RESILIENT AND SCALABLE FILE SERVICE ON VMWARE vSAN™ WITH COHESITY

White Paper

vmware®

Table of Contents

Executive Summary

IT organizations are under high pressure to deliver fast and reliable services with lower operational costs. The growing demands of new technology, multi-cloud integration, security, data protection and the fast pace of modern business push the legacy IT design to change. Separate servers, storage networks and storage arrays management need more IT staff with deep knowledge in each area. The silos created by traditional datacenter infrastructure often present difficulties to share data and integrate fully automated solutions, adding complexity to every step from ordering, deployment, management and backup. Data is a differentiator in the digital economy and the data has become the most valuable and the most targeted business asset. To truly protect their data, customers need a holistic solution that can proactively prevent, detect, and respond to virus and ransomware attacks.

With a hyperconverged infrastructure solution, it is easier to scale out, streamline the deployment and automate operations. It increases data protection and performance with lower cost and more efficient IT staff. VMware vSAN and Cohesity DataPlatform are the perfect hyperconverged infrastructure combination with turnkey software-defined solution for primary and backup workloads. vSAN provides a simple evolution to full stack HCI, broadest flexibility and multi-cloud capability.

vSAN allows to pool storage capability, automatically provision vm storage and dynamically scale performance and capacity as needed, through a policy driven control plane.

Cohesity provides a policy-driven approach for file services for VMs running on VMware vSphere and vSAN environments. Cohesity delivers scale-out file storage for the cloud era in a modern, software-defined solution that eliminates the complexity and costly management of a fragmented and inefficient storage environment.

The vSAN and Cohesity solution consolidates all file services workloads via multiprotocol access (NFS, SMB and S3) on a single platform that spans from core, to cloud and edge. This reduces the TCO drastically for enterprises as they do not have to spend on siloed infrastructure, which brings point solutions and exponentially increases the overall cost of the solution.

Software development and agility being paramount, high availability of infrastructure becomes a core requirement, with this combined solution clients can realize the benefits of a resilient file service, which provides durable and persistent file handles even in an outage scenario.

Solution Overview

To meet rapidly growing unstructured data volumes, businesses require an enterprise-class, scale-out NAS solution that not only supports common file and object protocols (NFS, SMB, S3) but also guarantees data resilience and offers global storage efficiency. Furthermore, the solution needs to pave an easy pathway into the cloud. Cohesity provides a modern web-scale software-defined solution that eliminates complexities and higher management costs stemming from a fragmented and inefficient storage environment. vSAN is a hyperconverged software defined storage that is ideal for resilient and scalable workloads. Cohesity DataPlatform is a data management platform built upon Google-like web-scale principles that consolidate all file services workloads via multiprotocol access (NFS, SMB/CIFS and S3) with unified permissions on a single platform that spans from core, to edge, and into the cloud. Cohesity deployed on vSAN guarantees data resiliency at scale with strict consistency and data efficiency with global variable block length deduplication and compression between different workloads, like VMs, physical machines, databases, NAS, and file share. As a software-defined solution which supports being deployed on any vSAN ready nodes, this solution also natively integrates with AWS, Azure, and Google Cloud to leverage the economics and elasticity of the public cloud.

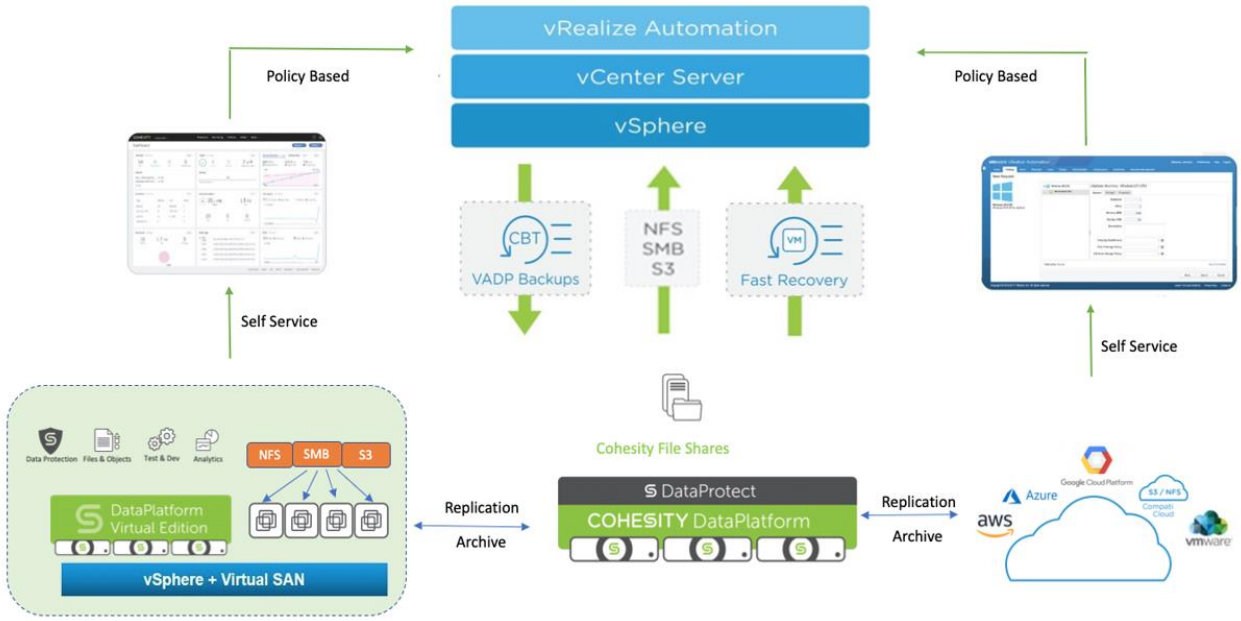

Figure 1. vSAN and Cohesity Solution Overview

This document covers the following areas:

- Deployment process and best practices for designing and implementing Cohesity file services on VMware vSAN hyperconverged solution.
- How this solution addresses the challenges around scalability, resiliency, and efficiency.
- Web-scale for simplicity: compatibility, multiprotocol access, limitless snapshots, strict consistency.
- Limitless scalability, on-prem and in the cloud: scale out architecture, cloud-ready, archival, replication and cloud tiering
- Support for workload diversity: global dedupe, zero-cost clone, QoS, and fast search.
- Designed for data protection: role-based access, policy-based backup, and quota.

Audience

This paper is intended for anyone who wants to design and deploy a scalable and highly available file service in VMware vSAN environment. The information in this paper is ideal for experienced system administrators who are familiar with software defined data center operations, especially in the area of storage.

Technology Overview

This section provides an overview of the technologies used in this solution:

- VMware vSAN
- Cohesity

Introduction to VMware vSAN

VMware vSAN is a storage virtualization solution built from the ground up for the vSphere environment. It abstracts and aggregates locally attached disks in a vSphere cluster to create a single storage pool shared by all hosts in this cluster. The vSAN software runs natively as part of the ESXi hypervisor, invisible to the end users. It can be easily provisioned and managed from vCenter and the vSphere Web Client. vSAN integrates with the entire VMware stack,

including features like vMotion, HA, and DRS. Virtual machine storage provisioning and day-to-day management of storage SLAs can all be controlled through virtual machine-level policies that can be set and modified on-the-fly. vSAN delivers enterprise-class features, scale and performance, making it the ideal storage platform for virtual machines.

Because vSAN software is tightly integrated in the hypervisor, it does not rely on a dedicated virtual appliance pinned to each host in a cluster to provide storage services. This unique architecture simplifies deployment and management, reduces resource overhead and ensures the consistent performance of workload virtual machines. It also allows the customers easily scale up (add more capacity to existing hosts) and scale out (add more hosts to the vSAN cluster) vSAN deployment.

vSAN is enterprise-class, storage virtualization software that, when combined with vSphere, allows you to manage compute and storage with a single platform. With the key benefits and features below, it provides an ideal infrastructure for running Cohesity File Service.

vSAN simplifies operations by allowing customers to quickly build and integrate cloud infrastructure with guided, exhaustive instructions for complex tasks, which makes it easy to get started with HCI. It also keeps infrastructure stable and secure with automated patching and upgrades. vSAN ensures consistent application performance and resiliency during maintenance operations and reduces time spent troubleshooting maintenance issues. vSAN's centralized health monitoring helps customers to quickly identify and resolve any performance issue.

vSAN automates space reclamation, dynamically reducing application storage usage over time, freeing up valuable resources as well as enhancing application performance. It also enables admins to size capacity needs correctly and incrementally to improve capacity management and planning.

Table 1. vSAN Key Features

Features	Description
Tightly Integrated with vSphere	Built into the vSphere kernel, optimizing the data I/O path to provide the highest levels of performance with minimal impact on CPU and memory.
VM-centric Policy-based Management	Part of the larger VMware SDDC stack that uniquely delivers consistent, VM-centric operations through policy-based management. Using simple policies, common tasks are automated and storage resources are balanced to reduce management time and optimize HCI efficiency.
Single Pane of Glass Management	Natively integrates with the user interface of the SDDC stack, removing the need for training and operating specialized storage interfaces. vSAN uses a modern HTML5-based web client. VMware vRealize® Operations™ within VMware vCenter® enables rapid visibility into a vSAN deployment with broad monitoring and deep analytics, all from vCenter.
Flash-optimized	Minimizes storage latency with built-in caching on server-side flash devices delivering up to 50% more IOPS than previously possible.
Granular Non-disruptive Scale-up or Scale-out	Non-disruptively expand capacity and performance by adding hosts to a cluster (scale-out) or just grow capacity by adding disks to a host (scale-up).
Deduplication and Compression	Software-based deduplication and compression optimizes all-flash storage capacity, providing as much as 7x data reduction with minimal CPU and memory overhead.
Erasure Coding	Erasure Coding increases usable storage capacity by up to 100% while keeping data resiliency unchanged. It is capable of tolerating one or two failures with single parity or double parity protection.
vSAN Encryption	Native to vSAN, vSAN Encryption provides data-at-rest security at the cluster level and supports all vSAN features, including space efficiency features like deduplication and compression. Enabled with a few clicks, vSAN Encryption is built for compliance requirements and offers simple key management with support for all KMIP compliant key managers, such as CloudLink, Hytrust, SafeNet, Thales and Vormetric. vSAN Encryption is FIPS 140-2 validated, meeting stringent US Federal Government standards.
Stretched Clusters with Local Protection	Create a robust stretched cluster with site and local protection between two geographically separate sites, synchronously replicating data between sites. It enables enterprise-level availability where an entire site failure can be tolerated as well as local

	component failures, with no data loss and near zero downtime. Users can set granular protection on a per-VM basis and non-disruptively change policies.
Quality of Service (QoS)	QoS controls, limits and monitors the IOPS consumed by specific virtual machines, eliminating noisy neighbor issues.
vSAN Health Service	Health Service provides integrated hardware compatibility checks, performance monitoring, storage capacity reporting and diagnostics directly from VMware vCenter Server.
vSAN Support Insight	vSAN Support Insight helps keep vSAN running in an optimal state, saving monitoring and troubleshooting time, by providing real-time support notifications and actionable recommendations. The analytics tool can also optimize performance for certain scenarios with recommended settings.

Introduction to Cohesity

Cohesity Web-Scale NAS goes beyond scale-out NAS. The Cohesity web-scale architecture eliminates limitations of traditional scale-out NAS – starting with as few as three nodes. Like the web, file and object services scale in multiple dimensions for boundless capacity, performance, cross-silo data deduplication, resiliency, and cloud-like management. File and object services run on-premises, in the cloud, and can connect to the cloud. Either deployment delivers the scale to consolidate multiple NAS silos and point solutions. Sliding window variable dedupe across multiple workload volumes is far more cost effective than legacy dedupe by silo. Data reduction is further enhanced with compression, small file optimization, and no-cost clones. Rich multiprotocol support for SMB, NFS, and S3 along with a data protection and cloud integration make Cohesity Web-Scale NAS a primary solution for enterprise file and object environments.

Cohesity DataPlatform is a hyperconverged, web-scale and Cloud ready platform with these features and benefits.

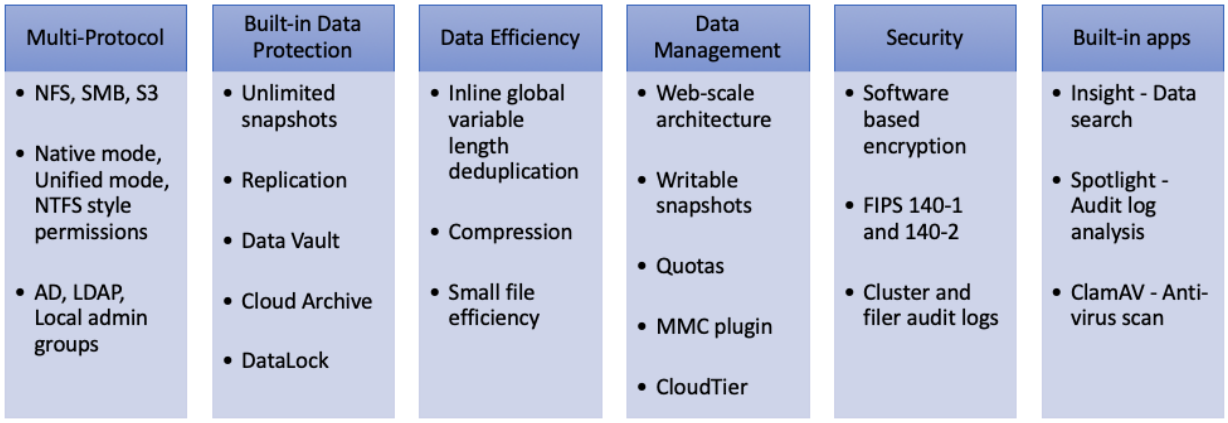

Table 2. Cohesity File Share Key Features

Features	Description
NFSv3, CIFS, SMB2.x, SMB 3.0, and S3 APIs	Multiprotocol access to same data allows support of applications across all major enterprise operating systems including Microsoft Windows, Linux, and S3 API
Strict Consistency	Guaranteed data resiliency at scale
SnapTree® snapshots and clones	Limitless and fully-hydrated snapshots for granular Cohesity Views (file systems) as well as writable snapshot clones that provide instant creation, testing and development of view-based datasets
Web-scale File System	Limitless scalability, always-on availability, non-disruptive upgrades, pay-as-you-grow model

Hyperconverged Secondary Storage	Single platform for data protection, files, objects, test/dev, and analytics.
Global deduplication and compression	Unparalleled storage efficiency with global deduplication and compression across all nodes of the cluster that significantly reduces data center footprint
Erasure coding	Data is protected against any individual node failure with erasure coding across nodes
Global indexing and search	File and object metadata is indexed upon ingest, enabling Google-like search across all files in a cluster
Mixed-mode permission mapping	Cohesity manages the permission mapping and also natively integrates with Centrify. Centrify allows Cohesity to directly access the ID mapping information stored in Centrify's AD. This eliminates the need for LDAP proxy and simplifies the user experience.
Windows Active Directory and Kerberos Integration with Role-Based Access Control (RBAC)	Simplify user and group access to data utilizing credentials and permissions with Windows AD and Kerberos mechanisms. Create and manage custom Cohesity cluster administration roles for domain users and groups
External KMS integration	Internal key management service (KMS) support and integration with external KMS for key management
Quotas	Easily establish user and file system quotas with audit logs
Policy-based backup protection	Integrated data protection software and SnapTree technology is available to allow simplified data protection of objects with fully-hydrated snapshots
QoS	QoS policies are provided that optimize performance for different types of workloads
Encryption	Cohesity solution provides data-at-rest as well as data-in-flight encryption using industry standard 256-bit Advanced Encryption Standard (AES) algorithm. The platform is also FIPS 140-2 compliant
Write Once Read Many (WORM)	Enables long-term retention of data that have compliance controls mandating a policy that objects cannot be modified during the lock time
Replication for Disaster Recovery	Built-in, granular, and secure replication services for geo redundancy
Cloud integration (CloudArchive, CloudTier, CloudReplicate)	Archive into public cloud services for long-term retention. Utilize cloud tiering for transparent capacity expansion into the cloud. Replicate into the cloud for disaster recovery and test/dev

Use Cases

Overview

Cohesity file service on VMware vSAN differentiates itself from traditional scale-out NAS with "Web-Scale NAS" as a next-gen product category. As a web-scale platform, this solution offers multiple benefits including unlimited scale-out and unparalleled storage efficiency with global deduplication and compression across the cluster. This is combined with resiliency during node failure and Cohesity NAS uses SMB 3 persistent file handles and maintains locks across nodes in a cluster.

This solution can be used in the following use cases:

- Corporate File Sharing
- Archiving and Tiering
- Home Directories
- Video and Image Stores
- Cold Data – for Splunk, Hadoop, and other applications
- Cloud Native Applications

Corporate File Sharing

Corporate file shares are files made available on the corporate network. Network "shares" are used to grant access to files based on user privileges. Cohesity file share provides a multi-protocol, (NFS, SMB and S3) unified permissions cost-effective secure access to corporate and department files for Windows, Linux, and Unix environments and is

ideal for large and small files, documents, images, objects, and video files. Cohesity DataPlatform scales seamlessly and limitlessly without the fear of hitting a capacity or performance "brick-wall" – or having to do forklift upgrades. Cohesity Web-Scale NAS also dedupes across workload volumes and allows for a single file search across multiple sites and clouds.

Archiving and Tiering

File archiving - Storage of files with less frequent access are often candidates for archiving. Lower cost and/or local space savings are primary reasons for archiving.

Tiering – To move less frequently used file data to the cloud or a lower cost data store. The benefit is fast local performance for files that are more frequently accessed - and file data that is infrequently accessed can still be retrieved with acceptable performance while costing less. Tiered data can reside on a cost-effective local storage tier, or in the cloud. Cohesity DataPlatform's easy integration of archiving and tiering with the cloud provides an easy path to hybrid IT.

Home Directories

Cohesity delivers the scale to consolidate multiple NAS silos and point solutions. Sliding window variable dedupe across multiple workload volumes is very cost effective along with compression, small file optimization, and fast, efficient, no-cost clones for maximum data reduction. Rich multiprotocol support for SMB, NFS, and S3 along with native data protection and cloud integration make Cohesity Web-Scale NAS the preferred choice for file and object workloads. Exceptional security and privacy through secure long-term retention and compliance with SEC17a-4, software encryption, cluster and file-access auditing (HIPPA), multifactor authentication (MFA), integrated anti-virus protection, and ransomware protection. Taken together, this is the rigid class of security required by the financial services industry - yet available for all industries and use cases.

Video and Image

Video and image files are generated by media applications that output in video and image formats. Video and large image files consume large amounts of storage capacity - by definition. Cost-effective and secure storage for video archives, surveillance data, and large image files. Cohesity Web-Scale NAS allows for a single file search across multiple views or volumes, and even across multiple sites and clouds through Helios. Cohesity file services are compatible with DICOM images used in medical PACS and VNA systems. Securely store images on-premises or in the cloud.

Splunk and Hadoop Cloud Data

Both Splunk and Hadoop create data that must be managed from the time the data is created ("hot") until the time it becomes "cold" or "frozen". Cohesity empowers organizations to simplify the management of data and apps at web-scale, and on a single software-defined data management platform. Complementary technologies, both Cohesity and Splunk, a purpose-built platform to make machine data accessible, usable, and valuable at scale to businesses, simplify the management of massive amounts of data and help enterprises achieve their data retention goals.

Cloud Native Applications

Cloud native applications are developed with services packaged in containers, architected as microservices, deployed on elastic cloud infrastructure and continuously delivered through agile DevOps processes. This new way of developing web-scale applications brings many advantages to businesses and drives the need for change in legacy IT design, thus gaining huge popularity in the recent years. It allows the applications to fully leverage the strength of cloud computing: elastic, scalable, fault-tolerant, and highly available. Applications are developed, tested and released much more rapidly, frequently, and consistently, reducing risks and accelerating time to market. This rapid development model also brings some challenges to the IT infrastructure that supports it. One of the challenges is a resilient and scalable storage solution for storing persistent data when containers are destroyed and started and for sharing data between services. An NFS volume is very commonly used storage for these purposes. The vSAN and Cohesity joint solution provides the scalability and resiliency unmatched by the traditional NFS storage.

File Service on VMware vSAN with Cohesity

This joint solution can be deployed on any vSAN ready nodes which can be found in VMware Compatibility Guide. In this solution, we used Supermicro SuperServer 2027R-AR24NV servers with direct-attached SSDs and HDDs on ESXi hosts to provide a vSAN datastore. Each ESXi host has two disk groups consisting of one cache-tier SSD and six capacity-tier HDDs.

Each ESXi server in the vSAN cluster has the following configuration:

Table 3. ESXi Host Hardware Configuration

PROPERTY	SPECIFICATION
Server Model	Supermicro® SuperServer 2027R-AR24NV
CPU	2x Intel® Xeon® CPU E5-2690 v2 @ 3.00GHz, 10 cores each
RAM	512GB
Network Adapter	Intel® 82599EB Dual Port 10GbE, 1x Intel® I350 GbE
Storage Adapter	2x Avago® (LSI Logic) Fusion-MPT 12GSAS SAS3008 PCI-Express
Disks	2 x SAS SSD 745GB 12 x SAS HDD 1TB

The software components are listed in the table below.

Table 4. Software Components

Software	Version
VMware vSphere	6.7 U2
VMware vSAN	6.7 U2
Cohesity DataPlatform	6.3

VMware vSAN Architecture Design

Each host participating in the vSAN cluster contributes a flash device for cache and at least one magnetic device for the data. vSAN can be implemented as hybrid or all-flash cluster. In Hybrid cluster, flash devices are used for cache and magnetic disks are used for capacity. Hosts without any local storage can also participate in the vSAN cluster if there are at least 3 ESXi hosts contributing storage capacity in the cluster. Flash devices designated for caching cannot be used for capacity and do not contribute to the datastore capacity.

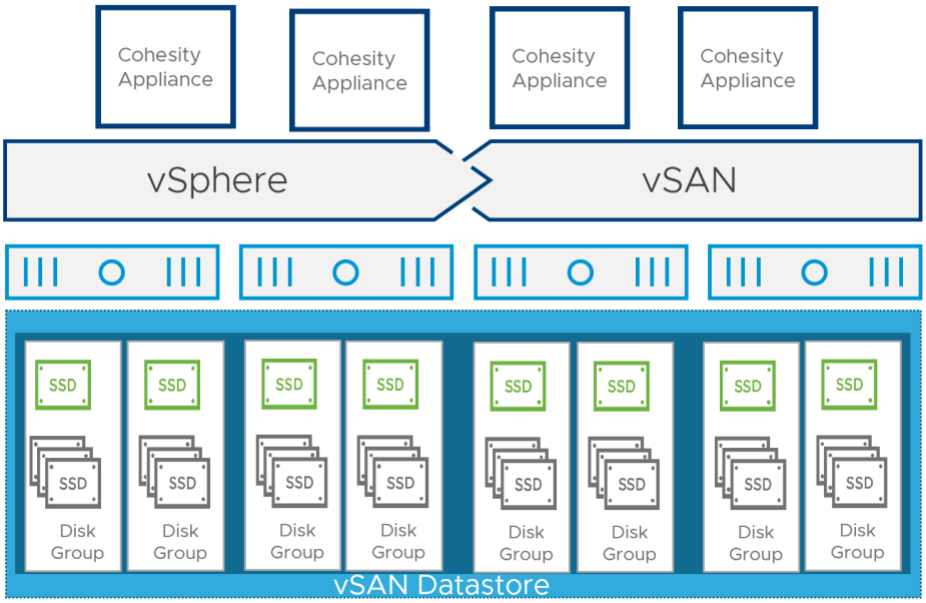

Figure 2. vSAN Architecture

vSAN Datastore
After you enable vSAN on a cluster, a single vSAN datastore is created. It appears as another type of datastore in the list of datastores that might be available, including Virtual Volume, VMFS, and NFS. A single vSAN datastore can provide different service levels for each virtual machine or each virtual disk. In vCenter Server®, storage characteristics of the vSAN datastore appear as a set of capabilities. You can reference these capabilities when defining a storage policy for virtual machines. When you later deploy virtual machines, vSAN uses this policy to place virtual machines in the optimal manner based on the requirements of each virtual machine.

Disk Group
In vSAN, a disk group is a unit of physical storage capacity on a host and a group of physical devices that provide performance and capacity to the vSAN cluster. On each ESXi host that contributes its local devices to a vSAN cluster, devices are organized into disk groups. Each disk group must have one flash cache device and one or multiple capacity devices. The devices used for caching cannot be shared across disk groups and cannot be used for other purposes. A single caching device must be dedicated to a single disk group. In hybrid clusters, flash devices are used for the cache layer and magnetic disks are used for the storage capacity layer. In an all-flash cluster, flash devices are used for both cache and capacity.

Storage Policy-Based Management (SPBM)
When you use vSAN, you can define virtual machine storage requirements, such as performance and availability, in the form of a policy. vSAN ensures that the virtual machines deployed to vSAN datastores are assigned at least one virtual machine storage policy. When you know the storage requirements of your virtual machines, you can define storage policies and assign the policies to your virtual machines. If you do not apply a storage policy when deploying virtual machines, vSAN automatically assigns a default vSAN policy with Primary level of failures to tolerate configured to one, a single disk stripe for each object, and thin provisioned virtual disk.

Object-Based Storage
vSAN stores and manages data in the form of flexible data containers called objects. An object is a logical volume that has its data and metadata distributed across the cluster. For example, every VMDK is an object, as is every snapshot. When you provision a virtual machine on a vSAN datastore, vSAN creates a set of objects comprised of multiple components for each virtual disk. It also creates the VM home namespace, which is a container object that stores all metadata files of your virtual machine. Based on the assigned virtual machine storage policy, vSAN

provisions and manages each object individually, which might also involve creating a RAID configuration for every object.

Objects and Components
Each object is composed of a set of components, determined by capabilities that are in use in the VM Storage Policy. For example, with Primary level of failures to tolerate set to 1, vSAN ensures that the protection components, such as replicas and witnesses, are placed on separate hosts in the vSAN cluster, where each replica is an object component. In addition, in the same policy, if the Number of disk stripes per object configured to two or more, vSAN also stripes the object across multiple capacity devices and each stripe is considered a component of the specified object. When needed, vSAN might also break large objects into multiple components.

Witness
A witness is a component that contains only metadata and does not contain any actual application data. It serves as a tiebreaker when a decision must be made regarding the availability of the surviving datastore components, after a potential failure.

Cohesity DataPlatform Architecture Design

Cohesity Cluster
A Cohesity cluster hosted on hardware consists of three or more nodes. A Cohesity cluster Virtual Edition consists of a single Node.

Storage Domains
A storage domain is a named storage location on a Cluster. A Storage Domain defines the policy and frequency for deduplication among other configuration such as compression, erasure coding, replication factor, and so on. A Storage Domain contains Views. When you configure a Protection Job, you specify a Storage Domain and during a Protection Job Run, the Cohesity cluster stores the Snapshots in that Storage Domain.

Views
A view provides a storage location with NFS, SMB, and S3 mount paths in a storage domain on the Cohesity cluster. The NFS mount path can be mounted onto systems using the NFS protocol. The SMB mount path can be mounted onto systems as an SMB share. You can use the view to store data such as files, backup snapshots, and cloned VMs. Views can only be mounted on any machines specified in the whitelisted subnets. A view can also act as a datastore during the VM cloning process and during VM recovery.

Cohesity DataPlatform provides the ability to create file shares that can be accessed via NFS or SMB/CIFS protocols and are called DataPlatform “Views.” These Views are members of the DataPlatform “Storage Domain”, which are logical data pools with defined storage policies for efficiency (deduplication and compression), replication factor and erasure coding, encryption, and cloud tiering.

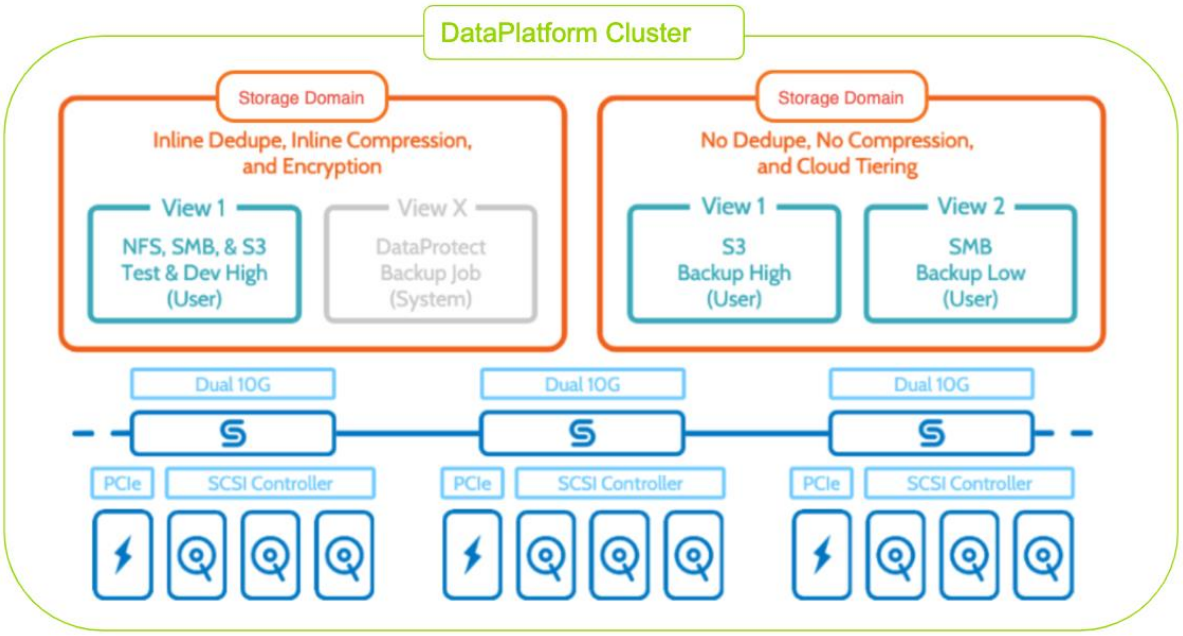

Figure 3. Cohesity DataPlatform Architecture

Cloud Integration
Long-term data and application retention are critical to organizations that seek to prevent data loss and meet security, legal, and compliance requirements. Cohesity DataPlatform provides a policy-based method to archive any VMs or files to public clouds (AWS, Azure, GCP), any S3-compatible storage, any NFS mount point. Achieving your long-term data retention and archival objectives for secondary data is simpler with Cohesity. The archived data is efficiently transferred and stored by sending only deduplicated, compressed incremental backups, thereby reducing network and storage utilization. Once the data is archived, administrators can also take advantage of Cohesity CloudRetrieve feature as a cost-effective alternative for disaster recovery, geo-redundancy, and business continuity. In the event the cluster you archived from becomes unavailable, you can retrieve your data onto a different cluster.

Cohesity integration with Cloud enables your organization to:

• Save time and lower TCO – Leverage cloud scalability for the long-term retention and archival of secondary data without cloud gateways and disparate point solutions connecting to the cloud.

• Improve efficiency – Use advanced Cohesity algorithms for true global deduplication— across clusters, workloads, and protocols and compression in the cloud to optimize capacity efficiency and lower the cost of cloud for archival.

• Derive greater value from your data – Gain fast access and retrieval of data from cloud to make data more useful to business teams seeking to uncover meaningful insights from previously untapped data.

Solution Advantages

Given that both vSAN cluster and Cohesity cluster are highly scalable and resilient, this joint solution gives the customers the best of both worlds. The file service offered by Cohesity on VMware vSAN has some unique characteristics comparing to other products.

Linear Scalability

In the vSAN architecture, storage is treated as a cluster resource, just like CPU and memory. This unique approach allows linear scalability that is not available in legacy architecture with external shared storage. For customers who want to add more storage capacity to their infrastructure, they have different ways to achieve the goal easily. First, storage performance and capacity can scale out relative to compute resource by adding more hosts to the existing cluster. Any new storage automatically joins the pool and shared by all hosts in the cluster. As many as 64 hosts per vSAN cluster are supported.

Second, new capacity disks can be added to the existing hosts to scale up storage. The new disks can be added to the existing disk groups or new disk groups can be created. As many as 35 disks and 5 disk groups are supported in each host.

Cohesity platform is a fully distributed file system spanning all nodes- SpanFS™. Cluster expansion, upgrade and remove nodes are non-disruptive. Compute and storage scale are independent.

Space Efficiency

Data deduplication is a storage efficiency feature cross volumes, cross clusters, cross protocols and cross workloads that frees up storage capacity by eliminating redundant data blocks. Different vendors implement deduplication at a file-level and a block-level of different sizes, which only works well across a single storage pool or within a single object (for example application or VM). In this solution, deduplication and compression can be achieved at both vSAN and Cohesity levels.

vSAN can perform block-level deduplication and compression to save storage space. When you enable deduplication and compression on a vSAN all-flash cluster, redundant data within each disk group is reduced. Deduplication removes redundant data blocks, whereas compression removes additional redundant data within each data block. These techniques work together to reduce the amount of space required to store the data. vSAN applies deduplication and then compression as it moves data from the cache tier to the capacity tier. When deduplication and compression is enabled on a vSAN cluster, redundant data within a particular disk group is reduced to a single copy. Deduplication and compression can be enabled when a new vSAN all-flash cluster is created or when an existing vSAN all-flash cluster is modified.

vSAN’s implementation of deduplication and compression feature makes it very simple to enable space efficiency for the whole cluster. Writes are ingested by the buffer and acknowledged back to the virtual machine prior to deduplication and compression occurring. This ensures that performance expectations are maintained. This is what is referred to as Nearline deduplication and compression. It does not have the performance penalty faced by inline approach. vSAN's deduplication occurs on a per disk group basis, avoiding strain on resources to maintain tables and naturally is more robust against exposure to global corruption should some issue arise. A significant advantage to vSANs implementation of deduplication is usage of 4KB fixed blocks. Other solutions use larger block sizes, which inherently are more challenging to find a match to deduplicate to. This yields very good effective deduplication rates to vSAN, even when deduplication occurs on a disk group basis. Compression is a conditional step for vSAN, and occurs after deduplication. If the block can be compressed from 4KB to below 2KB, then it will do so. If not, then it will remain as a 4KB block stored on persistent storage. This reduces unneeded processing for data that cannot compress very well in the first place.

Cohesity leverages a unique, variable-length data deduplication technology that spans clusters, volumes, workloads and protocols, resulting in significant savings across a customer’s entire storage footprint. With variable-length deduplication, the size is not fixed. Instead, the algorithm divides the data into chunks of varying sizes based on the data characteristics. The chunk size is varied in real time in response to the incoming data which results in greater data reduction than fixed-size deduplication. The efficiency benefit of variable-length deduplication compounds over time as additional backups are retained. Cohesity also allows customers to decide if their data should be deduplicated in-line (when the data is written to the system) or post-process (after the data is written to the system) to optimize the backup protection jobs against backup time windows. Cohesity also provides compression of the deduped blocks to further maximize space efficiency.

High Resiliency

Both vSAN and Cohesity employs multiple mechanisms to ensure resiliency for data stored in it, tolerating failures in many different levels.

vSAN administrators can leverage Storage Policy Based management to easily define desired outcome for applications, including Cohesity cluster. vSAN storage policies define the levels of protection and performance for stored objects. Policies are created and managed in virtual center and can be applied to one or more vSAN clusters. The policies can be applied to a virtual machine at any time and the virtual machine will adopt the new performance and protection settings without any down time.

Failure Tolerance Method defines how vSAN layout the data stored. There are two FTMs, Mirroring (RAID-1) and Erasure Coding (RAID 5/6), in vSAN. With RAID 1, the data is fully copied on a different host. There is also a small Witness component needed to determine quorum. This method is not as space efficient as RAID-5/6 but provides better performance. with RAID-5/RAID-6, the data (components) is striped across multiple hosts with parity information written to provide tolerance of a failure. Parity is striped across all hosts and is done inline, therefore there is no post-processing required. RAID-5 will offer a guaranteed 30% savings in capacity overhead compared to RAID-1 and RAID-6 will offer a guaranteed 50% savings.

VMware vSAN Failures to Toleration (FTT) defines the number of failures that the object can handle while still maintaining data availability (albeit with reduced redundancy). With FTM of RAID-5, FTT of 1 is implied and with FTM of RAID 6, FTT of 2 is implied. To support FTT of 3 or above, FTM can be only RAID 1/Mirroring. The default storage policy is RAID 1 with FTT of 1.

vSAN uses software checksum to detect and resolve silent disk errors. It checks both data in flight and at rest. If a checksum verification fails, vSAN fetches data from other copy if FTM is RAID 1 or rebuilds data if FTM is RAID 5/6. In addition, vSAN runs disk scrubbing in the background to proactively detect and correct data corruptions for all data.

Cohesity's highly available, fault-tolerant platform ensures organizations can continue conducting business in the event of an outage as well as during maintenance with support for rolling non-disruptive software upgrades.

The following table describes common component failures and their Cohesity remedies. During any of these failures, all operations of the cluster, such as backups and recovery, file IO, replication, and archival are unaffected. For most, the self-healer automatically heals the failure and brings the Cohesity cluster back to a normal state. The scenarios are listed in the increasing order of ease of resolution.

Table 4. Cohesity Failure Modes and Remedies

Failure	Remedy
Node Failure	Cohesity cluster can sustain up to two simultaneous node failures. Some aspects for a node or a disk failure are: • Node failure results in both file data and metadata failures. • File data is recovered from the erasure coding or replication scheme. Metadata is recovered from one of the replicas.
Capacity Disks Failure	Cohesity cluster can sustain up to two simultaneous capacity disk failures at any point in time. The number of capacity disk failures that can be tolerated depends on the RF or EC settings.
Metadata Disk Failure	Cohesity cluster can sustain up to two simultaneous metadata disk failures. Metadata disks are used to store the Distributed Metadata and the Distributed Journal and are replicated on either 2 or 3 nodes. Parts of the Distributed Metadata and Journal are stored on every node of the Cohesity cluster. A metadata disk failure can be considered the same as a node failure.
Network Failure	Each node gets its own Node IP address, a VIP address, and a hardware management IPMI address. The DNS is set-up to do round-robin, which gives each VIP a turn to satisfy client read or write requests and equally distributes them across all the nodes. **VIP Failover**: Cohesity cluster has a heartbeat mechanism that detects a node failure. If the directly connected network link goes down, the VIP of the failed node gets reassigned to another node, and node re-balancing similar to a node failure is initiated.

	For more details, see Optimal Network Designs with Cohesity.

Cohesity VIP based design is based on DNS load balancing. For high availability, Virtual IPs (VIPs) – One virtual IP is recommended for each cluster node. Cohesity enables storage I/O resiliency on the client side with this VIP based design.

To test the resilience of the solution, one node is shut down while a large file is being transferred to the file share in this test case. The transfer is not interrupted, and the persistent file handle and session is maintained with the Cohesity VIP, this is true for NFS/CIFS and S3 protocols. The cluster was working properly with the remaining two nodes and the file integrity is maintained. This proves the above statement of N+1 redundant at 3 VE nodes as forward read/write is still being made without any disruption. Test details is included in the appendix.

Deployment Guide

Best Practices

- Create a minimum of 3 virtual edition VMs to form a cluster.
- Metadata Disk Format Thick Provision Lazy Zeroed is recommended
- Configure vSphere Availability Settings for Virtual Edition for VMware
  - Turn on vSphere HA
  - Set the Host Failure Response to Restart VMs
  - Enable VM Monitoring of the heartbeat
- Use RF1 and vSAN resilience for storage saving.
- Enable deduplication and compression in both vSAN and Cohesity or at least one of them.
- Choose inline deduplication and compression for Cohesity.
- For SMB client 2.x (Windows Server 2008, Windows Server 2008 R2, Windows 7 and Windows Vista), set the HKLM\System\CurrentControlSet\Services\LanmanWorkstation\Parameters\DisableLargeMtu registry key value to 0 (zero) to enable 1 MB writes.
- On Windows 2008/Windows 2008 R2 the QoS Packet Scheduler's Limit Reservable Bandwidth default setting is not optimal. By default, it is set to 20% if not enabled or not set. For the best SMB throughput, this should be enabled and set to 100%.

Sizing Guide

To form a Cohesity cluster, virtual edition nodes need to start 3 nodes in one cluster. For a small setup, Cohesity supports single node VE. Increasing the disk sizes requires the number of vCPUs or memory to be increased. Below are two examples of how to increase capacity from a small configuration and a large configuration VE.

Table 5. Scaling up Cohesity Appliance

Configuration	Size of Metadata Disk in GB	Size of Data Tier Disk in GB	Minimum Number of Virtual CPUs	Minimum Memory in GB
SMALL Configuration Deployment Type	400 GB[1]	8,000 GB[1]	8	32
LARGE Configuration Deployment Type	800 GB[1]	16,000 GB[1]	12	51

Cohesity can co-exist with other types of workloads in the same vSAN cluster if the CPU, memory, and storage requirements listed above are met. It is not required to have a dedicated vSAN cluster for Cohesity.

The following table gives you an example of how to increase disk capacity by scaling up the disk space on the VE node or scaling out by adding VE nodes. Adding VE nodes also provides the benefits of increasing the throughput by using more nodes.

Table 6. Cohesity Cluster Sizing Guide

Proximate Raw Storage Capacity	Cohesity VE Nodes	Cohesity Meta Data Disk size	Cohesity Storage Disk Size
1TB	3	50GB	330GB
4TB	3	100GB	1.3TB
8TB	4	100GB	2TB
16TB	4	200GB	4TB
32TB	8	200GB	4TB
64TB	8	400GB	8TB
128TB	8	800GB	16TB

vSAN Storage Policy for Cohesity Cluster

General vSAN best practices applies to this solution, refer to vSAN Planning and Deployment guide. It is also a good practice to create a dedicated vSAN storage policy for Cohesity virtual appliances with the following values.

Table 7. vSAN Storage Policy for Cohesity Cluster

Setting	Value
Site disaster tolerance	None – standard cluster
Failures to tolerate	1 failure – RAID-5 (Erasure Coding)
Number of disk stripes per object	1
IOPS limit for object	0
Object space reservation	Thick provisioning
Flash read cache reservation (%)	0
Disable object checksum	OFF
Force provisioning	OFF

Cohesity Virtual Edition Cluster

DataPlatform Clustered Virtual Edition for VMware is a multiple node cohesity cluster that is hosted on virtual machines on ESXi hosts of a VMware vCenter Server.

Virtual Edition for VMware and Clustered Virtual Edition for VMware is supported on the following versions of VMware vSphere ESXi: 6.7,6.5, 6.0 and 5.5.

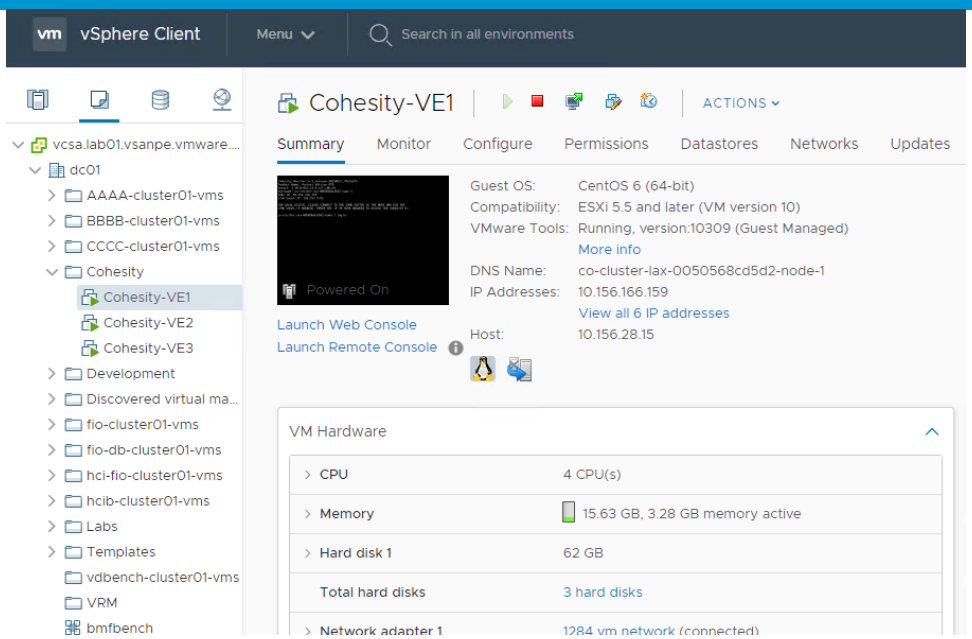

Figure 4. Deploy a Cohesity Cluster in vSphere

Privileges for Cohesity on the Source
The Cohesity cluster must perform various actions on the source. For the Cohesity cluster to perform these actions, the user specified to connect to the source (the one used to register the source) must have adequate privileges. Refer to Ensure Adequate Privileges for Cohesity on the Source for the user requirements of each hypervisor type.

Install and Configure VMware Virtual Machines for a Clustered Virtual Edition
For resilience, virtual edition nodes need to start 3 nodes in one cluster. We recommend using anti-affinity rule to place VE nodes on different ESXi hosts. This will improve redundancy and you can create anti-affinity rules to separate VE nodes.

A Clustered Virtual Edition runs on multiple VMware virtual machines. Create a Virtual Machine for each Node in the Cluster by deploying the Clustered Virtual Edition OVA file using a VMware vSphere Client. Cohesity provides the following two default deployment configurations: SMALL and Large configurations during the initial install. Virtual Edition only supports a single Metadata disk and a single Data Tier disk and required vCPUs and memories to support the workload performance. It is recommended to use SSD disks as Metadata disks. All disks should be attached to different SCSI controllers.

Table 8. Cohesity Appliance Default Configuration

VE Configuration	Size of Metadata Disk in GB	Size of Data Tier Disk in GB	Minimum Number of Virtual CPUs	Minimum Memory in GB
SMALL Default Configuration Deployment Type	50GB	60GB – 1,000GB	4	16
LARGE Default Configuration Deployment Type	400GB	8,000GB	8	32

Select the same configuration type for all nodes that will make up the Cluster. When the OVA file is deployed, the deployment process configures the memory, number of CPUs and a hard drive to store the Cohesity cluster software. In addition, you must create and attach Metadata and Data Tier disks to the Virtual Machine. The required settings for the Virtual Machine and the attached disks are summarized in the following table. Virtual Edition only supports a single Metadata disk and a single Data Tier disk. Cohesity recommends using a High IOPS Performance disk for the Operating System.

Table 9. Cohesity Appliance Deployment Settings

VE Setting	Values
Virtual Machine Name	When you create a new virtual machine, you must specify a unique Virtual Machine name for each Node in the Cluster. Each name must be unique to distinguish it from the existing VMs in the parent folder or Datacenter and must not exceed 80 characters.
Metadata Disk Format	Thick Provision Lazy Zeroed is recommended
Size of Metadata	For both the small and large configuration, Cohesity supports attaching up to two Metadata Tier disk of a size between 512 GB to 1 TB. At most, allocate 16 times more disk space to the Data Tier drive than the Metadata drive. The Metadata drive size must be smaller than the Data Tier drive size.
Mode of Metadata Disk and Data Disk	Independent - Persistent
Data Tier Disk Format	Thick Provision Lazy Zeroed is recommended
Size of Data Tier Disk	For the small configuration, Cohesity supports attaching one Data Tier disk with size between 1 TB to 8 TB. For the large configuration, Cohesity supports attaching two Data Tier disk with a size between 1 TB to 8 TB. You can choose to use disks of different sizes within the above constraints.

For the detailed steps on how to create a cluster, see document
https://docs.cohesity.com/6_3/Web/PDFs/SetupGuideVirtualEditionVMware.pdf

Storage View

To create a file share, first you need to create a storage view or use an existing storage view.

A view provides a storage location with NFS and SMB mount paths in a storage domain on the Cohesity cluster. The NFS mount path can be mounted onto systems using the NFS V3 protocol. The SMB mount path can be mounted onto systems as an SMB2 and SMB3 share. You can use the view to store data such as files, backup snapshots, and cloned VMs. Views can only be mounted on any machines specified in the whitelisted subnets. A view can also act as a datastore during the VM cloning process and during VM recovery.

1. In the Cohesity Dashboard, select Platform > Views.
2. To create a new view, click Create View at the top right of the page. To edit an existing view, click the actions menu and select Edit.
3. If you are creating or cloning a new view, specify a name for the view.

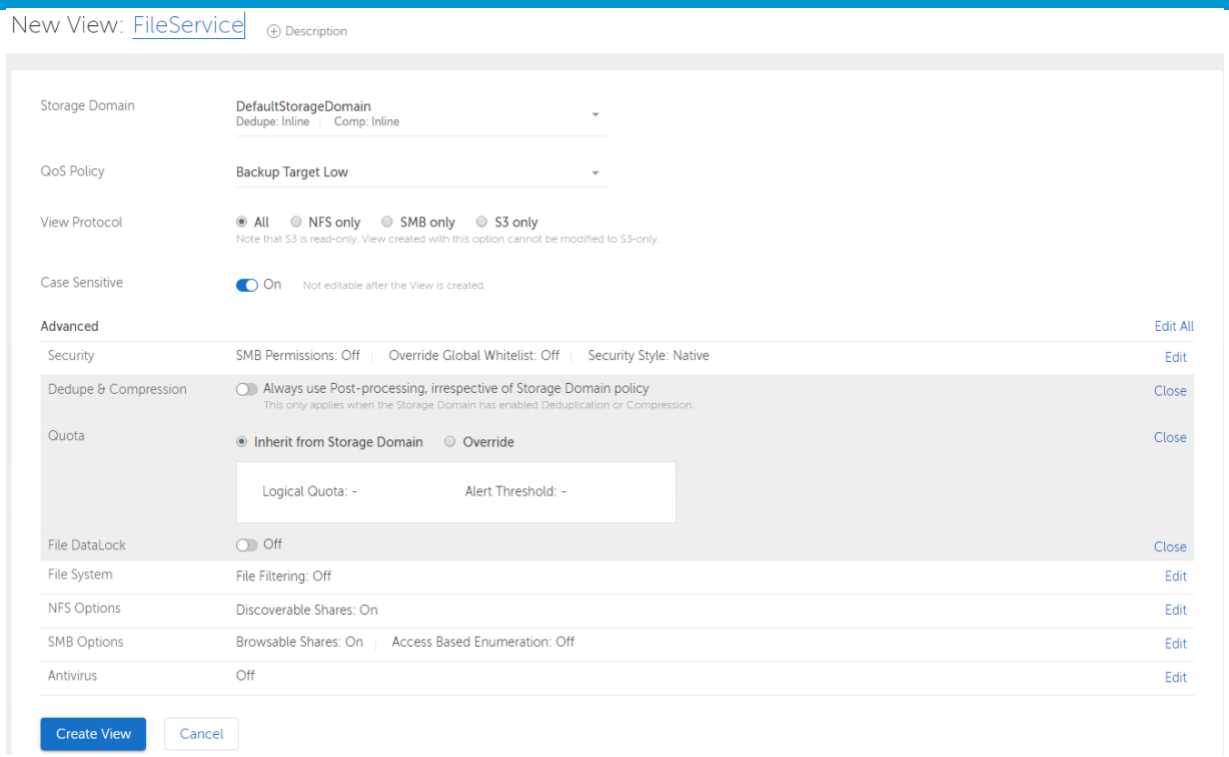

Figure 5. Cohesity Storage View

Understand View and Storage Domain Settings

A Cohesity View is created inside a storage domain. At the storage domain level, settings like fault tolerance, encryption, and storage efficiencies are applied.

Table 10 summarizes the features at the storage domain and view levels, and how they are inherited.

Table 10. Features at the Storage Domain and View Levels

Feature	Storage Domain	View	Recommendation
Fault Tolerance	Erasure Coding (EC) Replication Factor (RF)	Inherits storage domain fault tolerance settings and cannot be changed for the View.	RF1 EC2:1
Storage Efficiency	Deduplication (inline, post-process, or disabled) Compression (inline, post-process, or disabled)	Inherits storage domain efficiencies. OR Can be overwritten at the View level to be post process if at storage domain level is inline but cannot be disabled.	Enable Deduplication and Compression for storage saving.
Encryption	Enable or disable	Inherits storage domain's settings and cannot be changed for the View.	Use as needed.
QoS Policy	N/A	Set for each View.	Use TestAndDev in most cases

Quota	N/A	User quotas limit how much view capacity a user can use. When a user reaches 85% of the quota, an Alert (UserExceededQuotaAlertLimit) is triggered in the Cohesity cluster.	Set a quota based on a customer's requirement.

QoS Policy

Each Cohesity view is assigned a Quality of Service (QoS) Policy. A QoS Policy determines to which storage media data is written and the priority of I/O when contention occurs. Select an appropriate QoS policy for the view. There are two basic QoS policies, TestAndDev and Backup Target, each of which has variants by priority and storage media.

Table 11. Cohesity QoS Policy

QoS Policy	Optimized for I/O Workload Type	Priority ↓	Storage Media
TestAndDev	Random reads and writes (for NFS, SMB, Cohesity Views)	High Low	SSD
Backup Target	Sequential reads and writes (for backups using Cohesity DataProtect)	SSD High Low	SSD HDD

TestAndDev is the recommended QoS policy unless the VE is used for a backup target with large number of sequential streams.

Global Whitelists

To have privileges to mount and access a View, a system's IP Address must be specified in a Subnet that has been added to a Cohesity Whitelist. See diagram below:

Figure 6. Cohesity Whitelist

Role-Based Access Control

Simplify user and group access to data utilizing credentials and permissions with Windows AD and Kerberos mechanisms. Create and manage custom Cohesity cluster administration roles for domain users and groups. Cohesity can have two types of users: local and Active Directory principals. Both user types can log in to the Cohesity Dashboard to manage data on the Cohesity cluster and view statistics. Local users exist on the Cohesity cluster only. When the cluster joined an AD domain, you can assign roles to principals in the AD or the AD's trusted domains, which determines the cluster access.

Table 12. Cohesity Default System Roles

Role	Description
Admin	Users with the Admin role have full access to all actions and workflow within the Cohesity UI and Cohesity CLI except for changing DataLock Views and setting their expiration dates. For example, a user assigned the Admin role can create backups, recover Snapshots and files from backups, add sources, delete sources, set up and manage the Cohesity cluster, manage nodes, manage Views, resolve Alerts. A user with the Admin role can also add and delete users.
Operator	Users with the operator role have viewer role privileges and can run existing protection jobs and create recover tasks.
Viewer	Users with the Viewer role have read-only access for all workflows within the Cohesity UI.
SMB Security	Users with the SMB Security role have all SMB privileges when accessing data over SMB.
Self Service Data Protection	Users with the self-service data protection role have Viewer role privileges and can manage clones and protection jobs and policies, as well as create recover tasks.
Data Security	Users with the Data Security role have viewer role privileges and can create DataLock views and set DataLock expiration dates.

Security Style

Cohesity DataPlatform provides the following security styles:

Table 13. Cohesity Security Styles

Security Style	Description
Native	Windows and Linux permissions are stored separately.
Unified	Windows and Linux permissions are kept consistent such that when one changes, the other is modified automatically.
NTFS	Access via SMB and NFS is controlled by an SMB ACL present on the file.

File DataLock

DataLock is the DataPlatform WORM (write once read many) feature that locks and retains files in a view for compliance and regulatory purposes. File level DataLock is typically used for compliance and regulatory purposes. After File DataLock is enabled, it cannot be disabled by anyone; however, a Cohesity user with Data Security privileges can change the settings. Cohesity recommends adding at least one user with Data Security privileges.

Antivirus Scanning

With Cohesity cluster, you can enable antivirus software to protect your data on the Cohesity cluster from any malicious virus. Once an antivirus software scans a file and detects a threat, Cohesity cluster reports the threat on the Cohesity Dashboard. The administrators can either quarantine, unquarantine, reset, or delete the virus infected file.

File Share

A share is used to mount a View on a user's system. Cohesity automatically creates a default share for each View in the Cluster, using the View name as the share name. The default share provides access to the root View. To see a list of all the shares in the Cluster, select **Platform > Views** and click the **Shares** tab.
After a storageview is create, next step is to create a file share.

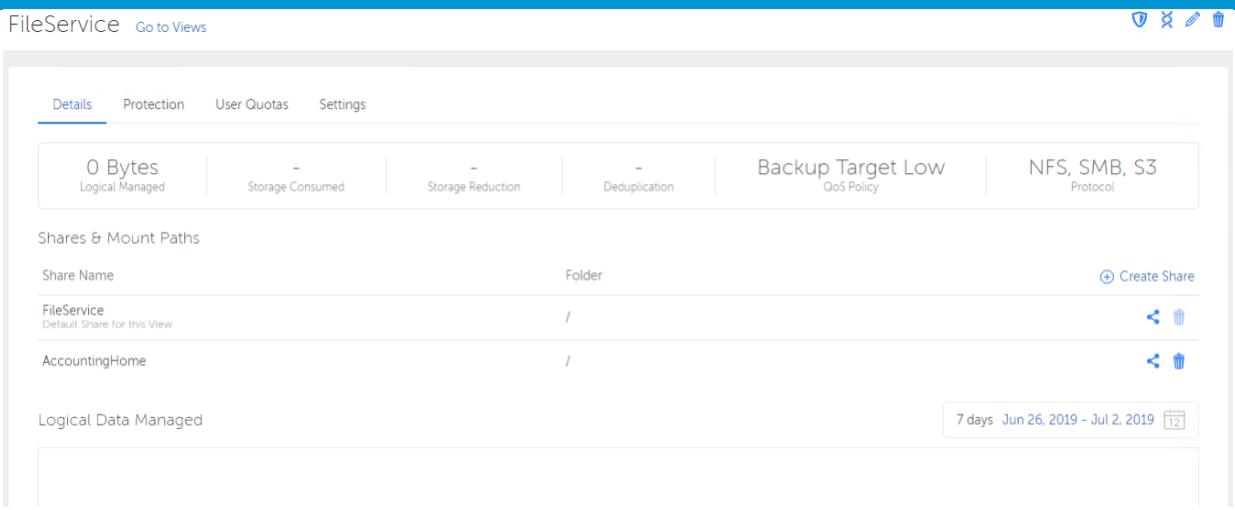

Figure 7. Cohesity File Share

Create a File Share

Perform the following procedures to create a file share:

1. In the Cohesity Cluster, select Platform **> Views**.
2. Click the highlighted View name to display the details page.
3. Click **Create Share**. The root/directory and any sub-directories in the View are displayed.
4. You can create the share on the root directory and click a directory and any sub-directories until the **Preview Panel** displays the path for which you want to create the share.
5. The **Share Name** field displays a modifiable name, based on the final directory in the path. Change the name as needed. The name must be unique within the cluster.
6. Click **Create Share**.
7. Mount points for the share are shown below.

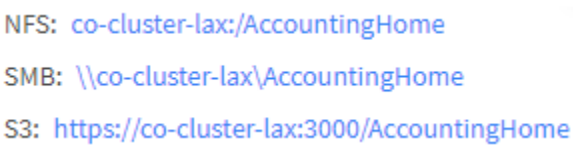

Figure 8. Cohesity File Share Mount Points

Mount a View Using NFS

On Linux client, follow the steps below to mount a view:

1. Install the nfs-common and rpcbind modules

```
Ubuntu: sudo apt-get install rpcbind nfs-common -y
CentOS: sudo yum install rpcbind nfs-common -y
```

2. Create a directory to mount the share

```
mount -t nfs -o
noatime,vers=3,proto=tcp,rsize=1048576,wsize=1048576,timeo=10000,hard,intr,nolock <VIP
IP Address of Cluster>:/<NFS_View_Name> <View Mount Path>
```

3. Make the mount persistent by adding an entry in /etc/fstab

```
echo coh01.cohesity.com:/NFS1 /home/cohesity_user1/nfsdir
noatime,vers=3,proto=tcp,rsize=1048576,wsize=1048576,timeo=10000,hard,intr,nolock 0 0 >>
/etc/fstab
```

Cohesity supports NFS share Windows client, check Cohesity TechDoc for details.

Mount a View with SMB

Cohesity DataPlatform supports SMB 3.0 and SMB 2.x.
To mount a View on a Windows system, you can use the Windows command prompt to run the following command:

```
net use <Drive Letter>: \\<View Mount Path>
```

If using Windows PowerShell, you do not have to mount the View. You can change the directory directly to a UNC path as long as your Windows login has privileges for the remote share. For example, you can do the following:

```
cd \\<View Mount Path>
```

Cohesity SMB Authentication

On Cohesity clusters that are not joined to AD, all Views are accessible as SMB shares and there is no authentication. However, if the Active Directory administrator blocks NTLM, the shares would be inaccessible.
After the Cluster is joined to an AD domain, users in Active Directory can access Cohesity cluster SMB shares based on their Active Directory credentials. Joining the Cohesity cluster to an AD domain is optional, and all accounts and Cluster SMB share privileges for users in Active Directory are managed through Active Directory or where the SMB shares are mounted. Non-admin AD users must be granted backup and restore privileges on Windows clients explicitly.

Cohesity SMB Share Access

Client access to an Active Directory-joined Cohesity cluster SMB share consists of the following actions:

1. A client requests access to a Cohesity cluster SMB share.
2. The client attempts to perform an operation on the Cohesity cluster SMB share.
3. Active Directory authenticates the client's membership within a group that can perform the requested operations on the Cohesity cluster SMB share.

Enable Active Directory authentication for SMB shares. Specify whitelisted subnets that define ranges of IP addresses for systems that have privileges to access all views in the Cohesity cluster.

Cohesity S3 Integration

Cohesity provides a S3-compartible storage platform in addition to NFS and SMB. Cohesity DataPlatform, an enterprise-ready, scale-out S3-compatible hyperconverged secondary storage solution provides unlimited scalability. Cohesity supports both Signature Version 2 and Version 4.

Cohesity S3 benefits:

- S3 compatible object storage with REST APIs for easy integration with apps and services
- Converged on Cohesity DataPlatform that can be used to consolidate objects, files, data protection, and test/dev copies
- Web-scale platform with unlimited scalability, nondisruptive upgrades, pay-as-you-grow scalability
- Industry’s only globally deduplicated, S3 object storage
- Global Google-like search on all object metadata
- Public cloud integration for archival, tiering and replication

Create a S3 Storage View on Cohesity Platform

1. In the Cohesity Dashboard, select **Platform** > **Views**.
2. Create View and select Protocol **S3 only**.
3. Click on **Create View.**

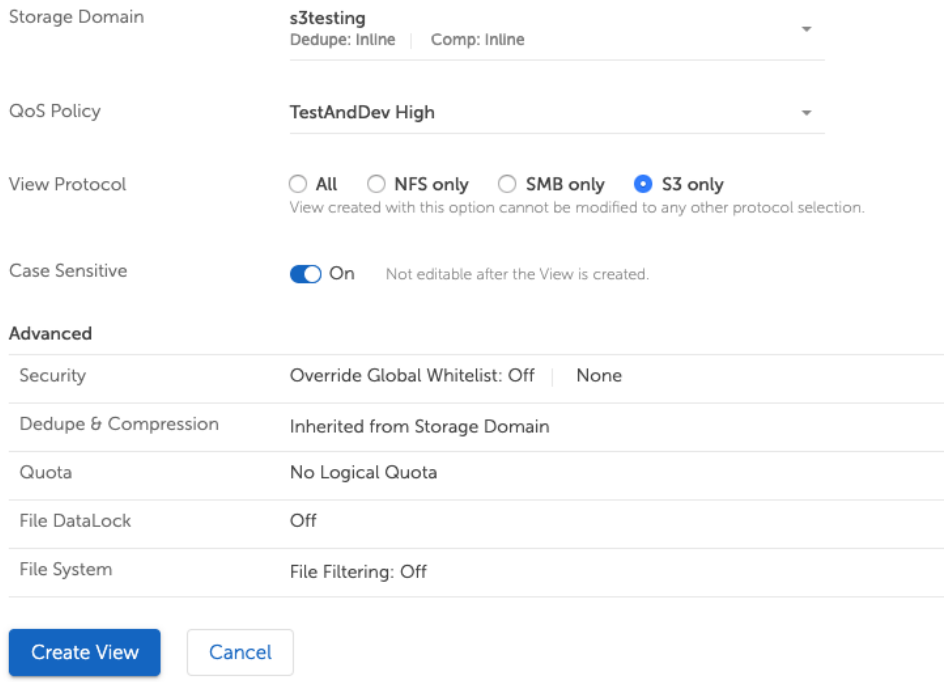

Figure 8. Create a Cohesity S3 Storage View

Use the AWS CLI to access Cohesity S3 buckets

1. Download and install AWS CLI. You can also check the S3 command sets and Cohesity support details.
   - AWS CLI: http://docs.aws.amazon.com/cli/latest/userguide/installing.html
   - S3 command sets: http://docs.aws.amazon.com/cli/latest/reference/s3api/index.html#cli-aws-s3api
   - Cohesity cluster S3 support details: https://docs.cohesity.com/6_3/Web/UserGuide/Content/Dashboard/Admin/S3.htm?Highlight=s3#Support
2. Get the access key and secret key from the command output below:
   Click on **Admin** > **Access Management**
3. From the endpoint, use **<Hostname or IP Address>:3000** where you can find the hostname or IP address from Cohesity DataPlatform **Platform**->**Networking** page.

Protection

Cohesity provides the ability to protect the data in a view by capturing snapshots of the view. A View Protection Job captures View Snapshots according to the protection schedule defined by the associated View Protection Policy. After View Snapshots are captured, these Snapshots can be used to complete the following tasks:

- The Snapshot can be cloned into a new View that is stored in the same Storage Domain as the original Snapshot. The new View contains the version of the files and directories as they were captured when the original Snapshot was created.
- The Snapshot can be replicated (copied) to another storage domain on a remote cluster. The replication schedule is defined by the View Protection Policy associated with the Protection Job.

Data can also be archived to tape and cloud.

File Share Restore

You can restore the previous versions of files and folders that are in Cohesity-protected SMB Views. This feature enables you to perform the restore from the system where the Views are mounted and provides the same functionality as the Windows "Restore previous versions" function.
Note: The file or folder you want to restore must have already been backed up by a previous View Snapshot on the Cohesity cluster. **To restore a previous version:**

1. On the system that contains the file or folder you want to restore, right-click the item and select **Restore previous versions**.
2. Click the **Previous Versions** tab.
3. Select the version you want to restore to and click **Restore**.

Conclusion

This combined solution can bring agility for enterprises by simplification of their infrastructure stack. The solution brings the core benefits of HCI such as simplicity, scalability, and resiliency for both primary and secondary storage platforms. This allows enterprises to focus on their business rather than spending time and effort to manage infrastructure. With the consolidation of data protection, file services and disaster recovery operations, the OPEX for managing this next generation architecture is reduced while providing more values with less overheads.

VMware vSAN is a hyperconverged, software defined storage platform for running virtualized workloads and Cohesity DataPlatform provides file services workloads with multiprotocol access (NFS, SMB/CIFS, and S3) with unified permissions on a single platform that spans from core, to edge, and into the cloud. This validated solution provides the best of the both worlds, we create a solution that is resilient and scalable in many different levels.

Appendix: Resiliency Test

1. Check VIP DNS load balancing configuration.

```
cohesity@view-damien-ubu:~$ nslookup sv16-mkt-vsan-ready01-vip
Server:         10.2.1.1
Address:    10.2.1.1#53

Non-authoritative answer:
Name:    sv16-mkt-vsan-ready01-vip.eng.cohesity.com
Address: 10.2.146.92
Name:    sv16-mkt-vsan-ready01-vip.eng.cohesity.com
Address: 10.2.146.91
Name:    sv16-mkt-vsan-ready01-vip.eng.cohesity.com
Address: 10.2.146.90
Name:    sv16-mkt-vsan-ready01-vip.eng.cohesity.com
Address: 10.2.146.89
```

2. VIP is distributed evenly on each Cohesity node. As seen below each node has 2 IP addresses assigned. One address is the bond0 interface and the other is VIP.

```
[cohesity@test-vsan-ready-node-1 ~]$ allssh.sh ip a | grep 10.2.146
=========== 10.2.146.57 ===========
    inet 10.2.146.57/20 brd 10.2.159.255 scope global bond0
    inet 10.2.146.89/20 brd 10.2.159.255 scope global secondary bond0:1
=========== 10.2.146.59 ===========
    inet 10.2.146.59/20 brd 10.2.159.255 scope global bond0
    inet 10.2.146.91/20 brd 10.2.159.255 scope global secondary bond0:1
=========== 10.2.146.61 ===========
    inet 10.2.146.61/20 brd 10.2.159.255 scope global bond0
    inet 10.2.146.90/20 brd 10.2.159.255 scope global secondary bond0:2
=========== 10.2.146.55 ===========
    inet 10.2.146.55/20 brd 10.2.159.255 scope global bond0
    inet 10.2.146.92/20 brd 10.2.159.255 scope global secondary bond0:2
```

3. The Linux client is using the VIP hostname to mount the NFS storage from Cohesity DataPlatform.

```
[root@localhost ~]# mount -t nfs
sv16-mkt-vsan-ready01-vip:/phy-fs1 on /nfs/nfs01 type nfs
(rw,relatime,vers=3,rsize=1048576,wsize=1048576,namlen=255,hard,proto=tcp,timeo=600,retr
ans=2,sec=sys,mountaddr=10.2.146.90,mountvers=3,mountport=2049,mountproto=tcp,local_lock
=none,addr=10.2.146.90)
```

4. Run file server workloads using FileBench to simulate file server traffic from a Linux client.

```
[root@localhost nfs01]# filebench
Filebench Version 1.4.9.1
9868: 0.000: Allocated 170MB of shared memory
filebench> load fileserver
9868: 7.524: File-server Version 3.0 personality successfully loaded
9868: 7.524: Usage: set $dir=<dir>
9868: 7.524:        set $meanfilesize=<size>     defaults to 131072
9868: 7.524:        set $nfiles=<value>      defaults to 10000
9868: 7.524:        set $nthreads=<value>    defaults to 50
9868: 7.524:        set $meanappendsize=<value>  defaults to 16384
9868: 7.524:        set $iosize=<size>  defaults to 1048576
```

```
9868: 7.524:        set $meandirwidth=<size> defaults to 20
9868: 7.524:        run runtime (e.g. run 60)
filebench> set $dir=/nfs/nfs01
filebench> run 3600
9868: 65.439: Creating/pre-allocating files and filesets
9868: 65.454: Fileset bigfileset: 10000 files, 0 leafdirs, avg dir width = 20, avg dir
depth = 3.1, 1240.757MB
9868: 269.641: Removed any existing fileset bigfileset in 205 seconds
9868: 269.641: making tree for filset /nfs/nfs01/bigfileset
9868: 278.948: Creating fileset bigfileset...
9868: 462.995: Preallocated 7979 of 10000 of fileset bigfileset in 185 seconds
9868: 462.995: waiting for fileset pre-allocation to finish
10021: 462.995: Starting 1 filereader instances
10022: 462.997: Starting 50 filereaderthread threads
9868: 464.129: Running...
```

5. Shutdown one node in the cluster to simulate an outage.

```
[cohesity@cohesity-hxvsan-bqkp50545261-node-2 ~]$ iris_cli cluster status
CLUSTER ID                    : 4753914011867584
CLUSTER NAME                  : cohesity-HXvSAN
CLUSTER INCARNATION ID        : 1537769063609
SERVICE STATE SYNC            : DONE
CLUSTER ACTIVE OPERATION      :
CLUSTER HEAL STATUS           : NORMAL

NODE ID                       : 130593348060
NODE IPS                      : 10.2.146.57


NODE ID                       : 130593348052
NODE IPS                      : 10.2.146.59


NODE ID                       : 130593348032
NODE IPS                      : 10.2.146.55
SOFTWARE VERSION              :
ACTIVE OPERATION              :
MESSAGE                       : Request timed out.
NODE STATUS                   : DOWN

NODE ID                       : 130593341976
NODE IPS                      : 10.2.146.61
```

6. Check resilience:
   a. File transfer is still going.
   b. Two VIPs are on one VE.

```
[cohesity@cohesity-hxvsan-bqkp50545261-node-2 ~]$ allssh.sh ip a | grep 10.2.146
=========== 10.2.146.55 ===========
ssh: connect to host 10.2.146.55 port 22: Connection timed out
=========== 10.2.146.59 ===========
    inet 10.2.146.59/20 brd 10.2.159.255 scope global bond0

    inet 10.2.146.91/20 brd 10.2.159.255 scope global secondary bond0:1
    inet 10.2.146.92/20 brd 10.2.159.255 scope global secondary bond0:136
```

```
========== 10.2.146.61 ==========
    inet 10.2.146.61/20 brd 10.2.159.255 scope global bond0

    inet 10.2.146.90/20 brd 10.2.159.255 scope global secondary bond0:2
========== 10.2.146.57 ==========
    inet 10.2.146.57/20 brd 10.2.159.255 scope global bond0
    inet 10.2.146.89/20 brd 10.2.159.255 scope global secondary bond0:1
```

c. The failed node VIP can still be reached.

```
cohesity@view-damien-ubu:~$ ping sv16-mkt-vsan-ready01-vip
PING sv16-mkt-vsan-ready01-vip.eng.cohesity.com (10.2.146.92) 56(84) bytes of data.
64 bytes from 10.2.146.92: icmp_seq=1 ttl=63 time=0.183 ms
64 bytes from 10.2.146.92: icmp_seq=2 ttl=63 time=0.242 ms
64 bytes from 10.2.146.92: icmp_seq=3 ttl=63 time=0.156 ms
64 bytes from 10.2.146.92: icmp_seq=4 ttl=63 time=0.193 ms
```

d. All file shares are still accessible.

Reference

See more vSAN details and customer stories:

- vSAN
- Storagehub
- VMware Compatibility Guide
- vSAN Planning and Deployment
- Cohesity Docs

About the Authors

Jia Dai, Senior Solutions Architect in the Product Enablement team of the Hyperconverged Infrastructure Business Unit wrote the original version of this paper.

Rachel Zhu, Technology Evangelist at Cohesity, A technology evangelist, a solution architect and a product manager focus on next-generation technologies and revenue growth.

Damien Philip, Principal Solutions Architect with a demonstrated history of building solutions with industry technology partners and helping solve data center challenges for Enterprises.

Other essential contributors included:

- Vibhor Guptor, Cohesity Product Manager
- Ganesh Shanmuganathan, Cohesity Principal Engineer
- Scott Owens, Cohesity Technical Marketing Engineer

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com
Copyright © 2019 VMware, Inc. All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws. VMware products are covered by one or more patents listed at http://www.vmware.com/go/patents. VMware is a registered trademark or trademark of VMware, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

©Cohesity, Inc. 2019. All Rights Reserved. This document is for informational purposes only and Cohesity, Inc. assumes no responsibility for any inaccuracies. Cohesity, Inc. reserves the right to modify this publication without notice.