WHITE PAPER – **AUGUST 2017**

ENTERPRISE PROTECTION AND RECOVERABILITY FOR VMWARE CLOUD FOUNDATION (VCF) WITH COHESITY

vmware®

Table of Contents

Executive Summary

With the Software-Defined Data Center (SDDC), VMware laid out the vision for the architecture of the hybrid cloud. SDDC redefines the architecture and operational model of the data center, enabling IT to complete the transition to hybrid cloud and maximize its benefits. SDDC aims to decouple compute, storage, and networking services from underlying hardware infrastructure and abstract them into logical pools of resources that can be more flexibly provisioned and managed.

To accelerate the customer journey to SDDC, VMware has introduced VMware Cloud Foundation™, a new unified SDDC platform for the private and public cloud. Cloud Foundation brings together VMware's compute, storage, and network virtualization into a natively integrated stack that can be deployed on premises or run as a service from the public cloud.

For this white paper, VMware and Cohesity worked closely together as technology partners to deliver a joint solution. The solution consists of extending VMware Cloud Foundation with complementary data protection and business continuity technology from Cohesity. The core pillars of this SDDC solution are resiliency, high availability, and effortless data protection. This combined solution will enable organizations to achieve greater business agility and scalability and protect their investment while supporting the next generation of applications.

While public cloud services can be a good fit for many applications, private data centers continue to play a critical role, especially for those mission-critical applications that require greater control and security. As a result, organizations are looking to shift to a more agile, service-oriented IT model that leverages both private and public clouds. Cloud Foundation allows CIOs to enable their business to achieve the operational and cost efficiency of web-scale cloud service providers while maintaining the control they need.

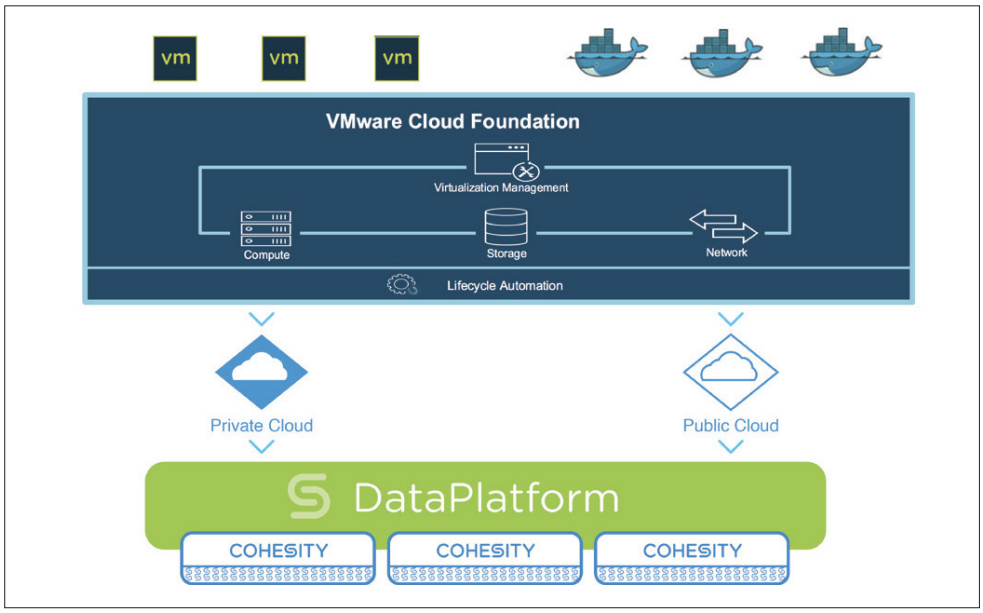


Figure 1: VMware Cloud Foundation and Cohesity Joint Solution

Cohesity provides simple data protection and instant recovery for Cloud Foundation. With Cohesity, organizations can recover an entire private cloud environment in just a few seconds. The joint solution enables organizations to recover both their cloud management functions as well as their applications running in the cloud.

Together, VMware and Cohesity are delivering a software-defined, web-scale solution that spans both primary and secondary storage and private and public clouds.

Evolution of VMware's Software-Defined Data Center (SDDC)

To respond to the ever-increasing demand for faster innovation, organizations are looking to shift to a more agile, service-oriented IT model that leverages both private and public clouds.

Highly dynamic, agile, available, and programmatic compute, network, storage, and security services are no longer a business advantage, but are simply "table stakes" to remain competitive. While customers recognize the need to complete the journey to the hybrid cloud and SDDC, they are faced with significant challenges:

- Manage and control diverse infrastructure which creates operational complexity
- Improve security to face cybersecurity threats
- Deliver enterprise-level SLA to mission-critical apps while keeping cost under control
- Manage public cloud sprawl driven by shadow IT
- Manage risk and cost by avoiding vendor or cloud lock-in

VMware Cloud Foundation Vision and Strategy

VMware Cloud Foundation's vision and strategy for IT is centered around simplifying the journey to the hybrid cloud that it aims to deliver.

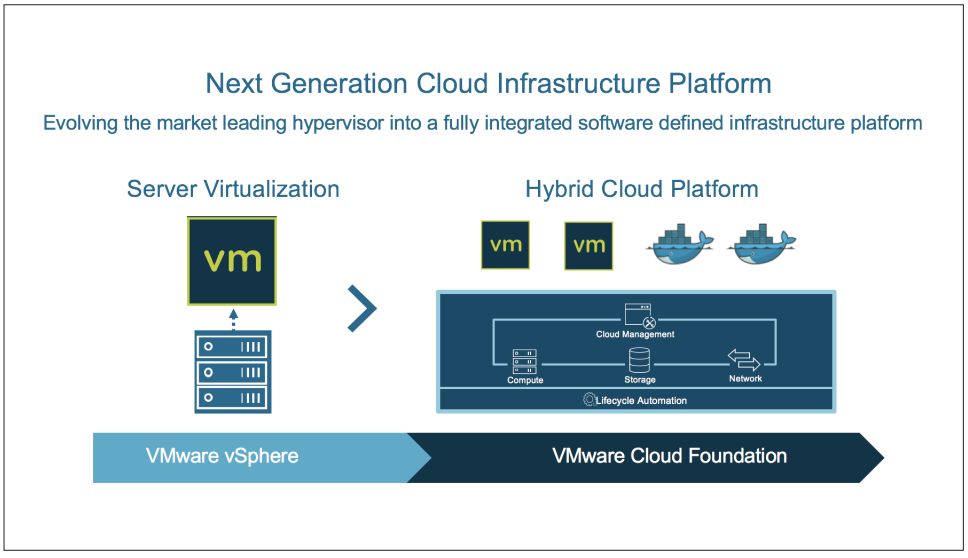


Figure 2: Evolution of VMware Cloud Foundation

1. Modernize data centers

To build a private cloud, IT needs to:

a) modernize infrastructure by virtualizing compute, storage, network, and security with a software-defined approach that removes dependencies from underlying hardware.

b) automate the delivery and ongoing management of the virtualized infrastructure, enabling end users to consume it as a service at scale for any application. VMware Cloud Foundation delivers a complete, enterprise-ready private cloud by integrating the cloud infrastructure with cloud management into a single platform that is simple to operate with built-in automated lifecycle management.

2. Extension to the public cloud

Give customers the flexibility to run this same platform in their private cloud on premises or to consume it as a service through public cloud partners, leveraging a common foundation that delivers a consistent operational model across private and public and can be managed using existing skill set and processes.

Today, Cloud Foundation is available as a service on IBM Bluemix and will power the service delivered on VMC on AWS, Centurylink, and Rackspace.

3. Enable true hybrid cloud

In the near future, integrate Cloud Foundation's cloud infrastructure with VMware's cloud management to create a single hybrid cloud platform that delivers common management, policies, networking, and security across your private and public clouds, removing cloud silos and offering you the ultimate in cloud flexibility and freedom, ready for traditional and containerized applications.

Introduction to VMware Cloud Foundation (VCF)

VMware Cloud Foundation is the industry's most advanced cloud infrastructure platform. It provides a complete set of software-defined services for compute, storage, networking, and security to run enterprise apps—traditional or containerized—in private or public environments. Cloud Foundation drastically simplifies the path to the hybrid cloud by delivering a single, integrated solution that is easy to operate thanks to new built-in automated lifecycle management.

Cloud Foundation is the next generation VMware cloud infrastructure platform. It evolves VMware's market-leading server virtualization, vSphere, by extending the core hypervisor with integrated, software-defined storage, networking, and security capabilities that can be consumed flexibly on premises or as a service in the public cloud (VMware Cloud on AWS, VMware Cloud Foundation on IBM Bluemix, vCloud Air Network).

When coupled with a cloud management platform (vRealize Suite), the end result is a hybrid cloud platform than can span private and public environments, offering a consistent operational model based on well-known vSphere tools and processes, and freedom to run applications anywhere without the complexity of application re-writing.

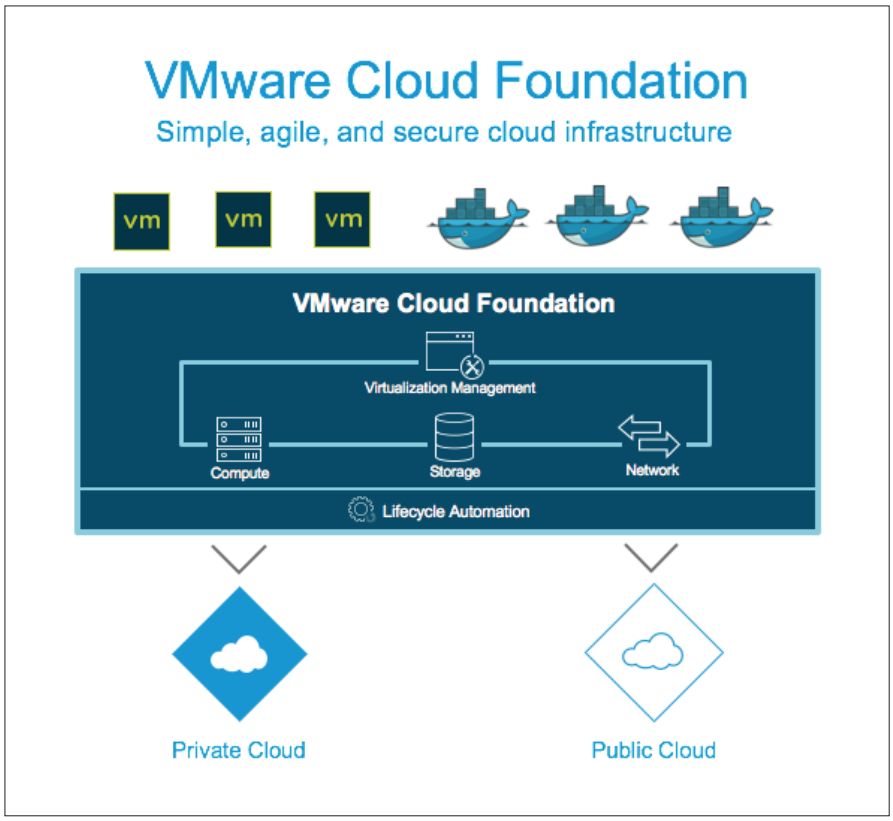


Figure 3: VMware Cloud Foundation

VCF is engineered to be simple to operate by VMware system engineers and unlocks greater agility and productivity by eliminating the operational bottlenecks of legacy infrastructure. VCF accomplishes this by delivering:

- Integrated Stack—Engineered integration into a single solution for the entire software-defined stack with guaranteed interoperability. No more complex interoperability matrixes to deal with.
- Standardized Architecture—VCF resources are automatically deployed using a standard architecture that is based on VMware's Validated Design. This ensures quick, repeatable deployments while eliminating risk of faulty configurations.
- Automated Lifecycle Management—VCF includes unique lifecycle management services that automate day 0 to day 2 operations from bring up to configuration, resource provisioning, and updates/patches.

Cloud Foundation-based infrastructure is entirely defined in software and so inherently highly dynamic, scalable, and hardware independent. Underlying physical hardware is completely abstracted into logical pools that can be flexibly allocated to individual tenants or applications.

Deployments can start from as little as four nodes and non-disruptively scaled to hundreds of nodes with linear, predictable scalability of capacity and performance. Service levels can be dynamically controlled through policy at the VM level.

Components in the VMware Cloud Foundation (VCF) Software Stack

For the case of private cloud deployments, the Cloud Foundation stack includes VMware vSphere, vSAN, NSX, and VMware SDDC Manager. Customers that possess unused licenses of these software components (vSphere, vSAN, or NSX) can bring them to the environment and acquire only the missing components to complete the licensing of the Cloud Foundation stack.

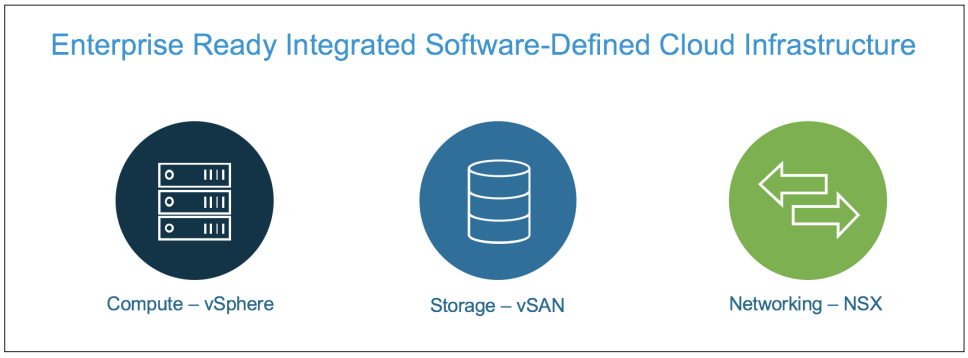


Figure 4: Components of VMware Cloud Foundation

SDDC Manager is a new innovative system management automation component specifically developed by VMware for Cloud Foundation.

SDDC Manager complements the existing suite of VMware management software, namely VMware vCenter Server® and vRealize Suite, to provide simplified operational experience at the system level, allowing customers to manage a highly distributed architecture as a single logical system.

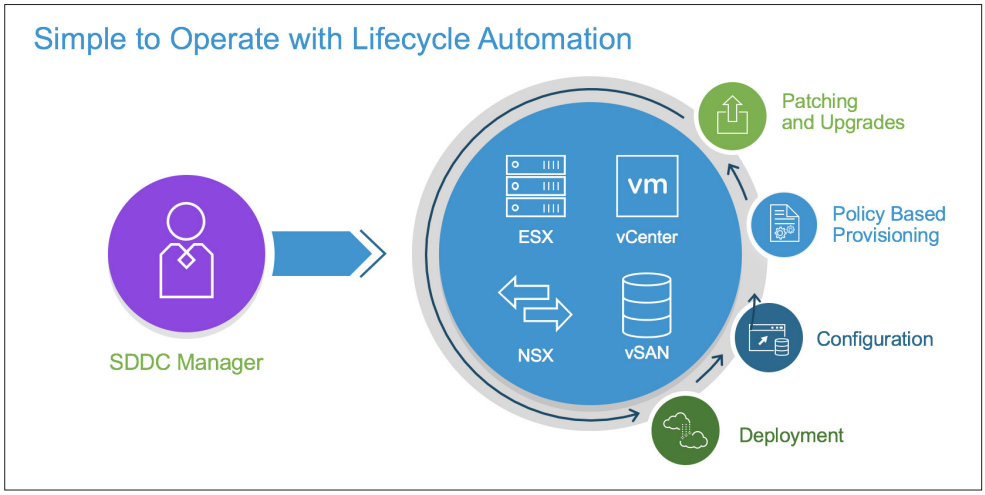


Figure 5: Lifecycle Automation of VMware Cloud Foundation

Cloud Foundation integrates with the VMware stack, including VMware's virtual desktop solution, Horizon, and cloud management platform, vRealize Suite.

When these additional components are deployed on top of Cloud Foundation, the SDDC Manager can extend its lifecycle automation to the Horizon Suite, vRealize Operations Manager, and vRealize Log Insight, with more software components of the management suite to be integrated soon.

Resources	Supported services and capabilities
Compute	• Hypervisor (vSphere) – Live migrations – High availability – Dynamic load balancing
Storage	• Hyper-converged object store (vSAN) – All flash or hybrid – High-performance linear scalability – Data efficiency (dedupe/compression) – Data protection (snaps)
Networking	• Network Virtualization (NSX) – Distributed switching and routing – Load balancing – L4-L7 network services – Remote gateway
Security	• Security (NSX, vSAN, vSphere) – Micro-segmentation – Distributed firewall – VPN – Compliance – vSphere encryption – vSAN encryption
Core Virtualization Management	• Intelligent Operations (vRealize log Insight/vRealize Operations) – Monitoring – Capacity planning – Troubleshooting – Log analytics and performance
IT Automation (optional)	• IaaS (vRealize Automation) – Service composition and automation – Self-service portal and APIs – Governance – Security policy – Approval policy – Provisioning automation
VDI (optional)	• Virtual desktops (Horizon)

VMware Cloud Foundation Architecture—Integrated and Standardized

VMware Cloud Foundation is based on a standardized architecture that automatically deploys a VMware Validated Design. This ensures quick, repeatable deployments of cloud infrastructure while eliminating risk of invalid configurations.

Cloud Foundation is not just a bundle, but a truly integrated stack with engineered integration of software-defined compute, storage, network, and security services in a single solution that guarantees interoperability of the entire stack, eliminating the need for complex design and interoperability matrixes.

VMware Cloud Foundation Concepts

The VMware Cloud Foundation system introduces the concept of "Workload Domain" to address user-consumable infrastructure resources.

A workload domain is a collection of compute, memory, disk, and host networking resources—a set of physical nodes. In the current version of VCF, this set of physical nodes explicitly maps to a vSphere cluster. And each vSphere cluster is tied to an instance of vCenter and to a standalone instance of NSX Manager.

This workload domain may then be used to serve three distinct functions.

Primarily, there are three kinds of consumable infrastructure resources:

1. Infrastructure as a Service (IaaS) Workload Domain
2. Management Workload Domain
3. Virtual Desktop Infrastructure (VDI) Workload Domain

Infrastructure as a Service (IaaS) Workload Domain

An IaaS Workload Domain (WD) is a set of physical hosts configured as a vSphere cluster. An IaaS WD must have a minimum of three similar physical hosts to accommodate vSAN requirements and a maximum of 64 similar physical hosts.

Management Workload Domain

A Management Workload Domain is a specialized workload domain with the specific objective of providing management functionality. It is suggested that a management WD have at least four physical hosts for availability.

Virtual Desktop Infrastructure (VDI) Workload Domain

VDI Workload Domain is a specialized overlay domain atop IaaS Workload Domain providing Virtual Desktop Infrastructure and related management functionality.

VCF System Networking Architecture

VCF employs Leaf/Spine L2 network topology for internal system physical network. Each physical rack in VMware Cloud Foundation has two top-of-the-rack switches (ToRs), a management switch (Mgmt) and many physical servers (up to 32).

The servers are dual 10Gbps connected to the ToRs using LAG. The server IPMI interfaces are connected to the rack local management switch. The management interfaces of both ToRs are also connected to the rack local management switch.

Both the ToRs are connected to each other using a 4 x 10 Gbps LAG/VPC.

The ToRs may also be connected to a pair of Spine switches, which allows for scaling out the system. The Spine switch pair is connected using a 2 x 40 Gbps LAG.

The Spine switches enable the "SCALE OUT" of the VCF system. The actual scale out capacity is determined by the number of available physical ports on the Spine switches. Using a 32-port Spine switch should easily allow for scale out up to eight physical racks OR up to 256 physical servers. The Spines ply the "east/west" traffic between the racks within VMware Cloud Foundation.

For "north/south" traffic, that is ingress/egress traffic from/to the VMware Cloud Foundation system, connectivity is established through the ToRs in the first physical rack.

For "north/south" traffic from the Nth physical rack, that is any physical rack other than the 1st, traffic is forwarded from the Nth rack local ToRs to the Spine switches and finally forwarded from the ToRs in the 1st physical rack out to the Upstream Network.

Introduction to Cohesity

Cohesity provides a web-scale data management platform designed to eliminate secondary storage silos by converging all secondary storage and associated data services on one unified solution—including backups, cloud gateway, files, objects, test/dev copies, and analytics data. Cohesity is a software-defined solution that spans from the edge to the data center and the cloud.

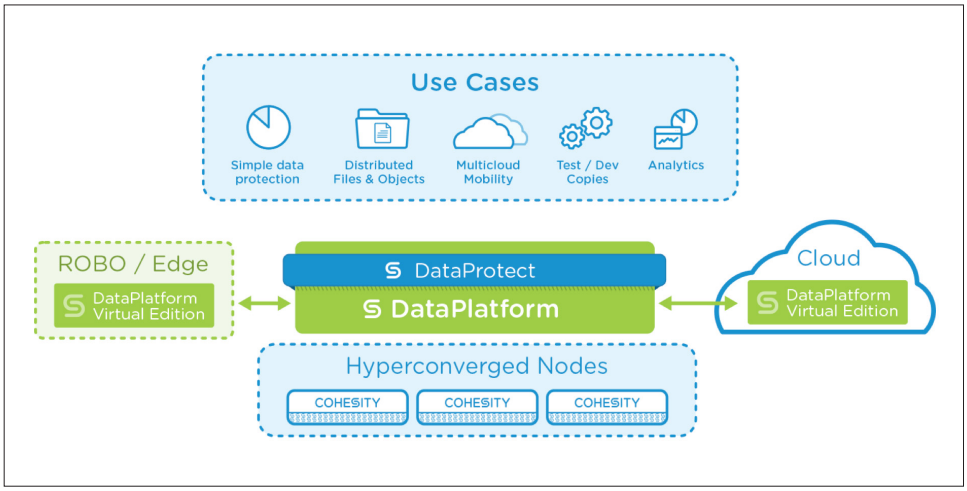


Figure 6: Use Cases for Cohesity DataPlatform

With Cohesity, enterprises can:

Consolidate secondary storage: Eliminate storage silos by consolidating secondary data, including backups, files, and test/dev copies, on a scale-out, globally deduped storage platform. Increase space and cost efficiency, simplify management and capacity planning, and eliminate the need for costly data migrations.

Simplify data protection and recovery: Simplify your data protection infrastructure with an end-to-end backup and disaster recovery solution that is fully converged on the Cohesity platform. Infinite backups. Instant recoveries. Integrated replication for multi-site disaster recovery.

Improve economics with public cloud integration: Extend your data platform to the public cloud for long-term archival, tiering of storage capacity, and disaster recovery. Make use of public cloud economics and flexibility without complicated gateways.

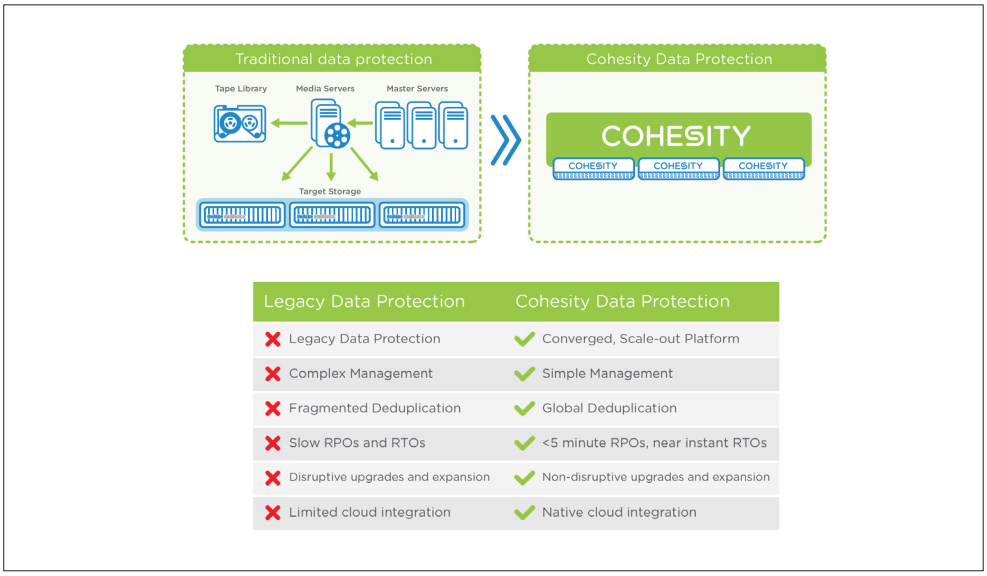
Scale-out files and objects with multiprotocol access: NFSv3, SMB 2.x, and SMB 3.0 multi-protocol access to data allows support of applications across all major enterprise operating systems including Microsoft Windows and Linux.

Gain visibility into your dark data: Shine a light on your dark data with real-time analytics on data utilization. Extract valuable insight from your data by running custom queries directly on the Cohesity platform.

Accelerate application development: Release applications faster by instantly cloning and provisioning test and dev environments on the Cohesity platform.

Cohesity for Simple Data Protection

Cohesity provides the only hyperconverged platform that eliminates the complexity of traditional data protection solutions by unifying your end-to-end data protection infrastructure, including target storage, backup, replication, disaster recovery, and cloud tiering. Eliminate data protection silos by converging all your backup infrastructure on the Cohesity scale-out platform. Simplify management with a single UI and policy-based automation. Accelerate your recovery points and recovery times while cutting data protection costs by 50%. Integrate with all the leading public clouds for archival, tiering, and replication.



Legacy Data Protection	Cohesity Data Protection
✗ Legacy Data Protection	✓ Converged, Scale-out Platform
✗ Complex Management	✓ Simple Management
✗ Fragmented Deduplication	✓ Global Deduplication
✗ Slow RPOs and RTOs	✓ <5 minute RPOs, near instant RTOs
✗ Disruptive upgrades and expansion	✓ Non-disruptive upgrades and expansion
✗ Limited cloud integration	✓ Native cloud integration

Figure 7: Advantages of Cohesity for Data Protection

Web-Scale with Best-In-Class Space Efficiency

- Scale-out platform: Replace individual storage appliances with a scale-out platform. Ensure always-on availability, pay-as-you-grow by scaling performance and capacity linearly, and eliminate expensive forklift upgrades.
- Global deduplication, erasure coding, and compression: Increase storage efficiency with compression and variable-length, global deduplication that spans an entire cluster. Provide data availability without sacrificing storage efficiency with erasure coding across nodes.

Simple, Automated Data Protection

- Converge backup + target storage: Simplify your data protection infrastructure with a complete backup and recovery solution that is fully converged on Cohesity DataPlatform. Eliminate the need for separate backup software, proxy servers, media servers, and target storage.
- End-to-end workflows through single UI: Manage and provision your data protection services through a single pane of glass. Orchestrate data protection with a complete set of REST APIs and associated documentation. Integrate with your existing orchestration and DevOps tools.
- Policy-based automation: Create policies that specify your application requirements including RPO, retention periods, off-site replication, and archival targets. Automate data protection by assigning policies to individual applications based on SLA requirements.

Broad Application and Infrastructure Support

- Supports all the leading hypervisors: Backup VMs running on all the leading hypervisors, including VMware vSphere, Microsoft Hyper-V, Nutanix AHV, and KVM. Management integration enables Cohesity to have a full view of VMs running on the hypervisor. DataProtect uses hypervisor-specific APIs (e.g., VMware VADP, Microsoft RCT) to perform changed-block tracking and app-consistent backups.
- App-consistent backups for common OS and databases: Perform application-consistent backups with application adapters for physical Windows, Linux, and SQL Server, with support for Windows clustering and SQL AAG. Provision test/dev copies of SQL Server by automating clone and copy attach of SQL databases to any SQL Server. Provide fully managed Oracle RMAN-based data protection with source-side dedupe, support for Oracle RAC and ASM, and log backups for any-point-in-time recovery.
- Native protection of storage devices: Integrate with Pure Storage FlashArray to automatically tier snapshots from Pure Storage down to Cohesity. Protect Pure Storage FlashBlade, NetApp, and Dell EMC Isilon with NAS volume snapshots, identification of changed data on the volume, and a high-performance multi-stream data mover. Protect any generic NAS device.

Fast Recovery Points and Recovery Times

- Sub-five minute recovery points: Reduce your Recovery Points to sub-five minutes by taking an unlimited number of incremental backups, storing each backup as a fully hydrated snap on Cohesity DataPlatform, and leveraging parallelized data ingest.
- Full catalog of always-ready snaps: Access each backup instantaneously for cloning, application, or file-level recovery, with a full catalog of always-ready snapshots.
- Near-Instantaneous Recovery Times: Recover large, multi-tier applications in seconds by instantly provisioning clones from backup snapshots. Run recovered applications directly on Cohesity DataPlatform until data is copied back to primary storage.

Granular Search and Recovery

- Instant Google-like search: Instantly find your virtual machine and file data with Google-like wild-card search on Virtual Machines and individual files.
- Granular VM, file, and object-level recovery: Recover individual VMs, restore files to source VMs, and recover individual application objects for Exchange, SQL, and SharePoint.

Off-Site and Long-Term Data Protection

- Remote replication for disaster recovery and migrations: Replicate between Cohesity clusters for off-site data protection, disaster recovery, and application migrations. Leverage flexible topologies including one-to-many and many-to-one replication.
- Cloud: Integrate natively with all the leading public cloud providers including Google Cloud Storage Nearline, Microsoft Azure, Amazon S3, and Glacier. Use the cloud for long-term archival, data tiering, and replication.
- Tape archival: Support external tape libraries for long-term data archival.

Built-In Security

- Software-based encryption of data at-rest and in-flight: Encrypt data on Cohesity with software-based encryption using the AES-256 standard, with optional FIPS certification. Data is encrypted at-rest on the platform, and in-flight when replicated or archived to the cloud. Keys are automatically rotated and managed either by an external key management system or by the Cohesity cluster.
- Role-Based Access Control: Integrate with Windows AD to support role-based permissions. Customize permission levels by type of user (admin vs. end-user) and by source of protected data.
- Ransomware protection: Provide comprehensive protection against ransomware attacks. Take frequent backups, as often as every five minutes, to enable quick recovery of data on primary storage systems with minimal data loss. In addition, Cohesity backups are stored on immutable snaps to prevent malicious alterations.

Cohesity Architecture

Cohesity's architecture is based on its unique file system designed to consolidate secondary storage at web-scale, called SpanFS.

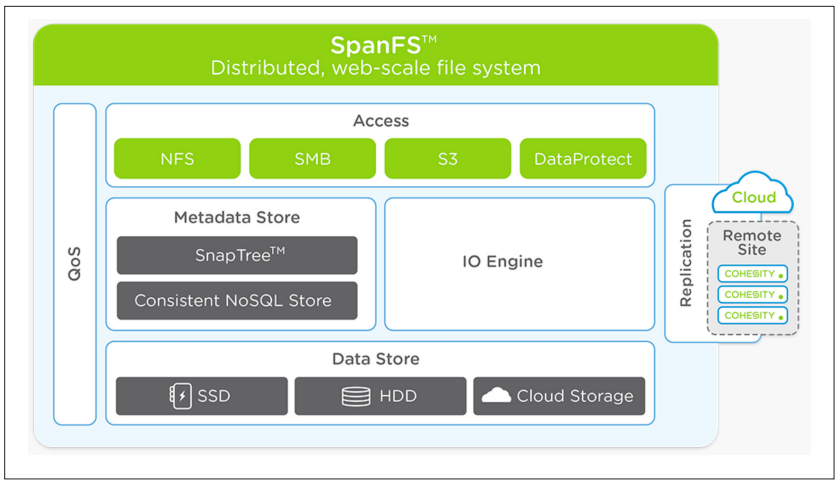


Figure 8: SpanFS Architecture

SpanFS is a completely new file system designed specifically for secondary storage consolidation. At the topmost layer, SpanFS exposes industry-standard, globally distributed NFS, SMB, and S3 interfaces. It also provides a proprietary interface to store backup data protected with Cohesity DataProtect.

The IO Engine manages IO operations for all the data written to or read from the system. It detects random vs. sequential IO profiles, then splits the data into chunks, performs deduplication, and directs the data to the most appropriate storage tier (SSD, HDD, cloud storage) based on the IO profile.

Random IO is placed on the SSD tier, sequential IO is sent straight to HDD or SSD based on QoS, and colder data may be sent to the cloud if cloud storage is in use. To keep track of and manage the data sitting across nodes, Cohesity also had to build a completely new metadata store. The metadata store incorporates a consistent, distributed NoSQL store for fast IO operations at scale, and SnapTree provides a distributed metadata structure based on B+ tree concepts.

SnapTree is unique in its ability to support unlimited, frequent snapshots with no performance degradation. SpanFS has QoS controls built at all layers of the stack to support workload and tenant-based QoS, and can replicate, archive and tier data to another Cohesity cluster or to the cloud.

The file system is distributed on hyperconverged nodes, built with commodity x86 servers, available from Cohesity or third-party hardware partners such as Cisco and Hewlett-Packard Enterprise. SpanFS can also be deployed in the public cloud, on cloud VMs, and is available in public cloud service catalogs.

Cohesity Architecture Components

Access Layer: SpanFS exposes industry-standard NFS, SMB, and S3 protocols. Any number of volumes or object buckets can be configured simultaneously on a single Cohesity cluster. The volumes are completely distributed with no single choke point. The data is spread out across all the nodes in the cluster.

Volumes are accessed through a virtual IP mount point, and user access and IO are distributed across the nodes using the virtual IP address. Each of these volumes benefits from all the unique SpanFS capabilities such as global deduplication, encryption, replication, unlimited snapshots, and file/object level indexing and search.

IO Engine: The IO Engine manages read and write IOs as well as data operations like deduplication. The IO Engine automatically detects whether the workload is sequential or random in nature, and directs IO to the appropriate data path and media tier based on the profile. Sequential IOs may go straight to HDDs or may use the SSDs based on QoS policies. Random IOs are directed to a distributed data journal that resides on SSDs. This mechanism enables SpanFS to effectively manage both sequential and random IOs with high throughput and low latency. The IO Engine splits the data into chunks and spreads the chunks across nodes to maximize performance and capacity utilization.

Each chunk is protected against node failures either by replicating the chunks across nodes or by using erasure coding across nodes. The IO Engine is also responsible for performing data operations that are required prior to storing the chunks of data, such as variable-length deduplication and indexing.

Deduplication is performed using a unique, variable-length data deduplication technology that spans an entire cluster, resulting in significant savings across a customer's entire storage footprint. SpanFS creates variable-length chunks of data, which optimizes the level of deduplication no matter the type of file.

In addition to providing global data deduplication, Cohesity allows customers to decide if their data should be deduplicated in-line (when the data is written to the system), post-process (after the data is written to the system), or not at all.

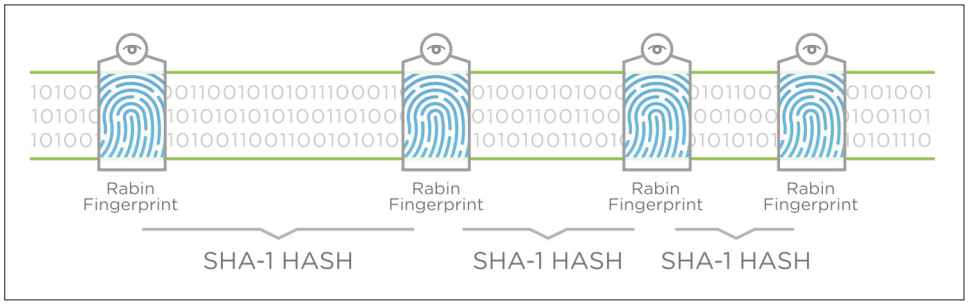


Figure 9: SpanFS Global Deduplication Architecture

Data Store: The data store is responsible for storing data on HDD, SSD, and cloud storage. The data is spread out across the nodes in the cluster to maximize throughput and performance, and is protected either with multi-node replication or with erasure coding. Sequential IOs may go straight to HDDs or to SSDs based on QoS policies.

Random IOs are directed to a distributed data journal that resides on SSDs. As the data becomes colder, the data store can tier the data down from SSD to HDD. And hot data can be up-tiered to SSD. Customers may set up cloud storage, in which case the data store can automatically move colder chunks of data to reside in the cloud, and bring the chunks back to HDD or SSD once they become hot again.

SnapTree: In legacy storage, snapshots form a linked chain, with each link containing the changes from the prior snapshot. Every time a new snapshot is done, an additional link is added to the chain. As the chain grows, the performance overhead required to access the data increases proportionally because the file system must traverse the chain to access the data.

Hence, snapshots introduce performance overhead and are limited in scope. SnapTree introduces a completely new approach to managing metadata at scale, and enables SpanFS to provide unlimited snapshots with no performance overhead. SnapTree is based on a B+ tree metadata structure, but adds multiple innovations including:

- Distributes the tree across nodes
- Provides concurrent access from multiple nodes
- Supports the creation of instantaneous clones and snaps
- Garbage collects unreferenced nodes in the background using Map-Reduce
- Ensures consistent access performance regardless of the number of snapshots and clones

- Stores only one value per leaf node, as opposed to multiple values in traditional B+ trees, which avoids unnecessary snapshotting of multiple values
- Supports a variable fan-out factor that increases further down the tree, which avoids making any given sub-part of the tree too hot while at the same time keeping tree balancing costs low

With SnapTree, Views (volumes) and files are represented by a tree of pointers to the underlying data. The root node represents the View or individual files. The root node points to some intermediary nodes, which in turn point to the leaf nodes which contain the location of the chunks of data.

Customers can take snapshots of entire Views (volumes) or individual files within the Views. As snapshots are taken, the number of hops from the root to the leaves does not increase. Customers can take snapshots as frequently as desired without ever experiencing performance degradations.

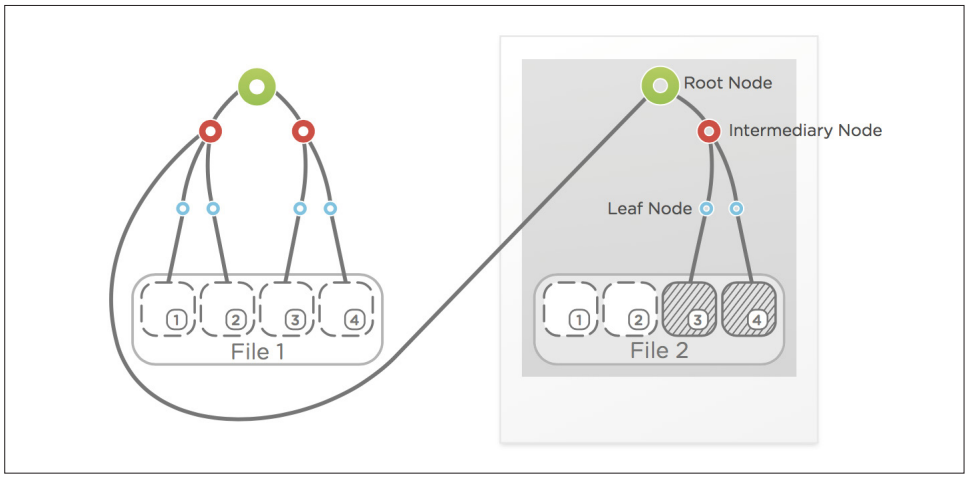


Figure 10: SnapTree Architecture

The figure above shows how the metadata is structured in SnapTree. In this example, the original file is called File 1 with its own root node and metadata tree. File 2 is a snapshot of File 1, and gets its own root node. The root node points to the original intermediary node for unchanged chunks of data (chunks 1 and 2). But a new intermediary node is created for File 2, to point to the new changed chunks of data (chunks 3 and 4).

Consistent NoSQL Store: The metadata store uses a distributed NoSQL store that stores the metadata on the SSD tier. It's optimized for fast IO operations, provides data resiliency across nodes, and is continually balanced across all the nodes. However, the key-value store by itself provides only "eventual consistency." To achieve strict consistency, the NoSQL store is complemented with Paxos algorithms. With Paxos, the NoSQL store provides strictly consistent access to the value associated with each key.

QoS: Quality of Service is designed into every component of the system. As data is processed by the IO Engine, Metadata Store, or Data Store, each operation is prioritized based on QoS. High priority requests are moved ahead in subsystem queues, and are given priority placement on the SSD tier.

Replication and Cloud: SpanFS can replicate data to another Cohesity cluster for disaster recovery, and archive data to third-party storage like tape libraries, NFS volumes, and S3 storage. SpanFS has been designed to interoperate seamlessly with all the leading public clouds (AWS, Microsoft Azure, Google Cloud). SpanFS makes it simple to use the cloud in three different ways. First, CloudArchive enables long-term archival to the cloud, providing a more manageable alternative to tape.

Next, CloudTier supports data bursting to the cloud. Cold chunks of data are automatically stored in the cloud, and can be tiered back to the Cohesity cluster once they become hot. Finally, CloudReplicate provides replication to a Cohesity Cloud Edition cluster running in the cloud. The Cohesity cluster in the cloud manages the data to provide instant access for disaster recovery, test/dev, and analytics use cases.

Instant Data Provisioning

SpanFS enables enterprises to consolidate all their secondary data on a single platform. But SpanFS goes beyond consolidation—it also enables enterprises to provision data instantly to support secondary storage use cases. SpanFS doesn't just do Copy Data Management—it eliminates the need to make data copies. Users can instantly provision clones of backup data, files, objects, or entire Views and present those clones to support a variety of use cases.

For example , the clones can scan for instant recovery for test/dev copies or to support analytics. All these use cases can be supported directly on SpanFS serving as the active storage system, or the data can first be moved to another storage system such as primary storage or test/dev storage.

The snapshots and clones are very efficient. They don't consume space until data is modified, in which case they only need to store the deltas from the original copy. They can be created instantly without having to move data between storage devices. This is in stark contrast to the inefficiency of traditional secondary storage, where full copies of data are created between storage silos, wasting lots of storage capacity, time, and IO bandwidth.

Next Generation Data Protection and Recovery for VMware Cloud Foundation (VCF)

Case for Data Protection

Often times, recovery from local backup is much faster and more cost-efficient compared to recovery from a secondary site. This isn't to say a true DR scenario isn't required, but local backups provide an additional level of redundancy/availability recovery efficiency and timeliness to business continuity.

Also, not all workloads may require DR. Some workloads can be tiered for DR while others be structured to be recovered locally at the site.

VMware and Cohesity have combined efforts to provide the next generation data protection and recovery solution for VMware Cloud Foundation. The following are the use cases and test scenarios that have been tested and validated as part of this solution.

- Simple protection of both Management VMs and Workload VMs
- Recovery of Management VMs and/or Workload VMs from a point-in-time backup

- Storage Migration of Workload VMs from a point-in-time backup to a different infrastructure
- Recovery of file/folder and Workload VMs from a point-in-time backup after catastrophic loss

The joint Cohesity and VMware VCF solution delivers the following benefits:

- Web-scale, pay-as-you grow primary and secondary storage architecture
- Dynamic, application-centric operations through integration with storage policy-based management
- Simple end-to-end data protection for both Management and Workload VMs
- Ensure fast recovery of an entire VCF environment, including Management VMs and Workload VMs
- Lower total cost of ownership
- Consolidate backups, files, and test/dev copies on a single web-scale platform
- Accelerate application time-to-market with instantaneous provisioning of clones for test/dev
- Get deep visibility into your secondary data with built-in analytics capabilities

The Cohesity DataPlatform provides the capability for IT administrators to bridge the traditional "islands of secondary storage" by leveraging the truly global file system and storage efficiency technologies built into the platform. These capabilities help customers in transforming their data centers from silos of dark data to highly efficient, next-generation webscale Enterprise IT.

Cohesity is being used to protect the VCF domains in the graphic on the following page.

The best practice would be to isolate the data protection infrastructure from VCF infrastructure. This design would localize failures and therefore would have a viable (as defined with RPO/RTO policies) copy of data in the event of a catastrophic failure of either infrastructure.

The Cohesity data platform can be configured for L2 or L3 networking. In this paper, we discuss L3 design guidelines, which have the following benefits :

1. Consolidates data protections across multiple VCF and traditional VMware environments in the same on-premises location.
2. Isolates primary storage traffic from data protection traffic and by extension also localizes failures.
3. Allows for scalability within on-premises infrastructure and/or across geographical regions.

In the following diagram, the two VCF domains are “Humboldt” (rack1) and “Plumas” (rack2).

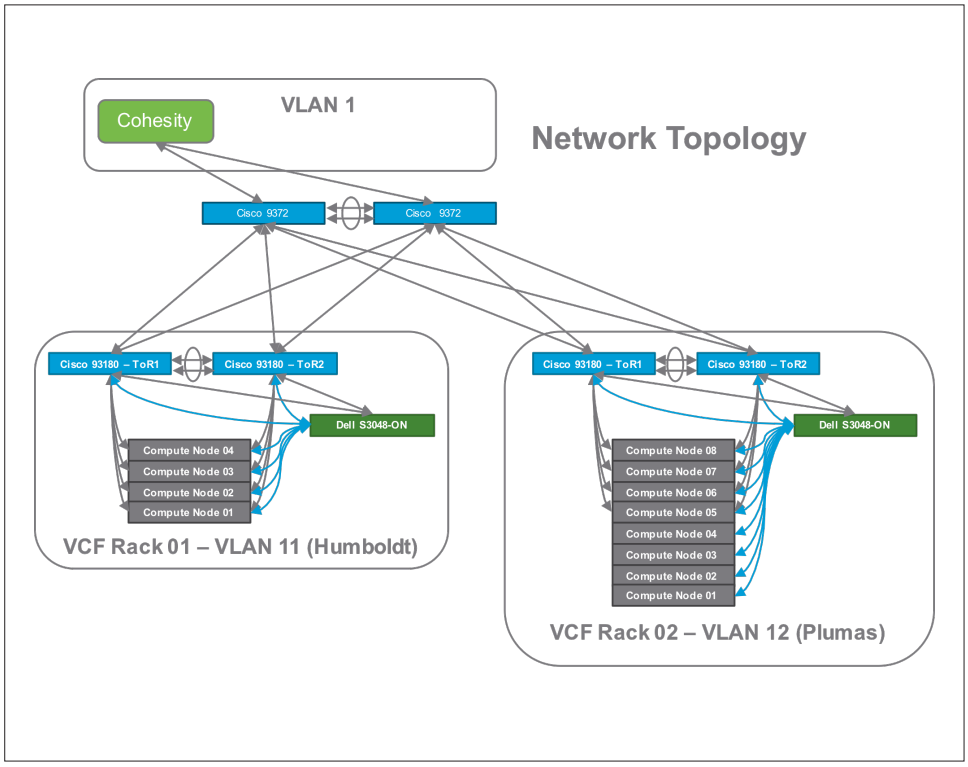


Figure 11: VCF Domain Topology

As seen in the topology above, the multiple VCF domains were separated by locating them in different parts of the data center.

Use case 1 : Recovery of Management VMs and/or Workload VMs from a point-in-time backup

The following VCF SDDC environment was used for testing and validating this use case.

This VCF Management domain is made up of four hosts connected to two TOR switches for network and storage IO. All the four hosts of the management domain also participate in a vSAN, which provides the primary storage for the VCF domain. The Cohesity DataPlatform is connected to this VCF environment through the 2x Cisco 9372 network core switches. Cohesity is being used to protect all the components that make up the management and workloads in this VCF domain.

The following graphic shows the dashboard for the VCF domain, which was used for testing this combined solution.

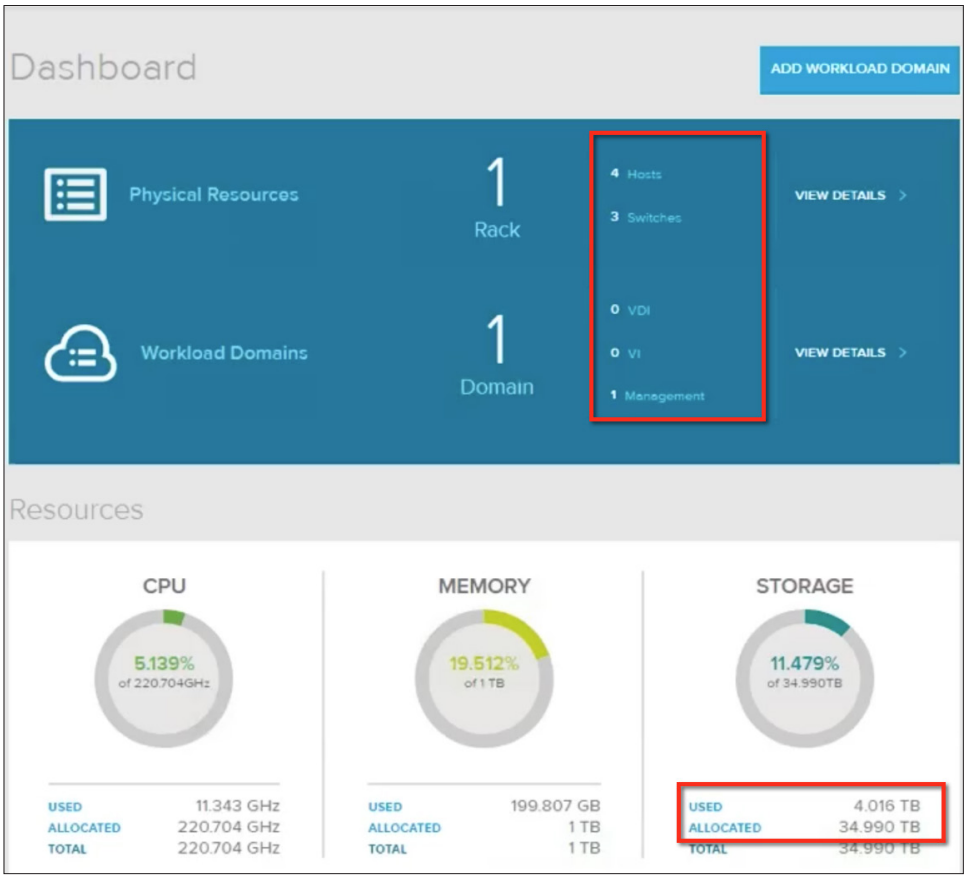


Figure 12: VCF SDDC Manager Dashboard

The physical resources of the VCF rack "Humboldt" are shown below.

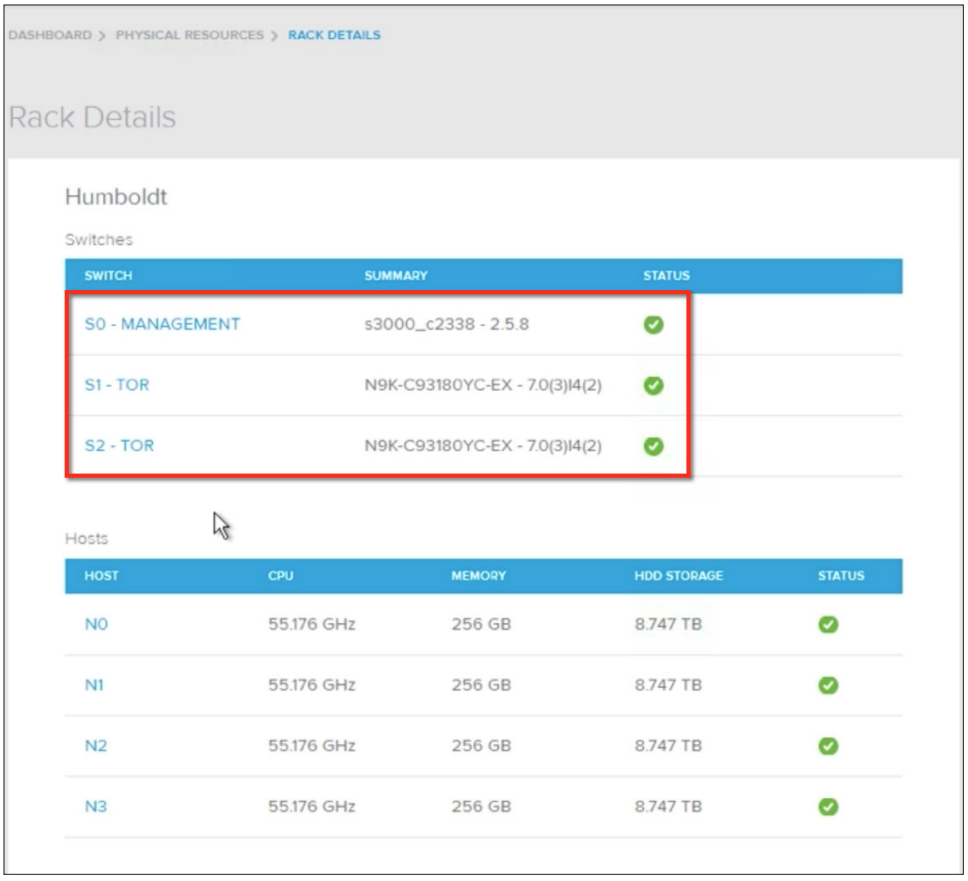


Figure 13: VCF Domain Physical Resources

The following are the HW specifications that were used :

- Management switch—Dell S3048-ON. 48 x 1Gbps + 4 x 10Gbps ports.
- TOR Switches—Cisco 93180YC-EX 48 x 10/25Gbps + 6 x 40/100Gbps ports
- Servers—Dell R630 with Intel Xeon

The management domain details and resources of Humboldt are seen below.

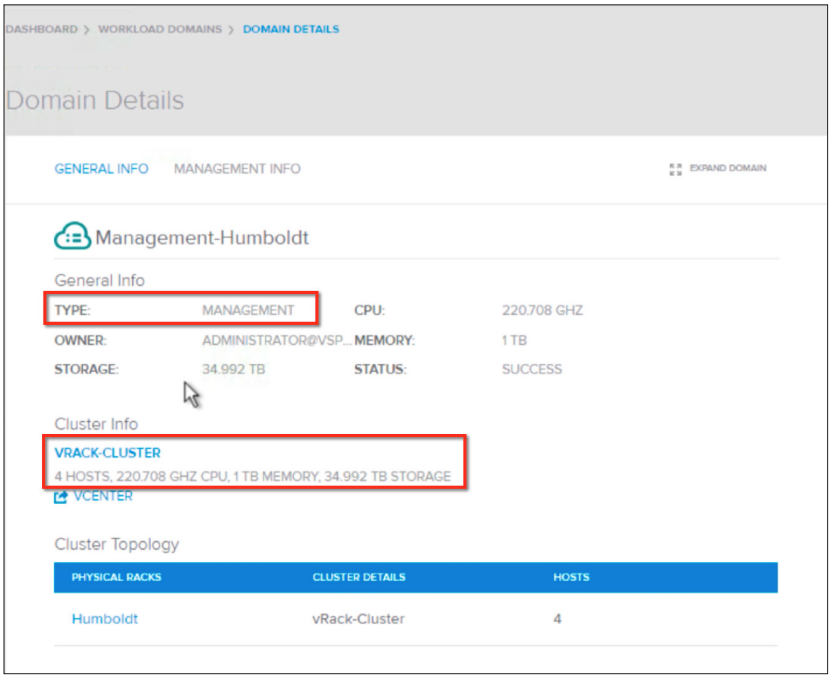


Figure 14: VCF Management Domain Details

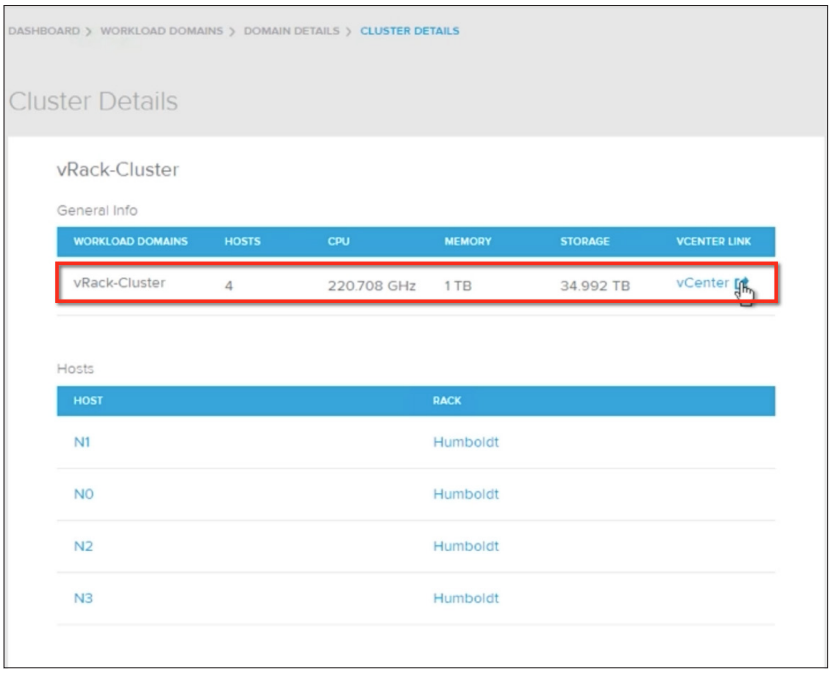


Figure 15: VCF Management Domain Details (continued)

The management stack of VCF is made up of vCenter, PSC, SDDC Manager, NSX Manager, vRealize Loginsight, and vRealize Operations. All the VMs in this stack work in tandem to enable the SDDC functionality in VCF.

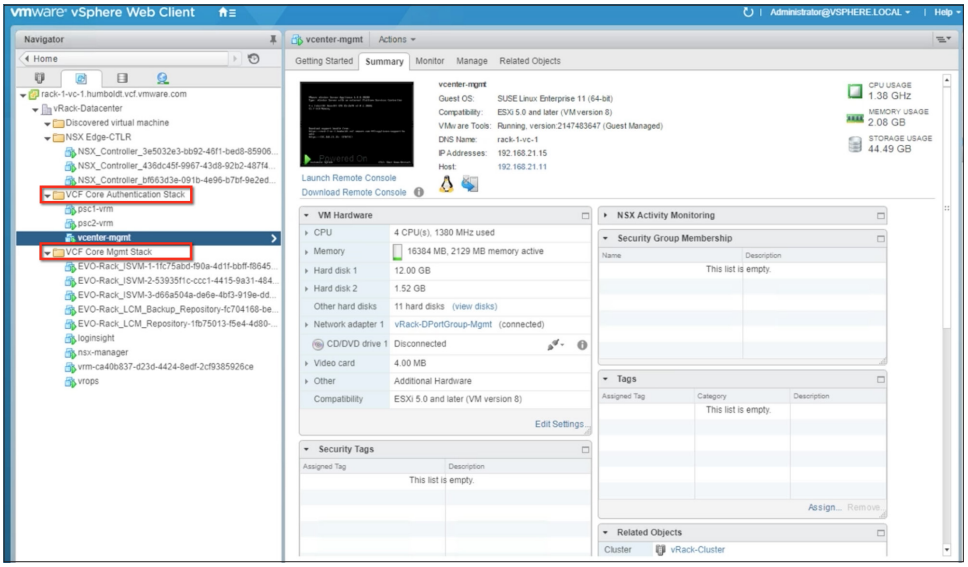


Figure 16: VCF Domain Folder Hierarchy

These VMs are working in tandem to provide the SDDC services to enable workload domains.

Total provisioned and used space on primary storage vSAN:

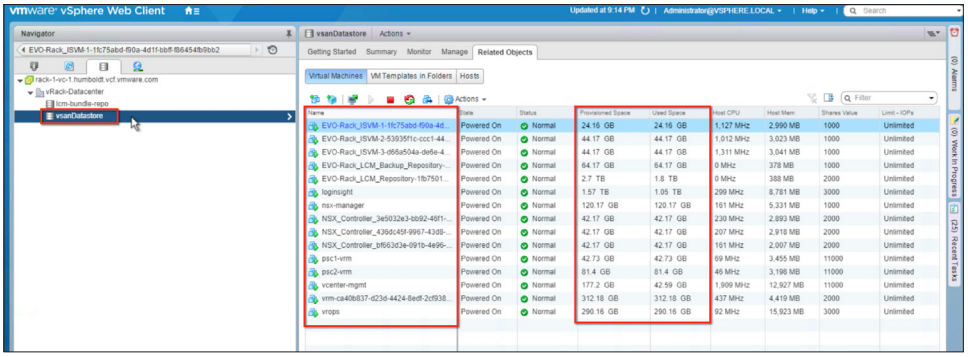


Figure 17: VCF Management VMs on vSAN Datastore

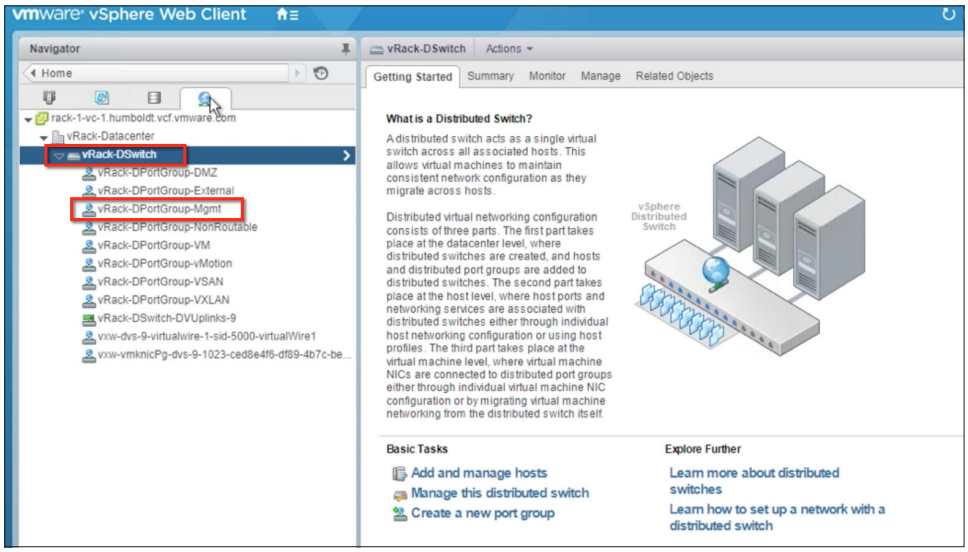


Figure 18: VCF Domain DVS Port Group configuration

Each Workload Domain is configured with a single DVS with up to seven Port Groups. These seven Port Groups are configured with a unique VLAN and designated for the following purpose:

1. Mgmt—Management Traffic
2. NonRoutable—Internal Management Traffic
3. vMotion—Workload Domain vMotion Traffic
4. VSAN—Workload Domain VSAN Traffic
5. VXLAN—Workload Domain VXLAN Traffic
6. DMZ / VM—Unmapped to any VLAN/Native VLAN Traffic
7. External—Workload Domain VM Traffic

Mgmt, NonRoutable, DMZ/VM Portgroups span all the WLDs in a VCF instance.

External Portgroup may be to unique to a specific WLD or be shared across some or all of the WLDs.

All of the above Port Groups/VLANs are created over a single LAG comprising of 2 x10 Gbps network adapters. Further, for all the Management components such as VC, PSC, and VRM, the physical hosts are connected on the Management Network. Management-related traffic on this network is nominal. As such, the Management Network was used for the purpose of recovery.

Cohesity Data Protection Policies and Job Details

Cohesity Dashboard

The Dashboard provides a wealth of information about the cluster at a glance.

1. Tabs on the top provide information about last 24 hours Job Run information along with SLA violations and Errors if any, plus number of VMs and NFS/SMB Views (mount points exported from Cohesity) being protected.

2. In the middle, users would see information about the cluster (i.e., number of nodes, alerts in the last 24 hours, and amount of RAW storage and usage).
3. The bottom tabs provide information about Data Reduction (dedupe + compression) and performance for the last 24 hours.

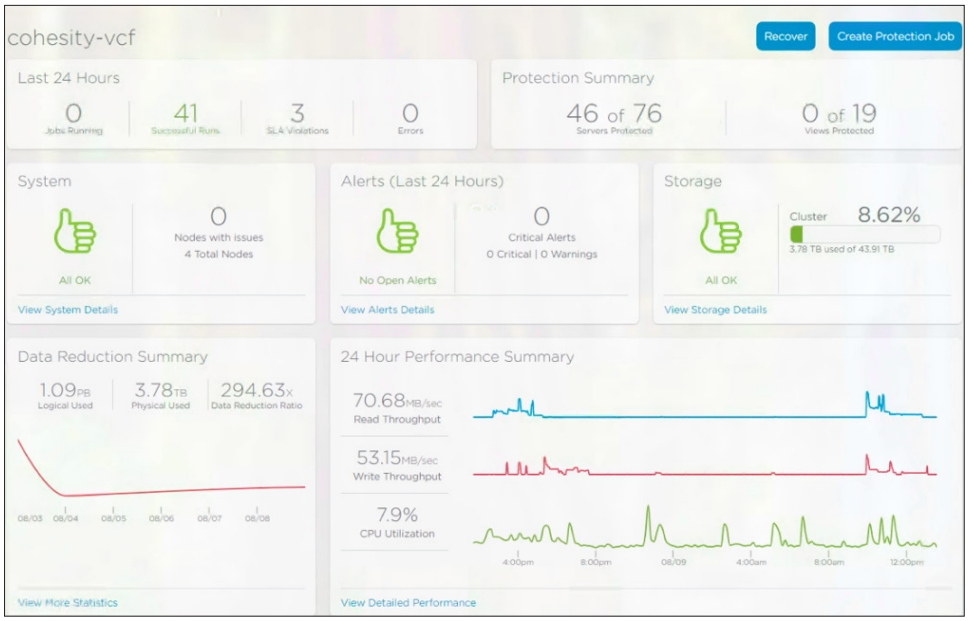


Figure 19: Cohesity Dashboard

Users can see advanced system details by clicking any one of these tabs.

We can initiate recovery workflows for files or folders from VMs being protected and VM level recovery.

We will look at each of these recovery options in detail in the following sections.

The following policies were used on the Cohesity platform to protect the VCF Management VMs.

The policies were created separately as the restore process needs to be executed in a certain sequence, and the VMs need be brought up in the following order.

The vCenter and the PSC VMs provide SSO and other authentication functions to all the VMs in the Management Stack. The “VCF Authentication Stack” policy was created so these three VMs could be protected together, and in case of a failure or vulnerability such as Ransomware, could be restored back together.

Restoring these VMs back together allows users to maximize the availability and other SLAs which could be affected due to authentication services being unavailable.

Following are the VCF Authentication Stack Job details which are associated with the above policy.

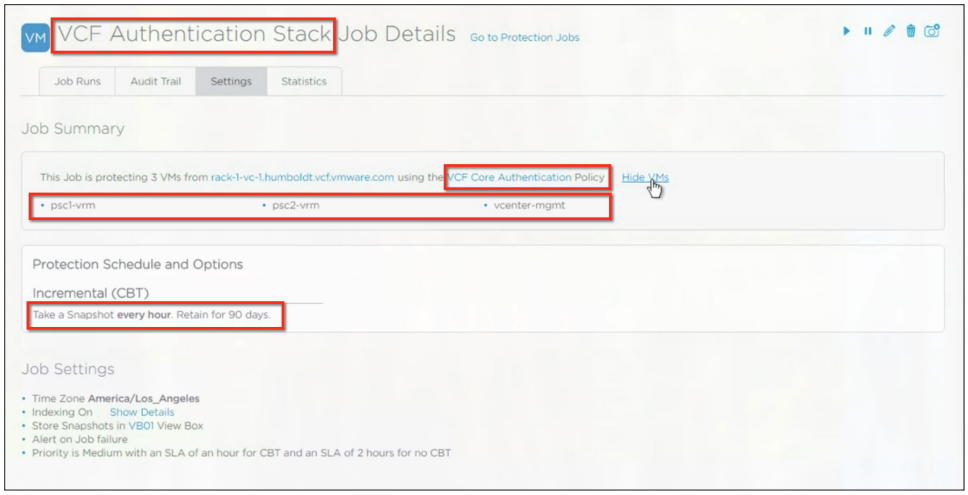


Figure 20: VCF Authentication Stack Protection Job Details

The rest of the core management VMs (i.e., VROPs, LCM vms, ISVMs, SDDC Manager, NSX manager and LogInsight VMs) are all protected together with a single policy called “VCF Core Mgmt” policy.

Following are the VCF Core Mgmt Stack job details which are associated with the above policy.

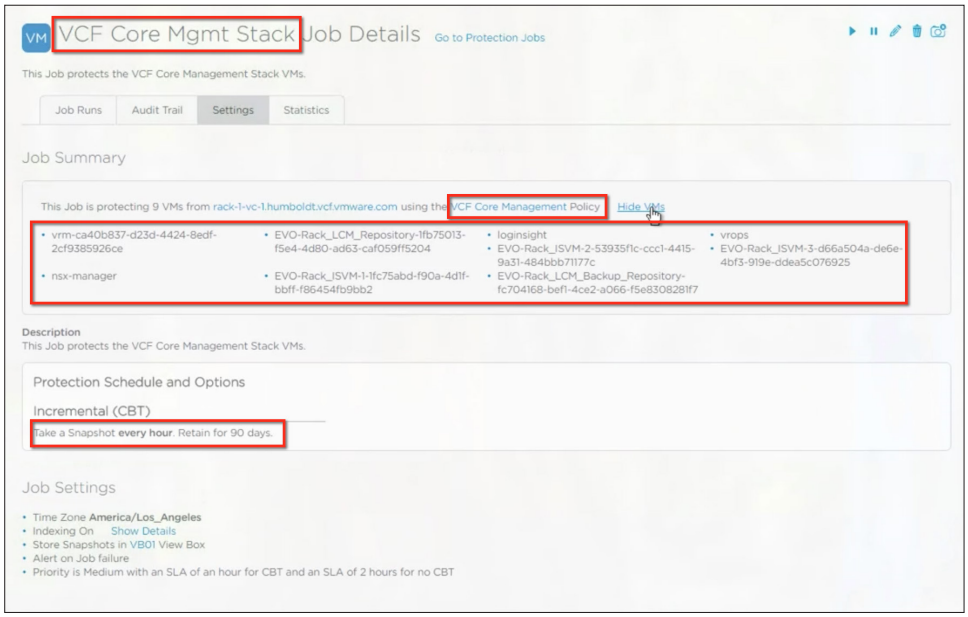


Figure 21: VCF Management Stack Protection Job Details

The SDDC Manager and the ISVMs are protected with the same job, as the ISVMs are running a distributed NoSQL database, which the SDDC Manager uses to store all the UUID and other VM identification information for both the management domains and workload domains in the VCF environment.

A Cohesity backup job coordinates actions on the vCenter, ESXi host, the guest OS, and all nodes in the Cohesity cluster.

The key elements and order of operations are:

- Step 1: Cohesity DataProtect software triggers a scheduled backup job run or the job is manually triggered by a user.
- Step 2: The cluster automatically distributes the task of backing up individual VMs across the entire cluster.
- Step 3: Cohesity cluster contacts the vCenter to gather current inventory and obtain virtual disk and ESXi host information for the target VMs to be protected.
- Step 4: vCenter passes the requested information back to Cohesity.
- Step 5: Cohesity contacts the ESXi host and gathers host resource status as well as currently running backup tasks. If resources and backup tasks are within Cohesity thresholds, the backup job run will begin. If the ESXi host is already loaded with backup tasks, then the Cohesity cluster will poll for host availability and defer the backup job to another time.
- Step 6: When application-consistent backups are set, the Cohesity cluster checks for the presence of VMTools on the guest OS, which is required to invoke Windows VSS Snapshots for application consistency.
- Step 7: Cohesity then contacts the vCenter and requests a VM snapshot. If application-consistency is requested, VMware invokes VSS Snapshots for Windows. Optionally, VMware will also run pre/post scripts for Linux as an additional step before invoking VMware snapshot.
- Step 8: VMware takes a snapshot honoring any requested VSS snapshot or pre/post scripts configured.
- Step 9: Cohesity validates the snapshot.
- Step 10: Cohesity backs up all of the snapshots from all ESXi hosts in parallel across all Cohesity cluster nodes optimized for both VMware and Cohesity parameters such as VMs per data store and maximum backups per Cohesity cluster.
- Step 11: Cohesity requests that the snapshot be deleted.
- Step 12: VMware host releases the snapshot.

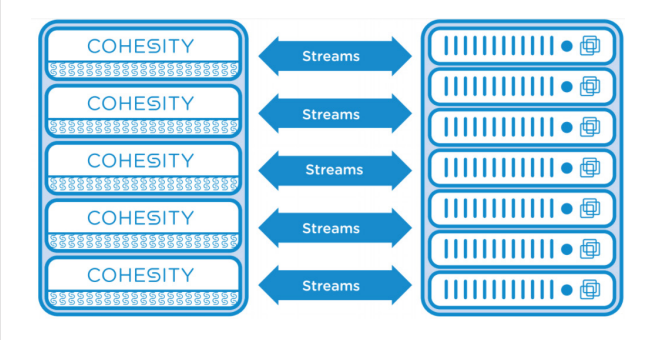


Figure 22: Cohesity Distributed Ingest

To validate the use case, a catastrophic failure was initiated in the VCF domain, by using a Powershell script to wipe out all the Management VMs which make up this domain.

Before we ran the script, we validated that all the VMs in the stack were available and online.

TIMESTAMP	VM_NAME	IP_ADDRESS	STATUS
08-09-2017.13:42:01	SDDC-Manager	192.168.21.14	<<UP>>
08-09-2017.13:42:02	vCENTER	192.168.21.15	<<UP>>
08-09-2017.13:42:03	LogInsight	192.168.21.16	<<UP>>
08-09-2017.13:42:04	vROPs	192.168.21.17	<<UP>>
08-09-2017.13:42:05	NSX-Manager	192.168.21.18	<<UP>>
08-09-2017.13:42:06	Controller-1	192.168.21.19	<<UP>>
08-09-2017.13:42:07	Controller-2	192.168.21.20	<<UP>>
08-09-2017.13:42:08	Controller-3	192.168.21.21	<<UP>>
08-09-2017.13:42:09	PSC-1	192.168.21.23	<<UP>>
08-09-2017.13:42:10	PSC-2	192.168.21.24	<<UP>>
08-09-2017.13:42:11	LCM-Repo	192.168.21.25	<<UP>>
08-09-2017.13:42:12	LCM-Backup	192.168.21.26	<<UP>>

Figure 23: Ping test to check uptime of VCF Management VMs

The "kill-vm-v2.ps1" script deleted all the VMs in the management domain and rendered the VCF domain completely unavailable.

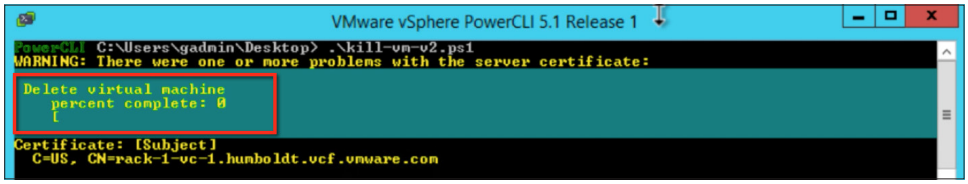


Figure 24: PowerCLI script to delete all VCF Management VMs

The vCenter and SDDC manager were confirmed to be unavailable, and also all the rest of the VMs were validated to be down, by running a ping test across all the IP addresses which are assigned to respective management VMs of the VCF stack.

TIMESTAMP	VM_NAME	IP_ADDRESS	STATUS
08-09-2017.13:42:50	SDDC-Manager	192.168.21.14	>>DN<<
08-09-2017.13:42:51	vCENTER	192.168.21.15	>>DN<<
08-09-2017.13:42:52	LogInsight	192.168.21.16	>>DN<<
08-09-2017.13:42:53	vROPs	192.168.21.17	>>DN<<
08-09-2017.13:42:54	NSX-Manager	192.168.21.18	>>DN<<
08-09-2017.13:42:55	Controller-1	192.168.21.19	<<UP>>
08-09-2017.13:42:56	Controller-2	192.168.21.20	<<UP>>
08-09-2017.13:42:57	Controller-3	192.168.21.21	<<UP>>
08-09-2017.13:42:58	PSC-1	192.168.21.23	>>DN<<
08-09-2017.13:42:59	PSC-2	192.168.21.24	>>DN<<
08-09-2017.13:43:00	LCM-Repo	192.168.21.25	>>DN<<
08-09-2017.13:43:01	LCM-Backup	192.168.21.26	>>DN<<

Figure 25: Unavailability of VMs after deletion

Both the vCenter and the SDDC manager reported that the application was unavailable after the script had executed.

The respective UIs for both SDDC Manager and vCenter reported, "This site can't be reached."

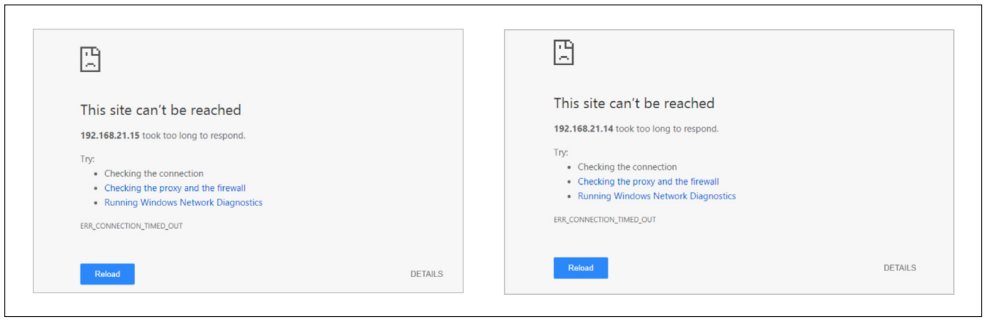


Figure 26: Application Unavailability in SDDC Manager and vCenter

The ping test reported that except for the NSX controllers, all the core infrastructure VMs were down.

Note: NSX controllers were not backed up as part of this validation. All NSX controllers are backed up as part of the NSX manager database backups. See link here for more details.

The following steps were taken to recover the environment from the Cohesity backups.

With incumbent backup and recovery solutions, the recovery operation is dependent on the vCenter. Hence, when the vCenter is down, recovery becomes a series of steps and adds complexity.

The first challenge to recover the infrastructure is that the vCenter and the PSC VMs are both down. Without the vCenter and PSC VMs, SSO and authentication would not be possible and any incumbent backup solution would not be able to communicate to the vCenter to restore the VMs.

Cohesity allows restoring VMs to an individual standalone ESXi host, by means of exposing a NFS data store to it, and powering on the VMs from the backups from the selected RPO. The restore process is almost instantaneous as the VM is powered on from the Cohesity NFS data store. This allows the authentication services to come online as the first step in the overall infrastructure recovery procedure.

Follow the steps below to restore the VCF Authentication Stack.

Click on the Recover tab in the Cohesity dashboard to be presented with the Recovery workflow.

1. Search for the job using a regex search string.
2. Select the job to be restored.

3. Pick the desired RPO from the drop down and "Add to Cart" and then hit Continue.

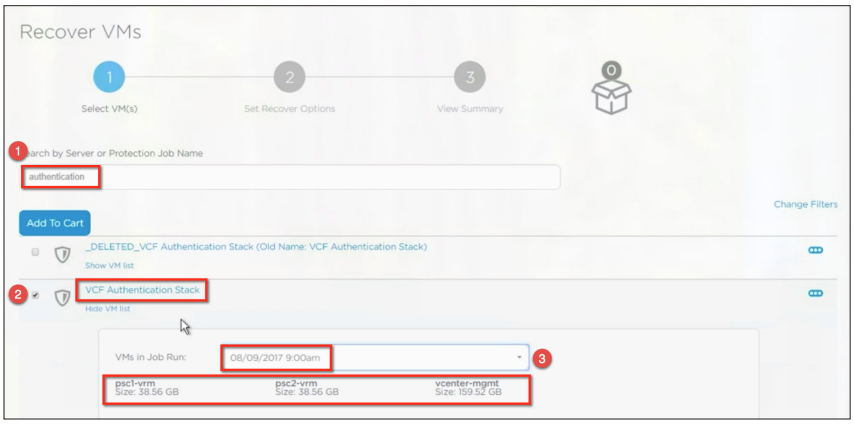


Figure 27: VCF Authentication Stack Restore

4. Select "Restore to a new location" radio button.
5. From the drop down, select the previously registered ESXi standalone host.

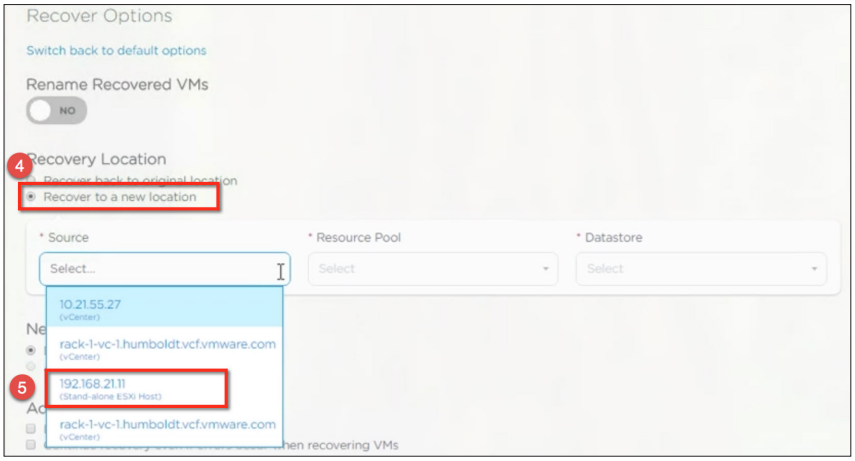


Figure 28: VCF Authentication Stack Restore (continued)

6. Select the radio button to "Attach to a new network".
7. Select the respective network port group the bring up the three VMs with.

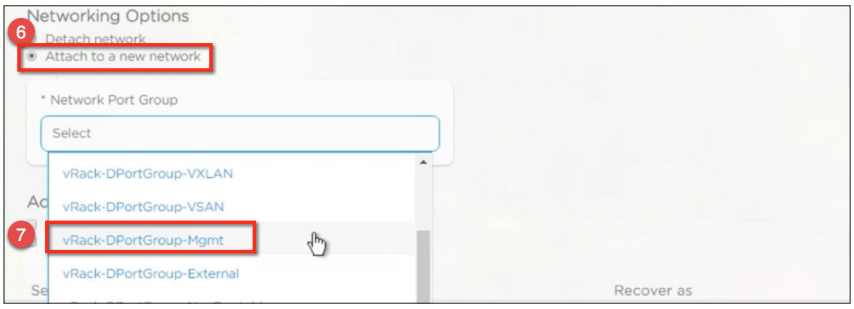


Figure 29: VCF Authentication Stack Restore (continued)

8. Click Finish to start the restore process.

The three VMs are recovered from a NFS data store presented from the Cohesity secondary storage platform, and the VMs come online on the network in less than two minutes.

When using Cohesity, "time to application availability" to restore VMs is drastically reduced because of the almost-instant RTO capabilities of the platform.

TIMESTAMP	VM_NAME	IP_ADDRESS	STATUS
08-09-2017.13:47:21	SDDC-Manager	192.168.21.14	>>DN<<
08-09-2017.13:47:22	vCENTER	192.168.21.15	<<UP>>
08-09-2017.13:47:23	Loginsight	192.168.21.16	>>DN<<
08-09-2017.13:47:24	vROPs	192.168.21.17	>>DN<<
08-09-2017.13:47:25	NSX-Manager	192.168.21.18	>>DN<<
08-09-2017.13:47:26	Controller-1	192.168.21.19	<<UP>>
08-09-2017.13:47:27	Controller-2	192.168.21.20	<<UP>>
08-09-2017.13:47:28	Controller-3	192.168.21.21	<<UP>>
08-09-2017.13:47:29	PSC-1	192.168.21.23	<<UP>>
08-09-2017.13:47:30	PSC-2	192.168.21.24	<<UP>>
08-09-2017.13:47:31	LCM-Repo	192.168.21.25	>>DN<<
08-09-2017.13:47:33	LCM-Backup	192.168.21.26	>>DN<<

Figure 30: VCF Authentication Stack vCenter comes back online

The vCenter and PSC VMs come online and once the vCenter DB is online, full access to all authentication services for VCF are restored.

Now the rest of the VMs can be restored, as the vCenter is back online, and Cohesity can communicate to it directly and restore all the VMs back to source.

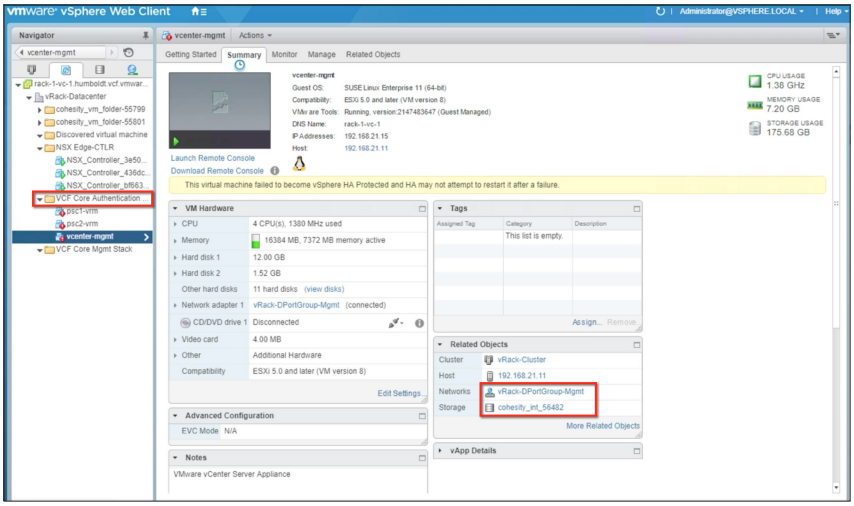


Figure 31: VCF Authentication Stack vCenter comes back online (continued)

Recover rest of the VCF management VMs

Search for the VCF Core Mgmt Stack job that was created previously to protect rest of the Mgmt VMs.

Select the job to restore all VMs in the job to a desired RPO.

Note: Since the SDDC Manager and ISVMs running a distributed NoSQL DB were all protected in quick succession, the simultaneous restore of SDDC Manager and ISVMs allow the SDDC Manager application to come online without any consistency issues or any intervention required to manually recover the NoSQL DB. This applies to VCF 2.1.x. In VCF 2.2, SDDC Manager and ISVMs are consolidated into a single VM.

To recover the VCF Core Mgmt Stack, follow steps 1 through 3 below and then use defaults, as now with the vCenter and PSC VMs already online, the rest of the VMs can be restored to the original vSAN database with all the default networking configuration.

Once the admin clicks on Finish, the following tasks are performed for a successful recovery of all VMs:

- Step 1: User manually triggers a Cohesity VM recovery task and selects snapshot, target, networking settings, VM name, target data store.
- Step 2: Cohesity contacts VMware endpoint to validate current inventory and chosen recovery task settings.
- Step 3: Cohesity creates an internal view and clones the VM snapshot and mounts the view to the target ESXi host(s).
- Step 4: Create a new VM object using original VM configuration file and chosen recovery settings. Network configuration changes take place at this step.
- Step 5: VM is (optionally) powered on (Note that the VM is now available for use).
- Step 6: Storage vMotion is initiated to move the data store from the Cohesity cluster to the primary data store.
- Step 7: Storage vMotion completes, VMware non-disruptively migrates data store access from the Cohesity cluster snapshot to the primary data store.
- Step 8: Cohesity requests the data store to unmount.
- Step 9: ESXi host unmounts data store.
- Step 10: Cohesity releases the view.

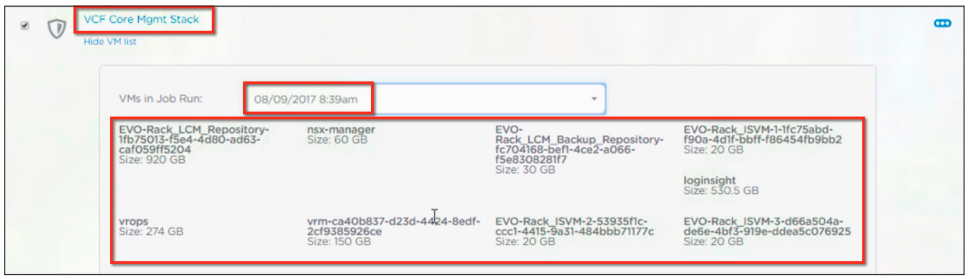


Figure 32: VCF Management Stack VMs Restore Task

Note: As Cohesity uses its patented snapshot technology SnapTree to protect the VMs, they can be restored from any RPO instantaneously because each of the Cohesity snapshots is a fully hydrated snap. Unlike other incumbent solutions which

can only restore instantly from the last snap, Cohesity snaps are full snaps and do not need to be hydrated at restore. This drastically reduces the "time to application liveness."

Now that the vCenter is already online, there is no need to select an alternate location or network port group selection for the subsequent restore workflow. All the VMs can be restored back to the original location with the original network settings.

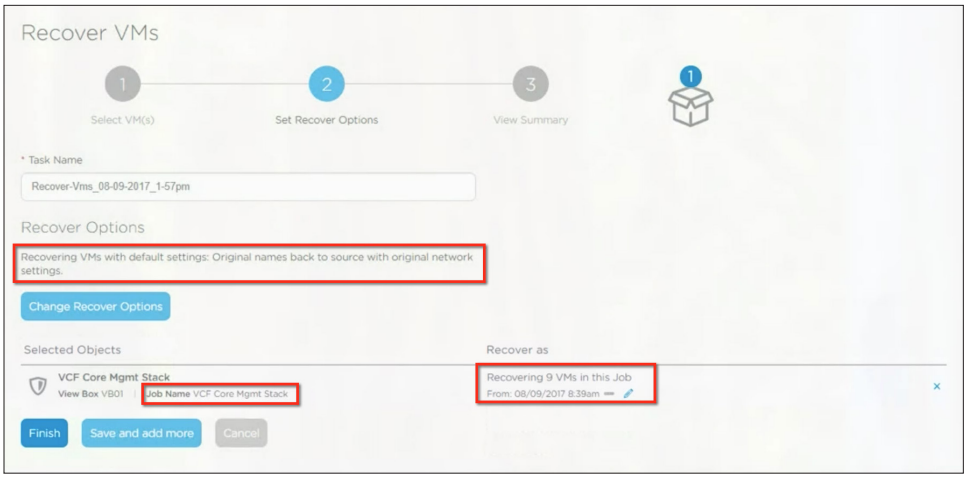


Figure 33: VCF Management Stack VMs Restore Task (continued)

After the restore was triggered, it took less than five minutes to restore and bring all the VMs back online on the network and start up their respective applications.

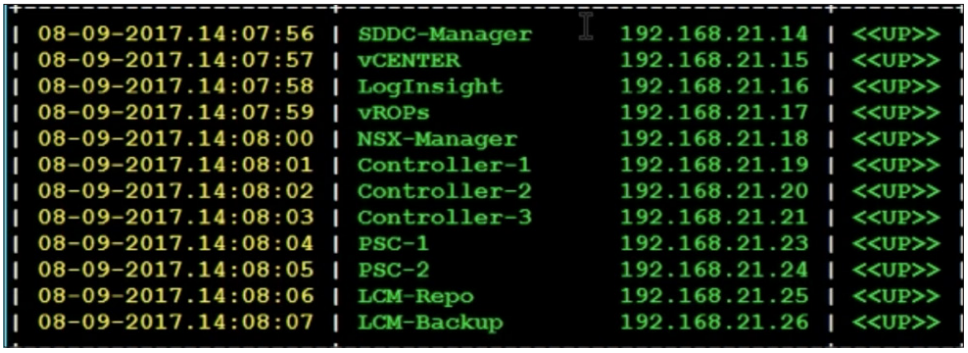


Figure 34: VCF Management Stack VMs back Online

vCenter shows all the VMs powered on and the SDDC manager online. When the restore is in progress, the VMs are being relocated from the Cohesity data store (where they are initially powered on) back to the primary vSAN data store in the background using storage vmotion.

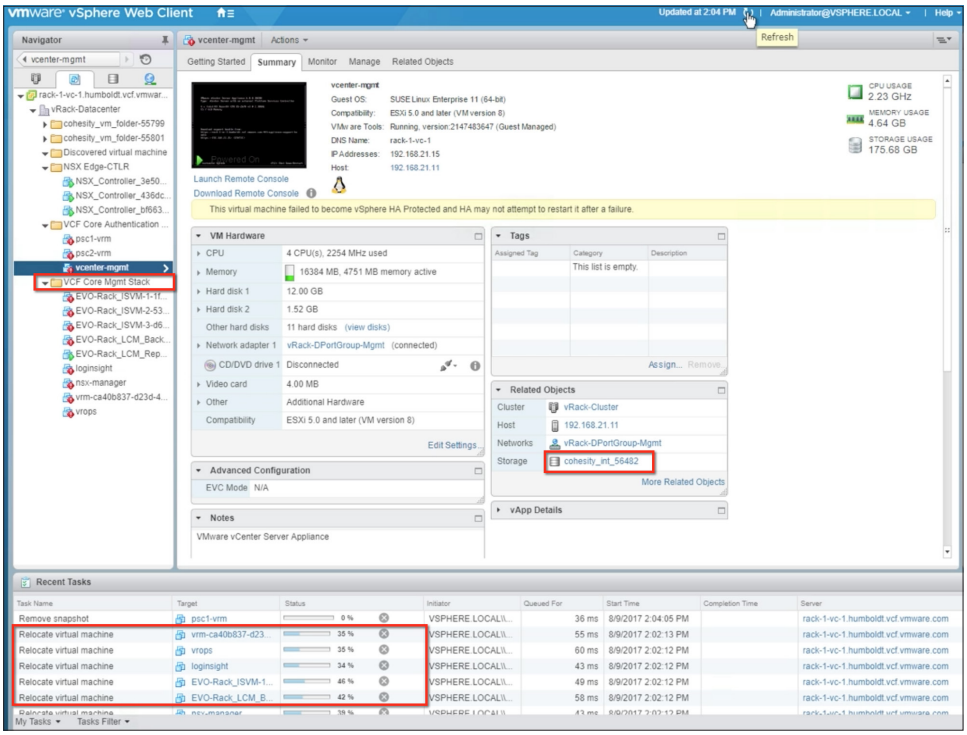


Figure 35: VCF Management Stack VMs Storage vMotioned back to vSAN Primary Datastore

As a final step, the vCenter and PSC VMs which were restored to the individual host can now be manually migrated back to the original vSAN data store. The storage vMotion was unavailable initially because storage vMotion is not available on an ESXi host, and is only packaged with vCenter.

Conclusion

We have demonstrated what is possible with next-generation architectures built with resiliency and redundancy from the ground up, which provides granular access and recovery capability to VM/files/folders/objects from the protected dataset.

The resiliency and highly available architecture of VCF combined with the end-to-end data protection and secondary storage capabilities of Cohesity make implementing, managing, and protecting on-premises infrastructure simple. This allows customers to focus on their revenue-generating applications and providing their end customers value by safeguarding their investments and resources.

VMware and Cohesity are both focused on providing cloud-like simplicity on premises and also enabling the journey of adoption to the cloud. This is achieved by extending VCF and data protection from on premises to the cloud for disaster recovery while enabling the data to be leveraged in the cloud for Devops and other secondary storage use cases.

Contributing Authors

VEN IMMANI
TECHNICAL PRODUCT MANAGER
VMWARE, INC., VMWARE CLOUD FOUNDATION
Ven started out running an ISP out of his bedroom in 1998. He has since been immersed in technology related to developing test software, networked and distributed storage, high performance computing, converged architecture, performance tuning/competitive benchmarking, enterprise networking, hyper-converged public cloud/private cloud architecture, and generally anything related to data center infrastructure.

ANIL KAPUR
DIRECTOR PRODUCT MANAGEMENT
VMWARE, INC., VMWARE CLOUD FOUNDATION
Anil is the Director of Product Management for VMware Cloud Foundation. He has led, defined, and delivered innovative hardware and software infrastructure products to businesses worldwide. Anil has served in product management, architect, and R&D engineering roles in both the network/telecom (Cisco, Qualcomm) and SaaS space.

VIVEK AGARWAL
HEAD OF CORPORATE AND BUSINESS DEVELOPMENT
COHESITY, INC.
Vivek brings vast technology experience, having held varying roles as a startup founder, technology banker, and management consultant. He was a banker in Credit Suisse's Technology Investment Banking Group, where he was involved in transactions totaling several billion dollars and his clients included several storage companies such as Commvault, 3Par, and EMC, among others.

DAMIEN PHILIP
PRINCIPAL SOLUTIONS ARCHITECT
COHESITY, INC.
Principal Solutions Architect responsible for building solutions with industry technology partners and helping solve data center challenges for enterprises. As the 1st Solution Architect for Cisco UCS, Damien was responsible for building the initial Reference Architecture which later was used to create the vBlock. He has been in the technology industry for more than 20 years and is a practitioner with hands-on expertise in Operating Systems, Data Storage (i.e., SAN, NAS, DAS, Network, and Distributed storage), and Converged, Hyper-Converged, and Public Cloud/Private Cloud infrastructures.

RAWLINSON RIVERA
GLOBAL FIELD CHIEF TECHNOLOGY OFFICER
COHESITY, INC.
Rawlinson is well known as an industry thought leader on cloud enterprise architectures and hyper-converged infrastructures. Rawlinson recently joined Cohesity's leadership team after 10 years at VMware, where he was most recently the Principal Architect working in the Office of CTO for the Storage and Availability Business Unit. Focused on defining and communicating Cohesity's product vision and strategy, Rawlinson will be working closely with the leadership, engineering, and product teams to provide customer insights and use cases for feature enhancements and future roadmap prioritization. He is the author of the popular blog PunchingClouds, which can be followed on Twitter via @PunchingClouds.

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com