Technical Report

NetApp Converged Infrastructure with Veeam Backup & Replication

Alan Cowles, NetApp
May 2019 | TR-4773

Abstract

This document provides an overview of the Veeam Backup & Replication solution and it's tight integration with NetApp® ONTAP® and the NetApp converged infrastructure.

NetApp®

TABLE OF CONTENTS

1 Introduction to NetApp Converged Infrastructure

A common challenge in the IT infrastructure world is the complications that can arise from the subtle differences in the way individual products in different technology groups interact with each other. It's often difficult to make sure that networking, compute, and storage systems are all going 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 back-bone 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 to support end user data access to both virtualized and bare metal resources.

2 Backup Methods and Terminologies

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 over time 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 Snapshot 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 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 taken also references the current file system state and only retains 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, that is available on all NetApp storage arrays running ONTAP. However, there are a few caveats when creating file backups used to support virtual machines (VMs) hosted by a converged infrastructure. These machines often run specific applications like databases and mail servers, and you must consider the behavior of these applications during a backup window when creating a Snapshot copy. A backup is only as useful as the data that can be restored from it.

A point-in-time Snapshot copy of a VM created by NetApp Snapshot technology is the state of a machine at a particular moment. It 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. 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 runtimes are available for VMs and specific applications that quiesce all operations for a moment and prepare them for backup. Use of these plug-ins or runtimes within VMware vCenter, NetApp SnapCenter®, or Veeam Backup & Replication creates a pristine copy of the data for capture by the Snapshot process (an application-consistent Snapshot copy). These time points are ideal for VM restoration should a disaster occur 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 in the event of 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 (RPO) and recovery time objectives (RTO) should a disaster occur. 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, an additional layer of protection is created by replicating data to the remote site using NetApp SnapMirror® technology. This feature creates either synchronous and 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 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 data stores at both sites.

To meet RPOs, you can manage manual replication operations with tools such as NetApp SnapCenter and Veeam Backup & Replication. These tools place VMs in the environment 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 original primary site.

3 Introduction to Veeam Replication and Recovery

Veeam has developed and sells the Veeam Backup & Replication application for virtual and physical environments. This proprietary application is available with a trialware license to customers that would like to perform full or incremental backups and restoration of the VMs in their environment.

It currently supports two primary hypervisors: Microsoft's Hyper-V and VMware's vSphere, which are both validated hypervisors in the NetApp converged infrastructure portfolio. In addition to using virtual snapshot technology, Veeam Backup & Replication has specialized integrations that allow it to interact directly with NetApp Snapshot technology in vSphere environments, to create storage snapshots. It can also access these storage snapshots directly with a built in NFS agent. The ability to mount NAS volumes from a NetApp controller and access snapshots directly allows Veeam Backup & Recovery to bypass individual hosts when performing backup, restore, and scanning operations.

As was first announced at NetApp Insight in October 2017, NetApp and Veeam have entered into a cooperative reseller agreement in March of 2018. This agreement allows NetApp to resell the complete Veeam Backup & Replication solution to NetApp converged infrastructure customers that would like to backup and restore virtual and physical machines in their environment.

4 Using Veeam to Protect Converged Infrastructure

This section provides a review of the Veeam Backup & Replication solution as it pertains to NetApp converged infrastructure solutions.

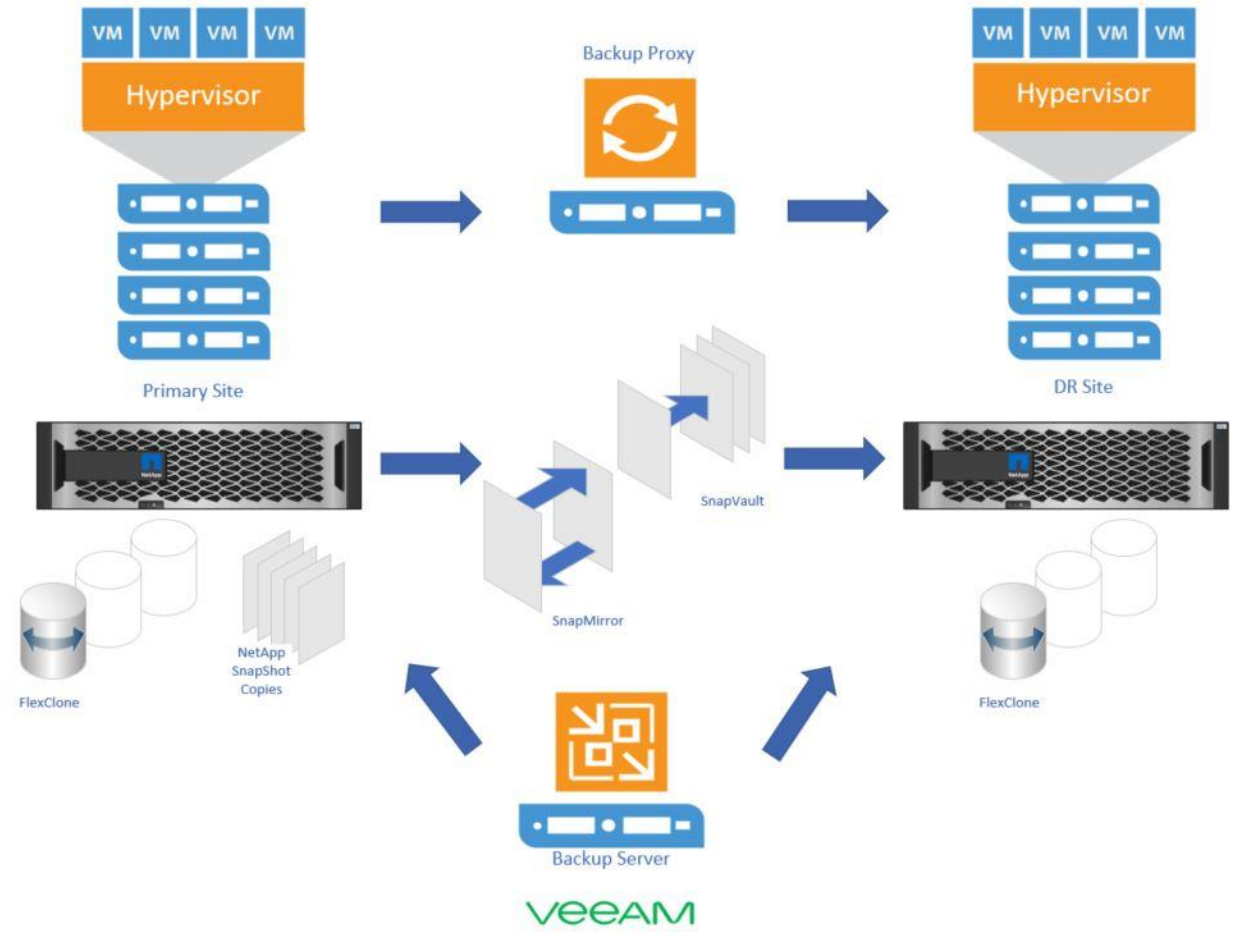
4.1 Veeam Architecture

The architecture of a Veeam Backup & Replication deployment consists of the following major components:

- Backup Server
- Backup & Replication Console
- Backup Proxy
- Runtime coordination process

- ONTAP integration

Figure 1) Veeam architecture.



Backup Server

This is the Microsoft Windows-based host where Veeam Backup & Replication is installed. The host can be physical or virtual, and it is hosted within the NetApp Converged Infrastructure. This server provides core functionality for the Veeam deployment and performs most administrative actions including coordinating all data protection activities (backup, replication, recovery verification, and restoration of protected systems). It controls all job scheduling, it is used to configure and manage all backup components in the infrastructure, and it can modify global settings.

Backup and Replication Console

The backup and replication console is the client-facing portion of a Veeam Backup and Replication installation. It is installed on the same server hosting the Backup Server by default, but it can also be installed in standalone mode on it's own server for a distributed deployment.

A large organization might want to deploy multiple remote consoles to ease management of the data protection environment. However, the Veeam Backup Server only permits a single active session in which modifications can be made. This can lead to an issue if multiple storage administrators are working at the same time. When the most recent save of configuration is committed, all other logged-in users are encouraged to refresh their session to see the updated configuration. If a user is disconnected from the

remote console before a change can be committed, the existing session can be resumed for a period of five minutes. Otherwise, the user is asked to log in again.

Backup Proxy

Backup proxy servers are installed to distribute the load of backup and recovery operations. When Veeam Backup & Replication is first installed to a converged infrastructure, all job activities and data movement passes through the backup server. Passing all data-protection traffic through the backup server in a large-scale environment can have adverse effects on the server itself. Backup proxies can be installed on dedicated physical or virtual Microsoft Windows servers and made available at either the primary or disaster recovery datacenter to move data between sites.

Runtime Coordination Process

Veeam's backup for virtual guests is an agentless process that uses a runtime process that is not persistent on VMs. If a backup job must be application-aware, the guest filesystem of a VM must be indexed, or transation logs must be processed before backup completion, runtime is enabled for that backup operation. After the backup operation is complete, runtime is removed. If the guest being backed up is a Microsoft Windows VM, a guest interaction proxy must be deployed in the environment. All other guest operating systems have their runtime functionality managed by the backup server.

ONTAP Integration

Integration between Veeam Backup & Replication and a NetApp Converged Infrastructure environment is managed with a wizard in the Storage Infrastructure view. You must provide the credentials required to access your cluster, and select the volumes that contain your VMware datastores. Volume selection can be performed automatically by having Veeam scan the system for VMware Virtual Machine File System (VMFS) volumes, or it can be selected individually from the wizard.

In a converged infrastructure environment using SnapMirror or NetApp SnapVault® technology for data protection, these relationships must be established outside of the Veeam console. The SAN or network must be configured correctly to allow the backup proxy to access the storage system. The backup proxy requires no special configuration to access NFS or ISCSI datastores. However if you are using Fiber Channel datastores, then you should create and configure a dedicated igroup to include the WWN ID of the backup proxy servers at the site.

For direct NFS access to datastore contents, you must create an export for Veeam to use it's mounting functionality to load virtual guests for testing, development, or restoration.

4.2 Sample Backup and Recovery Workflow

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

- Verify connectivity with storage systems
- Identify VM datastores
- Identify the RTO and RPO objectives
- Licensing the appropriate ONTAP features
- Create mirror and vault relationships
- Validate the data protection solution

Verify Connnectivity With Storage Systems

To manage a NetApp ONTAP system as a part of your converged infrastructure, you must verify that it is discovered by Veeam Backup & Replication. This process starts by performing an initial system scan which identifies all cluster members, networking information, volumes, LUNS, and existing NetApp Snapshot copies. Veeam Backup & Replication then creates a list of all of the snapshots for each volume and a list of all LUNs and NFS exports. It then validates whether or not the backup proxies can communicate with the storage system. It does this by validating the protocol licenses enabled on the storage system, and it confirms whether the respective servers are running within the ONTAP environment. The backup proxies then test each protocol—ISCSI, NFS, or FC—that it is planning to use to access storage data for protection and verifies that the connection is available.

Identify the Virtual Machine Datastores

Veeam Backup & Replication then scans each ESXi host in the converged infrastructure that has been added to its configuration. This scan validates each datastore that is exported from the ONTAP storage ,system and catalogs each VM present. A Veeam storage snapshot is created for each VM found. The storage snapshot is then used to create a rough list of the VMs that Veeam is responsible for protecting.

This list is validated by mounting the newly created storage snapshots and verifying their contents against those of the datastore exports from the ONTAP system. After this, Veeam purges any orphaned snapshots, that is to say, snapshots that are not locked as a part of a snapshot chain and listed within the database.

Identify the RPO and RTO Objectives

The type of backup taken depends on the RPO and RTO, as defined by the organization. If a VM must be restored to a point in time before a data corruption event, you must verify that your backup and retention policies account for those requirements when you perform your backups. Local NetApp 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, but you need to be aware of the time it takes to restore data from a remote archive to an active running state.

Licensing the Appropriate ONTAP Features

Veeam Backup & Replication requires that ONTAP has available entitlements for SnapRestore and NetApp FlexClone® technology enabled on the storage system. Without these features licensed and enabled, the core functionality of Veeam will not function as intended. It uses SnapRestore to perform single-file restores from NetApp Snapshot copies. In addition, it uses FlexClone to access a block protocol datastore by creating a thin clone of an existing LUN from a base snapshot and mounting it to an ESXi host in the converged infrastructure for access.

Create Mirror and Vault Relationships

When using Veeam Backup & Replication to protect VM datastores in your converged infrastructure thought remote replication, you must manually create SnapMirror or SnapVault relationships for each volume that you wish to protect. In ONTAP 9.3 and later, NetApp supports SnapMirror Unified Replication, in which all relationship types are XDP (extended data protection), and the default retention policy is MirrorAndVault. This configuration allows you to keep a synchronized mirror at your disaster recovery site and also to keep a number of legacy NetApp Snapshot copies to restore VMs from should a disaster occur at your primary site.

Validate the Data Protection Solution

After a data protection relationship is established between your primary and disaster recovery datacenters, 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 writeable 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/or websites:

- Veeam and NetApp Take Relationship to New Level with Resell Agreement https://www.veeam.com/news/veeam-and-netapp-take-relationship-to-new-level-with-resell-agreement.html
- Veeam + NetApp – Delivering the Always-On Next-Gen Data Center https://www.veeam.com/executive-blog/netapp-next-generation-data-center-solution.html

Information about Veeam and Netapp Integration:

- Configuration Guide and Best Practices for NetApp and Veeam Backup & Replication 9.5 https://www.veeam.com/netapp-configuration-best-practices-guide_wpp.pdf

Acknowledgments

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