COHESITY | cisco

Cohesity and Cisco HyperFlex Reference Architecture

Table of Contents

Executive Summary

Hyperconvergence is becoming a norm in data centers today. Companies adopting this next generation infrastructure have realized significant savings in total cost of ownership (TCO). These savings are the result of vastly simplified architectures, lower power and cooling needs, workload consolidation, smaller hardware footprint and "pay as you grow" consumption model.

What ties all these benefits together is the simplicity of managing these web-scale platforms from a single UI. The design principles of distributed control and data planes that eliminate complexities in infrastructure and management make hyperconverged architectures attractive and bring overall value to end customers.

Until now hyperconvergence has focused largely on primary storage workloads. However, the principles of hyperconvergence can be productively applied to secondary storage environments as well. As per Gartner, data volume is set to grow 800% over the next 5 years and 80% of it will reside as unstructured data. IDC forecasts that by 2025, the global datasphere will grow to 163ZB (that is a trillion gigabytes).

Companies struggle to maintain control over a sprawling landscape of secondary storage point solutions that waste resources with redundant data copies across backup, test/dev, file services and analytic workloads,and it gets further complicated as cloud infrastructure is added to the equation. A consolidated, platform-first approach to secondary storage allows for consolidation of all secondary storage use cases into a distributed, web-scale, highly-efficient, intelligent and cost-effective infrastructure.

As seen below, most current enterprise data center storage landscapes resemble this architecture, with complexities built into every aspect of data management because of the point solutions and multiple redundant copies of data being stored in each of these point solutions.

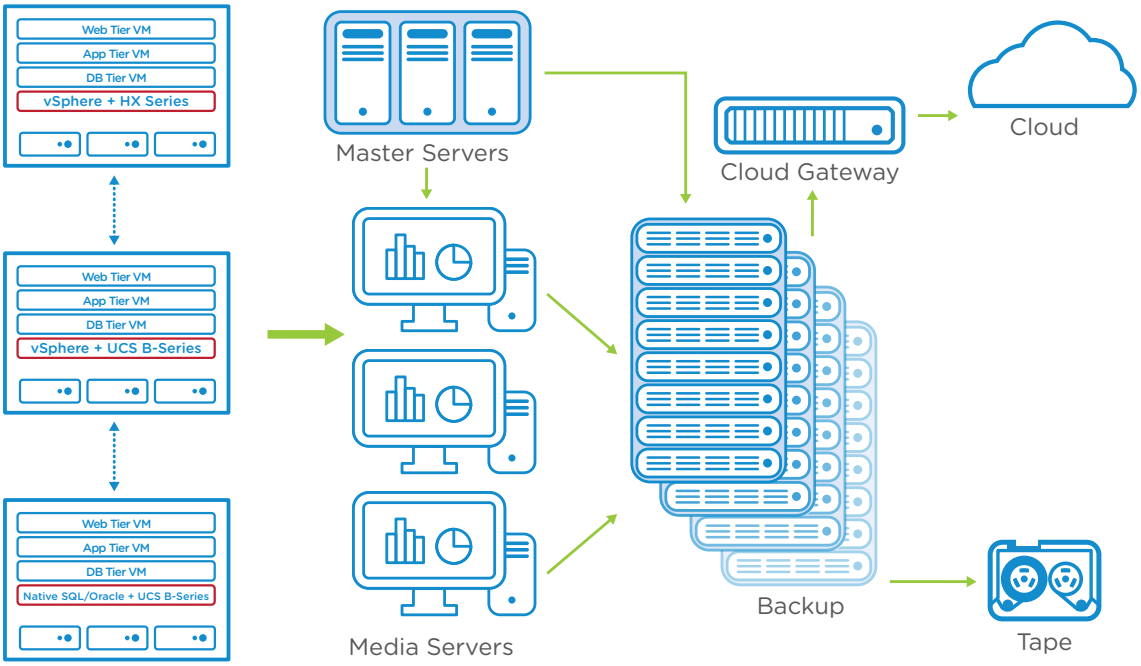


Figure 1: Legacy Enterprise Data Center

This causes the issues of sprawl, islands of fragmented storage and security concerns with the introduction of dark data.

In this Reference Architecture we will discuss the Cisco HyperFlex primary storage solution combined with the Cohesity data protection and secondary storage subsystem. Bringing both these platforms together under a single holistic solution allows to further simplify implementation of an already simple hyperconverged compute/network/storage and drastically reduce the storage management and end to end lifecycle of data management by seamlessly combining primary virtualization workloads and subsequent secondary storage use cases of backups, DR, test/dev, file shares, archiving and analytics under a single fully integrated solution. By reducing the storage platforms for all the above use cases from typically 4-5 in a traditional architecture to 2 platforms simplifies implementation and also overall manageability in the long term with this new approach and next generation architecture.

Cohesity, has created a hyperconverged secondary storage platform that offers a radically efficient solution to fragmented data silos by unifying all secondary use cases on a single, infinitely scalable platform. Cohesity is a true Software Defined Storage System that enables policy-based provisioning, management of data storage independent of the underlying hardware and data reusability. This approach aims to change physical storage hardware without disrupting application access and to give users the flexibility in terms of hardware choice to build datacenter storage services with lower costs. Cisco Hyperflex System with Cisco UCS is a best-of-breed solution that is the perfect answer for customers requiring an advanced, enterprise-class data availability solution for their virtual environments that is simple to order, deploy and manage. In addition, it can easily expand over time as the customer's need increases. It provides instant and reliable recovery of virtualized applications and data bringing virtual machine backup and replication together in a single solution with award-winning support. Together the solution provides the greatest scalability, with simplest administration, at lowest cost.

This integrated solution is specifically designed to address the infrastructure challenges, operational complexities, data management challenges and fragmentation that persist across secondary storage silos within enterprise data centers today.

Business Case

The combined solution aims to provide an order-of-magnitude additional value over existing Hyperconverged solutions. Cisco HyperFlex simplifies the implementation, management and reduces the operational costs of managing HW/SW. It unifies network fabric and computing resources in a next-generation data platform which increases operational efficiency and adaptability to more workloads in a customer's data center. Cohesity delivers a web-scale platform that consolidates all secondary storage and data services on to one unified, efficient solution. Cohesity simplifies data protection, consolidates file and object services, provides instant access to test/dev copies, and performs in-place searches and analytics. All on a software-defined platform that spans from the edge to the cloud. The solution significantly extends the value of secondary storage by 'operationalizing' stored data.

Cohesity's offers several key benefits including:

- Introduces an extremely performant enterprise-class hardware supporting a Software Defined Storage platform that can scales-out limitlessly
- Reduces data center footprint by consolidating islands of secondary storage
- Simplifies administration by managing all secondary storage workloads through a single pane of glass
- Eliminates over provisioning with huge upfront CapEx investment as customers can deploy storage in a pay-as-you-grow model
- Achieves comprehensive data protection and recovery, onsite, offsite and in the cloud
- Brings compute to data - instant access to all stored data for Test/Dev and search and analytics
- Provides full visibility to stored data with native storage analytics

Cisco HyperFlex HX Data Platform has the following benefits :

- Unify fabric network and computing technology in a next-generation data platform.
- Gain high performance and storage longevity with a file system built for hyperconvergence.
- Scale nodes, computing, and capacity independently in HyperFlex clusters as needs change.
- Increase capacity use while maintaining performance with data deduplication and compression.

Intended Audience

This document describes the solution reference architecture for Cisco HyperFlex and Cohesity DataPlatform. It focuses on technical and business benefits of the combined solution. This document is intended for Sales Engineers, Partners, IT professional and customers who are interested in learning about and deploying Cisco HyperFlex and Cohesity DataPlatform for Tier 1 through Tier 4 storage use cases i.e hyperconverged high performance virtualized workloads, backup & restore, archiving, file services with NFS,SMB/CIFS and S3 backed object storage.

Solution Overview

HyperFlex Overview

The Cisco HyperFlex solution delivers next-generation hyperconvergence in a data platform to offer end-to-end simplicity for faster IT deployments, unifying computing, networking, and storage resources. The Cisco HyperFlex solution is built on the Cisco Unified Computing System™ (Cisco UCS®) platform and adheres to a data center architecture supporting traditional, converged, and hyperconverged systems with common policies and infrastructure management. The Cisco HyperFlex HX Data Platform is a purpose-built, high-performance, distributed file system delivering a wide range of enterprise-class data management and optimization services. This platform redefines distributed storage technology, expanding the boundaries of hyperconverged infrastructure with its independent scaling, continuous data optimization, simplified data management, and dynamic data distribution for increased data availability. This agile system is easy to deploy and manage, scales as your business needs change, and provides the first level of data availability.

Cohesity Overview

Cohesity enables enterprises to take control of their increasingly complex storage environments through a hyperconverged secondary storage infrastructure. The Cohesity DataPlatform can consolidate multiple use cases such as data protection, test/dev, file services and analytics onto a single web-scale platform. Today, 80% of enterprise data resides in siloed secondary storage systems. These islands of storage create inefficiency, waste and overhead expense. Cohesity consolidates silos into a highly resilient and fully redundant, web scale storage fabric, dramatically improving IT productivity and reducing long-term storage operating costs.

Cohesity DataPlatform, integrated with Cisco Unified Computing System (Cisco UCS) maintains all the value that comes from secondary storage hyperconverged while providing enterprise-class hardware for added protection and performance. Cisco UCS unifies computing, networking, management, and storage access into a single integrated physical architecture that can enable end-to-end server visibility, management and control in both bare-metal and virtualized environments. Cohesity and Cisco created this joint solution by integrating the Cohesity software onto the Cisco UCS C240-M4 rack server. The UCS C240-M4 is an enterprise-class server designed to deliver exceptional performance, expandability and efficiency for storage and I/O-intensive infrastructure workloads. Cisco UCS nodes provide a robust physical layer that has compute, memory, NVMe flash, HDDs and networking built into them. This makes this platform best suited for big data analytics, virtualization graphics-rich and bare-metal applications. The C240-M4 server can be deployed standalone or as part of the Cisco Unified Computing System (UCS).

Cohesity complements Cisco data-center stack with a distributed file system software architecture that is designed for high availability. The nodes have shared-nothing topology and there is no single point of failure or inherent bottlenecks. Consequently both performance and capacity can scale linearly as more physical nodes are added to the cluster. The distributed file system spans across all nodes in the cluster and natively provides global deduplication, compression and encryption.

In addition, Copy Data Management is built into Cohesity's DataPlatform. It allows for creation of zero-space instant writeable clones from the backups and can provision and orchestrate test & development environments. The entire platform is exposed through a single pane of glass that manages all secondary storage workflows spanning multiple clusters, geographical sites and even the cloud.

By consolidating all the Tier 1-4 storage workflows into a single unified scale-out web scale platform, Cohesity and Cisco together significantly reduce TCO and improve business agility for enterprise customers. The solution eliminates fragmentation, significantly improves storage efficiency and with its inbuilt MapReduce based analytics, shines light into data which otherwise would have remained dark.

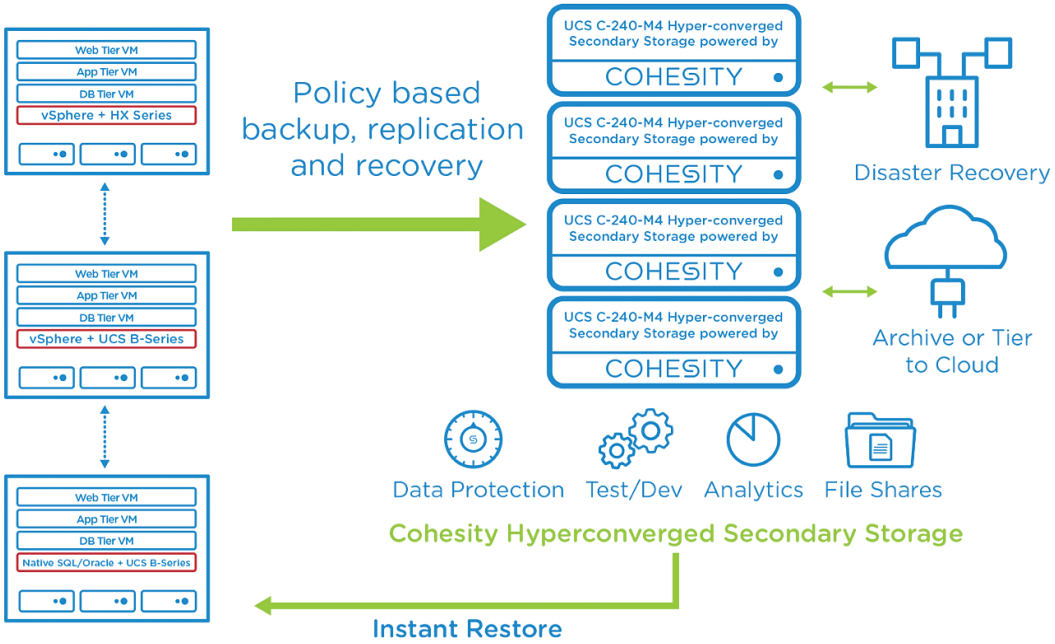


Figure 2: Cohesity Hyperconverged Secondary Storage

Cohesity and Cisco HyperFlex Architecture and Concepts
Cohesity DataPlatform

Cohesity DataPlatform couples Cisco UCS C240-M4 hardware with intelligent, extensible software, enabling customers to spend less time worrying about how to retrofit their legacy solutions with future needs, and more time focusing on the core functions of the enterprise.

The Cohesity DataPlatform is built on a file system that combines infinite scalability with an open architecture flexibility that can consolidate multiple business workloads on a single platform. With built-in, native applications to support data protection, copy data management, test and development, and in-place analytics, customers experience the benefits of consolidation right out of the box. This filesystem was built from the ground up to be the most robust and fully distributed system in the market. Unlike traditional distributed file systems that are 'Eventually Consistent', Cohesity distributed file system leverages a purpose-built noSQL store, combined with Paxos protocols, that delivers Strict Consistency with the ability to make decisions rapidly, at massive scale, and without performance penalties.

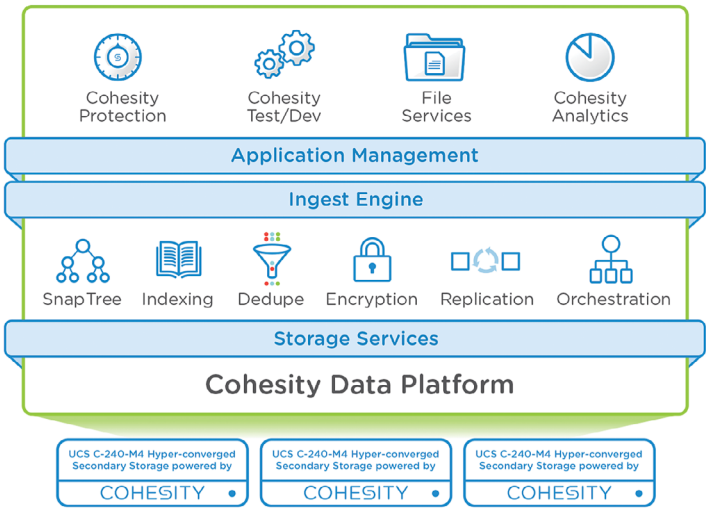


Figure 3: Cohesity Software Defined Storage Platform

The file system is comprised of several services, each one handling a key function to provide an infinitely scalable architecture while optimizing performance enabling the consolidation of multiple workloads.

Cluster Manager controls all the core services that run on a Cohesity Cluster. This layer is responsible for maintaining all configuration information, networking information, and the general state of all other components in the system.

I/O Engine is responsible for all read and write operations that take place on the cluster. It is comprised of the write cache, which lives in SSD, and the tiered storage layers that span across both SSD and spinning disk. For write operations, as data is streamed into the system, it is broken down into smaller chunks, which are optimally placed onto the tier that best suits the profile of that particular chunk. The I/O Engine also ensures that all data is written to two nodes concurrently, providing write fault tolerance. This enables completely non-disruptive operations, even if a node were to become unavailable during a given operation. For read operations, the I/O Engine receives the location information of the data from the Distributed Metadata Store and fetches the associated chunk(s). If a particular chunk of data is frequently requested in a short period of time, that chunk is kept in SSD to ensure quick access and optimized performance on subsequent requests.

Metadata Store is a consistent key value store that serves as the file system metadata storage repository. Optimized for quick retrieval of file system metadata, the Metadata Store is continually balanced across all nodes within the cluster (accounting for nodes that are added or removed from the cluster). The Metadata Store ensures that three copies are maintained at any point in time, so that data is always protected, even in the event of a failure.

Cohesity Storage Services

The next layer in the Cohesity DataPlatform architecture consists of the Cohesity Storage Services, which provide the storage efficiency capabilities that customers depend on at a scale that no other solution can achieve.

Snapshots

Cohesity SnapTree™ for managing data copies: In legacy storage solutions, snapshots (of a file system at a particular given point in time) form a chain, tracking the changes made to a set of data and form the basis for organizing and storing copies of data. Every time a change is captured, a new link is added to the chain. As these chains grow with each and every snapshot, the time it takes to retrieve data on a given request grows because the system must re-link the chain to access that data. Cohesity's patented SnapTree™ technology creates a tree of pointers that limits the number of hops it takes to retrieve blocks of data, regardless of the number of snapshots that have been taken. SnapTree uses a B+ tree data structure such that access to any point in the tree takes a fixed number of hops no matter how many snapshots there are, without having to rebuild any chain linkage. Because SnapTree is implemented on a distributed file system, every node sees the same nested structure of the chain with a fixed depth independent of where the actual data is stored in the cluster. This is what allows the platform to keep snapshots fully hydrated. Keeping the snapshots fully hydrated improves the recovery times of any snapshot from t0 to tn because it does not incur the time penalty of traversing the entire chain of changes (Figure 6). This capability is available with the Integrated Cohesity DataProtect solution.

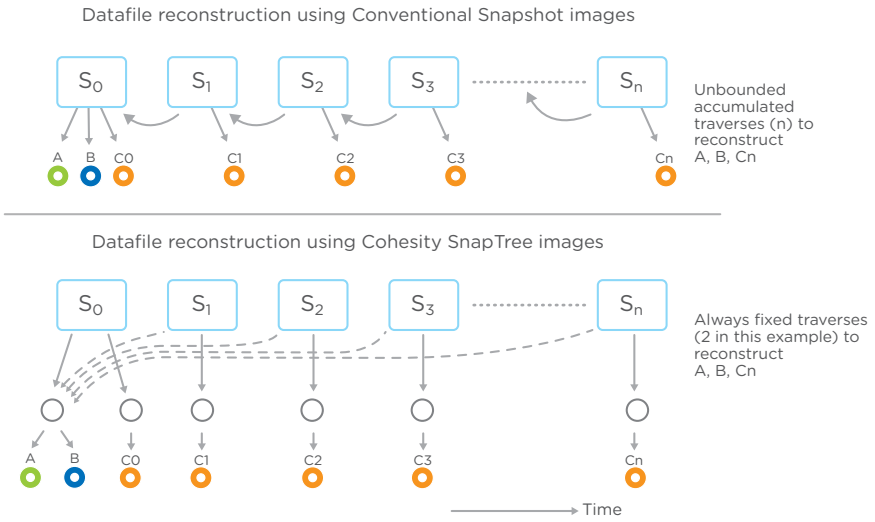


Figure 4: Cohesity's SnapTree™ powered snapshots vs. traditional snapshots

Global deduplication and compression

Cohesity leverages a unique, variable-length data deduplication technology that spans an entire cluster, 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. See figure below for further clarification.

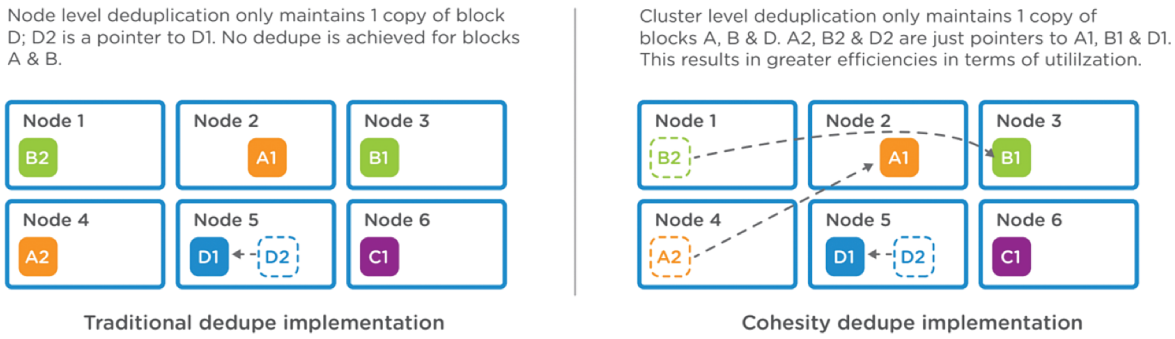


Figure 5: Cohesity's global deduplication across all nodes in a cluster results in less storage consumed compared to just node level deduplication used in several other in market data protection solutions

Figure 5 : Cluster level deduplication

Cohesity encryption engine

Cohesity DataPlatform also provides encryption of data at rest and in transit over the network with AES 256-bit encryption to secure data. Encryption in flight is applicable to data that is replicated to a remote Cohesity cluster or when data is tiered/archived to the cloud from the Cohesity DataPlatform. This ensures that data stored on the Cohesity cluster is protected well from malicious attacks.

The system provides full software-based encryption that is hardware-accelerated through the latest Intel processors. With hardware acceleration, the software-based encryption route has become much faster (in the order of several GB/s), and thus has minimal impact on performance. While the file system also takes advantage of self-encrypting drives (SED), full software encryption provides greater flexibility for the enterprise customers. Setting up encryption on the cluster takes just selecting a checkbox.

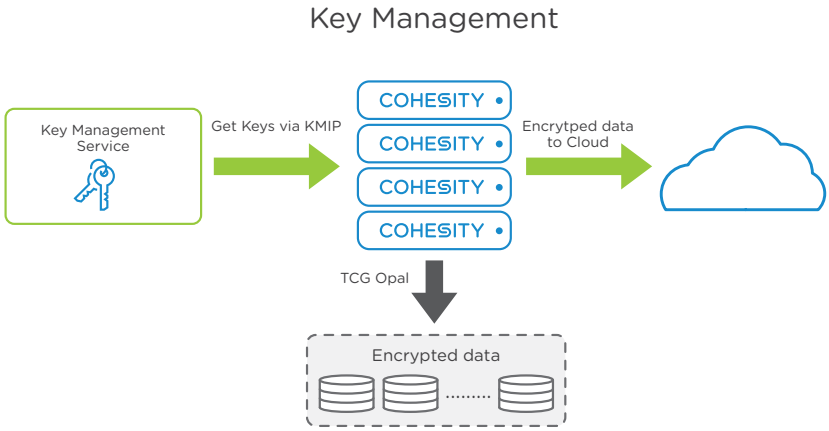
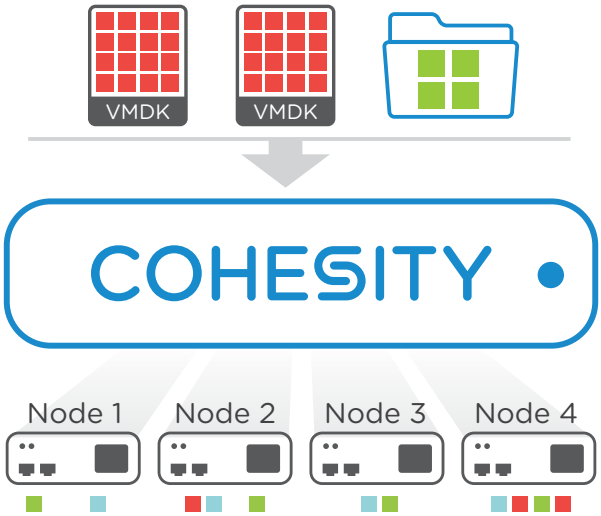


Figure 6: Key management implementation in Cohesity's encryption engine

Intelligent Data Placement

Intelligent data placement ensures that data is always available, even in the event of a node failure. When data is written to a Cohesity Cluster, a second copy of that data is instantly replicated to another node within the cluster. For customers who have multiple solution Blocks (a chassis with one or more UCS nodes) or racks, Cohesity will always optimize the placement of the data by placing the second copy on a different block or in a different rack, providing an even higher level of fault tolerance. For customers with stricter fault tolerance requirements, the Replication Factor (RF), or number of copies of data that are replicated within a Cluster, can be adjusted to fit their needs. This intelligent placement, or sharding, of data also enhances the performance characteristics of the data placed on the cluster. As the data hits the cluster, it is broken down into smaller bite-sized chunks (typically 8K to 24K). By spreading data across all nodes of a cluster the I/O load is shared across all available resources and eliminates the notion of a 'Hot Node' or 'Hot Disk' which would get accessed more frequently and would create an I/O bottleneck.



As data is ingested into OASIS it is evenly distributed across the available nodes of the cluster. This reduces the notion of 'Hot Nodes' or 'Hot Disks' which plague systems that keep an entire copy of the object on a single node.

Figure 7: Intelligent data placement across nodes for fast access and availability through node failures

Indexing Engine

The Indexing Engine is responsible for inspecting the data that is stored in a cluster. On its first pass, the Indexing Engine grabs high-level indices for quick data retrieval around top-level objects, such as Virtual Machine (VM) names and metadata. On its second pass, the Indexing Engine cracks open individual data objects, such as Virtual Machine Disks (VMDKs), and scans individual files within those data objects. This native indexing enables rapid search-and-recover capabilities to quickly find and restore files stored within higher-level data objects such as VMs.

Cisco Unified Computing System

Cisco brings 30 years of breadth, leadership, and vision to guide businesses through networking and infrastructure challenges. Cisco's Unified Computing System (UCS) continues Cisco's long history of innovation in delivering integrated systems that deliver business results. Cisco UCS integrated infrastructure solutions speed up IT operations today and create the modern technology foundation needed for the critical business initiatives of tomorrow.

The Cisco Unified Computing System is a next-generation data center platform that unites compute, network, and storage access. The platform, optimized for virtual environments, is designed using open industry-standard technologies and aims to reduce total cost of ownership (TCO) and increase business agility. The system integrates a low-latency, lossless 10 Gigabit Ethernet unified network fabric with enterprise-class, x86-architecture servers. It is an integrated, scalable, multi chassis platform in which all resources participate in a unified management domain.

The main components of Cisco Unified Computing System are:

- Computing—The system is based on an entirely new class of computing system that incorporates rack-mount and blade servers based on Intel Xeon Processors.

- Network—The system is integrated onto a low-latency, lossless, 10-Gbps or 40-Gbps unified network fabric. This network foundation consolidates LANs, SANs, and high-performance computing networks which are separate networks today. The unified fabric lowers costs by reducing the number of network adapters, switches, and cables, and by decreasing the power and cooling requirements.
- Virtualization—The system unleashes the full potential of virtualization by enhancing the scalability, performance, and operational control of virtual environments. Cisco security, policy enforcement, and diagnostic features are now extended into virtualized environments to better support changing business and IT requirements.
- Storage access—The system provides consolidated access to both (Storage Area Network) SAN storage and Network Attached Storage (NAS) over the unified fabric. By unifying the storage access the Cisco Unified Computing System can access storage over Ethernet, Fibre Channel, Fibre Channel over Ethernet (FCoE), and iSCSI. This provides customers with choice for storage access and investment protection. In addition, the server administrators can pre-assign storage-access policies for system connectivity to storage resources, simplifying storage connectivity, and management for increased productivity.
- Management—The system uniquely integrates all system components which enable the entire solution to be managed as a single entity by the Cisco UCS Manager (UCSM). The Cisco UCS Manager has an intuitive graphical user interface (GUI), a command-line interface (CLI), and a robust application programming interface (API) to manage all system configuration and operations.

The Cisco Unified Computing System is designed to deliver:

- A reduced Total Cost of Ownership and increased business agility.
- Increased IT staff productivity through just-in-time provisioning and mobility support.
- A cohesive, integrated system which unifies the technology in the data center. The system is managed, serviced and tested as a whole.
- Scalability through a design for hundreds of discrete servers and thousands of virtual machines and the capability to scale I/O bandwidth to match demand.
- Industry standards supported by a partner ecosystem of industry leaders

Cisco HyperFlex Converged Data Platform Software

The Cisco HyperFlex HX Data Platform is a purpose-built, high-performance, distributed file system with a wide array of enterprise-class data management services. The data platform's innovations redefine distributed storage technology, exceeding the boundaries of first-generation hyperconverged infrastructures. The data platform has all the features that you would expect of an enterprise shared storage system, eliminating the need to configure and maintain complex Fibre Channel storage networks and devices. The platform simplifies operations and helps ensure data availability. Enterprise-class storage features include the following:

- **Replication** of all written data across the cluster so that data availability is not affected if single or multiple components fail (depending on the replication factor configured).
- **Deduplication** is always on, helping reduce storage requirements in which multiple operating system instances in client virtual machines result in large amounts of duplicate data.
- **Compression** further reduces storage requirements, reducing costs, and the log- structured file system is designed to store variable-sized blocks, reducing internal fragmentation.
- **Thin provisioning** allows large volumes to be created without requiring storage to support them until the need arises, simplifying data volume growth and making storage a "pay as you grow" proposition.
- **Fast, space-efficient clones** rapidly replicate virtual machines simply through metadata operations.
- **Snapshots** help facilitate backup and remote-replication operations: needed in enterprises that require always-on data availability.

Cisco HyperFlex HX Data Platform Administration

The Cisco HyperFlex HX Data Platform is administered through a VMware vSphere web client plug-in or through an HTML5 interface called HyperFlex Connect. Through this centralized point of control for the cluster, administrators can create datastores, monitor the data platform health, and manage resource use. Administrators can also use this data to predict when the cluster will need to be scaled.

Cisco HyperFlex HX Data Platform Controller

A Cisco HyperFlex HX Data Platform controller resides on each node and implements the Cisco HyperFlex HX Distributed Filesystem. The storage controller runs in user space within a virtual machine intercepting and handling all I/O from guest virtual machines. The storage controller VM uses the VMDirectPath I/O feature to provide PCI pass-through control of the physical server's SAS disk controller. This method gives the controller VM full control of the physical disk resources, utilizing the SSD drives as a read/write caching layer, and the HDDs or SSDs as a capacity layer for distributed storage. The controller integrates the data platform into VMware software through two pre-installed VMware ESXi vSphere Installation Bundles (VIBs):

- **IOvisor:** The IOvisor is deployed on each node of the cluster and acts as a stateless NFS proxy that looks at each IO request and determines which cache vNode it belongs to and routes the IO to the physical node that owns that cache vNode. In the event of a failure, the IOvisor transparently handles it and will retry the same request to another copy of the data based on new information it receives. Decoupling the IOvisor from the controller VM enables access to the distributed filesystem and prevents hotspots. Compute-only nodes, and VMs continue to perform storage IO in the event of a disk, SSD, or even a storage controller failure.
- **VMware API for Array Integration (VAAI)**: This storage offload API allows vSphere to request advanced file system operations such as snapshots and cloning. The controller implements these operations through the manipulation of metadata rather than actual data copying, providing rapid response, and thus rapid deployment of new environments.

Data Operations and Distribution

The Cisco HyperFlex HX Data Platform controllers handle all read and write operation requests from the guest VMs to their virtual disks (VMDK) stored in the distributed datastores in the cluster. The data platform distributes the data across multiple nodes of the cluster, and across multiple capacity disks of each node, per the replication level policy selected during the cluster setup. This method avoids storage hotspots on specific nodes, and on specific disks of the nodes, and thereby also avoids networking hotspots or congestion from accessing more data on some nodes versus others.

Replication Factor

Enterprise class hyperconverged solutions should have three copies of data blocks across any three data nodes. This helps to ensure high availability during rare failure events such as single node failure and disk failure or during software and firmware upgrades, performed on a HX System. Thus three copies, or a replication factor of 3 (RF=3), is a default setting and also a recommended best practice for HyperFlex systems.

- Replication Factor 3: For every I/O write committed to the storage layer, 2 additional copies of the blocks written will be created and stored in separate locations, for a total of 3 copies of the blocks. Blocks are distributed in such a way as to ensure multiple copies of the blocks are not stored on the same disks, nor on the same nodes of the cluster. This setting can tolerate simultaneous failures of 2 disks, or 2 entire nodes without losing data and resorting to restore from backup or other recovery processes.
- Replication Factor 2: For every I/O write committed to the storage layer, 1 additional copy of the blocks written will be created and stored in separate locations, for a total of 2 copies of the blocks. Blocks are distributed in such a way as to ensure multiple copies of the blocks are not stored on the same disks, nor on the same nodes of the cluster. This setting can tolerate a failure of 1 disk, or 1 entire node without losing data and resorting to restore from backup or other recovery processes.

Data Write Operations

For each write operation, data is written to the local caching SSD on the node where the write originated, and replica copies of that write are written to the caching SSD of the remote nodes in the cluster, per the replication factor setting. For example, at RF=3 a write will be written locally where the VM originated the write, and two additional writes will be committed in parallel on two other nodes. The write operation will not be acknowledged until all three copies are written to the caching layer SSDs. Written data is also cached in a write log area resident in memory in the controller VM, along with the write log on the caching SSDs (Figure 8). This process speeds up read requests when reads are requested of data that has recently been written.

Data Destaging, Deduplication and Compression

The Cisco HyperFlex HX Data Platform constructs multiple write caching segments on the caching SSDs of each node in the distributed cluster. As write cache segments become full, and based on policies accounting for I/O load and access patterns, those write cache segments are locked and new writes roll over to a new write cache segment. The data in the now locked cache segment is destaged to the HDD capacity layer of the node. During the destaging process, data is deduplicated and compressed before being written to the HDD capacity layer. The resulting data after deduplication and compression can now be written in a single sequential operation to the capacity drives of the server accomplishing the task in the minimal amount of time (Figure 8). Since the data is already deduplicated and compressed before being written, the platform avoids additional I/O overhead often seen on competing systems, which must later do a read/dedupe/compress/write cycle.

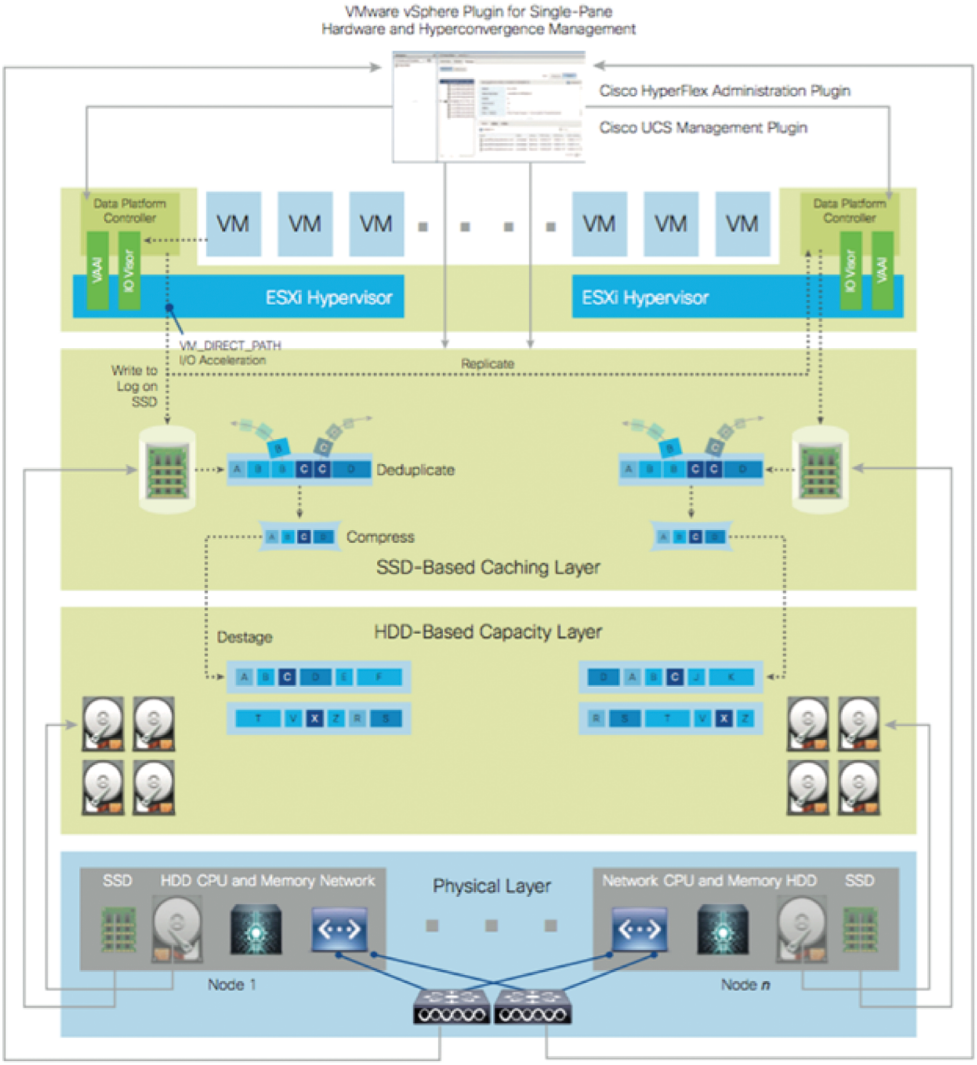


Figure 8 : HyperFlex HX Data Platform Data Movement

Use cases tested in this Reference Architecture

Data Protection

Scale-out data protection is fully baked into Cohesity's DataPlatform. Backup, recovery and DR are provided through a single unified platform eliminating complexity.

Some key features of the integrated Data Protection solution are:

- Fast, application-consistent backups: The scale-out nature allows parallelization of jobs thus minimizing backup windows. Moreover application-level integration allows for application-consistent backups.
- Remote replication and DR orchestration: Built-in remote replication protect data off-site and enables disaster recovery / migrations to remote sites.
- Policy-based management: Allows policies that specify application SLA requirements including RPO, retention policies, off-site replication and cloud archival.
- De-minimis recovery points and near-instantaneous recovery times: The zero-cost snapshots support fast 15-minute RPOs and SnapTree™ technology allows for near-instantaneous RTOs.
- Indexing engine for rapid search and recovery: Cohesity DataPlatform also includes an indexing engine that lets user quickly find and restore files stored within higher-level data objects such as VMs.
- Cloud integration: The unified platform fully leverages public or private clouds. It can can tier data to the cloud to extend storage capacity of the cluster for colder data. It can also use the cloud as a replication destination for disaster recovery. And finally the unified platform can archive data in the cloud for long-term retention.

To know more about the Data Protection solution, please read Cohesity Data Protection White Paper

Test and Development

Cohesity enables copy data management by allowing enterprises to leverage backups and accelerate time-to-market with the rapid deployment and management of test/dev environments.

- Instant, zero-space clones enable businesses to quickly spin up test/dev environments from a backup or existing file share, enabling rapid test and development from actual data without any capacity overhead.
- Workflow automation provides end-to-end orchestration for test/dev environments, automating everything from creating a clone to mounting the test/dev environment in a virtual host, all with a single click.
- Instant access to data enables fast provisioning of test and development environments, drastically reducing time and complexity associated with development environment staging times.

File Services

Cohesity's scale-out system presents itself through industry standard protocols such as SMB, NFS and S3 to become an infinitely scalable, globally deduplicated and fully searchable file services platform. It delivers storage incrementally, eliminating expensive forklift upgrades.

- Integrated support for NFS and SMB protocols and pay-as you-grow scalability enables simple growth of customer's NAS storage with built-in integration for Microsoft applications and file services.
- Inline and post process data reduction technologies that are configurable on a per workload basis combined with copy data management capabilities ensures efficient storage and usage of data.
- Automated global indexing powers Google-like search, enabling instant wildcard searches for any VM, file, or object ingested into the system. This index is fully distributed across the entire cluster and is served from the cluster's flash memory, ensuring extremely fast access to the data in the index collections.
- Authentication using Active Directory integration.

Solution Validation

The goal of this solution is to illustrate the complementary benefit of combining both the Cisco HyperFlex primary HyperConverged platform with Cohesity Data Protection and Data management solution. Both solutions together drastically simplify the storage stack from Tier 1 through Tier 4 workloads.

The design illustrates the real world use case when using a hyperconverged storage solution for primary data and applications and secondary storage providing tier 2-4 storage, i.e. file share, test/dev and backup & archive in a single storage platform.

The Cohesity solution running alongside Cisco HyperFlex is designed to be simple to deploy and offers simple scalability. The reference architecture specifically validates the integration performed deploying the backup,recovery and file services on the combine Cisco HyperFlex and Cohesity clusters. Operational tasks are validated including performing VM-centric data migration, setting up VM and application backups and recovery, dealing with various failure scenarios, and running application workloads. The solution comprises a Cohesity cluster of (4) Cisco UCS C240 M4 servers connected to a (4) node Cisco HyperFlex cluster. The Cohesity nodes are separated from the Cisco Fabric Interconnects to show a multi-tiered networking environment. This network design is the recommended best practice as the Cohesity is a backup platform for HyperFlex primary storage and virtual machines while also being a NFS/CIFS/S3 platform for backing up FI configuration and switching configurations. Being a data protection platform, Cohesity should be outside of the fault domain which it protects, so it is unaffected by failures in a primary fault domain, and can reliably service recovery of all primary data-center stack components.

Topology and Design Considerations

The following diagram illustrates the solution topology.

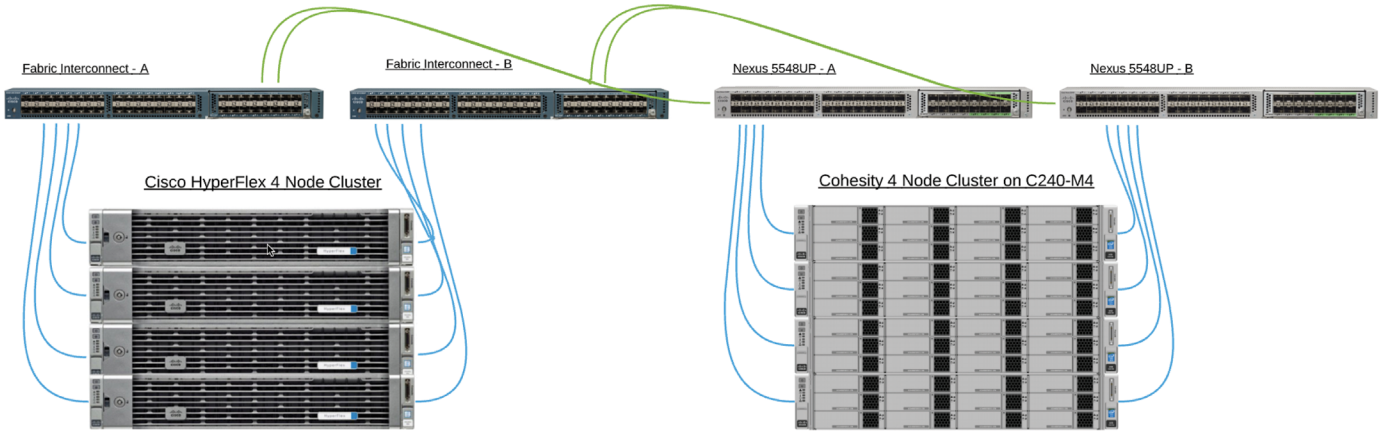


Figure 9 : Cisco UCS with Cohesity hardware topology

Cisco UCS_Cohesity Hardware specifications

The following Cisco UCS hardware SKUs have been qualified for Cohesity SDS platform.

Model	UCS C240-M4 Model 1		UCS C240-M4 Model 2	
	Per node*	Per 3-node cluster*	Per node*	Per 3-node cluster*
CPU	2 x Intel 16 Core, 2.10 GHz Broadwell	6 x Intel 16 Core, 2.10 GHz Broadwell	2 x Intel 16 Core, 2.10 GHz Broadwell	6 x Intel 16 Core, 2.10 GHz Broadwell
Memory	128GB DDR4	384GB DDR4	128GB DDR4	384GB DDR4
SSD	2 x 1.6TB PCI NVMe	9.6TB PCI NVMe	2 x 3.2TB PCI NVMe	19.2TB PCI NVMe
HDD	12 x 4TB SAS (48TB total)	144TB SAS	12 x 10TB SAS (120TB total)	360TB SAS
Network	2 x 1GbE management port	6 x 1GbE management port	2 x 1GbE management port	6 x 1GbE management port
	1 x Dual-port 10GbE SFP+ NIC	3 x Dual-port 10GbE SFP+ NIC	1 x Dual-port 10GbE SFP+ NIC	3 x Dual-port 10GbE SFP+ NIC

* A minimum of 3 nodes are required to form a Cohesity cluster; Beyond 3 nodes, Cohesity cluster can be scaled up on a per node basis.

Note : Please contact Cohesity Sales to spec out a specific BOM prior to ordering.

Cisco HyperFlex Setup and Initial Cluster Configuration

Cisco HyperFlex cluster installation uses an installation wizard to automate the setup of the physical network, virtual network, cluster nodes and cluster software. The administration simply fills out the required customized details for the cluster settings in the installer VM then the installation continues automatically. Please refer to the Cisco HyperFlex Installation Guide on Cisco.com for full details of the installer process.

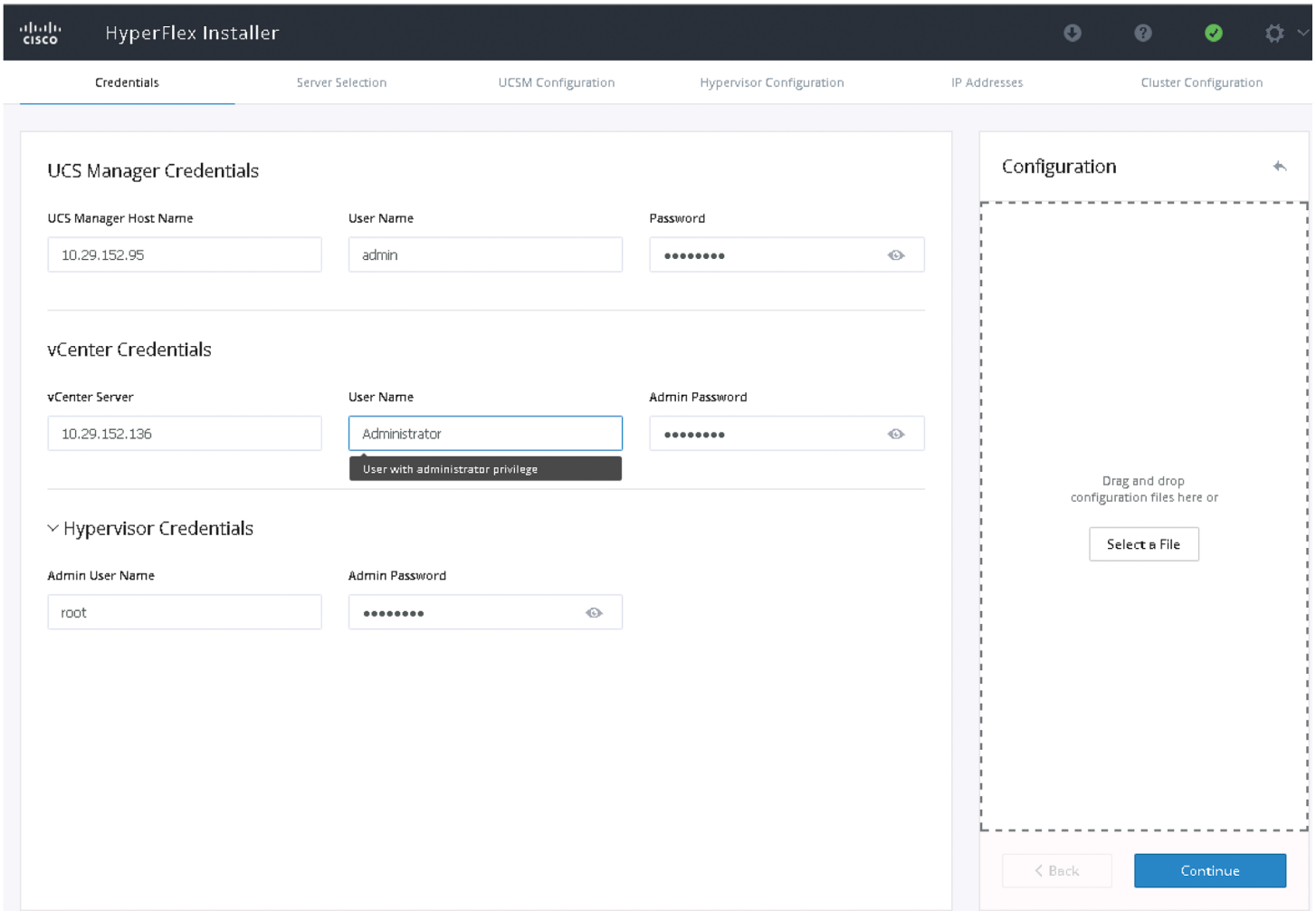


Figure 10 : HyperFlex Installer wizard showing customized host and cluster information being input

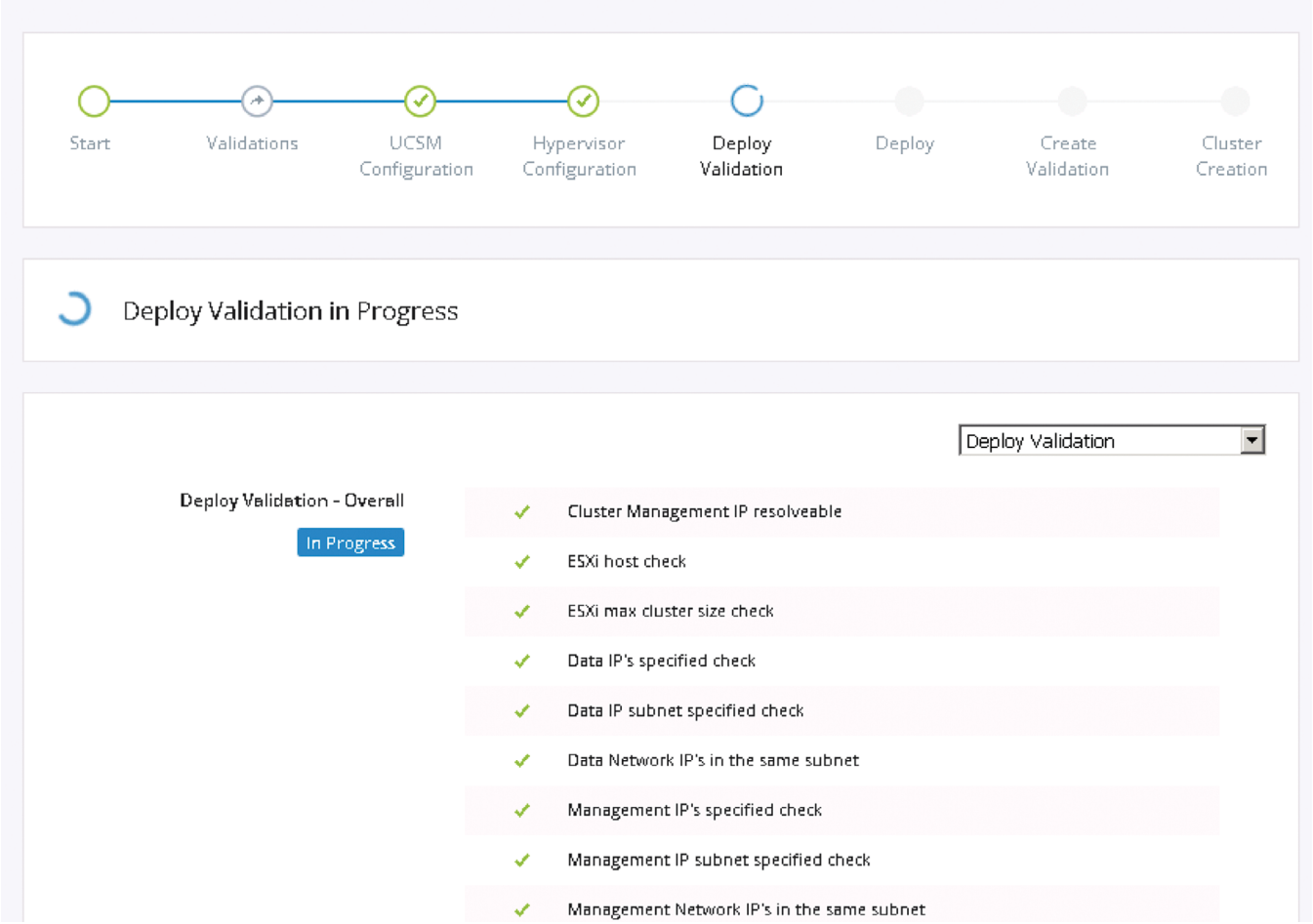


Figure 11 : HyperFlex Installer view of automated process, configuring UCS Manager, physical networks, virtual networks, verifying dependencies and creating the cluster.

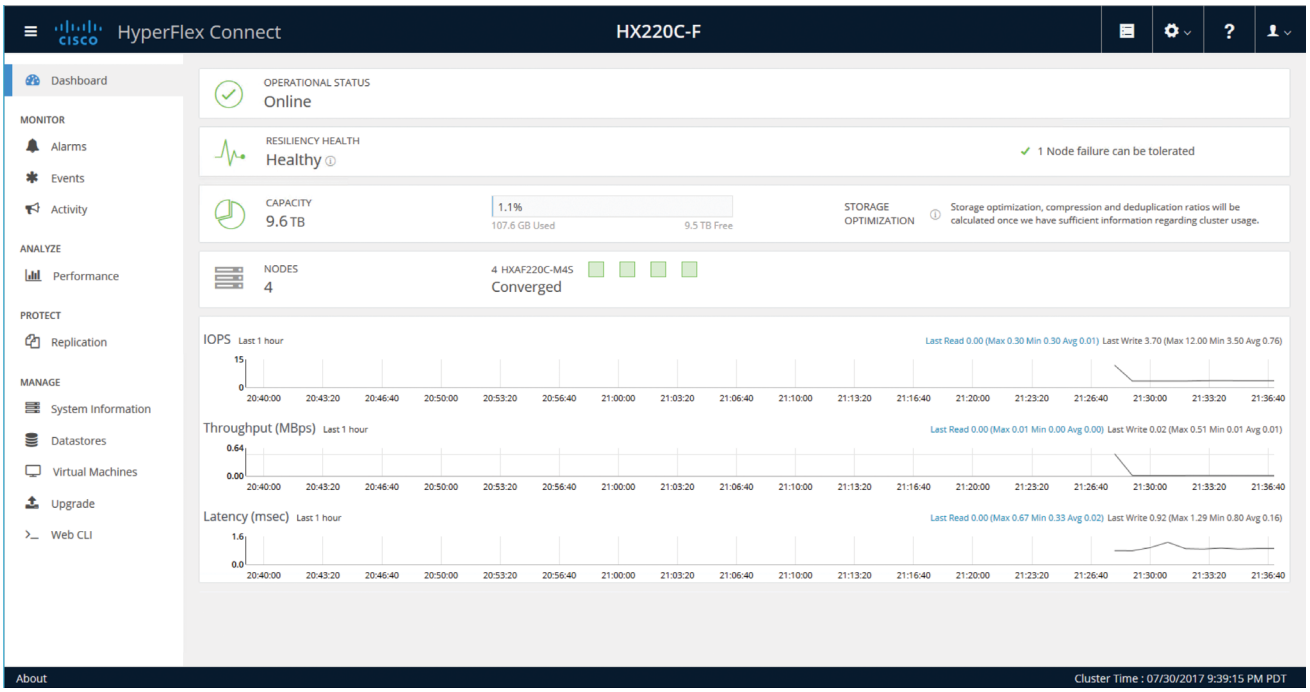


Figure 12 : HyperFlex Connect UI displaying a new configured cluster

Cohesity Setup and Cluster Installation

The Cohesity software running on Cisco UCS C240 M4 servers offers an integrated solution for backup and tier 2 storage needs. The installation starts with mounting the installer ISO on the (4) C240 M4 Cisco rack servers and letting the OS installation complete on each node. Some basic OS networking configuration may be required if static IP addresses are used. After the network is configured, cluster configuration resumes by connecting to the mgmt IP address of one of the nodes.

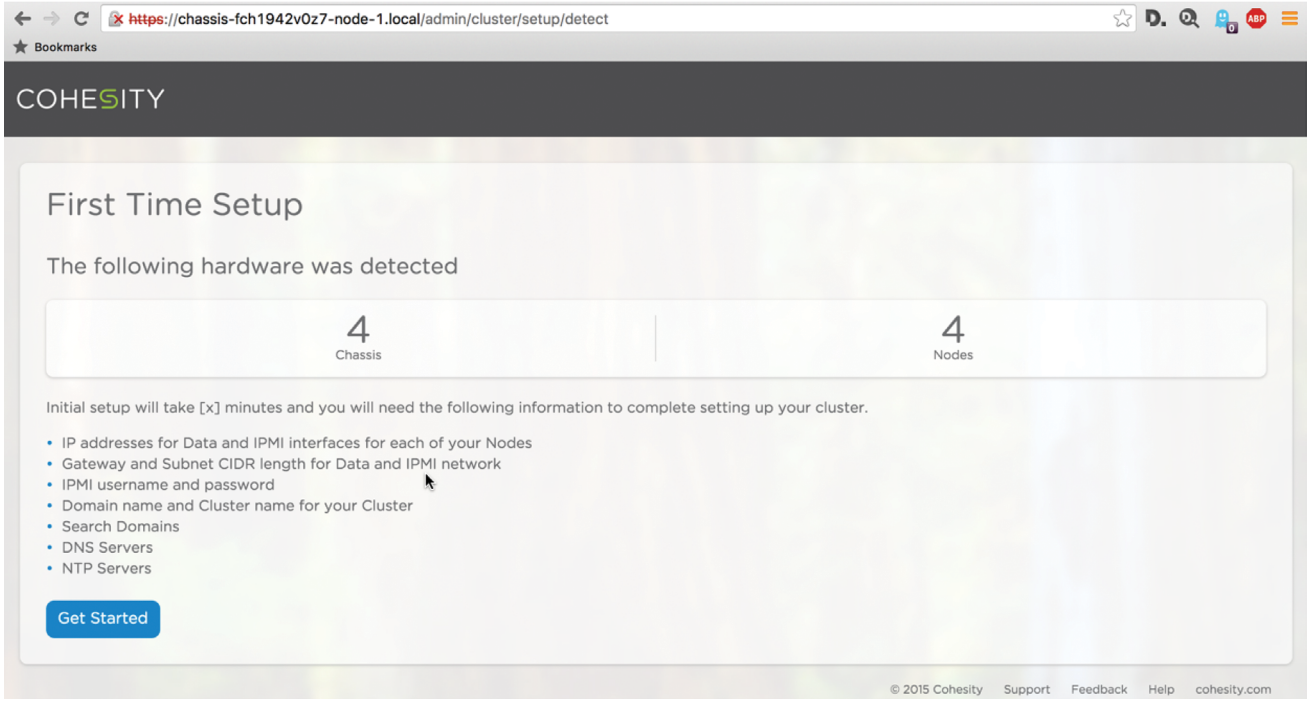


Figure 13 : Cohesity installation wizard and discovery of new unconfigured nodes

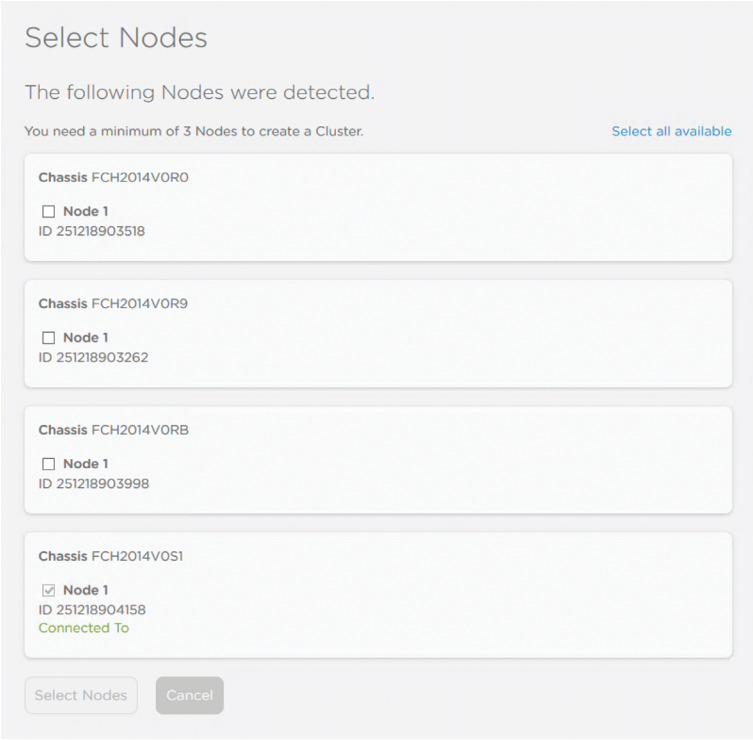


Figure 14 : Cohesity installation wizard and discovery of new unconfigured nodes

Explore the Cluster interface

Using the built in user interface all cluster storage operations are performed including: cluster configuration, management and backup tasks, and setting up file services. The user interface is built into the cluster software and accessed via a web browser and provides cluster setup and defining all backup jobs. Multiple operations such as creating different logical containers (View Boxes), creating data protection jobs, performing backups to Cisco HyperFlex Cluster, creating filesystem mount points and presenting these to external ESXi hosts, and defining recovery tasks.

Cohesity Dashboard

Cohesity's policy-based storage management approach is immediately viewable. Upon supplying credentials to the uni ed management web console hosted on the DataPlatform, users are presented with an overall health dashboard of the cluster:

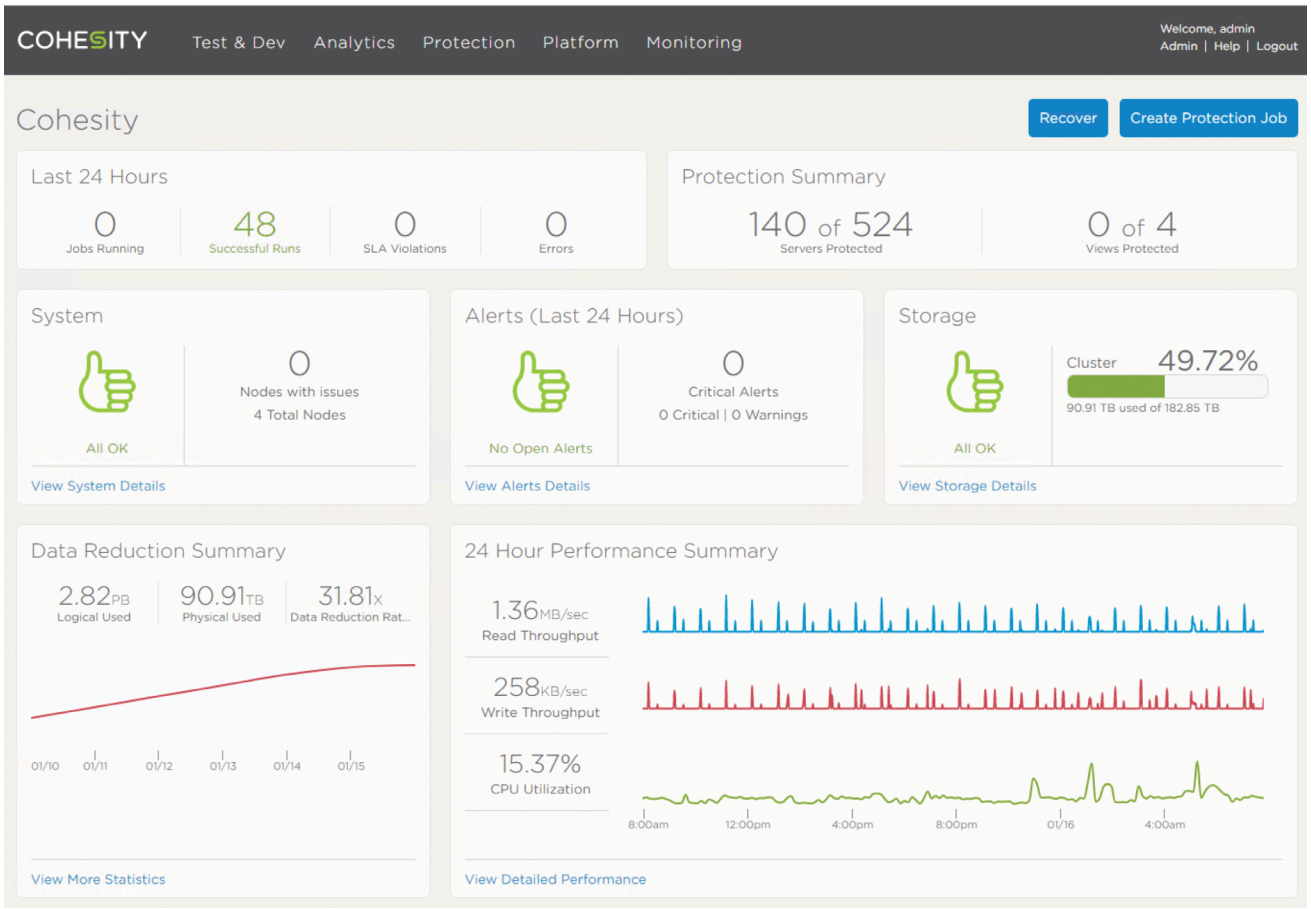


Figure 15: Dashboard for the Cohesity Data Management Platform

This dashboard reflects the overall health and state of the cluster including the number of jobs that have run, any SLA violations, errors or alerts as well as a brief data reduction and performance summary. Each of these items can be clicked on to review further detail.

Storage, backup and virtualization administrators may utilize the dashboard for cluster status reviews, however the platform also provides built-in alerting mechanisms as well as SNMP support for noti cation of certain conditions.

Finally, all cluster management tasks can be driven via the native REST API. Full monitoring and administration capabilities are available via the REST API and the documentation contained on the cluster.

Cohesity Storage Protection Policies

This test scenario outlines how to configure Cohesity to protect VMs running on a HyperFlex cluster of 4 nodes under a shared datastore. Cohesity can be configured with multiple types of protection jobs with different schedule time frames, full and incremental backup policies and various retention policies.

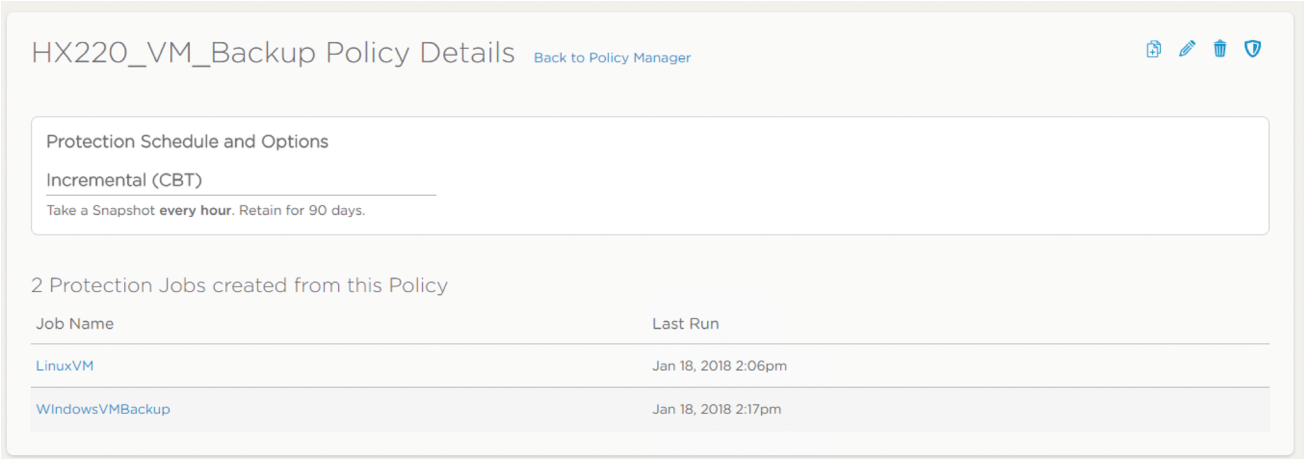


Figure 16 : VM Protection Policy

Protection policies are very flexible and provide a number of extended rules and monitoring options.
In the example above, the standard rules call for snapshots every 1 hour with 90 day retention.

The following are some other policy configuration parameters that can be optionally edited :

- Schedule End Date: a policy can be retired on a given date
- Custom Retry Settings: set job retry intervals and number of attempts
- Blackout Periods: set periodic pauses in associated job runs that use the policy
- Alerts and SLAs: ensure IT administrators are aware of job success, failures or SLA violations for SLA time limits.
- Priority: set a relative priority for jobs that use the policy

Policies are usually created to ensure RPO compliance and are made available to protection jobs across the entire cluster. Define policies based around data protection requirements that can be shared across protection jobs. Implement a naming scheme that allows easy identification of policy attributes.

Cohesity Protection Jobs

Protection jobs define one or more groupings of VMs for protection that comply with a specific Protection Policy (see above). Jobs also set the time and timezone of the job as well as the desired viewbox that will contain the protected data.

Jobs are primarily used to assign policies to groups of VMs that have similar attributes. Most commonly, jobs are used to group VMs that have a common SLA and assigned a common protection policy. Jobs can also be used to capture VMs that serve the same application as a group. In other instances, jobs are used to capture VMs from logical groupings such as organizational user groups, departments or geographical locations.

Jobs protect VMs in one vCenter, thus VMs that span multiple vCenters will require multiple jobs. When doing so, it is best to use a naming convention that allows quick identification of related jobs.

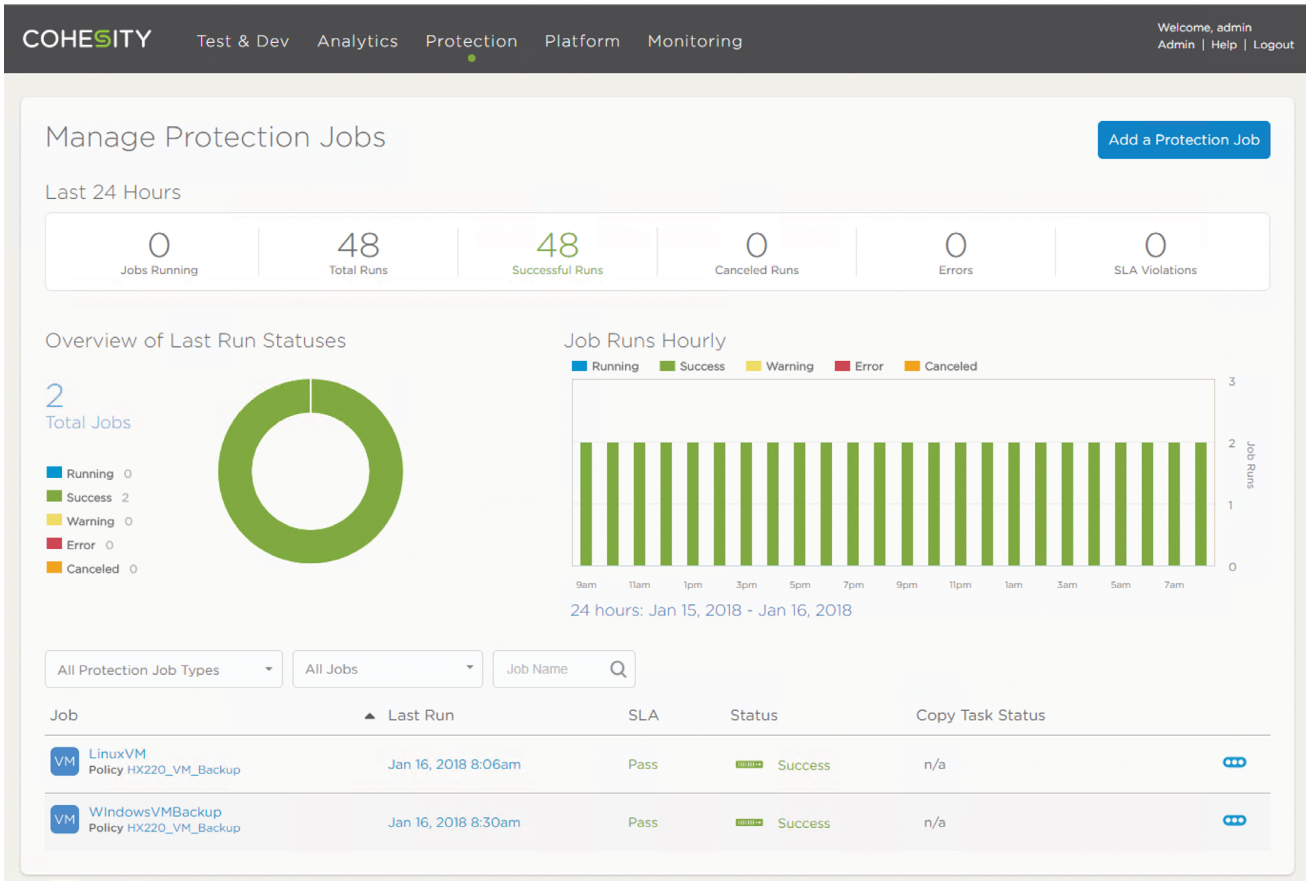


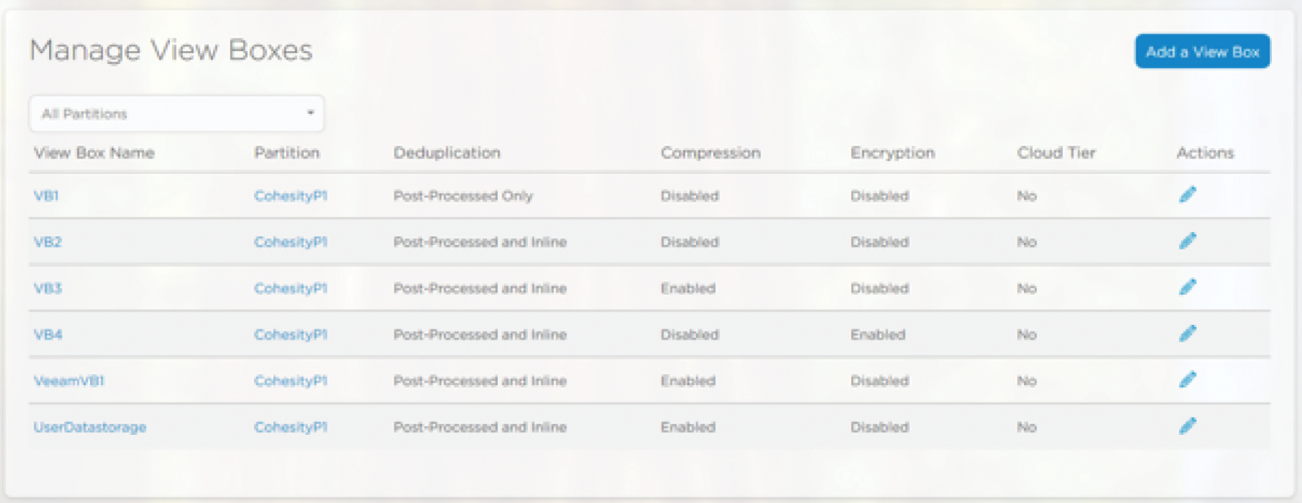
Figure 17 : Dashboard displaying all the protection jobs configured in the system

Cohesity View Box

View boxes represent storage efficiency domains within the Cohesity cluster and can optionally associate a view box with a specific cloud tier. When view boxes are created, the administrator assigns deduplication, compression and encryption attributes to the view box. View boxes can be set for inline, post-process dedupe or no dedupe as well as inline post-process or no compression. These settings can be modified at creation and over the life of the view box.

Encryption is enabled at creation and is only chosen at creation. Similarly, the Cloud Tier option is only available at creation time.

For testing to create this reference architecture multiple view boxes were configured with various storage policies for dedup, compression, and encryption. All of these tasks are performed from the Cohesity GUI and there is no interaction with the HyperFlex primary datastore.



View Box Name	Partition	Deduplication	Compression	Encryption	Cloud Tier	Actions
VB1	CohesityP1	Post-Processed Only	Disabled	Disabled	No	
VB2	CohesityP1	Post-Processed and Inline	Disabled	Disabled	No	
VB3	CohesityP1	Post-Processed and Inline	Enabled	Disabled	No	
VB4	CohesityP1	Post-Processed and Inline	Disabled	Enabled	No	
VeeamVB1	CohesityP1	Post-Processed and Inline	Enabled	Disabled	No	
UserDatastorage	CohesityP1	Post-Processed and Inline	Enabled	Disabled	No	

Figure 18: View Boxes can be configured with different protection policies

Restore VM from backup to original location

A significant feature of the Cohesity platform is instant restore of a virtually unlimited number of VMs and applications, to any recovery point. Because the Cohesity DataPlatform can present an NFS mount point for any file, every backup image is immediately and directly accessible to all ESX hosts. Instead of instantly restoring a few VMs like most traditional backup solutions, administrators can instantly restore an unlimited number of VMs, or even an entire vCenter. The PCIe flash in each Cohesity node ensures high performance for restored VMs. Instant restore is automated by the Cohesity console. Administrators can mount a specific image stored on Cohesity as an NFS data store, redirect the VMware ESX host to that image, and then initiate the VM boot process and application recovery. The end-user returns to full operations in the time it takes to recover the VM, without the time-consuming rehydration and data transfer that most data protection schemas require. While the restore is happening Cohesity leverages vMotion to transparently migrate the VMs back to the primary data store all without any manual intervention from the administrator.

This use case will outline the steps needed for the recovery of VMs that were deleted from the Cisco HyperFlex datastore and recovery to place the VMs back into the original storage location.

Cohesity View Box

View boxes represent storage efficiency domains within the Cohesity cluster and can optionally associate a view box with a specific cloud tier. When view boxes are created, the administrator assigns deduplication, compression and encryption attributes to the view box. View boxes can be set for inline, post-process dedupe or no dedupe as well as inline post-process or no compression. These settings can be modified at creation and over the life of the view box.

Encryption is enabled at creation and is only chosen at creation. Similarly, the Cloud Tier option is only available at creation time.

For testing to create this reference architecture multiple view boxes were configured with various storage policies for dedup, compression, and encryption. All of these tasks are performed from the Cohesity GUI and there is no interaction with the HyperFlex primary datastore.

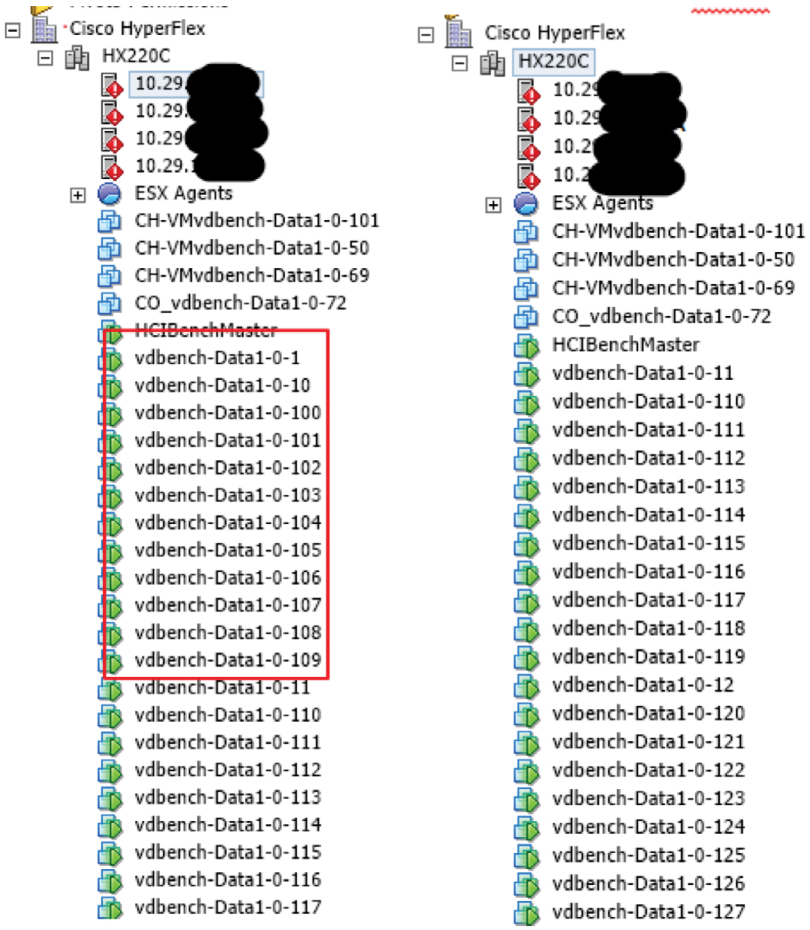


Figure 19 : View of VMs before and after deletion

- 11 different VMs were deleted from the Cisco HyperFlex datastore and removed from vCenter.
- Using the Cohesity GUI a recovery task was started and the 11 missing VMs were selected from one of the backup View Boxes where the backups are stored.
- The deleted VMs were marked to be restored to their original location with the original VM names.
- Cohesity Instant Mass Restore mounts an NFS datastore to the ESXi hosts, and each VM was booted and powered up. All of the VM data was available and VMs were ready to boot within a few seconds, and began booting on vSphere.
- In the background all the VMs are now servicing active applications and are automatically migrated to the original HyperFlex datastore using vMotion.
- Each of the VMs are running at the recovery point with no missing data or issues.

Step 1:

Perform a regex search for the VM names and select the 11 VMs that were deleted from the vCenter.

Step 2: Define the recover options for VM recovery

As the original VMs were deleted on the vCenter the default recovery options are used, which ensure that the VMs are restored back to the HyperFlex Platform with all original settings.

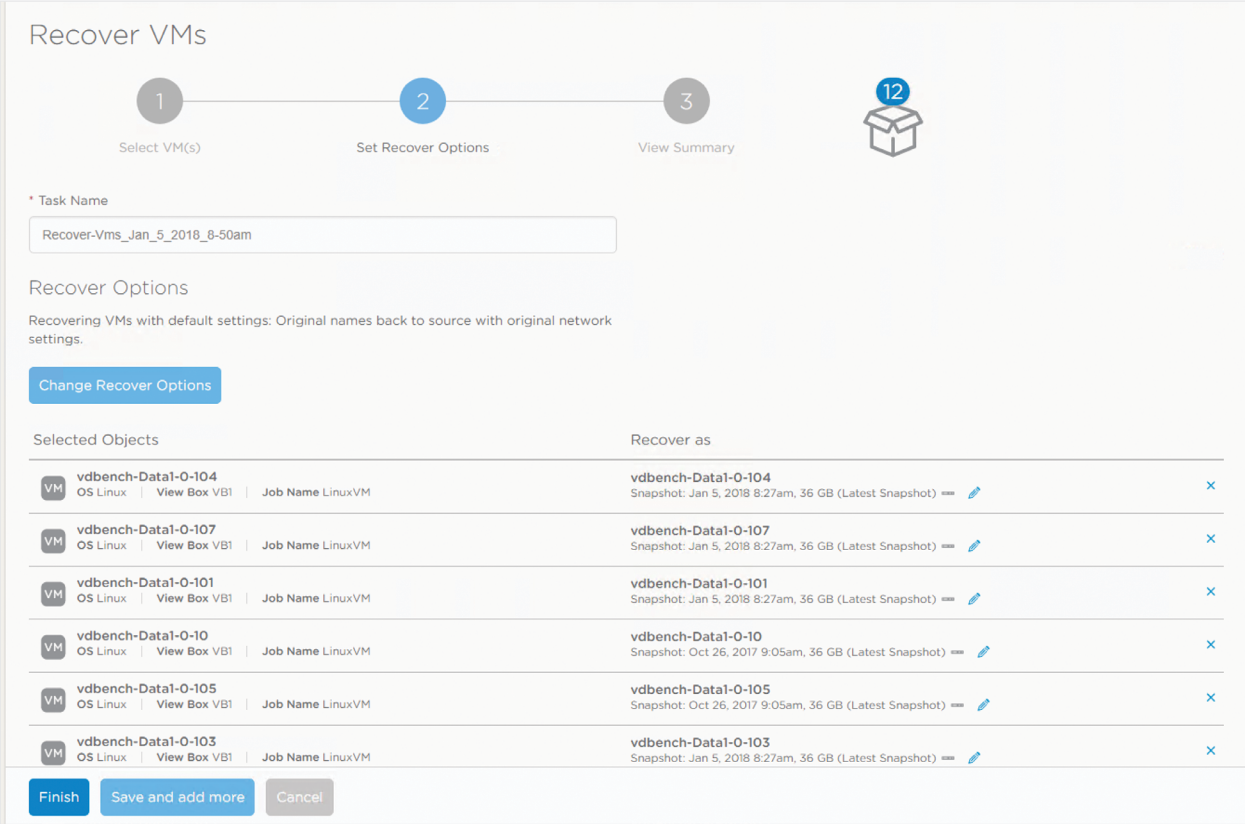


Figure 20 : Restoring VMs to the original location

Step 3: Once the RPO has been selected for the VMs hit Finish to Instantly recover all 11 VMs

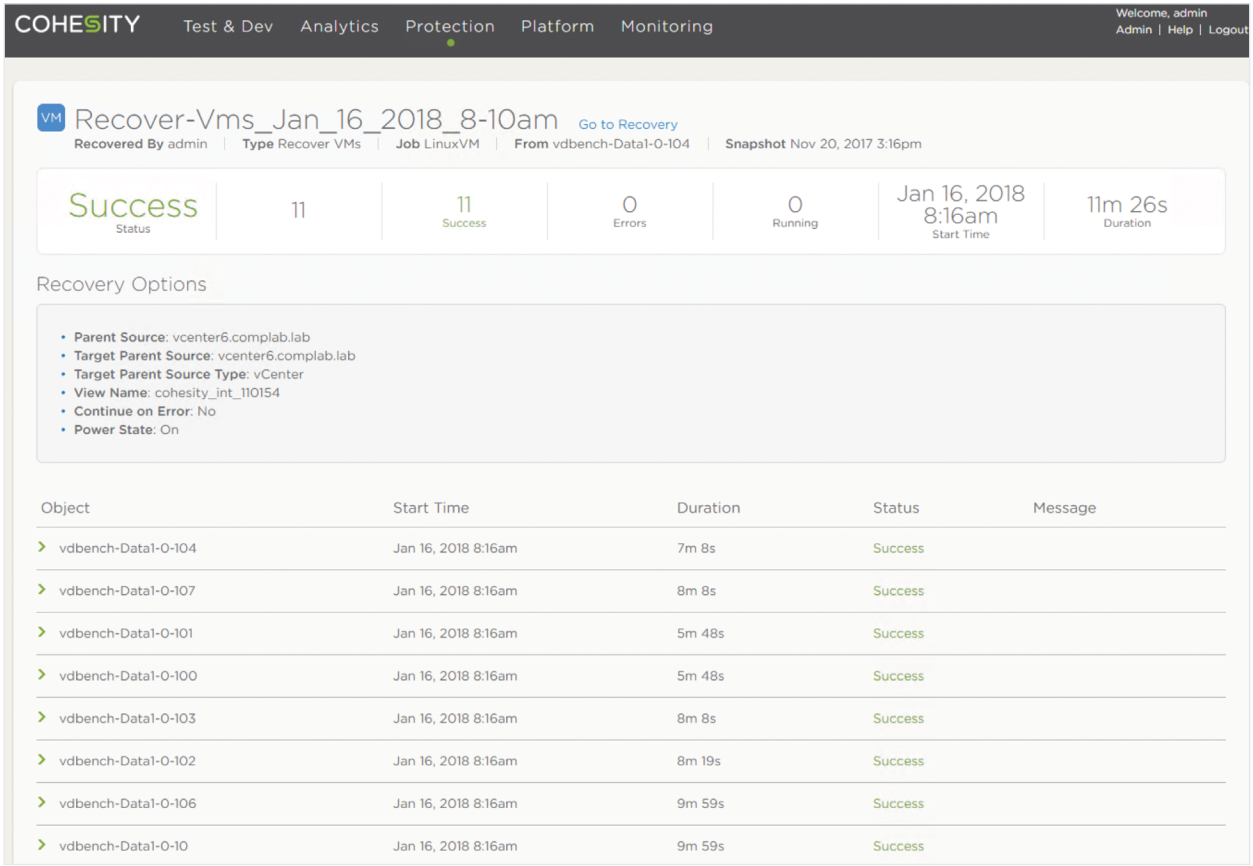


Figure 21 : Successful recovery of VMs backed up on Cohesity and restored on Cisco HyperFlex

Anatomy of a VM Restore Job

Step 1: User manually triggers a Cohesity VM recovery task and selects snapshot, target, networking settings, VM name, target datastore…

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 access)

Step 6: Storage vMotion is initiated to move the datastore from the Cohesity cluster to the primary data store.

Step 7: Storage vMotion completes, VMware non-disruptively migrates datastore access from the Cohesity cluster snapshot to the primary data store.

Step 8: Cohesity requests the datastore to unmount

Step 9: ESXi host unmounts datastore

Step 10: Cohesity releases the view

Restore individual VM files from backup

Cohesity can restore an individual file and place it back to the original location or alternate location on a VM. This test was setup to simulate an accidental deletion of multiple files from a Windows VM on a Cisco HyperFlex cluster. The deleted files need to be restored to the original location on the VM OS. Using the Cohesity GUI the protection job and recovery task are selected and then a File or Folder will need to be recovered. There are two ways to find the deleted files, either by file or folder name or by selecting the VM and manually searching for the missing files. The latter method was used to point to the VM and select the missing file and folder with multiple files. The restore operation completes within a few seconds the files were fully recovered onto the VM in the original directory.

Step1: Select the VM to browse to the file location

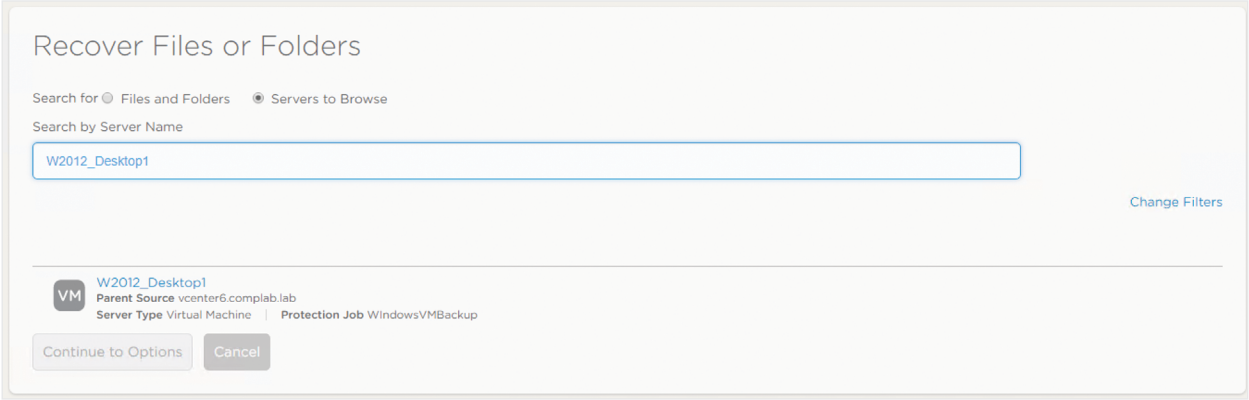


Figure 22 : File recovery using the browse server option

Step 2: Locate the files and folders for restore

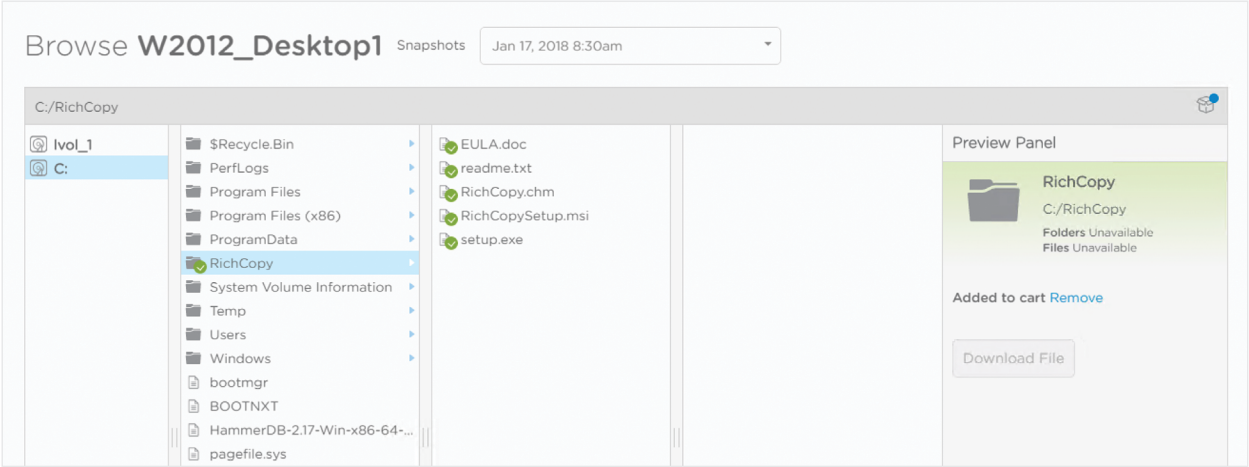


Figure 23 : Navigate the disks of the VM to select files or folders to be recovered

Step 3: Customize the recovery and input os credentials

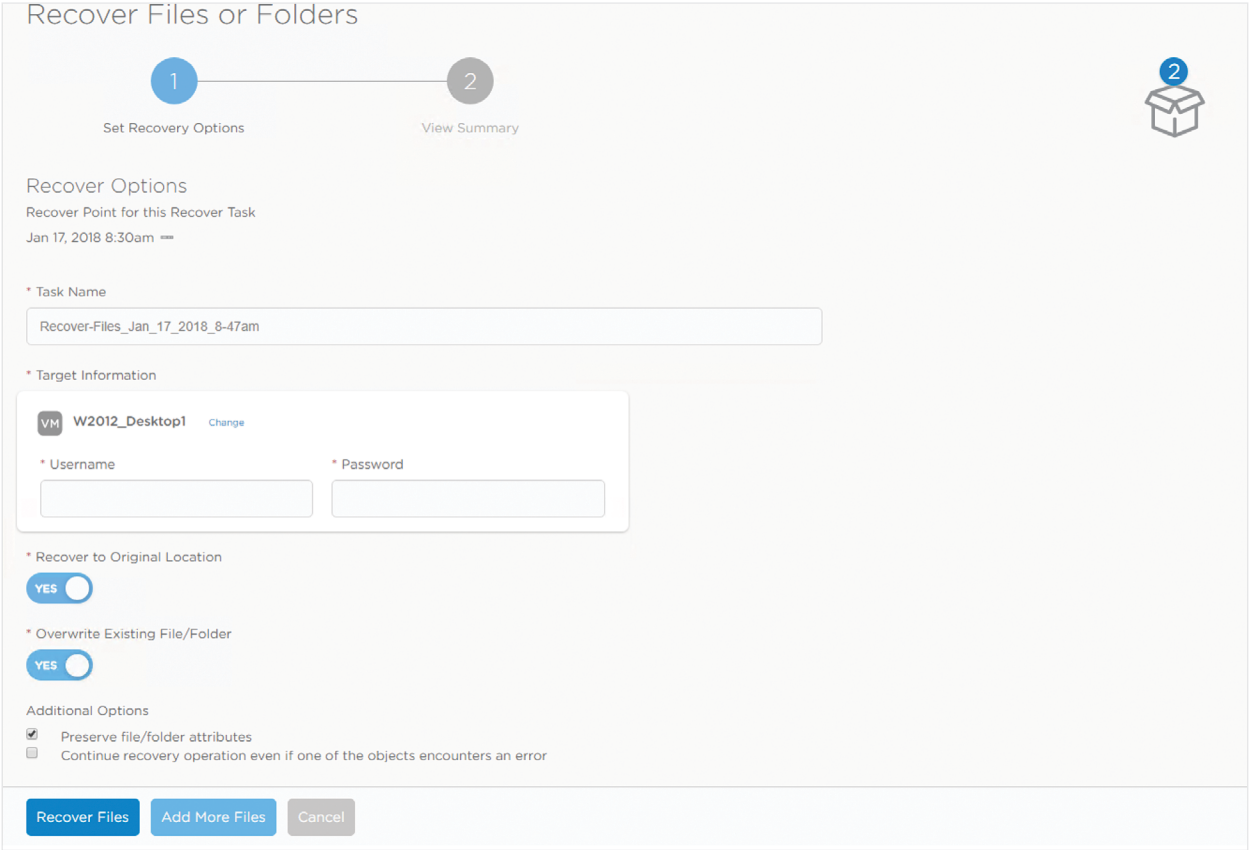


Figure 24 : Provide the credentials and select to recover files to original local and overwrite the existing files while preserving the original file/folder attributes

Step 4: Files and folder are restored to the VM original location

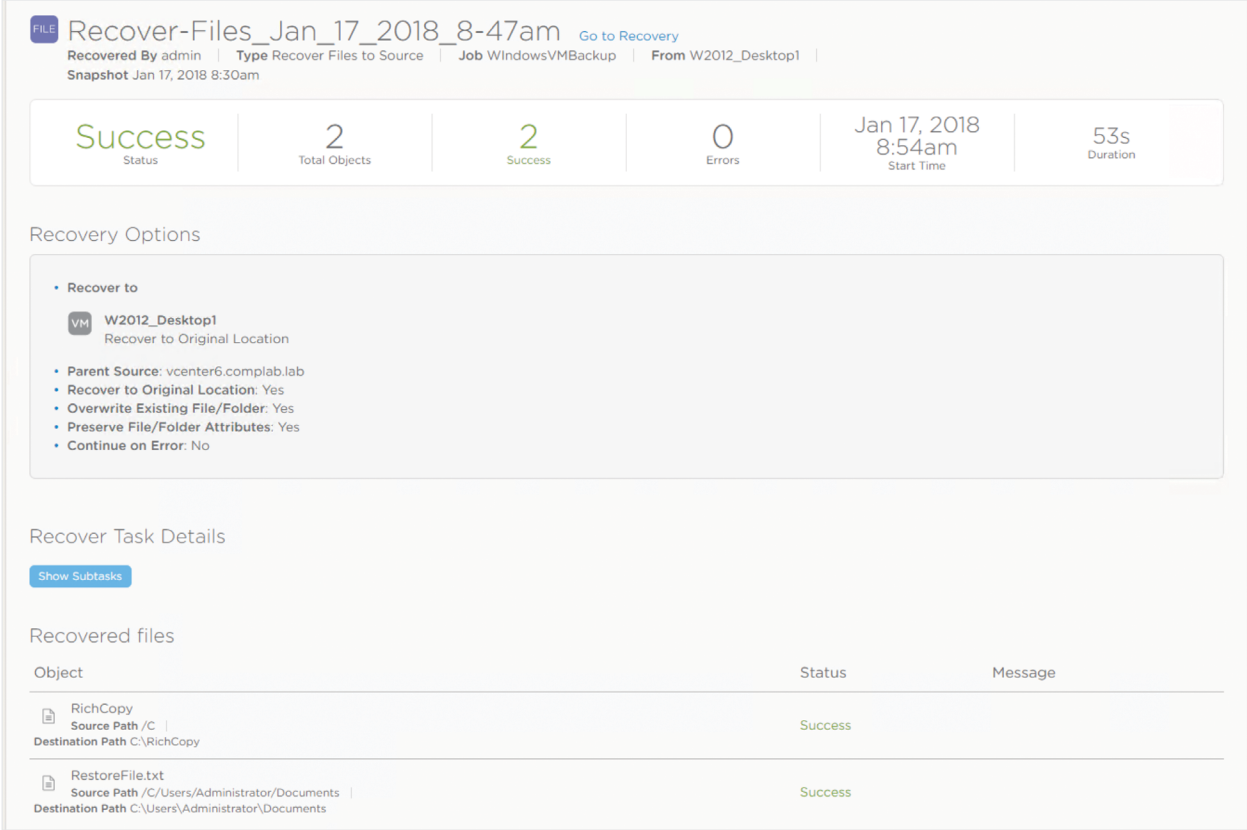


Figure 25 : Files and successfully restored back to the VM with original file/folder attributes

Anatomy of a File Restore Job

Step 1: User manually triggers a le/folder recovery task either by searching the les through the elasticsearch database or via browsing VMs and it's volumes.

Step 2: Cohesity creates an internal view and clones the VM snapshot and mounts the view to the target ESXi host(s)

Step 3: Cohesity attaches the cloned VMDK les to the target VM to which the les are being recovered.

Step 4: Cohesity deploys a helper utility onto the VM and triggers the restore process

Step 5: The restore helper utility performs le copy from the attached disks (originally from the backup) onto the recovery location. The utility additionally preserves the le attributes and other properties based on the user preferences.

Step 6: Once the le/folder copy completes, the disks are detached from the VM Step 7: Cohesity requests the datastore to unmount

Step 8: ESXi host unmounts the datastore

Step 9: Cohesity releases the view

Presentation of storage for Test/Dev

Cohesity can be used as a test/dev storage cluster so that production VMs may be cloned from Cisco HyperFlex primary storage. The cloned VMs on Cohesity are completely independent from the production datastore and any testing may be performed without impact to the production applications. The Cohesity cluster is configured to present NFS datastore to the Cisco HyperFlex cluster in a simple and flexible way. In the GUI multiple NFS and SMB shares were created and presented to ESXi host on the HyperFlex cluster as well as a Windows virtual machine running on another ESXi host with direct mapping. This process was simple and normal file system mounting procedures for each ESXi host was followed with no issues.

In the following screen HA_DRS is a Cohesity View that was created in the Cohesity UI, and it can now be exposed with NFS/SMB 3.0 and S3 protocols directly to client systems.

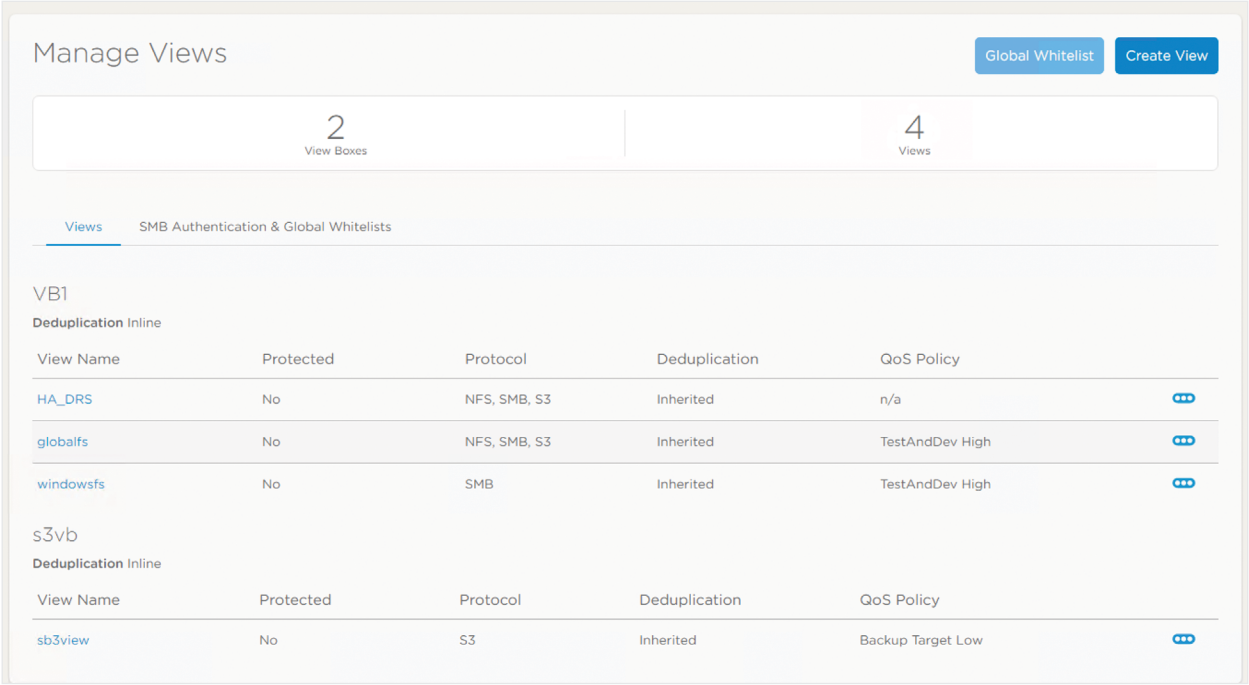
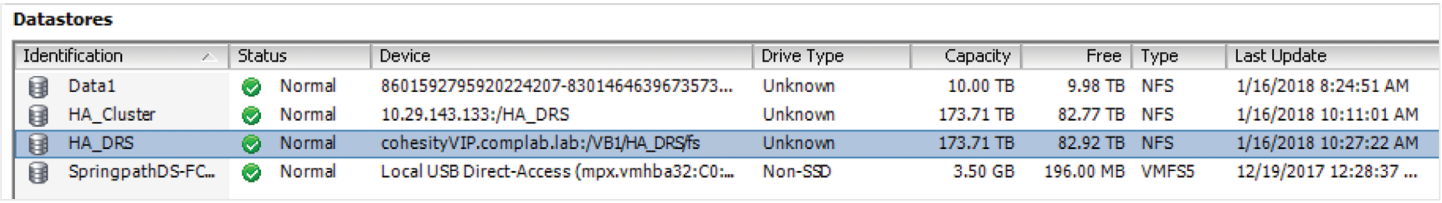


Figure 26 : Presenting a view to external host through NFS

The next step is to mount the newly created Cohesity View as an NFS Datastore on each HyperFlex ESSi host, to complement the HyperFlex primary storage system for Tier 2 scale out storage to run 2nd tier VMs.



Datastores

Identification	Status	Device	Drive Type	Capacity	Free	Type	Last Update
Data1	Normal	860159279592022420 7-8301464639673573...	Unknown	10.00 TB	9.98 TB	NFS	1/16/2018 8:24:51 AM
HA_Cluster	Normal	10.29.143.133:/HA_DRS	Unknown	173.71 TB	82.77 TB	NFS	1/16/2018 10:11:01 AM
HA_DRS	Normal	cohesityVIP.complab.lab:/VB1/HA_DRS/fs	Unknown	173.71 TB	82.92 TB	NFS	1/16/2018 10:27:22 AM
SpringpathDS-FC...	Normal	Local USB Direct-Access (mpx.vmhba32:C0:...	Non-SSD	3.50 GB	196.00 MB	VMFS5	12/19/2017 12:28:37 ...

Figure 27 : Datastore details with new Cohesity NFS storage presented to the ESXi host

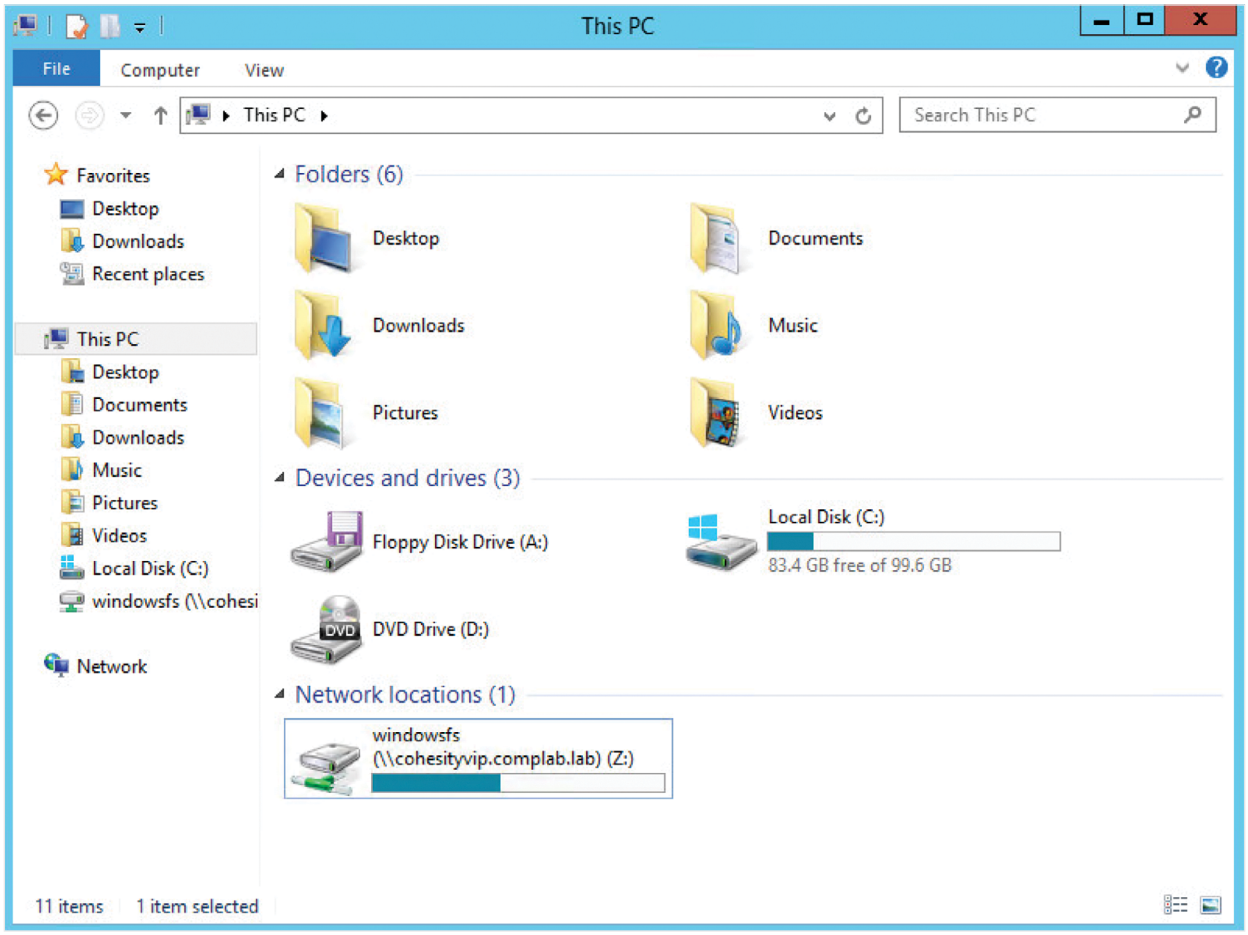


Figure 28 : Cohesity cluster presented as a remote network drive on Windows

Clone Protected VMs for Test/Dev

Cohesity empowers developers to instantiate the latest backup of their production application stack and run it directly off the Cohesity platform, providing a unified foundation for copy data management.

Cohesity can be used to create clones of running production VMs and present them back to the Cisco HyperFlex ESXi cluster for testing and development. This enables the use of real live production writable clone VMs for testing with complete separation of the parent VMs and completely independent of the HyperFlex production datastore.

An existing backup job was used to select multiple VMs from the Cohesity storage platform to instantly create zero-copy clones for test & dev use. The clones are created instantly because Cohesity keeps all backed up data sets fully hydrated at all times. These VMs were restored to the HyperFlex cluster into a datastore that can be automatically created and mounted to each of the HyperFlex nodes.

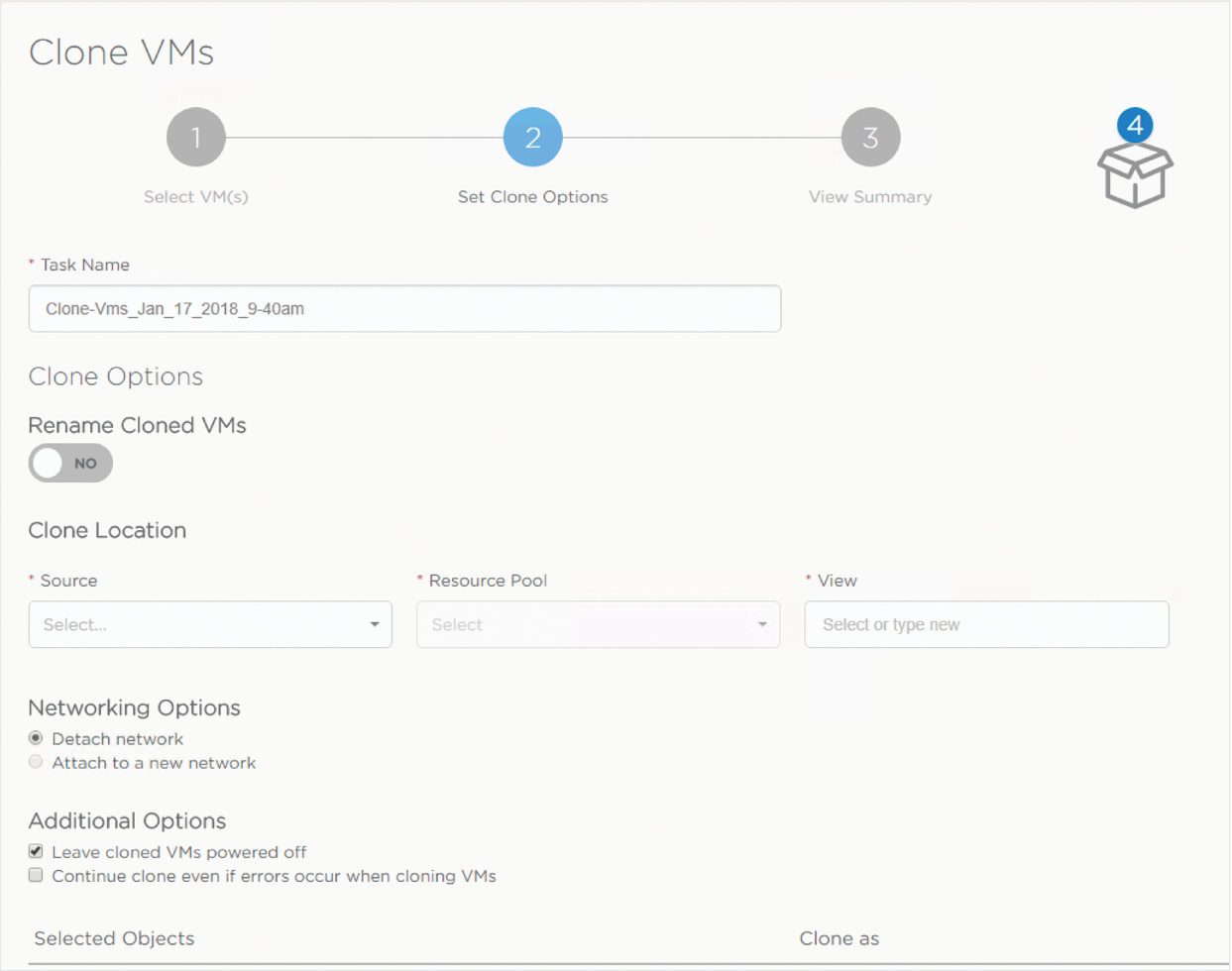


Figure 29 : Cohesity cluster presented as a remote network drive on Windows

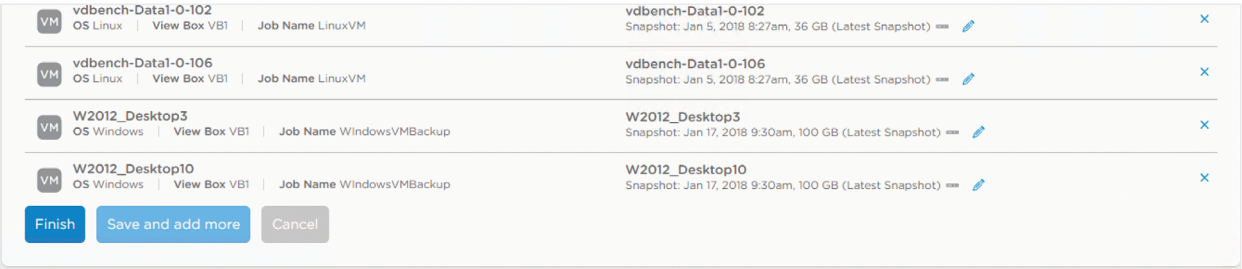


Figure 30: Select and customize the clone VMs to be used for Test/Dev

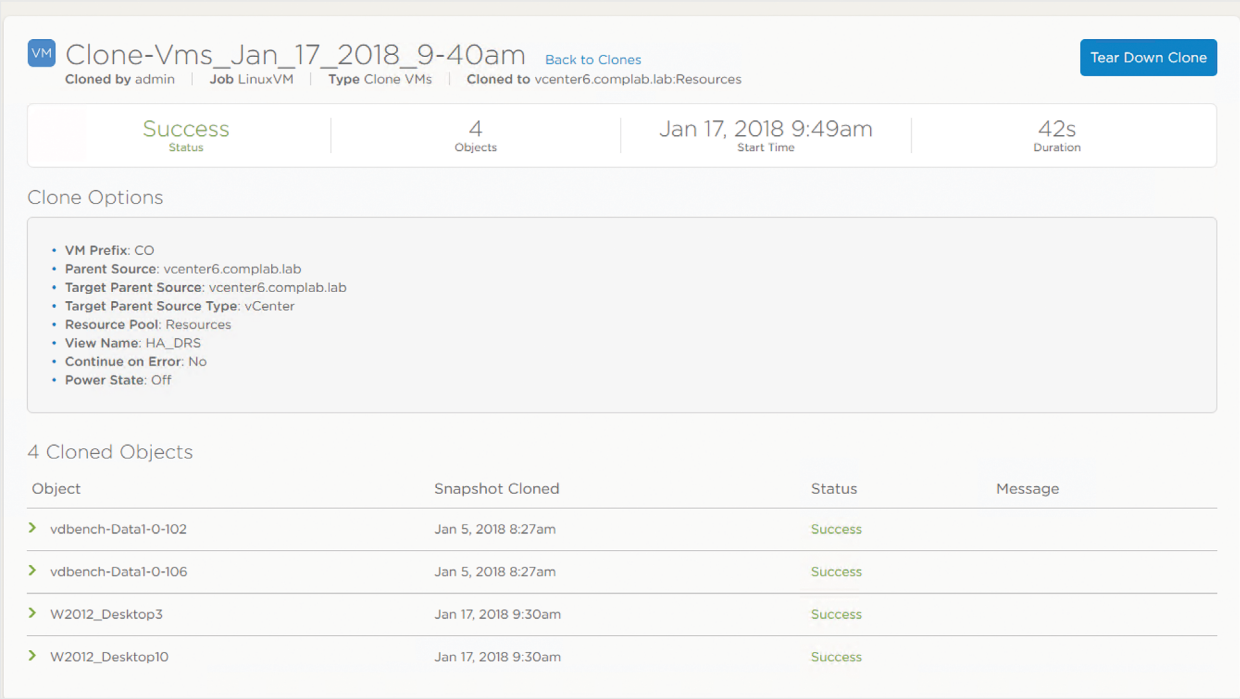


Figure 31 : Cloning summary

Figure 32 : Cloned VMs presented to HyperFlex

Cluster scalability and growth

Validation of scaling the Cohesity cluster by adding a fourth node to an existing 3 node cluster while running a protection job was done. The performance of the cluster backup process increased linearly with the new node dynamically starts taking load from a currently running protection job. Adding the new node was a simple process utilizing the web UI.

Figure 33 : Status of the three nodes in a cluster

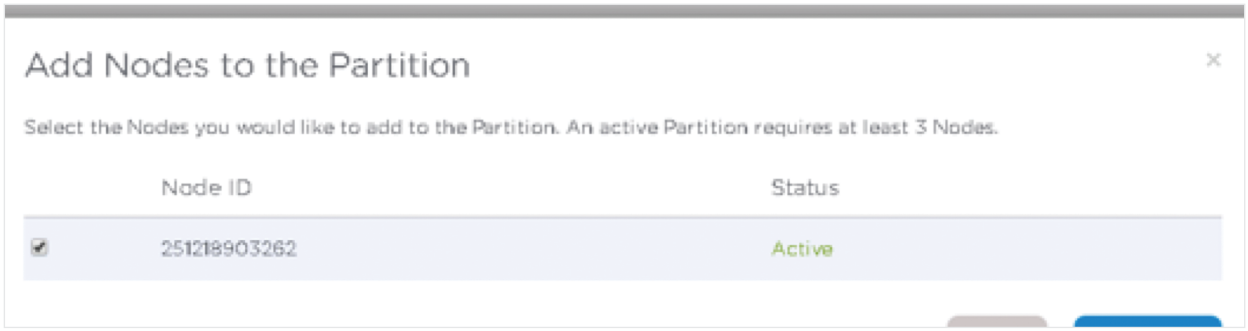


Figure 34 : Adding a fourth node to the cluster

After new nodes are added, the disk capacity of the Cohesity cluster is immediately increased and available for use by Protection Jobs, Clone task and Recovery tasks. This addition of more nodes improves performance because additional CPUs, memory, and disk bandwidth are available to the cluster. Also, the cluster distributes the blocks from existing nodes to the new nodes. In addition, the cluster replicates the metadata about the block distribution onto these new Nodes.

Cluster backup performance

Multiple backup jobs were executed with different types of storage policies to observe the backup performance. The expectation is that with 10Gb networking the backup jobs would complete with no issues and in a timely manner. There were 2 different view boxes created with different policies configured for compression, variable length dedup (post process and inline), and encryption. Multiple protection jobs were created to backup 144 VMs on the HyperFlex Cluster and the same VMs were used throughout the comparative testing. Each protection job was run individually so that there were no overlapping jobs competing for resources. The testing incorporated full state backups and incrementals across the 4 different storage policies. All test results completed with no errors on the backup jobs and backups were completed within expected time frames.

Figure 35: Backup job performance and status

Conclusion

This reference architecture demonstrated how we have leveraged the key strengths of the Cisco UCS architecture with Cisco HyperFlex and Cohesity DataPlatform to consolidate secondary storage workflows. The architecture offers the industry's top compute infrastructure running Cisco HyperFlex and Cisco UCS networking fabric running Cohesity's scale-out secondary storage platform with its integrated backup and recovery software. All of the secondary storage workflows functionality including backup & recovery, test/dev and SMB/NFS/S3 file shares is delivered through a single pane of glass. Cohesity can backup VMs and provide a storage repository for Cisco HyperFlex converged infrastructure solutions as well as bare metal systems. By exposing itself as an NFS or SMB file share, Cohesity can store datasets being backed up by 3rd party backup software. Cohesity can also spin up the backed up VMs instantly to create test and development environments demonstrating how it significantly reduces the storage footprint while making the development environments considerably more agile. This provides enterprises the significant operational agility, the key motivation for deploying HCI infrastructures. Using Cohesity on Cisco HyperFlex greatly simplifies the protection of VMs and applications in the datacenter and offers expanded storage for testing and development.

Authors:

Chris Dunk - Technical Marketing Engineer UCS

Damien Philip - Principal Solutions Architect - Cohesity

About Cohesity

Cohesity delivers the industry's first solution for secondary storage

consolidation. Cohesity enables companies of all sizes to bring order to their data chaos by converging storage workloads, including file services, data protection, Test/Dev, and in-place analytics, onto an infinitely scalable, intelligent data platform.

With Cohesity, customers can manage and protect data seamlessly, use it efficiently, and learn from it instantly. Cohesity is headquartered in the heart of Silicon Valley, California with a global presence across the Americas, EMEA, and APAC.

152018

Cohesity, Inc.
Address 300 Park Ave., Suite 300, San Jose, CA 95110
Email contact@cohesity.com www.cohesity.com
@cohesity