



What's New in vSphere 7 Core Storage

VMware Storage

Table of contents

What's New in vSphere 7 Core Storage	4
vSphere 7.0 U3	4
NVMe/TCP	4
Host Scale support increase for NFS and VMFS	6
Affinity 3.0 Improvements for CNS	7
vVols Batch Snapshots	7
vSphere 7.0 U2	9
iSCSI Path Limit increase	9
RDM support for RHEL HA	9
VMFS SESparse Snapshot Improvements	9
Multiple Paravirtual RDMA (PVRDMA) adapter support	9
Performance Improvements on VMFS	9
NFS Improvements	10
HPP Fast Path Support for Fabric Devices	10
HPP as the default plugin for vSAN	10
VOMA improvements	10
vVols	10
vVols Enhancements and Updates	10
Support for Higher Queue Depth with vVols Protocol Endpoints	10
Create larger than 4GB Config vVol	11
vVols with CNS and Tanzu	11
vVols with CNS and Tanzu	11
SPBM Multiple Snapshot Rule Enhancements	11
32 Snapshot support for Cloud Native Storage (CNS) for First Class Disks	11
CNS PV to vVol mapping	11
vSphere 7.0 U1	12
Features and enhancements for vSphere 7.0 Update 1	12
VMFS Enhancements	12
SESparse Snapshot consolidation bloat reduction.	12
Reduced vSphere Snapshot stun time.	12
NFS Enhancements	12
NAS VAAI Plugins	12
NVDK cloning of LZT disk could fail with unsupported disk type. We have updated the code to support all disk types.	12
RDM	12

- pRDM extension with Microsoft WSFC 12
- NVMeoF 12
- NVMeoF Support for Oracle RAC 12
- vVols as Principal Storage in VCF 4.1 12
- Support for Virtual Volumes (vVols) as principal storage in VMware Cloud Foundation 4.1 13
- vSphere 7.0 18
- Core Storage features and enhancements included in vSphere 7.0 18
 - NVMeoF 18
 - NVMeoF Insight 18
 - VMware NVMeoF 18
 - Clustered VMDK 20
 - Shared VMDK or Clustered VMDK in VMFS6 20
 - Demo and details on migrating WSFC using RDMS to Shared VMDK on VMFS 21
 - Enabling Clustered VMDK 21
 - Enabling Shared VMDK support on VMFS Datastores 22
 - Perennially Reserved Flag for RDMS 22
 - Using the Perennially Reserved Flag for WSFC RDMS 22
 - Affinity 2.0 25
 - VMFS Affinity Manager 2.0 25
 - New Affinity 2.0 Manager 26
 - End Of Availability (EOA) vFRC and CBRC 1.0 27
 - vSphere Flash Read Cache (vFRC) EOL 27
 - Content-Based Read Cache (CBRC) 1.0 EOA 27
 - vVols Interoperability 27
 - vVols Interoperability with VMware Products 27
 - vVols Support in SRM 8.3 27
 - Site Recovery Manager - SRM 27
 - vVols Support for CNS 29
 - Cloud-Native Storage - CNS 29
 - vVols Support in vRealize Operations 8.1 29
 - vRealize Operations - vROps 29
 - vVols as Supplemental Storage in VCF 30
 - VMware Cloud Foundation - VCF 30

What's New in vSphere 7 Core Storage

vSphere 7.0 U3

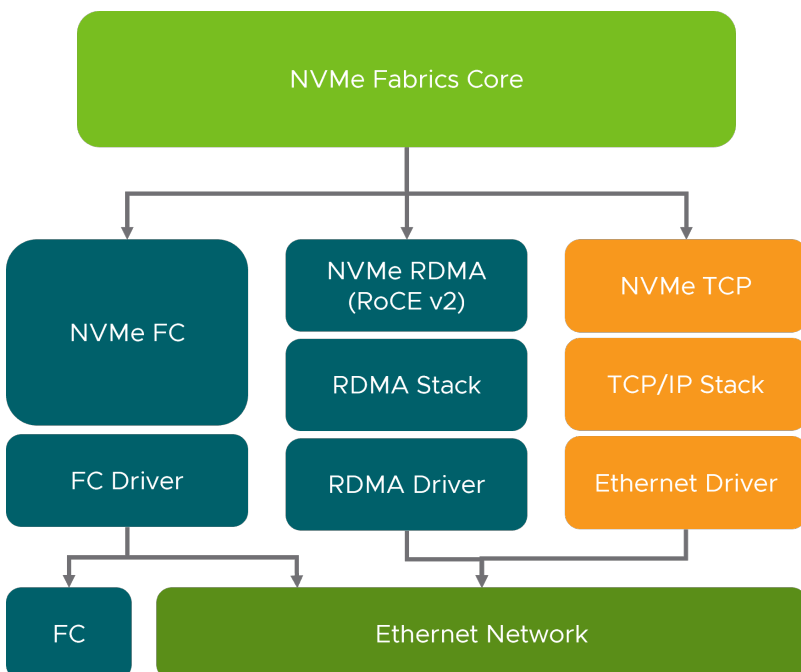
Core Storage features and enhancements for vSphere 7 Update 3.

NVMe/TCP



NVMe-oF TCP/IP Support

NVMe over fabric extends NVMe from local storage to shared network storage. With the release of vSphere 7, the supported protocols for NVMeoF were FC and RDMA. With the release of vSphere 7 U3, we are adding support for TCP. One of the benefits of NVMe/TCP is there is no need for specialized HBAs or RNICs (RDMA NIC) for connectivity. Subsequently, standard Ethernet networks and HW may be used. Of course, having the necessary bandwidth for the additional overhead is imperative. With the ability to use standard Ethernet HW, the cost of entry for NVMeo TCP/IP is less than with FC and RDMA. This allows many existing customers access to new storage technologies like NVMeoF.



If you're wondering which protocol performs best, a better question is what does your application requires. Also, what is the

current configuration or your environment? There can be additional costs for FC or RDMA, where TCP can usually use existing HW. In general, RDMA will be the highest performing, then FC, and finally TCP. If you need the absolutely lowest latency and highest performance, RDMA is the way to go. If you're interested in NVMeoF and without adding additional costs, TCP maybe your best option. Of course if you already have an FC environment, that would be the most obvious choice.

The screenshot displays the vSphere configuration interface for Storage Adapters. The 'Configure' tab is selected. The left-hand navigation pane shows 'Storage Adapters' under the 'Storage' category. The main content area shows a table of installed storage adapters. The 'vmhba67' adapter is selected, and its properties are shown below the table.

Adapter	Model	Type
vmhba05	VMware NVMe over RDMA Storage Adapt...	RDMA
vmhba66	VMware NVMe over RDMA Storage Adapt...	RDMA
vmhba67	VMware NVMe over TCP Storage Adapter	NVMe over TCP
vmhba68	VMware NVMe over TCP Storage Adapter	NVMe over TCP

Properties	Devices	Paths	Namespaces	Controllers
General				
Name	vmhba67			
Model	VMware NVMe over TCP Storage Adapter			

With the addition of NVMe/TCP we have also added storage vmknix tagging in NIOC for NVMe/TCP. This allows customers to tune network resources for the NVMe/TCP to ensure enough bandwidth is allocated. When NIOC is enabled, distributed switch traffic is divided into the following predefined network resource pools:

- Fault Tolerance traffic
- iSCSI traffic
- vMotion traffic
- management traffic
- vSphere Replication (VR) traffic
- NFS traffic
- virtual machine traffic
- vSphere Data Protection traffic
- backup NFC traffic
- **Newly added:** NVMe/TCP traffic

DVS-NVMe | ACTIONS

Summary Monitor **Configure** Permissions Ports Hosts VMs Networks

Settings >

Resource Allocation v

System traffic

Network resource pools

Alarm Definitions

0 Gbit/s 18.75 Gbit/s 25.00 Gbit/s

Network I/O Control Enabled
Version 3
Physical network adapters 4
Minimum link speed 25 Gbit/s

Total bandwidth capacity	25.00 Gbit/s
Maximum reservation allowed ⓘ	18.75 Gbit/s
<input checked="" type="checkbox"/> Configured reservation	0.00 Gbit/s
<input type="checkbox"/> Available bandwidth	25.00 Gbit/s

EDIT

Traffic Type	Shares	Shares Value	Reservation	Limit
Management Traffic	Normal	50	0 Mbit/s	Unlimited
Fault Tolerance (FT) Traffic	Normal	50	0 Mbit/s	Unlimited
vMotion Traffic	Normal	50	0 Mbit/s	Unlimited
Virtual Machine Traffic	High	100	0 Mbit/s	Unlimited
iSCSI Traffic	Normal	50	0 Mbit/s	Unlimited
NFS Traffic	Normal	50	0 Mbit/s	Unlimited
vSphere Replication (VR) Traffic	Normal	50	0 Mbit/s	Unlimited
vSAN Traffic	Normal	50	0 Mbit/s	Unlimited
vSphere Data Protection Backup Traffic	Normal	50	0 Mbit/s	Unlimited
vSphere Backup NFC Traffic	Normal	50	0 Mbit/s	Unlimited
NVMe over TCP	Normal	50	0 Mbit/s	Unlimited

Host Scale support increase for NFS and VMFS



VMFS-6 & NFS Scale to 128 Nodes

Many larger enterprises, service providers, and cloud deployments often reach the vSphere limit of 64 hosts per VMFS or NFS datastore. With the release of update 3, we have increased the number of hosts that may connect to a VMFS-6 or NFS datastore

from 64 to **128**. This will alleviate the need for special approval for a larger number of hosts accessing VMFS or NFS datastores.

Note: This is not a hosts-per-cluster increase; this is a number of hosts that can access a single VMFS or NFS datastore.

Affinity 3.0 Improvements for CNS



Affinity 3.0 Enhancement for CNS

In vSphere 7, VMware updated the Affinity Manager, which handles first writes with thin or lazy thick provision. The new [Affinity Manager, 2.0](#), maintains a map of all free storage Resource Cluster. Resource Clusters are available space for new writes, which enables quicker first writes.

In U3, we add additional enhancements to Affinity 3.0 which now adds support for CNS persistent volumes or FCDs (First Class Disks). We have also added support for the higher number of vSphere hosts per cluster.

vVols Batch Snapshots



Batch Snapshots on vVols

With the potential scale vVols offers, ensuring operational efficiency is key. As engineering continues to enhance and develop vVols, we have enhanced the procedure for processing large numbers of vVol snapshots by making snapshot operations into a batch process. By grouping large amounts of snapshot operations, we reduce the serialize actions used for snapshots making the process more efficient and reducing the effect on the VMs and storage environment.

vSphere 7.0 U2

Core Storage features and enhancements for vSphere 7 Update 2.

iSCSI Path Limit increase

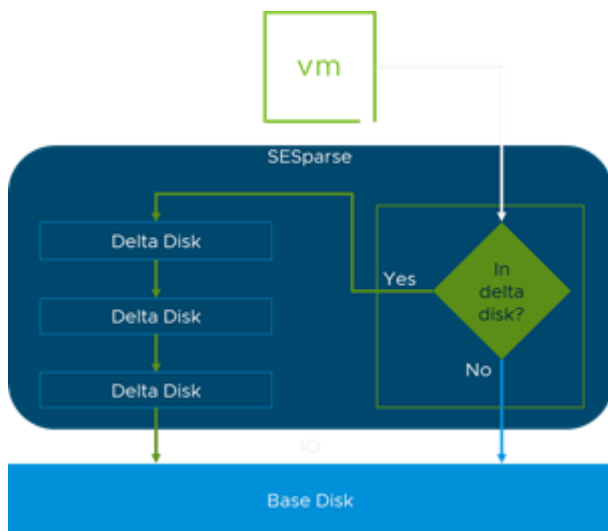
One of the enhancements from the [vSphere 7 Update 2](#) release I'm sure many customers will be thrilled about is the iSCSI path limit increase. Until this release, the iSCSI path limit was 8 paths per LUN, and many customers end up going over this. Whether it's from multiple VMKernels or targets, customers often ended up with 16 or 24 paths. I'm excited to announce that with the vSphere 7.0 U2, the new iSCSI path limit is now 32 paths per LUN.

RDM support for RHEL HA

There were a few changes that needed to be made to enable support for Red Hat Enterprise HA to be able to use RDMs in vSphere. With the release of vSphere 7.0 U2, RHEL HA is now supported on RDMs.

VMFS SESparse Snapshot Improvements

Read performance improvements by using a technique for directing the reads to where the data resides rather than traversing the delta disk snapshot chain every time. Previously, if a read came into the Virtual Machine and the VM had snapshots, the reads would traverse the snapshot chain then to the base disk. Now, when a read comes in a filter will direct the read to either the snapshot chain or base disk reducing the read latency.



Multiple Paravirtual RDMA (PVRDMA) adapter support

In vSphere 6.7, we announced support for RDMA in vSphere. One of the limitations was only a single PVRDMA adapter was supported per Virtual Machine. With the release of vSphere 7.0 U2, we now support multiple PVRDMA adapters per VM.

Performance Improvements on VMFS

With the release of vSphere 7.0 U2, we have made performance improvements to VMFS. The performance was improved for first writes on thin disks. These changes improve performance for backup and restore, copy operations, and Storage vMotion in certain instances. With this improvement and the enhancements with [Affinity 2.0 in vSphere 7](#), the first write impact has further been reduced. These improvements help to reduce the potential effects of first writes when using thin-provisioned disks.

NFS Improvements

NFS required a clone to be created first for a newly created VM and the subsequent snapshots would be offloaded to the array. With the release of vSphere 7.0 U2, we have enabled NFS array snapshots of full, non-cloned VMs to not use redo logs but instead, use the snapshot technology of the NFS array in order to provide better snapshot performance. The improvement here will remove the requirement/limitation of creating a clone and enables the first snapshot also to be offloaded to the array.

HPP Fast Path Support for Fabric Devices

The High Performance Plugin (HPP) is a new Multi-Pathing Plugin (MPP) for ESXi that VMware has developed for very fast devices. HPP is a leaner MPP than NMP, but achieves some of this by dropping support for sub-plugins like Storage Array Type Plugins (SATPs) and Path Selection Plugins (PSPs). HPP started shipping with ESXi 6.7 but was not the default plugin for any devices and only supported single-pathed local devices. This made it relevant for special NVMe-PCIe based use cases where it could reach much higher IOPS than the ESXi Native Multi-Pathing Plugin (NMP).

With the release of vSphere 7.0 U2, HPP is now the default plugin for NVMe devices. The plugin comes with 2 options – SlowPath with legacy behavior, VM fairness capabilities, and the newly added FastPath, designed to provide better performance as compared to SlowPath with some restrictions. Even in SlowPath mode HPP can often perform better than NMP for the same device due to IOs being handled in batch mode by helping to reduce lock contention and CPU overhead in the IO path. There are some limitations to when FastPath will apply, so it is mostly intended for limited use cases. The FastPath is enabled by setting a Latency Sensitive Threshold, which is the threshold below which we allow operation of FastPath. Once the device latency goes above the threshold we will move to SlowPath and thus ensure that fairness is respected when latency has a higher impact.

To learn see VMware Docs article [Set Latency Sensitive Threshold](#).

HPP as the default plugin for vSAN

With the release of vSphere 7.0 U2, HPP is now the default MPP for all devices (SAS/SATA/NVMe) used with vSAN. Note that HPP is also the default plugin for NVMe fabric devices. This is an infrastructure improvement to ensure vSAN uses the improved storage plugin and can take advantage of the improvements.

VOMA improvements

vSphere On-disk Metadata Analyzer (VOMA) is used to identify and fix metadata corruption affecting the file system or underlying logical volumes. With the release of vSphere 7.0 U2, VOMA support has now been enabled for spanned VMFS volumes.

For more information on VOMA,

VMware Docs article [Checking Metadata Consistency with VOMA](#)

VMware KB article [Using vSphere On-disk Metadata Analyzer \(VOMA\) to check VMFS metadata consistency \(2036767\)](#)

vVols

vVols Enhancements and Updates

Support for Higher Queue Depth with vVols Protocol Endpoints

In some cases, the Disk.SchedNumReqOutstanding (DSNRO) configuration parameter did not match the queue depth of the vVols Protocol Endpoint (PE) (VVolPESNRO). With the release of vSphere 7.0 U2, the default QD for the PE will now be 256 or the maxQueueDepth of the exposed LUN. Subsequently, the default minimum PE QD is now 256.

VMware KB article: [Changing the queue depth for QLogic, Emulex, and Brocade HBAs \(1267\)](#)

Create larger than 4GB Config vVol

This allows the Config vVol to be larger than the default 4GB for partners to be able to store images for automatic builds. Some of our partners needed to be able to store images and build files in the Config vVol which was previously limited to 4GB. Now the Config vVol can be increased similar to the Data vVol.

vVols with CNS and Tanzu

vVols with CNS and Tanzu

SPBM Multiple Snapshot Rule Enhancements

With vVols, Storage Policy-Based Management (SPBM) gives the VI admin autonomy to manage storage capabilities, at a VM level, via policy. With the release of vSphere 7.0 U2, we have enabled our vVols partners to support multiple snapshot rules in a single SPBM storage policy. This feature will need to be supported in the respective VASA providers that enable snapshot policies to be constructed. When supported by our vVols partners, it will be possible to have a single policy with multiple rules with different snapshot intervals.

32 Snapshot support for Cloud Native Storage (CNS) for First Class Disks

Persistent Volumes (PV) are created in vSphere as First-Class Disks (FCD). FCDs are independent disks with no VM attached. With the release of vSphere 7.0 U2, we are adding snapshot support of up to 32 snapshots for FCDs. This enables you to create snapshots of your K8s PVs which goes along with the SPBM multiple snapshot rules.

CNS PV to vVol mapping

In some cases, customers may want to see which vVols is associated with which CNS Persistent Volume (PV). With the release of vSphere 7.0 U2 in the CNS UI, you can now see a mapping of the PV to its corresponding vVol FCD.

The screenshot shows the vSphere CNS interface for 'Container providers: Kubernetes'. It features a table of vVols and a detailed view for the selected volume 'blkvol-927968-11-587'. The table lists various vVol names, and the details pane shows metadata such as Type (Block), Volume ID, Volume Path, and Volume Backing Object ID. The Volume Path and Volume Backing Object ID are highlighted with a red box.

Volume Name
blkvol-927968-10-120
blkvol-927968-2-511
blkvol-927968-10-444
blkvol-696655-5-1
blkvol-927968-11-587
blkvol-822416-8-24
blkvol-927968-1-376
pvmongo-2

blkvol-927968-11-587

Basics | Kubernetes objects | Related Pages

Type: Block

Volume ID: fd726e00-180f-11e8-8fa1-0050568e3cc9

Volume Path: 1234567/aineirgnerignhiulwhgbog/wehfowheofwebf123erlg/dcdc.vmdk

Volume Backing Object ID: abs-wer223nti4-1235

Datastore: vsanDatastore

Storage policy: Default vSAN policy

Compliance status: ✔ compliant

Accessibility: ✔ Accessible

vSphere 7.0 U1

Features and enhancements for vSphere 7.0 Update 1

VMFS Enhancements

- **SESparse Snapshot consolidation bloat reduction.**
 - We have optimized the SESparse snapshot process reducing bloat. When using thin VMDKs, there can be an increase in disk usage when consolidating vSphere snapshots and unmapping the deleted data. We have enhanced the process by optimizing the consolidation process.
- **Reduced vSphere Snapshot stun time.**
 - We have optimized the snapshot process reducing to help reduce stun time during snapshot creation and deletion. By updating the way the [Affinity Manager](#) updates Resource Clusters (RC), we have reduced snapshot creation and deletion stun times. We have also enhanced the reporting of the snapshot consolidation progress.

NFS Enhancements

- **NAS VAAI Plugins**
 - Previously, the installation of NAS VAAI plugins required a reboot of the ESXi host. In vSphere 7 Update 1, we have enabled the ability to install VAAI NAS plugins without requiring a reboot.
- **NVDM cloning of LZT disk could fail with unsupported disk type. We have updated the code to support all disk types.**

RDM

- **pRDM extension with Microsoft WSFC**
 - One function RDM can currently provide over other shared disk applications is being able to hot extend shared disks. In vSphere 7 Update 1, we have validated the support for online disk/LUN expansion for pass-through RDM used with Windows Server Failover Clustering (WSFC).

NVMeoF

NVMeoF Support for Oracle RAC

There are numerous customers using clustered applications; Oracle RAC for example. As NVMeoF continues to gain support, especially for database instances, we want to ensure we validate the various deployments.

- With vSphere 7 Update 1, we have extended support for Oracle RAC using NVMeoF targets.

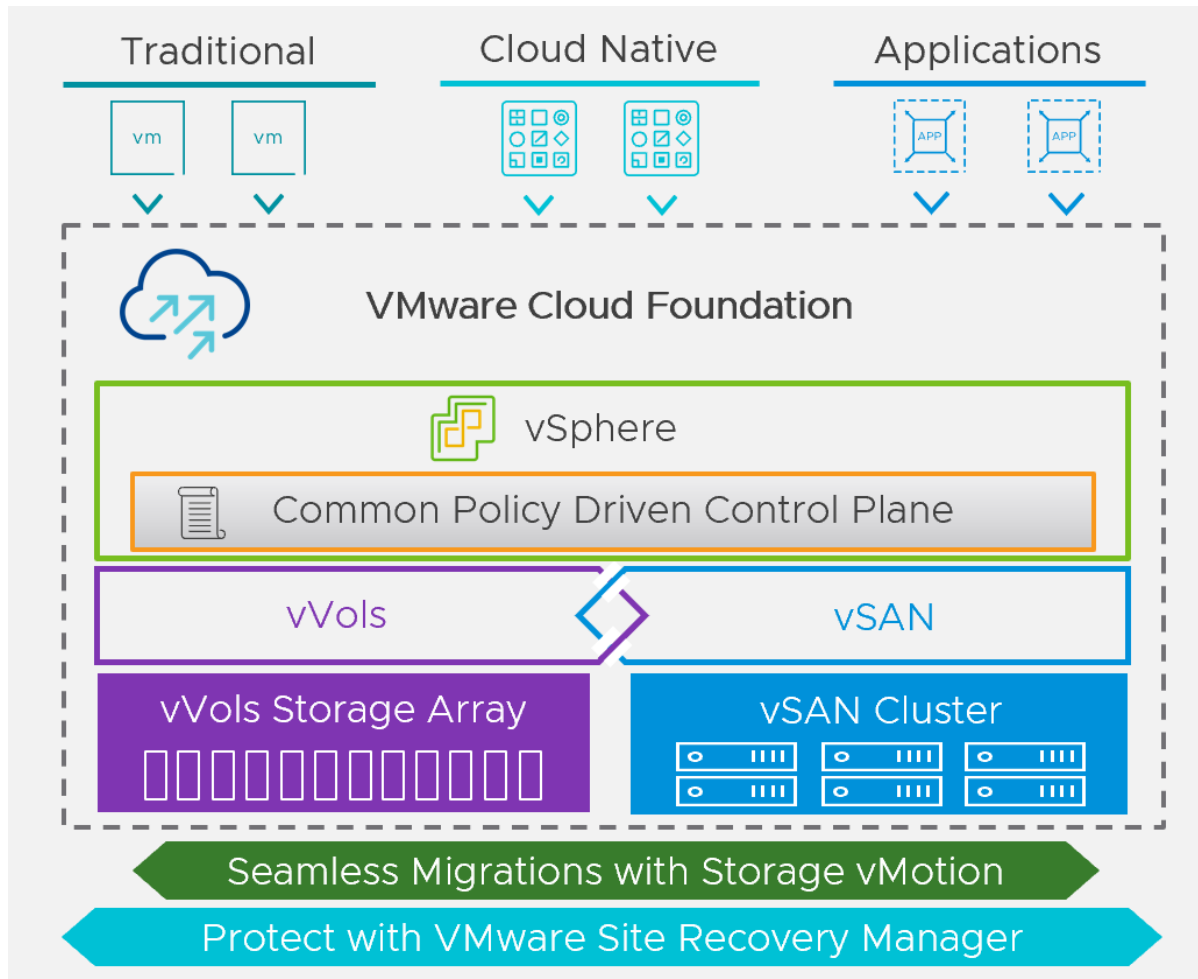
vVols as Principal Storage in VCF 4.1

Support for Virtual Volumes (vVols) as principal storage in VMware Cloud Foundation 4.1

With the release of vSphere 7.0 Update 1, there were also updates to Cloud Foundation and Tanzu.

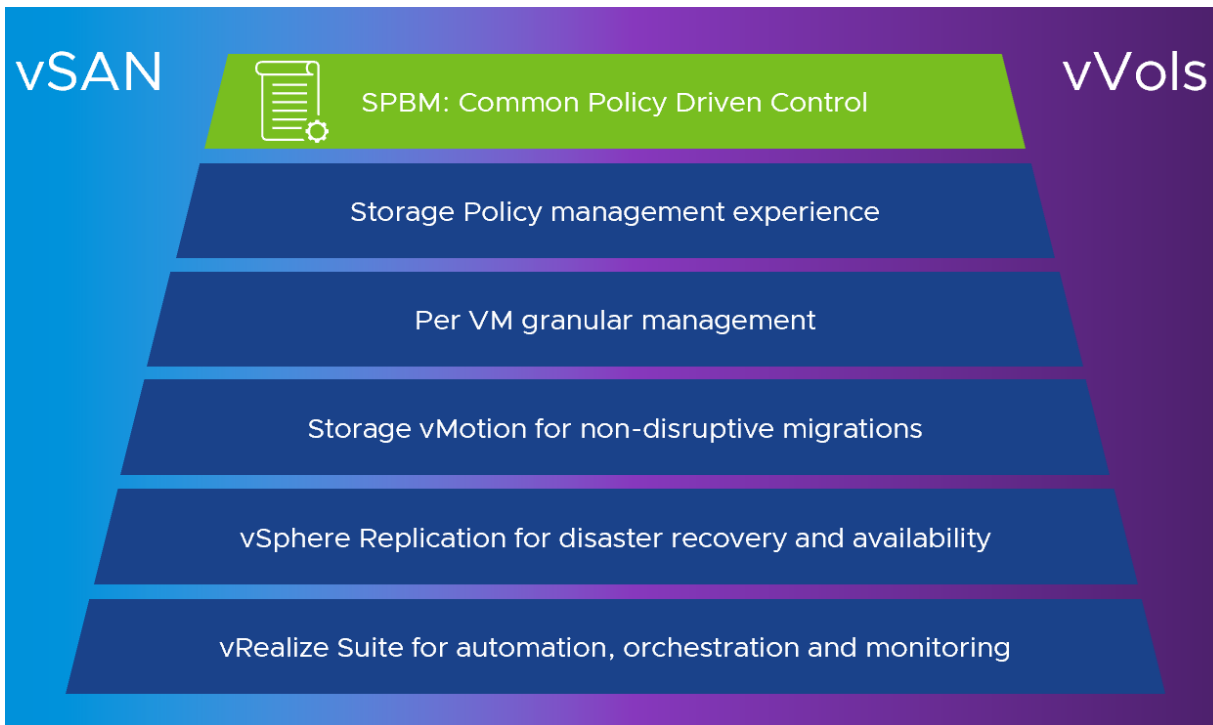
VMware Cloud Foundation 4.1

- In VCF 4.1, we have added support for vVols as principal storage in workload domains.



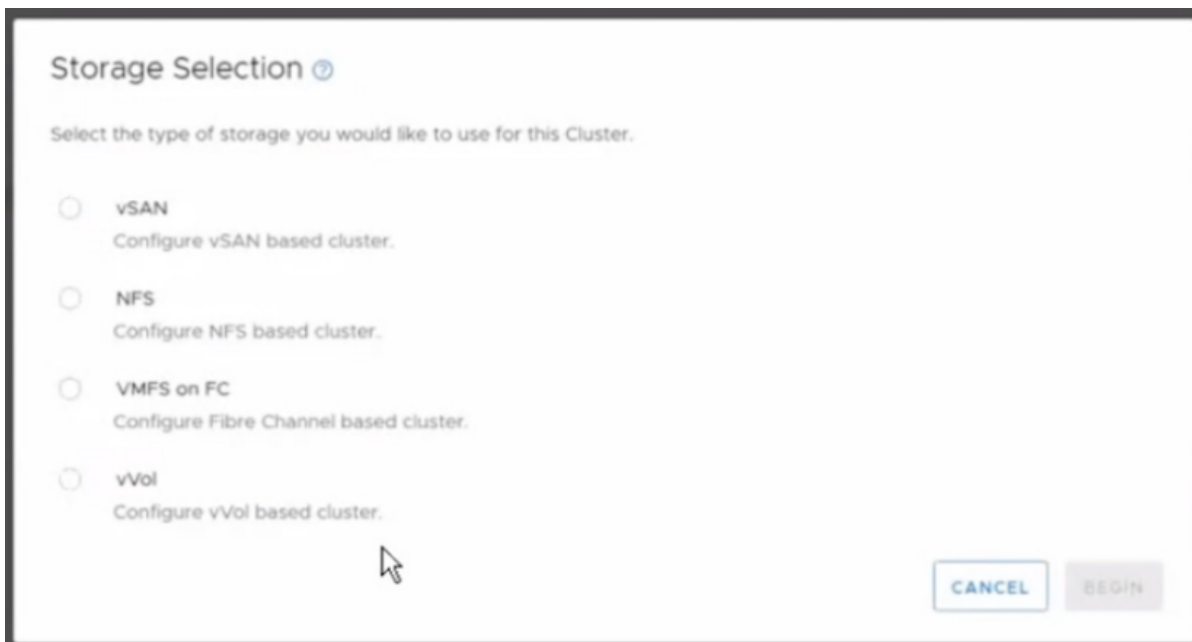
With VMware Cloud Foundation (VCF), your management domain requires vSAN, which can easily be managed using policy-based management or SPBM. SPBM allows simplified operational management of your storage capabilities. Although you can use tag-based policies with external storage for VCF, it is not something that scales easily and requires quite a bit of manual operations. When you think about the possible scale VCF enables, manually tagging datastores can become daunting. Subsequently, being able to programmatically manage all your VCF storage simplifies your operations, freeing valuable time for other tasks.

In VCF 4.0 we supported vSAN, NFS, and VMFS on FC for Principal Storage for newly created Workload Domain clusters. VCF 4.1 expands principal storage options by adding support for vVols. vVols was previously supported for Supplemental Storage only. The big difference between the two is principal storage is the initial storage option selected when creating new clusters in VCF, the setup of principal storage is automated through VCF workflows. Supplemental storage is added to a cluster manually through the vSphere Web UI after it has been created. With vVols, numerous benefits may now be utilized in VCF and, you can use the same SPBM management plane for your vSAN and external arrays. vVols enables you to use all of your array's capabilities such as array-based snapshots, cloning, and replication. All managed via policies, on a single vVols datastore, with VM granularity.

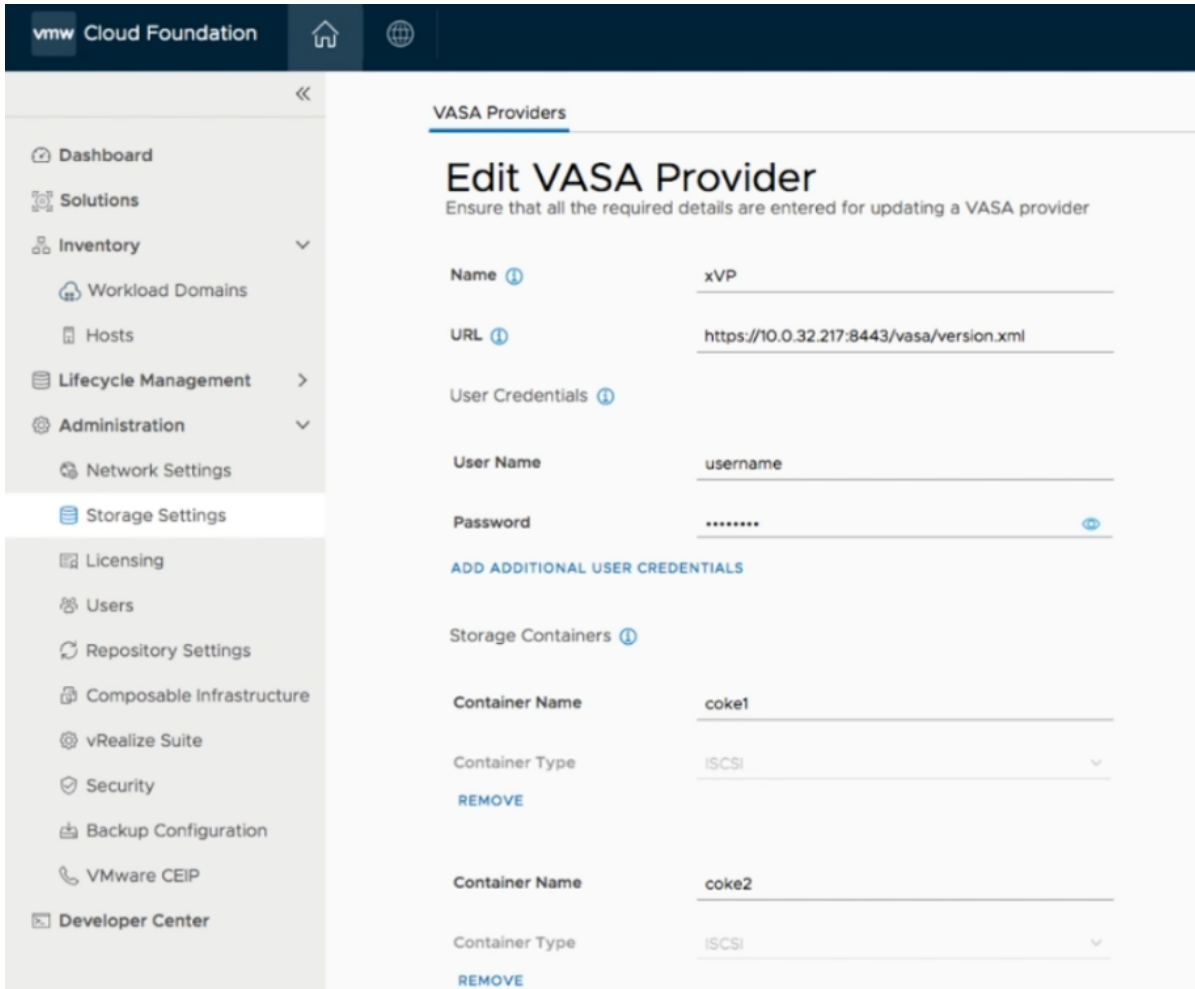


Setting up vVols as principal storage in VCF 4.1

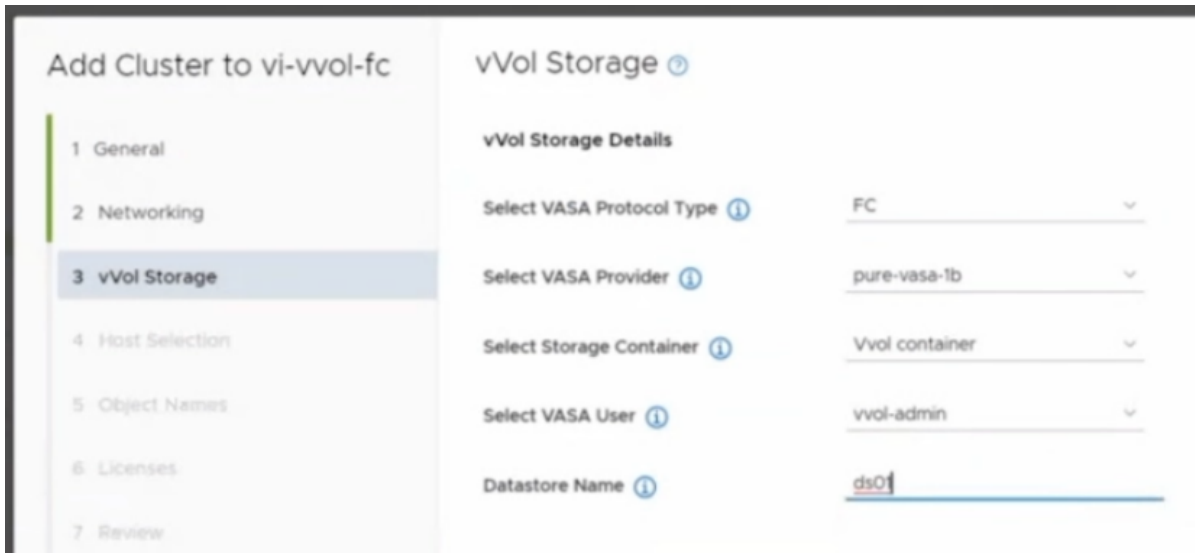
The VCF engineering team has been diligently working internally and with our storage partners to enable vVols as principal storage. With the 4.1 release, we support NFS 3.x, FC, and limited iSCSI protocols for vVols. For iSCSI, there are a few pre-tasks that must be completed. Setting up your SW iSCSI initiator on all your hosts in the new WLD, and your VASA must be listed as a Dynamic Target.



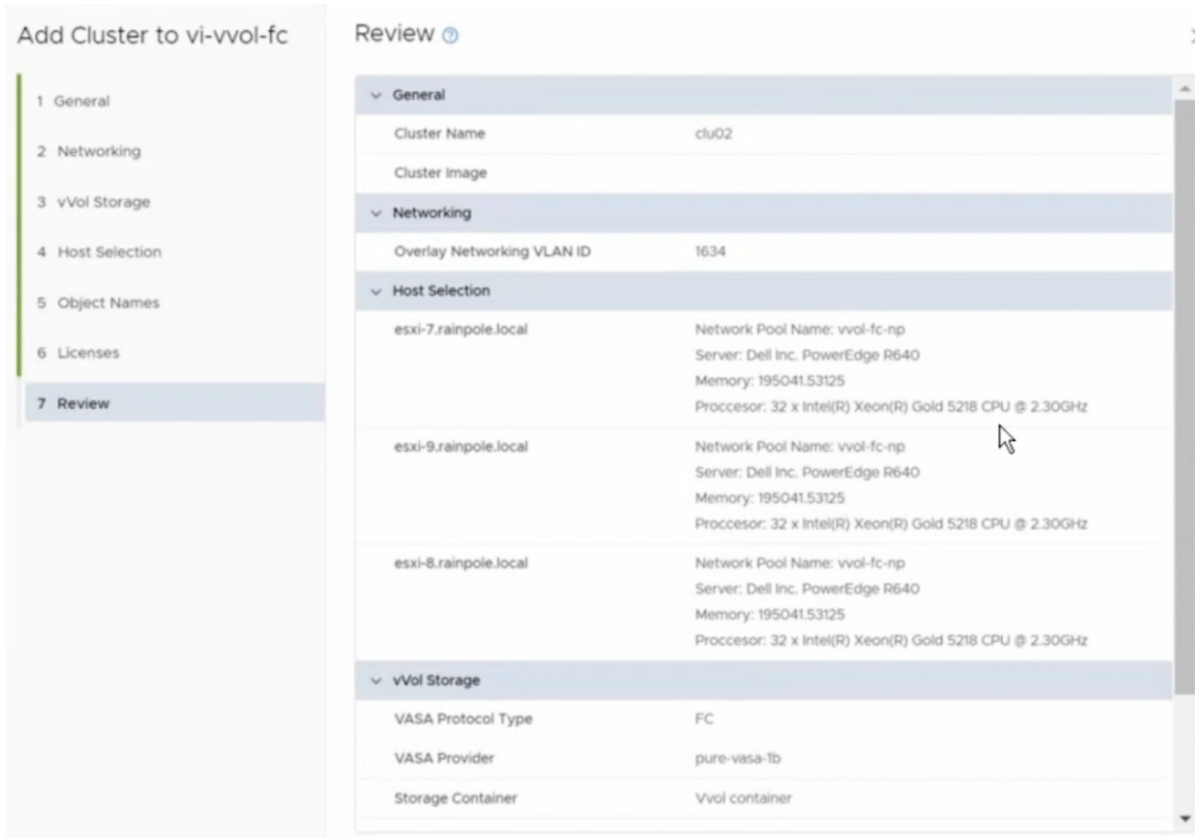
The VASA registration has been enabled outside the workflow in the event incorrect VASA details are entered. This way, the workflow doesn't fail, allowing you to update incorrect information for the VASA registration.



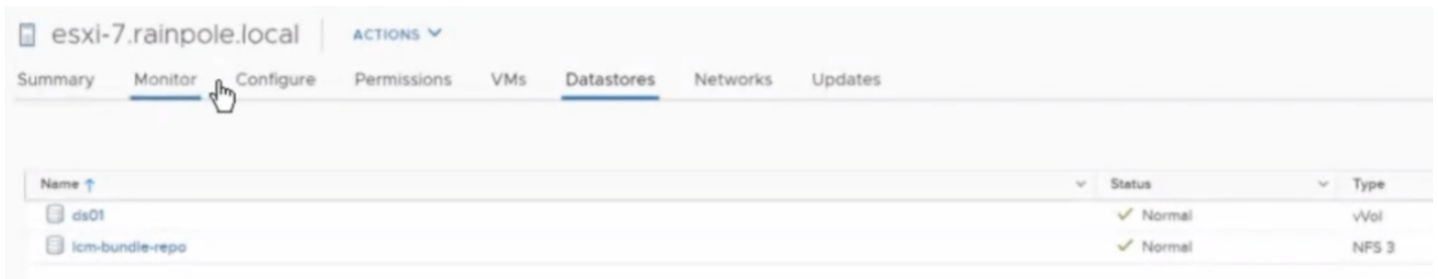
Once you get to the storage for the new WLD, you can enter the details for the vVols datastore.



After going through the rest of the details for the new WLD, you will see we now have vVols storage.



With the WLD build completed, you can then go into your hosts, and you will see a vVols datastore connected to the hosts.



As a default, an SPBM policy, “vVols No Requirement Policy,” is created. We do not create any other SPBM policies because there are too many variables between array types and customer requirements. There is no way to generate an advanced policy, tailored to the requirements needed, without input from the customer. This allows the customer to create specific and tailored SPBM policies that meet application, organization, or security requirements.

vVols continues to be developed for many of VMware’s products and our partners are also continue to enable more and more features for vVols. To learn more about vVols or VCF, head over to the [vVols](#) or [VCF](#) pages at core.vmware.com.

To learn more about VCF and vVols, make sure to attend Todd Simmons and my VMworld session.
[VCF and vVols: Empower Your External Storage \[HCI2270\]](#)

Be sure to check out the VCF announcement blog.
[What’s New with VMware Cloud Foundation 4.1](#)

More VMworld sessions

- [VMware Cloud Foundation at VMworld 2020](#)
- [VMworld Core Storage & vVols Track](#)

vSphere 7 U1 storage related blogs

- [vSphere 7 Update 1 - vSphere Clustering Service \(vCLS\)](#)

vSphere 7.0

Core Storage features and enhancements included in vSphere 7.0

NVMeoF

NVMeoF Insight

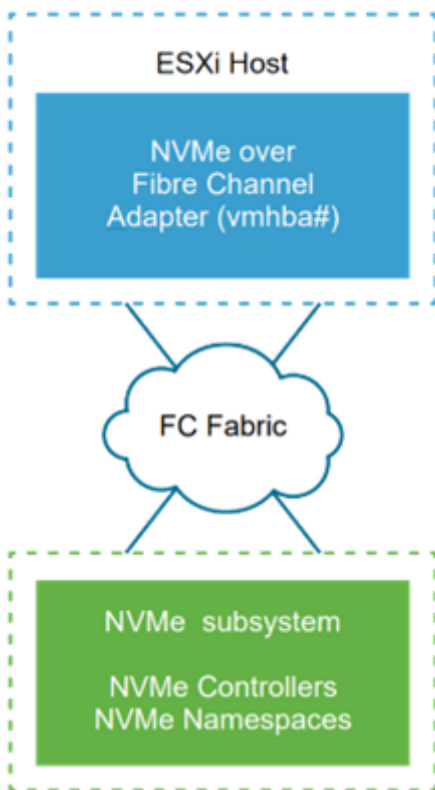
NVMe continues to become more and more popular because of its low latency and high throughput. Industries, such as Artificial Intelligence, Machine Learning, and IT, continue to advance, and the need for increased performance continues to grow. Typically, NVMe devices are local using the PCIe bus. So how can you take advantage of NVMe devices in an external array? The industry has been advancing external connectivity options using NVMe over Fabrics (NVMeoF). Connectivity can be either IP or FC based. There are some requirements for external connectivity to maintain the performance benefits of NVMe as typical connectivity is not fast enough.

VMware NVMeoF

In vSphere 7, VMware added support for shared NVMe storage using NVMeoF. For external connectivity, NVMe over Fibre Channel and NVMe over RDMA (RoCE v2) are supported.

With NVMeoF, targets are presented as namespaces, which is equivalent to SCSI LUNs, to a host in Active/Active or Asymmetrical Access modes. This enables ESXi hosts to discover, and use the presented NVMe namespaces. ESXi emulates NVMeoF targets as SCSI targets internally and presents them as active/active SCSI targets or implicit SCSI ALUA targets.

NVMe over Fibre Channel



This technology maps NVMe onto the FC protocol enabling the transfer of data and commands between a host computer and a target storage device. This transport requires an FC infrastructure that supports NVMe.

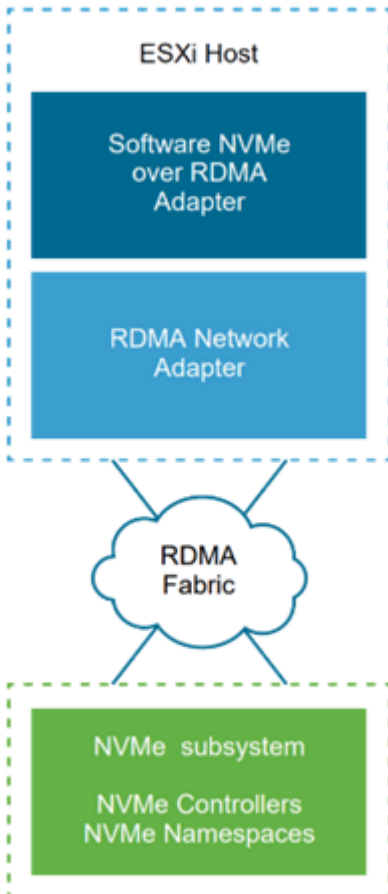
To enable and access NVMe over FC storage, install an FC adapter supporting NVMe in your ESXi host. There is no configuration required for the adapter; it will automatically connect to an appropriate NVMe subsystem and discovers all shared NVMe storage

devices. You may, at a later time, reconfigure the adapter and disconnect its controllers or connect other controllers.

NVMe over FC Requirements

- NVMe array supporting FC
- Compatible vSphere 7 ESXi host
- HW NVMe adapter (HBA supporting NVMe)
- NVMe controller

NVMeoF over RDMA



This technology uses Remote Direct Memory Access (RDMA) transport between two systems on the network. The transport enables in memory data exchange, bypassing the operating system or processor of either system. ESXi supports RDMA over Converged Ethernet v2 (RoCE v2).

To enable and access NVMe storage using RDMA, the ESXi host uses an RNIC adapter on your host and a SW NVMe over RDMA storage adapter. You must configure both adapters to use them for NVMe storage discovery.

NVMe over RDMA requirements:

- NVMe array supporting RDMA (RoCE v2) transport
- Compatible ESXi host
- Ethernet switches supporting a lossless network.
- Network adapter supporting RoCE v2
- SW NVMe over RDMA adapter
- NVMe controller

- [Network Requirements for RDMA over Converged Ethernet](#)

NVMeoF Setup Prerequisites

When setting up NVMeoF, there are a few practices that should be followed.

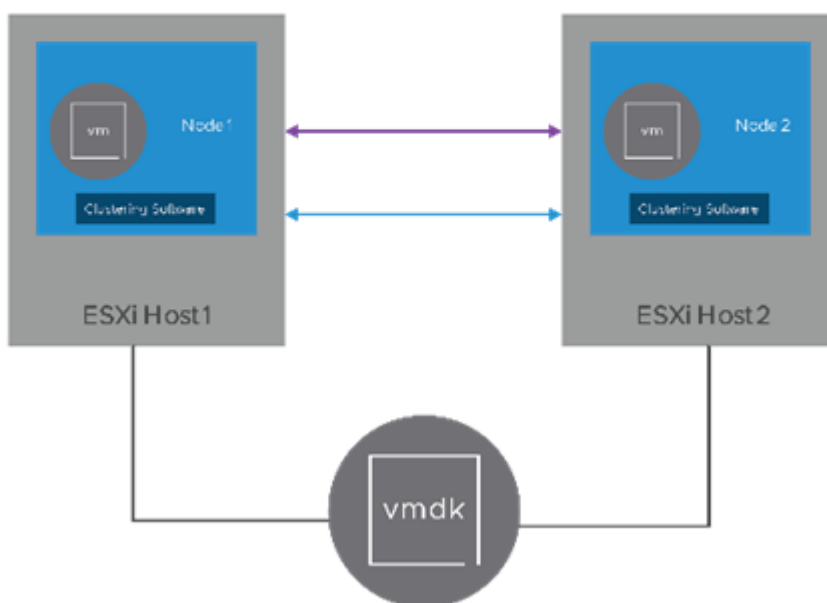
- Do not mix transport types to access the same namespace.
- Ensure all active paths are presented to the host.
- NMP is not used/support; instead, HPP (High-Performance Plugin) is used for NVMe targets.
- You must have dedicated links, VMkernels, and RDMA adapters to your NVMe targets.
- Dedicated layer 3 VLAN or layer 2 connectivity
- Limits:
 - Namespaces=32
 - Paths=128 (max 4 paths/namespace on a host)

Clustered VMDK

Shared VMDK or Clustered VMDK in VMFS6

In vSphere 7, VMware added support for SCSI-3 Persistent Reservations (SCSI-3 PR) at the virtual disk (VMDK) level. What does this mean? You now have the ability to deploy a Windows Server Failover Cluster (WSFC), using shared disks, on VMFS. This is yet another move to reduce the requirement of RDMs for clustered systems. With supported HW, you may now enable support for clustered virtual disks (VMDK) on a specific datastore. Allowing you to migrate off your RDMs to VMFS and regain much of the virtualization benefits lost with RDMs.

Microsoft Windows Server Failover Clustering (WSFC)



Clustered/Shared VMDKs on VMFS6 Prerequisites

- Your array must support ATS, SCSI-3 PR type Write Exclusive-All Registrant (WEAR).

- Only supported with arrays using Fibre Channel (FC) for connectivity.
- Only VMFS6 datastores.
- Storage devices can be claimed by NMP or any other third-party (non-VMware) plugins (MPPs). But please check with the vendor regarding the support for Shared VMDK before using their plugin (MPP).
- VMDKs must be Eager Zeroed Thick (EZT) Provisioned.
- Clustered VMDKs must be attached to a virtual SCSI controller with bus sharing set to "physical."
- A DRS anti-affinity rule is required to ensure VMs, nodes of a WSFC, run on separate hosts.
- Change/increase the WSFC Parameter "QuorumArbitrationTimeMax" to 60.

Other Caveats

- Windows Server 2012R2/2016/2019. SQL Server 2016/2017 was used to validate the configuration.
- The boot disk (and all non-shared disks) should be attached to a separate virtual SCSI controller with bus sharing set to "none."
- Mixing clustered and non-shared disks on a single virtual SCSI controller is not supported.
- The datastore cannot be expanded or span multiple extents.
- All hosts and vCenter must be vSphere 7 or above
- A mix of clustered VMDKs and other disk types (vVols, RDMs) are not supported
- Limits:
 - Support for up to 5 WSFC nodes (Same as RDMs)
 - 128 clustered VMDKs per host
- Only Cluster across Box (CaB) is supported, Cluster in a Box (CiB) is not supported.

vSphere Features

- Supported Features:
 - vMotion to supported hosts meeting the same requirements.
- Unsupported Features:
 - Snapshots, cloning, and Storage vMotion
 - Fault Tolerance (FT)
 - Hot change to VM HW or hot expansion of clustered disks

Here's a link to VMware's document on [Microsoft Clusters](#).

Enabling Clustered VMDK

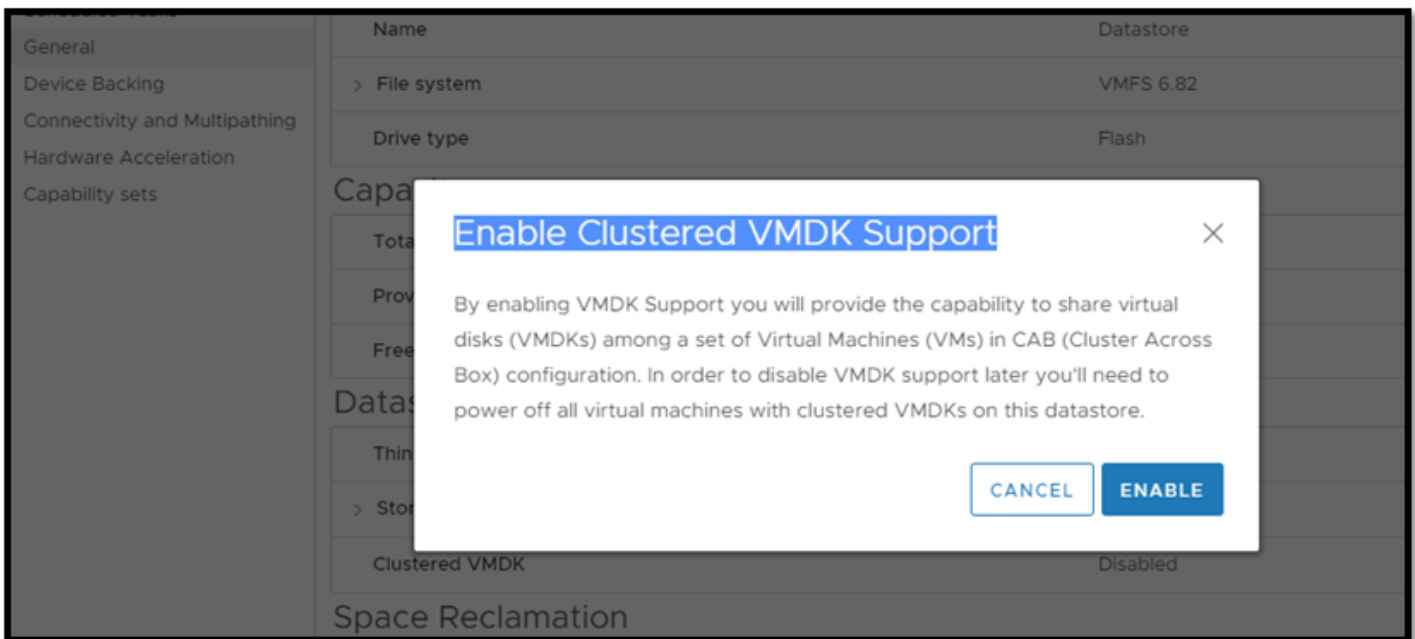
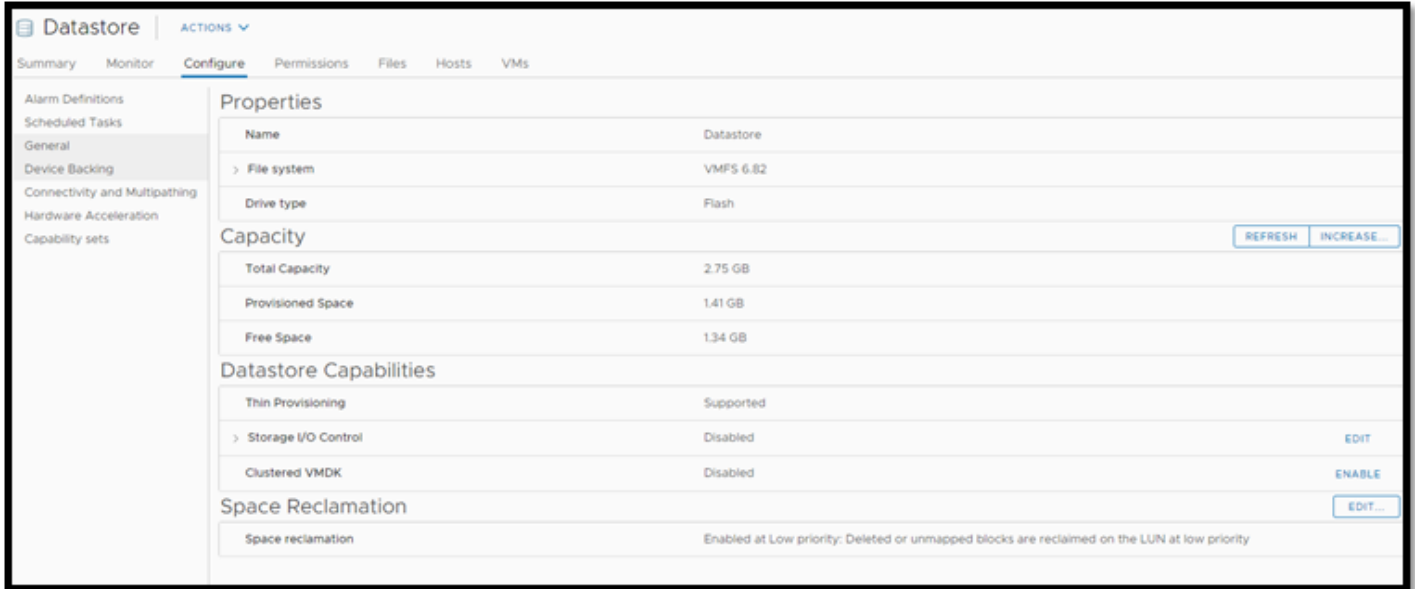
When you navigate to your supported datastore, under the Configure tab, you will see a new option to enable Clustered VMDK. If you are going to migrate or deploy a Microsoft WSFC cluster using shared disks, then you would enable this feature. Once the feature is enabled, you can then follow the [Setup for Windows Server Failover Clustering](#) documentation to deploy your WSFC on the VMFS6 datastore.

Demo and details on migrating WSFC using RDMs to Shared VMDK on VMFS
[vSphere 7 WSFC RDM to Shared VMDK Migration](#)

Enabling Clustered VMDK

Enabling Shared VMDK support on VMFS Datastores

When you navigate to your supported datastore, under the Configure tab, you will see a new option to enable Clustered VMDK. If you are going to migrate or deploy a Microsoft WSFC cluster using shared disks, then you would enable this feature. Once the feature is enabled, you can then follow the [Setup for Windows Server Failover Clustering](#) documentation to deploy your WSFC on the VMFS6 datastore.



Perennially Reserved Flag for RDMs

Using the Perennially Reserved Flag for WSFC RDMs

In cases where customers are using numerous pRDMs in their environment, host boot times or storage rescans can take a long time. The reason for the longer scan times is each LUN attached to a host is scanned at boot or during a storage rescan. Typically, RDMs are provisioned to VMs for Microsoft WSFC and are not directly used by the host. During the scan, ESXi attempts to read the partitions on all the disks but it is unable to for devices persistently reserved by the WSFC. Subsequently, the longer it can take a host to boot or rescan storage. The WSFC uses SCSI-3 persistent reservation to control locking between the nodes of the WSFC

which, blocks the hosts from being able to read them.

VMware recommends implementing perennial reservations for all ESXi hosts hosting VM nodes with pRDMs. Check the following [kb1016106](#) or more details

The question then arises; How can you get the host not to scan these RDMs and reduce boot or rescan times? I'm glad you asked!

There is a device flag called "Perennially Reserved" which tells the host the RDM should not be scanned and is used elsewhere (perennially) in the environment. Before vSphere 7, this flag is enabled via CLI and requires the UUID (naa.ID).

The command to set the flag to true:

```
esxcli storage core device setconfig -d naa.id --perennially-reserved=true
```

To verify the setting:

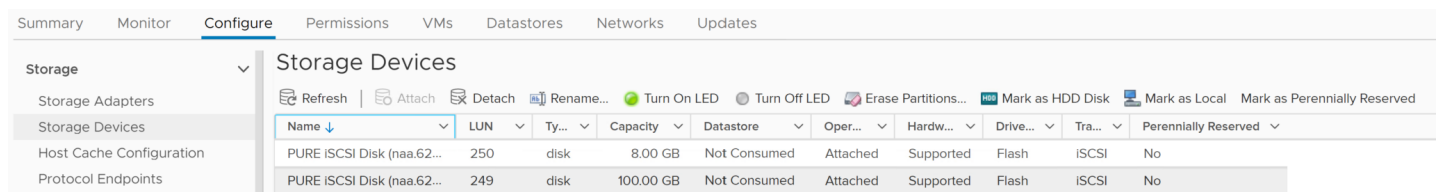
```
esxcli storage core device list -d naa.id
```

In the device list you should see:

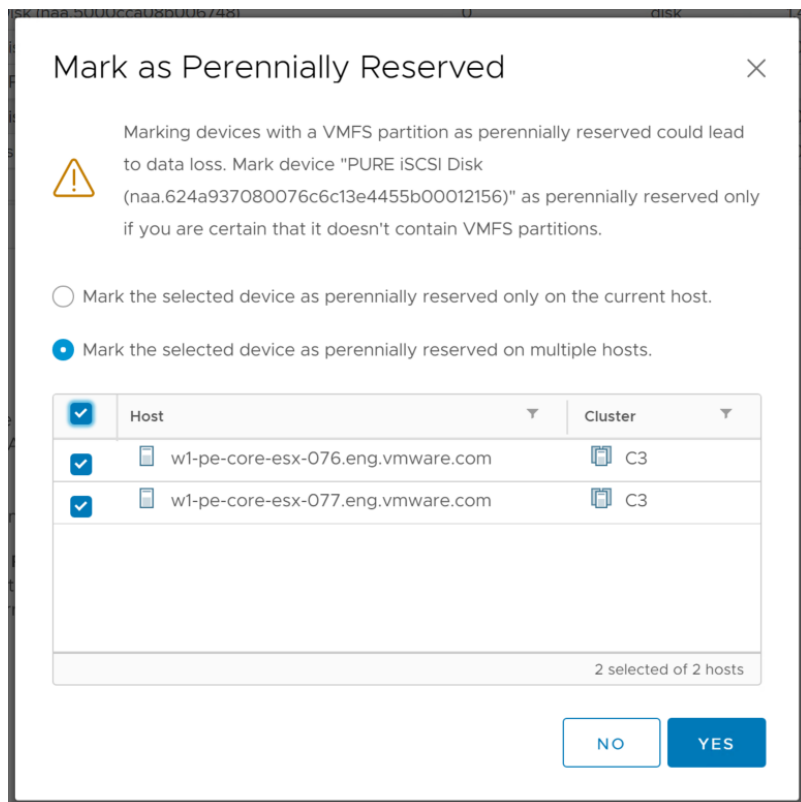
Is Perennially Reserved: true

When setting this option, it must be run for each relevant RDM used by the WSFC and on every host with access to that RDM. You can also set the Perennially Reserved flag in Host Profiles.

With the release of vSphere 7, setting the Perennially Reserved flag to true was added to the UI under storage devices. There has also been a field added to show the current setting for the Perennially Reserved flag.



Once you select the RDM to be Perennially Reserved, you have the option to "Mark the selected device as perennially reserved" for the current host or multiple hosts in the cluster. This eliminates the manual process of setting the option per host via CLI. If preferred, you can still use ESXCLI, PowerCLI, or Host Profiles.



Once you click YES, the flag will be set to true and the UI updated.

What's New in vSphere 7 Core Storage

Summary Monitor **Configure** Permissions VMs Datastores Networks Updates

Storage

Storage Adapters

Storage Devices

Host Cache Configuration

Protocol Endpoints

Storage Devices

Refresh Attach Detach Rename... Turn On LED Turn Off LED Erase Partitions... Mark as HDD Disk Mark as Local Unmark as Perennially Reserved

Name	LUN...	Ty...	Capacity	Datastore	Oper...	Hardw...	Drive...	Tra...	Perennially Reserved
PURE ISCSI Disk (naa.624a937054ce6d72d23042cd000117e5)	1	disk	1.00 MB	Not Consumed	Attached	Supported	Flash	ISCSI	No
PURE ISCSI Disk (naa.624a937054ce6d72d23042cd000117e5)	249	disk	100.00 GB	Not Consumed	Attached	Supported	Flash	ISCSI	Yes

You may also unmark the device using the same process.

Unmark as Perennially Reserved

Unmark the selected device as perennially reserved only on the current host.

 Unmark the selected device as perennially reserved on multiple hosts.

Host	Cluster
<input checked="" type="checkbox"/> w1-pe-core-esx-076.eng.vmware.com	<input checked="" type="checkbox"/> C3
<input checked="" type="checkbox"/> w1-pe-core-esx-077.eng.vmware.com	<input checked="" type="checkbox"/> C3

2 selected of 2 hosts

NO YES

Short Clip showing process:

w1-pe-core-esx-076.eng.vmware.com ACTIONS

Summary Monitor **Configure** Permissions VMs Datastores Networks Updates

Storage

Storage Adapters

Storage Devices

Host Cache Configuration

Protocol Endpoints

I/O Filters

Networking

Virtual switches

VMkernel adapters

Physical adapters

RDMA adapters

TCP/IP configuration

Virtual Machines

VM Startup/Shutdown

Agent VM Settings

Default VM Compatibility

Swap File Location

System

Licensing

Host Profile

Time Configuration

Authentication Services

Certificate

Power Management

Advanced System Settings

System Resource Reservation

Firewall

Services

Security Profile

Storage Devices

Refresh Attach Detach Rename... Turn On LED Turn Off LED Erase Partitions... Mark as HDD Disk Mark as Local Mark as Perennially Reserved

Name	LUN...	Type	Capacity	Datastore	Operatio...	Hardware Accelerat...	Drive T...	Transport	Perennially Reserved
Local HP Enclosure 3VC Dev (naa.60060160cb128028398787c5...)	0	disk	8.00 GB	Not Consumed	Attached	Not supported	Flash	ISCSI	No
PURE ISCSI Disk (naa.624a937080076c6c13e4455b...)	250	disk	1.00 MB	Not Consumed	Attached	Supported	Flash	ISCSI	No
DGC ISCSI Disk (naa.60060160cb128028398787c5...)	511	disk	512.00 B	Not Consumed	Attached	Supported	HDD	ISCSI	No
PURE ISCSI Disk (naa.624a937080076c6c13e4455b...)	253	disk	1.00 MB	Not Consumed	Attached	Supported	Flash	ISCSI	No
Local HGST Disk (naa.5000cca08b006740)	0	disk	1.46 TB	Not Consumed	Attached	Not supported	Flash	SAS	No
Local HGST Disk (naa.5000cca08b006740)	0	disk	1.46 TB	Not Consumed	Attached	Not supported	Flash	SAS	No
Local HGST Disk (naa.5000cca08b006748)	0	disk	1.46 TB	Not Consumed	Attached	Not supported	Flash	SAS	No
PURE ISCSI Disk (naa.624a937080076c6c13e4455b...)	254	disk	50.00 GB	Not Consumed	Attached	Supported	Flash	ISCSI	No
Local Marvell Processor (eui.0050430000000000)	0	scsi pr...	Not Consumed	Not Consumed	Attached	Not supported	HDD	Block Adapter	No
PURE ISCSI Disk (naa.624a937080076c6c13e4455b...)	249	disk	100.00 GB	Not Consumed	Attached	Supported	Flash	ISCSI	No
DGC ISCSI Disk (naa.60060160cb128028a687a85e...)	0	disk	100.00 GB	Not Consumed	Attached	Supported	HDD	ISCSI	No

Copy All 23 items

Properties Paths Partition Details

General

Name PURE ISCSI Disk (naa.624a937054ce6d72d23042cd000117e5)

Identifier naa.624a937054ce6d72d23042cd000117e5

Type disk

Location /vmfs/devices/disks/naa.624a937054ce6d72d23042cd000117e5

Capacity 4.50 TB

Drive Type Flash

Hardware Acceleration Supported

Transport ISCSI

Owner NMP

Sector Format 512n

Multipathing Policies

Path Selection Policy Round Robin (VMware)

Storage Array Type Policy VMW_SATP_ALUA

Edit Multipathing...

Setting the Perennially Reserved flag on your pRDMs used by your WSFC is recommended in the clustering guides. When set, ESXi no longer tries to scan the devices and this can reduce boot and storage rescan times. I have added links below to resources on clustering guides and the use of this flag. Another benefit of flagging RDMs is you can easily see which devices are RDMs and which are not.

Resources:

- ESXi/ESX hosts with visibility to RDM LUNs being used by MSCS/WSFC nodes with RDMs may take a long time to start or during LUN rescan ([1016106](#))
- Guidelines for Microsoft Clustering on vSphere ([1037959](#))
- Microsoft Windows Server Failover Clustering (WSFC) with shared disks on VMware vSphere 7.x: Guidelines for supported configurations ([79616](#))
- Microsoft Windows Server Failover Clustering (WSFC) with shared disks on VMware vSphere 6.x: Guidelines for supported configurations ([2147661](#))

Affinity 2.0

VMFS Affinity Manager 2.0

Overview

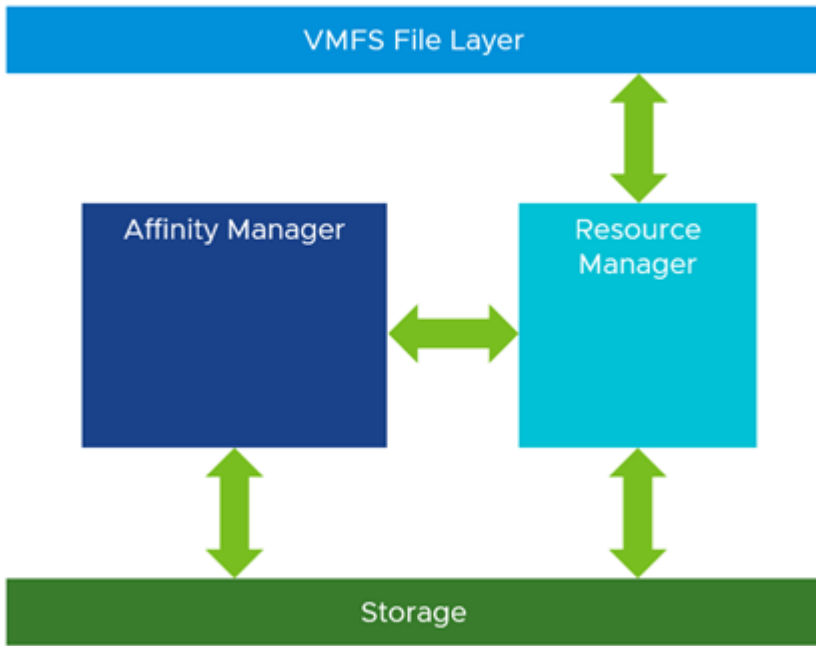
Some of the benefits of deploying VMs as Thin Provisioned VMDKs are the effective use of space, and space reclamation. A thin VMDK is a file on VMFS, where Small File Blocks (SFB) are allocated on-demand at the time of first write IO. There can be an overhead cost to this process, which can affect performance. In some cases, for maximum performance, it is recommended Eager Zero Thick (EZT) disks are utilized to avoid the overhead of allocating space for new data.

First Write Process

In VMFS, resources are organized, on-disk, in groups called “Resource Clusters” (RC). When a file requires new storage resources, (new data to be written), there are two decisions made.

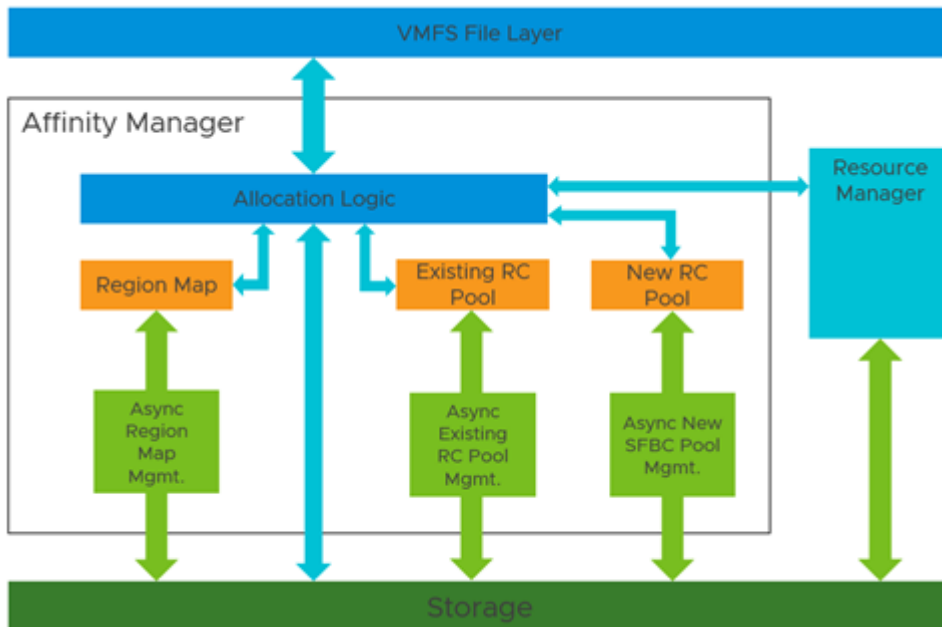
1. Which RC to fetch the resources.
2. Which resources from the RC to allocate.

The “Affinity Manager” is a component of VMFS6 responsible for RC allocation. The Affinity Manager allocates RCs, particularly Small File Block Cluster (SFBC), such that minimal files share the same RC. The process of mapping a resource to an RC is called affinity. VMFS is shared storage, and multiple hosts can request resources in parallel. Resource allocations are synchronized using an on-disk lock (ATS) with no other communication between hosts. This can lead to issues during allocation. The Resource Manager requests RCs from the Affinity manager, who then responds to the VMFS file layer.



New Affinity 2.0 Manager

With vSphere 7, we are introducing Affinity 2.0, which is designed to be smarter and more efficient in the allocation of resources, minimizing overhead in the IO path. This is accomplished by creating a cluster-wide view of the resource allocation status of the RCs, dividing disk space into “regions.” Then ranking them based on the number of blocks allocated by the current host as well as other hosts. This information is called a “Region Map.” With the Region Map, the Affinity Manager can now quickly supply information on available RCs to the Resource Manager.



Benefit

What does all this mean? When the Resource Manager requests RCs from the new Affinity Manager, when a first write IO occurs, the allocations are no longer a “go and find” an RC to allocate avoiding the back and forth overhead. With the Region Map, the Affinity Manager knows where and what RC resources are available and can quickly direct the Resource Manager. The Resource request, from the VMFS file layer, now goes directly to the Affinity Manager. The Affinity Manager relies on the generated Region Map to find an available RC to allocate. The Affinity Manager then locks the RC on-disk, checks for free resources, and hands it over to the Resource Manager. This reduces the repeated back and forth between the Affinity Manager and Resource Managers trying to find an available RC. Therefore, reducing metadata IO and the overhead required for the first write on thin-provisioned

disks. It can also improve the allocation of space for EZT or LZT provisioning.

End Of Availability (EOA) vFRC and CBRC 1.0

With the release of vSphere 7, there are a few antiquated features that have reached End of Availability.

vSphere Flash Read Cache (vFRC) EOL

vFRC currently has a minimal customer base. With VAIO, it allows 3rd party vendors to create custom caching solutions. When you upgrade to vSphere 7, you will receive a warning message that vFRC will no longer be available. "vFRC will be gone with this upgrade, please deactivate vFRC on a VM if using it."

Remediation Pre-check | Cluster-1

⚠️ Cluster has issues that may prevent completion of remediation ⓘ

Issue	Issue status
> DRS should be enabled, if available	⚠️ User attention suggested

EXPORT 1 Actions

✅ All hosts are ready to remediate ⓘ

HOST SUMMARY DETAILS

Host	Issue	Issue status	VM
10.161.56.6	vFRC will be gone with this upgrade	⚠️ User attention suggested	vm

Open Documentation for Remediation Precheck

RE-RUN PRE-CHECK DONE

Content-Based Read Cache (CBRC) 1.0 EOA

CBRC 1.0 has a maximum cache size of 2GB, whereas 2.0 has a maximum of 32GB. As of vSphere 6.5, CBRC 2.0 is the default for Content-Based Read Cache. Starting in vSphere 7, CBRC 1.0 has been removed to ensure it is not used, especially in Horizon environments. This will also eliminate the building and compiling of un-used code.

vVols Interoperability

vVols Interoperability with VMware Products

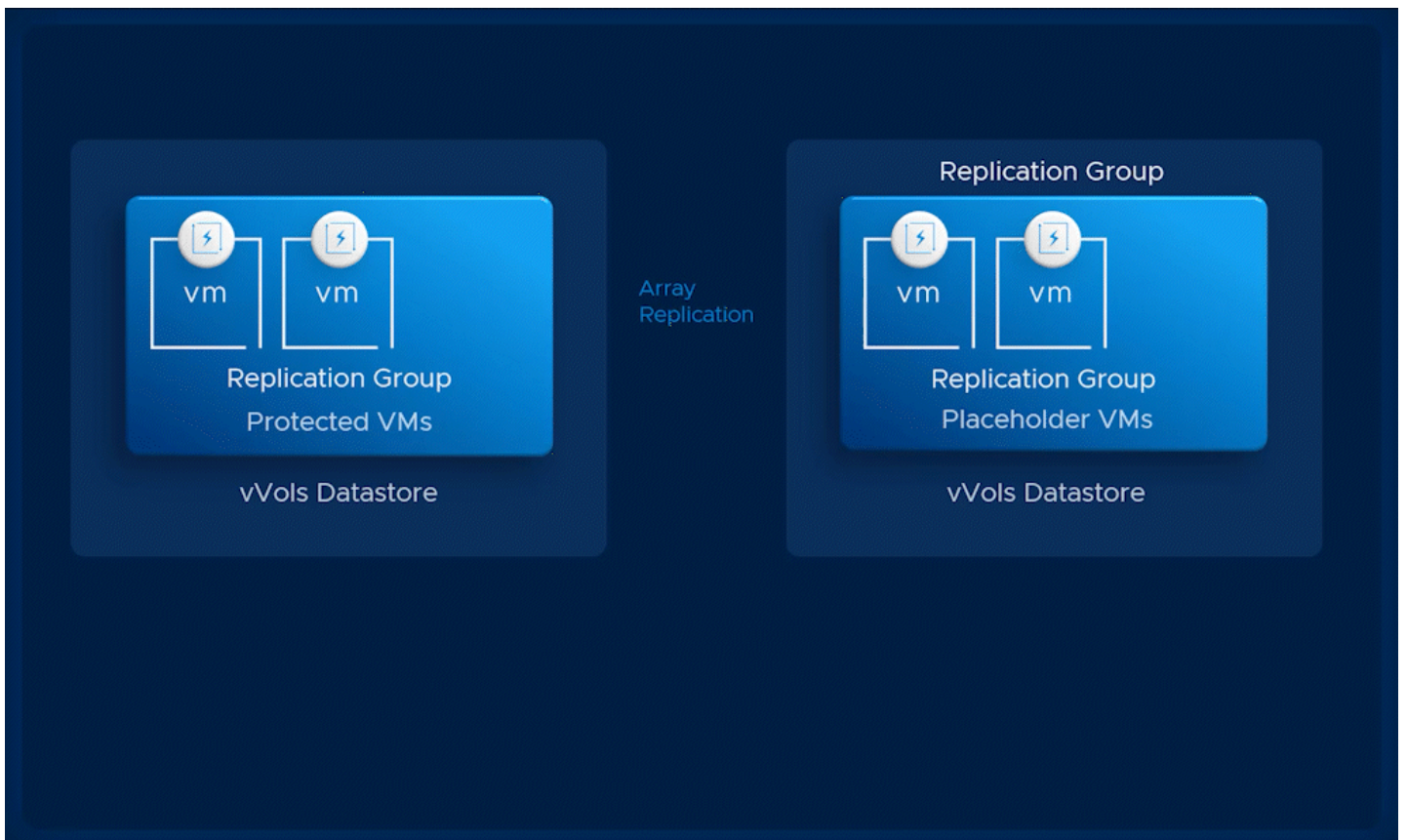
VMware Virtual Volumes (vVols) adoption continues to grow and is accelerating in 2020, and it's easy to see why. vVols eliminates LUN management, accelerates deployments, simplifies operations, and enables utilization of all of your array's functionality. VMware and our storage partners continue to develop and advance vVols, and its functionality. In vSphere 7, more features and enhancements have been added, showing the continued commitment to the program.

vVols Support in SRM 8.3

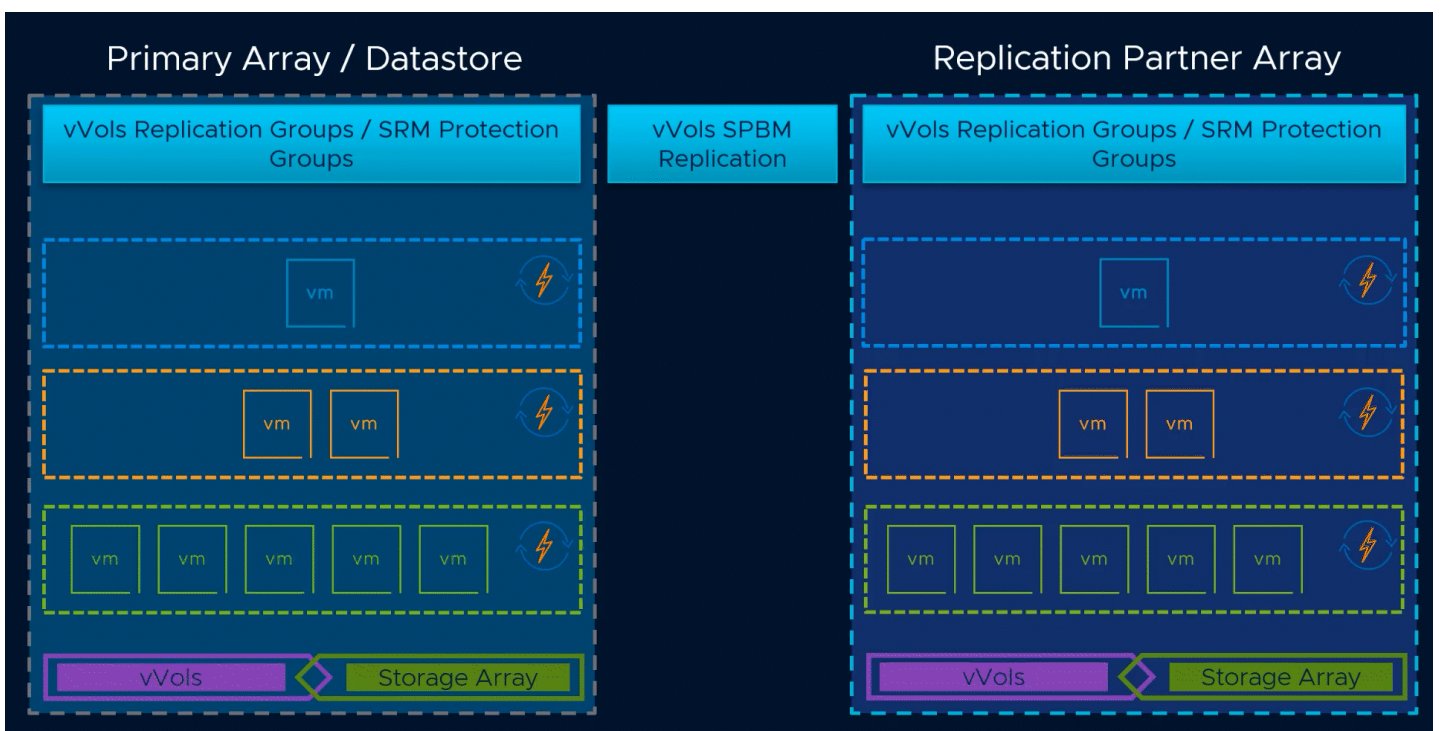
Site Recovery Manager - SRM

Because vVols uses array-based replication, it is very efficient. Array-based replication is a preferred method of replicating data between arrays. With vVols and SPBM, you can easily manage which VMs are replicated rather than everything in a volume or LUN. With the release of Site Recovery Manager 8.3, you can now manage your DR process with SRM while using the replication efficiency and granularity of vVols and SPBM.

Here's a link to the [announcement blog for SRM 8.3](#)



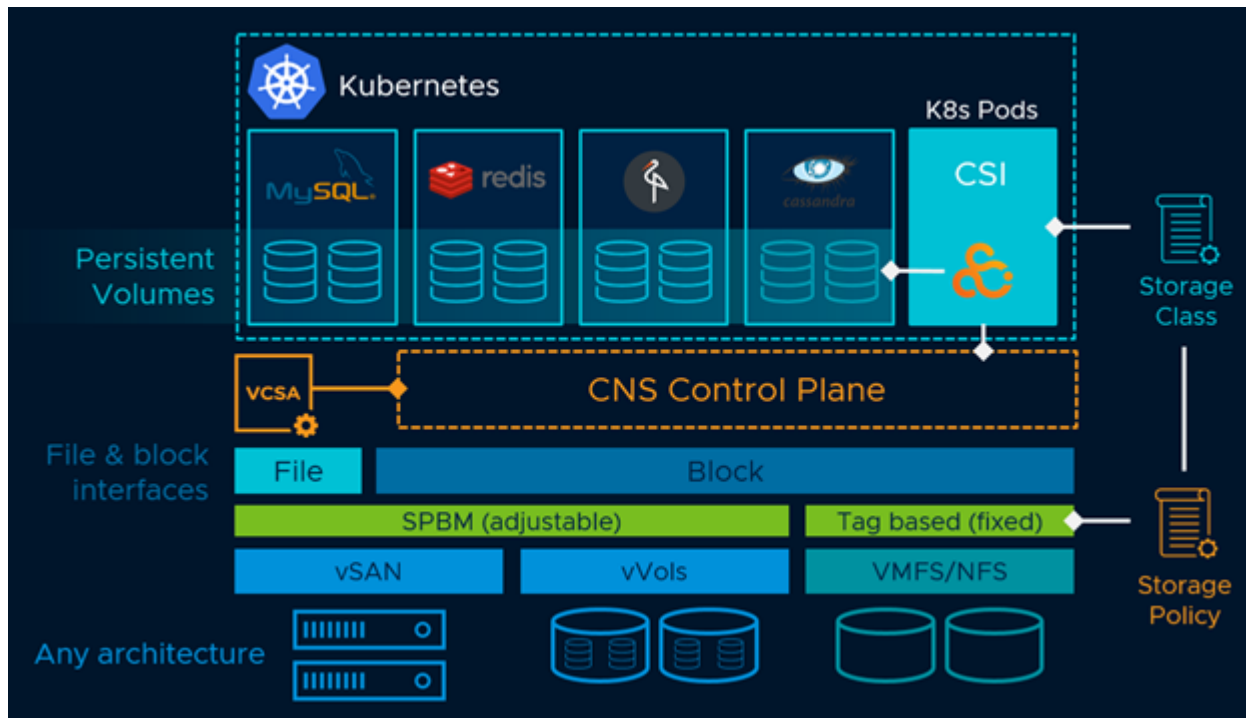
With vVols and SRM, you can have independent vVols replication-groups/SRM protection-groups for a single VM, application, or group of VMs. Another benefit is each replication-group/protection-group can have different RPOs, and all use array-based replication.



vVols Support for CNS

Cloud-Native Storage - CNS

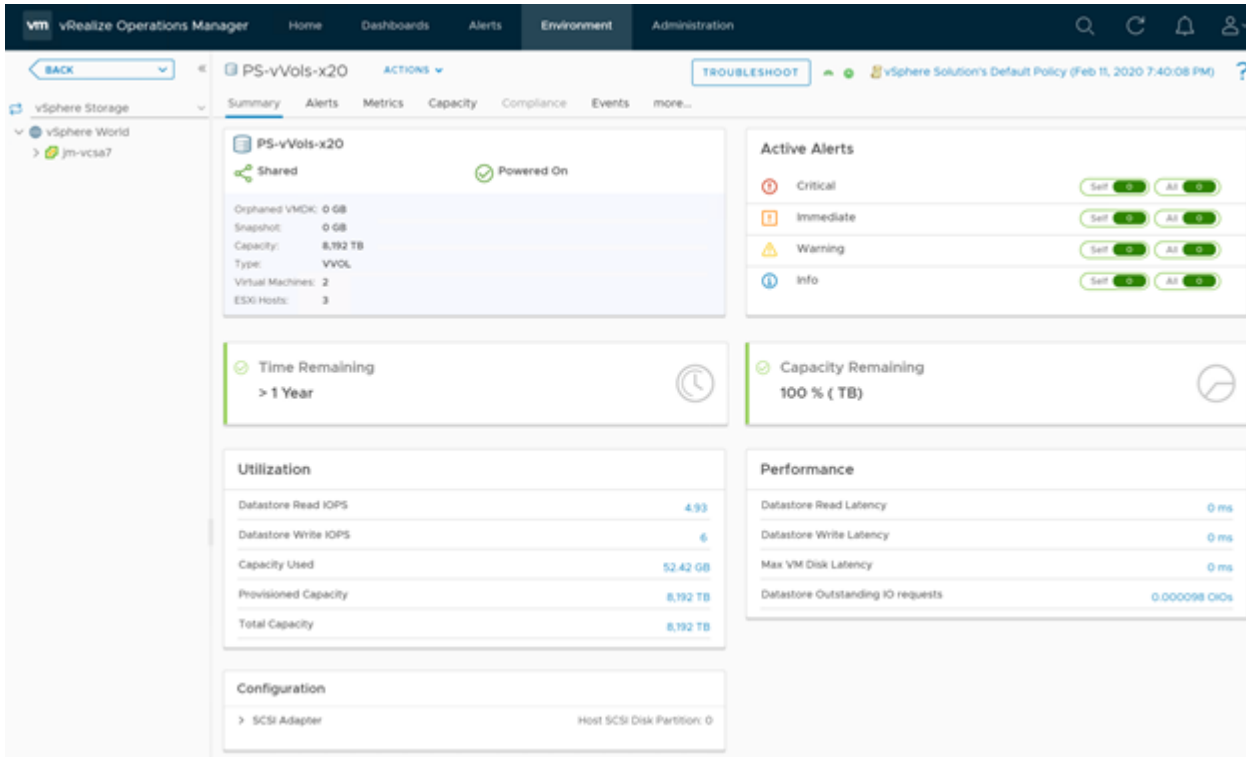
Kubernetes continues to grow in adoption, and VMware is on the forefront. One of K8s' requirements is persistent storage, and until now, that included vSAN, NFS, and VMFS. The thing is vVols couldn't be more suited for K8s storage because a vVol is its own entity. Now, deploy an FCD as a vVol and you've got a first-class disk as a first-class citizen that has additional benefits like mobility and CSI to SPBM policy mapping. With the initial release, snapshots and replication with vVols will not be supported.



vVols Support in vRealize Operations 8.1

vRealize Operations - vROps

A feature that has been requested for a while is finally available; support for vVols datastores in vROps! With the release of vROps 8.1, you can now utilize vROps monitoring on your vVols datastores the same as any other datastore. Giving your alerting, planning, troubleshooting, and more for your vVols datastores. For more information here's the link to [vROps](#). Make sure to read about the new release on the [vROps 8.1 announcement blog](#).



vVols as Supplemental Storage in VCF

VMware Cloud Foundation - VCF

VMware Cloud Foundation allows organizations to deploy and manage their private and public clouds. VCF currently supports vSAN, VMFS, and NFS principal storage. Customers are asking for the support of vVols as principle storage, and while the VCF team continues to evaluate and develop that option, it is not available. In the meantime, vVols can be used as supplemental storage after the Workload Domain build has completed. Support for vVols as supplemental storage is a partner supported option.

Please work with your storage array vendor for the supported processes and procedures in setting up vVols with VCF as supplemental storage.

For more information, here's the link to [VCF](#).

Here's a link to the blog on [What's New in VCF 4.0](#)

