Technical Report

NetApp Converged Infrastructure with Commvault Complete Backup & Recovery

Alan Cowles, NetApp
May 2019 | TR-4772

Abstract

This document provides an overview of the Commvault Complete Backup & Recovery solution, its tight integration with NetApp® ONTAP® software, and NetApp converged infrastructure.

NetApp®

TABLE OF CONTENTS

1 Introduction to NetApp Converged Infrastructure 3

2 Introduction to Backup and Recovery 3

3 Introduction to Commvault Complete Backup & Recovery 4

4 Using Commvault to Protect Your Converged Infrastructure 4

4.1 Commvault Architecture 4

4.2 Sample Backup and Recovery Workflow 6

Where to Find Additional Information 7

Acknowledgments 8

1 Introduction to NetApp Converged Infrastructure

A common challenge in the IT infrastructure world is the complication that can arise from the subtle differences in the way products in different technology groups interact. It's often difficult to make sure that networking, compute, and storage systems will interoperate in a predictable way from one deployment to the next. To address this issue, various technology vendors have collaborated to develop converged infrastructure.

NetApp is one such company, and NetApp provides the storage backbone for these collaborative deployments. NetApp® storage systems support SAN-based boot for the converged deployment, allowing each server to be extremely dynamic. NetApp also provides data storage to the deployed host OS through several block-based and IP-based protocols. This storage supports end-user data access to both virtualized and bare-metal resources.

2 Introduction to Backup and Recovery

Backup is an overarching term used to describe data protection strategies for enterprise deployments. There are various methods for backing up data, whether by a manual copy or by an integrated mechanism in the hypervisor or storage system.

Many backup engines require a one-time complete backup of a resource, followed by additional incremental backups as data deviates from the baseline copy. The initial copying of data can take a long time, depending on the size of the source data. However, NetApp Snapshot™ technology is a patented system that can instantaneously create a copy of an entire NetApp storage volume.

The initial Snapshot copy is created rapidly as a metadata reference to blocks that exist in the current state of the NetApp WAFL® file system. As these blocks are modified or deleted, their retrospective place in the Snapshot copy is occupied by the information in the original block. Each additional Snapshot copy also references the current file system state and retains only the changes in data since the previous Snapshot copy. In the same manner, you can use the NetApp SnapRestore® feature to instantly restore a single file or an entire volume to a specific place in time as referenced by an existing Snapshot copy.

Creating instant Snapshot copies for volumes or datastores is a useful feature, especially for static or file data on all NetApp storage arrays that are running NetApp ONTAP software. However, there are a few caveats about creating file backups used to support virtual machines hosted by a converged infrastructure. These machines often run applications like databases and mail servers, and you must consider the behavior of these applications during a backup window. A backup is only as useful as the data that can be restored from it.

A point-in-time Snapshot copy of a virtual machine is the state of a machine at a particular moment. A Snapshot copy captures the contents of system memory, however inconsistent, and all actions currently taking place at that moment in time. This is called a crash-consistent Snapshot copy. Although it is effective for brute-force backup of system data, such a backup can result in the corruption or permanent loss of data at the OS or application level.

To correct this problem, specialized plug-ins or agents are available for virtual machines and specific applications that quiesce all operations for a moment and prepare them for backup. Use of these plug-ins or agents in VMware vCenter, NetApp SnapCenter®, or Commvault Complete Backup & Recovery software for NetApp Storage creates a pristine copy of the data for capture by the Snapshot process (an application-consistent Snapshot copy). These time points are ideal for virtual machine restoration if a disaster occurs, because an application-consistent backup makes sure that all applications and the machine itself are in an idle, stable state.

NetApp Snapshot copies provide the first level of defense against data loss. However, they cannot protect data against a sitewide disaster. Often, enterprise-level deployments have both a primary site and a disaster recovery site. Both sites are online and in sync to meet defined recovery point objectives (RPOs)

and recovery time objectives (RTOs) if a disaster occurs. Meeting RPOs and RTOs became much easier with the advent of converged infrastructure, because identical stacks could be scoped and deployed at multiple sites more easily.

Emulating infrastructure at another site is ideal for the restoration of operations, but you do need dependable copies of data available at the mirrored location as well. Therefore, you can create an additional layer of protection by replicating data to the remote site through NetApp SnapMirror® technology. This feature creates either synchronous or asynchronous copies of data at remote sites. Therefore, SnapMirror protects enterprise operations against a loss of access or even the complete loss of data at a site, and it allows you to resume standard operations within your defined RPO and RTO.

In a manner analogous to the local Snapshot copy process, SnapMirror takes a baseline Snapshot copy and transfers this baseline when the site-mirroring relationship is initialized. The ONTAP system can run regular updates to the mirror to make sure that copies of the data are available at both sites. The system can be configured in a synchronous manner to create identical datastores at both sites.

To meet RPOs, you can manage manual replication operations with tools such as NetApp SnapCenter and Commvault Complete Backup & Recovery. These tools place virtual machines in an application-consistent state before Snapshot copies are created locally and mirrored. During a disaster, you can elevate the remote site to the primary site to meet your defined RTO. After normal operations have resumed at the primary site, you can use SnapMirror resynchronization to restore the relationship and copy any changed data back to the primary site.

3 Introduction to Commvault Complete Backup & Recovery

Commvault provides solutions that ease all aspects of data management for customers of enterprise storage solutions, including backup, restore, archiving, and replication. The latest release of its flagship product, Commvault Complete Backup & Recovery, was introduced in July 2018 and greatly simplifies each of these functions. It works seamlessly across many converged infrastructures, including infrastructures that depend on NetApp storage systems.

Commvault and NetApp have partnered in the past to deliver the NetApp SnapProtect® and Commvault IntelliSnap software solutions. These solutions use Commvault's native UI and the patented NetApp Snapshot and SnapMirror technologies to perform backup and replication operations. Furthermore, these companies entered into a cooperative reseller agreement in October 2018. This agreement enables NetApp to resell Commvault solutions as part of the NetApp ONTAP data protection solution for both storage systems and converged infrastructures.

4 Using Commvault to Protect Your Converged Infrastructure

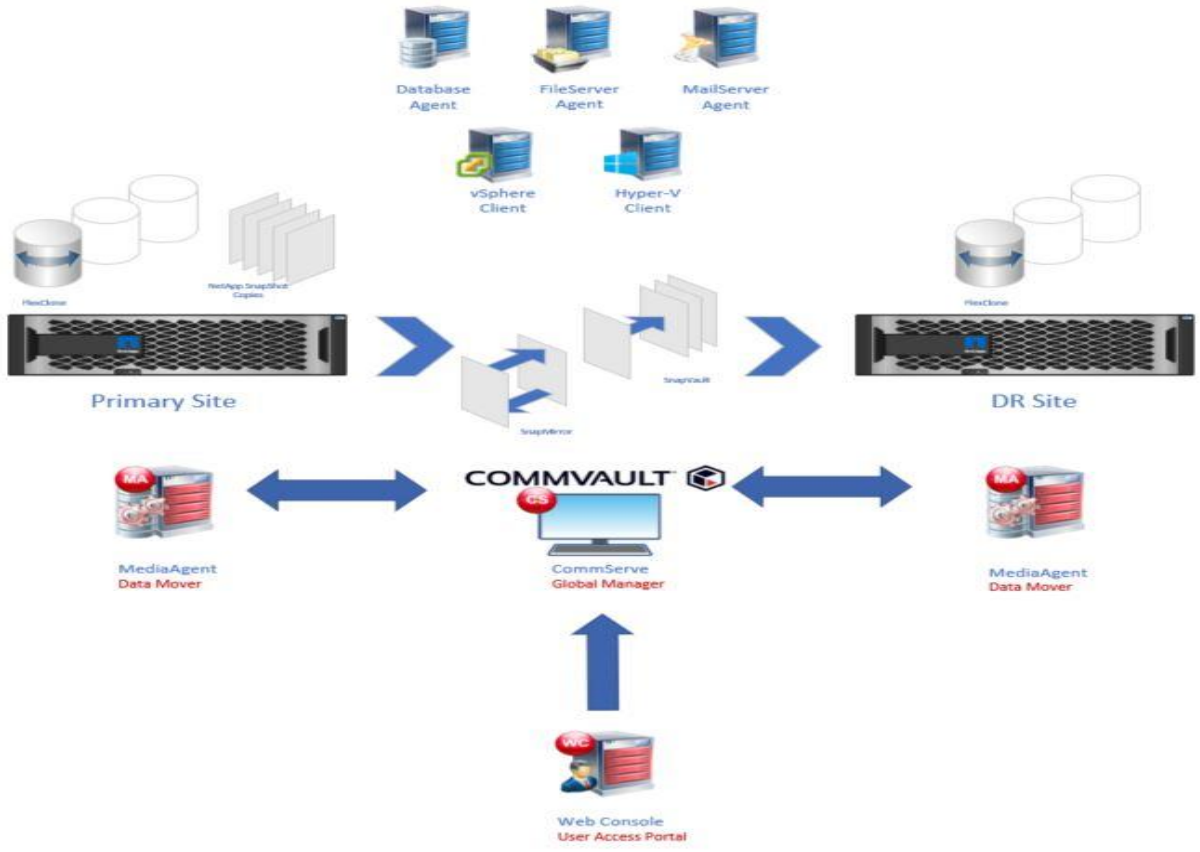
The following section reviews the CommVault Complete Backup & Recovery solution as it pertains to NetApp converged infrastructure solutions.

4.1 Commvault Architecture

A Commvault Complete Backup & Recovery system consists of the following major components (Figure 1):

- CommServe Server
- MediaAgent
- CommCell Console
- WebServer and WebConsole
- Clients and agents
- NetApp ONTAP integration

Figure 1) Commvault architecture.



CommServe Server

The CommServe host is a physical or virtual system in an environment that acts as the central management engine for the CommCell architecture. A CommVault Complete Backup & Recovery deployment allows the creation of only a single CommServe server in a protected environment. A CommServe server hosts an internal MySQL database that contains all necessary information for the protection of the converged infrastructure, including all configuration, security, and operational history for the environment.

MediaAgent

The MediaAgent manages the data storage libraries and oversees all data transmission operations in the protected environment. Like the CommServe server, it can be deployed in the converged infrastructure as a physical or virtual host. However, unlike the CommServe server, you can deploy multiple MediaAgent servers to further scale operations and provide high-performance data movement.

CommCell Console

The CommCell Console is a simple UI for the centralized management of all operations in the protected environment. It enables administrators to monitor and control all active jobs, and it enables users to review all activities and event logs.

WebServer and WebConsole

WebServer and WebConsole are installed by default along with the CommServe server. They provide front-end and back-end functionality for management of the protected environment. These components allow end users to remotely manage their data through an HTML5-based web interface known as the Command Center. Through this interface, users can create reports, modify backup jobs and schedules, and manage the virtual machines deployed in the converged infrastructure.

Clients and Agents

Clients are the targets of backup operations in the CommCell environment, whether they are physical or virtual entities deployed in the converged infrastructure. Installing agents on each client provides advanced backup functionality. Some agents provide backup and recovery for simple file systems, whereas other agents enable application-aware backups to create consistent restore points. Another class of agents enables database backup and recovery by quiescing database operations to create a stable state before a new backup copy is taken.

To define the client and type of backup to be created, an end user selects an agent. The user can then further refine the backup operation by defining backup sets, subclients, and storage policies for each data protection action. This step might be needed because a single application can have multiple dependencies and data types when deployed in a converged infrastructure environment. All of this data should be backed up in tandem to provide a consistent copy from which to restore the application if a disaster occurs.

ONTAP Integration

Commvault Complete Backup & Recovery integrates with both modern and legacy ONTAP systems for the management of data protection workloads. For local protection operations, a CommCell can integrate directly with NetApp FAS and AFF systems to provide NetApp FlexClone®, Snapshot, and SnapRestore functionality. Before ONTAP 8.3.2, NetApp OnCommand® Unified Manager had to be deployed in the environment to manage NetApp SnapMirror and SnapVault relationships between primary and disaster recovery sites. Although you can still use OnCommand Unified Manager to manage these relationships, in more recent ONTAP versions, you can manage these functions with NetApp Open Replication if the mirroring relationships between ONTAP systems have been created in advance.

4.2 Sample Backup and Recovery Workflow

A sample workflow for the backup and recovery of an application in your converged infrastructure consists of the following steps:

- Verify connectivity with the storage systems.
- Identify the system or application.
- Identify the RTO and RPO.
- Create mirror and vault relationships.
- Create subclients, backup sets, and storage policies.
- Validate the data protection solution.

Verify Connectivity with Storage Systems

To manage a NetApp ONTAP system as a part of your converged infrastructure, you must verify that it is loaded into the CommCell as a storage array. If you also want to manage a storage system as a disaster recovery replication target, you must add that system separately. If you are configuring ONTAP systems earlier than 8.3.2, you must also deploy and configure OnCommand Unified Manager to manage SnapMirror and SnapVault relationships. You must also make the CommCell aware of OnCommand

Unified Manager deployment. For any deployments that use NetApp Open Replication, you do not need to use OnCommand Unified Manager. However, you must manually create SnapMirror and SnapVault relationships with ONTAP.

Identify the System or Application

As mentioned earlier, you can create a data protection relationship for a virtual guest at the hypervisor or OS level by installing that host as a client on a CommCell. You can create a simple client-based backup by quiescing the guest OS and making a Snapshot copy at that moment. However, the Snapshot copy might not provide the granularity needed for individual systems that are running applications such as databases or mail servers. For these systems, you must install an agent on the host to facilitate an intelligent backup operation that quiesces application operations and enables you to create a consistent Snapshot copy.

Identify the RPO and RTO

The type of backup needed depends on the RPO and RTO defined by your organization. If a dataset must be restored to a particular point in time before a data corruption event, you must verify that your backup and retention policies support those requirements. Local Snapshot copies might not meet this requirement if the site becomes unavailable or if the recovery point is beyond the current Snapshot retention window. Using a SnapVault relationship on a remote converged infrastructure to keep archived copies can help meet this requirement. However, keep in mind the time it takes to restore data from a remote archive to an active running state.

Create Mirror and Vault Relationships

When using NetApp Open Replication, you must manually create the SnapMirror or SnapVault relationship for each converged infrastructure volume that you want to protect. In ONTAP 9.3 and later, NetApp supports SnapMirror unified replication, in which all relationship types are extended data protection (XDP) and the default retention policy is MirrorAndVault.

Create Subclients, Backup Sets, and Storage Policies

Within the CommCell for each client, you must decide what data is backed up and how often. You can refine this process for each client by creating subclients, which are logical containers that identify specific data to back up. Multiple subclients can be grouped into a backup set, which defines the protection mode (backup, archive, or replication) and makes sure that all subclients are backed up in a consistent manner. The storage policy enables you to determine the lifecycle of the data being backed up so that you have the correct number of copies and replicas to meet your RPO and RTO.

Validate the Data Protection Solution

After a data protection relationship is established between your primary and disaster recovery data centers, you can validate the data in your local converged infrastructure or at the disaster recovery site. Features like NetApp SnapRestore then allow you to recover data incrementally from a Snapshot copy. In addition, if your converged infrastructure is hosting test or development infrastructure, you can use NetApp FlexClone to create a writable clone of a client. You can then use it to validate code or software upgrades without affecting your production environment.

Where to Find Additional Information

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

- Commvault and NetApp Expand Partnership to Offer Powerfully Simple Backup and Recovery Solutions from NetApp and NetApp Channel Partners https://www.commvault.com/news/2018/october/commvault-and-netapp-expand-partnership-to-offer-powerfully-simple-backup
- Commvault Complete Backup & Recovery Optimized for NetApp Storage Systems https://cloud.kapostcontent.net/pub/c3643411-a9d3-4423-b7cc-acb0e1e716ea/netapp-solution-brief-for-commvault-complete-backup-and-recovery?kui=quYSx6dN9QhIKldB9eNyBg
- Commvault Documentation Center: Getting Started with the NetApp Storage Array http://documentation.commvault.com/commvault/v11_sp15/article?p=33779.htm

Acknowledgments

I would like to give special thanks to Lindsey Street and Ankita Dhawale for their assistance in the preparation, validation, and editing of this document.

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.

Copyright Information

Trademark Information