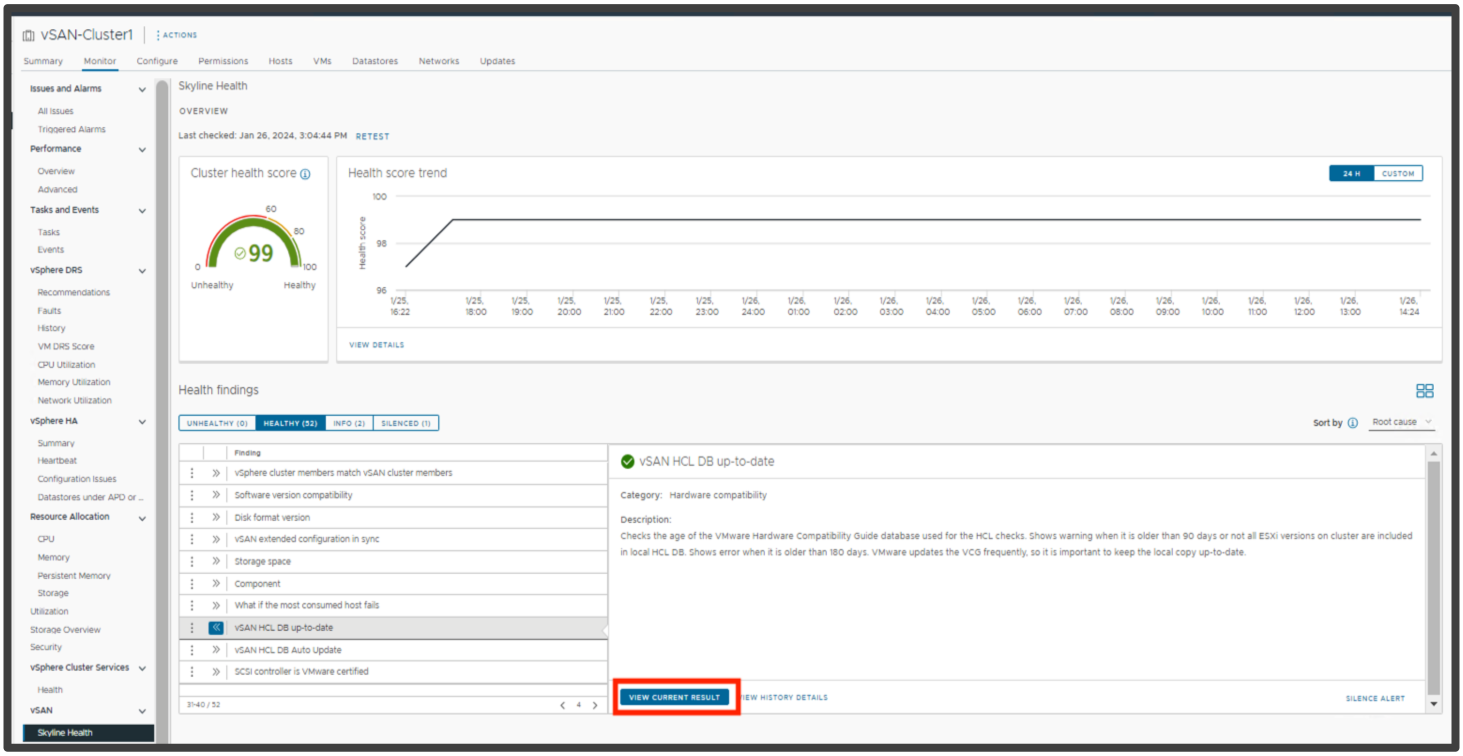
Finding	Status	Score impact	Occurred on	Category	Impact a	Description	Likely impacts	Risk if no action taken
Hosts with connectivity issues	✓ Healthy	--	--	Network		This check ret	--	--
vSAN cluster partition	✓ Healthy	--	--	Network		To ensure prc	--	--
All hosts have a vSAN vmknic configured	✓ Healthy	--	--	Network		In order to pa	--	--
Hosts disconnected from VC	✓ Healthy	--	--	Network		This check ret	--	--
vSAN: Basic (unicast) connectivity check	✓ Healthy	--	--	Network		Performs a pi	--	--
vSAN: MTU check (ping with large packet size)	✓ Healthy	--	--	Network		Performs a pi	--	--
vMotion: Basic (unicast) connectivity check	✓ Healthy	--	--	Network		Performs a pi	--	--
vMotion: MTU check (ping with large packet size)	✓ Healthy	--	--	Network		Performs a pi	--	--
Network latency check	✓ Healthy	--	--	Network		Performs a ne	--	--
Hosts with duplicate IP addresses	✓ Healthy	--	--	Network		Check if there	--	--

If any of the health checks fail, select the appropriate finding, and then click on the Troubleshoot button for information on how to resolve the issue. This brings up a new screen providing details on the cause and recommended fix. The screen also contains an **Ask VMware** button where appropriate, which will take you to a VMware Knowledge Base article detailing the issues, troubleshooting steps, and potential resolutions.



Last checked: Jan 26, 2024, 1:24:26 PM

Finding	Status	Score impact	Occurred on	Category	Impact a	Description	Likely impacts	Risk if no action taken
vSphere cluster members match vSAN cluster members	✓ Healthy	--	--	Cluster		Checks that v	--	--
Software version compatibility	✓ Healthy	--	--	Cluster		Once a vSAN	--	--
Disk format version	✓ Healthy	--	--	Cluster		vSAN cluster	--	--
vSAN extended configuration in sync	✓ Healthy	--	--	Cluster		Checks if all o	--	--
Storage space	✓ Healthy	--	--	Capacity ...		Check if the u	--	--
Component	✓ Healthy	--	--	Capacity ...		Checks for th	--	--
What if the most consumed host fails	✓ Healthy	--	--	Capacity ...		This health ch	--	--
vSAN HCL DB up-to-date	✓ Healthy	--	--	Hardware...		Checks the ag	--	--
vSAN HCL DB Auto Update	✓ Healthy	--	--	Hardware...		Checks the vS	--	--
SCSI controller is VMware certified	✓ Healthy	--	--	Hardware...		Displays infor	--	--



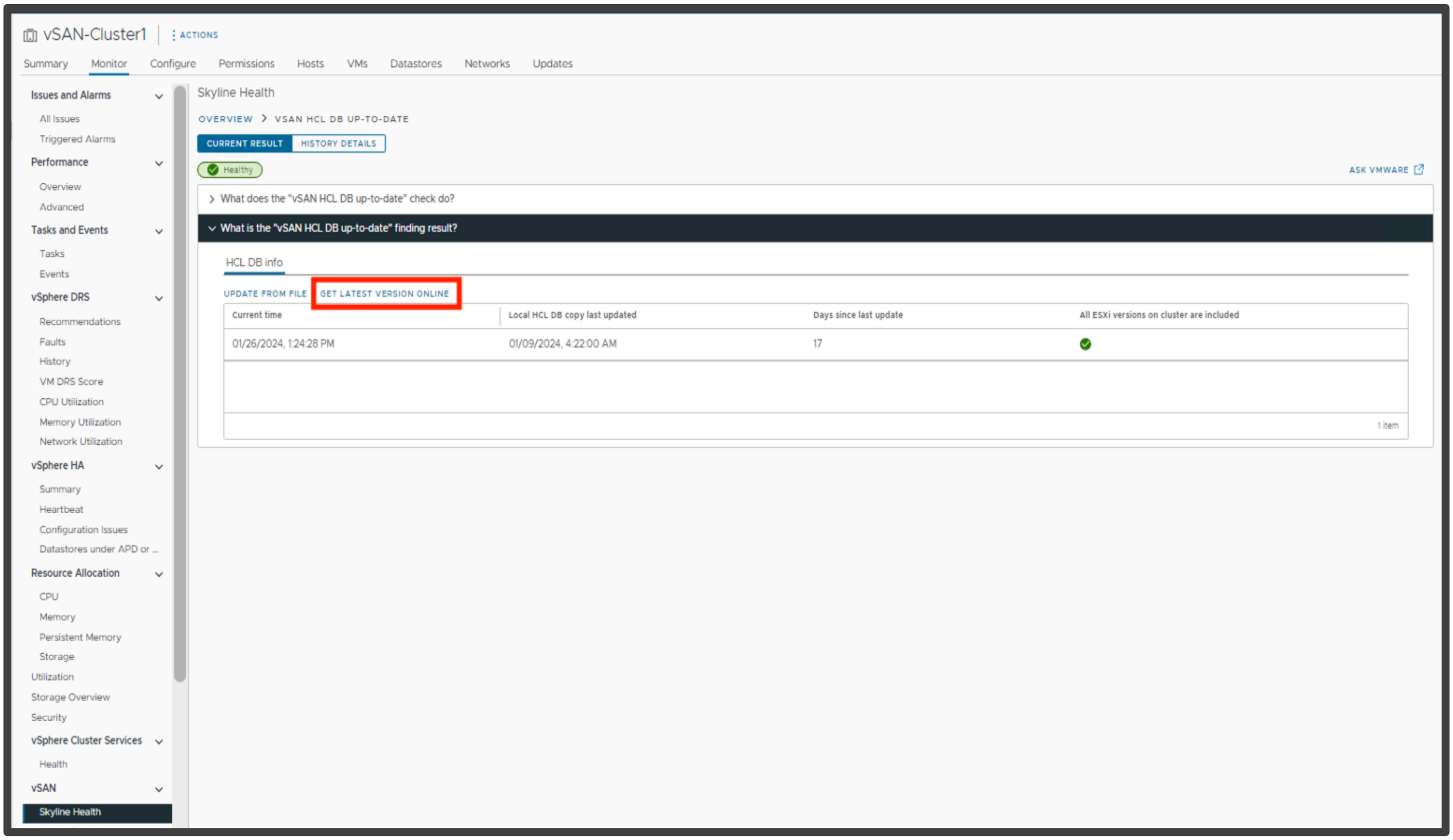
Last checked: Jan 26, 2024, 3:04:44 PM

Finding
vSphere cluster members match vSAN cluster members
Software version compatibility
Disk format version
vSAN extended configuration in sync
Storage space
Component
What if the most consumed host fails
vSAN HCL DB up-to-date
vSAN HCL DB Auto Update
SCSI controller is VMware certified

✓ vSAN HCL DB up-to-date

Category: Hardware compatibility

Description:
Checks the age of the VMware Hardware Compatibility Guide database used for the HCL checks. Shows warning when it is older than 90 days or not all ESXi versions on cluster are included in local HCL DB. Shows error when it is older than 180 days. VMware updates the VCG frequently, so it is important to keep the local copy up-to-date.



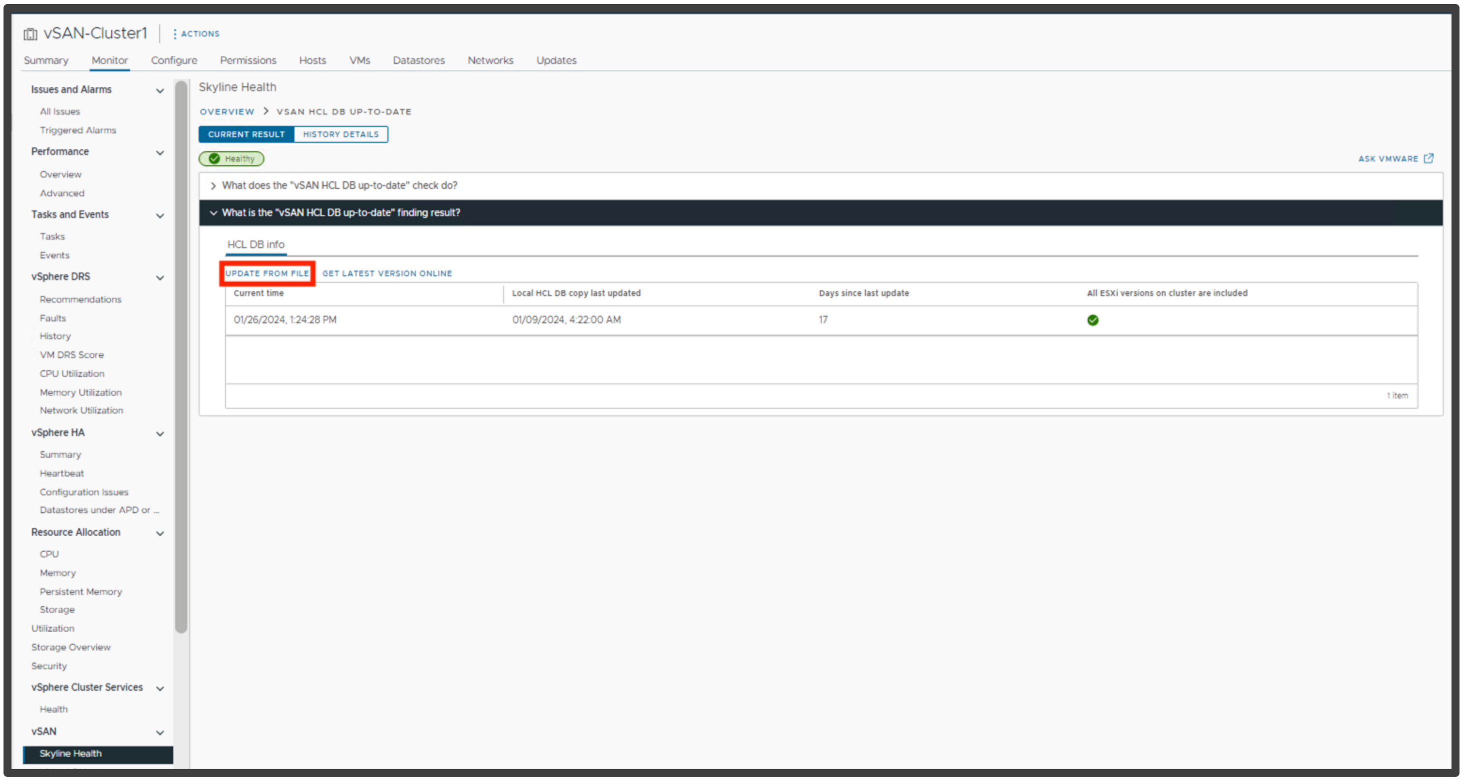
> What does the "vSAN HCL DB up-to-date" check do?

⌄ What is the "vSAN HCL DB up-to-date" finding result?

Current time	Local HCL DB copy last updated	Days since last update	All ESXi versions on cluster are included
01/26/2024, 1:24:28 PM	01/09/2024, 4:22:00 AM	17	✓

Ensure that the latest version of the HCL has been downloaded and run a **RETEST** on the Health check screen. Navigate to the vSAN HCL DB up-to-date health finding, expanding the finding, then click View Current Result. On the Result Screen select **GET LATEST VERSION ONLINE**.

If there is no internet connectivity, download the latest JSON from https://partnerweb.vmware.com/service/vsan/all.json (see https://kb.vmware.com/s/article/2145116 for more details) and select **UPDATE FROM FILE…**

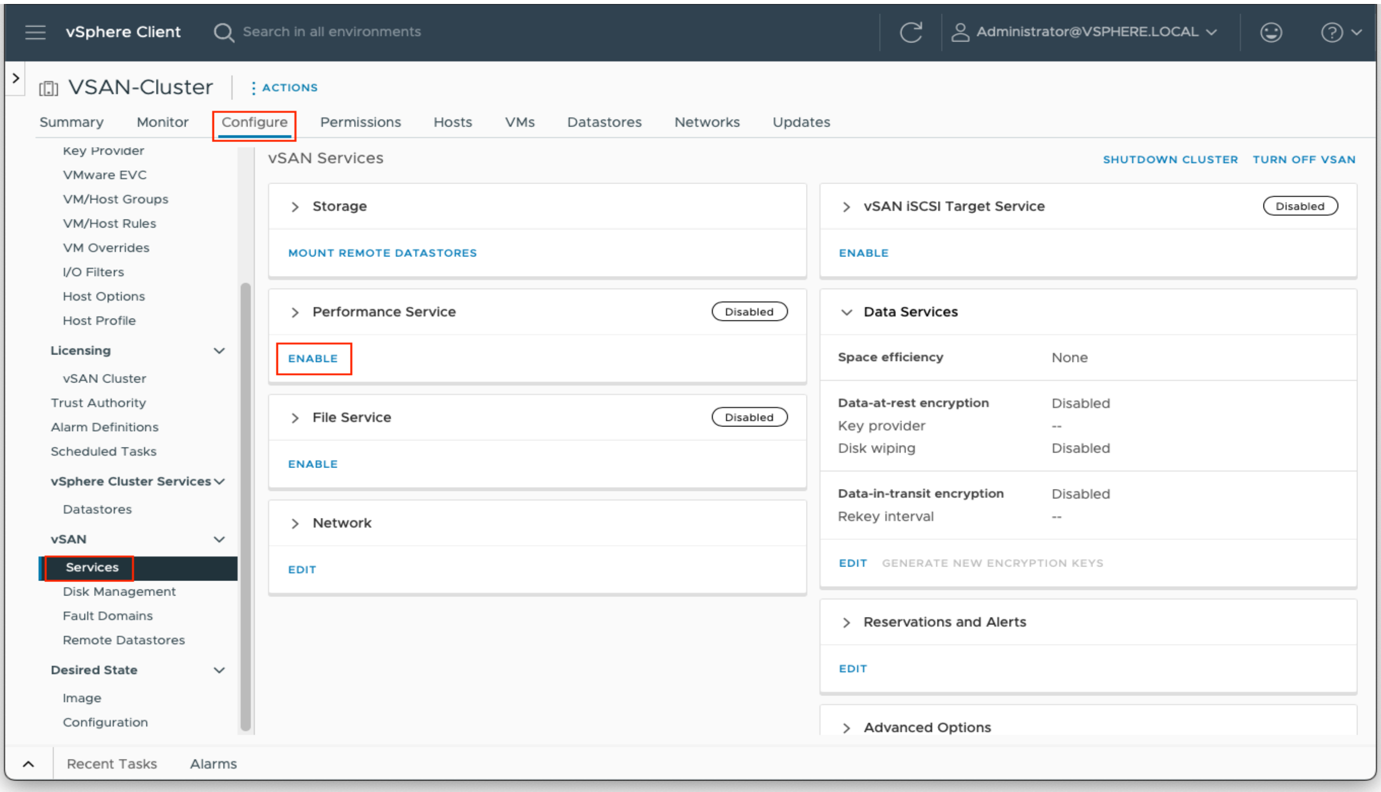


What does the "vSAN HCL DB up-to-date" check do?

What is the "vSAN HCL DB up-to-date" finding result?

Current time	Local HCL DB copy last updated	Days since last update	All ESXi versions on cluster are included
01/26/2024, 1:24:28 PM	01/09/2024, 4:22:00 AM	17	✓

The Performance Service is enabled by default. You can check its status from **[vSAN Cluster] > Configure > vSAN > Services**. If it needs to be manually enabled, click the **EDIT** button next to **Performance Service** and turn it on using the defaults. The Performance Service provides vSAN performance metrics to vCenter and other tools like Aria Operations.



Space efficiency None

Data-at-rest encryption Disabled
Key provider --
Disk wiping Disabled

Data-in-transit encryption Disabled
Rekey interval --

To ensure everything in the cluster is optimal, the health service will also check the hardware against the VMware Compatibility Guide (VCG) for vSAN, verify that the networking is functional, and that there are no underlying disk problems or vSAN integrity issues.

Manually Checking against the VCG

The following commands are useful to help identify firmware and drivers in ESXi for comparison with the VCG. First, log in to an ESXi host via SSH, then run the following commands to obtain the information from the server:

See the controller details:

```
esxcli vsan debug controller list
```

List VID DID SVID SSID of a storage controller or network adapter:

```
vmkchdev -l | egrep 'vmnic|vmhba'
```

Show which NIC driver is loaded:

```
esxcli network nic list
```

Show which storage controller driver is loaded:

```
esxcli storage core adapter list
```

Display driver version information:

```
vmkload_mod -s <driver-name> | grep -i version
```

Display NVMe driver information:

```
esxcli system module get -m nvme_pcie
```

For NVMe device info (replace X with the appropriate value):

```
esxcli nvme device get -A vmhbaX | egrep "Serial|Model|Firmware"
```

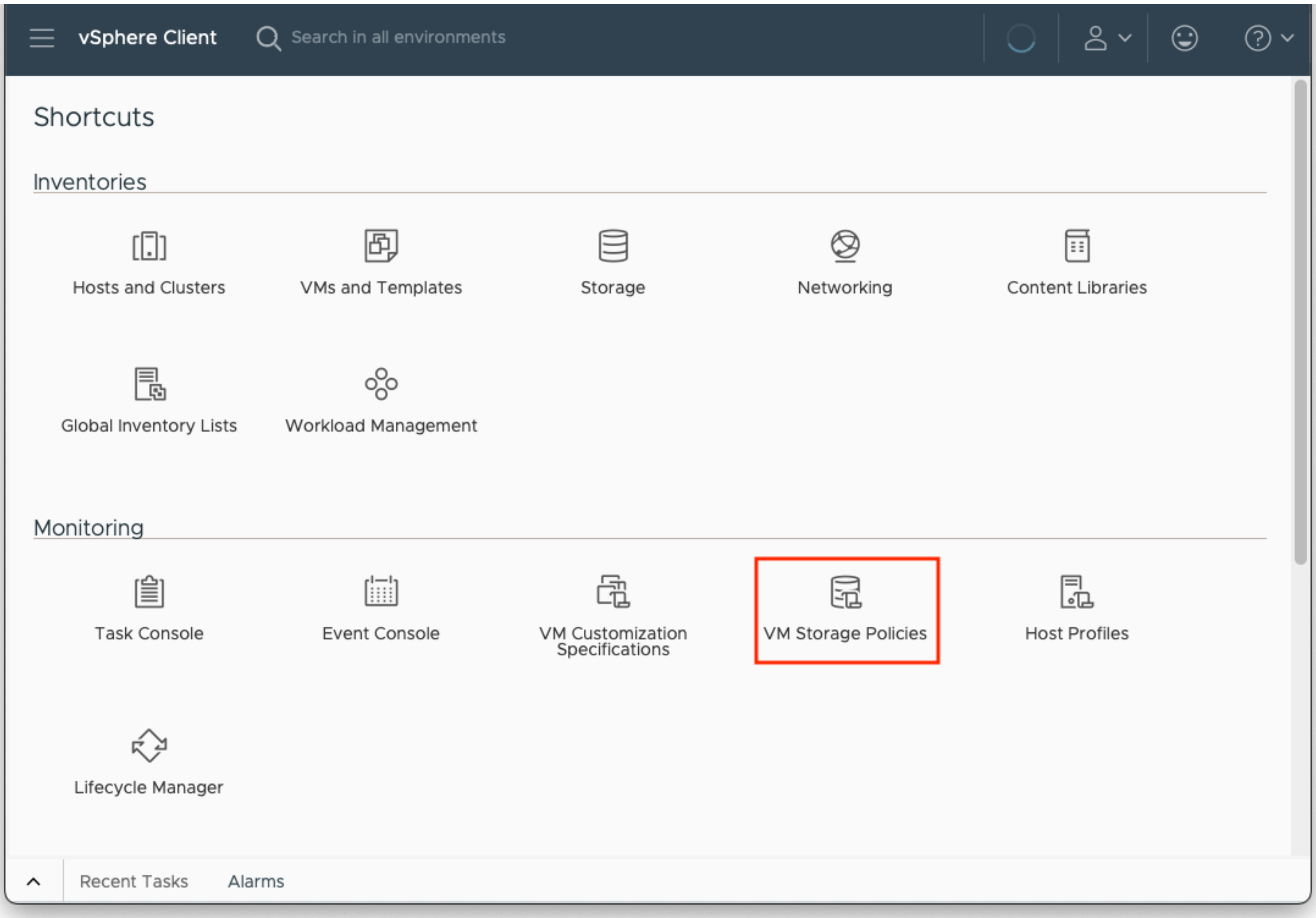
vSAN Basics

Deploy your first Virtual Machine

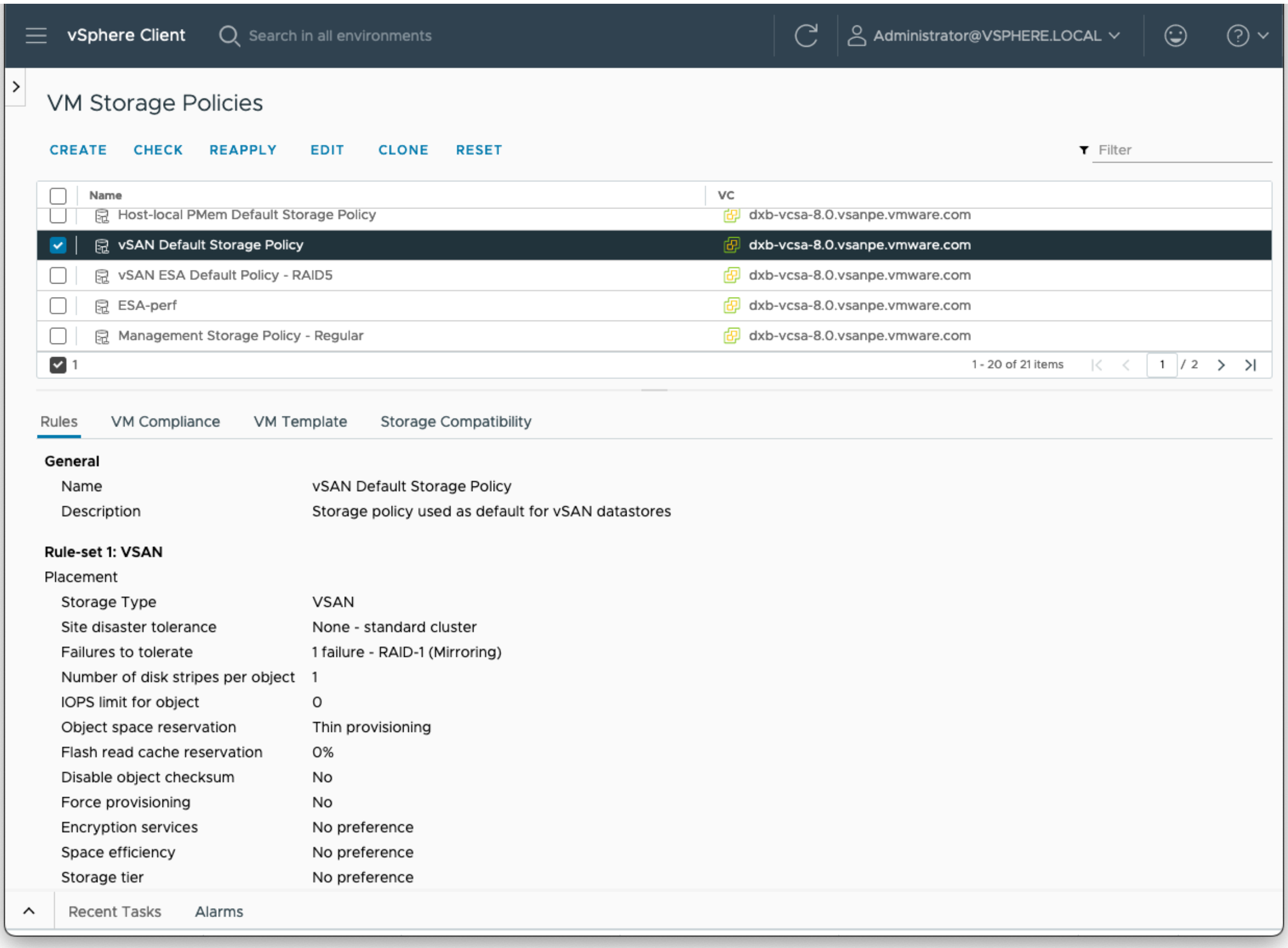
In this section, a VM is deployed to the vSAN datastore using the 'vSAN default storage policy', which stipulates a simple RAID-1 mirror.

Note: Due to the way data is structured in vSAN ESA, it recommended in most circumstances to define a RAID-5 policy for VMs.

To examine the default policy settings, navigate to **Menu** > **Shortcuts** >**VM Storage Policies**.



From there, select **vSAN Default Storage Policy**. Look under the Rules tab to see the settings on the policy:



VM Storage Policies

	Name	VC
□	Host-local PMem Default Storage Policy	dxb-vcsa-8.0.vsanpe.vmware.com
☑	vSAN Default Storage Policy	dxb-vcsa-8.0.vsanpe.vmware.com
□	vSAN ESA Default Policy - RAID5	dxb-vcsa-8.0.vsanpe.vmware.com
□	ESA-perf	dxb-vcsa-8.0.vsanpe.vmware.com
□	Management Storage Policy - Regular	dxb-vcsa-8.0.vsanpe.vmware.com

☑ 1 1 - 20 of 21 items 1 / 2

General

Name	vSAN Default Storage Policy
Description	Storage policy used as default for vSAN datastores

Rule-set 1: VSAN

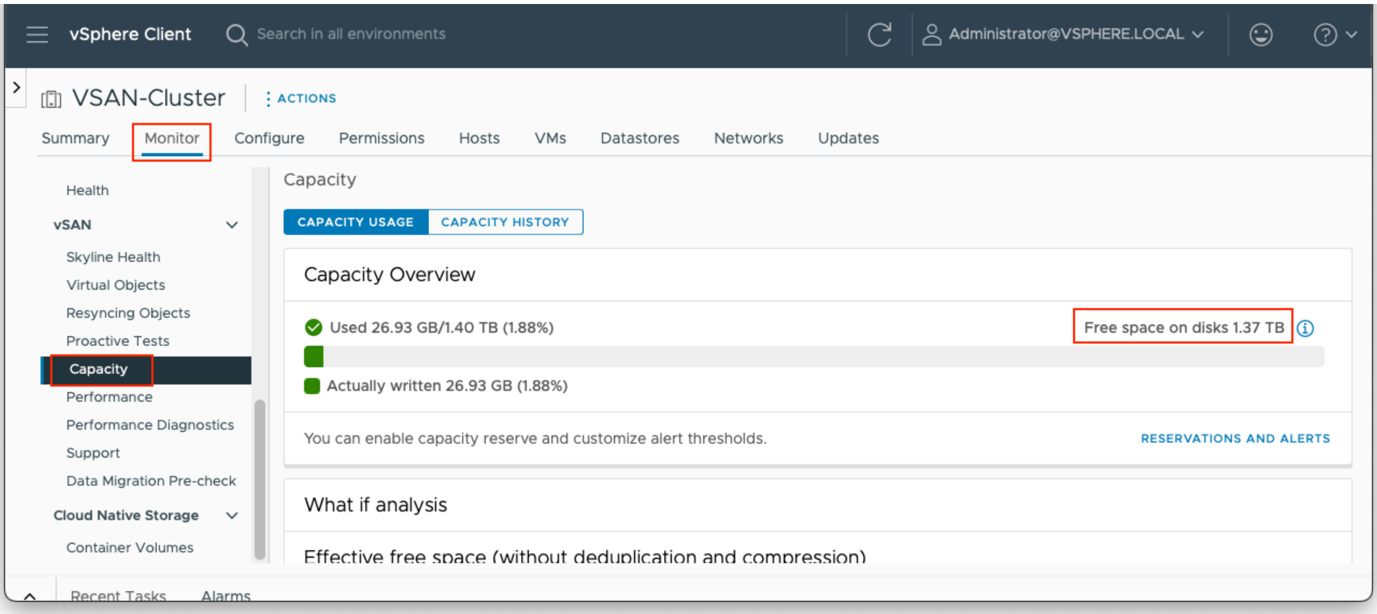
Placement

Storage Type	VSAN
Site disaster tolerance	None - standard cluster
Failures to tolerate	1 failure - RAID-1 (Mirroring)
Number of disk stripes per object	1
IOPS limit for object	0
Object space reservation	Thin provisioning
Flash read cache reservation	0%
Disable object checksum	No
Force provisioning	No
Encryption services	No preference
Space efficiency	No preference
Storage tier	No preference

We will return to VM Storage Policies in more detail later, but when a VM is deployed with the default policy, it should have a mirror copy of the VM data created. This second copy of the VM data is placed on storage on a different host or *fault domain* to enable the VM to tolerate any single failure.

Also note that object space reservation is set to 'Thin provisioning', meaning that the object should be deployed as "thin". After we have deployed the VM, we will verify that vSAN adheres to both capabilities.

One final item to check before we deploy the VM is the current free capacity on the vSAN datastore. This can be viewed from the [vSAN Cluster] > Monitor > vSAN > Capacity. In this example, it is 1.37 TB.

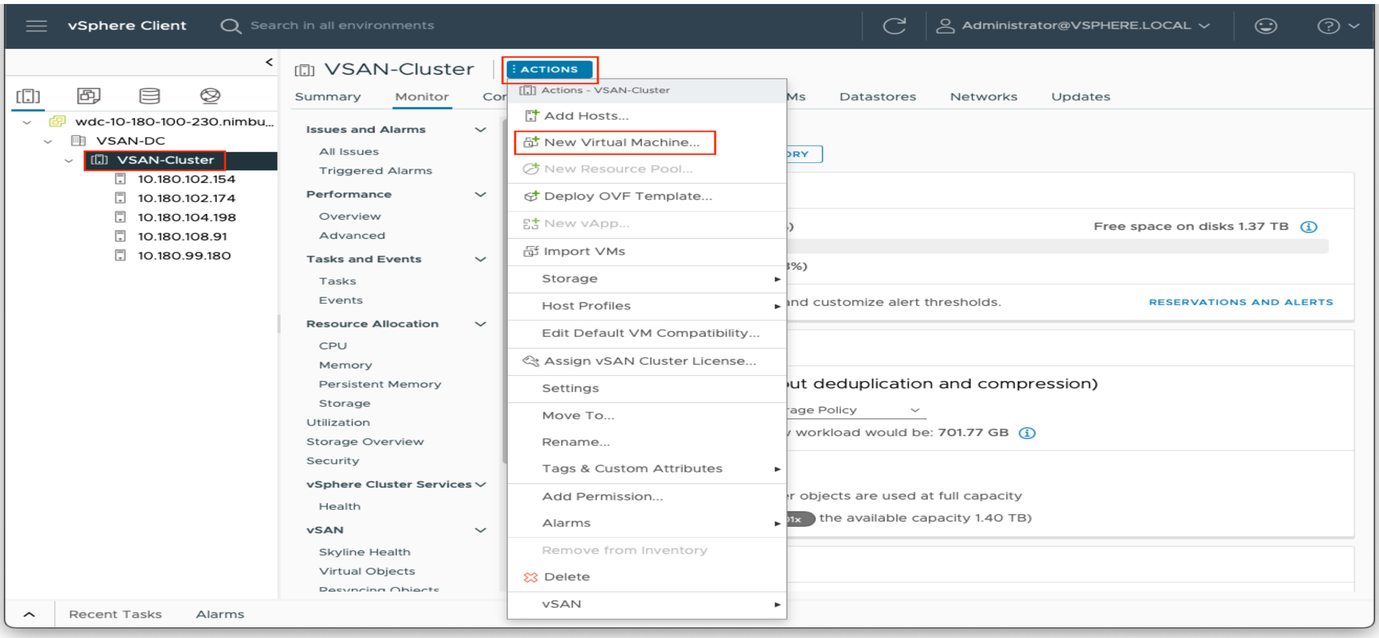


Make a note of the free capacity in your environment before continuing with the deploy VM exercise.

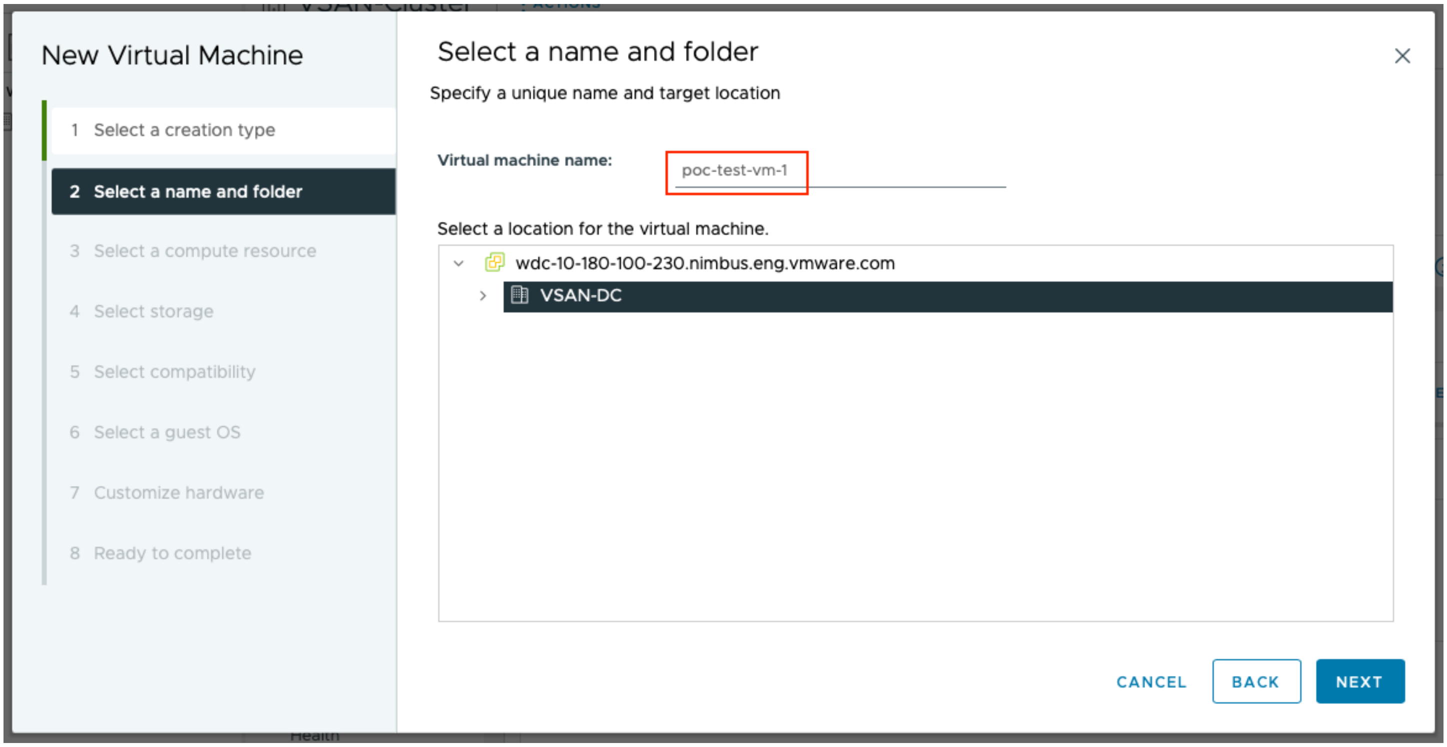
To deploy the VM, simply follow the steps provided in the wizard.

Select **New Virtual Machine** from the **Actions** Menu.

Select **Create a new virtual machine**.



Provide a name for the VM:



Select a name and folder

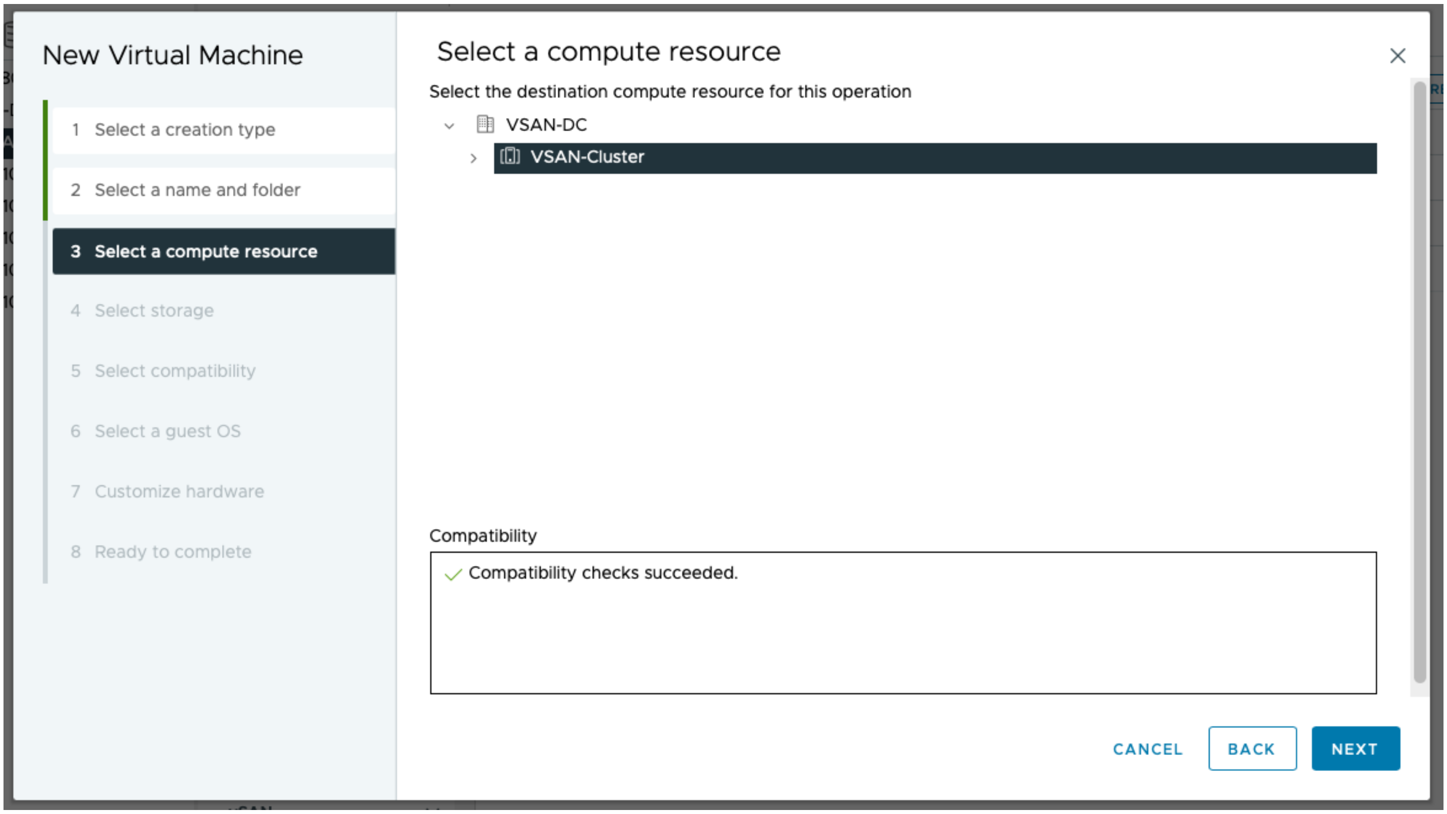
Specify a unique name and target location

Virtual machine name: poc-test-vm-1

Select a location for the virtual machine.

- wdc-10-180-100-230.nimbus.eng.vmware.com
  - VSAN-DC

Select a compute resource (if DRS is enabled on the cluster, select the cluster itself, otherwise select one of the hosts):



New Virtual Machine

1 Select a creation type
2 Select a name and folder
3 Select a compute resource
4 Select storage
5 Select compatibility
6 Select a guest OS
7 Customize hardware
8 Ready to complete

Select a compute resource

Select the destination compute resource for this operation

VSAN-DC
VSAN-Cluster

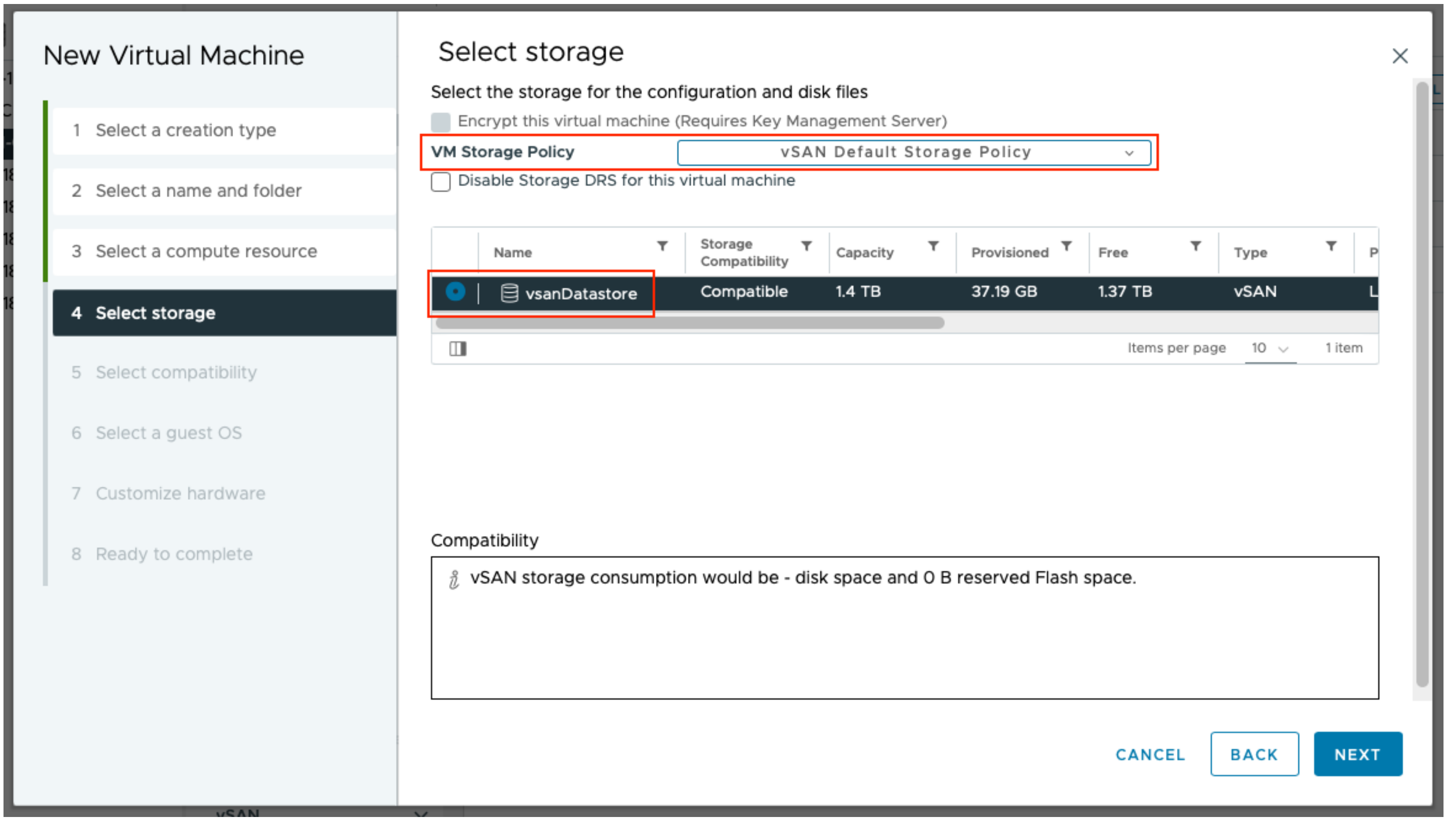
Compatibility

✓ Compatibility checks succeeded.

Up to this point, the virtual machine deployment process is identical to all other virtual machine deployments. It is the next section that may be unfamiliar: this is where a policy for the virtual machine is chosen.

As per the screenshot below, select change the VM storage policy to 'vSAN Default Storage Policy'.

Once the policy has been chosen, datastores are split into those that are either compliant or non-compliant with the selected policy. As seen below, only the vSAN datastore can utilize the policy settings in the vSAN Default Storage Policy, so it is the only one that shows up as Compatible in the list of datastores.



Select the storage for the configuration and disk files

□ Encrypt this virtual machine (Requires Key Management Server)

VM Storage Policy: vSAN Default Storage Policy

□ Disable Storage DRS for this virtual machine

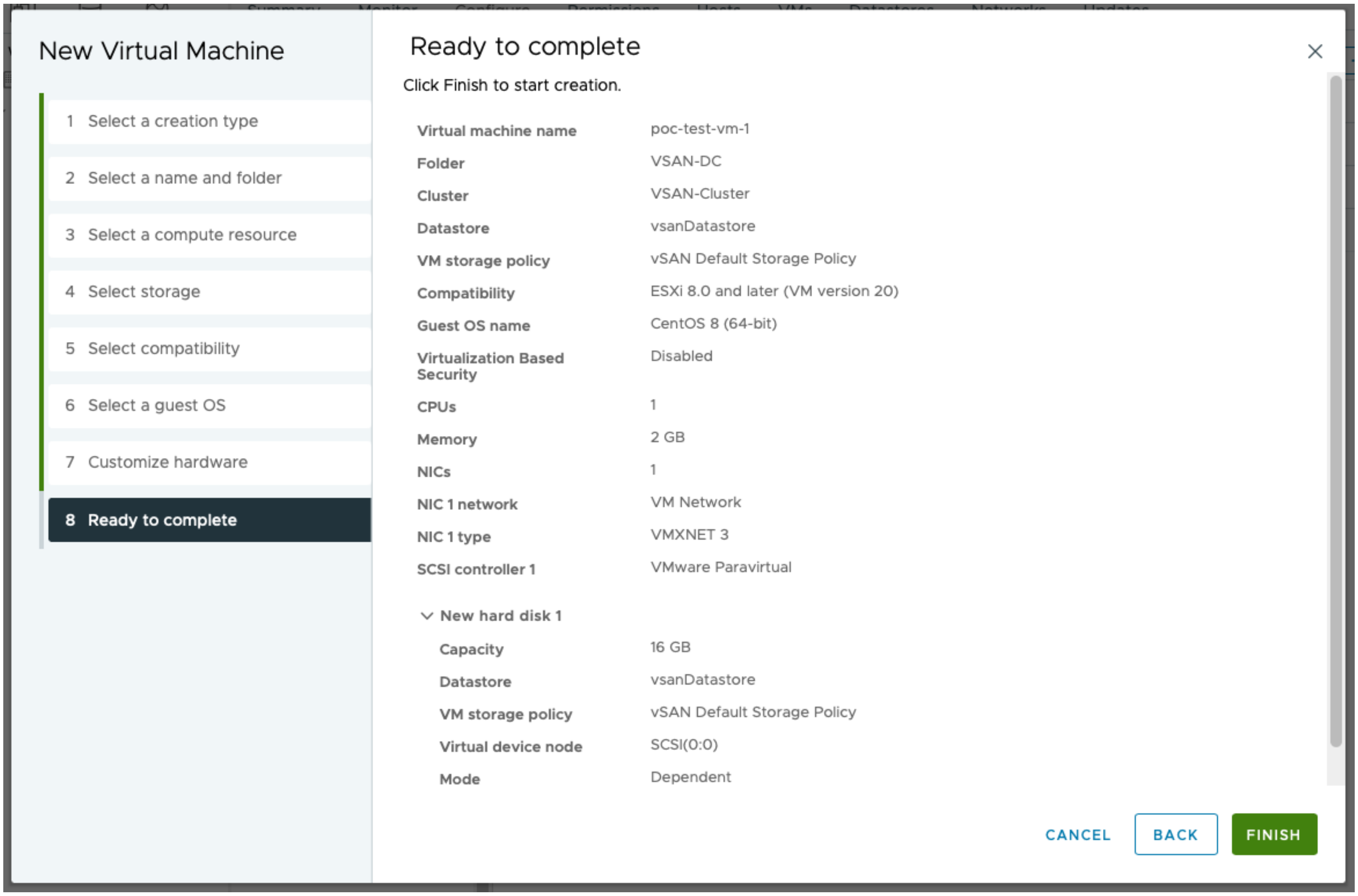
	Name	Storage Compatibility	Capacity	Provisioned	Free	Type
●	vsanDatastore	Compatible	1.4 TB	37.19 GB	1.37 TB	vSAN

Items per page 10 1 item

Compatibility

vSAN storage consumption would be - disk space and 0 B reserved Flash space.

The rest of the VM deployment steps in the wizard are quite straightforward, and simply entail selecting ESXi version compatibility (leave at default), a guest OS, and customize hardware (no changes needed).



Ready to complete

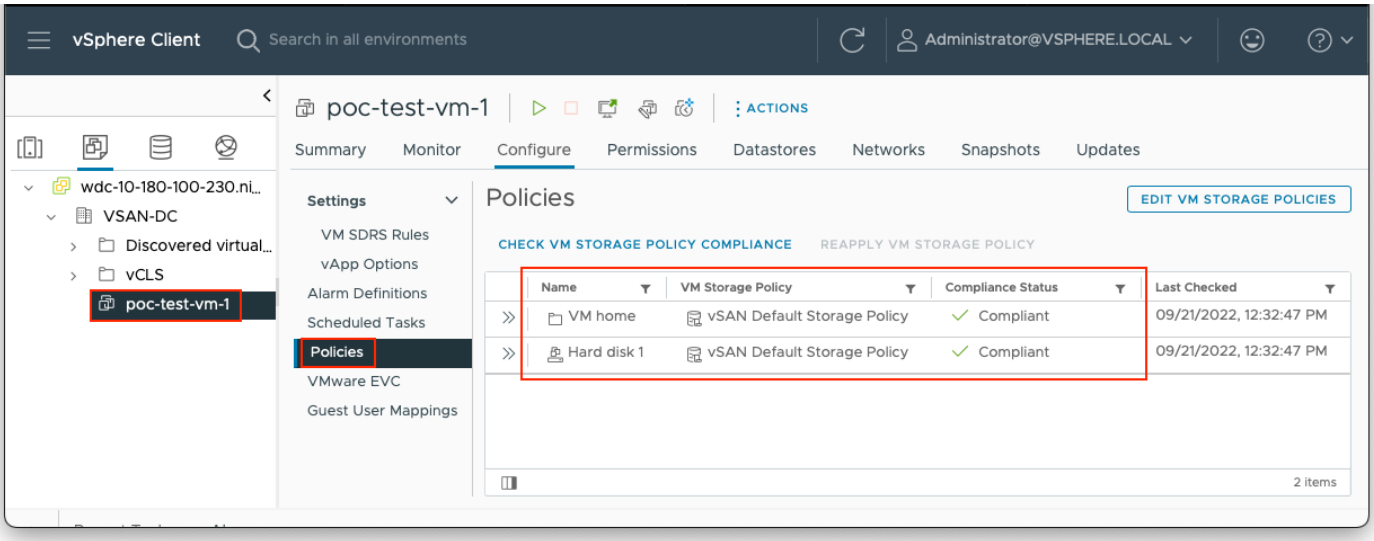
Click Finish to start creation.

Virtual machine name	poc-test-vm-1
Folder	VSAN-DC
Cluster	VSAN-Cluster
Datastore	vsanDatastore
VM storage policy	vSAN Default Storage Policy
Compatibility	ESXi 8.0 and later (VM version 20)
Guest OS name	CentOS 8 (64-bit)
Virtualization Based Security	Disabled
CPUs	1
Memory	2 GB
NICs	1
NIC 1 network	VM Network
NIC 1 type	VMXNET 3
SCSI controller 1	VMware Paravirtual
New hard disk 1	
Capacity	16 GB
Datastore	vsanDatastore
VM storage policy	vSAN Default Storage Policy
Virtual device node	SCSI(0:0)
Mode	Dependent

Verifying Disk Layout of a VM on vSAN

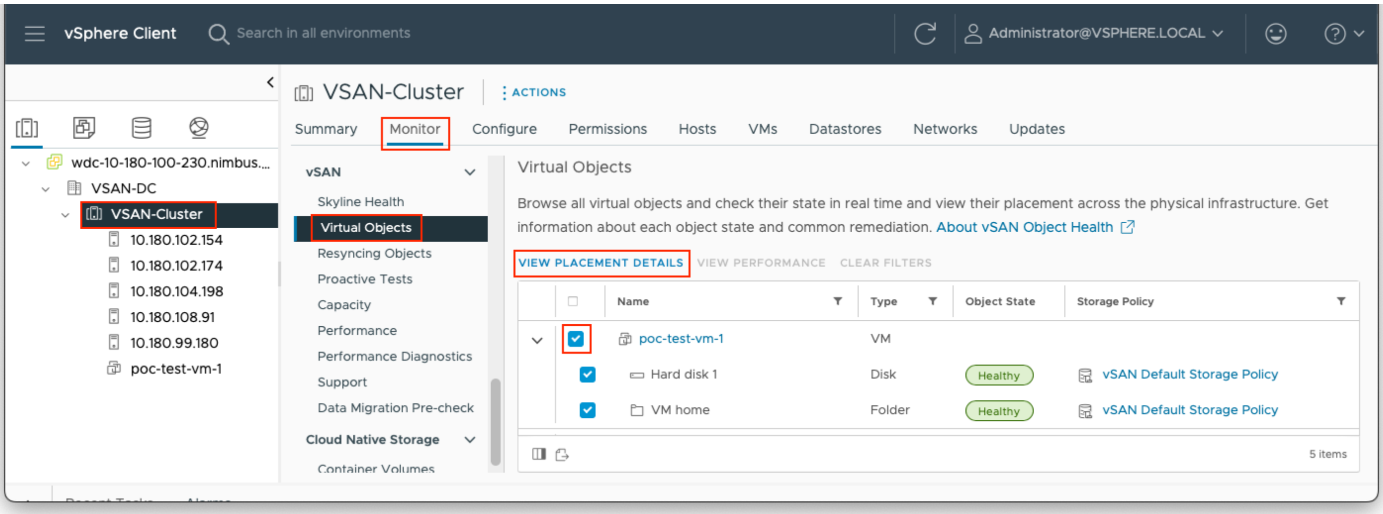
Physical placement of data on vSAN is defined by storage polices, which we will look at in more detail in another section. Here, we look at how the data is placed with vSAN default policy.

Once the VM is created, select the new VM in the inventory, navigate to the **Configure** tab, and then select **Policies**. There should be two objects shown, "VM home" and "Hard disk 1". Both should show a compliance status of *Compliant* meaning that vSAN was able to deploy these objects in accordance with the policy settings.



Name	VM Storage Policy	Compliance Status	Last Checked
VM home	vSAN Default Storage Policy	✓ Compliant	09/21/2022, 12:32:47 PM
Hard disk 1	vSAN Default Storage Policy	✓ Compliant	09/21/2022, 12:32:47 PM

To verify this, navigate to the **[vSAN Cluster] > Monitor > Virtual Objects**. Once again, both the “VM home” and “Hard disk 1” should be displayed. Select the VM, followed by **View Placement Details**.



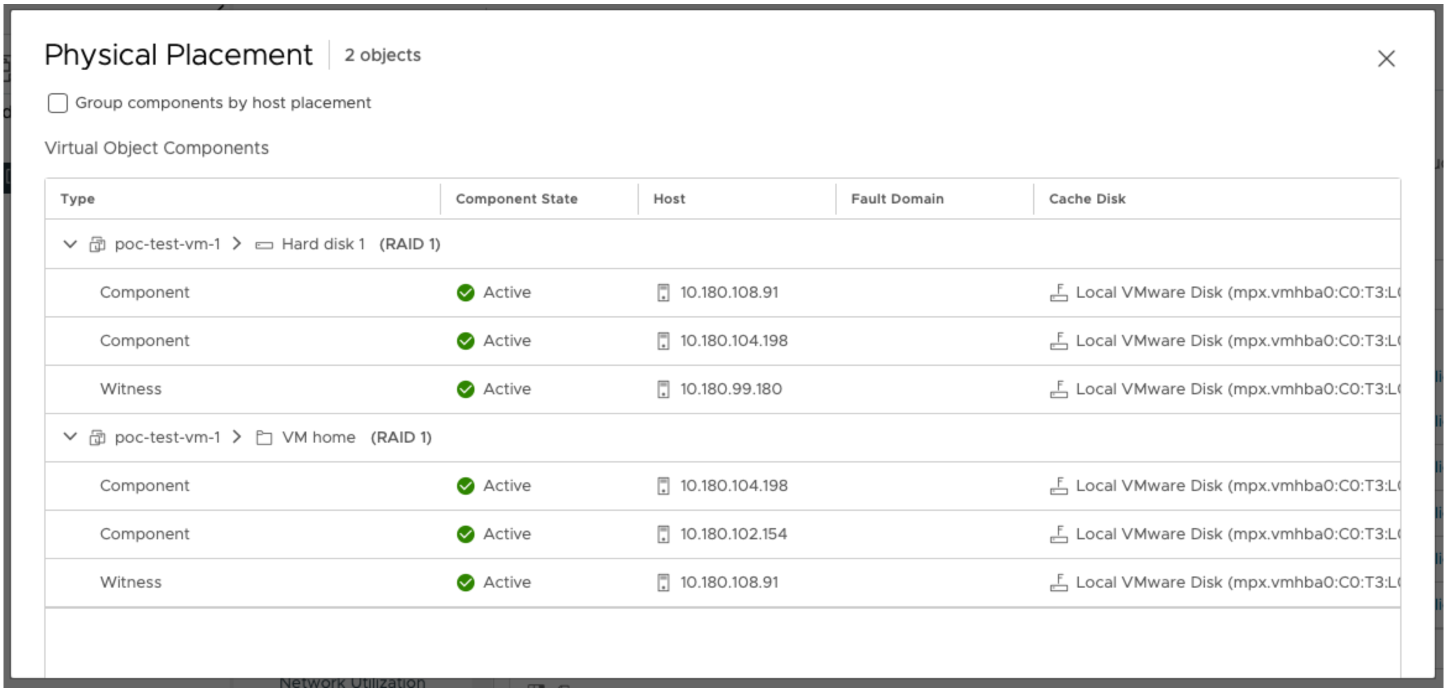
Browse all virtual objects and check their state in real time and view their placement across the physical infrastructure. Get information about each object state and common remediation. About vSAN Object Health

☐	Name	Type	Object State	Storage Policy
☑	poc-test-vm-1	VM		
☑	Hard disk 1	Disk	Healthy	vSAN Default Storage Policy
☑	VM home	Folder	Healthy	vSAN Default Storage Policy

This should display a physical placement of RAID 1 configuration with two components, each component representing a mirrored copy of the virtual disk. It should also be noted that the components are located on different hosts or *fault domains*. This implies that the policy setting to tolerate 1 failure is being adhered to, as each host is an implicit fault domain. Further, fault domains can be explicitly defined: for instance, hosts within a single rack. Thus, data is resilient to failure of the entire rack. For details on how to create fault domains, review Managing Fault Domains in vSAN Clusters - https://docs.vmware.com/en/VMware-vSphere/8.0/vsan-administration/GUID-8491C4B0-6F94-4023-8C7A-FD7B40D0368D.html

Physical Placement: vSAN OSA

In a vSAN OSA cluster, the 'witness' component is used to maintain a quorum on a per-object basis. For more information, refer to the VMware vSAN Design Guide on core.vmware.com



Physical Placement | 2 objects

□ Group components by host placement

Virtual Object Components

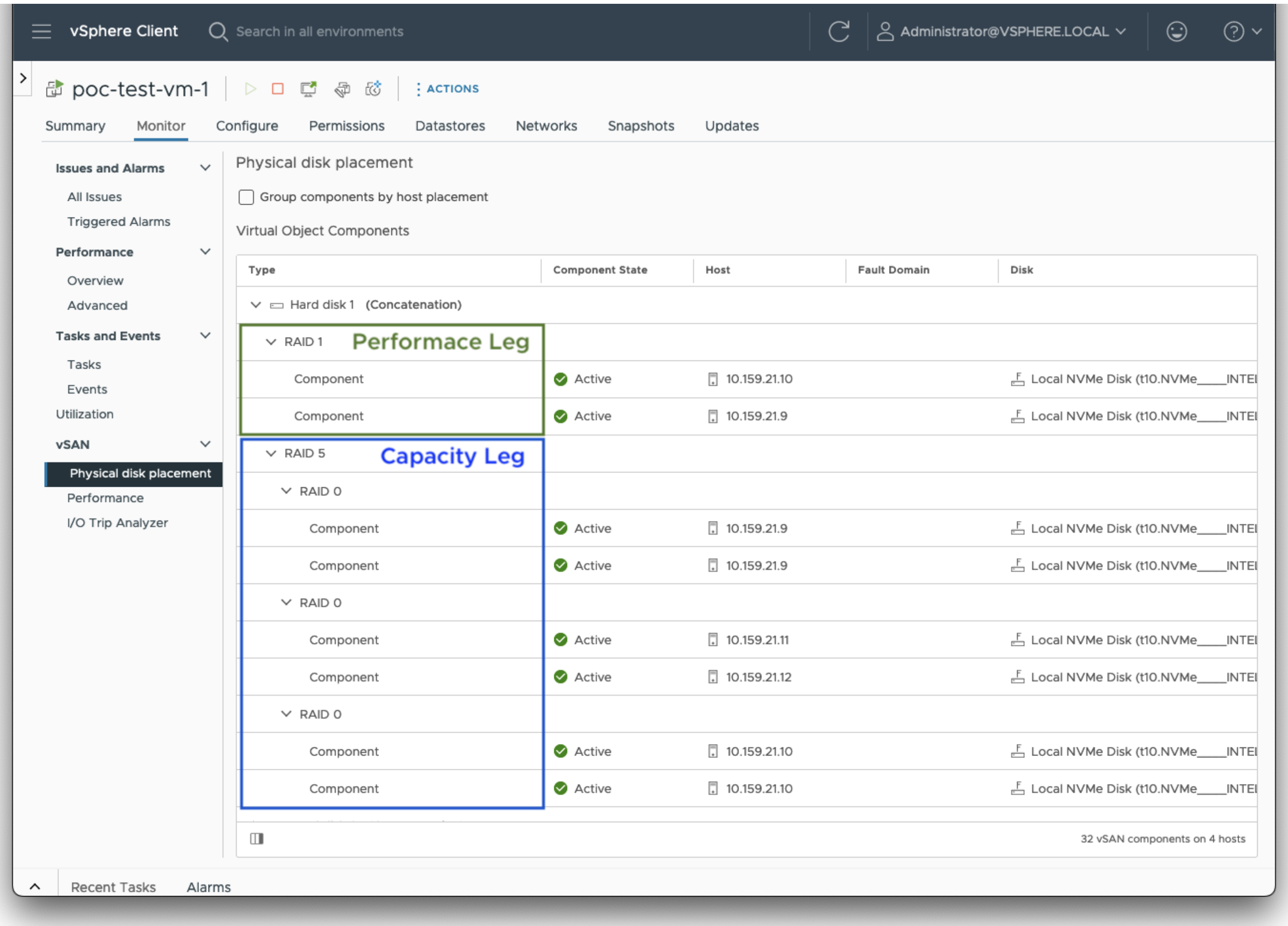
Type	Component State	Host	Fault Domain	Cache Disk
poc-test-vm-1 > Hard disk 1 (RAID 1)				
Component	Active	10.180.108.91		Local VMware Disk (mpx.vmhba0:C0:T3:L
Component	Active	10.180.104.198		Local VMware Disk (mpx.vmhba0:C0:T3:L
Witness	Active	10.180.99.180		Local VMware Disk (mpx.vmhba0:C0:T3:L
poc-test-vm-1 > VM home (RAID 1)				
Component	Active	10.180.104.198		Local VMware Disk (mpx.vmhba0:C0:T3:L
Component	Active	10.180.102.154		Local VMware Disk (mpx.vmhba0:C0:T3:L
Witness	Active	10.180.108.91		Local VMware Disk (mpx.vmhba0:C0:T3:L

Physical Placement: vSAN ESA

In vSAN ESA, physical placement is a little different. Data is written into two legs: writes are first ingested into a performance leg and then coalesced and written to a capacity leg. Whilst the distribution of data on the capacity leg reflects the storage policy setting (RAID 1 vs. RAID 5, etc.), the performance leg is *always* a RAID-1 mirror (and the FTT of the policy is followed). For this reason, RAID-5 performance in vSAN ESA is at least on-par with RAID-1 performance on vSAN OSA. **Thus, it is recommended, for most workloads, to define a RAID-5 policy for VMs on vSAN ESA.**

For more information on object placement in vSAN ESA, visit https://core.vmware.com/vsan-esa

Below we see how the vSAN default policy (RAID-1, FTT-1) distributes the objects (physical disk placement can also be seen per VM, by selecting the VM and navigating to **Monitor > vSAN > Physical disk placement**). Also note that no witness components are needed (as opposed to vSAN OSA) as there are enough data objects to maintain quorum:



Physical disk placement

☐ Group components by host placement

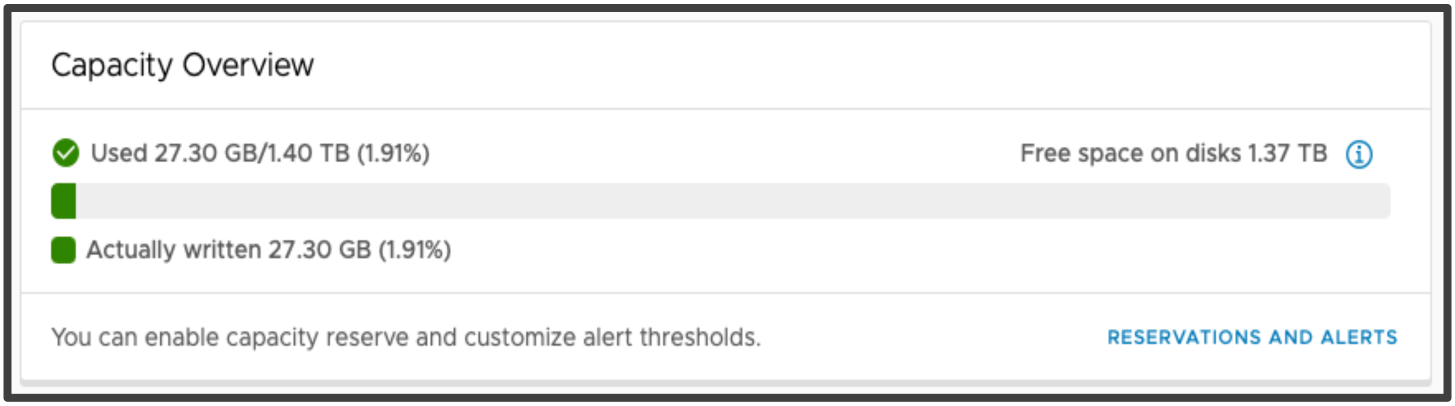
Virtual Object Components

Type	Component State	Host	Fault Domain	Disk
Hard disk 1 (Concatenation)				
RAID 1 Performace Leg				
Component	Active	10.159.21.10		Local NVMe Disk (t10.NVMe____INTEI
Component	Active	10.159.21.9		Local NVMe Disk (t10.NVMe____INTEI
RAID 5 Capacity Leg				
RAID 0				
Component	Active	10.159.21.9		Local NVMe Disk (t10.NVMe____INTEI
Component	Active	10.159.21.9		Local NVMe Disk (t10.NVMe____INTEI
RAID 0				
Component	Active	10.159.21.11		Local NVMe Disk (t10.NVMe____INTEI
Component	Active	10.159.21.12		Local NVMe Disk (t10.NVMe____INTEI
RAID 0				
Component	Active	10.159.21.10		Local NVMe Disk (t10.NVMe____INTEI
Component	Active	10.159.21.10		Local NVMe Disk (t10.NVMe____INTEI

32 vSAN components on 4 hosts

Physical Space Requirements

The “object space reservation” policy setting defines how much space is initially reserved on the vSAN datastore for a VM's objects. By default, it is set to "thin provisioning", implying that the VM’s storage objects are entirely “thin” and consume no unnecessary space. Note the free capacity in the vSAN datastore after deploying the VM, we see that the free capacity is very close to what it was before the VM was deployed, as displayed:



Capacity Overview

✓ Used 27.30 GB/1.40 TB (1.91%)

Free space on disks 1.37 TB ⓘ

🟩 Actually written 27.30 GB (1.91%)

You can enable capacity reserve and customize alert thresholds.

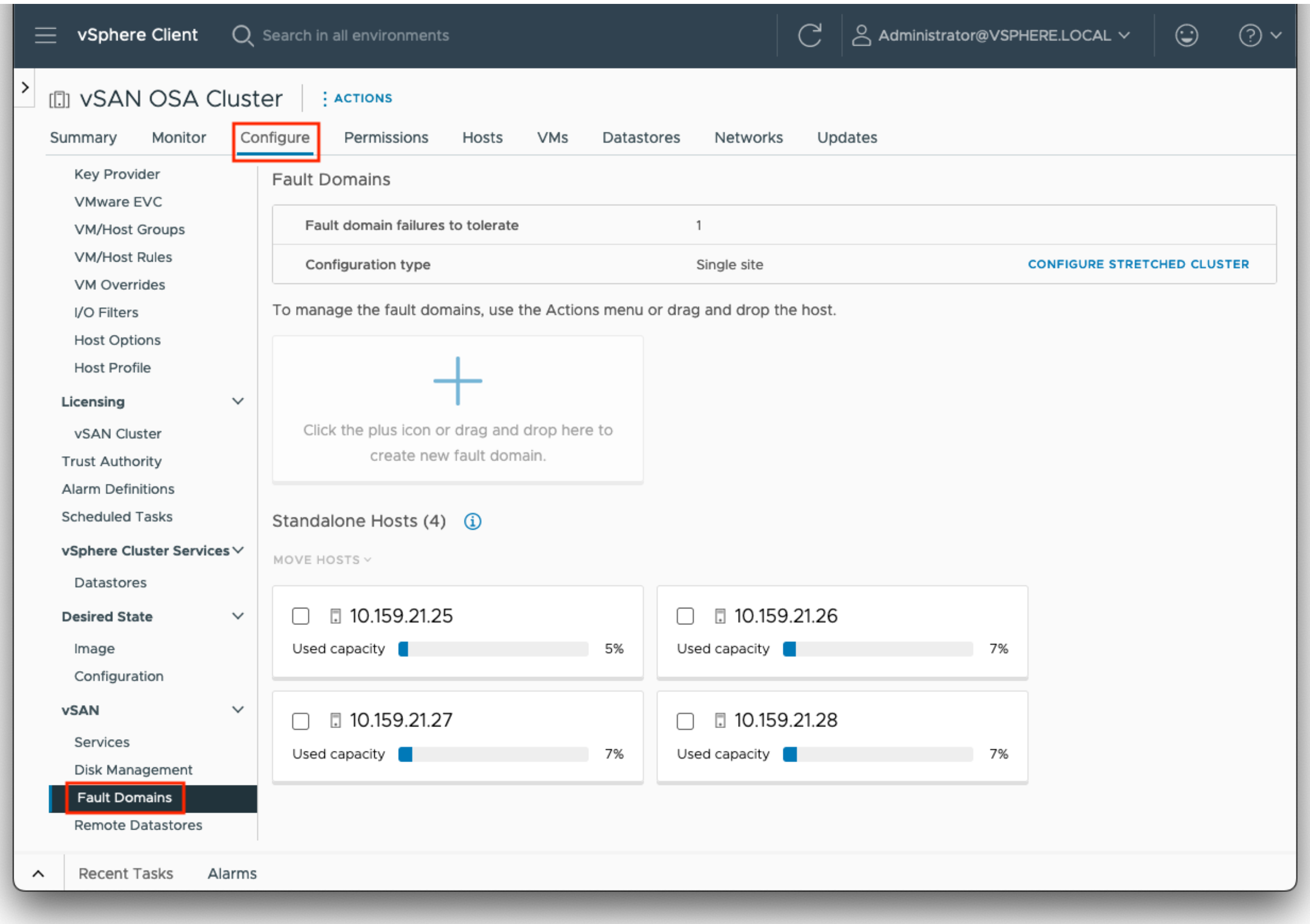
Because we have not installed anything in the VM (such as a guest OS) - it shows that only a tiny portion of the vSAN datastore has so far been used, verifying that the object space reservation setting of "Thin provisioning" is working correctly (observe that the "Virtual disks" and "VM home objects" consume less than 1GB in total, as highlighted in the "Used Capacity Breakdown" section).

Do not delete this VM as we will use it for other tests going forward.

Configuring Fault Domains

As mentioned above, a single host is a *fault domain*, i.e. data is separated such that the failure of a host does not lead to data loss. On failure, data can be rebuilt elsewhere. We can also group hosts so that data is further spread, and data is better protected.

Fault domains can be defined at cluster creation time (see the sections above). To define (or alter) fault domains, thereafter, navigate to **[vSAN Cluster] > Configure > vSAN > Fault Domains**.



Fault Domains

Fault domain failures to tolerate	1
Configuration type	Single site

To manage the fault domains, use the Actions menu or drag and drop the host.

Click the plus icon or drag and drop here to create new fault domain.

Standalone Hosts (4)

Host	Used capacity
10.159.21.25	5%
10.159.21.26	7%
10.159.21.27	7%
10.159.21.28	7%

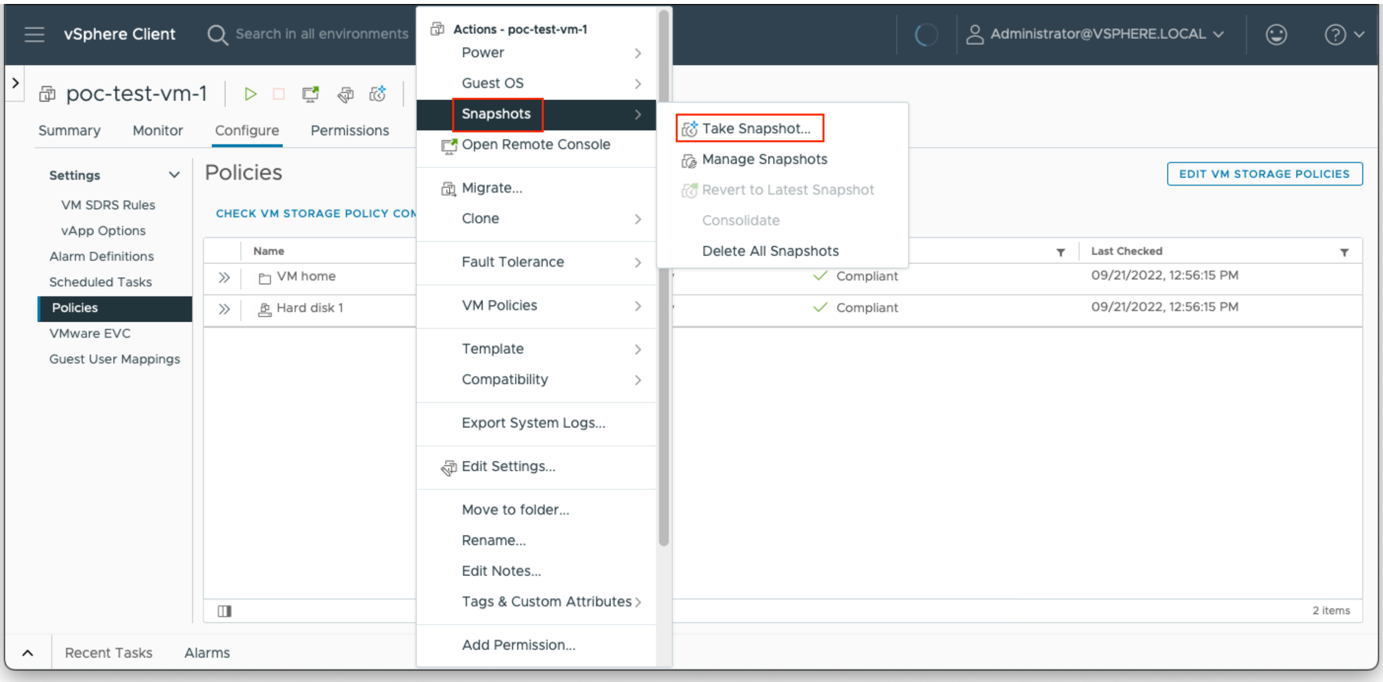
Fault domains can then be created by either dragging and dropping into the 'plus' icon, or by ticking the appropriate hosts and selecting **Move Hosts**.

Creating a Snapshot

Using the virtual machine created previously, take a snapshot of it. The snapshot can be taken when the VM is powered on or powered off. The objectives are to see that:

- No setup is needed to make vSAN handle snapshots
- The process for creating a VM snapshot is unchanged with vSAN
- A successful snapshot delta object is created
- The policy settings of the delta object are inherited directly from the base disk object

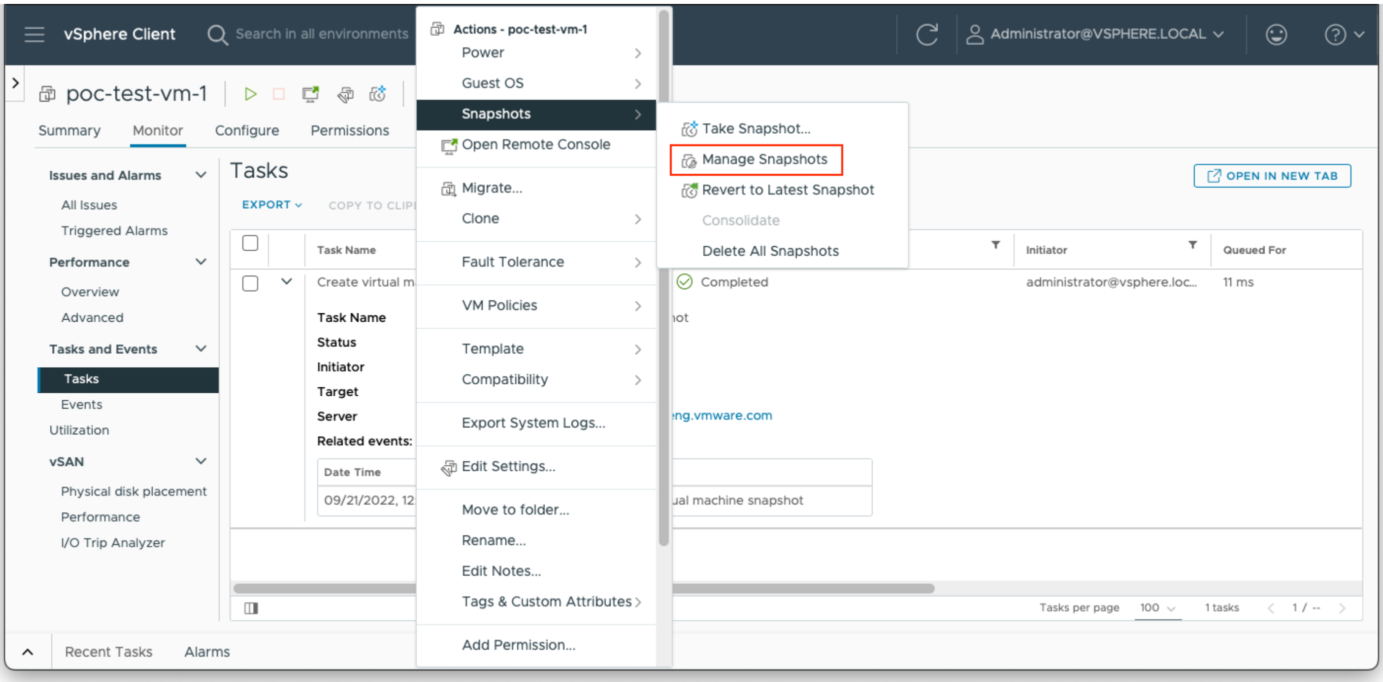
From the VM object in vCenter, click **Actions > Snapshots > Take Snapshot…**



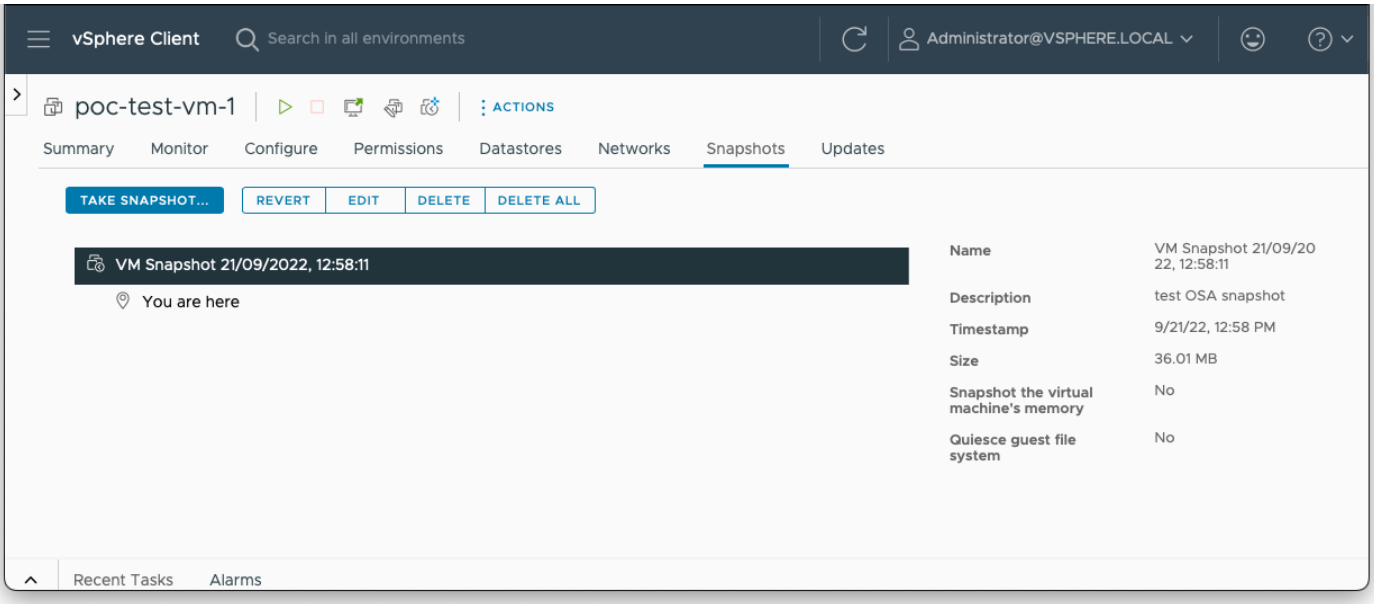
Take a Snapshot of the virtual machine created in the earlier step.

Provide a name for the snapshot and optional description.

Once the snapshot has been requested, monitor tasks and events to ensure that it has been successfully captured. Once the snapshot creation has completed, additional actions will become available in the snapshot drop-down window. For example, there is a new action to **Revert to Latest Snapshot** and another action to **Manage Snapshots**.

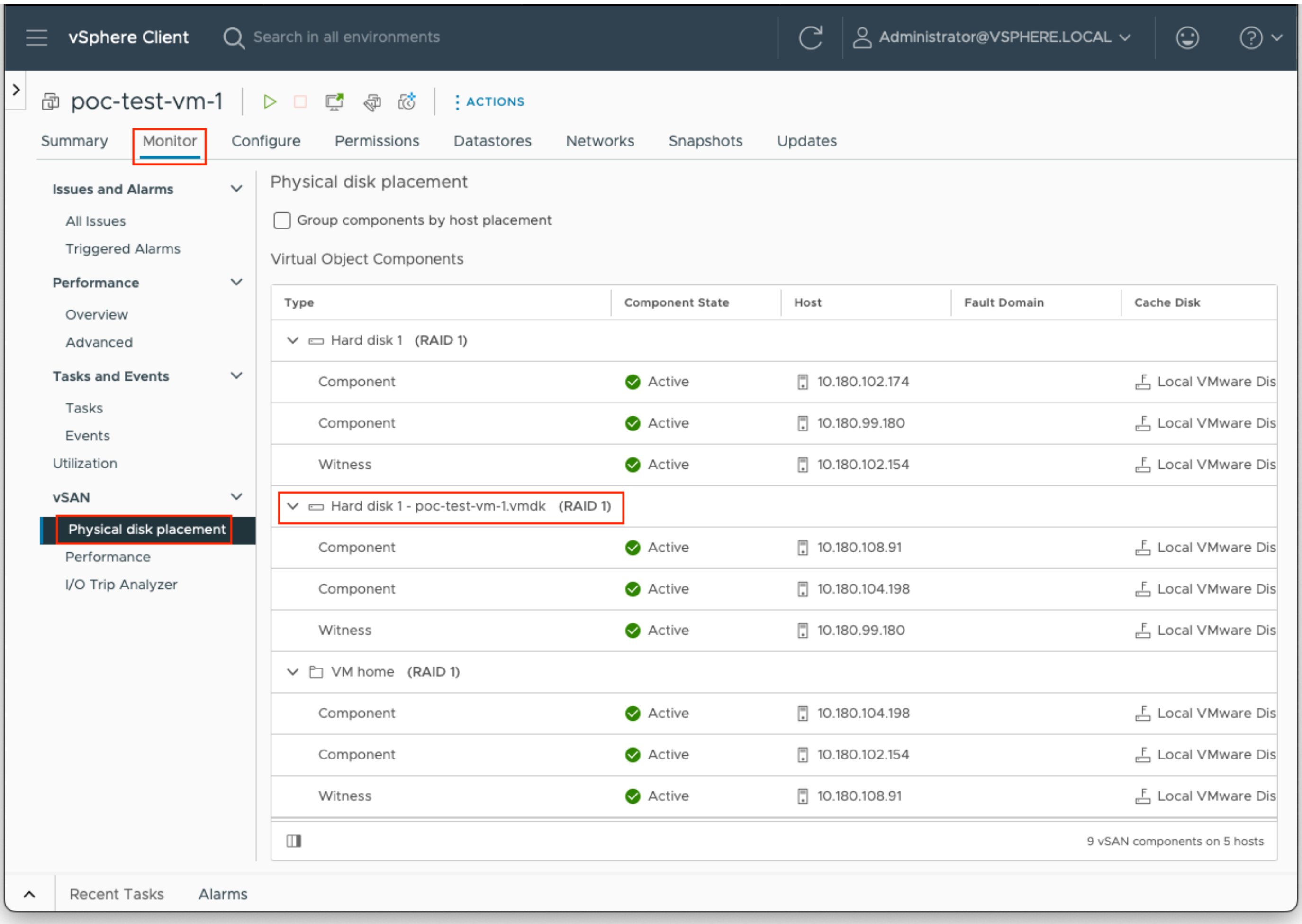


Choose the **Manage Snapshots** option. The following is displayed. It includes details regarding all snapshots in the chain, the ability to delete one or all of them, as well as the ability to revert to a particular snapshot.



Name	VM Snapshot 21/09/20 22, 12:58:11
Description	test OSA snapshot
Timestamp	9/21/22, 12:58 PM
Size	36.01 MB
Snapshot the virtual machine's memory	No
Quiesce guest file system	No

To see snapshot delta object information from the UI, navigate to **Monitor > vSAN > Physical disk placement**.



Physical disk placement

☐ Group components by host placement

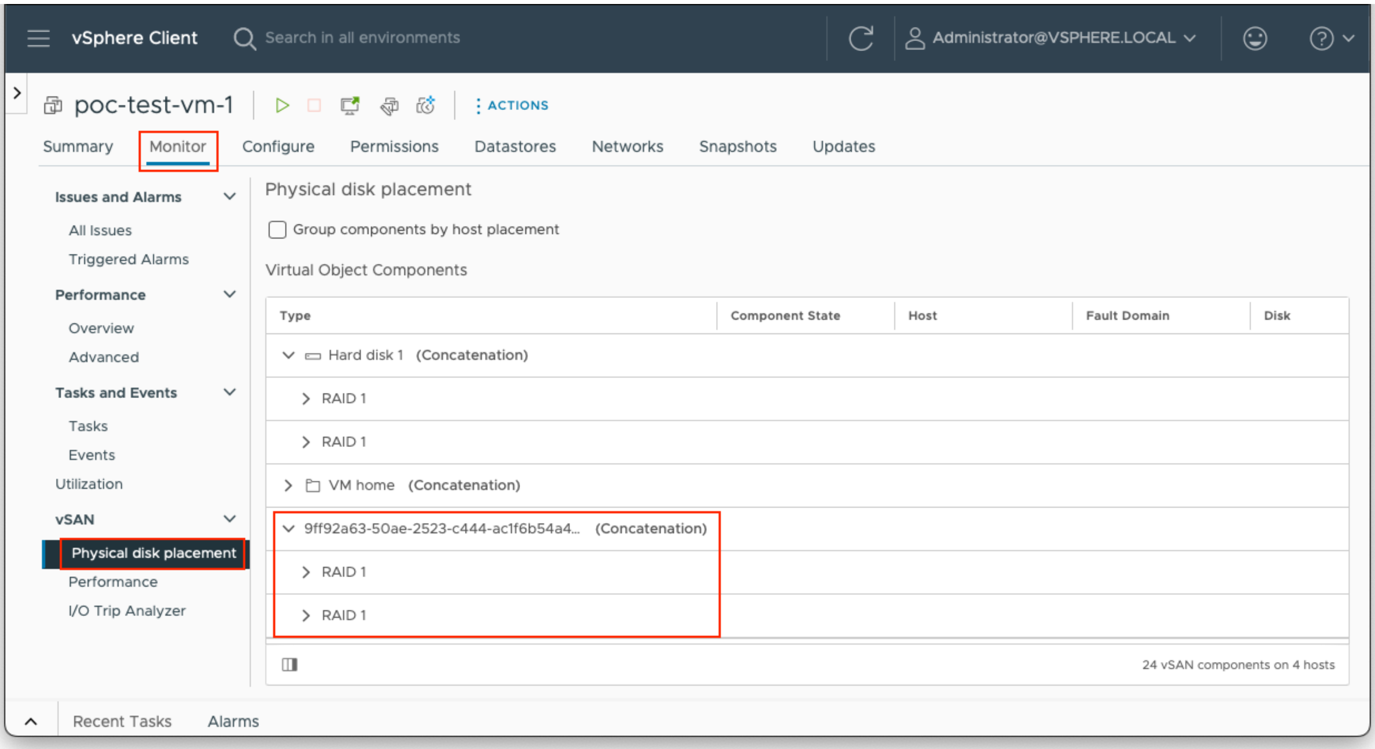
Virtual Object Components

Type	Component State	Host	Fault Domain	Cache Disk
Hard disk 1 (RAID 1)				
Component	✓ Active	10.180.102.174		Local VMware Dis
Component	✓ Active	10.180.99.180		Local VMware Dis
Witness	✓ Active	10.180.102.154		Local VMware Dis
Hard disk 1 - poc-test-vm-1.vmdk (RAID 1)				
Component	✓ Active	10.180.108.91		Local VMware Dis
Component	✓ Active	10.180.104.198		Local VMware Dis
Witness	✓ Active	10.180.99.180		Local VMware Dis
VM home (RAID 1)				
Component	✓ Active	10.180.104.198		Local VMware Dis
Component	✓ Active	10.180.102.154		Local VMware Dis
Witness	✓ Active	10.180.108.91		Local VMware Dis

9 vSAN components on 5 hosts

There are now three objects that are associated with that virtual machine. First is the "VM Home" namespace. "Hard disk 1" is the base virtual disk, and "Hard disk 1 - poc-test-vm1.vmdk" is the snapshot delta. Notice the snapshot delta inherits its policy settings from the base disk that needs to adhere to the vSAN Default Storage Policy.

For vSAN ESA, another object is created (with the same layout as Hard disk 1):



Physical disk placement

☐ Group components by host placement

Virtual Object Components

Type	Component State	Host	Fault Domain	Disk
Hard disk 1 (Concatenation)				
RAID 1				
RAID 1				
VM home (Concatenation)				
9ff92a63-50ae-2523-c444-ac1f6b54a4... (Concatenation)				
RAID 1				
RAID 1				

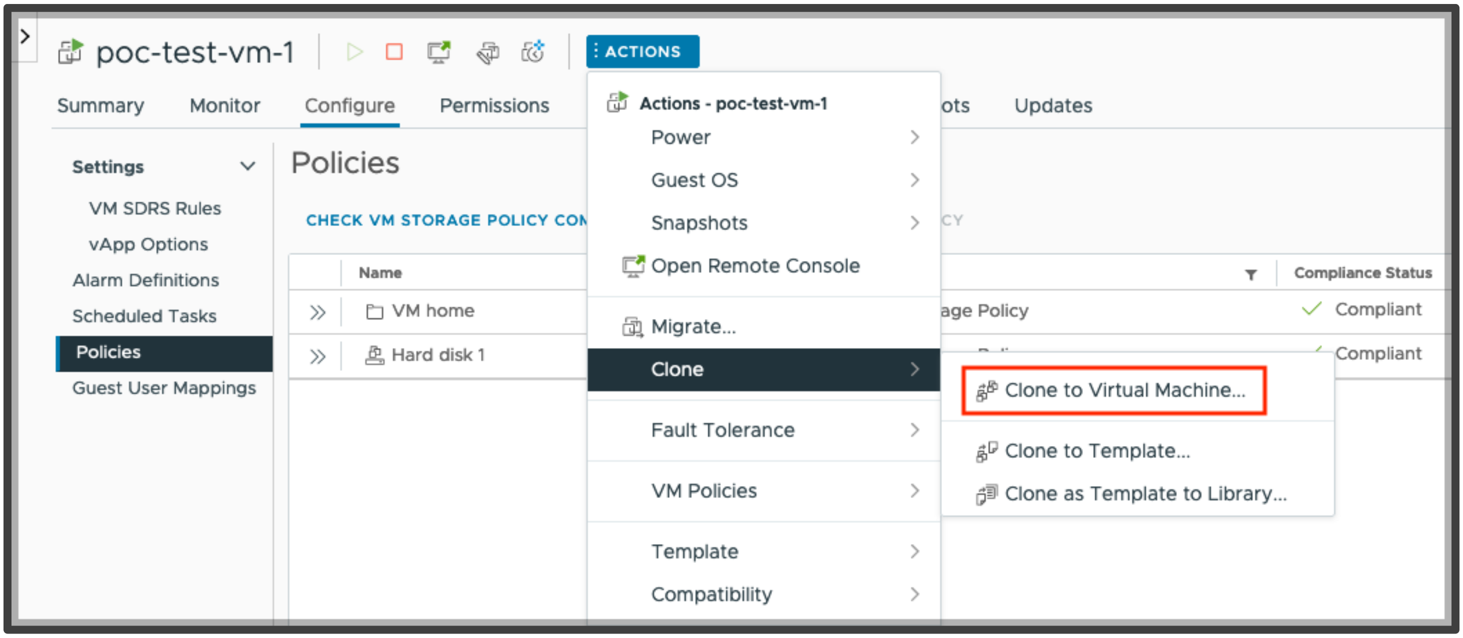
24 vSAN components on 4 hosts

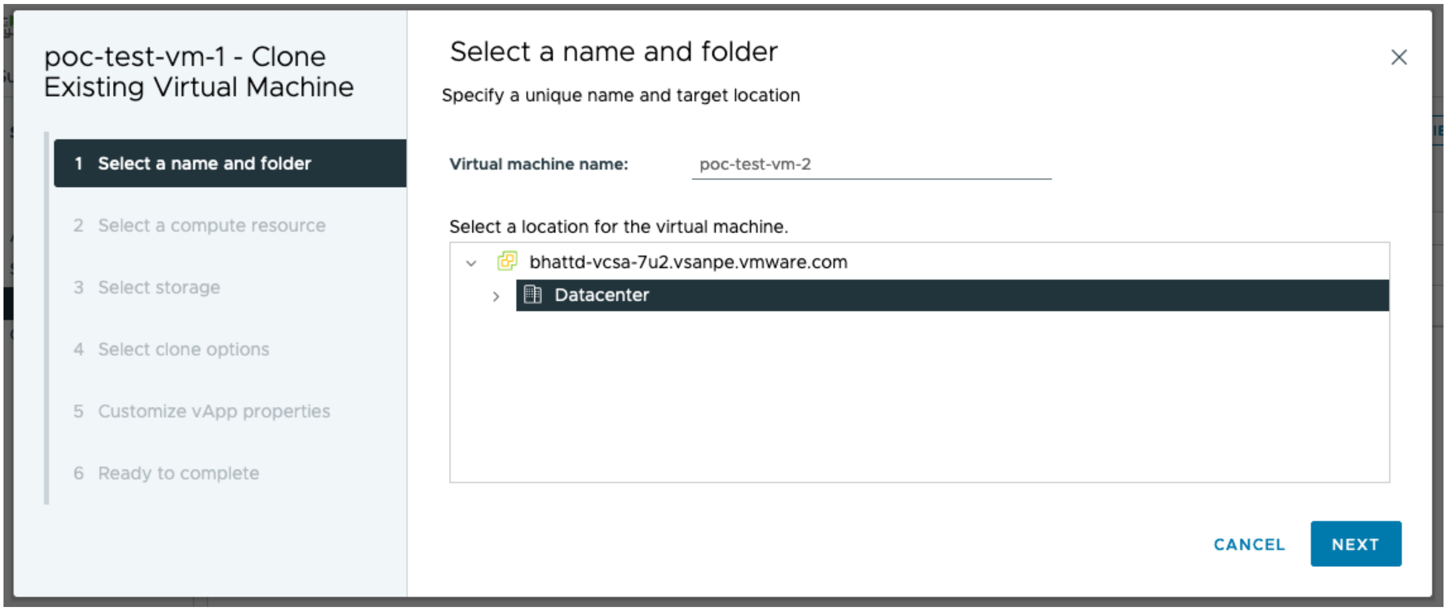
The snapshot can now be deleted from the VM. Monitor the VM's tasks and ensure that it deletes successfully.

Clone a Virtual Machine

We will continue to use the same VM as before. This time make sure the VM is powered on first.

There are several different cloning operations available with vSAN. Here we will "Clone to Virtual Machine". The cloning operation is a straightforward click-through operation. This next screen is the only one that requires human interaction. Simply provide the name for the newly cloned VM, and a folder if desired.




poc-test-vm-1 - Clone Existing Virtual Machine

1 Select a name and folder
2 Select a compute resource
3 Select storage
4 Select clone options
5 Customize vApp properties
6 Ready to complete

Select a name and folder

Specify a unique name and target location

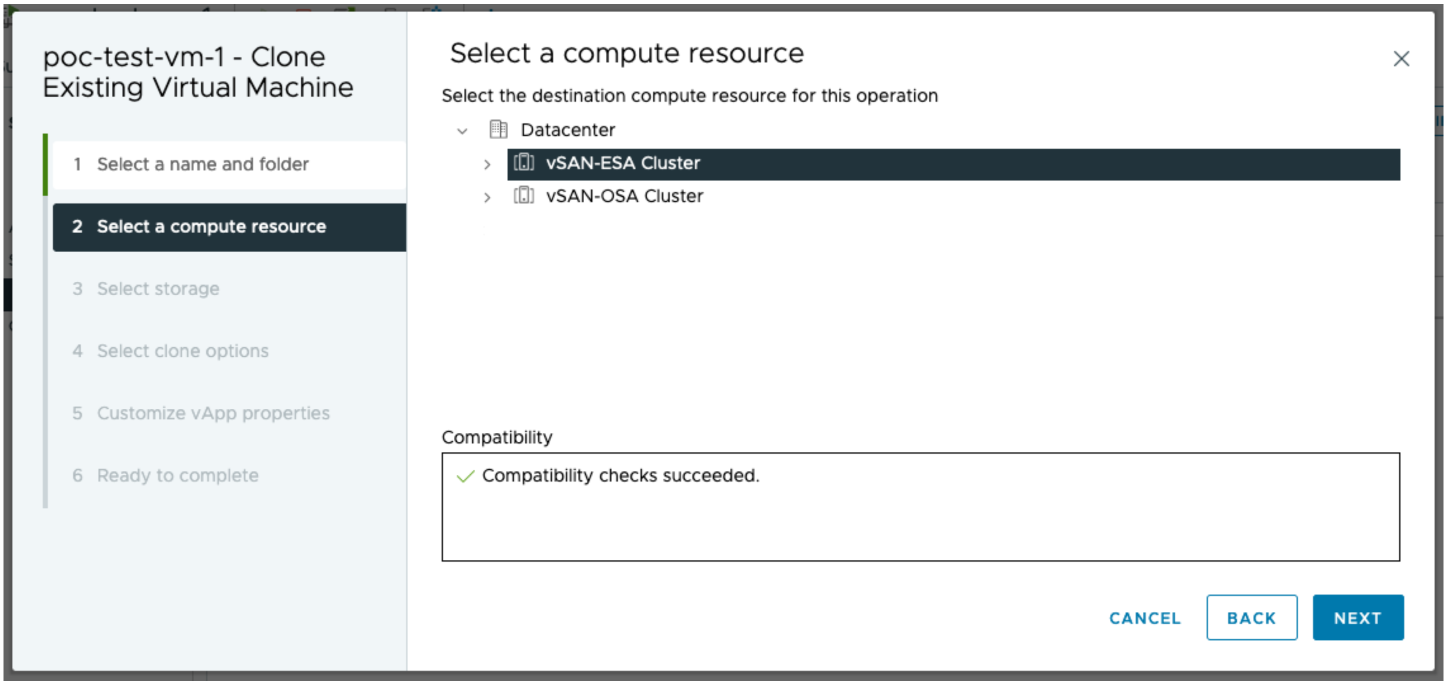
Virtual machine name: poc-test-vm-2

Select a location for the virtual machine.

bhattd-vcsa-7u2.vsanpe.vmware.com

Datacenter

We are going to clone the VM in the vSAN cluster, so this must be selected as the compute resource.



poc-test-vm-1 - Clone Existing Virtual Machine

1 Select a name and folder
2 Select a compute resource
3 Select storage
4 Select clone options
5 Customize vApp properties
6 Ready to complete

Select a compute resource

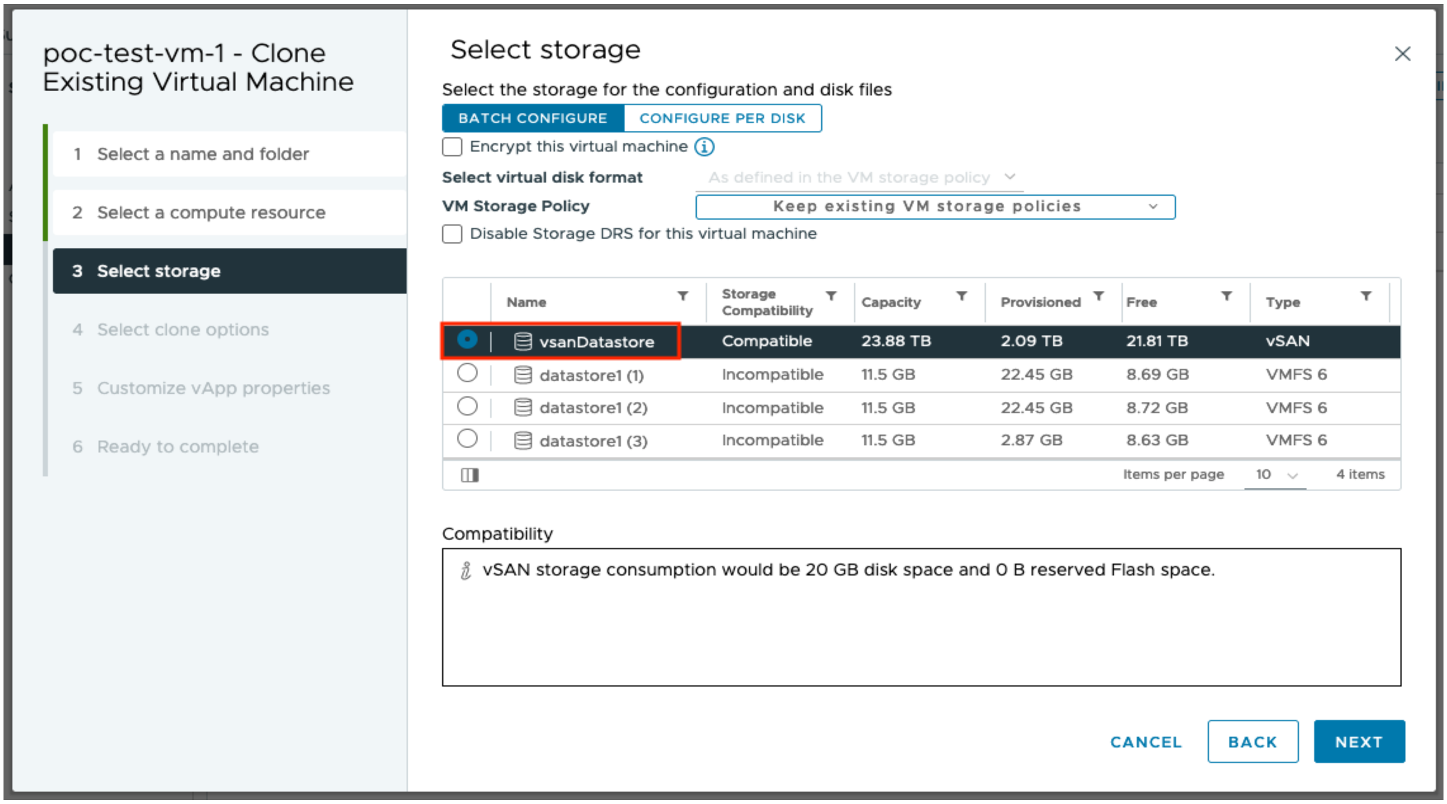
Select the destination compute resource for this operation

- Datacenter
  - vSAN-ESA Cluster
  - vSAN-OSA Cluster

Compatibility

✓ Compatibility checks succeeded.

On the “Select Storage” screen select the source datastore for the VM, “vsanDatastore”. This will all be pre-selected for you if the VM being cloned also resides on the vsanDatastore.


poc-test-vm-1 - Clone Existing Virtual Machine
1 Select a name and folder
2 Select a compute resource
3 Select storage
4 Select clone options
5 Customize vApp properties
6 Ready to complete
BATCH CONFIGURE
CONFIGURE PER DISK
CANCEL
BACK
NEXT

Select storage

Select the storage for the configuration and disk files

□ Encrypt this virtual machine

Select virtual disk format As defined in the VM storage policy

VM Storage Policy Keep existing VM storage policies

□ Disable Storage DRS for this virtual machine

	Name	Storage Compatibility	Capacity	Provisioned	Free	Type
●	vsanDatastore	Compatible	23.88 TB	2.09 TB	21.81 TB	vSAN
○	datastore1 (1)	Incompatible	11.5 GB	22.45 GB	8.69 GB	VMFS 6
○	datastore1 (2)	Incompatible	11.5 GB	22.45 GB	8.72 GB	VMFS 6
○	datastore1 (3)	Incompatible	11.5 GB	2.87 GB	8.63 GB	VMFS 6

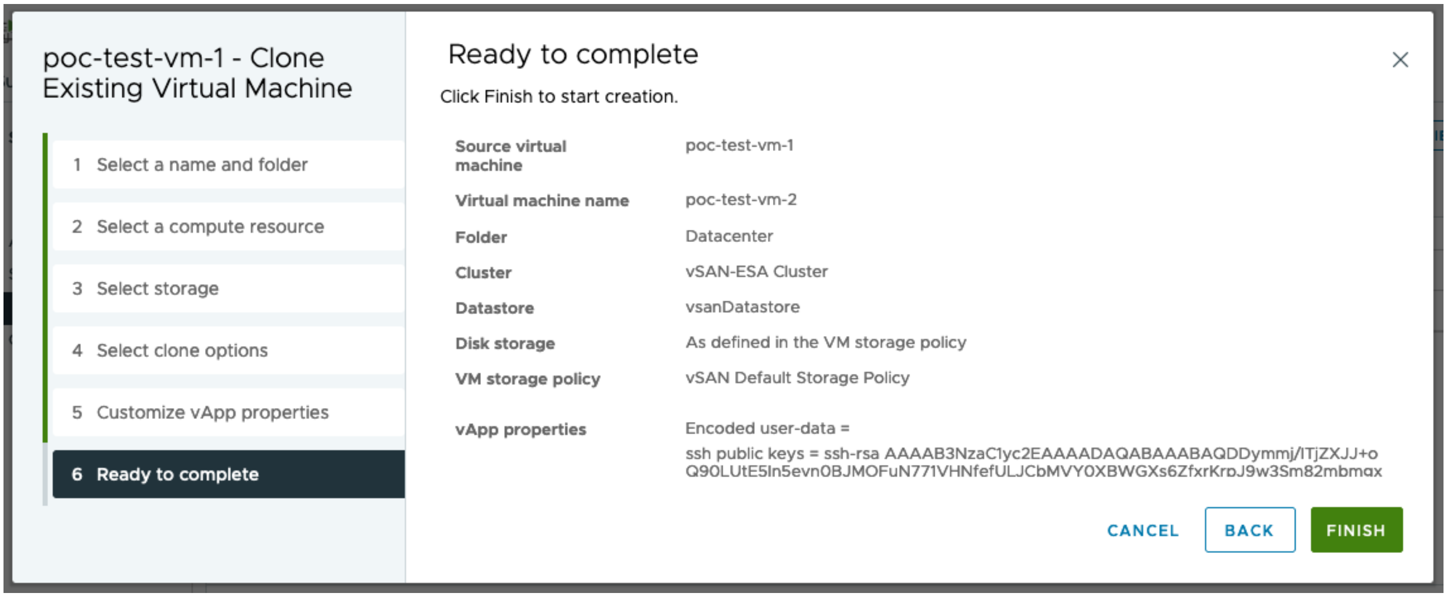
Items per page 10 4 items

Compatibility

vSAN storage consumption would be 20 GB disk space and 0 B reserved Flash space.

Select from the available options (leave unchecked - default)

This will take you to the "Ready to complete" screen. If everything is as expected, click **FINISH** to commence the clone operation. Monitor the VM tasks for the status of the clone operation.


poc-test-vm-1 - Clone Existing Virtual Machine
1 Select a name and folder
2 Select a compute resource
3 Select storage
4 Select clone options
5 Customize vApp properties
6 Ready to complete
CANCEL
BACK
FINISH

Ready to complete

Click Finish to start creation.

Source virtual machine	poc-test-vm-1
Virtual machine name	poc-test-vm-2
Folder	Datacenter
Cluster	vSAN-ESA Cluster
Datastore	vsanDatastore
Disk storage	As defined in the VM storage policy
VM storage policy	vSAN Default Storage Policy
vApp properties	Encoded user-data = ssh public keys = ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAABAQDDymmj/ITjZXJJ+o Q90LUtE5In5evn0BJMOFuN771VHNfefULJCbMVY0XBWGXs6ZfxrKrpJ9w3Sm82mbmax

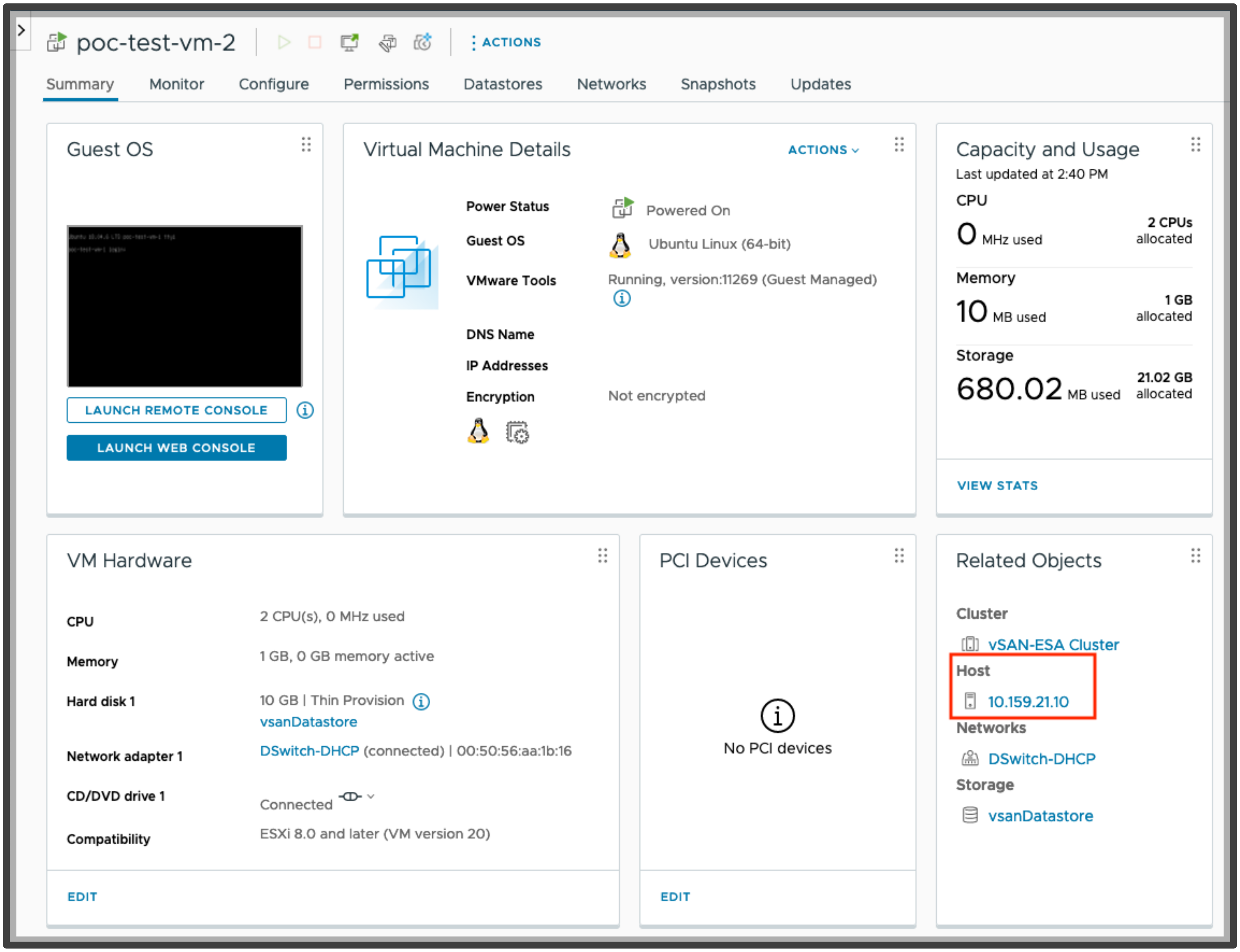
This completes the cloning section of this guide. Do not delete the newly cloned VM, we will be using it in subsequent tests.

vMotion a Virtual Machine Between Hosts

The first step is to power-on the newly cloned virtual machine. We will migrate this VM from one vSAN host to another vSAN host using vMotion.

Note: Take a moment to revisit the network configuration and ensure that the vMotion network is distinct from the vSAN network. If these features share the same network, performance will not be optimal.

First, determine which ESXi host the VM currently resides on. Selecting the **Summary** tab of the VM shows this, in the 'related objects window'.



Guest OS

Virtual Machine Details

Power Status	Powered On
Guest OS	Ubuntu Linux (64-bit)
VMware Tools	Running, version:11269 (Guest Managed)
DNS Name	
IP Addresses	
Encryption	Not encrypted

Capacity and Usage

Last updated at 2:40 PM

CPU: 0 MHz used, 2 CPUs allocated

Memory: 10 MB used, 1 GB allocated

Storage: 680.02 MB used, 21.02 GB allocated

VM Hardware

CPU	2 CPU(s), 0 MHz used
Memory	1 GB, 0 GB memory active
Hard disk 1	10 GB | Thin Provision vsanDatastore
Network adapter 1	DSwitch-DHCP (connected) | 00:50:56:aa:1b:16
CD/DVD drive 1	Connected
Compatibility	ESXi 8.0 and later (VM version 20)

PCI Devices

No PCI devices

Related Objects

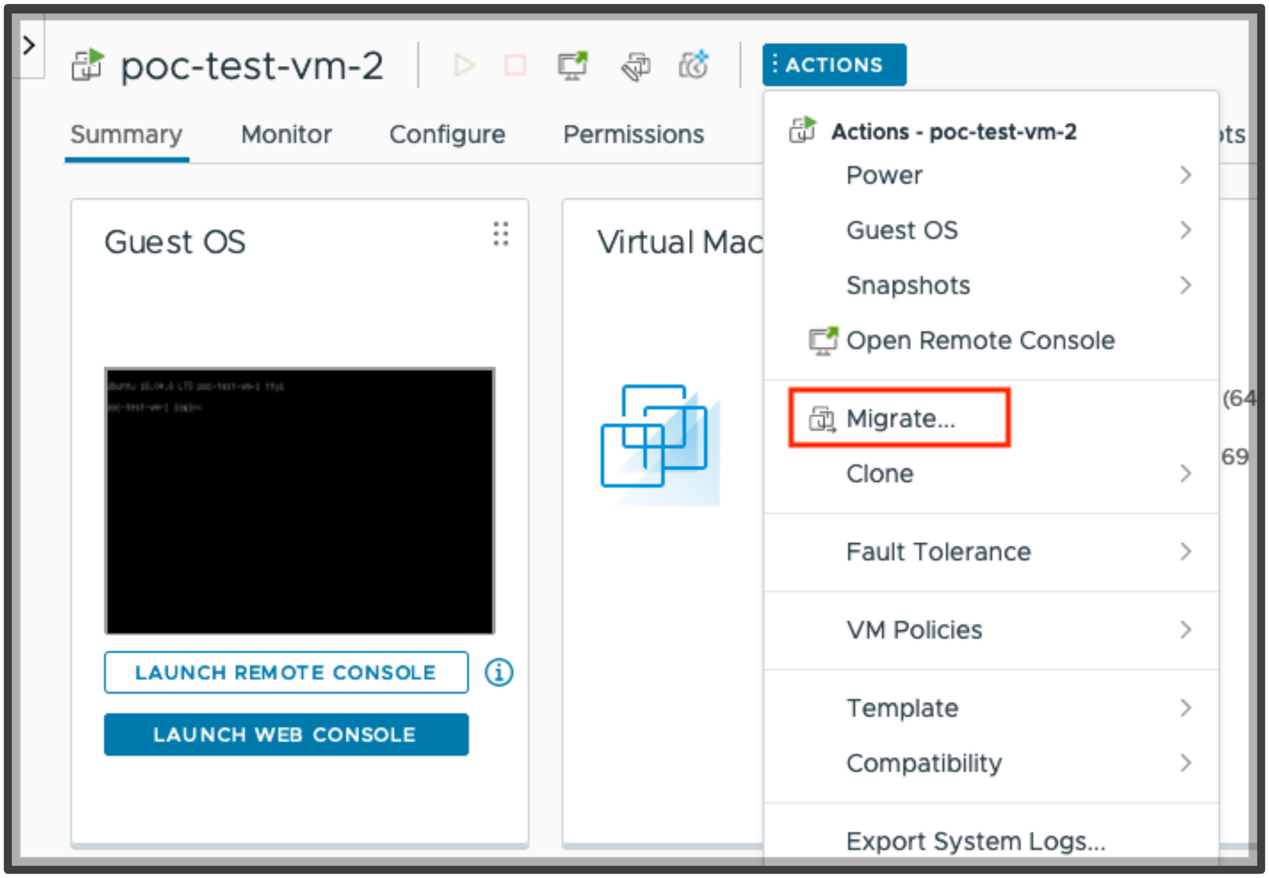
Cluster: vSAN-ESA Cluster

Host: 10.159.21.10

Networks: DSwitch-DHCP

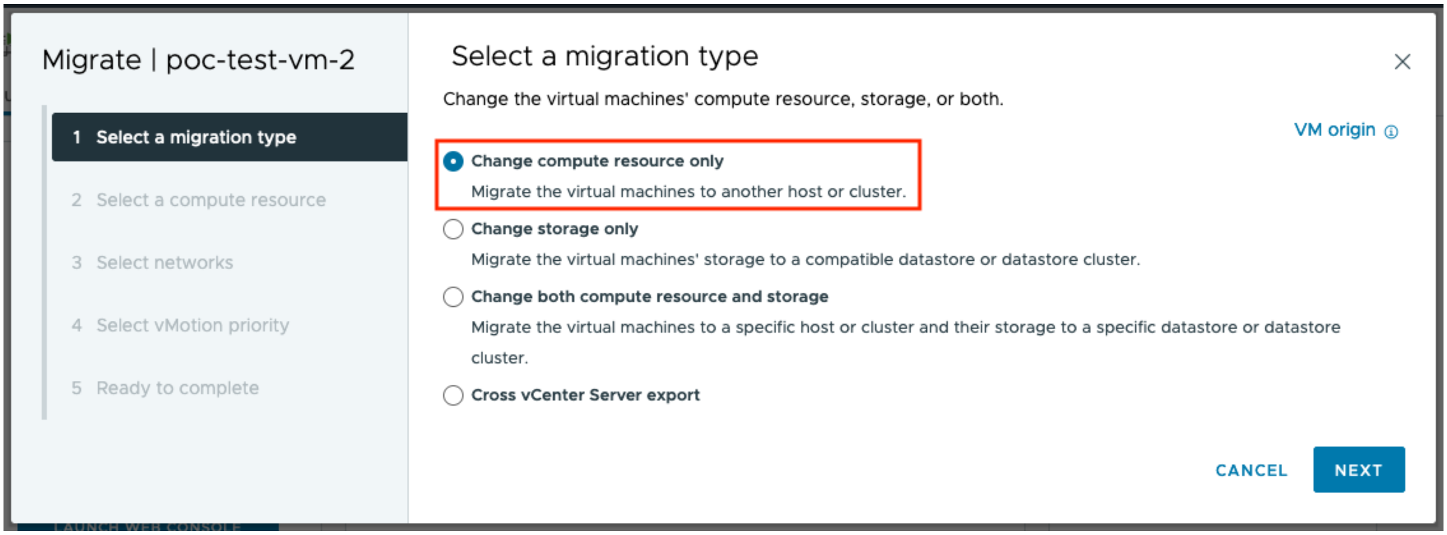
Storage: vsanDatastore

Next, select 'migrate', either from the VM 'actions' menu or by right-clicking on the VM:



"Migrate" allows you to migrate to a different compute resource (host), a different datastore or both at the same time. In this initial test, we are simply migrating the VM to another host in the cluster, so this initial screen should be left at the default of "Change compute resource only".

Select **Change compute resource only**:


Migrate | poc-test-vm-2

1 Select a migration type
2 Select a compute resource
3 Select networks
4 Select vMotion priority
5 Ready to complete

Select a migration type

Change the virtual machines' compute resource, storage, or both.

VM origin

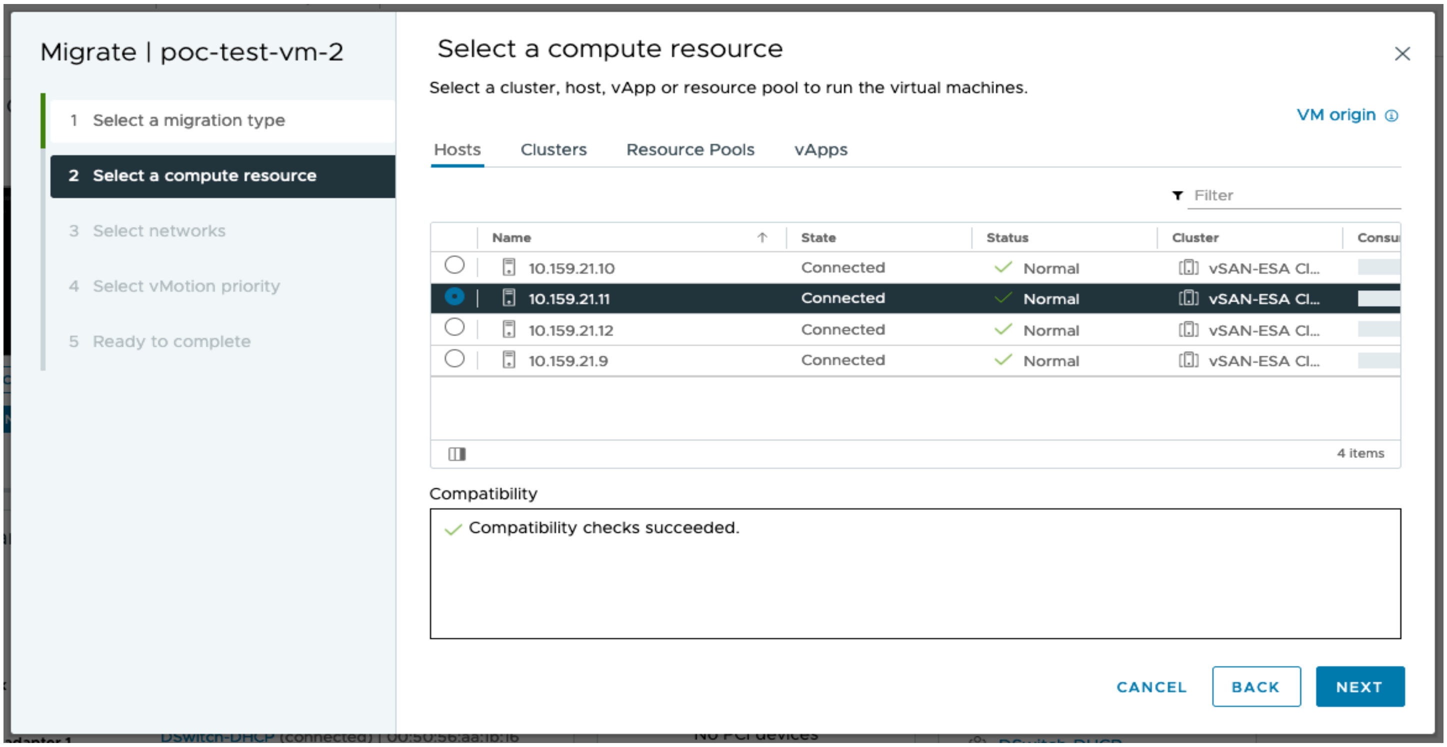
● **Change compute resource only**
Migrate the virtual machines to another host or cluster.

○ **Change storage only**
Migrate the virtual machines' storage to a compatible datastore or datastore cluster.

○ **Change both compute resource and storage**
Migrate the virtual machines to a specific host or cluster and their storage to a specific datastore or datastore cluster.

○ **Cross vCenter Server export**

Then select a destination host:



Migrate | poc-test-vm-2

1 Select a migration type

2 Select a compute resource

3 Select networks

4 Select vMotion priority

5 Ready to complete

Select a compute resource

Select a cluster, host, vApp or resource pool to run the virtual machines.

VM origin

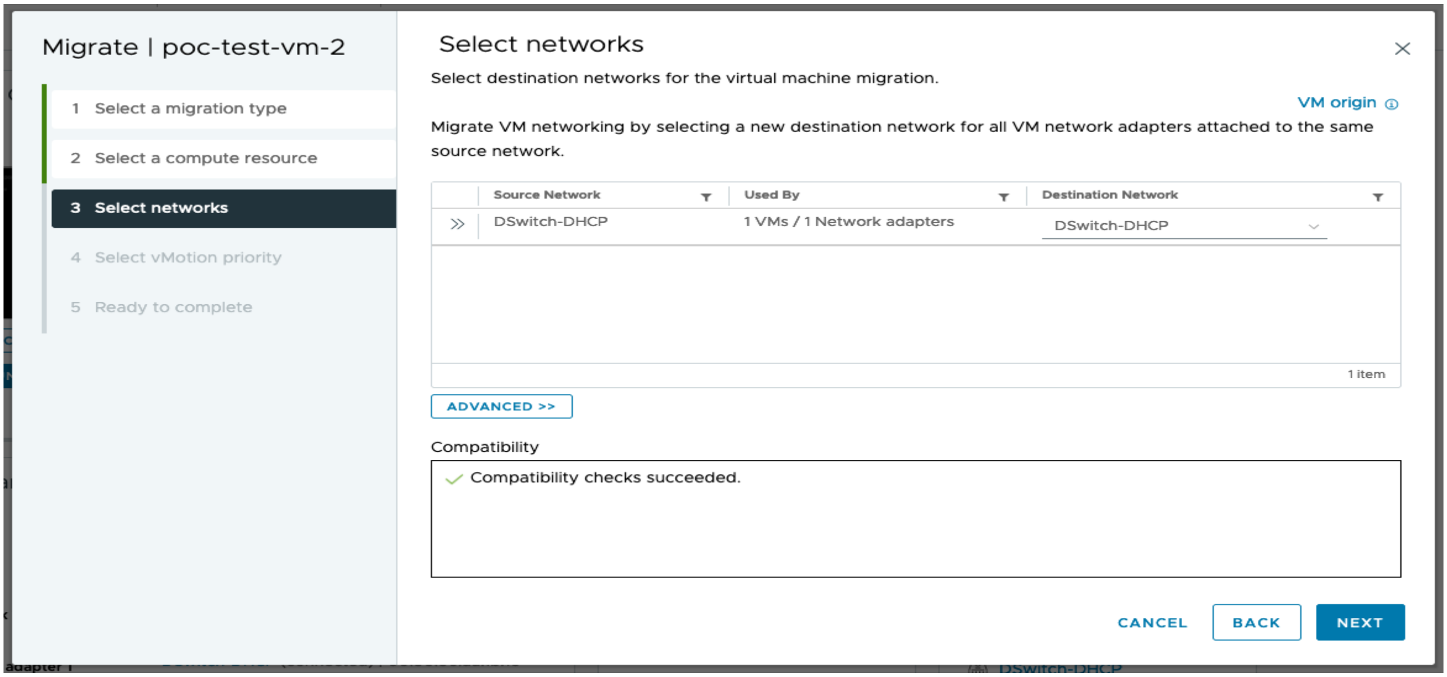
	Name	State	Status	Cluster	Consu
○	10.159.21.10	Connected	✓ Normal	vSAN-ESA Cl...	
●	10.159.21.11	Connected	✓ Normal	vSAN-ESA Cl...	
○	10.159.21.12	Connected	✓ Normal	vSAN-ESA Cl...	
○	10.159.21.9	Connected	✓ Normal	vSAN-ESA Cl...	

4 items

Compatibility

✓ Compatibility checks succeeded.

Select a destination network and click **Next**.



Migrate | poc-test-vm-2

1 Select a migration type

2 Select a compute resource

3 Select networks

4 Select vMotion priority

5 Ready to complete

Select networks

Select destination networks for the virtual machine migration.

VM origin

Migrate VM networking by selecting a new destination network for all VM network adapters attached to the same source network.

	Source Network	Used By	Destination Network
»	DSwitch-DHCP	1 VMs / 1 Network adapters	DSwitch-DHCP

1 item

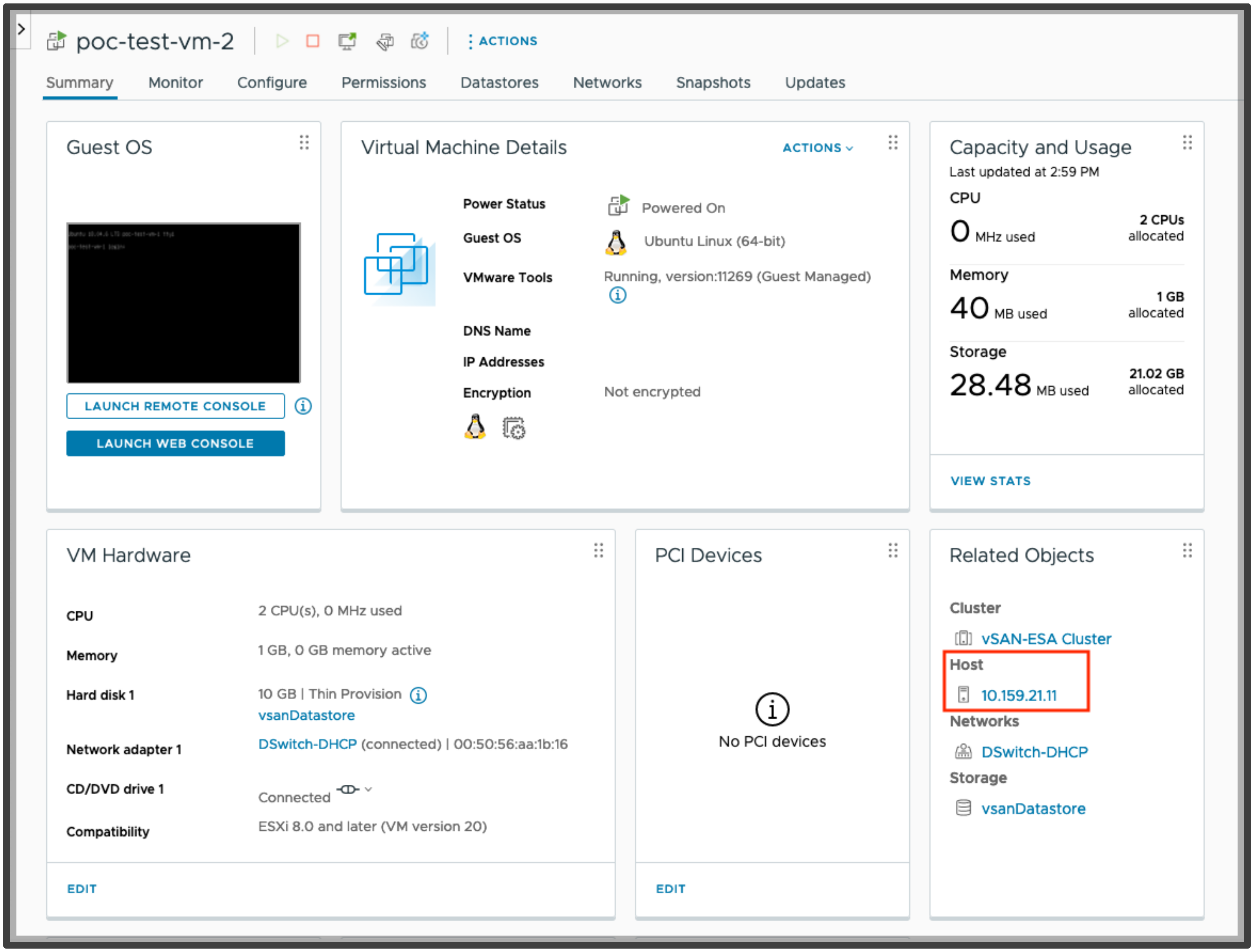
Compatibility

✓ Compatibility checks succeeded.

Leave the default (high) on the vMotion Priority window, click **Next**.

At the “Ready to complete” window, click on FINISH to initiate the migration. If the migration is successful, the summary tab of the virtual machine should show that the VM now resides on a different host.

Verify that the VM has been migrated to a new host:



Power Status: Powered On
Guest OS: Ubuntu Linux (64-bit)
VMware Tools: Running, version:11269 (Guest Managed)
DNS Name
IP Addresses
Encryption: Not encrypted

Last updated at 2:59 PM
CPU: 0 MHz used, 2 CPUs allocated
Memory: 40 MB used, 1 GB allocated
Storage: 28.48 MB used, 21.02 GB allocated

CPU: 2 CPU(s), 0 MHz used
Memory: 1 GB, 0 GB memory active
Hard disk 1: 10 GB | Thin Provision, vsanDatastore
Network adapter 1: DSwitch-DHCP (connected) | 00:50:56:aa:1b:16
CD/DVD drive 1: Connected
Compatibility: ESXi 8.0 and later (VM version 20)

No PCI devices

Cluster: vSAN-ESA Cluster
Host: 10.159.21.11
Networks: DSwitch-DHCP
Storage: vsanDatastore

This completes the “VM migration using vMotion” section of this guide. As you can see, vMotion works just great with vSAN. Do not delete the migrated VM: we will be using it in subsequent tests.

VMware vSphere® Storage vMotion® (Storage vMotion) VM Between Datastores

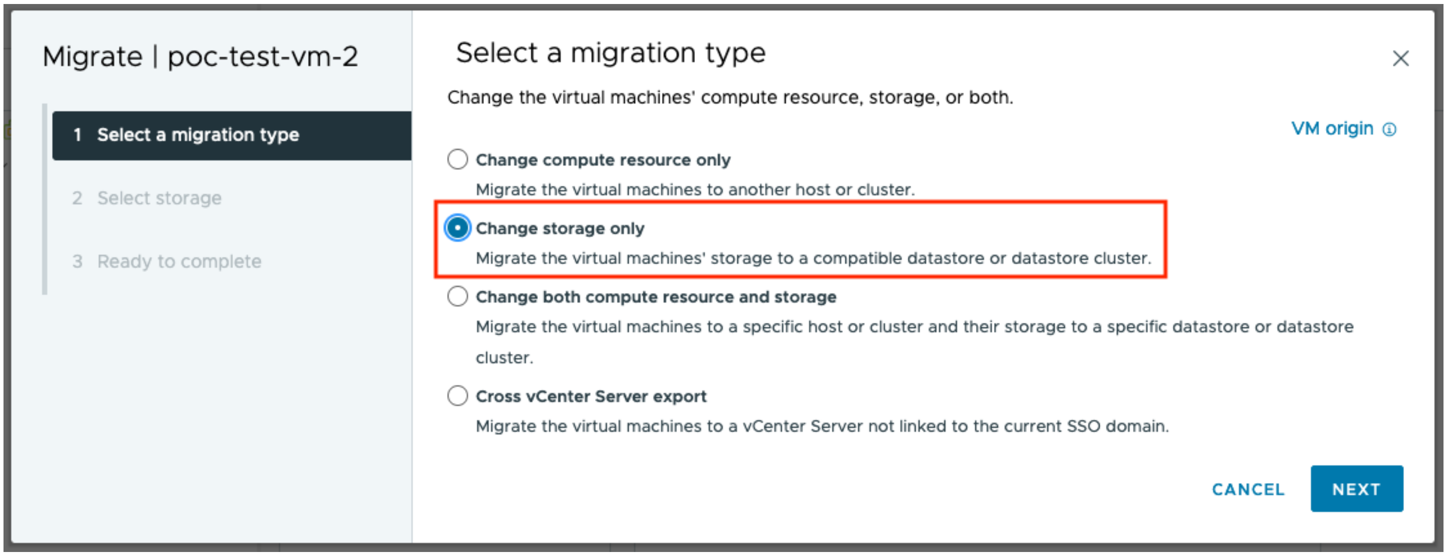
This test will only be possible if there is space on a VMFS datastore (such as the boot volume) on the host housing the VM or you have another datastore type available, such as an NFS share. The objective of this test is to successfully migrate a VM from another datastore type into vSAN and vice versa.

Mount an NFS Datastore to the Hosts (optional)

The steps to mount an NFS datastore to multiple ESXi hosts are described in the vSphere Administrators Guide. See the Create NFS Datastore in the vSphere Client topic for detailed steps.

Storage vMotion a VM from vSAN to another Datastore Type

Currently, the VM resides on the vSAN datastore. As we did before, launch the migrate wizard, however, on this occasion move the VM from the vSAN datastore to another datastore type by selecting **Change storage only**.


Migrate | poc-test-vm-2

1 Select a migration type
2 Select storage
3 Ready to complete

Select a migration type

Change the virtual machines' compute resource, storage, or both.

VM origin ⓘ

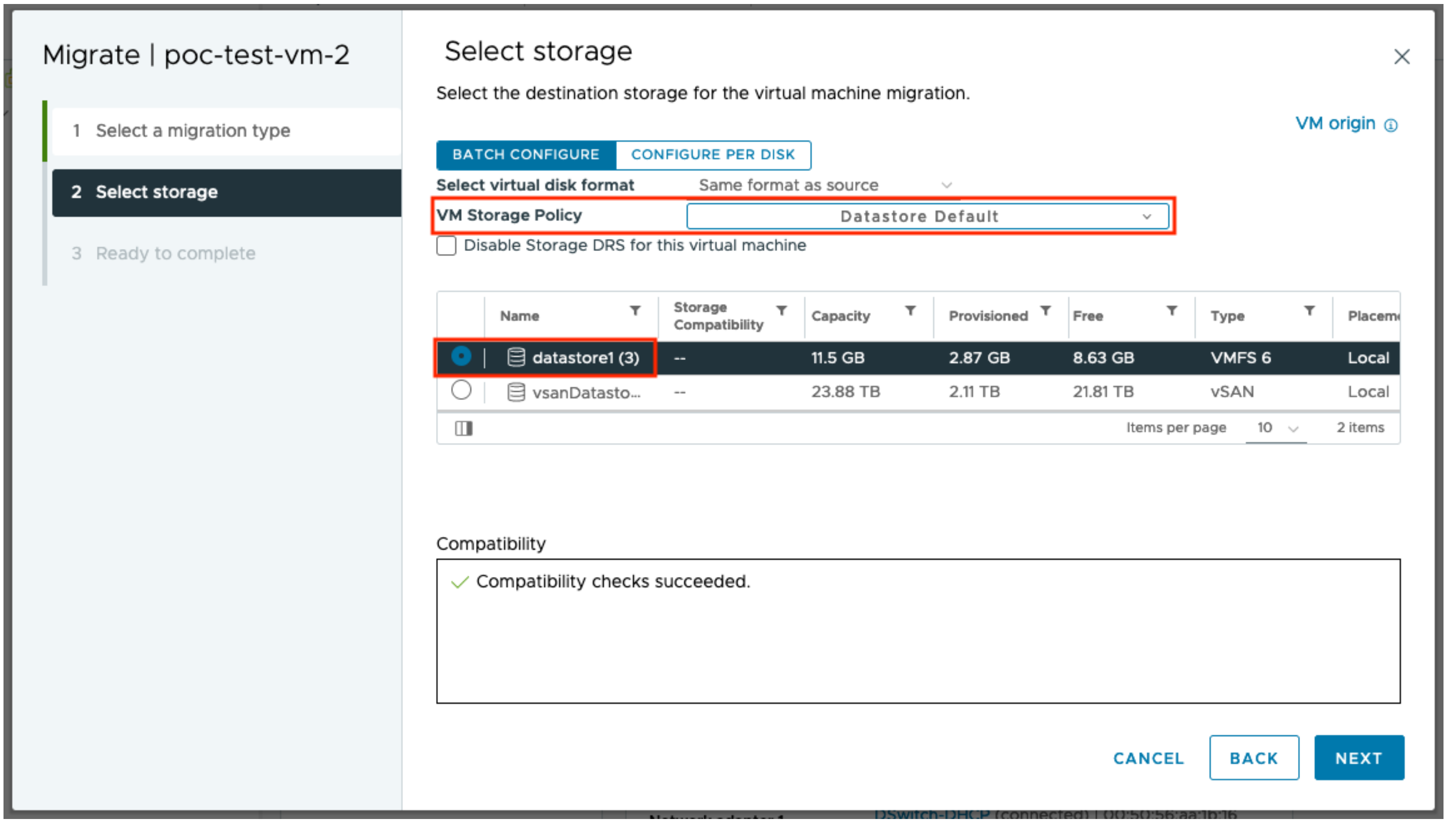
○ **Change compute resource only**
Migrate the virtual machines to another host or cluster.

● **Change storage only**
Migrate the virtual machines' storage to a compatible datastore or datastore cluster.

○ **Change both compute resource and storage**
Migrate the virtual machines to a specific host or cluster and their storage to a specific datastore or datastore cluster.

○ **Cross vCenter Server export**
Migrate the virtual machines to a vCenter Server not linked to the current SSO domain.

Select destination datastore and change the VM Storage Policy to **Datastore Default** as the vSAN policy will not apply to a VMFS (or NFS) datastore. Here we are migrating to the local VMFS datastore on the host:



Migrate | poc-test-vm-2

1 Select a migration type

2 Select storage

3 Ready to complete

Select storage

Select the destination storage for the virtual machine migration.

Select virtual disk format: Same format as source

VM Storage Policy: Datastore Default

□ Disable Storage DRS for this virtual machine

	Name	Storage Compatibility	Capacity	Provisioned	Free	Type	Placem
●	datastore1 (3)	--	11.5 GB	2.87 GB	8.63 GB	VMFS 6	Local
○	vsanDatasto...	--	23.88 TB	2.11 TB	21.81 TB	vSAN	Local

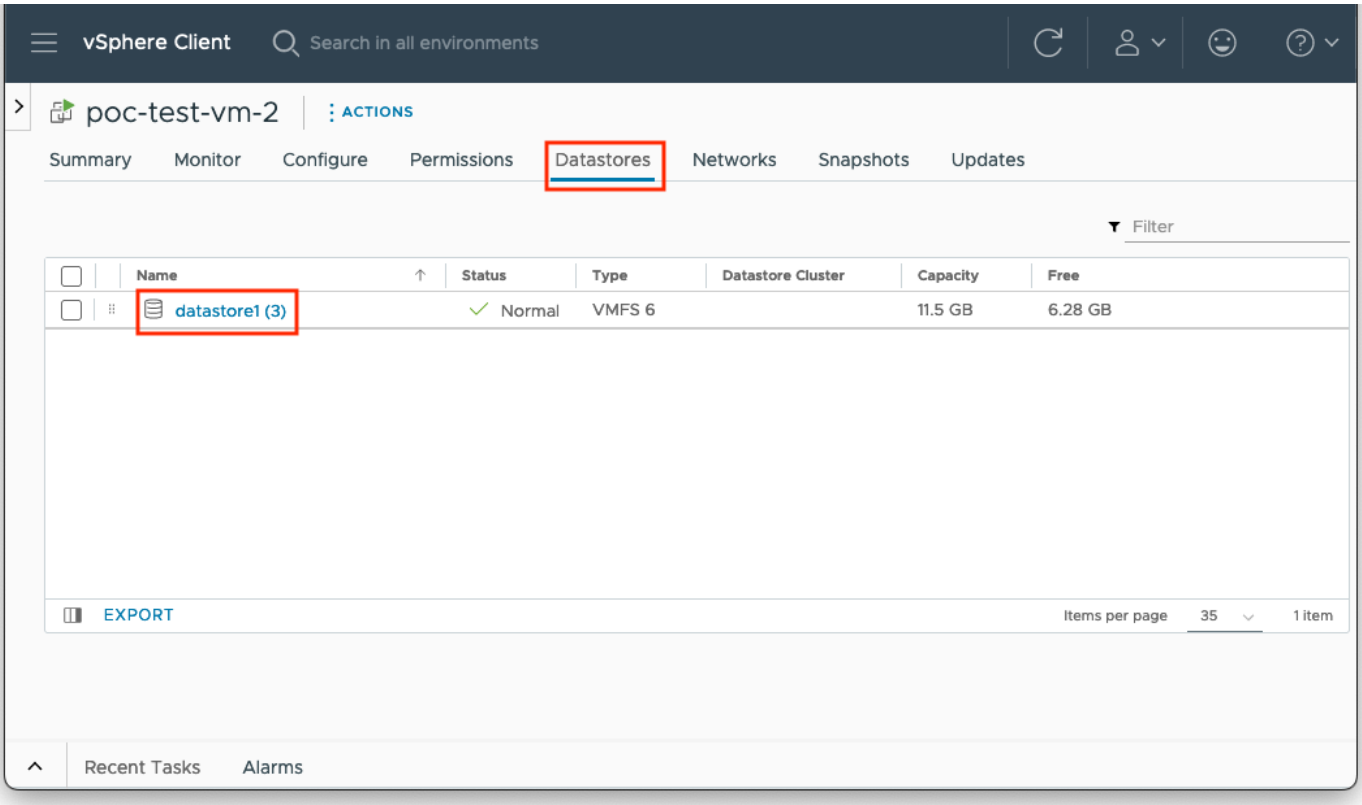
Items per page 10 2 items

Compatibility

✓ Compatibility checks succeeded.

On the “Ready to complete” screen, click **FINISH** to initiate the migration.

Once the migration completes, the ‘Datastores’ tab can be used to examine the datastore on which the VM resides.



Name	Status	Type	Datastore Cluster	Capacity	Free
datastore1 (3)	✓ Normal	VMFS 6		11.5 GB	6.28 GB

Verify that the VM has been moved to the new storage.

VM Storage Policies and vSAN

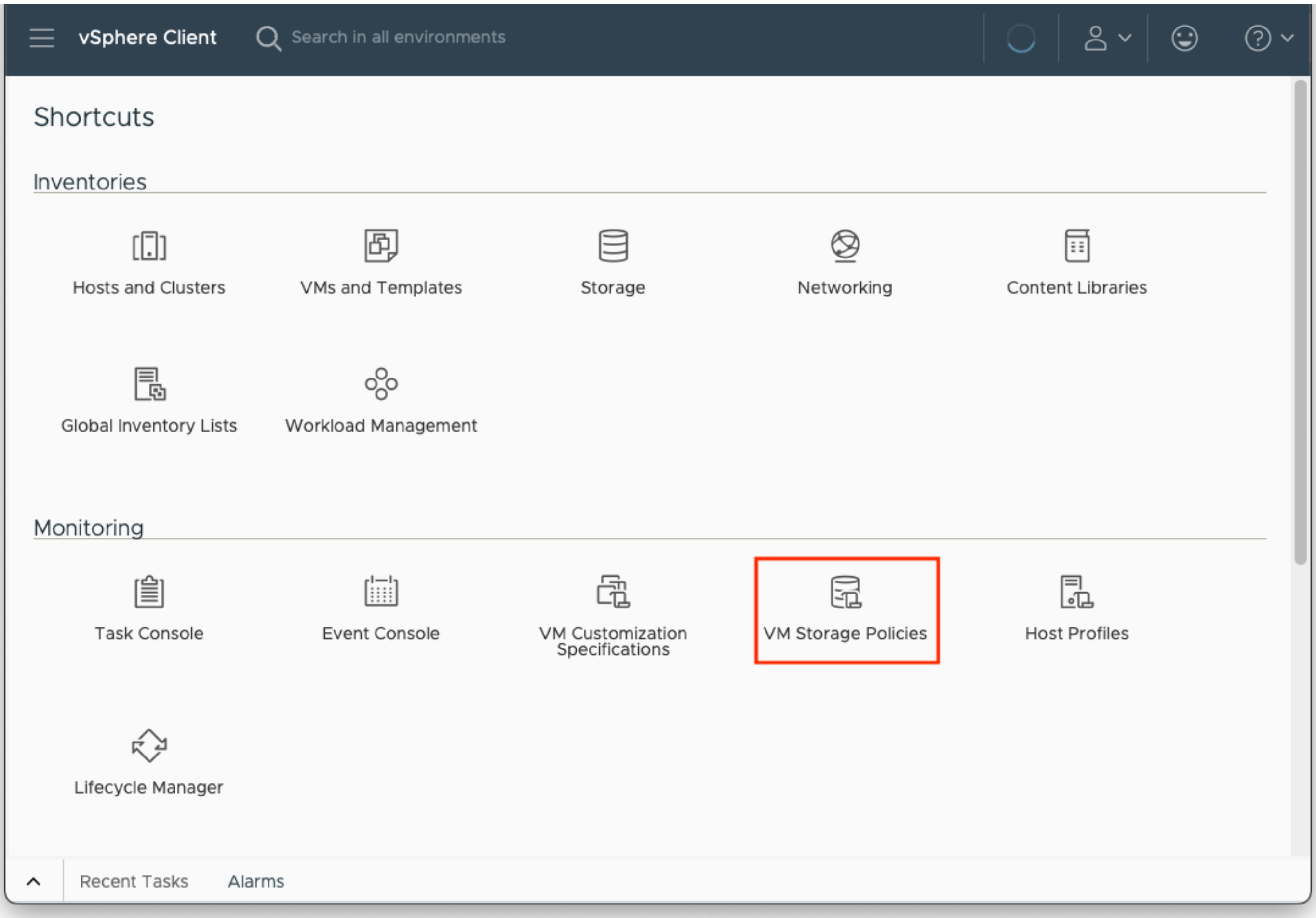
VM Storage Policies form the basis of VMware's Software-Defined Storage vision. A VM Storage Policy dictates how vSAN should place data (as well as some other features) across the physical resources, such as RAID type, number of stripes and compression (for vSAN ESA). Previously, we have deployed our VMs onto the 'vSAN Default Storage Policy'.

The actual status of the data may not reflect the policy definition (for instance, if there is a failure, or if the policy has recently been changed). Thus, VM disks can be either compliant or non-compliant with the assigned storage policy. The latter is usually a temporary state, until the system stabilizes. Once the rebuild (or other operation) is complete, compliance is automatically regained.

Note: Storage policies are applied per VMDK in vSAN OSA and per VM in vSAN ESA. Further, the recommended storage policy for vSAN ESA clusters is RAID-5 (see the vSAN features guide for more information).

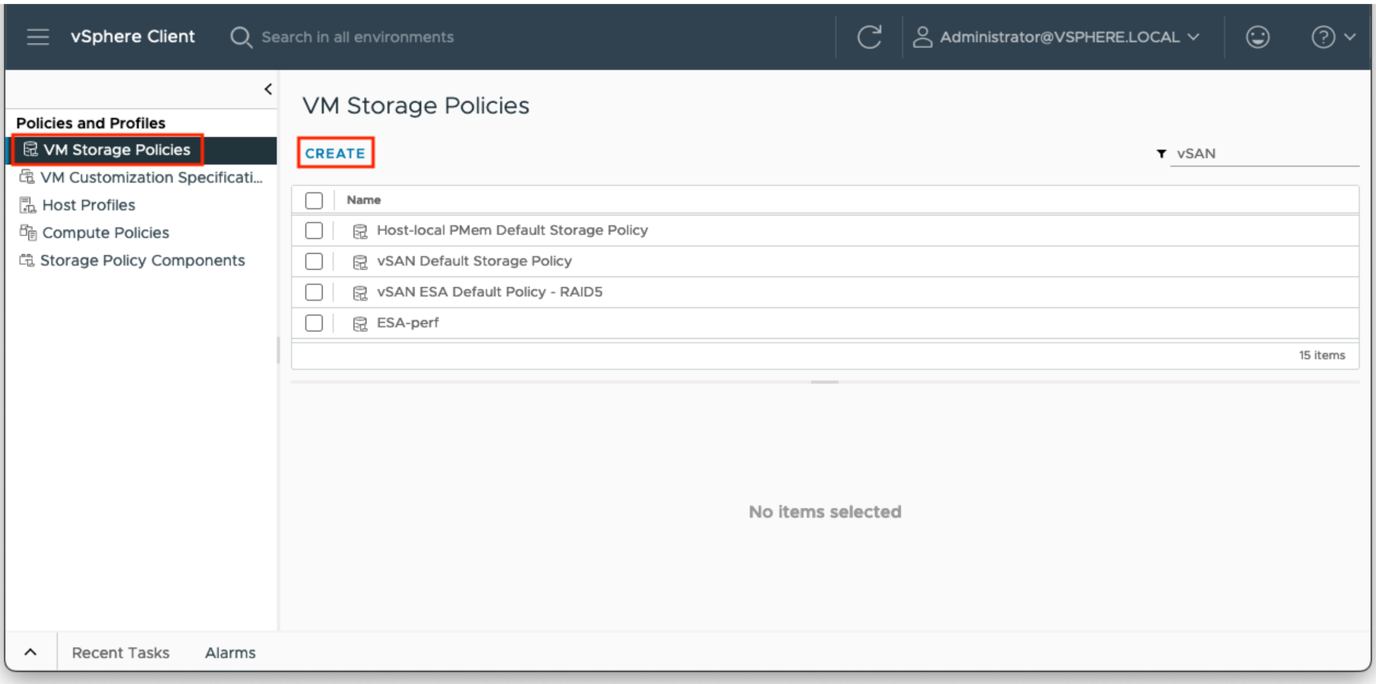
Create a New VM Storage Policy

We will build a policy with RAID 1 and a stripe width of two. The VM Storage Policies can be accessed from the 'Shortcuts' page on the vSphere client, as shown below.



Here we see the existing policies already in place, such as the ‘vSAN Default Storage Policy’ (already used to deploy VMs in the ‘Basic vSphere Functionality’ section of this guide).

To create a new policy, click on **Create**.



vSphere Client | Search in all environments | Administrator@VSPHERE.LOCAL

Policies and Profiles
- VM Storage Policies
- VM Customization Specificati...
- Host Profiles
- Compute Policies
- Storage Policy Components

VM Storage Policies

CREATE | vSAN

Name
Host-local PMem Default Storage Policy
vSAN Default Storage Policy
vSAN ESA Default Policy - RAID5
ESA-perf

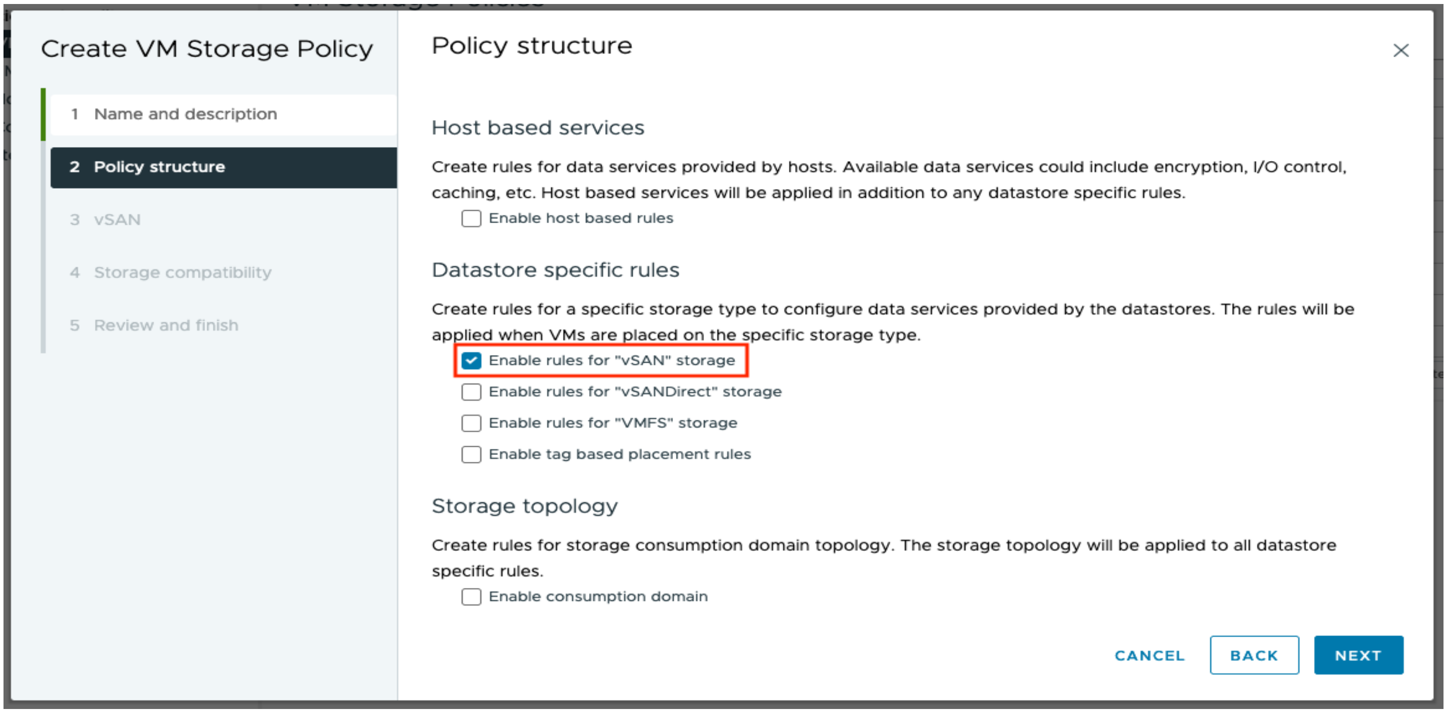
15 items

No items selected

Recent Tasks | Alarms

The next step is to provide a name and an optional description for the new VM Storage Policy. Since this policy will contain a stripe width of two, we have given it a name to reflect this. You may also give it a name to reflect that it is a vSAN policy.

The next section sets the policy structure. We select **Enable rules for "vSAN" Storage** to set a vSAN specific policy:



Create VM Storage Policy

1 Name and description
2 Policy structure
3 vSAN
4 Storage compatibility
5 Review and finish

Policy structure

Host based services

Create rules for data services provided by hosts. Available data services could include encryption, I/O control, caching, etc. Host based services will be applied in addition to any datastore specific rules.

☐ Enable host based rules

Datastore specific rules

Create rules for a specific storage type to configure data services provided by the datastores. The rules will be applied when VMs are placed on the specific storage type.

☑ Enable rules for "vSAN" storage
☐ Enable rules for "vSANDirect" storage
☐ Enable rules for "VMFS" storage
☐ Enable tag based placement rules

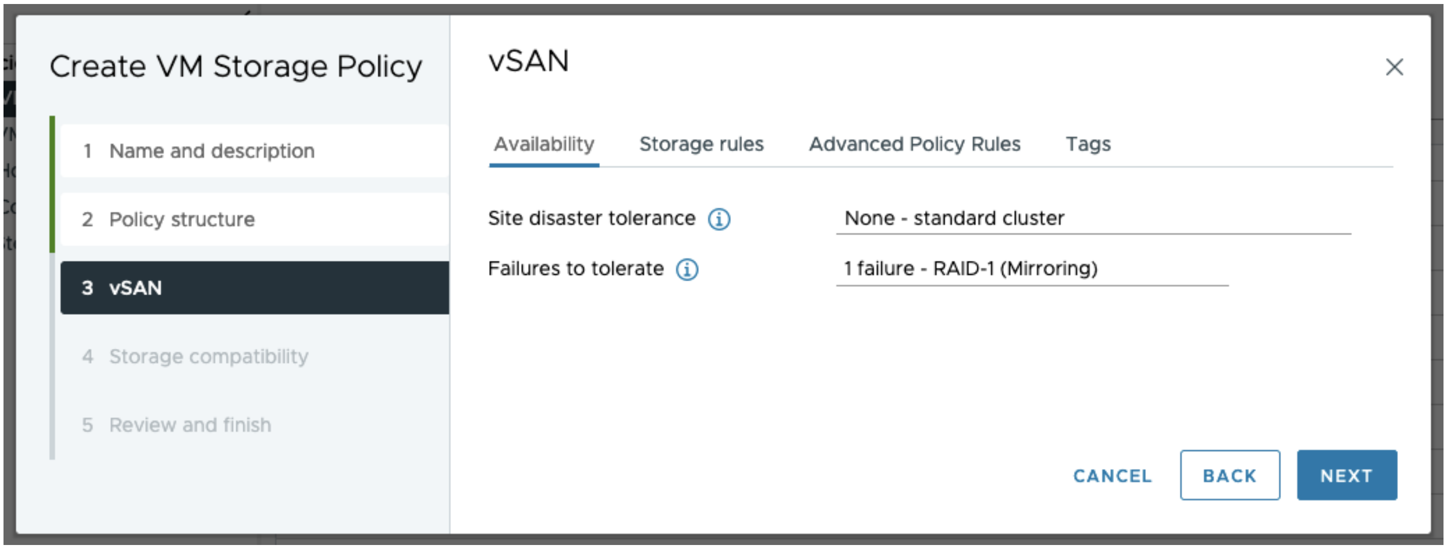
Storage topology

Create rules for storage consumption domain topology. The storage topology will be applied to all datastore specific rules.

☐ Enable consumption domain

CANCEL | BACK | NEXT

Now we get to the point where we create a set of rules. The first step is to select the availability of the objects associated with this rule. Set the failures to tolerate to one failure (RAID 1)

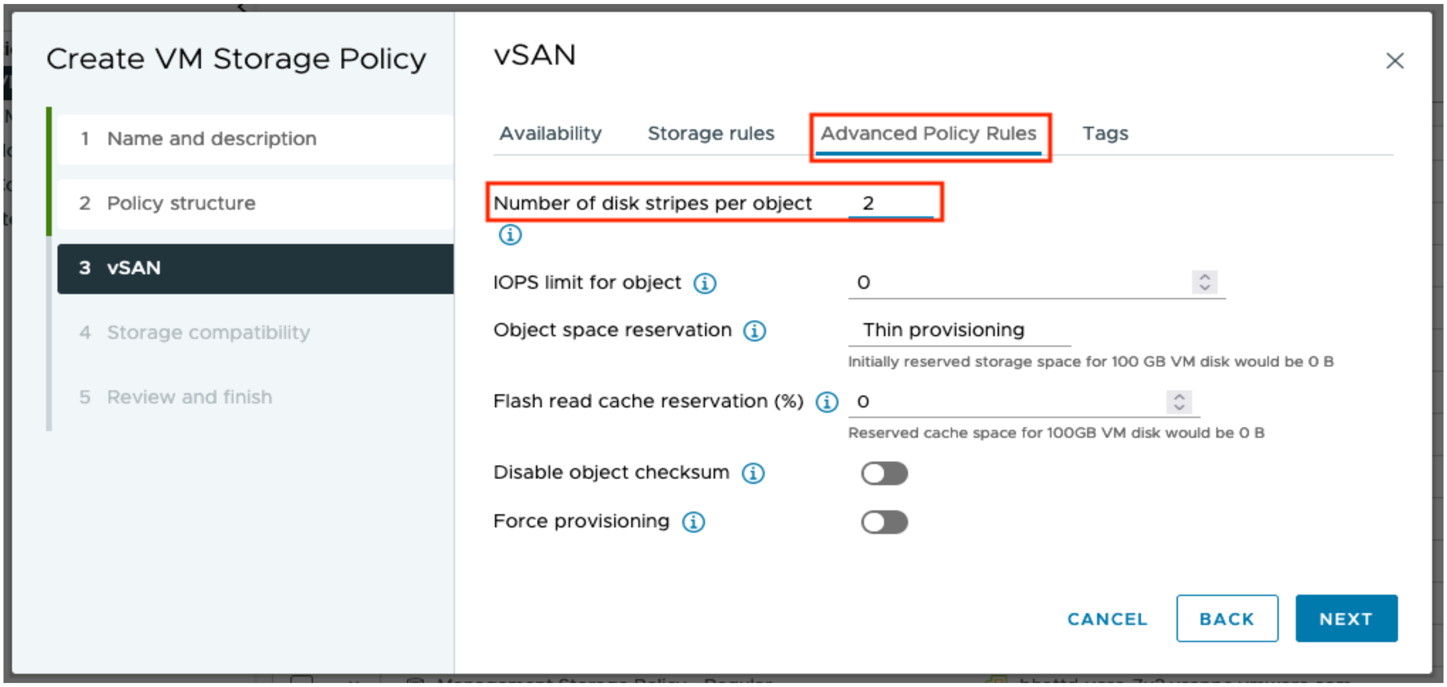


Site disaster tolerance: None - standard cluster

Failures to tolerate: 1 failure - RAID-1 (Mirroring)

We then set the Advanced Policy Rules. Once this is selected, the six customizable capabilities associated with vSAN are exposed. Since this VM Storage Policy is going to have a requirement where the stripe width of an object is set to two, this is what we select from the list of rules. It is officially called "*Number of disk stripes per object*".

Note: The general recommendation is to keep the number of stripes at the default, one (unless troubleshooting performance or for specific scenarios). Here, we are using this setting to clearly demonstrate how a storage policy affects storage components.



Number of disk stripes per object: 2

IOPS limit for object: 0

Object space reservation: Thin provisioning

Initially reserved storage space for 100 GB VM disk would be 0 B

Flash read cache reservation (%): 0

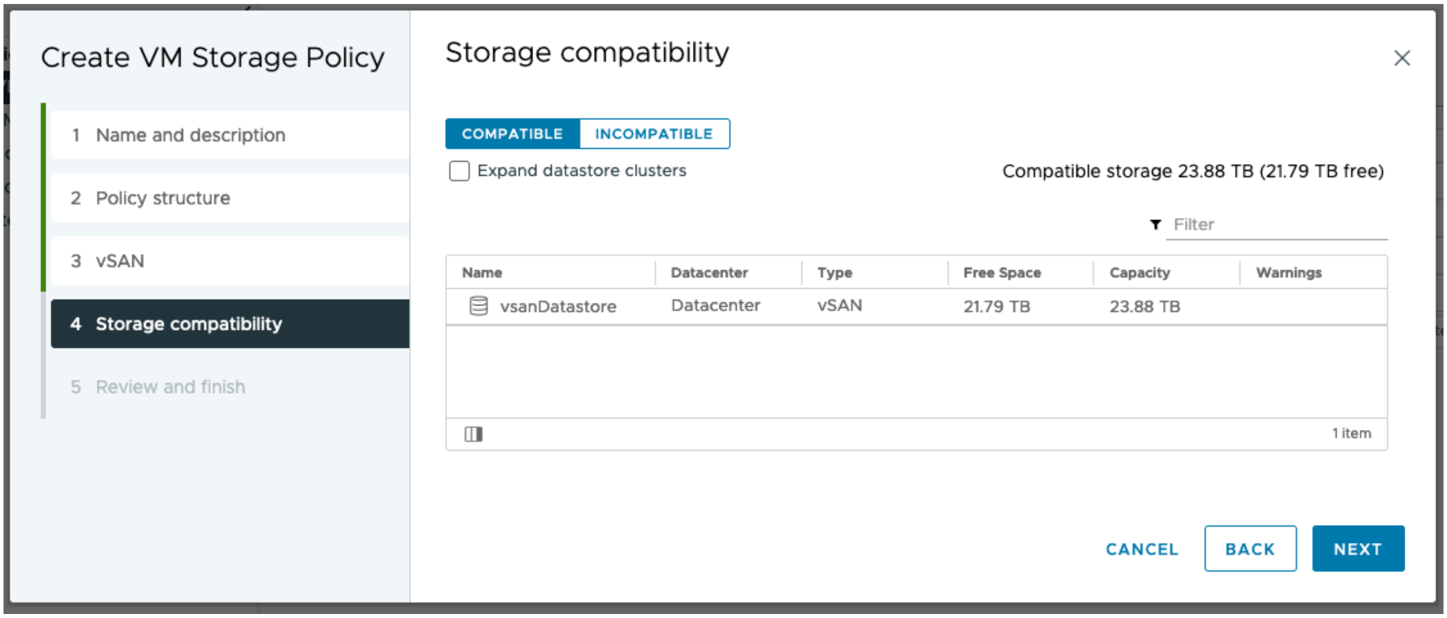
Reserved cache space for 100GB VM disk would be 0 B

Disable object checksum: ○

Force provisioning: ○

The next screen shows the datastores that are compatible with the policy. In this case, only the vsanDatastore is compatible with the policy settings.

Note: This does not mean that the vSAN datastore can successfully deploy a VM with this policy. It simply means that the vSAN datastore understands the rules or requirements in the policy. The 'force provisioning' option will try and apply the policy without first checking if it can be fulfilled by the cluster.



Storage compatibility

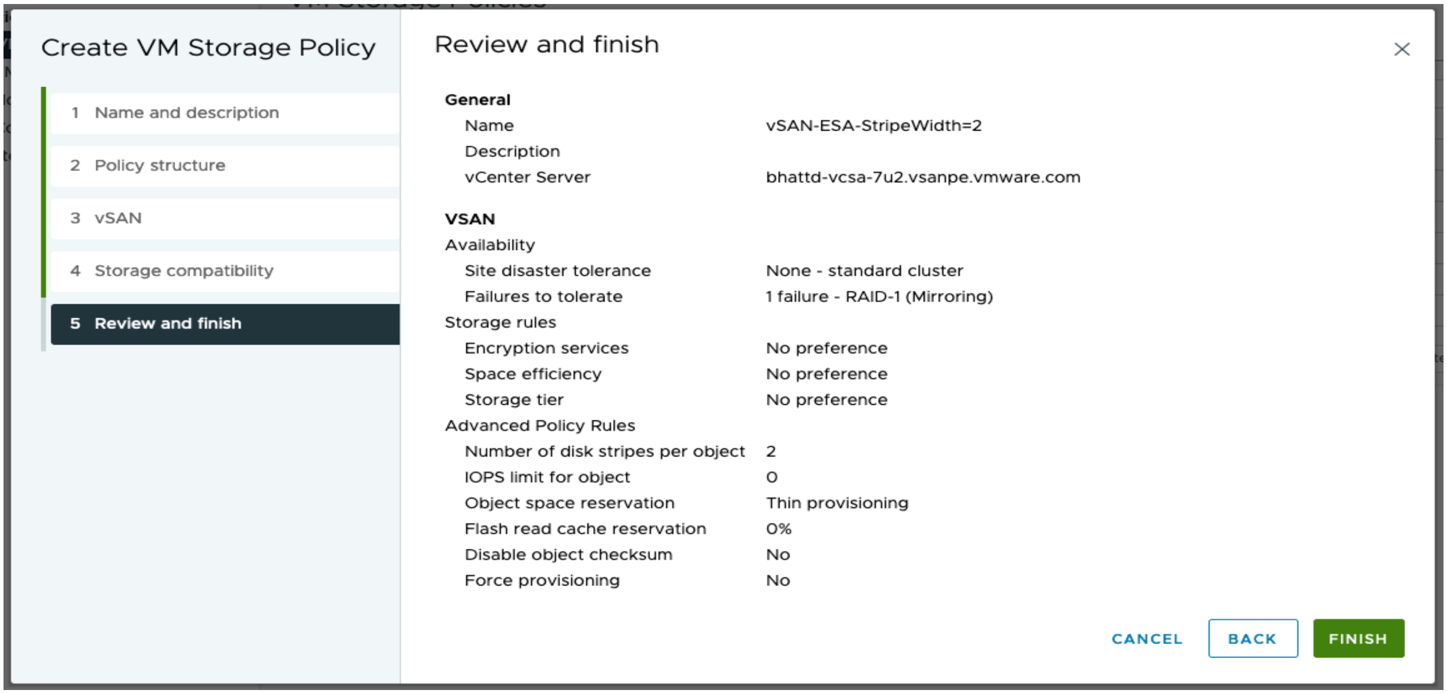
□ Expand datastore clusters

Compatible storage 23.88 TB (21.79 TB free)

Name	Datacenter	Type	Free Space	Capacity	Warnings
vsanDatastore	Datacenter	vSAN	21.79 TB	23.88 TB	

1 item

Review the settings on the next screen and click **FINISH** to create the policy:



Review and finish

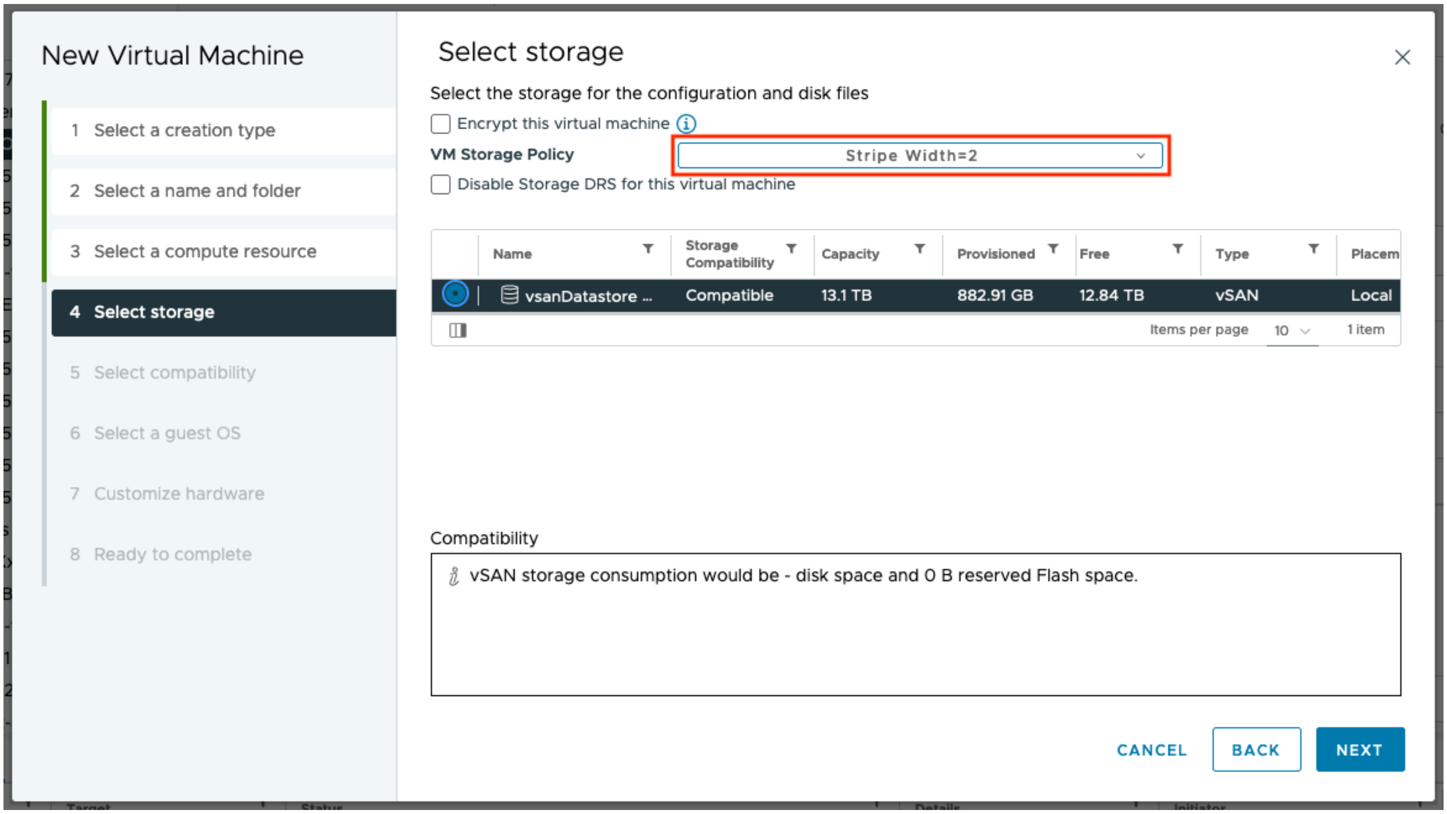
General	
Name	vSAN-ESA-StripeWidth=2
Description	
vCenter Server	bhattd-vcsa-7u2.vsanpe.vmware.com
VSAN	
Availability	
Site disaster tolerance	None - standard cluster
Failures to tolerate	1 failure - RAID-1 (Mirroring)
Storage rules	
Encryption services	No preference
Space efficiency	No preference
Storage tier	No preference
Advanced Policy Rules	
Number of disk stripes per object	2
IOPS limit for object	0
Object space reservation	Thin provisioning
Flash read cache reservation	0%
Disable object checksum	No
Force provisioning	No

We can now go ahead and deploy a VM with this new policy and see what effect it has on the layout of the underlying storage objects.

Note: vSAN includes some pre-defined storage policies for the vSAN File Service, named 'FSVM_Profile_DO_NOT_MODIFY'. These policies are used internally by vSAN for vSAN file services and should not be modified.

Deploy a new VM with a New Storage Policy

The workflow to deploy a New VM remains the same until we get to the point where the VM Storage Policy is chosen. This time, instead of selecting the default policy, select the newly created policy, as shown below. As before, the vsanDatastore should show up as the compatible datastore, and thus the one to which this VM should be provisioned. To illustrate clearly, we will see how this works on a vSAN OSA cluster to begin with.



Select storage

Select the storage for the configuration and disk files

☐ Encrypt this virtual machine

VM Storage Policy: Stripe Width=2

☐ Disable Storage DRS for this virtual machine

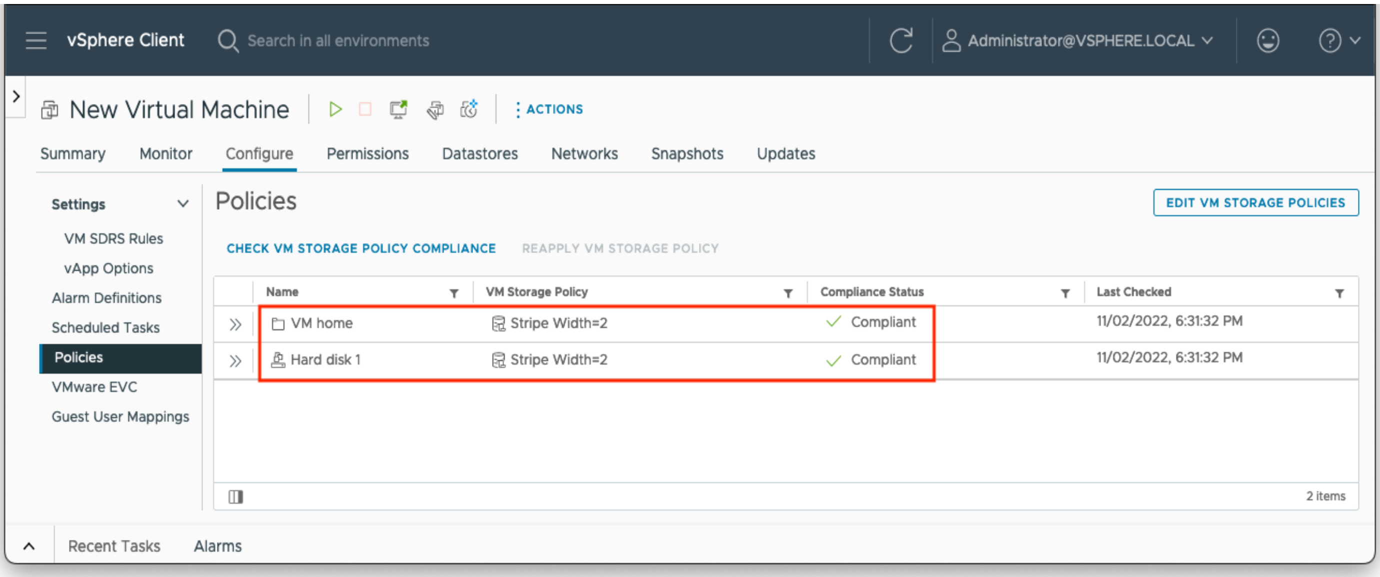
	Name	Storage Compatibility	Capacity	Provisioned	Free	Type	Placem
●	vsanDatastore ...	Compatible	13.1 TB	882.91 GB	12.84 TB	vSAN	Local

Items per page 10 1 item

Compatibility

vSAN storage consumption would be - disk space and 0 B reserved Flash space.

Now let's examine the layout of this virtual machine, and see if the policy requirements are met, i.e. do the storage objects of this VM have a stripe width of 2. First, ensure that the VM is compliant with the policy by navigating to **[VM] > Configure > Policies**, as shown here:

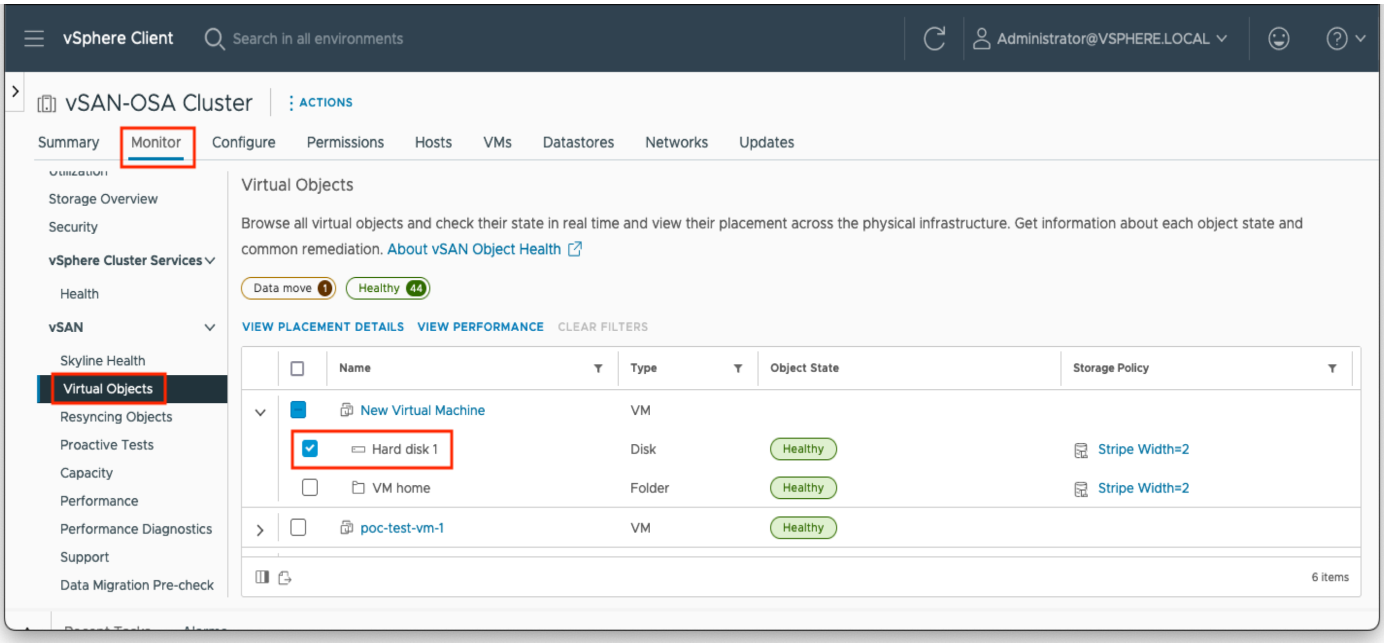


Policies

Name	VM Storage Policy	Compliance Status	Last Checked
VM home	Stripe Width=2	✓ Compliant	11/02/2022, 6:31:32 PM
Hard disk 1	Stripe Width=2	✓ Compliant	11/02/2022, 6:31:32 PM

2 items

The next step is to select the **[vSAN Cluster] > Monitor > vSAN > Virtual Objects** and check the layout of the VM's storage objects. Select the "Hard Disk 1" object and click on the **View Placement Details**:



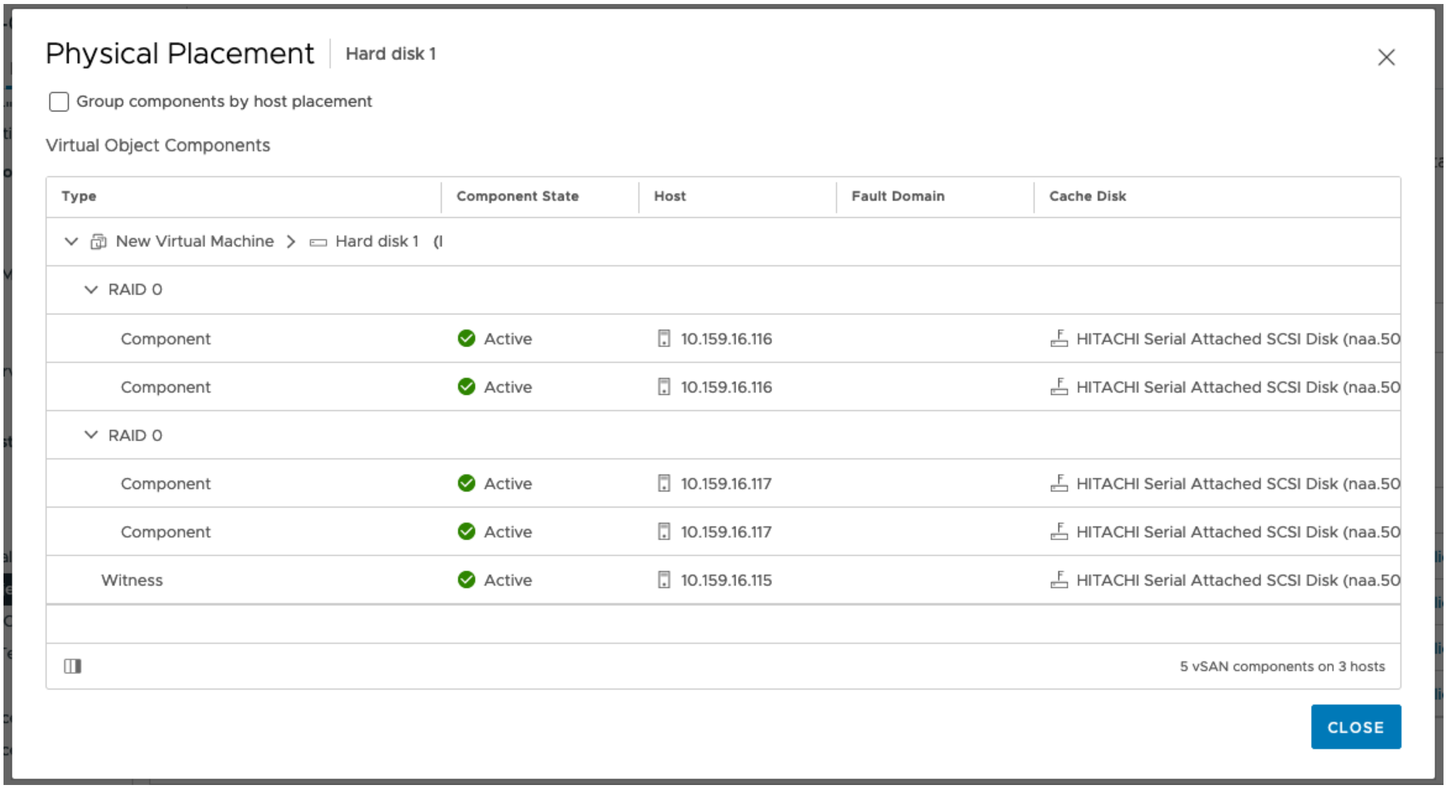
Virtual Objects

Browse all virtual objects and check their state in real time and view their placement across the physical infrastructure. Get information about each object state and common remediation. About vSAN Object Health

	Name	Type	Object State	Storage Policy
■	New Virtual Machine	VM		
☑	Hard disk 1	Disk	Healthy	Stripe Width=2
□	VM home	Folder	Healthy	Stripe Width=2
□	poc-test-vm-1	VM	Healthy	

6 items

Here, in our vSAN OSA cluster, we can see the two (RAID 0) stripe components, as defined by the policy. Each striped component must be placed on its own physical (capacity) disk. There are enough physical disks to meet this requirement here. However, a request for a larger stripe width would not be possible in this configuration. The stripes are across physical disks, and not necessarily disk groups.



Physical Placement | Hard disk 1

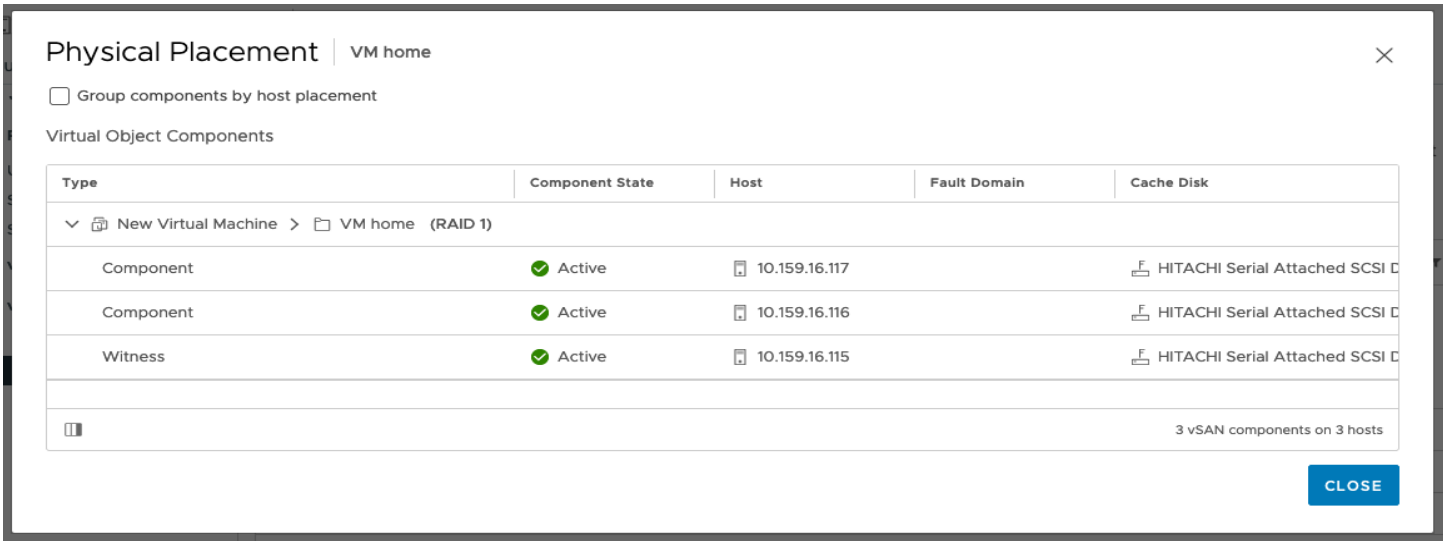
□ Group components by host placement

Virtual Object Components

Type	Component State	Host	Fault Domain	Cache Disk
New Virtual Machine > Hard disk 1 (I				
RAID 0				
Component	✓ Active	10.159.16.116		HITACHI Serial Attached SCSI Disk (naa.50
Component	✓ Active	10.159.16.116		HITACHI Serial Attached SCSI Disk (naa.50
RAID 0				
Component	✓ Active	10.159.16.117		HITACHI Serial Attached SCSI Disk (naa.50
Component	✓ Active	10.159.16.117		HITACHI Serial Attached SCSI Disk (naa.50
Witness	✓ Active	10.159.16.115		HITACHI Serial Attached SCSI Disk (naa.50

5 vSAN components on 3 hosts

However, on examining the "VM Home" object, we see an apparent policy violation – there is no striping seen. This is by design, as there is no performance gain by striping this object, so the policy setting is ignored.



Physical Placement | VM home

□ Group components by host placement

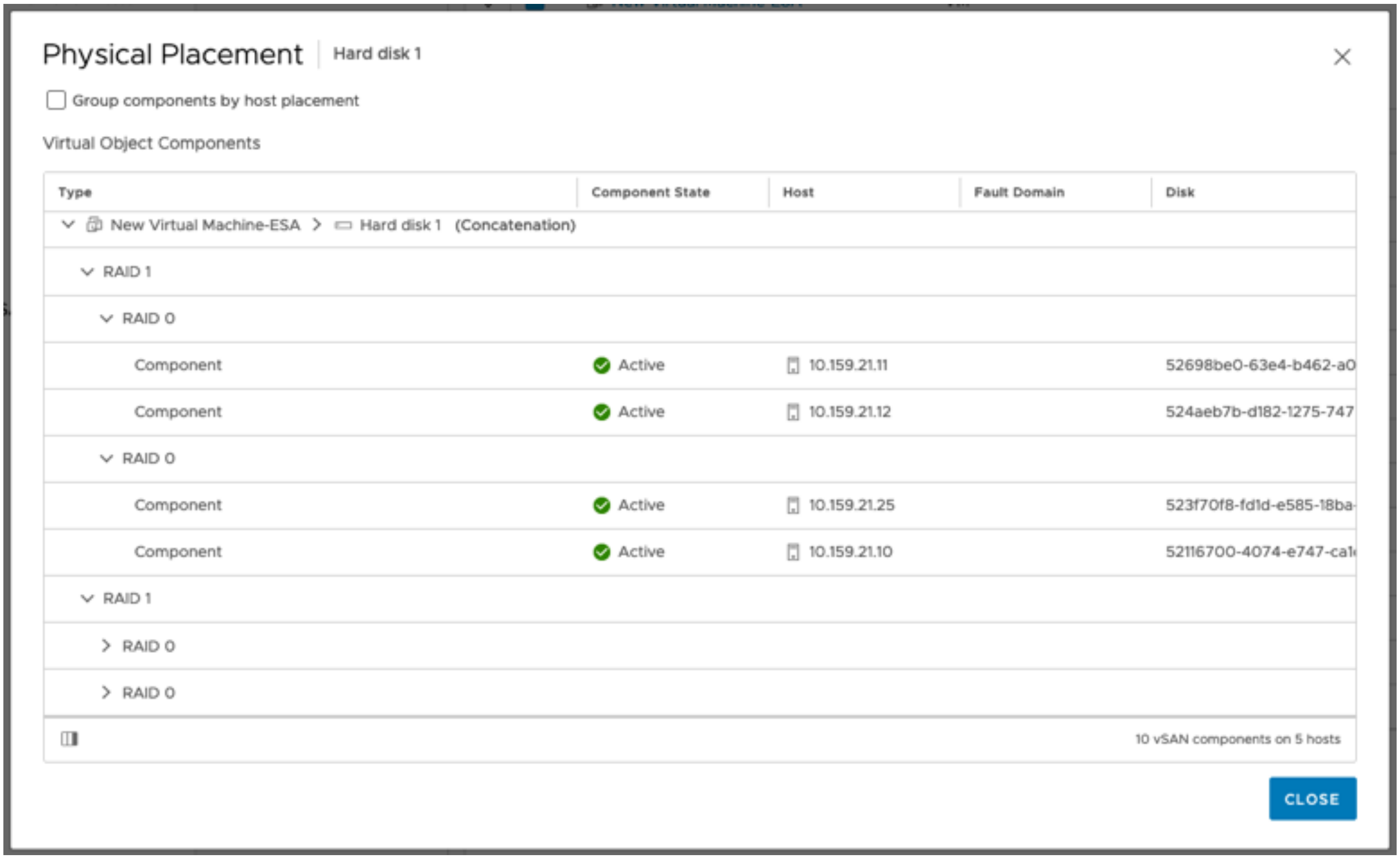
Virtual Object Components

Type	Component State	Host	Fault Domain	Cache Disk
New Virtual Machine > VM home (RAID 1)				
Component	✓ Active	10.159.16.117		HITACHI Serial Attached SCSI D
Component	✓ Active	10.159.16.116		HITACHI Serial Attached SCSI D
Witness	✓ Active	10.159.16.115		HITACHI Serial Attached SCSI D

3 vSAN components on 3 hosts

It should also be noted that snapshots taken of this base disk continue to inherit the policy of the base disk, implying that the snapshot delta objects will also be striped.

With a vSAN ESA cluster, the "Hard Disk 1" object is striped, as per the policy on the capacity leg.



Physical Placement | Hard disk 1

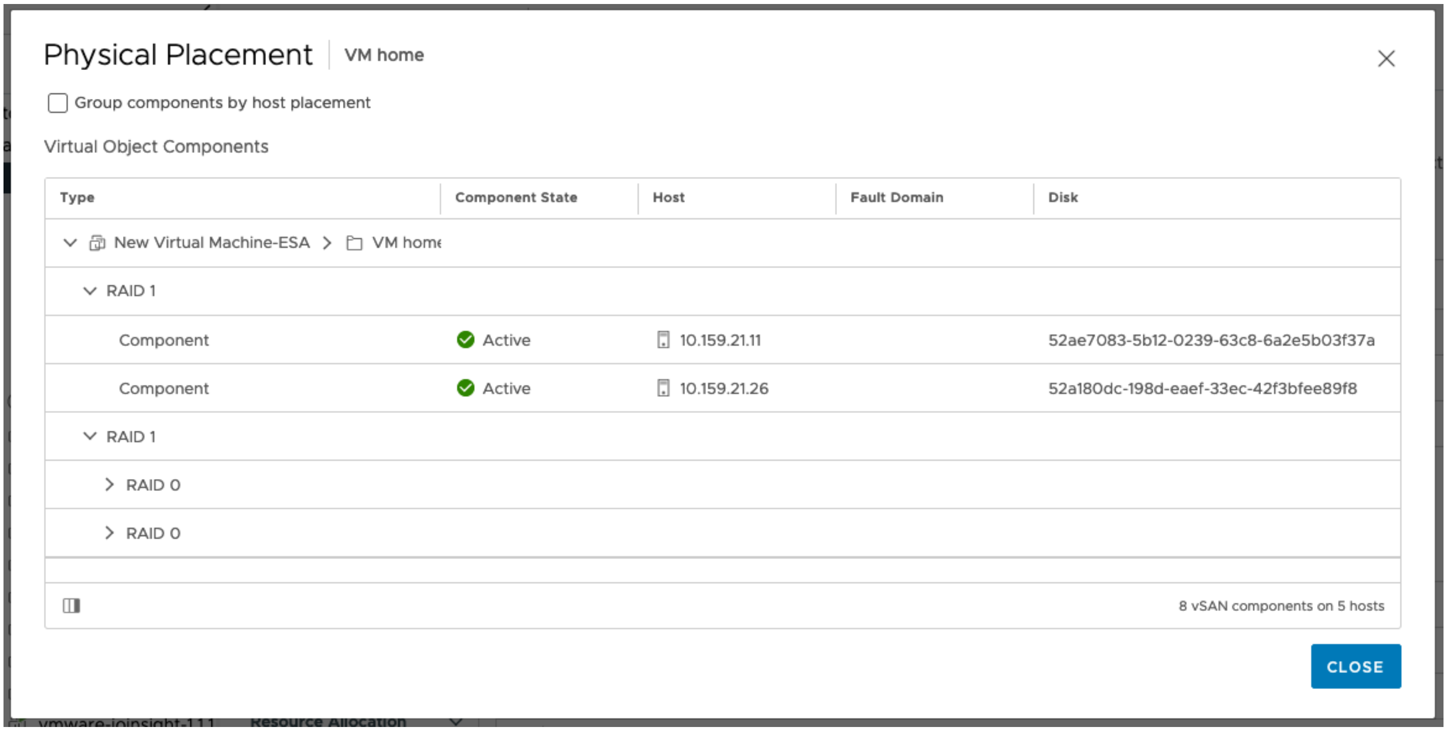
□ Group components by host placement

Virtual Object Components

Type	Component State	Host	Fault Domain	Disk
New Virtual Machine-ESA > Hard disk 1 (Concatenation)				
RAID 1				
RAID 0				
Component	Active	10.159.21.11		52698be0-63e4-b462-a0
Component	Active	10.159.21.12		524aeb7b-d182-1275-747
RAID 0				
Component	Active	10.159.21.25		523f70f8-fd1d-e585-18ba-
Component	Active	10.159.21.10		52116700-4074-e747-ca1
RAID 1				
RAID 0				
RAID 0				

10 vSAN components on 5 hosts

Again, the “VM home” object ignores the stripe width from the policy, and the capacity leg is still only RAID 1, single stripe:



Physical Placement | VM home

□ Group components by host placement

Virtual Object Components

Type	Component State	Host	Fault Domain	Disk
New Virtual Machine-ESA > VM home				
RAID 1				
Component	Active	10.159.21.11		52ae7083-5b12-0239-63c8-6a2e5b03f37a
Component	Active	10.159.21.26		52a180dc-198d-eaef-33ec-42f3bfee89f8
RAID 1				
RAID 0				
RAID 0				

8 vSAN components on 5 hosts

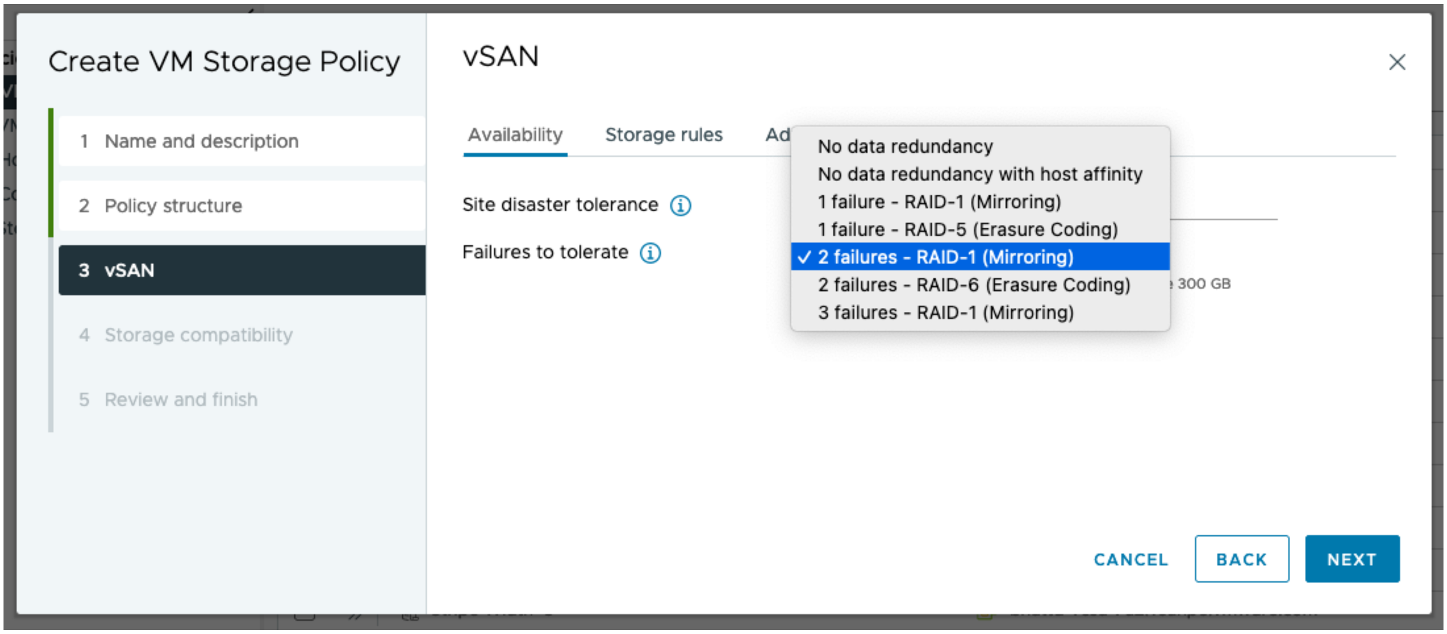
Note: The stripe width setting may have unintended consequences if used on an ESA cluster. For more information, visit: https://core.vmware.com/blog/stripe-width-storage-policy-rule-vsan-esa

Assign a new Storage Policy to an Existing VM

You can choose to modify the VM Storage Policy mapping of an existing VM deployed on the vSAN datastore. The configuration of the objects associated with the VM will be modified to comply with the newer policy. For example, if the number of failures to tolerate (FTT) is increased, newer components would be created, synchronized with the existing object, and subsequently, the original object is discarded. VM Storage policies can also be applied to individual objects.

Here, we will illustrate this (on a vSAN ESA cluster) by applying a new policy to the VM, increasing the failures to tolerate to FTT=2 (keeping RAID 1 and the stripes=2). Once again, we create a new policy. Below, we have named the policy “R1 FTT=2 Stripe Width=2’.

First, we set the FTT value:



Create VM Storage Policy

1 Name and description
2 Policy structure
3 vSAN
4 Storage compatibility
5 Review and finish

vSAN

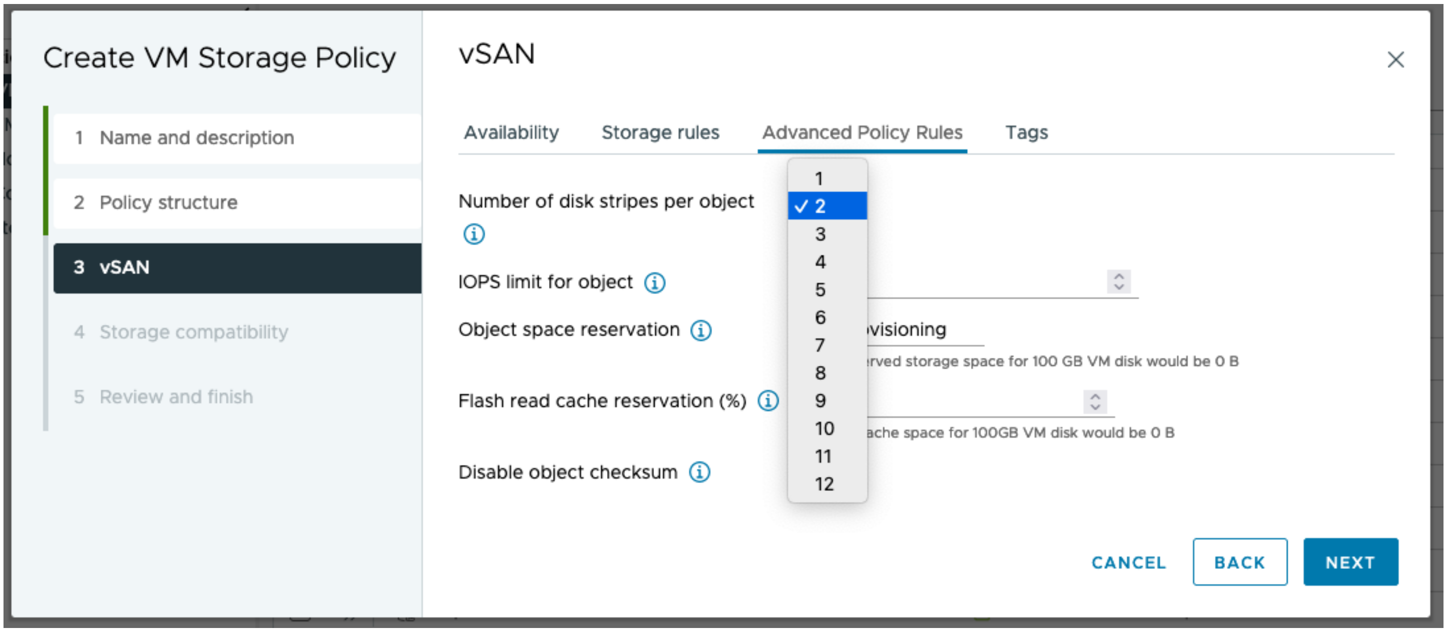
Availability
Storage rules
Ad

Site disaster tolerance

Failures to tolerate: ✓ 2 failures - RAID-1 (Mirroring)

No data redundancy
No data redundancy with host affinity
1 failure - RAID-1 (Mirroring)
1 failure - RAID-5 (Erasure Coding)
2 failures - RAID-1 (Mirroring)
2 failures - RAID-6 (Erasure Coding)
3 failures - RAID-1 (Mirroring)
300 GB
CANCEL
BACK
NEXT

Then, as before, the stripe width:



Create VM Storage Policy

1 Name and description
2 Policy structure
3 vSAN
4 Storage compatibility
5 Review and finish

vSAN

Number of disk stripes per object: ✓ 2

IOPS limit for object

Object space reservation: ...visioning

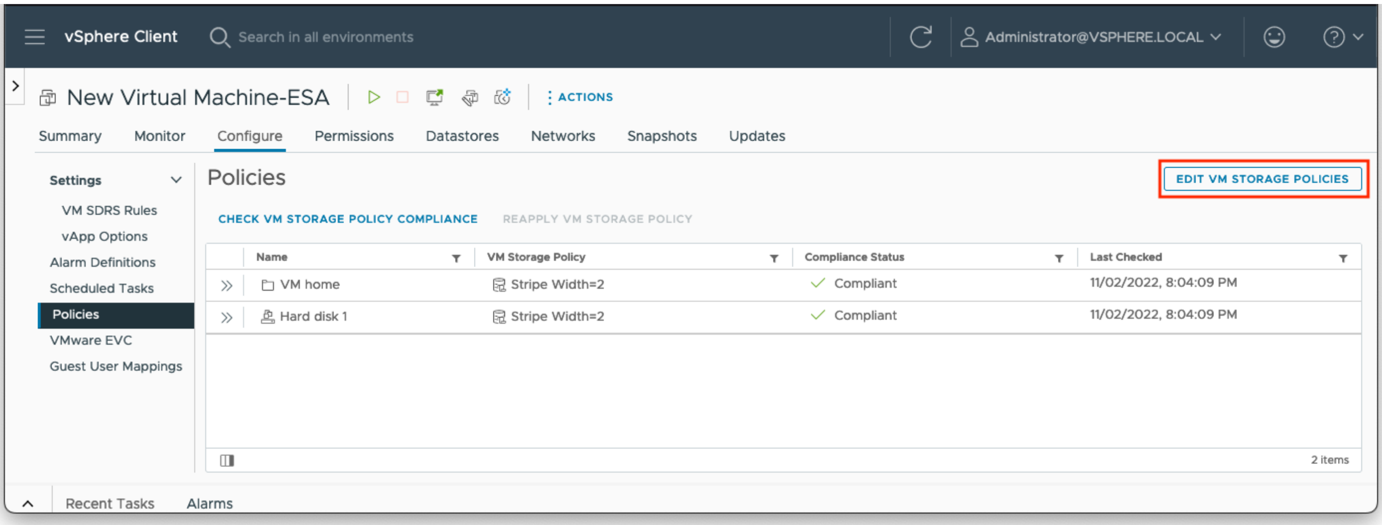
...rved storage space for 100 GB VM disk would be 0 B

Flash read cache reservation (%)

...ache space for 100GB VM disk would be 0 B

Disable object checksum

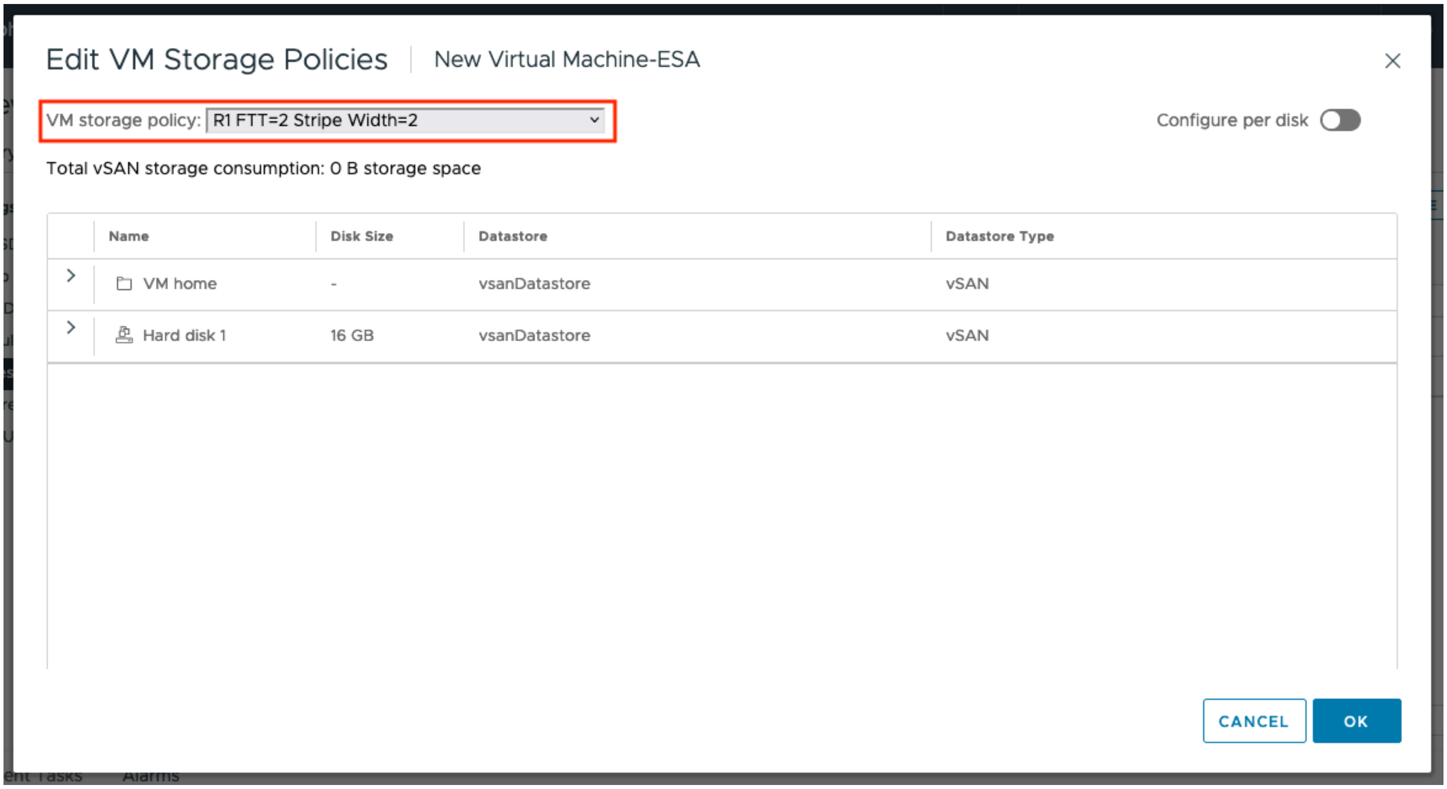
Then, we navigate to our VM, then **Configure > Policies** and click on **EDIT VM STORAGE POLICES**



Policies

Name	VM Storage Policy	Compliance Status	Last Checked
VM home	Stripe Width=2	✓ Compliant	11/02/2022, 8:04:09 PM
Hard disk 1	Stripe Width=2	✓ Compliant	11/02/2022, 8:04:09 PM

This takes you to the edit screen, where the policy can be changed. The new policy can then be selected from the drop-down list



Edit VM Storage Policies | New Virtual Machine-ESA

VM storage policy: R1 FTT=2 Stripe Width=2

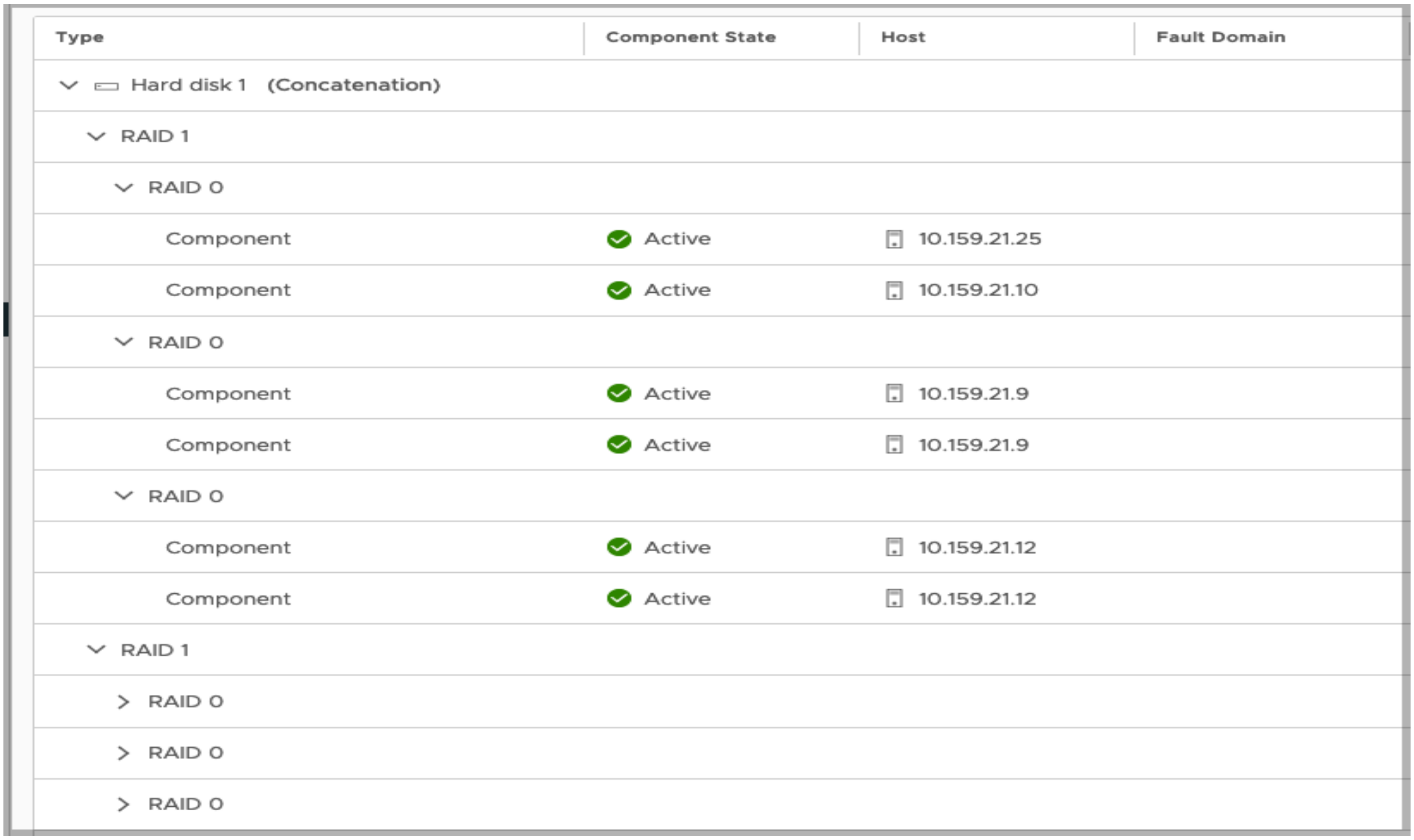
Configure per disk ○

Total vSAN storage consumption: 0 B storage space

Name	Disk Size	Datastore	Datastore Type
VM home	-	vsanDatastore	vSAN
Hard disk 1	16 GB	vsanDatastore	vSAN

Once the policy is selected, click the **OK** button as shown above to ensure the policy gets applied to all storage objects. The VM Storage Policy should now appear updated for all objects. Now when you revisit the **Configure > Policies** view, you should see the changes in the process of taking effect (Reconfiguring) or completed.

Looking at the physical disk placement, the capacity leg now has three sets of stripes (as expected). Moreover, as we have increased the FTT, the performance leg now has an extra stripe set:



Type	Component State	Host	Fault Domain
Hard disk 1 (Concatenation)			
RAID 1			
RAID 0			
Component	✓ Active	10.159.21.25	
Component	✓ Active	10.159.21.10	
RAID 0			
Component	✓ Active	10.159.21.9	
Component	✓ Active	10.159.21.9	
RAID 0			
Component	✓ Active	10.159.21.12	
Component	✓ Active	10.159.21.12	
RAID 1			
RAID 0			
RAID 0			
RAID 0			

Modify a VM Storage Policy

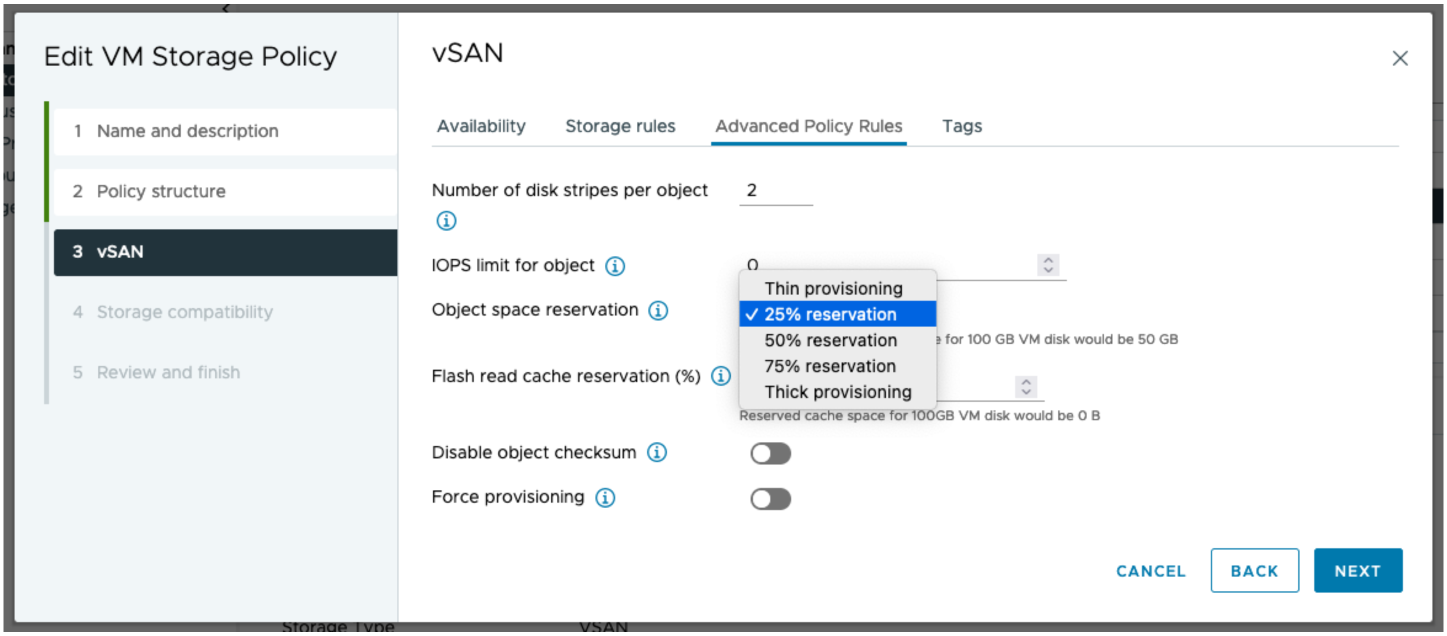
The workflow above is useful when you only need to modify the policy of one or two VMs, but if you need to change the VM Storage Policy of a significant number of VMs then this can be a little onerous. Instead, we can update by simply changing the policy used by those VMs. All VMs using those policies can then be "brought to compliance" by reconfiguring their storage object layout to make them compliant with the policy. We shall look at this next.

Note: Modifying or applying a new VM Storage Policy leads to additional backend IO as the objects are being synchronized.

In this task, we shall modify an existing VM Storage policy to set the 'Object Space Reservation' parameter to 25%. This means that each storage object will now reserve 25% of the VMDK size on the vSAN datastore. Since all VMs were deployed with 40GB VMDKs with *Failures to tolerate=1 failure - RAID-1 (Mirroring)*, the reservation value will be 20 GB.

As the first step, note the amount of free space in the vSAN datastore. This would help ascertain the impact of the change in the policy.

Select **StripeWidth=2** policy from the list of available policies, and then the **Edit Settings** option. Navigate to **vSAN > Advanced Policy Rules** and modify the **Object space reservation** setting to 25%, as shown below:



Edit VM Storage Policy

1 Name and description
2 Policy structure
3 vSAN
4 Storage compatibility
5 Review and finish

vSAN

Availability | Storage rules | Advanced Policy Rules | Tags

Number of disk stripes per object: 2

IOPS limit for object: 0

Object space reservation: ✓ 25% reservation (options: Thin provisioning, 25% reservation, 50% reservation, 75% reservation, Thick provisioning)

… for 100 GB VM disk would be 50 GB

Flash read cache reservation (%)

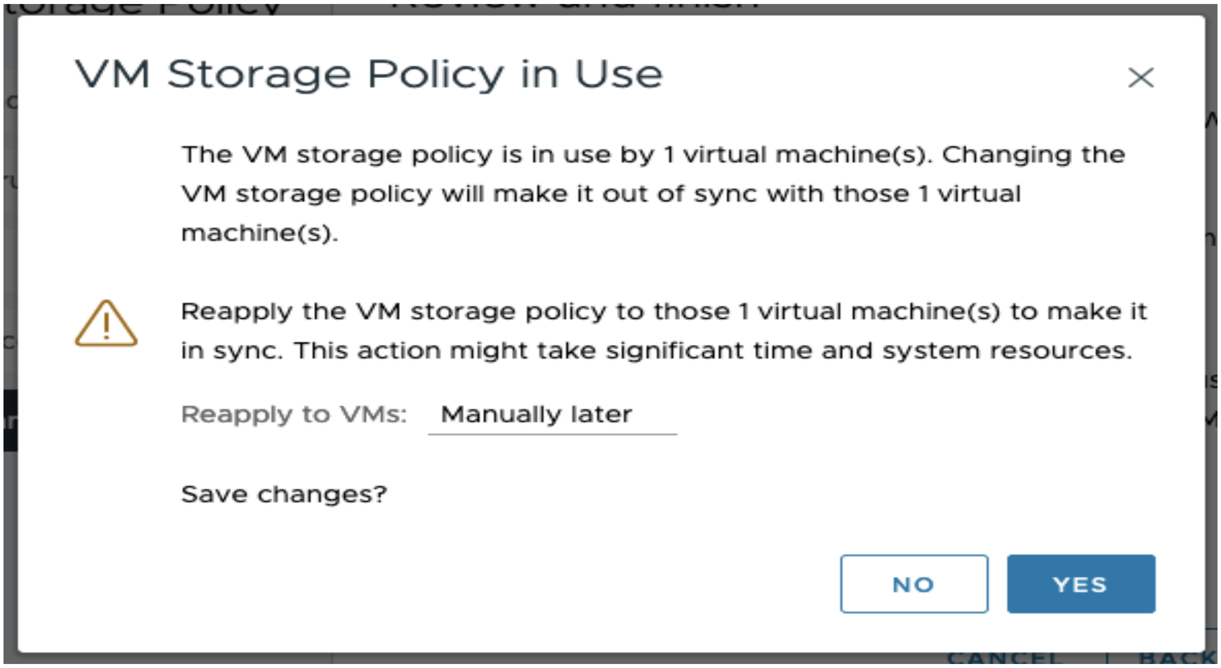
Reserved cache space for 100GB VM disk would be 0 B

Disable object checksum

Force provisioning

CANCEL | BACK | NEXT

Proceed to complete the wizard with default values and click **FINISH**. A pop-up message requiring user input appears with details of the number of VMs using the policy being modified. This is to ascertain the impact of the policy change. Typically, such changes are recommended to be performed during a maintenance window. You can choose to enforce a policy change immediately or defer it to be changed manually at a later point. Leave it at the default, which is "Manually later", by clicking Yes as shown below:



VM Storage Policy in Use

The VM storage policy is in use by 1 virtual machine(s). Changing the VM storage policy will make it out of sync with those 1 virtual machine(s).

⚠ Reapply the VM storage policy to those 1 virtual machine(s) to make it in sync. This action might take significant time and system resources.

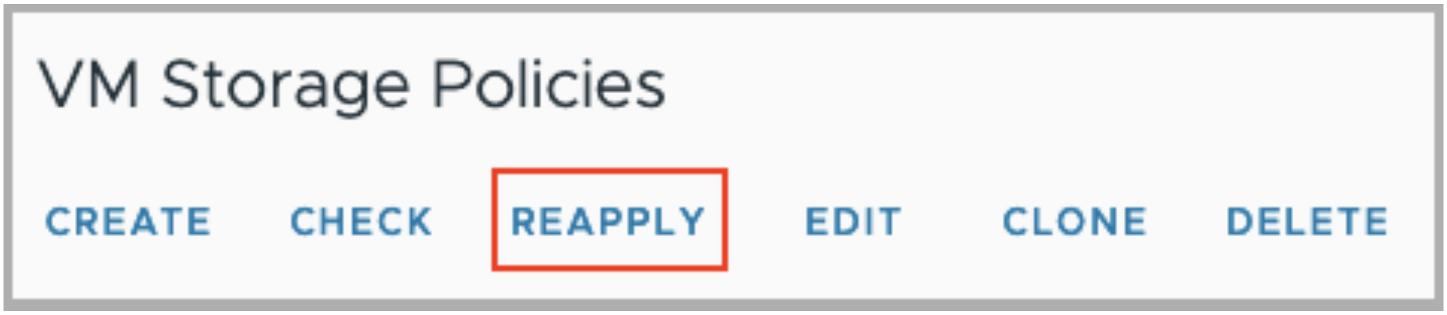
Reapply to VMs: Manually later

Save changes?

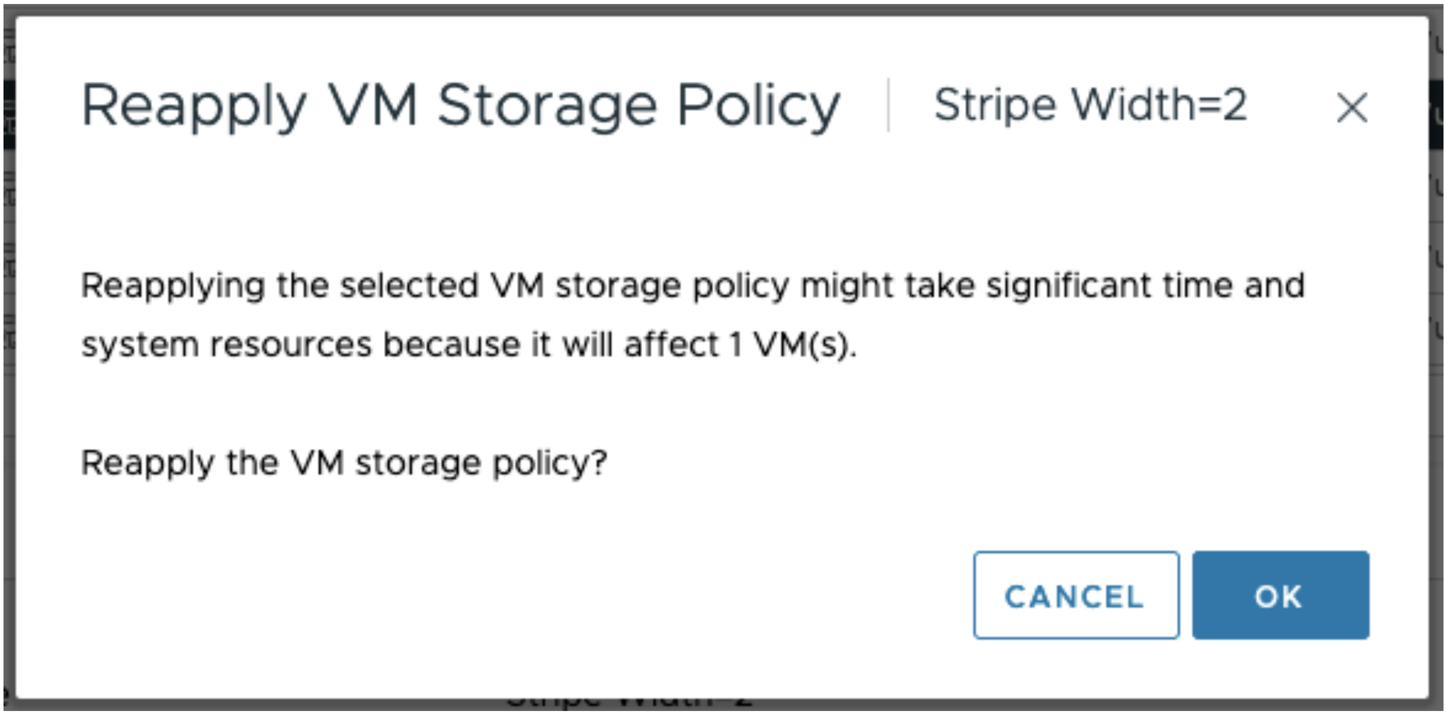
NO | YES

Next, select the Storage policy -*StripeWidth=2* and click on the **VM Compliance** tab in the bottom pane. It will display the two VMs along with their storage objects, and the fact that they are no longer compliant with the policy. They are in an "Out of Date" compliance state as the policy has now been changed.

You can now enforce a policy change by navigating to **[VM Storage Policies]** and clicking on **Reapply VM Storage Policy**:



When this button is clicked, the following popup appears.



Reapplying the selected VM storage policy might take significant time and system resources because it will affect 1 VM(s).

Reapply the VM storage policy?

When the reconfigure activity completes against the storage objects, and the compliance state is once again checked, everything should show as *Compliant*.

Since we have now included an *ObjectSpaceReservation* value in the policy, you may notice corresponding capacity reduction from the vSAN datastore.

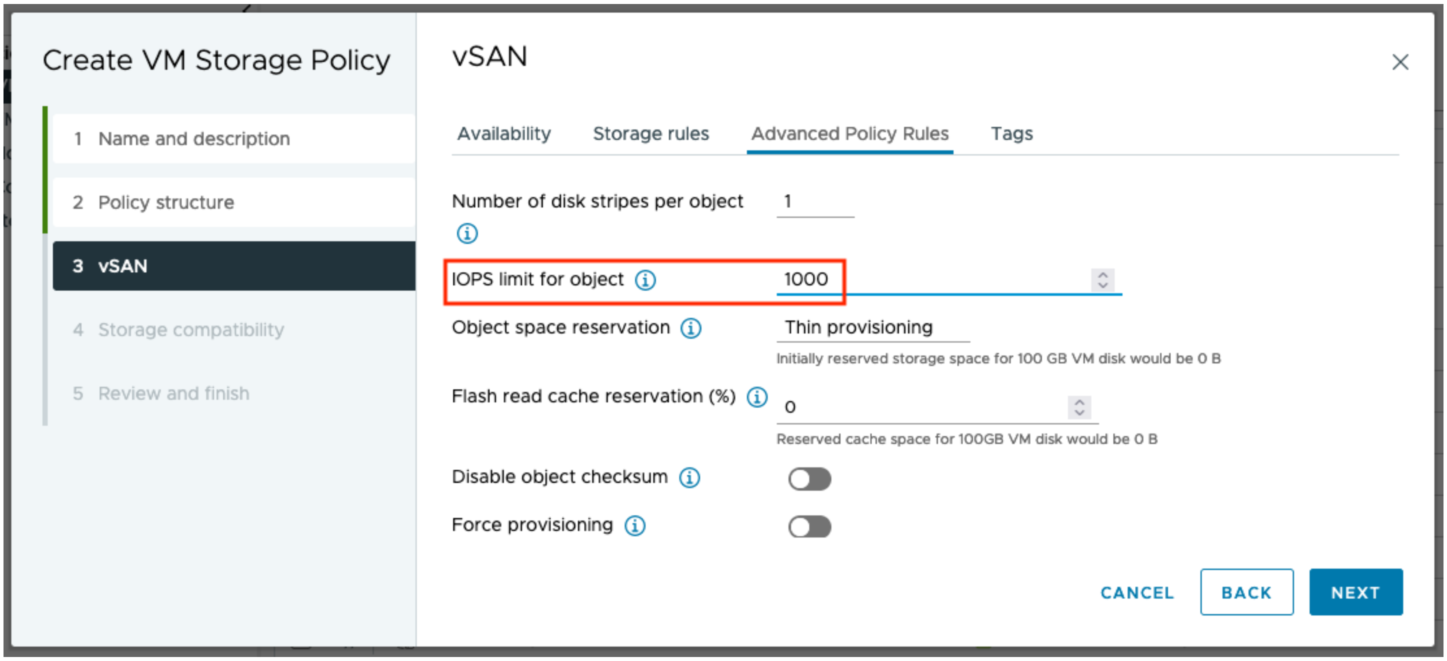
For example, the two VMs with the new policy change have 40GB storage objects. Therefore, there is a 25% ObjectSpaceReservation implying 10GB is reserved per VMDK. So that's 10GB per VMDK, 1 VMDK per VM, 2 VMs equals 20

GB reserved space, right? However, since the VMDK is also mirrored, there is a total of 40GB reserved on the vSAN datastore.

IOPS Limits

vSAN incorporates a quality-of-service feature that can limit the number of IOPS an object may consume. IOPS limits are enabled and applied via a policy setting. The setting can be used to ensure that a particular virtual machine does not consume more than its fair share of resources or negatively impact the performance of the cluster.

Here, we demonstrate setting IOPS limits by creating a new policy (as above). We can then set the IOPS limit by navigating to the 'Advanced Policy Rules' tab. In our example, we have set the IOPS limit to 1000:



Create VM Storage Policy

1 Name and description
2 Policy structure
3 vSAN
4 Storage compatibility
5 Review and finish

vSAN

Number of disk stripes per object: 1

IOPS limit for object: 1000

Object space reservation: Thin provisioning

Initially reserved storage space for 100 GB VM disk would be 0 B

Flash read cache reservation (%): 0

Reserved cache space for 100GB VM disk would be 0 B

Disable object checksum

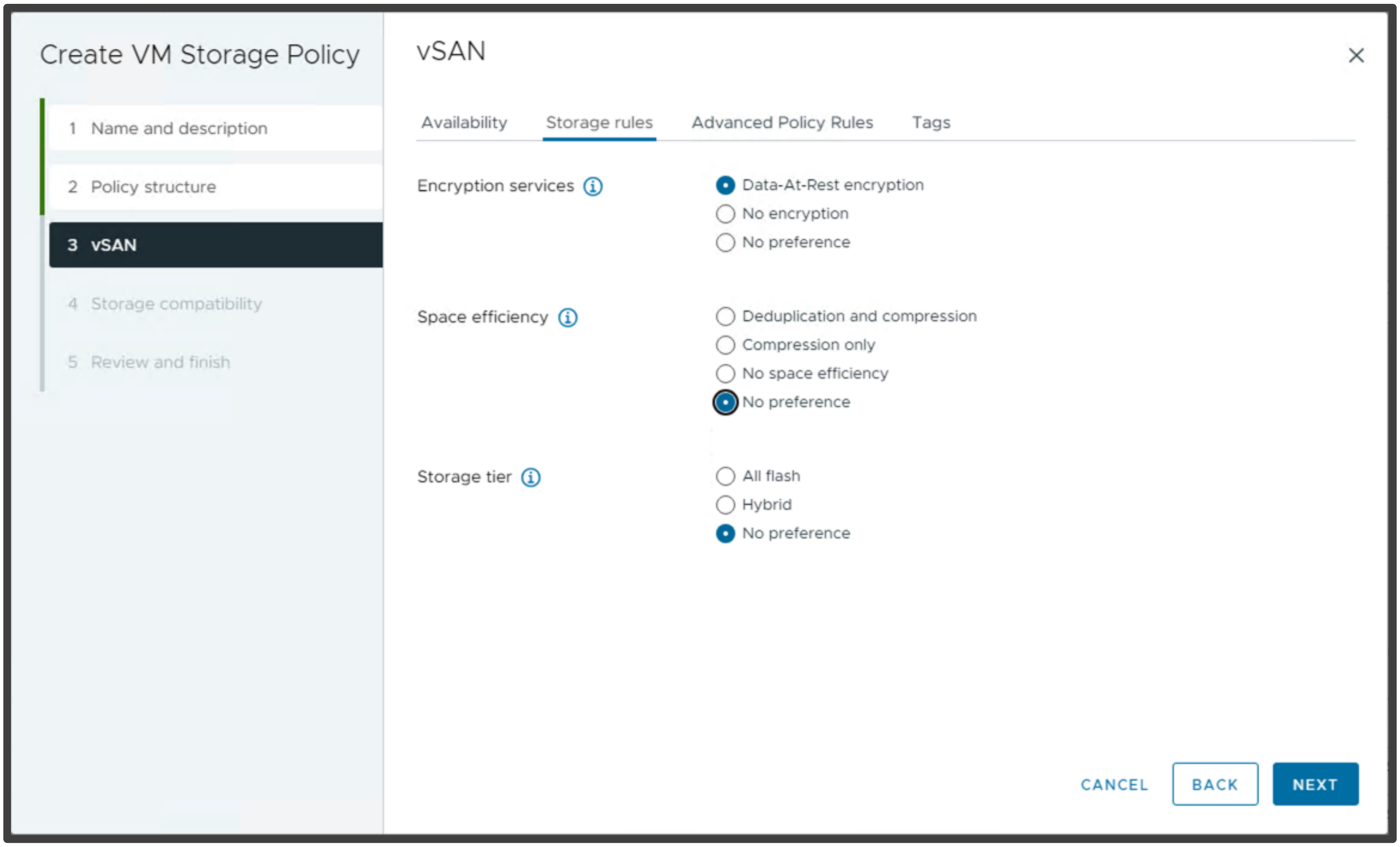
Force provisioning

The IOPS limit value is calculated as the number of 32KB IOs per second. Therefore, in this example, where we have a value of 1000, the IOPS limit is 1000x32KB=32MB/s. If I/O against the VM or VMDK should rise above the threshold, the additional I/O will be throttled. *Note that any I/O incurred by a snapshot is counted too.*

Space Efficiency & Encryption Defined by Policy (vSAN ESA)

In vSAN ESA compression can be enabled/disabled by VM storage policy. The policy applies to any new data, so turning off compression will only affect any new writes. **Note that compression is turned on by default in vSAN ESA.**

In the example below, we create a new policy (as per above) and then navigate to the 'Storage rules' tab to configure the services. Here, we enable encryption and disable compression:



Create VM Storage Policy

1 Name and description

2 Policy structure

3 vSAN

4 Storage compatibility

5 Review and finish

vSAN

Availability | Storage rules | Advanced Policy Rules | Tags

Encryption services
● Data-At-Rest encryption
○ No encryption
○ No preference

Space efficiency
○ Deduplication and compression
○ Compression only
○ No space efficiency
● No preference

Storage tier
○ All flash
○ Hybrid
● No preference

Note: Turning off compression will affect new writes only. Existing data is not affected.

APPENDIX A: vSAN RDMA Configuration & Troubleshooting

Overview

RDMA (Remote Direct Memory Access) provides a more efficient network transport layer than traditional TCP connections. For more details visit https://core.vmware.com/blog/vsan-7-update-2-rdma-support

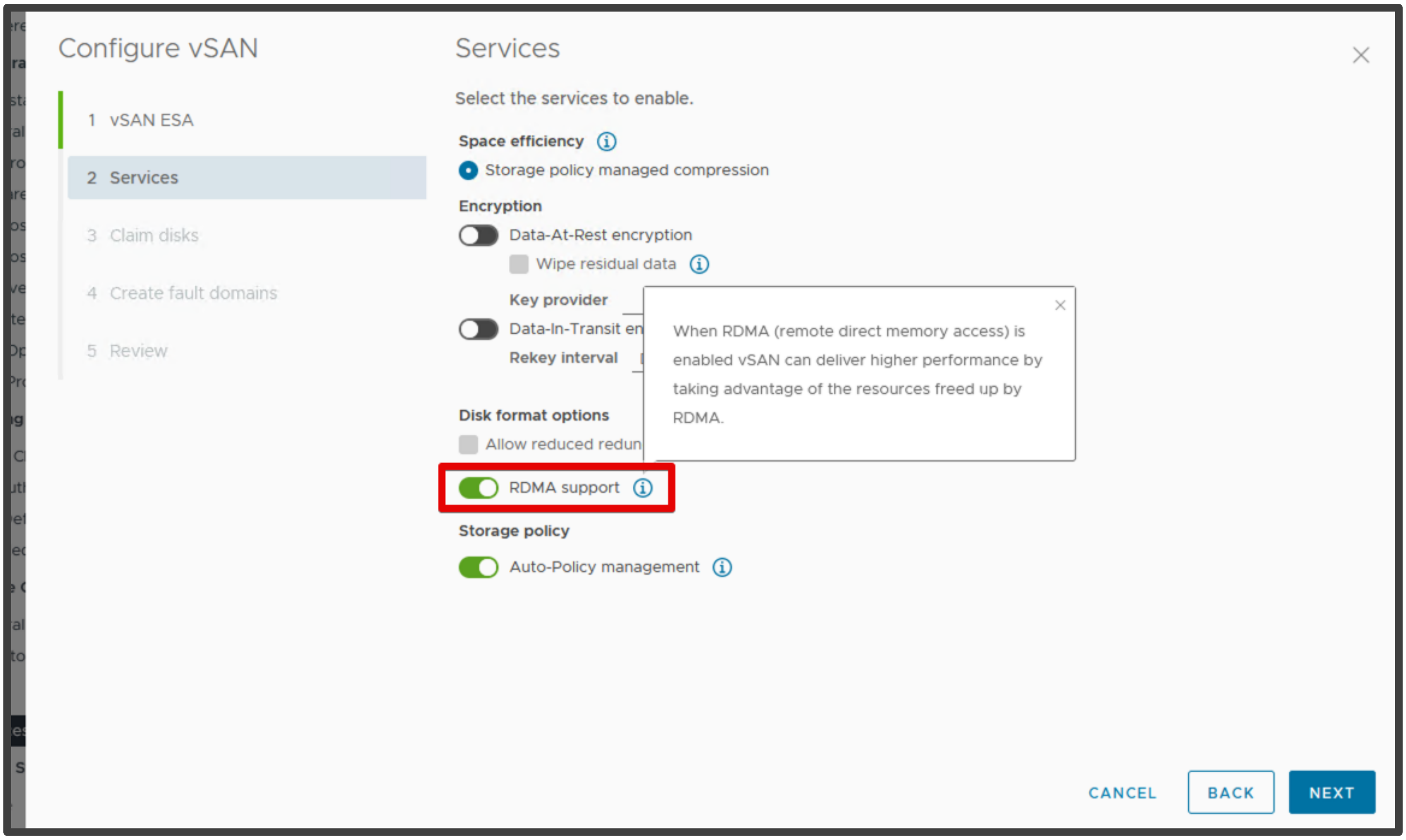
Requirements:

- A supported physical network adapter that has RDMA RoCEv2 capabilities; see the vSAN VCG
- A switch that supports RDMA RoCEv2 and associated configuration.

*Note that neither **vSAN Stretched Clustering nor two-node clusters are supported with RDMA and that LACP should not be configured on the network uplinks.***

To enable RDMA support, it can be achieved using HCI Quick Start wizard

Toggle to enable and select next



Configure vSAN

1 vSAN ESA
2 Services
3 Claim disks
4 Create fault domains
5 Review

Services

Select the services to enable.

Space efficiency

● Storage policy managed compression

Encryption

Data-At-Rest encryption

□ Wipe residual data

Key provider

Data-In-Transit en

Rekey interval

Disk format options

□ Allow reduced redun

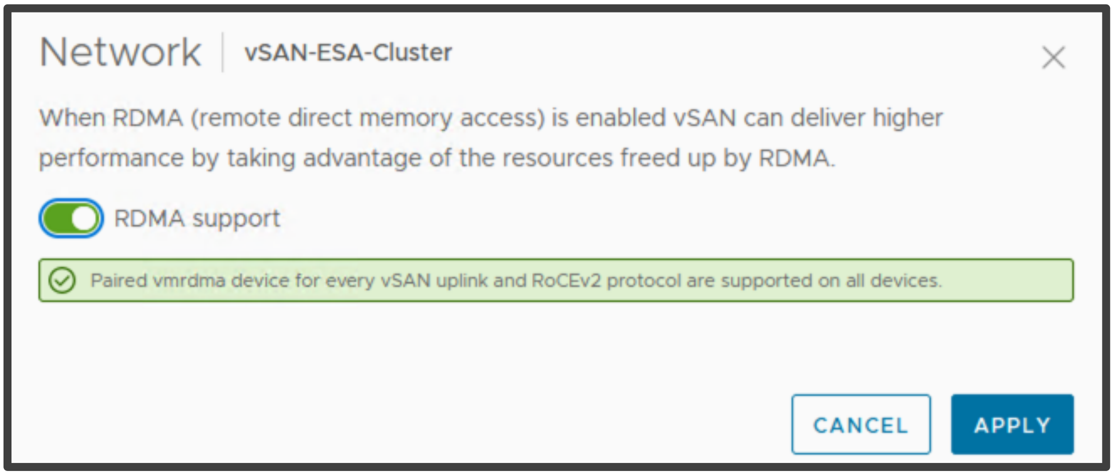
RDMA support (enabled)

Storage policy

Auto-Policy management (enabled)

When RDMA (remote direct memory access) is enabled vSAN can deliver higher performance by taking advantage of the resources freed up by RDMA.

Alternatively on an existing cluster, toggle to enable and click on **Apply**:



Network | vSAN-ESA-Cluster

When RDMA (remote direct memory access) is enabled vSAN can deliver higher performance by taking advantage of the resources freed up by RDMA.

RDMA support

Paired vmrdma device for every vSAN uplink and RoCEv2 protocol are supported on all devices.

Configuration Example

First, enable PFC support on the switch. The example below shows the configuration steps on a Mellanox SN2100:

Enable PFC:

```
switch01 [standalone: master] (config) # dcb priority-flow-control enable
This action might cause traffic loss while shutting down a port with priority-flow-control mode on
Type 'yes' to confirm enable pfc globally: yes

Enable Switch PFC for priority 4
dcb priority-flow-control priority 4 enable
```

Assign PFC to port (ESXi uplink):

```
switch01 [standalone: master] (config) # interface ethernet 1/9 dcb priority-flow-control mode on
force
```

Verify RDMA available adapter through ESXi shell:

```
[root@localhost:~] esxcli rdma device list


Name     Driver      State    MTU  Speed     Paired Uplink  Description
-------  ----------  ------  ----  --------  -------------  -----------
vmrdma0  nmlx5_rdma  Active  4096  100 Gbps  vmnic4         MT27700 Family
vmrdma1  nmlx5_rdma  Active  4096  100 Gbps  vmnic5         MT27700 Family
```

Looking at each virtual RDMA adapter, we see details on state, MTU size (see hardware specific documentation) and the linked adapter.

Note: To take advantage of RDMA you must have jumbo frames enabled on the physical switch. The RDMA adapter provides <= 4096 (maximum) MTU size.

Verify ESXi RDMA PFC status:

```
[root@localhost:~]esxcli network nic dcb status get -n vmnic4

Nic Name: vmnic4
Mode: 3 - IEEE Mode
Enabled: true
Capabilities:
Priority Group: true
Priority Flow Control: true
PG Traffic Classes: 8
PFC Traffic Classes: 8
PFC Enabled: true
PFC Configuration: 0 0 0 0 1 1 0 0
```

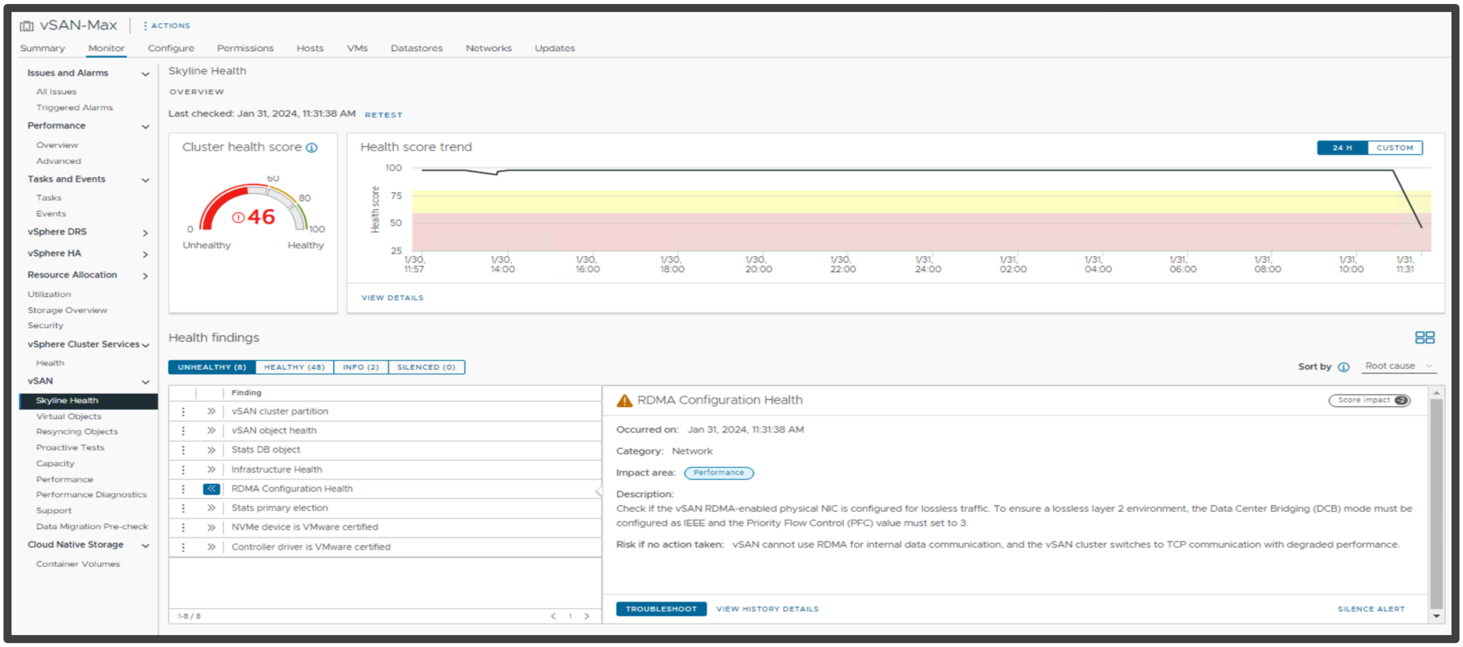
If we receive an error here, double check the driver/firmware combination as per vSphere HCL. The vSAN Health check invokes a similar process to query the device DCB status.

Verify ESXi RDMA available protocols:

```
[root@localhost:~] esxcli rdma device protocol list

Device       RoCE v1      RoCE v2      iWARP
-------      -------      -------      -----
vmrdma0             true         true         false
vmrdma1             true         true         false
```

Verify vSAN Skyline Health check in vCenter:

In the screenshot below, we see a RDMA Configuration Health issue:



Skyline Health

OVERVIEW

Last checked: Jan 31, 2024, 11:31:38 AM RETEST

Health findings

Finding
vSAN cluster partition
vSAN object health
Stats DB object
Infrastructure Health
RDMA Configuration Health
Stats primary election
NVMe device is VMware certified
Controller driver is VMware certified

RDMA Configuration Health

Occurred on: Jan 31, 2024, 11:31:38 AM

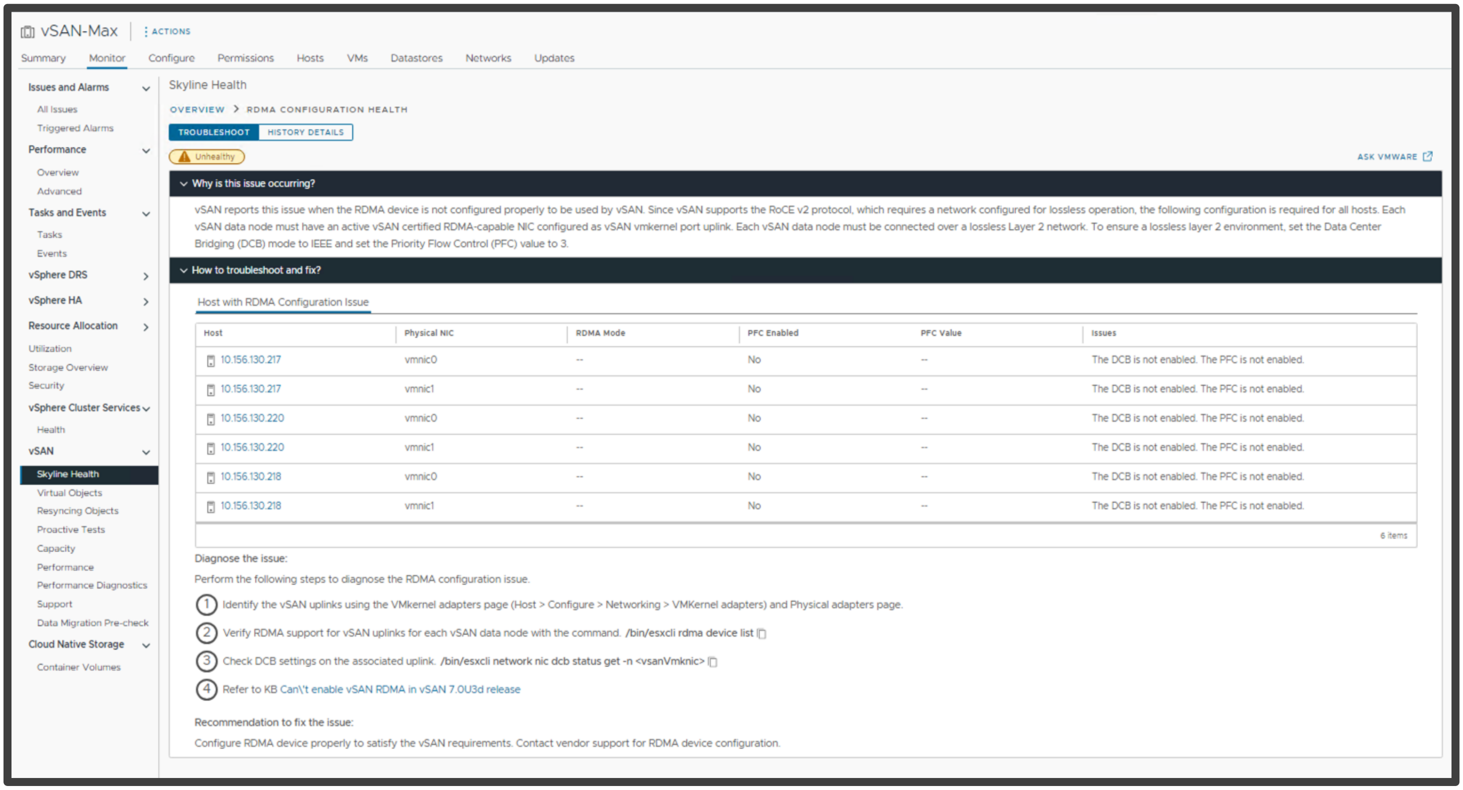
Category: Network

Impact area: Performance

Description:
Check if the vSAN RDMA-enabled physical NIC is configured for lossless traffic. To ensure a lossless layer 2 environment, the Data Center Bridging (DCB) mode must be configured as IEEE and the Priority Flow Control (PFC) value must set to 3.

Risk if no action taken: vSAN cannot use RDMA for internal data communication, and the vSAN cluster switches to TCP communication with degraded performance.

Here we see that PFC is not enabled on the switch:



Skyline Health

OVERVIEW > RDMA CONFIGURATION HEALTH

Why is this issue occurring?

vSAN reports this issue when the RDMA device is not configured properly to be used by vSAN. Since vSAN supports the RoCE v2 protocol, which requires a network configured for lossless operation, the following configuration is required for all hosts. Each vSAN data node must have an active vSAN certified RDMA-capable NIC configured as vSAN vmkernel port uplink. Each vSAN data node must be connected over a lossless Layer 2 network. To ensure a lossless layer 2 environment, set the Data Center Bridging (DCB) mode to IEEE and set the Priority Flow Control (PFC) value to 3.

How to troubleshoot and fix?

Host with RDMA Configuration Issue

Host	Physical NIC	RDMA Mode	PFC Enabled	PFC Value	Issues
10.156.130.217	vmnic0	--	No	--	The DCB is not enabled. The PFC is not enabled.
10.156.130.217	vmnic1	--	No	--	The DCB is not enabled. The PFC is not enabled.
10.156.130.220	vmnic0	--	No	--	The DCB is not enabled. The PFC is not enabled.
10.156.130.220	vmnic1	--	No	--	The DCB is not enabled. The PFC is not enabled.
10.156.130.218	vmnic0	--	No	--	The DCB is not enabled. The PFC is not enabled.
10.156.130.218	vmnic1	--	No	--	The DCB is not enabled. The PFC is not enabled.

Diagnose the issue:

Perform the following steps to diagnose the RDMA configuration issue.

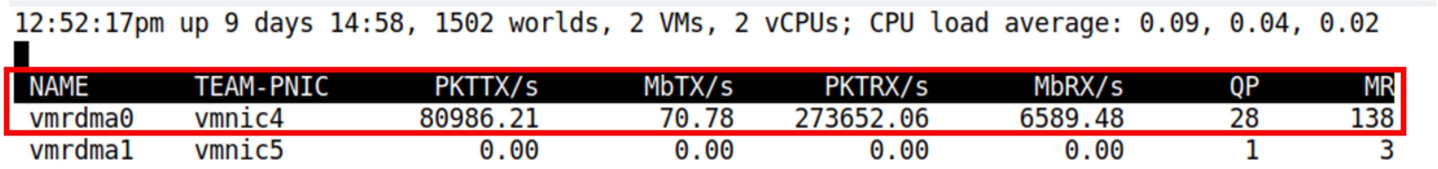
1. Identify the vSAN uplinks using the VMkernel adapters page (Host > Configure > Networking > VMKernel adapters) and Physical adapters page.
2. Verify RDMA support for vSAN uplinks for each vSAN data node with the command. **/bin/esxcli rdma device list**
3. Check DCB settings on the associated uplink. **/bin/esxcli network nic dcb status get -n <vsanVmknic>**
4. Refer to KB Can\'t enable vSAN RDMA in vSAN 7.0U3d release

Recommendation to fix the issue:

Configure RDMA device properly to satisfy the vSAN requirements. Contact vendor support for RDMA device configuration.

Verify virtual RDMA adapter performance with esxtop:

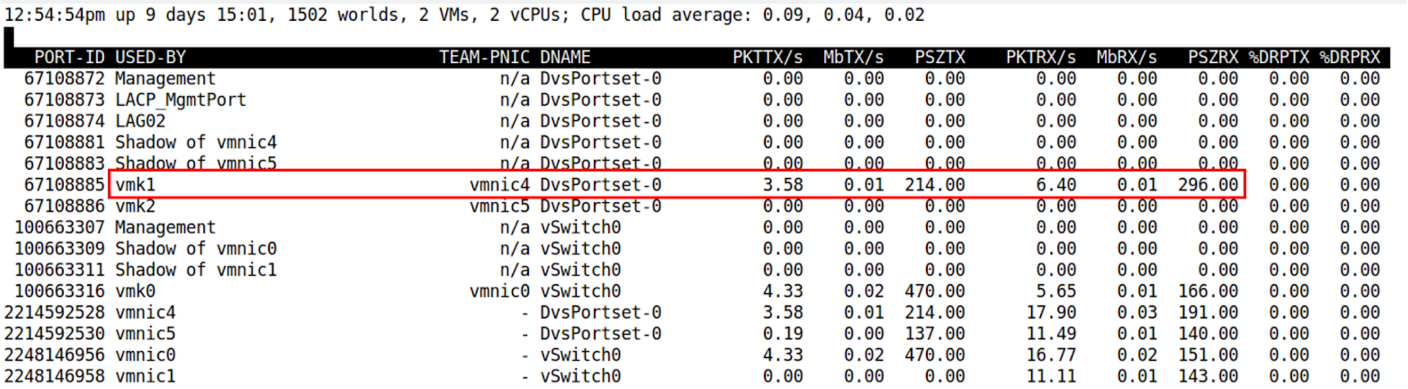
SSH to one of the hosts and launch esxtop. Press ‘r’ to show the RDMA screen:



```
12:52:17pm up 9 days 14:58, 1502 worlds, 2 VMs, 2 vCPUs; CPU load average: 0.09, 0.04, 0.02

 NAME       TEAM-PNIC       PKTTX/s      MbTX/s     PKTRX/s      MbRX/s      QP     MR
 vmrdma0    vmnic4         80986.21       70.78   273652.06     6589.48      28    138
 vmrdma1    vmnic5             0.00        0.00        0.00        0.00       1      3
```

This shows us the throughput for each virtual adapter. Here we see the traffic traversing vmrdma0.

Pressing ‘n’ for the network view shows that there is minimal traffic on the vmkernel adaptor:



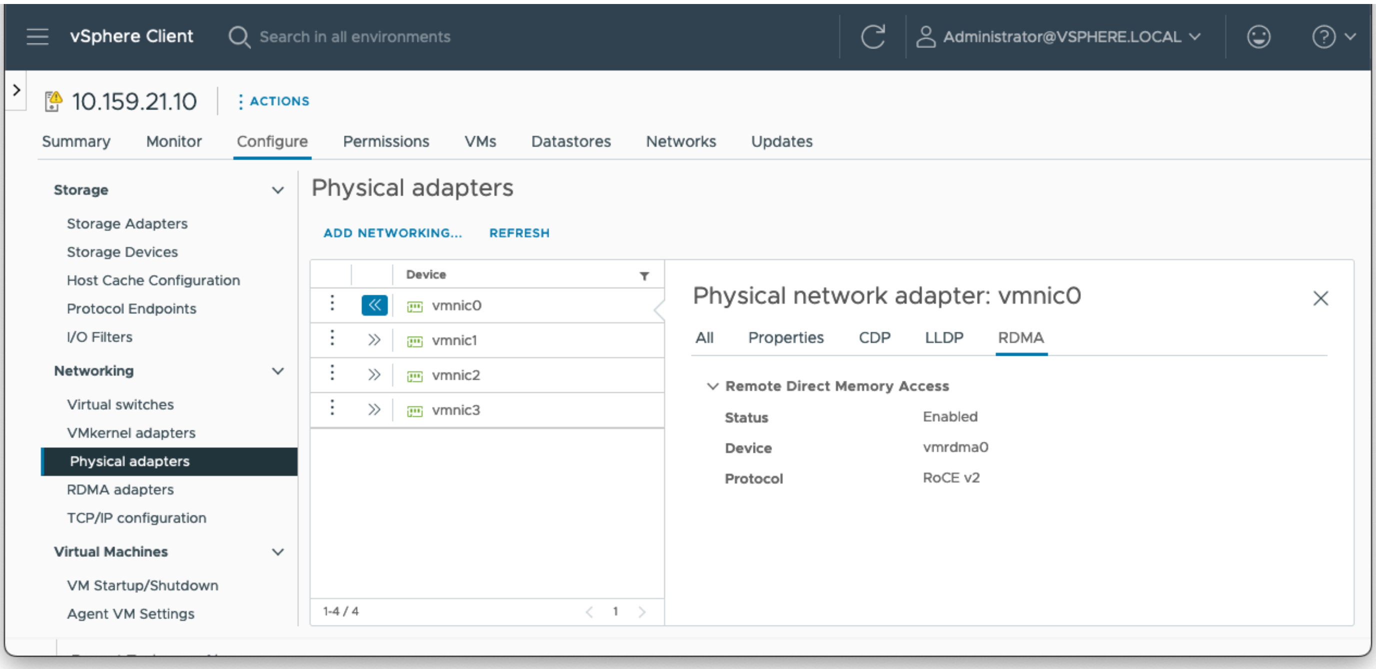
```
12:54:54pm up 9 days 15:01, 1502 worlds, 2 VMs, 2 vCPUs; CPU load average: 0.09, 0.04, 0.02

   PORT-ID USED-BY                            TEAM-PNIC DNAME             PKTTX/s  MbTX/s   PSZTX    PKTRX/s  MbRX/s   PSZRX %DRPTX %DRPRX
  67108872 Management                               n/a DvsPortset-0         0.00    0.00    0.00       0.00    0.00    0.00   0.00   0.00
  67108873 LACP_MgmtPort                            n/a DvsPortset-0         0.00    0.00    0.00       0.00    0.00    0.00   0.00   0.00
  67108874 LAG02                                    n/a DvsPortset-0         0.00    0.00    0.00       0.00    0.00    0.00   0.00   0.00
  67108881 Shadow of vmnic4                         n/a DvsPortset-0         0.00    0.00    0.00       0.00    0.00    0.00   0.00   0.00
  67108883 Shadow of vmnic5                         n/a DvsPortset-0         0.00    0.00    0.00       0.00    0.00    0.00   0.00   0.00
  67108885 vmk1                                  vmnic4 DvsPortset-0         3.58    0.01  214.00       6.40    0.01  296.00   0.00   0.00
  67108886 vmk2                                  vmnic5 DvsPortset-0         0.00    0.00    0.00       0.00    0.00    0.00   0.00   0.00
 100663307 Management                               n/a vSwitch0             0.00    0.00    0.00       0.00    0.00    0.00   0.00   0.00
 100663309 Shadow of vmnic0                         n/a vSwitch0             0.00    0.00    0.00       0.00    0.00    0.00   0.00   0.00
 100663311 Shadow of vmnic1                         n/a vSwitch0             0.00    0.00    0.00       0.00    0.00    0.00   0.00   0.00
 100663316 vmk0                                  vmnic0 vSwitch0             4.33    0.02  470.00       5.65    0.01  166.00   0.00   0.00
2214592528 vmnic4                                     - DvsPortset-0         3.58    0.01  214.00      17.90    0.03  191.00   0.00   0.00
2214592530 vmnic5                                     - DvsPortset-0         0.19    0.00  137.00      11.49    0.01  140.00   0.00   0.00
2248146956 vmnic0                                     - vSwitch0             4.33    0.02  470.00      16.77    0.02  151.00   0.00   0.00
2248146958 vmnic1                                     - vSwitch0             0.00    0.00    0.00      11.11    0.01  143.00   0.00   0.00
```

Verify functionality of RDMA and TCP/IP on the same physical vmnic:
In the setup below we can verify that RDMA transport layer is not used for standard TCP/IP protocol and handled separately on the vmnic card layer:

- Enable vSAN RDMA
- Prepare DVS / vSwitch portgroup for VMs using the RDMA adapter
- Configure two VMs with the iperf/iperf3 package installed
- Place both VMs on different hosts
- Run one VM as iperf server (`iperf3 -s`)
- Run the second VM as client (`iperf3 -H <IP address of server>`)
- On each host, run esxtop and look at the difference between the network (‘n’) and RDMA (‘r’) screens during the iperf3 test

RDMA troubleshooting

First, verify that the physical network adapters support RDMA. On each host, navigate to **Configure > Networking > Physical adapters > [adapter] > RDMA**:



Physical adapters

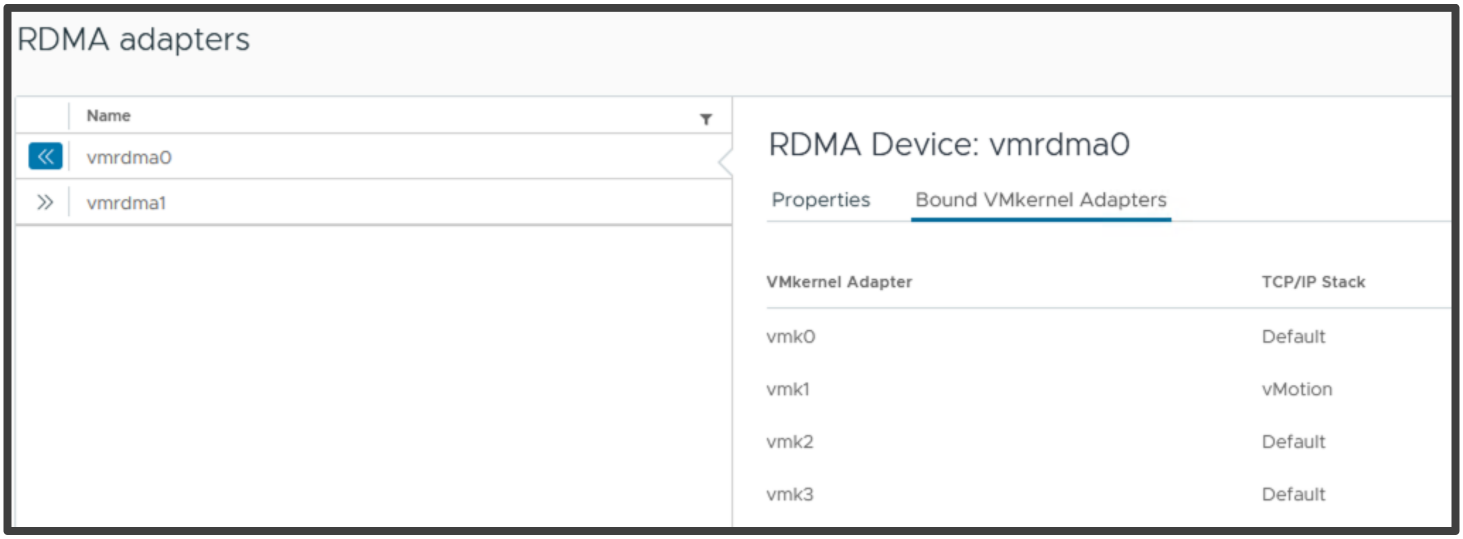
Physical network adapter: vmnic0

Remote Direct Memory Access

Status	Enabled
Device	vmrdma0
Protocol	RoCE v2

Verify if RDMA adapters are bound to VMkernel interface

On each host, navigate to **Configure > Networking > RDMA adapters > [vmrdma#] > Bound VMkernel Adapter**:



RDMA adapters

RDMA Device: vmrdma0

VMkernel Adapter	TCP/IP Stack
vmk0	Default
vmk1	vMotion
vmk2	Default
vmk3	Default

Note: The RDMA support flag is required for the final setup to enable vSAN RDMA transport (if the RDMA flag is not visible, then double check the hardware specification of the adapter along with driver/firmware versions and the vSphere HCL).

vSphere vSAN RDMA uses RoCEV2 as its protocol layer. When there is no RDMA support available on the physical link or setup, communication falls back to standard legacy TCP/IP automatically.

Esxtop provides additional fields for enablement through the ‘f’ key:

```
* A:  NAME =  Name of device
B:  DRIVER =  driver
C:  STATE =    State
* D:  TEAM-PNIC = Team Uplink Physical NIC Name
* E:  PKTTX/s =  Packets Tx/s
* F:  MbTX/s =  Megabits Tx/s
* G:  PKTRX/s =  Packets Rx/s
* H:  MbRX/s =  Megabits Rx/s
I:  %PKTDTX =  % Packets Dropped (Tx)
J:  %PKTDRX =  % Packets Dropped (Rx)
* K:  QP =  Number of Queue Pairs Allocated
L:  CQ =  Number of Completion Queue Pairs Allocated
M:  SRQ =  Number of Shared Receive Queues Allocated
* N:  MR =  Memory Regions Allocated
```

Toggle fields with a-n, any other key to return

Default setup enables only the minimum requirement for performance for MB/s, queue pairs (QP) and allocated memory regions verbs (MR). For in-depth RDMA functionality, please contact your hardware vendor.

Run the following to obtain detailed adapter statistics. Check for any errors here. Queue pairs are adjusted automatically by requirement:

```
[root@localhost:~] esxcli rdma device stats get -d vmrdma0
   Packets received: 1576258135
   Packets sent: 899769661
   Bytes received: 40653333761546
   Bytes sent: 1079424621290
   Error packets received: 0
   Error packets sent: 0
   Error length packets received: 0
```

APPENDIX B: Cleanly Removing vSAN Configuration

vCLS Retreat Mode

On occasion, it may become necessary to remove a vSAN cluster and reset hosts to a 'clean' state.

To expedite the process, it is advisable to first put vCLS into retreat mode. This will delete the vCLS VMs and make it easier to remove the vSAN datastore and put hosts into maintenance mode, etc.

To achieve this, an vCenter advanced setting, '`config.vcls.clusters.[domain].enabled`' needs to be set.

The procedure to do this is detailed in the documentation here: https://docs.vmware.com/en/VMware-vSphere/7.0/com.vmware.vsphere.resmgmt.doc/GUID-F98C3C93-875D-4570-852B-37A38878CE0F.html

To make this easier a script is available here to use (download to a Linux or Mac host, uses govc): https://github.com/vmware-tanzu-experiments/vsphere-with-tanzu-proof-of-concept-samples/blob/main/VCF/vCLS.sh

Remove vSAN Partitions and Clear Data

The next step is to turn off vSAN from vCenter, under [cluster] > Configure > Services > vSAN. If for some reason this step encounters errors, the method below may be useful.

First, open an SSH session to all hosts in the cluster and list the disks used by vSAN by using the command:

```
vdq -iH
```

The next step depends on the type of cluster

OSA Clusters

Remove the cache device from each disk group, using the command:

```
esxcli vsan storage remove -s [cache device]
```

ESA Clusters

Remove disks from the storage pool, using the command:

```
esxcli vsan storagepool remove -d [device]
```

Next, relabel the disks:

```
partedUtil mklabel /vmfs/devices/disks/[disk] gpt
```

Again, to make this easier, a script is available to help with this:

OSA: https://github.com/vmware-tanzu-experiments/vsphere-with-tanzu-proof-of-concept-samples/blob/main/VCF/vsan-remove-esa.sh

ESA: https://github.com/vmware-tanzu-experiments/vsphere-with-tanzu-proof-of-concept-samples/blob/main/VCF/vsan-remove-esa.sh