Technical Report

NetApp Scale-out Data Protection with Commvault

James Bradshaw, Jeannine Walter, Savita Kumari, NetApp
Bryan Clarke, Nigel Tozer, Sarbjit Singh, Commvault
May 2020 | TR-4831 | Version 1.1

Abstract

This paper introduces NetApp® Scale-out Data Protection, a combined NetApp and Commvault solution that offers expanded rapid recovery support for critical data across a data fabric powered by NetApp technology. NetApp Scale-out Data Protection (SDP) takes advantage of the NetApp portfolio, and is powered by Commvault Complete Backup & Recovery software to provide NetApp Snapshot™ management, backup, and disaster recovery functionality.

In partnership with

COMMVAULT ®

■ **NetApp**®

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Executive Summary

With many organizations struggling to meet the availability and rapid recovery goals demanded by their business, NetApp and Commvault have partnered to create an appliance-like verified architecture specifically for this purpose. Running on NetApp HCI and powered by Commvault software, the primary recovery media is fast all-flash storage, with a cloud tier based on NetApp StorageGRID® software for long-term retention. This clever design meets the most demanding recovery needs and keeps costs under control, while exceeding modern ease-of-use requirements for scalability and day-to-day management.

1.1 The Challenge

The enterprise IT landscape is going through a digital transformation as it shifts from handcrafted, single-purpose data center designs to new technologies that deliver agility, scalability, and greater efficiency. The hybrid cloud relies on the availability and resiliency of data. Enterprises require a complete data protection solution that is reliable, flexible, and easy to use. Although virtualizing an environment provides an increased level of data availability, meeting aggressive recovery point objectives (RPOs) and recovery time objectives (RTOs) becomes increasingly difficult.

Legacy and traditional approaches to backup and recovery just don't meet the needs of the modern enterprise—they're broken. Whether it's because of a lack of automation that slows response times, an inability to scale cost effectively, a lack of support for today's modern cloud environments, or an inability to meet governance regulations like the California Consumer Privacy Act (CCPA) or the European Union's General Data Protection Regulation (GDPR), yesterday's solutions are not addressing your needs today.

Despite these challenges, the value of your data continues to increase, making protection of your data more important than ever. You need a modern data protection solution that allows your company to innovate and grow while keeping the business running uninterrupted, proactively enabling rapid change, and applying effective governance in the face of regulatory shifts. Additionally, you want a solution that reduces administrative complexity while being delivered at the same or lower cost as your current solution.

Industry data that supports this view can be found in a survey conducted by the Enterprise Strategy Group in February 2020 (ESG 2020 Technology Spending Intentions), which showed that 64% of respondents said IT was more complex than two years ago. It also revealed that "improving data backup and recovery" was the #1 priority for data center modernization plans.

Ultimately, you need a data protection strategy that helps you address these challenges and increase your ability to respond across the following areas:

- **Recovery readiness:** What is your response when the business asks how ready you are to recover in the face of a disaster?
- **Data governance:** How do you adhere to complex data privacy and protection laws?
- **Constant change:** How fast can you respond to change in the business?
- **Management complexity:** How do you properly manage remote, physical, cloud, and virtual environments?
- **Optimizing costs:** How can you successfully optimize your current and future infrastructure utilization?

Additionally, as your company grows and your needs expand, you can simplify disaster recovery, development and testing, and workload migration through automation. You can also extract data insights for better data governance and business outcomes.

Commvault software also employs machine learning to help you meet your most demanding SLAs, by providing smart reports and automation.

1.2   About NetApp

NetApp is the leader in cloud data services, empowering global organizations to change their world with data. Together with our partners, we are the only ones who can help you build your unique data fabric. Simplify hybrid multicloud and securely deliver the right data, services, and applications to the right people at the right time.

1.3   About Commvault

Commvault believes in data readiness. Commvault helps organizations of all sizes intelligently manage data through solutions that store, protect, manage, and use their most critical asset—their data. Commvault software, solutions, and services are available from the company and through a global ecosystem of trusted partners. Commvault employs more than 2,300 highly skilled individuals across markets worldwide, is publicly traded on NASDAQ (CVLT), and is headquartered in Tinton Falls, New Jersey, in the United States.

1.4   Target Audience

The target audience for this solution also includes the following groups:

- Solution and cloud providers with data centers that require an easily deployable solution scaling to petabytes of data
- Channel partners and global systems integrators who assist customers to transform and modernize their data protection solution
- Executives and IT managers who want to achieve business continuity that can scale at a low cost
- Healthcare providers and businesses with GDPR
- Disaster recovery and virtual environment administrators, engineers, and architects
- Professional services and sales engineers

2   Solution Overview

To meet these challenges, Commvault and NetApp have collaborated to create NetApp Scale-out Data Protection (SDP), based on NetApp HCI and Commvault Complete Backup & Recovery for NetApp storage. The solution combines NetApp HCI for easy deployment and scaling with NetApp AFF technology for high-performance backup, plus NetApp StorageGRID object storage for long-term retention. Commvault software manages the backup and recovery processes, including data tiering from flash to object storage tiers.

Figure 1) Solution description.
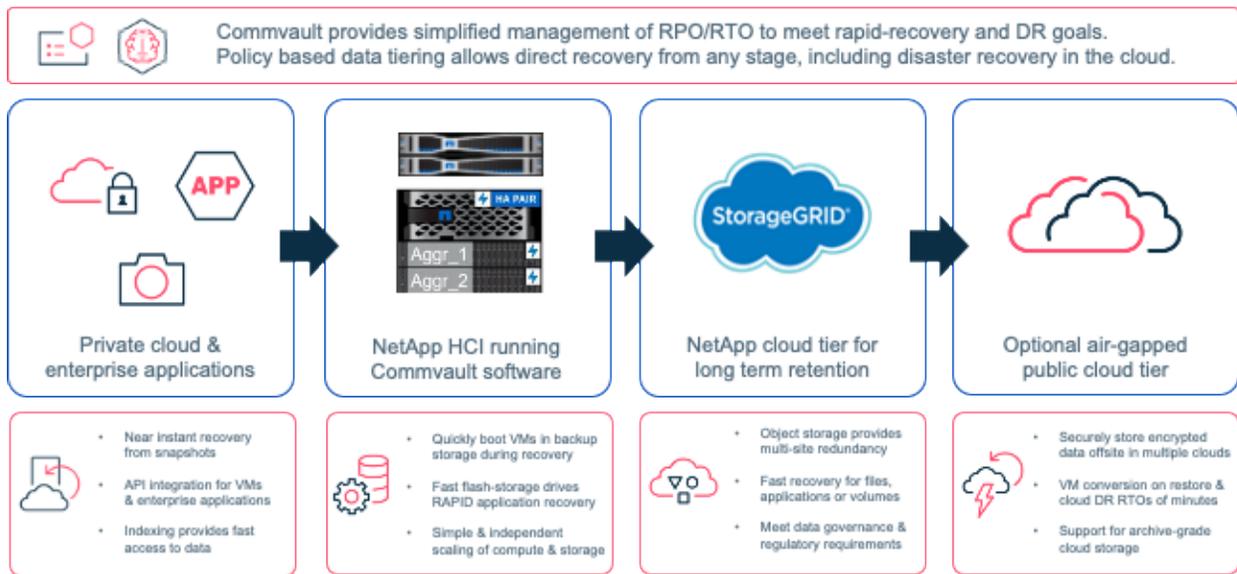
2.1 Fast, Cost-Effective, and High Capacity

Under normal circumstances, you would be forced to compromise and pick the two features that meet your most pressing demands. However, because we combined NetApp HCI, AFF, and StorageGRID with Commvault software, the normal compromises you would expect to make for your recovery systems simply evaporate.

This partnership combining NetApp SDP and Commvault's vast support for NetApp Snapshot™ copies from arrays and the cloud and a full set of application and virtual machine (VM) APIs provides a near-instant first line of recovery. Depending on business needs, the Snapshot copies can then be moved to the NetApp all-flash tier of the backup storage. If a system needs restoring, the all-flash NetApp recovery tier enables the VM to be quickly started by using Commvault's LiveMount and LiveRecovery in the deduplicated backup store while the recovery is taking place. This capability can help you meet your SLAs with time to spare.

NetApp HCI is an enterprise-scale hybrid cloud infrastructure solution that delivers compute and storage resources in an agile, scalable, easy-to-manage two-rack-unit, four-node building block. The NetApp Deployment Engine (NDE) and NetApp Element® plug-in for vCenter deploys and manages NetApp HCI.

NetApp HCI is focused on helping organizations manage multiple hybrid cloud environments in an elegant and simplified manner. Commvault software adds advanced data management.

This combination creates a complete backup solution. It takes a holistic approach to data management that helps enterprises reduce risk, improve operational efficiencies, cut costs, and derive more value from their data and information.

2.2 Benefits

Key features of the NetApp SDP solution include:

- Architecture:
    - NetApp Verified Architecture guaranteeing quick and effective deployment.
    - Commvault IntelliSnap software for managing Snapshot copies uses NetApp hardware orchestration to provide enhanced RPOs and RTOs.

- Complete data protection solution supporting all major operating systems, applications, and databases on virtual and physical servers, NAS shares, cloud-based infrastructures, and endpoint/mobile devices.
- Scalability and efficiency:
  - Scale compute, the primary recovery flash tier, and long-term storage tier components individually to meet the unique demands of your organization.
  - Perform deduplication for the NetApp flash tier and object storage tier.
  - Infinitely scale StorageGRID for long-term retention with the option to scale to public cloud.
- API integration/automation:
  - Full integration with VMware technologies—adds value by providing automation. The Commvault Virtual Server Agent (VSA) uses vStorage APIs for Data Protection (VADP), with support for all guest operating systems supported by VADP.
  - Reduce complexity by using prebuilt automation workflows and an extensive API library.
- Flexibility and performance:
  - All-flash tier for faster recoveries.
  - Granular recovery of VMs and files, including support for advanced recovery options like LiveBrowse, LiveMount, and LiveRecovery.
  - Built-in workload portability—when restoring a VM, you can choose to restore to a different hypervisor. This capability can be used to migrate VMs to a different virtualization platform (including public cloud providers).
  - Built-in backup, Snapshot management, replication, and archiving, with support for content indexing for e-discovery/search.
  - Cutting-edge end-user experience empowering users to protect, find, and recover their own data by using common tools such as web browsers, Microsoft Outlook, and File Explorer.
  - Policy-based data management: managing data according to business needs and not physical location.
- Security:
  - Multilayered protection from ransomware: Commvault software combines immutable recovery tiers from NetApp with proactive AI-based anomaly detection and other built-in ransomware protection technologies to detect attacks, keep your data safe, and enable fast recovery.
  - Advanced security capabilities to limit access to critical data, provide granular management capabilities, and provide single sign-on access for Active Directory users.

2.3   Data Fabric Integration Points

The data fabric is NetApp's vision for the future of data management. The data fabric seamlessly connects different clouds, whether they are private, public, or hybrid environments. A data fabric unifies data management across distributed resources to allow consistency and control of data mobility, security, visibility, protection, and access.

The previous section outlined a few of the benefits of Scale-out Data Protection. Here are some simple questions that can help determine if Scale-out Data Protection is right for you.

- Are you tired of overrunning your data protection windows?
- Do you hate how backup processes consume precious resources?
- Do you want to give those resources back to the apps that need them?
- Do you spend lots of human hours managing, planning, and optimizing?
- Do you want to crush the complexity of multiple data silos for backup, replication, archiving?

Scale-out Data Protection with NetApp and Commvault satisfies those needs and enables you to have completely integrated endpoints and a complete API-driven data services solution. It's your data fabric; it's your time.

Combining data fabric services into a virtual consumption layer enables your data to be protected and secured throughout its lifecycle.

- File services for high-performance cache (short-term retention and quick recovery)
- Object services that use StorageGRID technology (highly durable long-term retention)
- Integrated replication services (reduce backup time, recovery time, and nearly eliminate compute resource demands)
- Data orchestration and lifecycle management with Commvault software (single integrated approach, ease of management, automatic data tiering)
- Acceleration of the journey to the cloud by expanding the integrated data protection policies to cloud workloads running on NetApp Cloud Volumes ONTAP® and Azure NetApp Files

3   Solution Design and Best Practices

3.1   Solution Testing Configuration—Physical Setup Overview

This section illustrates the solution testing configuration and details the hardware and software versions used for validation testing. See the Appendix for details about the environment used to produce the workload.

Figure 2) Solution Testing configuration.
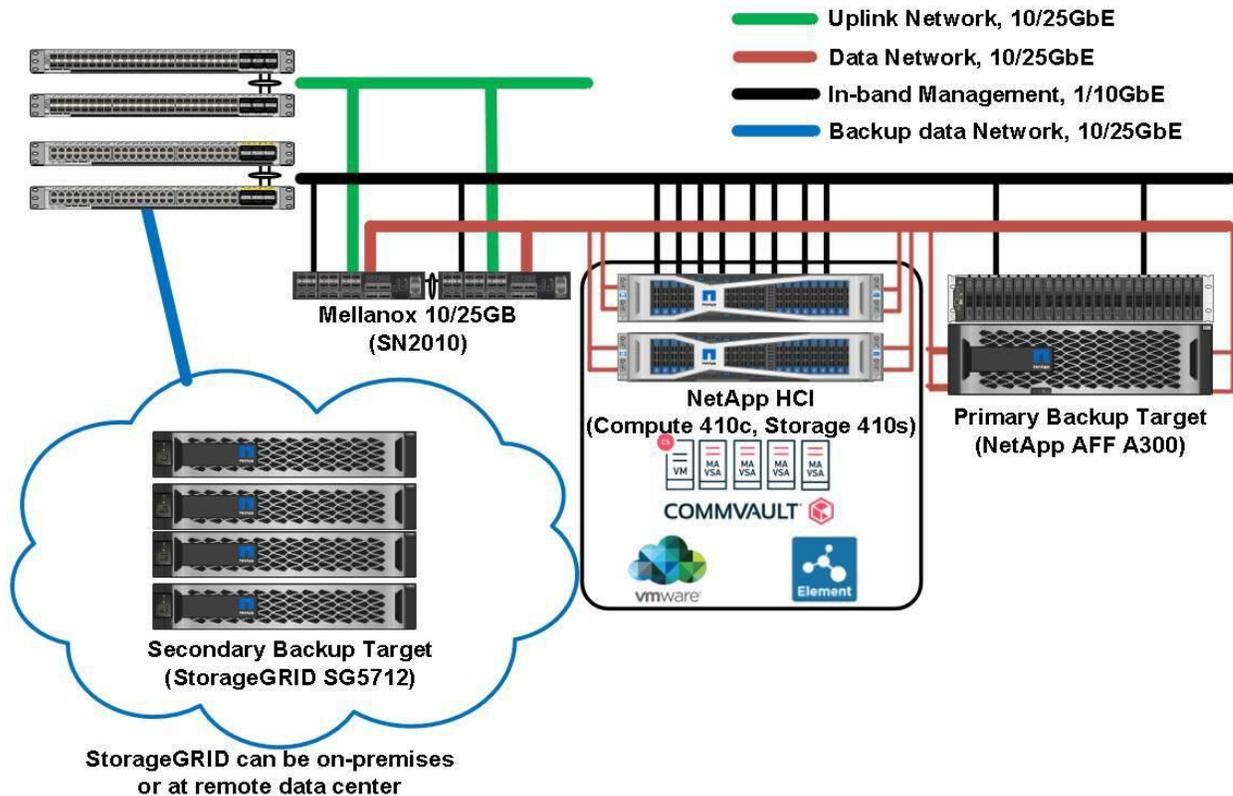
3.2 Solution Validation Hardware

This section lists the hardware used for this solution's validation testing.

Table 1) NetApp HCI hardware.

Hardware	Quantity	CPU	Cores	Memory	Disk	Usable Capacity
Compute: NetApp H410C	4	Intel Xeon Gold 5120 CPU@ 2.20GHz (Cascade Lake)	24	512GB	n/a	n/a
Storage: NetApp H410S	4	Intel Xeon Gold 5120 CPU@ 2.20GHz (Cascade Lake)	24	256GB	x24, 960GB	19TB

Table 2) Networking.

Hardware	Quantity	Speeds	Ports
Mellanox SN2010	2	10/25GbE	36

Table 3) Primary backup target.

Hardware	Quantity	Disk	Usable Capacity
NetApp AFF	1	n/a	n/a
StorageShelf DS224C	1	x24, 3.8TB (SAS SSD)	64TB

Table 4) Secondary backup target.

Hardware	Quantity	Disk	Usable Capacity
StorageGRID SG5712	4	x48, 4TB (NL-SAS HDD)	136TB

3.3 Solution Validation Software

This section lists software versions used for this solution's validation testing.

Table 5) Software versions.

Product Family	Product Name	Product Version	Operating System
NetApp HCI	Element software	11.7	n/a
	NDE	1.7 P1	n/a
	NetApp ONTAP	9.7	n/a
	Element vCenter Plug-in	4.3.0	n/a
VMware vSphere Enterprise Plus	ESXi	6.7.0, 10302608	n/a
	vCenter Server Appliance	6.7.0, 10244857	n/a
Mellanox	SN2010	3.7.1134	Onyx

Product Family	Product Name	Product Version	Operating System
Commvault	CommServe	v11 SP18 or later	See System Requirements
	MediaAgent	v11 SP18 or later	See System Requirements
NetApp StorageGRID	StorageGRID object storage	11.3.0.1	n/a

4 Sizing

Consult your NetApp data protection specialists for specific sizing for your environment. NetApp data protection specialists can use the Commvault Total Backup Storage Calculator tool to estimate the backup infrastructure requirements. The tool requires Commvault Partner Portal access. Sign up for access, if needed.

4.1 Commvault Sizing Inputs

The following tasks can be used to perform discovery for sizing of the data protection solution:

- Identify the system or application/database workloads and corresponding front-end capacity (in terabytes [TB]) that will need to be protected.
- Identify the VM/file workload and similar front-end capacity (TB) that will need to be protected.
- Identify short-term and long-term retention requirements.
- Identify the daily % change rate for the datasets/workloads identified.
- Identify projected data growth over the next 12, 24, and 36 months.
- Define the RTO and RPO for data protection/recovery according to business needs.

When this information is available, the backup infrastructure sizing can be done resulting in a breakdown of required storage capacities for the AFF tier and the StorageGRID tier. The solution provides the flexibility to size the flash and object storage tiers independently. The flash tier is typically used for short-term retention and for workloads requiring faster recovery, typically Tier0 or Tier1 workloads. Tier2 or Tier3 workloads can be directly backed up to StorageGRID tier for short-term and long-term retention.

4.2 NetApp HCI

For this solution, the Commvault management server (CommServe) and MediaAgent VMs are deployed on a NetApp HCI system with four compute and four storage nodes. The system provides 112 total virtual CPU (vCPU) cores, 19TB of usable capacity, and 200,000 total IOPS supporting small, medium, large, and extra-large system requirements.

Note: This solution is intended to be a closed system. Only the Commvault CommServe and MediaAgents are intended to run on the NetApp HCI system. No other workloads should be run on this system.

These are the overall system requirements for NetApp HCI and Commvault:

Note: Consult your NetApp data protection specialist for specific sizing for your environment.

- For CommServe server sizing, see Hardware Specifications for the CommServe Server.
- For NetApp HCI specifications, see the NetApp HCI datasheet.
- For MediaAgent sizing, see the Deduplication Building Block Guide.

4.3 NetApp StorageGRID

In order to perform StorageGRID sizing, you must understand the workload:

- Usable capacity
- Average object size
- Performance requirements
- ILM policy applied

The amount of usable capacity can be found by reviewing the customer's requirement for the size of the backup workload tiered to StorageGRID.

Average object size is an input parameter that helps with sizing for performance in a StorageGRID environment. The average object sizes used for a Commvault workload depends on the type of backup. Table 6 lists the different object sizes based on the type of backup.

Table 6) Average object size used in Commvault backups to StorageGRID.

Backup Type	Average Object Size
Aux-Copy to StorageGRID	32MB
Direct backup to StorageGRID (deduplication enabled)	8MB
Direct backup to StorageGRID (deduplication disabled)	32MB

In addition, understanding the customer's performance requirements for full backups and incremental backups can help determine sizing for the number of StorageGRID storage nodes.

StorageGRID information lifecycle management (ILM) policy data protection methods determine the capacity overhead required for storing Commvault backups and it will have an impact on sizing a grid for capacity.

StorageGRID ILM replication is one of two mechanisms used by StorageGRID to store object data. When StorageGRID matches objects to an ILM rule that is configured to create replicated copies, the system creates exact copies of object data and stores the copies on storage nodes.

Erasure coding is the second method used by StorageGRID to store object data. When StorageGRID matches objects to an ILM rule that is configured to create erasure-coded copies, it slices object data into data fragments, computes additional parity fragments, and stores each fragment on a different storage node. When an object is accessed, it is reassembled using the stored fragments. If a data or a parity fragment becomes corrupt or lost, the erasure-coding algorithm can recreate that fragment using a subset of the remaining data and parity fragments.

For example:

- Storing two whole replica copies of backup objects will result in 2x storage overhead
- Storing 2+1 erasure coded copy of backup objects will result in 1.5x storage overhead

For this solution, StorageGRID is used as secondary storage. We chose to implement a baseline system that represents an entry-level StorageGRID deployment on a single site:

- Admin node: VMware VM
- Load balancer: VMware VM
- Storage nodes: 4x SG5712 with 4TB drives
- In this test environment, the primary admin node and gateway node were deployed as VMware VMs with the minimum production workload requirements listed in Table 7.

Table 7) StorageGRID admin node and gateway node requirements.

Node Type	Quantity	vCPU	RAM	Storage
Gateway node	1	8	24GB	100GB LUN for OS
Admin node	1	8	24GB	100GB LUN for OS 200GB LUN for admin node tables 200GB LUN for admin node audit log

StorageGRID is typically deployed in two or more sites with data protection policies that replicate data to protect against node and site level failures. It is important to understand that when data is committed to StorageGRID, it is protected by making multiple copies of the data or by using erasure coding to protect in a single site or across many. StorageGRID can also use capacity in other cloud storage systems such as AWS, S3 compatible, and Azure blob.

5   Scaling

5.1   Commvault

NetApp SDP allows the flexibility to scale out compute and storage independently. The Commvault CommServe and MediaAgent VMs running on NetApp HCI can be scaled with data growth. As the amount of data to be protected grows in your environment, additional MediaAgent VMs can easily be provisioned on the NetApp HCI. The MediaAgent VMs can be scaled as a building block per the overall backup storage capacity being managed. Please refer the maximum supported grid backend storage as guideline for scaling MediaAgent VMs.

See the Commvault documentation for recommended specifications for the CommServe and MediaAgent VMs. For the NetApp SDP solution, the extra-large CommServe and MediaAgent specifications are recommended.

Review these Commvault documentation links for additional reference:

* Recommended specifications for the CommServe VM
* Recommended specifications for MediaAgent VM
* Recommended specifications for the Virtual Server Agent VM

For more details, see the CommCell performance tuning and scalability guides.

5.2   NetApp HCI

Expanding your NetApp HCI installation is easy. Using the NetApp Hybrid Cloud Control manageability suite, you can expand NetApp HCI storage and compute resources either separately or at the same time. NetApp HCI detects the existing network configuration and offers you configuration options within the existing networks and virtual LANs (VLANs).

For more information, see Expanding Your NetApp HCI Installation.

5.3   AFF System

With your AFF system, to get more capacity, you can hot-add disk shelves with IOM12 modules to an existing stack of disk shelves with IOM12 modules. Alternatively, you can hot-add a stack of disk shelves with IOM12 modules directly to a SAS host bus adapter (HBA) or an onboard SAS port on the controller.

For more information about shelf expansions, see SAS Shelves with IOM12 Modules.

5.4 StorageGRID

You can expand a NetApp StorageGRID system by adding storage to storage nodes, adding new grid nodes to an existing site, or adding a new data center site. You can perform expansions without interrupting the operation of your current system.

StorageGRID scales performance by using either higher performance nodes for storage nodes or the physical appliance which runs the load balancer and the admin nodes or by simply adding additional nodes.

Note: StorageGRID also supports third-party load balancers.

For more information about expanding the StorageGRID system, see StorageGRID 11.3 Expansion Guide.

6 Solution Deployment

This section provides an overview of deploying Commvault Complete Backup & Recovery on NetApp HCI. Completing all sizing and proper planning exercises is key for a successful deployment of this backup solution.

6.1 NetApp HCI

NetApp Deployment Engine

NetApp HCI simplifies Day 0 deployment with the NetApp Deployment Engine (NDE), completing over 400 deployment and configuration tasks in a fraction of the time it takes to deploy manually. When deployment is complete, the NetApp HCI system provides compute and storage for the Commvault CommServe and MediaAgents.

Complete the prerequisite steps and deployment steps provided in the NetApp HCI 1.7P1 Deployment Guide.

Note: This solution is intended to be a closed system. Only the Commvault CommServe and MediaAgents are intended to run on the NetApp HCI system. No other workloads should be run on this system.

Post-Deployment Checklist

In vCenter, add the required distributed port-groups for the following VLANs:

- In-band management traffic
- VLANs from nonrouted private virtual environments (optional)

The NetApp Element OS Plug-in for vCenter was used to provision datastores for CommServe and Media Agents.

For the latest details on the Element Plug-in for vCenter Server, see the NetApp HCI VCP 4.3 User Guide.

Table 8 lists the operating systems that were used for deployment of the Commvault components:

Table 8) Operating system deployment used for the Commvault backup components.

Component	Quantity	OS	CPU	RAM	Datastore	Size
CommServe	1	Windows Server 2016	16	128GB	OS	500GB
					SQL	500GB

Component	Quantity	OS	CPU	RAM	Datastore	Size
MediaAgent	4	Windows Server 2016	16	128GB	OS	500GB
					Index cache	500GB
					DDB	1TB

6.2 NetApp ONTAP Deployment and Configuration Overview

The following steps are required for providing a backup target that uses either NFS or CIFS:

- Complete Systems Installation and Setup Instructions on the NetApp AFF array.
- Complete cluster administration to provision clusters, networking, physical, and logical storage.
- Complete either NFS, CIFS, or iSCSI configuration that will be used to connect to MediaAgents:
    - Provisioning for NAS protocols
    - Provisioning for SAN protocols

6.3 NetApp StorageGRID Deployment and Configuration

For this solution, we chose a single site instance consisting of four SG5712 storage nodes with 4TB drives, a load balancer node, and an admin node deployed as VMware VMs. This configuration represents a typical entry-level system, optimized for cost-effective object storage.

For the data protection policy, StorageGRID implements integrated lifecycle policy or ILM to manage and protect data. ILM rules are evaluated in a policy from top to bottom. We implemented the ILM policy shown in Table 9.

Table 9) StorageGRID ILM policy.

Rule	Qualifiers	Ingest Behavior
Erasure code 2+1	Object over 200KB	Balanced
2 copy	All objects	Dual commit

The ILM 2 copy policy is the default rule. The policy applied for testing was 2+1 erasure coding to any object 200KB or larger. The default rule was applied for objects smaller than 200KB—this is in accordance with StorageGRID best practices.

For the Amazon Simple Storage Service (Amazon S3) bucket configuration, the bucket created for Commvault was set to the consistency level "available."

6.4 Commvault Deployment and Configuration

To validate the NetApp SDP solution, Commvault components (CommServe and MediaAgent) were deployed as Windows VMs on NetApp HCI. One CommServe and four MediaAgent VMs were used as part of the testing.

Before installing the Commvault software, review the following requirements:

- System Requirements for the CommServe Server
- System Requirements for the MediaAgent

Other packages that are installed along with the Commvault software might require a license. To review the list of packages, see Additional Packages Installed with the CommServe Server.

For more information about licensing, see License Administration.

To install the Commvault CommServe and MediaAgent components, follow the steps in the Commvault documentation. You must install the Command Center package to have access to the Command Center and Web Console interfaces.

After you complete the core setup, the Guided Setup page appears on the navigation pane of the Command Center. On the Guided Setup page, you can configure backup and recovery for file servers, VMs, laptop and desktop computers, and databases. For the configuration procedure, refer to Configuring Solutions in the Command Center.

To add the NetApp AFF system as a disk target that uses CIFS/NFS, proceed to configuring disk storage. You can add a NetApp StorageGRID bucket as a cloud storage target for Commvault. To begin adding a StorageGRID bucket, proceed to configuring cloud storage. To start protecting data, create one or more server backup plans according to RPO/RTO needs. Review the NetApp system requirements for protecting primary workloads on NetApp technology.

On the CommServe, to enhance performance of data movement operations, the number of data readers were increased. Data Readers determines the number of parallel read operations while the data is backed up. Configuring multiple data readers per subclient on disk arrays can improve the backup performance of clients. The data readers were increased to 25 during the validation testing.

7 Solution Testing and Validation

We validated this solution with Commvault by scaling up the infrastructure and conducting performance and functional testing for the initial workload. Commvault's CommServe was deployed on the first NetApp HCI cluster and the Commvault MediaAgents on the second NetApp HCI cluster. The test workload ran on traditional infrastructure. For more information, see the Appendix in the back of this document. The primary backup target was the NetApp AFF system, and the secondary backup target was the NetApp StorageGRID.

This solution was tested to validate full functionality and scale and to ensure that the system could grow to meet demands.

In this solution, scripts were used to generate unique random data in a series of files. A 50/50 mix of Linux and Windows VMs that were used in testing. Random workloads were spread across the VMs to simulate load.

There were four types of operations performed in the lab:

- Full or baseline:
    - Baseline operations are a one-time operation per client. The Commvault MediaAgent has no other protected datasets, in the case of this lab from these clients or any other, therefore all data blocks will be transferred and stored for protection.
    - This operation will typically be the slowest because all client data must be transferred to the MediaAgent. Because this is a one-time event and is not continual, the baseline operation isn't used in the metrics for calculating the long-term speeds of the solution.
- Full-on-full operations:
    - The full-on-full protection operations are the only operations that are done with no data change rates.
    - Full-on-full protection operations are significantly faster than baseline operations, because there is no data transferred. Of course, the faster the clients, the faster the full-on-full operations are.
- Incremental:

- Each incremental operation has a 1% data change rate associated with it.
- Only the changed or "unique" blocks of data are sent across the network.
- DASH Full:
  - DASH Full operations provide the outcome of a traditional synthetic full operation without having to reprocess all the data. The previous baseline and incremental operations store all the relevant blocks that need to be protected but also store the metadata and index information about the changes during each operation. Processing these indexes and metadata components creates a new full protection point without having to reread and rehydrate all the backed-up data to protect it again.

7.1 MediaAgent Tests

Validations were run on the virtual MediaAgents to ensure that the loads generated by normal operation would not strain the underlying NetApp HCI environment enough to decrease overall performance. These tests included Commvault's SIDB2 simulation tool and IOMeter. The tests were performed on a single MediaAgent and then simultaneously on two MediaAgents. Then the tests were scaled up to four MediaAgents simultaneously to ensure that as the Commvault infrastructure grew, the underlying NetApp HCI environment could handle the additional load. With all four MediaAgents running under load, the underlying infrastructure was able meet the demand and had no impact on overall solution performance, leaving additional resources for growth of the Commvault environment.

7.2 Solution Validation Tests

Several jobs were run to validate that the most common operations ran successfully. Table 10 and Table 11 describe the tests that were run. All the tests used a range of 500 to 1,000 VMs with a 50/50 split of Windows and Linux VMs with 25 reader sets within the software.

Table 10) VSA streaming operations.

Test Number	Test Description
1	Run full VM backup to AFF
2	Run incremental VM backup to AFF
3	Run DASH full VM backup to AFF
4	Auxcopy from AFF to StorageGRID
5	Full VM restore from AFF
6	File-level restore from VM from AFF
7	Full VM restore from StorageGRID
8	File-level from VM from StorageGRID
9	LiveMount VM from AFF
10	LiveRecovery of VM from AFF

Table 11) VSA IntelliSnap operations.

Test Number	Test Description
1	Run full VM IntelliSnap backup of a datastore
2	Run incremental IntelliSnap VM backup of datastore

Test Number	Test Description
3	Run backup copy to AFF
4	Full VM restore from primary Snapshot copy
5	File-level restore from primary Snapshot copy
6	Full VM revert from primary Snapshot copy
7	LiveMount a VM from primary Snapshot copy
8	LiveMount several VMs from primary Snapshot copy
9	LiveRecovery of VM from primary Snapshot copy
10	LiveRecovery of several VMs from primary Snapshot copy

7.3   Load Testing

We ran various scenarios to generate loads on all systems. We wanted to test end-to-end functionality under loads to further validate that the solution can grow without sacrificing performance. The following tests focus on the back-end infrastructure. We've briefly described some of the tests we used to provide load on all the systems.

The tests shown in Table 12 were run to test load on StorageGRID. During the tests, the average auxcopy from the AFF to StorageGRID ranged from 2TB/hour to 7TB/hour with peaks reaching up to 23TB/hour.

Table 12) Load testing on StorageGRID.

Test Number	Test Description
1	Run 500 VM backup multistream
2	Run auxcopy to StorageGRID
3	Run 500 VM backup multistream
4	Run auxcopy to StorageGRID
5	Run 1,000 VM backup multistream
6	Run auxcopy to StorageGRID

The tests shown in Table 13 were run to produce load on the MediaAgents. The object was to ensure that as more resources were consumed, the MediaAgents could sustain the required performance and would continue to do so. The CPU and memory load were generated to produce 50% load on the MediaAgents while performing backup operations. A baseline backup was performed with 1,000 VMs; the backup had an average baseline throughput of 2.8TB/hour.

Table 13) Load testing on MediaAgents.

Test Number	Test Description
7	Run CPU load on all MediaAgents; run 1,000 VM backup
8	Run auxcopy to StorageGRID
9	Run mem load on all MediaAgents; run 1,000 VM backup

Test Number	Test Description
10	Run auxcopy to StorageGRID
11	Run CPU and mem load on all MediaAgents; run 1,000 VM backup
12	Run auxcopy to StorageGRID

The tests shown in Table 14 were run to ensure that any background process such as data aging (the process of removing data beyond its retention off disk) would not significantly affect other operations.

Table 14) Load-testing background process.

Test Number	Test Description
13	Age-off data from AFF from test 1 and 3; run 1,000 VM backup
14	Age-off data from StorageGRID test 2 and 4
15	Restore multiple VMs from AFF
16	Restore multiple VMs from StorageGRID
17	Live-mount multiple VMs from AFF

The tests shown Table 15 were run to simulate data buildup on the storage targets to ensure that performance was not affected as the back-end storage was consumed. Using the same baseline as before (1,000 VMs at 2.8TB/hour), even as capacity was filling up, tests showed that the average throughputs remained the same at about an average of 2.8TB/hour.

Table 15) Load-testing back-end storage.

Test Number	Test Description
18	Run Datagen on AFF to consume 50% capacity
19	Run Datagen on StorageGRID to consume 50% capacity
20	Run 1000 VM back multistream
21	Run auxcopy to StorageGRID
22	Run Datagen on AFF to consume 90% capacity
23	Run Datagen on StorageGRID to consume 90% capacity
24	Run 1,000 VM backup multistream
25	Run auxcopy to StorageGRID
26	Age-off all data from AFF
27	Run 1,000 VM backup
28	Age-off all data from StorageGRID
29	Run auxcopy of data from test 27

Comparison tests were also run between the virtual MediaAgents and physical MediaAgents by using the same 1,000 VMs. The four virtual Windows MediaAgents on NetApp HCI performed backups at an average of 2.7TB/hour; the four physical Windows MediaAgents performed backups at an average of

2.6TB/hour. Restores were also tested using 100 VMs with the same four virtual MediaAgents for an average of 1.3TB/hour; the four physical MediaAgents also restored at an average of 1.3TB/hour. Using these results as a comparison, the virtual and the physical MediaAgents results were nearly identical.

8 Conclusion

NetApp SDP with Commvault provides fast recovery and infinite scalability, meeting the ever-changing needs of enterprise customers. With Scale-out Data Protection, you can easily move, manage, protect, and orchestrate data across the data fabric. This solution enables you to transform your data protection strategy while providing a solid foundation to be data ready against threats like ransomware or natural disasters. For more information, contact your NetApp account team.

Appendix

Additional Testing and Solution Validation Details

This section lists the hardware and software used for in this solution to produce the workloads illustrated in this document. The following hardware and software produced a workload of 1,000 VMs used to create data that used for backup and restore testing. For details on how workloads were created, see section 7, "Solution Testing and Validation."

Table 16) Hardware used to produce the workload.

Hardware	Quantity
Cisco C220 M4	6
NetApp AFF	1
Cisco Nexus 93180YC	2

Table 17) Software used to produce the workload.

Product Family	Product Name	Product Version
VMware vSphere Enterprise Plus	ESXi	6.7.0 U2
	vCenter Server Appliance	6.7.0.20000 U2
NetApp	ONTAP	9.7
	vCenter Plug-in for ONTAP	4.3.0
Cisco	Nexus 93180YC	7.0.3.I7.8

NetApp HCI

Table 18) NetApp HCI hardware for the workload.

Hardware	Quantity	Memory	Disk	Usable Capacity
Compute node: NetApp H615C	8	512GB	n/a	n/a

Hardware	Quantity	Memory	Disk	Usable Capacity
Storage node: NetApp H610S	4	256GB	x24, 960GB	42.25TB
Network: Mellanox	2	n/a	n/a	n/a

Table 19) Software used for the workload.

Product Family	Product Name	Product Version
VMware vSphere Enterprise Plus	ESXi	6.7.0 U2
	vCenter Server Appliance	6.7.0.20000 U2
NetApp	NDE	1.7 P1
	vSphere Plug-in for NetApp SolidFire®	4.3.0
Mellanox	Onyx	9.4.2000

Acknowledgments

The authors would like to thank the following people for their support and contribution to the creation of this document:

- Michael Phelan, product manager, NetApp
- Alex Goldblatt, senior product manager, NetApp
- Steven Pruchniewski, technical marketing engineer, NetApp
- Charlotte Anderson, solutions engineer, NetApp
- Sucheth Ramgiri, technical lead, Commvault

Where to Find Additional Information

To learn more about the information that is described in this document, review the following documents and websites:

- Solutions Catalog for NetApp HCI
  https://docs.netapp.com/us-en/hci/solutions/index.html
- NetApp HCI datasheet
  https://www.netapp.com/us/media/ds-3881.pdf
- Maintaining Hardware
  http://docs.netapp.com/hci/index.jsp?topic=%2Fcom.netapp.nav.hm%2Fhome.html
- NetApp HCI Documentation Center
  http://docs.netapp.com/hci/index.jsp?topic=%2Fcom.netapp.doc.ic-base%2Fresources%2Fhome.html
- NetApp HCI 1.7P1 Deployment Guide
- NetApp Active IQ
  https://www.netapp.com/us/products/data-infrastructure-management/active-iq.aspx

- NetApp Element Plug-in for vCenter Server 4.3
  https://docs.netapp.com/hci/topic/com.netapp.doc.hci-vcp-ug-170/Using%20VCP%20to%20manage%20NetApp%20HCI%201.7.pdf
- Commvault product portfolio feature matrix
  http://documentation.commvault.com/commvault/v11/others/pdf/commvault-software-product-portfolio-feature-matrix.pdf
- Getting Started with NetApp Storage Array
  http://documentation.commvault.com/commvault/v11_sp18/article?p=33739.htm
- Commvault Sizing Best Practices
  http://documentation.commvault.com/commvault/v11_sp18/article?p=1644.htm
- Commvault Certifications and Compliance
  http://documentation.commvault.com/commvault/v11_sp18/article?p=110316.htm
- Commvault Security Best Practices
  http://documentation.commvault.com/commvault/v11_sp18/article?p=110316.htm
- Ransomware Protection
  http://documentation.commvault.com/commvault/v11_sp18/article?p=7877.htm
- Commvault Documentation
  http://documentation.commvault.com/commvault/v11_sp18/article?p=7877.htm

Version History

Version	Date	Document Version History
Version 1.0	April 2020	James Bradshaw, Jeannine Walter, Savita Kumari, Bryan Clarke, Nigel Tozer, Sarbjit Singh
Version 1.1	May 2020	Jeannine Walter

Refer to the [Interoperability Matrix Tool (IMT)](#) on the NetApp Support site to validate that the exact product and feature versions described in this document are supported for your specific environment. The NetApp IMT defines the product components and versions that can be used to construct configurations that are supported by NetApp. Specific results depend on each customer's installation in accordance with published specifications.

■ NetApp®