Veeam with Cohesity
Data Platform
About This Guide

This paper details the steps and best practices to deploy a Cohesity Data Platform as a storage target for the Veeam backup application. Combining Veeam with Cohesity Data platform provides a comprehensive, highly scalable and flexible backup solution that fits the data protection needs of any size organization whether SMB or enterprise.

Audience

This paper is written for storage and backup administrators familiar with administering and managing backup environments. You must also be familiar with:

- Veeam Backup & Replication
- Cohesity Data Platform
- Server Message Block (SMB)

Data Protection for VMware Environments

There have been many technological advances in the secondary landscape of the modern data center. With each evolutionary change, new efficiencies bring optimizations to the data center administrator which have made the tasks of storing and finding data easier and easier. The migration from tape based backups to disk based backups bring the efficiencies of speed to not only the backup process, but also to the recovery process. Next, the move to deduplication appliances bring a dramatic on-disk space savings, and coupled with virtualization, a massive reduction in overall rack space used in the data center. While each of these evolutions brought speed and efficiency to the data center, they all continued to ignore the problem of data silos. Without the ability to deduplicate between these appliances, data growth continues to increase at unprecedented rates.

Cohesity Data Platform Overview

Cohesity introduced the world’s first scale-out data management platform to enable organizations to standardize secondary workflows on a completely unified and fully distributed solution. Cohesity’s scale-out distributed file system (OASIS) was built from the ground up to ensure complete scalability to enable organizations to flexibly grow their environment by simply adding nodes to a cluster. With this scalability, organizations can eliminate the costs of data migrations and forklift upgrades, while benefiting from the simplicity of a homogenous solution.

In addition to ensuring complete scalability, Cohesity enables the delivery of all secondary workloads by providing a powerful, unified platform with industry-leading capacity and performance across the most common file system protocols. In a single 2U Block, Cohesity delivers 48TB of hybrid flash and HDD storage to efficiently balance performance, capacity, and price to give organizations the flexibility to offload secondary workloads from expensive primary storage arrays. Cohesity’s integrated backup, archive, replication, disaster recovery, and public/private cloud support, combined with its inherent context awareness, eliminate the need for cataloging software, backup software, and other ancillary backup infrastructure.

With Cohesity, organizations can benefit from stringent data protection policies by gaining valuable insight from their data. Cohesity’s context-aware file system makes it easy to understand storage utilization trends, reclaim unused resources, and create summary and detailed reports with just the click of a button.

Veeam Overview

Veeam Backup & Replication provides a set of features for building and maintaining a flexible backup infrastructure, performing data protection tasks (such as VM backup, replication, copying backup files), and carrying out disaster recovery procedures.
Benefits of using the Cohesity Data Platform with Veeam Backup & Replication

Cohesity data platform, when used as a storage target for Veeam, delivers unparalleled performance, scalability and availability for the backup environment. The Cohesity data platform

- Maximizes storage capacity with Cohesity’s advanced data reduction functionality.
- Modern webscale distributed system with limitless scaling of performance and capacity.
- Continuous availability architecture with a minimum replication factor of 2 for stored data. Any node can fail and the system continues to function.
- Deploying a global storage target is as easy as just a few button clicks.

The Cohesity Veeam Solution

Architecture

There are several components that make up the Veeam Backup and Replication Architecture. For virtualized environments, a VMware Backup Proxy manages the data, backup, and restores for that environment. The Backup Proxy software is deployed on dedicated servers that manage the transfer of data from the clients to storage media. The management of the Veeam architecture is managed by a dedicated master server with the Veeam Backup and Replication management console software installed.

In the Veeam Architecture the Cohesity Data Platform cluster is considered a storage repository. The Cohesity Data platform presents its storage to Veeam as an exported SMB share. Cohesity Data Platform terms

- **View** – A view is a file system that is created on the data platform. A view can be accessed by external devices as an exported SMB share.
- **View Box** – A View Box is a storage container that contains a set of views. Storage policies on are set at this...
Storage policies consist of enabling data deduplication, either n-line data deduplication or post process deduplication.

- **Virtual IPs (VIPs)** - Each node on the Cohesity storage cluster is assigned a VIP address. The SMB shares that is exported by the Cohesity Cluster can be accessed through the VIPs. All VIP addresses are always active. In case of a node failure the VIP of the failed node will be taken over by another node within the cluster. This ensures that VIPs are always active. Load sharing on the cluster is accomplished using the VIPs by spreading the workload between the VIPs.

**Verified Solution**

- Veeam Backup & Replication version 9
- Cohesity Data Platform version 2.5.1+

**The Cohesity Advantage**

Data protection is complex. Between navigating cluttered user interfaces, choosing storage appliances, tape libraries, and off-site storage solutions, more time is spent planning than actually protecting. By combining the best in class virtual machine backup software from Veeam and the best in class data platform from Cohesity customers can implement the most comprehensive data protection solution available while also removing all the complexities of legacy secondary storage environments. Traditionally, data protection relied on a complex network of infrastructure which took days of configuration and constant monitoring. As technology changed, and storage appliances improved, fragmentation in the data center increased exponentially. Veeam and Cohesity now provides a webscale approach to future proof the datacenter.

**Advanced Data Platforms for Today's Modern Data Centers**

Legacy deduplication appliances suffer from legacy problems. They act only as islands of storage, and are unable to work together to provide global data reduction across these silos. These appliances are inefficient, and costly to the end user. Today's webscale data centers not only need flexible solutions for primary workloads, they need it where up to 80% of the data resides, in the secondary workloads of the data center. Data growth in the secondary storage space has increased exponentially when compared to data growth on primary workloads. Both the Veeam Backup and Replication Suite and the Cohesity Data Platform implement advanced data deduplication algorithms which will drastically reduce the on-disk footprint of any data protection environment. Allowing Cohesity to deduplicate all backup data being streamed from any number Veeam Backup Proxy servers will result in fewer bytes written to disk saving hard dollars on expensive appliances and time searching for which silo data resides on. No matter how big the environment gets, the Cohesity Data Platform can grow to fit the need.

**Pay for what you need, when you need it**

Webscale data platforms enable a pay-as-you-grow model. No longer do storage administrators need to pay today's prices for disks they won't use for years to come. Built on trusted x86 hardware and a fully distributed software stack, the Cohesity Data Platform infinitely scales from as few as three nodes, to racks and racks of nodes. Built with the same internals that power today's biggest webscale companies, the Cohesity Data Platform is resilient to failures. Cohesity's self-healing software is constantly scanning to insure that all data stored is uncorrupted and fully available at all times. This means that no matter what happens to the hardware, Veeam Backup or Recovery jobs can continue to chug along. The intelligences of a fully distributed architecture doesn't stop at self-healing. Fine-grained Quality of Service (QoS) controls allow backup administrators to weigh individual backup jobs with a priority setting. These QoS controls get applied to every write or read operation that lands on the platform, and stick with it as it traverses the layers of software, a first of it's kind in data policy control.

**Elastically expand into the Cloud with Cloud Archive and Replication**

The job of the backup administrator doesn’t end at the daily backup job. Succinct off-site replication and archival strategies need to be implemented. These can lead to lots of vendor fragmentation and several disparate user interfaces to manage. By leveraging the Cohesity Data Platform’s native site-to-site replication features and the ability to archive to the cloud, Veeam and Cohesity become a single stack solution to tackle the end-to-end requirements of the entire Data Protection program. Replication between Cohesity Data Platforms can be setup in a one-to-one or many-to-one fashion which can allow the Backup Administrator to conform to all off-site data storage needs for any number of data center locations. Attaching a cloud-based archival service such as Amazon Glacier or Google Near Line provides an off-site archiving strategy.
which can be accomplished without needing a secondary data center at all. Both the replication of backup data and the archival of backups happen based on policies set up when a pool of storage is created for Veeam. These data transfers happen automatically and continually without any need for manual intervention. With the index of data locations held locally to the platform, locating data is as easy as if it were all local.

In-Place restores for Rapid Recovery and Testing

Backups are only as good as their recoveries, and with a Cohesity backed Veeam implementation backup administrators are able to simulate Disasters, restore virtual machines in-place, or leverage the platform to host virtual machines in the event of an actual disaster or data center loss. The Cohesity Data Platform can not only serve as a target disk repository for backup jobs, but also serve as a VMware Datastore to host virtual machines. By recovering directly to the same platform, Veeam is able to restore in a fraction of the time it takes to restore from a legacy tape library or storage appliance. This flexibility reduces the overall need for multiple user interfaces and hardware stacks to perform end-to-end data protection tasks.

Storage Configurations for Veeam Backups

When configuring the Cohesity Data Platform as a target Storage Repository, a few storage specific settings should be applied to optimize the performance of the Cohesity Platform. When creating the CIFS based storage repository, select the Advanced button under the Repository section of the wizard. From this window, select the Use per-VM backup files option, and deselect the Limit maximum concurrent connections option. This will enable more concurrent streams to the Cohesity Platform, increasing the overall throughput.
It is also strongly recommended the Veeam is configured with as many storage repositories as there are nodes participating in the platform. Each storage repository should be created with each of the VIPs that are being served by the platform. All the storage repositories should point to the same target SMB Share created on the Cohesity Platform. This will enable an even distribution of load from all Veeam backup jobs to the Cohesity Platform. Once all of the storage repositories have been created, jobs can be divided up amongst them. Even though backup jobs are distributed across disparate storage repositories, they will all be globally deduped by Cohesity’s Global Deduplication Engine.

**Job Configurations for Veeam Backups**

The Cohesity Data Platform can perform data deduplication globally, coupled with specific configurations to the Veeam Backup and Recovery Repository and Job settings can be optimized to take advantage of the distributed nature of the platform. As both the Cohesity Data Platform and the Veeam Backup and Replication software can provide deduplication services, it is recommended that the resource intensive tasks of in-line deduplication be off-loaded to the Cohesity Data Platform. This allows data across all Veeam jobs to be deduplicated against each other, resulting in fewer blocks written to disk. Since Cohesity is handling the deduplication, it is also recommended that Veeam’s Compression services be turned off. This allows the Cohesity deduplication engine to perform the data reduction tasks more effectively. These settings can be found at the job level settings under the **Advanced Storage** settings.
During the backup job configuration, it is a best practice to limit the number of virtual machines per-job to 12. This will allow for more concurrent streams of data to every node on the Cohesity Data Platform. These jobs can be spread out across any number of Veeam Backup Proxy Servers that are available in the Veeam Backup and Replication environment.

**Delivering Webscale Simplicity for Secondary Storage Consolidation**

Once the Storage and Job level configurations have been optimized, the combined Veeam and Cohesity Data Protection stack becomes a simple, scalable, and reliable platform to drive secondary storage consolidation. Delivering best in class data protection for virtualization, Veeam has become the market leader in easy to manage software, flexible enough to grow to any size virtualization environment. The Cohesity Data Platform brings unparalleled scalability on top of technical advancements such as: global deduplication, in-place restores for test and dev environments, and in-place analytics. This gives data center administrators a single pane of glass to hyperconverged secondary storage workflows.
Appendix A: Deploying the Cohesity Data Platform with Veeam Backup & Replication

This chapter describes the steps for deploying the Cohesity storage appliance with the Veeam backup application.

Deployment Prerequisites

Veeam Backup & Replication v9 infrastructure with at least four Backup Proxies
  • A Configured Cohesity Data Platform Cluster with at least four nodes

Configuration Overview

The configuration for Veeam and the Cohesity Data Platform cluster is divided into a few simple steps.
  • Create a SMB share on the Cohesity Data Platform – The Cohesity Data Platform presents itself as a storage target in the form of an SMB file share. This task will step through the process of creating and sharing the SMB export to Veeam Backup Proxy.

  • Add the Cohesity Cluster storage as a storage repository on the Veeam Backup and Replication server.
  • A Configured Cohesity Data Platform Cluster with at least four nodes

Creating a SMB Export on the Cohesity Data Platform

This task will create a file system that will be exported as a SMB share on the Cohesity Data Platform. In the first step of this task a storage container called a View Box will be created that will allow us to set the policies for the shares to be exported. We can set the data management policies such as enabling data reduction functionality and the type of data deduplication whether inline or post process. In the second step a file system will be created called a View. We will associate the View with the View Box created in step 1 which will inherit the policies from the View Box. In the last step we will grant access to the Backup Proxies by putting its IP address in the whitelist that grants access to the shares.

  1. Create a View Box
  2. Create a View
  3. Whitelist the IP subnet of the Veeam Backup Proxy
Creating a View Box
1. From the Cohesity UI, navigate to the **Add a View Box** Tab, select **Platform > View Boxes > Add a View Box**.
Creating a View

2. From the Cohesity UI, navigate to **Create View** Tab, select **Platform > View**. Then click on the **Create View** button.

**Figure 4. Navigation to View Creation Window**

**Figure 5. New View Box Creation Settings**
3. Whitelist the IP Subnet of the Veeam Backup Proxies

This step will allow access of the Backup Proxy to the newly created SMB share. You will need to have the IP address of the Backup Proxy added to the list, or the IP address range if there are multiple Backup Proxies that needs access.

From the Cohesity UI navigate to the Specify Subnet tab select Platform > Views > Whitelisted Subnets > Specify Subnet (Figure 6.) Enter the IP address and the subnet mask of for the network range for the Backup Proxy. In the case of a single host a 32bit mask should be used. Ex 255.255.255.255.
Configuring the Veeam Backup Proxy for Multiple Streams
From the Veeam Backup & Replication GUI select Backup Proxies from the left hand navigational menu.

Figure 8. Veeam Backup & Recovery Backup Proxy view

Next right-click on each of the Backup Proxies in the list and select **Properties**

Figure 9. Backup Proxy Property Settings
In the **Backup Proxy Settings** window, ensure that the number of concurrent connections is set to the number CPUs/vCPUs that are on the system. Cohesity recommends at least 6 vCPUs/Physical Cores per Backup Proxy. This setting will push more concurrent data to the Data Platform, and increase the overall throughput to the platform.

Figure 10. Edit Backup Proxy Server Settings. Ensure Max Concurrent Tasks is set to the number of vCPUs or CPU Cores on the host.
Select **Add Repository** from the menu options located on the top navigation menu. This will launch the **New Backup Repository** wizard which will walk you through adding a new storage repository.

---

**Figure 11. Select Select Backup Infrastructure and Backup Repositories from the Veeam Console**
Once the new storage repository has been named, click the **Next** button to select the type of storage repository to configure. Select **Shared folder** to configure the Cohesity SMB target storage.
After the **Type** has been selected, click next to configure the **Shared folder** path and the Backup Proxy to use as the **Gateway** server. Then click **Next**.

---

Figure 13. New Backup Repository Type setting.
Figure 13. Enter the Share Settings using the Cohesity VIP. A repository for each VIP will need to be created in order to fully distribute the load.

Once the **Shared folder** and **Gateway Server** have been configured accept the default **Repository** settings and continue on.
Once the Shared folder and Gateway Server have been configured accept the default Repository settings and continue on.

Figure 14a. Ensure Limit maximum concurrent tasks is un-checked.
To enable vPower NFS on the Mount server make sure the same Gateway Server from step 4 is selected, and the Enable vPower NFS check box is selected. (figure 13)
Repeat steps 2 through 7 for the remaining VIPs evenly spread across the Backup Proxy infrastructure. This will evenly distribute the backup load to the Cohesity Target Storage subsystem.

**Note:** To fully utilize all of the Cohesity Data Platform nodes, a Backup Repository will need to be created for each of the nodes of the cluster. For each Repository use a different VIP, this will ensure an even distribution of load across the Platform.

**Summary**

In summary, combining Veeam Backup & Replication with the Cohesity Data Platform delivers a comprehensive backup solution that is infinitely scalable, highly available that provides unparalleled performance.