Technical Report

Red Hat Enterprise Virtualization 3.1 and NetApp Clustered Data ONTAP Storage

Best Practices Guide

Jon Benedict, NetApp
April 2013 | TR-4159

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

1.1 Overview

This technical report describes the current best practices for deploying Red Hat® Enterprise Virtualization (RHEV) 3.1 and NetApp® clustered Data ONTAP® 8.1. NetApp recommends that the reader browse through the document to get an understanding of each chapter and then reference this document as needed for specific requirements.

1.2 Intended Audience

This document addresses the needs of system architects, system administrators, and storage administrators who are investigating the use of or are deploying RHEV on clustered Data ONTAP. This document should be used in conjunction with officially branded documents available at the Red Hat Customer Portal and the NetApp Support site.

1.3 Best Practices

Best practices offer the optimal balance of desirable features that have been shown to produce superior results; that are selected by a systematic process; and that are judged as exemplary, good, and/or successfully demonstrated. A balance is struck based on gaining optimal performance, reliability, and simplicity without sacrificing those characteristics that are targeted for improvement.

When incorporated in all areas of an organization, and particularly in the data center, the use of best practices can lead to world-class performance, create repeatability, and have positive effects on both operational and capital costs. Data center staff often use more than one best practice, but unless best practices are adopted consistently across all functions of the data center, the highest levels of performance remain out of reach.

1.4 Scope of This Document

The objectives of this document are to:

- Describe the best practices for deploying RHEV and clustered Data ONTAP together
- Illustrate the benefits of deploying these technologies together

The scope is limited to where Red Hat and NetApp technologies intersect or directly affect each other in the context of RHEV 3 and clustered Data ONTAP 8.1. Where appropriate, configuration examples are provided.

1.5 Topics Out of Scope

Both Red Hat and NetApp have extensive product documentation describing all means of deploying their respective products. This document does not attempt to duplicate product documentation. It assumes that the reader has an in-depth understanding of Red Hat Enterprise Linux®, Red Hat Enterprise Virtualization, and NetApp storage.

Sections of this document cover best practices for networking technologies, but they do not attempt to provide instructions specific to any particular brand or model of network gear.

1.6 Why Virtualize Storage?

The virtualization of a data center results in many physical systems being virtualized as part of a cost-saving effort to reduce capital expenses (capex) and operating expenses (opex) through infrastructure consolidation and increased operational efficiencies. However, this trend of server virtualization and

consolidation is occurring at the same time as a data explosion. By 2020, IDC expects business data to increase 44 times, to 35 zettabytes (1 zettabyte = 10^{21} bytes).

There is no doubt that virtualization had to happen; there was no other way to meet the demands of consolidating servers and modernizing applications. But virtualization has to happen everywhere in the data center to meet the demands of scalability in performance, capacity, and operations. Businesses demand this type of agility from their data centers to move the business forward. The scale, complexity, and pace of business put a premium on agility.

"Business today is putting enormous pressure on IT to respond almost instantly to very challenging demands like never before. The result is an urgency to think differently and to better allocate resources—both financial and operational—so that IT can quickly adapt to and drive change, creating value. Organizations will use agile data infrastructure to fundamentally rethink how they architect and manage their data storage and ultimately to advance and accelerate their business objectives," says Steve Duplessie, founder, Enterprise Strategy Group.

Server virtualization and network virtualization are still clear requirements, but they cover only two-thirds of the data center. The original core requirements that demand virtualization are still true today; applications must be able to be run more quickly and inexpensively while maintaining security, performance, and manageability. These constraints and requirements are not limited to the applications and servers. Any honest discussion about application modernization must include enterprise storage considerations to be complete and holistic in scope and to meet the demands of the data explosion.

1.7 Introducing Clustered Data ONTAP 8.1 and Red Hat Enterprise Virtualization 3.1

Clustered Data ONTAP 8.1 is the next-generation storage platform to deploy modern business-critical applications, both virtual and bare metal. NetApp has long offered the ability to efficiently scale up storage in modular increments. The introduction of clustered Data ONTAP provides the ability to scale out performance for both SAN and NAS environments, while maintaining NetApp's storage efficiencies, unified architecture, and built-in data protection technologies.

Red Hat Enterprise Virtualization 3.1 provides a high-performing, secure, and eminently manageable virtualization platform on which to deploy business-critical applications. RHEV 3.1 includes a secure web-based portal (Red Hat Enterprise Virtualization Manager, RHEV-M) to manage the virtual environment as well as a thin, small-footprint hypervisor (Red Hat Enterprise Virtualization Hypervisor, RHEV-H) that is built around the KVM hypervisor. Additionally, RHEV-M is able to manage thick hypervisors (RHEL 6) in conjunction with RHEV-H. Because of the architecture of the KVM hypervisor, virtualized applications run at near-native speeds, while gaining mobility, increased manageability, and security. Red Hat makes no distinction between native and virtualized applications.

This supportability extends to NetApp's Interoperability Matrix Tool (IMT), which offers in-depth testing and qualifications that span the entire stack. More importantly, Red Hat and NetApp extend this high level of interoperability into joint solutions and best practices, as shown by this document.

What does this mean for deploying virtualized applications on RHEV3.1 and clustered Data ONTAP? It means that the data center can literally scale on demand without losing flexibility, security, performance, or manageability. Virtual machines can be deployed in a highly accelerated manner from the storage array. Applications can operate within secure tenants that span the entire hardware, network, and storage stack. Major components of the storage infrastructure can be replaced or upgraded nondisruptively.

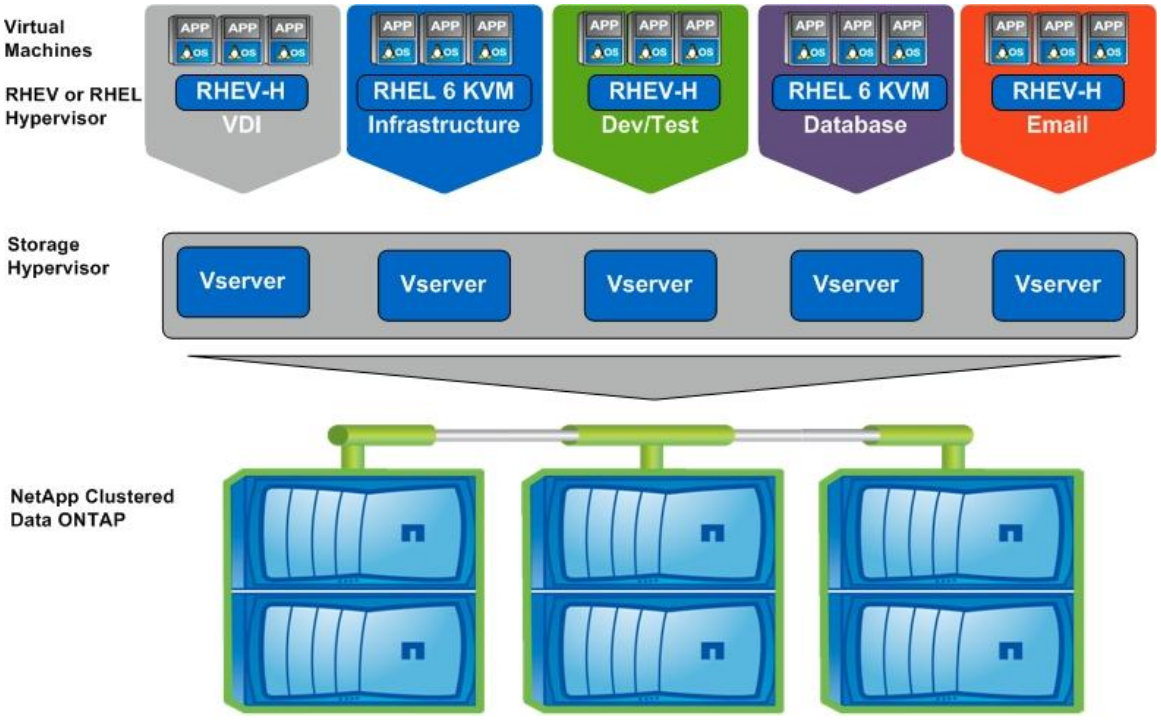
Ultimately, this supportability provides the foundation and confidence required to deploy an agile infrastructure.

Guiding Principles for Clusters

All clustering technologies follow a common set of guiding principles, which include the following:

- **Nondisruptive operations.** The key to efficiency and the linchpin of clustering is the ability to make sure that the cluster does not ever fail.
- **Virtualized access is the managed entity.** Direct interaction with the nodes that make up the cluster is in and of itself a violation of the term “cluster.” During the initial configuration of the cluster, direct node access is necessary; however, steady-state operations are abstracted from the nodes as the user interacts with the cluster as a single entity.
- **Data mobility and container transparency.** The end result of clustering—that is, the nondisruptive collection of independent nodes working together and presented as one holistic solution—is the ability of data to move freely within the boundaries of the cluster.
- **Delegated management and ubiquitous access.** In large complex clusters, the ability to delegate or segment features and functions into containers that can be acted upon independently of the cluster means that the workload can be isolated. Equally important is that the cluster must not place conditions on how its contents are accessed. This should not be confused with security concerns about the content being accessed.
- **Clustered Data ONTAP embodies these guiding principles.** Figure 1 shows how access is virtualized with NetApp clustered storage using virtual storage servers (Vservers), ubiquitous access through multiprotocol support, and the ability to move data within the cluster depending on workload needs and intracluster relationships.

Figure 1) NetApp clustered storage overview.



2 Clustered Data ONTAP Overview

Scaling performance while controlling costs is one of the most challenging efforts in the data center. High performance, technical computing, and digital media content applications place extreme demands on

storage systems. Compute clusters running these applications require multiple gigabytes per second of performance and many terabytes—or even petabytes—of capacity. To maintain peak application performance, users must be able to add storage and move data between systems and tiers of storage without disrupting ongoing operations. At the same time, to control costs, users must be able to effectively manage the storage environment.

Clustered Data ONTAP addresses these challenges and meets high-performance and high-capacity requirements. It enables organizations to achieve faster time to market by providing massive throughput and the scalability necessary to meet the demanding requirements of high-performance computing and virtualization infrastructures. These high performance levels address the growing demands of performance, manageability, and reliability for both virtualized and nonvirtualized workloads.

The clustered Data ONTAP operating system from NetApp includes:

- Nondisruptive operations based on a clustered file system hosted on interconnected nodes
- Multinode scaling with global namespacing technologies
- NetApp FlexVol® volumes for storage virtualization
- NetApp backup and recovery solutions based on local Snapshot™ copies, replication, and mirroring

2.1 Clustered Data ONTAP 8.1

Overview

With the release of clustered Data ONTAP 8.1, NetApp introduces enterprise-ready, unified scale-out storage. Developed from a solid foundation of proven Data ONTAP technology and innovation, clustered Data ONTAP is the basis for large virtualized shared storage infrastructures that are architected for nondisruptive operations over the lifetime of the system.

Scale-Out

Scale-out is a method of responding to growth in a storage environment. All storage controllers have physical limits to their expandability, such as number of CPUs, memory slots, and space for disk shelves that dictate the maximum capacity and controller performance. If more storage or performance capacity is needed, it may be possible to add CPUs and memory or to install additional disk shelves, but ultimately the controller becomes completely populated, with no further expansion possible. At this stage, the only option is to acquire another controller. One way to do this is to scale up; that is, to add controllers in such a way that each is a completely independent managed entity that does not provide any shared storage resources. If the original controller is to be completely replaced by the newer and larger controller, data migration is required to transfer the data from the old controller to the new one. This is time consuming, potentially disruptive, and usually requires configuration changes on all of the attached host systems.

If the newer controller can coexist with the original controller, there are now two storage controllers to be individually managed, and there are no native tools to balance or reassign workloads across them. The situation worsens as the number of controllers increases. If the scale-up approach is used, the operational burden increases consistently as the environment grows, and the result is an unbalanced and difficult-to-manage environment. Technology refresh cycles require substantial planning in advance, lengthy outages, and configuration changes, which introduce risk into the system.

By contrast, using scale-out means that as the storage environment grows, controllers are added seamlessly to the resource pool residing on a shared storage infrastructure. Host and client connections as well as datastores can move seamlessly and nondisruptively anywhere in the resource pool, so that existing workloads can be easily balanced over the available resources and new workloads can be easily deployed. Technology refreshes (replacing disk shelves, adding or completely replacing storage controllers) are accomplished while the environment remains online and is serving data.

Although scale-out products have been available for some time, they are typically subject to one or more of the following limitations:

- **Limited protocol support.** Support NAS only
- **Limited hardware support.** Support only a particular type of storage controller or a very limited set
- **Little or no storage efficiency.** No thin provisioning, deduplication, or compression
- **Little or no data replication capability**

Therefore, although these products are well positioned for certain specialized workloads, they are less flexible, capable, and robust for broad deployment throughout the enterprise.

Data ONTAP is the first product to offer a complete scale-out solution, and it offers an adaptable, always-available storage infrastructure for today's highly virtualized environments.

Multiprotocol Unified Architecture

Multiprotocol unified architecture is the ability to support multiple data access protocols concurrently in the same overall storage system, over a whole range of different controller and disk storage types. Data ONTAP 7G and Data ONTAP 8 operating in 7-Mode have been capable of this for a long time, and now clustered Data ONTAP 8.1 also supports a full range of data access protocols. The following data access protocols are supported:

- NFS v3, v4, and v4.1, including pNFS
- SMB 1 and 2
- iSCSI
- Fibre Channel
- FCoE

Data replication and storage efficiency features are seamlessly supported across all protocols in clustered Data ONTAP.

2.2 SAN Data Services

With the supported SAN protocols (Fibre Channel, FCoE, and iSCSI), Data ONTAP provides LUN services; that is, the ability to create LUNs and make them available to attached hosts. Because the cluster consists of multiple controllers, there are multiple logical paths to individual LUNs, and the best practice is to configure at least one path per node in the cluster. Asymmetric logical unit access (ALUA) is used on the hosts to make sure that the optimized path to a LUN is selected and made active for data transfer.

2.3 NAS Data Services

With the supported NAS protocols, CIFS (SMB) and NFS, Data ONTAP can provide a single namespace; that is, NAS clients can access a very large data container by using a single NFS mountpoint or CIFS share. Each client needs to mount only a single NFS file system mountpoint or access a single CIFS share, and only the standard NFS and CIFS client code for each operating system is required. In Data ONTAP, the namespace is composed of potentially thousands of volumes junctioned together by the cluster administrator. To the NAS clients, each volume appears as a folder or subdirectory, nested off the root of the NFS file system mountpoint or CIFS share. Volumes can be added at any time and are available to the clients immediately, with no remount required for visibility to the new storage. Furthermore, the clients are neither notified nor affected by the fact that they are traversing volume boundaries as they move about in the file system, since the underlying structure is completely transparent. Although Data ONTAP can be architected to provide a single namespace, it also supports the concept of multiple, securely partitioned namespaces to accommodate requirements for multi-tenancy

or isolation of particular sets of clients or applications. For more information, refer to section 2.8, “Logical Cluster Components.”

2.4 Nondisruptive Operations

In today’s 24/7 environments, shared storage infrastructures provide data services to thousands of individual clients or hosts and support many diverse applications and workloads across multiple business units or tenants. In such environments, downtime is not an option; storage infrastructures must always be on.

Nondisruptive operations (NDO) in Data ONTAP are intrinsic to its innovative scale-out architecture. NDO refers to the ability of the storage infrastructure to remain up and serve data through the execution of hardware and software maintenance operations, as well as during other IT lifecycle operations. The goal of NDO is to eliminate downtime, whether preventable, planned, or unplanned, and to allow changes to occur in the system at any time.

Data ONTAP is highly available by design and can transparently migrate data and network connections anywhere within the storage cluster. The ability to move individual data volumes by using NetApp DataMotion™ for volumes allows data to be redistributed across a cluster at any time and for any reason. DataMotion is transparent and nondisruptive to NAS and SAN hosts and enables the storage infrastructure to continue to serve data throughout these changes. DataMotion migration can be performed to rebalance capacity usage, to optimize for changing performance requirements, or to isolate one or more controllers or storage components in order to execute maintenance or lifecycle operations.

Table 1 lists the hardware and software maintenance operations that can be performed nondisruptively in a Data ONTAP 8.1 environment.

Table 1) Nondisruptive hardware and software maintenance operations.

Operation	Details
Upgrade software	Upgrade from one version of Data ONTAP to another
Upgrade firmware	Upgrade system, disk, and switch firmware
Replace a failed controller or a component within a controller	For example, replace network interface cards (NICs), host bus adapters (HBAs), power supplies, and so on
Replace failed storage components.	For example, replace cables, drives, I/O modules, and so on

Table 2 lists the lifecycle operations that can be performed nondisruptively in a Data ONTAP 8.1 environment.

Table 2) Nondisruptive lifecycle operations.

Operation	Details
Scale storage	Add storage (shelves or controllers) to a cluster and redistribute volumes for future growth
Scale hardware	Add hardware to controllers to increase scalability, performance, or capability (HBAs, NICs, NetApp Flash Cache™, NetApp Flash Pool™)
Refresh technology	Upgrade storage shelves, storage controllers, back-end switch

Operation	Details
Rebalance controller performance and storage utilization	Redistribute data across controllers to improve performance
Rebalance capacity	Redistribute data across controllers to account for future capacity growth
Rebalance disk performance and utilization.	Redistribute data across storage tiers within a cluster to optimize disk performance

2.5 Immortal Cluster

Clustered Data ONTAP 8.1 can deliver an effectively immortal cluster. In most environments, over time, the hardware infrastructure might be augmented and replaced many times, in addition to the typical software updates and configuration changes through the lifecycle of the system. Many years after the system was originally commissioned and the data has outlived the hardware, little or none of the original hardware may remain. Through nondisruptive operations, all of these changes would have been achieved with no outage, and without any effect on any of the applications or attached clients and hosts; the cluster entity would have persisted intact.

2.6 Data ONTAP Architecture

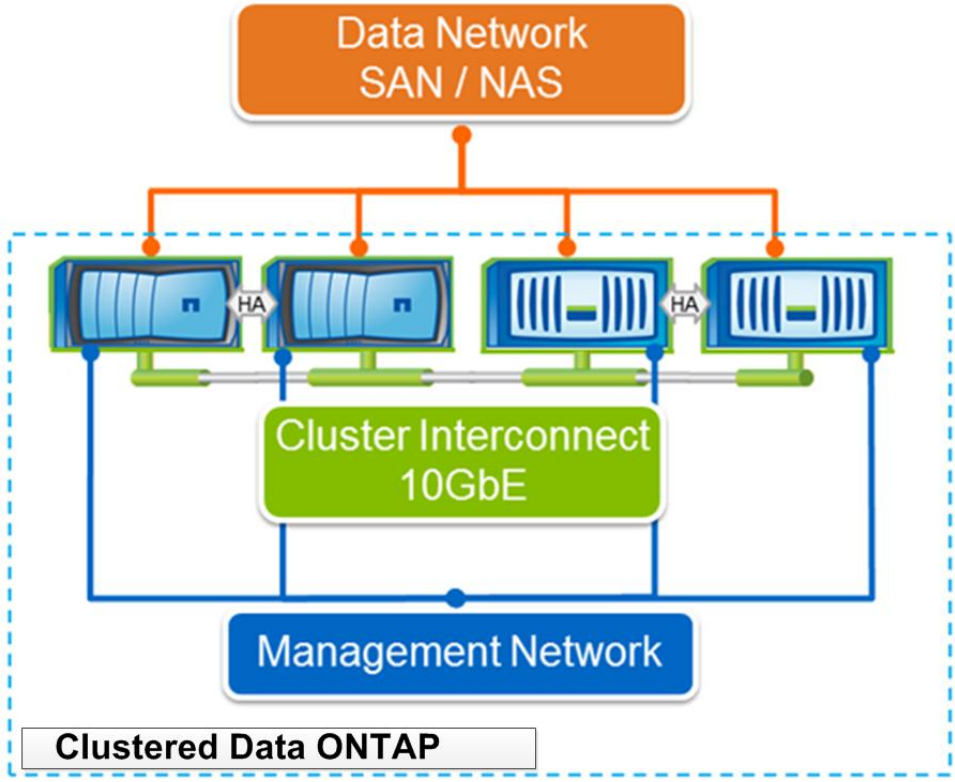
This section describes the architecture of Data ONTAP, with an emphasis on separation of physical resources and virtualized containers. Virtualization of storage and network physical resources is the basis for scale-out and nondisruptive operations.

Hardware Support and Basic System Overview

As Figure 2 shows, a clustered Data ONTAP system consists of two or more individual NetApp storage controllers (including V-Series) with attached disks. The basic building block is the high-availability (HA) pair, a term familiar from Data ONTAP 7G and 7-Mode environments. An HA pair consists of two identical controllers, each of which actively provides data services and has redundant cabled paths to the other controller's disk storage. If either controller is down for any planned or unplanned reason, its HA partner can take over its storage and maintain access to the data. When the downed system rejoins the cluster, the partner gives back the storage resources.

Note: Historically, the term *cluster* has been used to refer to an HA pair running Data ONTAP 7G or Data ONTAP operating in 7-Mode. This usage has been discontinued, and HA pair is the only correct term for this configuration. The term cluster now refers only to a configuration of one or more HA pairs within a clustered Data ONTAP deployment.

Figure 2) Data ONTAP cluster overview.



One of the key differentiators in a clustered Data ONTAP environment is that multiple HA pairs are combined into a cluster to form a shared pool of physical resources that are available to applications, SAN hosts, and NAS clients. For management purposes, the shared pool appears as a single system image. This means that there is a single common point of management, through either GUI or CLI tools, for the entire cluster. Although the members of each HA pair must be the same controller type, the cluster can consist of heterogeneous HA pairs. Over time, as the cluster grows and new controllers are released, it is likely to evolve into a combination of several different node types. All cluster capabilities are supported, regardless of the underlying controllers in the cluster.

Scalability

Data ONTAP 8.1 allows the inclusion of a wide variety of controller types in the same cluster, protecting the initial hardware investment and providing the flexibility to adapt resources to meet the business demands of the workloads. Similarly, support for different disk types, including SAS, SATA, and solid-state disk (SSD), makes it possible to deploy integrated storage tiering for different data types, together with the transparent DataMotion capabilities of Data ONTAP 8.1. Flash Cache cards can also be used to provide accelerated read performance for frequently accessed data. Data ONTAP 8.1.1 supports Flash Pool intelligent caching, which combines SSD with traditional hard drives for optimal performance and efficiency by using virtual storage tiering. The highly adaptable Data ONTAP architecture is the key to delivering maximum, on-demand flexibility for the shared IT infrastructure, and it offers flexible options to address needs for performance, price, and capacity.

Data ONTAP can scale both vertically and horizontally. High individual node capacities (for example, 4.3PB maximum storage on high-end controller HA pairs) mean that a cluster can scale to a large number of petabytes. This scalability, combined with protocol-neutral storage efficiency, can meet the needs of the most demanding workloads.

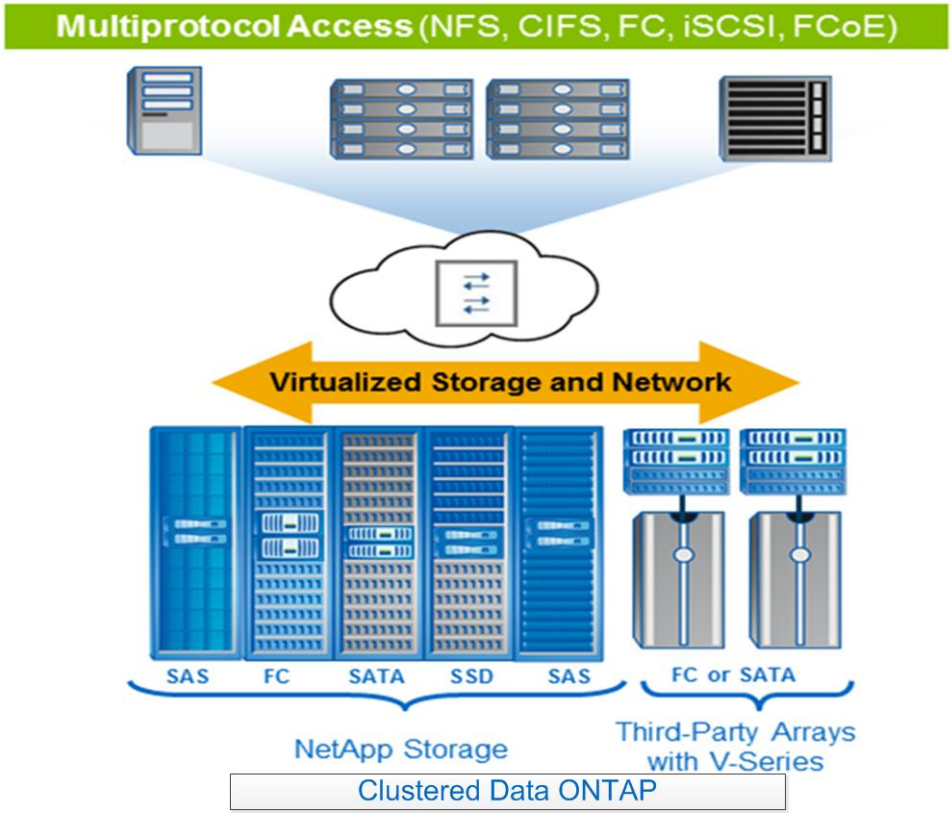
Networking in Data ONTAP 8.1

Figure 2 shows the underlying network architecture of Data ONTAP 8.1. Three networks are described.

- **Cluster interconnect.** A 10GbE/sec, private, dedicated, redundant, and high-throughput network used for communication between the cluster nodes and for DataMotion data migration. The cluster interconnect infrastructure is included with every Data ONTAP 8.1 configuration to support this network.
- **Management network.** All management traffic passes over this network. The management network switches are also included with every Data ONTAP 8.1 configuration. NetApp OnCommand® System Manager 2.0 and Unified Manager are available for management, configuration, and monitoring of Data ONTAP clusters, along with 7-Mode systems. These utilities provide GUI management, including a number of easy-to-use wizards for common tasks. In addition, a CLI, a set of APIs, and a software developer's kit are available for more specialized use.
- **Data networks.** These networks provide data access services over Ethernet or Fibre Channel to the SAN hosts and NAS clients. They are provided by the customer, according to the customer's requirements, and may include connections to other clusters acting as volume replication targets for data protection.

Figure 3 shows a large heterogeneous cluster consisting of different controller types, different disk types, and a mix of native NetApp FAS and V-Series controllers. V-Series makes it possible to front-end third-party storage with a NetApp controller, so that it can run Data ONTAP and participate in a cluster. Figure 3 also shows the client/host connections and the virtualized storage and network layer, which are explained in section 3, "Red Hat Enterprise Virtualization 3.1 Overview."

Figure 3) NetApp clustered Data ONTAP.



Storage Efficiency and Data Protection

The storage efficiency built into Data ONTAP offers substantial space savings, allowing more data to be stored at lower cost. Data protection provides replication services, making sure that valuable data is backed up and recoverable. The following features provide storage efficiency and data protection:

- **Thin provisioning.** Volumes are created by using "virtual" sizing. They appear to be provisioned at their full capacity, but they are actually created much smaller and use additional space only when it is needed. Extra unused storage is shared across all volumes, and the volumes can grow and shrink on demand.
- **Snapshot copies.** Automatically scheduled point-in-time copies that write only changed blocks, with no performance penalty. Snapshot copies consume minimal storage space, because only changes to the active file system are written. Individual files and directories can be easily recovered from any Snapshot copy, and the entire volume can be restored back to any Snapshot state in seconds.
- **FlexClone® volumes.** Near-zero space, instant virtual copies of datasets. The clones are writable, but only changes to the original are stored, so they provide rapid, space-efficient creation of additional data copies, ideally suited for test/dev environments.
- **Deduplication.** Removes redundant data blocks in primary and secondary storage with flexible policies to determine when the deduplication process is run.
- **Compression.** Compresses data blocks. Compression can be run whether or not deduplication is enabled and can provide additional space savings, whether run alone or together with deduplication.
- **SnapMirror®.** Volumes can be asynchronously replicated either within the cluster or to another cluster.

Cluster Virtualization and Multi-Tenancy Concepts

A cluster is composed of physical hardware; that is, storage controllers with attached disk shelves, NICs, and optional Flash Cache cards. Together, these create a physical resource pool that is virtualized as logical cluster resources to provide data access. Abstracting and virtualizing physical assets into logical resources offers the flexibility and potential multi-tenancy of Data ONTAP as well as the DataMotion ability at the heart of nondisruptive operations.

2.7 Physical Cluster Components

Although storage controllers can be of different types, by default they are considered equivalent in the cluster configuration because they are all presented and managed as cluster nodes. Individual disks are managed by defining them into aggregates—groups of disks of a particular type that are protected by using NetApp RAID-DP®. This is similar to the operation of Data ONTAP 7G and 7-Mode. NICs and HBAs provide physical ports (Ethernet and Fibre Channel) for connection to the management and data networks. The physical components are visible only to cluster administrators and not directly to the applications and hosts that are using the cluster. The physical components constitute a pool of resources from which the logical cluster resources are constructed. Applications and hosts access data only through virtual servers that contain volumes and logical interfaces (LIFs).

2.8 Logical Cluster Components

The primary logical cluster component is the virtual server. Data ONTAP supports from one to hundreds of virtual servers in a single cluster. Each virtual server enables one or more SAN and NAS access protocols and contains at least one volume and one LIF. The administration of each virtual server can also be delegated, so that separate administrators can be responsible for provisioning volumes and other virtual server–specific operations. This is particularly appropriate for multi-tenant environments or where workload separation is desired.

For NAS clients, the volumes are junctioned together into a namespace for CIFS and NFS access; for SAN hosts, LUNs are defined in the volumes and made available as described in section 2.2, “SAN Data Services.” The accessing hosts and clients connect to the virtual server by using a LIF. LIFs present either an IP address (which is used by NAS clients and iSCSI hosts) or a World Wide Name (WWN) (for FC and FCoE access). Each LIF is mapped to a home port on a NIC or HBA. LIFs are used to virtualize the NIC and HBA ports rather than mapping IP addresses or WWNs directly to the physical ports, because there are almost always more LIFs than physical ports in a cluster. Each virtual server requires its own dedicated set of LIFs, and up to 128 LIFs can be defined on any cluster node. A LIF defined for NAS access can temporarily move or migrate to another port on the same or a different controller, to preserve availability or to rebalance client performance.

Figure 4 shows a single virtual server that provides data services to SAN hosts and NAS clients. Each volume, as shown by an orange circle, is provisioned on an aggregate on a cluster node, and the combination of all the volumes constitutes the entire namespace or resource pool for LUNs. By default, volumes in a virtual server can be created in any of the defined aggregates and moved at any time from aggregate to aggregate as required. Delegated virtual server administrators can provision volumes in their own virtual servers. However, they cannot initiate the movement of volumes across the cluster, because this might affect the entire cluster. Only a cluster administrator can move volumes.

Note: Optionally, a cluster administrator can restrict the aggregates that can be used to provision volumes in a particular virtual server. This allows a virtual server to provide different classes of service—for example, by authorizing the virtual server to use only aggregates consisting of SSD or SATA drives or only aggregates on a particular subset of controllers.

Figure 4) One virtual server in a cluster.

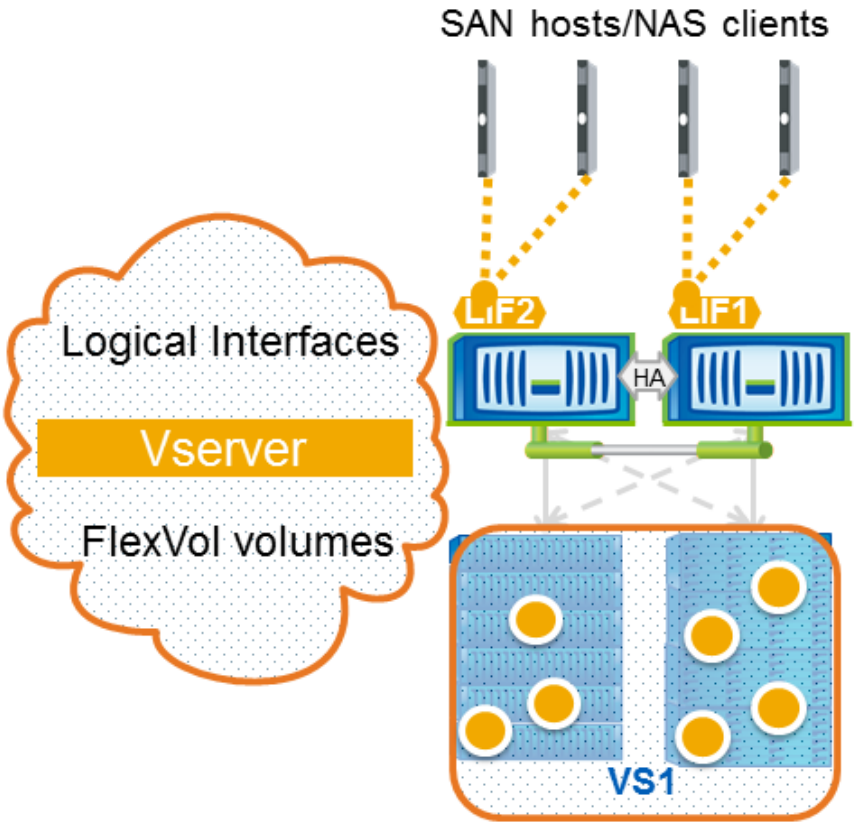
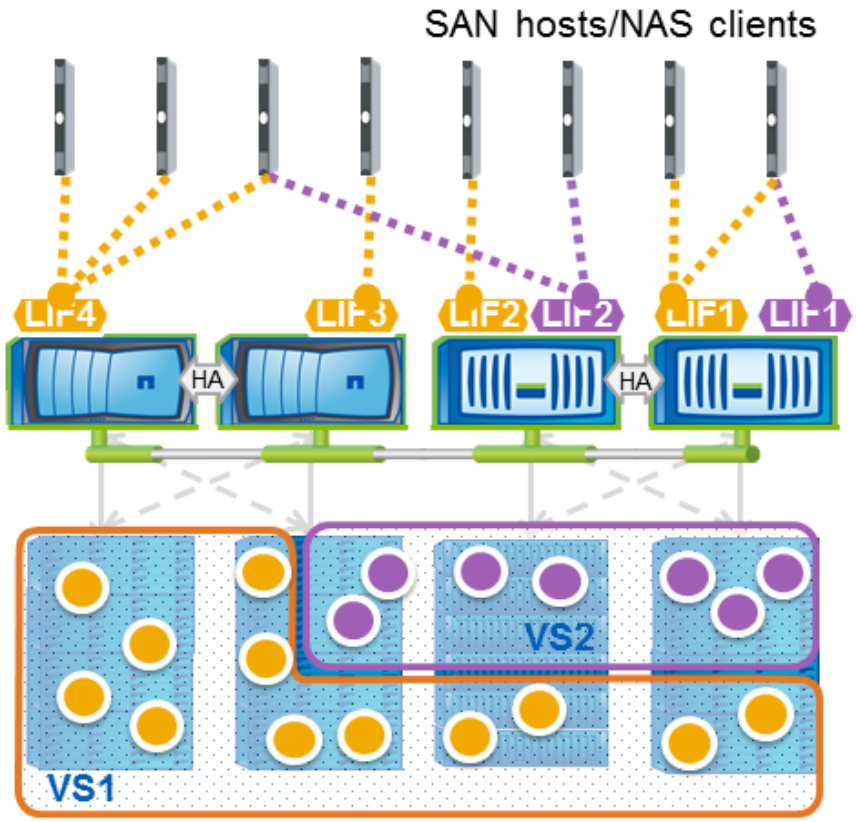


Figure 5 shows a more complex environment. There are four nodes in the cluster and two virtual servers providing SAN or NAS data access. Each virtual server consists of different volumes and LIFs, providing secure compartmentalized access. Although the volumes and LIFs in each virtual server share the same physical resources (network ports and storage aggregates), a host or client gets access to the data in VS1 only through a LIF defined in that virtual server, and the same rule applies to VS2. Administrative controls make sure that a delegated administrator with access to VS1 can see only the logical resources assigned to that virtual server, and a VS2-delegated administrator sees only the resources of VS2.

Figure 5) Multiple virtual servers in the same cluster.



By virtualizing physical resources into the virtual server construct, Data ONTAP implements multi-tenancy and scale-out and allows a cluster to host many independent workloads and applications.

2.9 Infinite Volume

Optionally, clustered Data ONTAP 8.1.1 can be configured to provide a single large volume (up to 20PB and 2 billion files) with NFSv3 client access, which is ideally suited for enterprise content repositories. For more information on Infinite Volumes, refer to TR-4037: Introduction to NetApp Infinite Volume.

3 Red Hat Enterprise Virtualization 3.1 Overview

Balancing business, technical, and budget requirements is a core issue for CIOs, engineers, and virtualization administrators. Even as data centers are consolidated by way of virtualization and converged infrastructures, individual layers of the stack must provide significant value when weighed

against the business, technical, and budgetary factors. Red Hat Enterprise Virtualization addresses these core requirements as follows:

- **Full support for virtualized applications.** From a support standpoint, Red Hat makes no distinction between a virtualized application and a native application. If an application has been certified to run on Red Hat Enterprise Linux (RHEL) 6, then that certification carries forward automatically to RHEL 6 KVM and RHEV-H.
- **Simple per-core licensing that includes all features and costs less.** Although some virtualization platforms require additional licenses in order to add functionality, RHEV includes all features, regardless of whether the customer chooses Basic or Premium support.
- **A constantly growing number of enterprise features.** RHEV 3.1 includes the addition of Live Storage Migration, Live Snapshots, user quotas, and an updated reporting dashboard.
- **Proven performance for KVM.** The KVM hypervisor that RHEV is built around currently holds 19 of the 27 published SPECVIRT_SC2010 performance benchmarks.

3.1 RHEV 3.1 Concepts

The following terms and concepts provide context to the Red Hat lexicon:

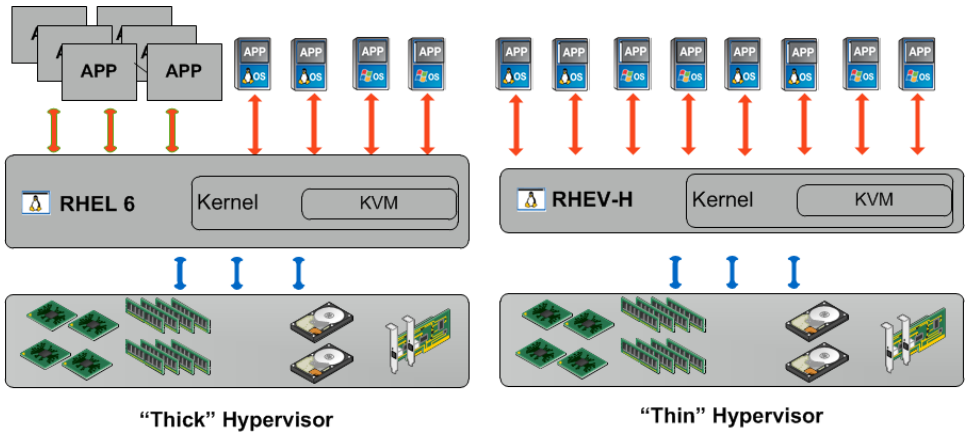
- **KVM.** Kernel-based Virtual Machine (KVM) is a loadable kernel module that turns the Linux kernel into a hypervisor. This is in contrast to many virtualization technologies that employ the hypervisor as a separate layer above the kernel or operating system.

  Note: Although RHEV is a Red Hat specific product, the KVM hypervisor is part of the Linux kernel and therefore is available from most Linux distributions, not just RHEL. Because of these differences, it is incorrect to use the terms KVM and RHEV interchangeably.

- **Thin hypervisor.** A thin hypervisor provides "just enough operating system" to manage server hardware and support virtual machines. In the case of the Red Hat Enterprise Virtualization Hypervisor 3.1(RHEV-H), it is an RHEL 6 distribution that has been stripped down to the bare minimum number of packages and has already been optimized to run and protect virtual machines. RHEV-H requires the use of the Red Hat Enterprise Virtualization Manager (RHEV-M) for day-to-day operations reconfiguration.
- **Thick hypervisor.** Thick hypervisor refers to the use of Red Hat Enterprise Linux 6 (RHEL) and KVM to run virtual machines. A thick hypervisor can run both native applications and virtual machines simultaneously.

Figure 6 illustrates thick and thin hypervisors.

Figure 6) Thick and thin hypervisors.



- **RHEV-M.** Red Hat Enterprise Virtualization Manager (RHEV-M) is a centralized management platform that provides a graphical interface.
- **Data center.** A data center is a logical container for clusters, hosts, virtual machines, and storage domains. A data center can contain multiple clusters, each of which can contain multiple hosts. Additionally, multiple data centers can exist in a single RHEV deployment.
- **Cluster.** A cluster is a logical container for hosts of the same CPU type. Virtual machines can live migrate only between hosts in the same cluster.
- **Logical network.** A logical network is a means of providing network connectivity between virtual resources and hosts. Logical networks can be used for management, virtual machine (VM) connectivity, or storage.
- **Host.** A host is a hypervisor that runs virtual machines. In the context of RHEV, a hypervisor can be of the thin type (RHEV-H) or the thick type (RHEL 6 KVM).
- **Virtual machine.** A virtual machine, or guest, is a computer that has been decoupled from physical resources. RHEV supports virtual machines that run Red Hat Enterprise Linux or Microsoft® Windows®.
- **Storage domain.** A storage domain is a container that is used to house virtual machines, RHEV-based Snapshot copies, or ISO images.

3.2 RHEV 3.1 Architecture

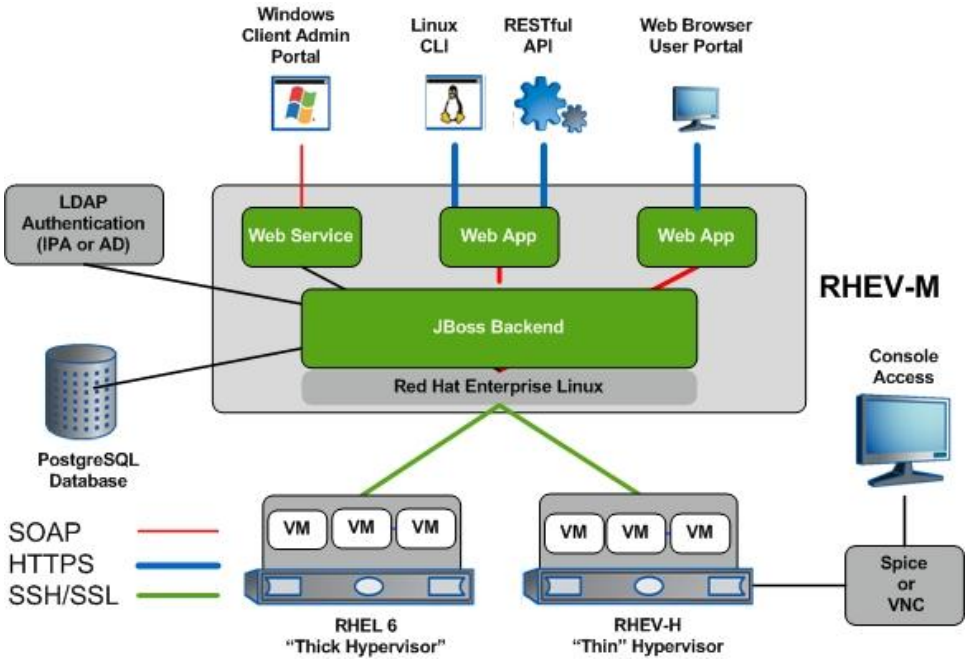
Major Components of RHEV

Red Hat Enterprise Virtualization (RHEV) includes two major components:

- A centralized management server, RHEV Manager, is built on JBoss® and a Postgres database, used to manage virtual resources, including virtual machines, logical networks, and storage domains, as well as thick and thin hypervisors.
- A thin hypervisor, RHEV Hypervisor (RHEV-H), is a small-footprint appliance-like hypervisor.

Figure 7 illustrates the RHEV architecture.

Figure 7) RHEV architecture.



4 RHEV-M Deployment

This section describes the best practices for deploying RHEV-M as well as other infrastructure applications.

4.1 Providing High Availability to RHEV-M and Infrastructure Applications

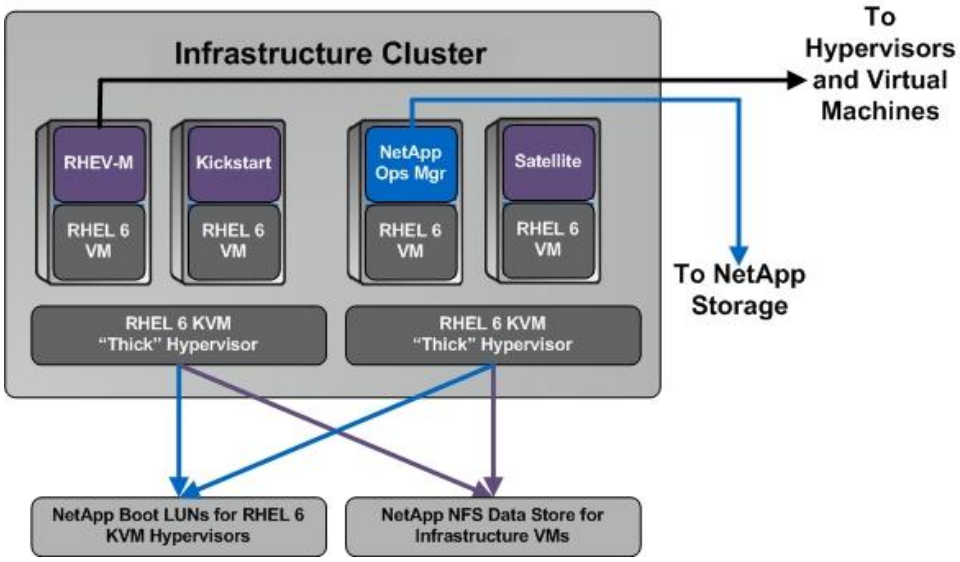
The best practice for deploying RHEV-M is to do so in a highly available manner. The simplest way to accomplish this is to deploy RHEV-M on a virtual machine that is hosted by two or more RHEL 6 KVM hosts. Extending that best practice to the Kickstart server, Red Hat Network Satellite server, NetApp OnCommand System Manager, and other infrastructure applications also makes good sense. Refer to TR-4104: Best Practices for RHEL 6, KVM, and Clustered Data ONTAP for guidelines for deploying the RHEL 6 KVM hypervisors.

Note: NetApp does not recommend including a Snap Creator® server in the same cluster that it is protecting, because it may pause the virtual machine that it is deployed on, with no way of resuming by itself.

Although RHEL 6 KVM supports both SAN and NAS datastores for virtual machines, using NFS keeps this infrastructure cluster as simplified as possible.

As shown in Figure 8, the infrastructure cluster exists separately from any hypervisors and virtual machines that support the application environment. The underlying NFS storage uses NetApp clustered Data ONTAP and is inherently fault tolerant and highly available. Additionally, the RHEL KVM hypervisors boot from SAN, as per best practices.

Figure 8) Infrastructure cluster deployment example.



4.2 RHEV-M Hardware Requirements

The following requirements are the same regardless of whether RHEV-M is virtualized or runs natively on RHEL 6:

- A dual core CPU; a quad core is preferred
- 4GB of RAM; 16GB is preferred
- 25GB of disk space; 50GB is preferred
- At least one 1GbE network interface card (NIC)

4.3 RHEV-M Software Requirements

- RHEV-M version 3.1 requires RHEL 6.3 or later as the operating system platform
- RHEV-M requires subscription to Red Hat Network (RHN)
- RHEV for desktops also requires access to either Microsoft Active Directory®, Red Hat Directory Server, or Red Hat IPA
- RHEV-M administrator access has the following requirements:
  - Mozilla Firefox 10 or later for access from RHEL
  - Internet Explorer® 8 or later for user access from Microsoft Windows
  - Internet Explorer 9 or later for administrator access from Microsoft Windows
- Virtual machine console access has the following requirements:
  - RHEL 5.8 or later
  - RHEL 6.2 or later
  - Microsoft Windows XP, Windows 7, or Windows 2008 R2
  - RHEV certified Linux-based thin client

Note: The spice-xpi package is available for Mozilla Firefox as a plug-in.

4.4 Storage Requirements for RHEV and Virtual Machines

RHEV supports the following protocols for virtual machine and ISO storage:

- NFS
- iSCSI
- Fibre Channel (FC)
- Fibre Channel over Ethernet (FCoE)

RHEV guests support the disk types shown in Table 3.

Table 3) Supported disk types.

NFS or SAN	Raw or Qcow2	Sparse or Preallocated
NFS	Raw	Preallocated
NFS	Raw	Sparse
NFS	Qcow2	Sparse
SAN	Raw	Preallocated
SAN	Qcow2	Preallocated
SAN	Qcow2	Sparse

A preallocated virtual disk in RHEV is thick provisioned in the sense that the space equal to the presented size is allocated at the time of deployment.

A sparse file in RHEV is thin provisioned in the sense that more space is presented to the guest than what was actually allocated at the time of deployment. As the guest writes to disk, additional space is allocated automatically as necessary.

In the case of SAN (iSCSI, FCoE, and FC) storage in RHEV, the entire LUN is initialized with LVM2, as one large volume group. Individual disks are handled as logical volumes inside the volume group. For preallocated disks, the logical volume is the same size as the disk presented to the virtual guest. For sparse disks, the logical volume is much smaller than what is presented to the guest (it starts at 1GB) and grows automatically as needed.

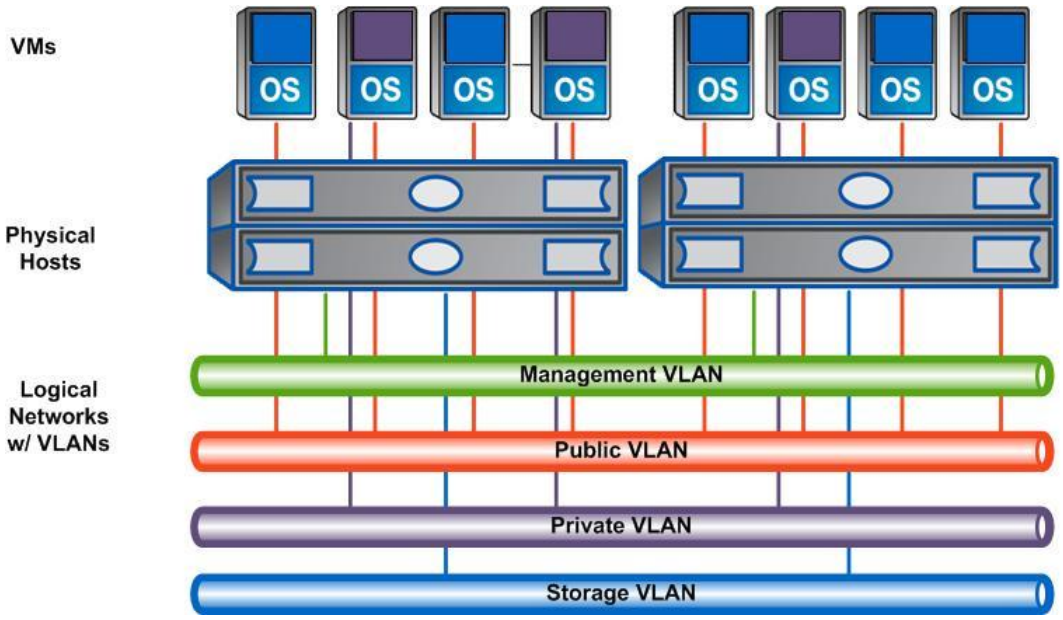
Virtual disks backed by NFS storage are typically of raw type, but Qcow2 is also supported. A Qcow2 file is initially created as close to zero in size as possible, but it has an initial raw format. As the disk size grows, additional layers are added to the disk, which are all formatted with Qcow2. In contrast to Qcow2, a raw disk has no additional layer of abstraction.

4.5 Network Layout for RHEV and NetApp Storage

There are multiple correct ways to implement a sound storage network. Figure 9 illustrates NetApp best practices to implement a storage network layout. Note that the storage network is on a completely separate VLAN from the other data traffic.

Note: See Appendix B: Ports to Allow Through the Firewall B, “Ports to Allow Through the Firewall,” for the list of ports that RHEV-M requires access through.

Figure 9) Network layout best practices.



4.6 Deploying RHEV-M

Install and Configure RHEV-M

1. Follow best practices for installing Red Hat Enterprise Linux 6 (RHEL 6) as a guest virtual machine on a RHEL 6 KVM host. Additionally, the fully qualified domain name must be listed in the `/etc/hosts` file and be fully resolvable in DNS.
2. Log in to the RHEL 6 virtual machine on which RHEV-M will be installed.
3. In addition to the base RHEL 6 software channel, subscribe the host to the following channels:
   - rhel-x86_64-server-6
   - rhel-x86_64-server-supplementary-6
   - rhel-x86_64-server-6-rhevm-3.1
   - jbappplatform-6-x86_64-server-6-rpm
4. Update all installed packages.

```
# yum -y update
```

5. Install and then configure RHEV-M with the following commands.

```
# yum -y install rhevm
# rhevm-setup
```

Note: Be prepared with an administrative password, a database password, and an organization name; the configuration tool will prompt for these items.

6. Accept the defaults for all options, including default storage, database location, iptables configuration, and confirmation of configuration. You can decline the creation of any storage domains at the time of installation.

Note: NFS-based storage domains should not be created at this time, because there are a few extra steps required when mounting the NFS exports for the first time. This is described in detail in section 6.9, “Create an ISO Domain.”

Connect to the RHEV-M Portal for the First Time

Before connecting to the RHEV-M portal, be sure that the minimum requirements for the browser and client have been met.

1. In a web browser, enter:

   `http://rhevm.domain.com/webadmin`

   Note: Replace `rhevm.domain.com` with your fully qualified domain name.
2. Install the certificate that is presented and click Next.
3. Log in as user admin and use the administrative password that was created at the time of installation.

4.7 Deploying Resources in Red Hat Enterprise Virtualization 3.1

All of the following resources are required before attaching storage and creating virtual machines.

Create an RHEV Data Center

1. Log in to the RHEV-M portal and click the Data Centers tab.
2. Click the New button to display the data center dialog box.
3. Enter a name and description for the new data center.
4. Select the storage type: NFS, iSCSI, or FCP. (If using FCoE, select FCP.)
5. Select 3.1 for compatibility, then click OK.
6. When prompted with the Guide Me dialog box, select Configure Later.

Create an RHEV Cluster

1. Log in to the RHEV-M portal and click the Clusters tab.
2. Click the New button to display the cluster dialog box.
3. Select the data center that will own the new cluster.
4. Enter a name and description for the new cluster.
5. Select the CPU name (Intel® or AMD chipset family) and compatibility level of 3.1.
6. Under Memory Optimization, select the appropriate option (server, desktop, or none).
7. Click OK.
8. When prompted with the Guide Me dialog box, select Configure Later.

Create a Logical Network

To apply the best practice of VLAN tagging, logical networks must be created and attached to the appropriate RHEV clusters.

1. Log in to the RHEV-M portal and click the Data Centers tab.
2. Select the data center that will own the new logical network.
3. Click the Logical Networks tab.
4. Click the New button to view the Logical Networks dialog box.
5. Enter a name and a description. Select the Enable VLAN tagging box and enter the VLAN number.

   Note: It is a best practice to use VLANs to segregate public, private, and storage networks.
6. If the new logical network should not be accessible to virtual machines, uncheck the VM Network box.
7. Select the RHEV cluster or clusters that should have access to the logical network and click OK.

5 Hypervisor Installation

RHEV-M can manage both thin (RHEV-H) and thick (Red Hat KVM) hypervisors in the same environment. The following sections outline the considerations when determining the best choice to make based on use case.

In both cases, the hypervisor should be installed to and boot from a SAN LUN. This allows the NetApp controller to provide the best possible storage efficiency through deduplication and thin provisioning. For example, 10 hypervisors that have been installed on thin provisioned and deduplicated LUNs take the amount of space equivalent to a little more than one hypervisor. Additionally, SAN booting the hypervisors means that the NetApp controller can easily clone a hypervisor template for rapid deployment.

5.1 RHEV-H and RHEL 6 Hypervisor Hardware Requirements

RHEV hypervisors have the following hardware requirements:

- Support for either the Intel VT or AMD-V CPU extension
- At least 2GB of RAM plus enough RAM to support the applications being virtualized
- At least one 1GbE network interface card (NIC); 10GbE is the best practice
- At least 2GB of storage to install RHEV-H

RHEV hypervisors have the following swap space requirements:

- Multiply the amount of system RAM by the expected overcommit ratio and add:
  - 2GB of swap space for systems with 4GB of RAM or less
  - 4GB of swap space for systems with between 4GB and 16GB of RAM
  - 8GB of swap space for systems with between 16GB and 64GB of RAM
  - 16GB of swap space for systems with between 64GB and 256GB of RAM

RHEV hypervisors have the following software requirements:

- RHEV 3.1 requires version 6.3 or later for RHEL and RHEV hypervisors

5.2 RHEV-H Deployment and Best Practices

Use Case for RHEV-H

Red Hat has taken the appliance approach to hypervisors in their creation of RHEV-H. In RHEV 3.0, RHEV-H is an extremely stripped-down RHEL 6 distribution. The entire RHEV-H ISO image takes up about 100MB of space because the vast majority of RHEL packages have been removed.

The only configurable files are related to host name and networking; all other files and directories are read-only. Security pieces such as iptables and SELinux are preconfigured to allow only the proper amount of access for storage, virtual machine, and management traffic.

Because the majority of the file system on RHEV-H is read-only and there are not many libraries installed, no other packages can be updated or installed. From a security perspective, this is preferred, because there is not much to exploit and so provides a use case where security is of highest importance.

However, if a backup or additional monitoring agent is required, it is not possible to deploy them on RHEV-H.

Deploy RHEV-H

Refer to the Red Hat Enterprise Virtualization Installation Guide for step-by-step instructions to download the RHEV-H ISO image as well as instructions for deploying RHEV-H, other than booting from a CD ISO image.

1. Make sure that the server BIOS is configured to allow booting from a CD.
2. Make sure that a boot LUN has been created on NetApp storage and has been properly zoned.
3. After the server is powered on, the RHEV-H installation splash screen appears. Select Install or Upgrade and press Enter.
4. Select the desired keyboard language.
5. Select the boot LUN to install to and accept the defaults for disk layout.
6. Enter a password.
7. Select Install and press Enter.
8. When the installation is complete, click the Restart button and then press Enter.

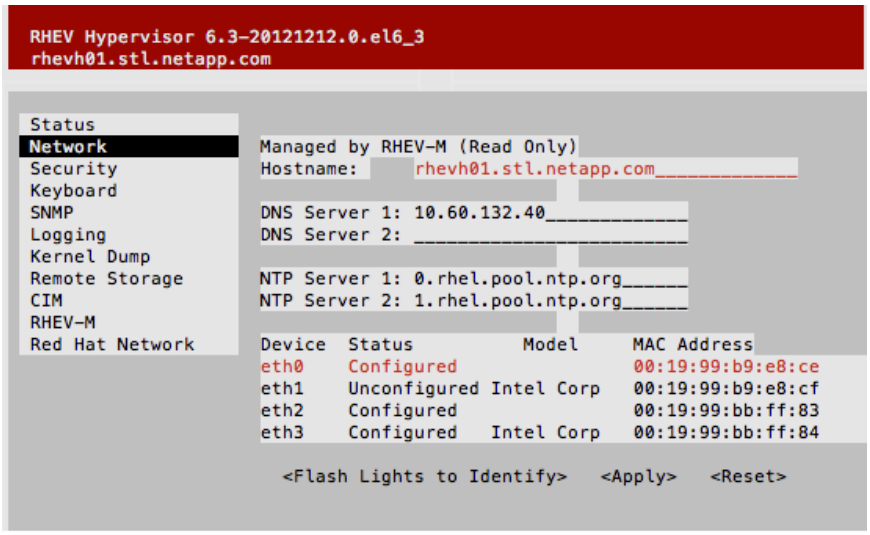
Security Best Practices Applicable to RHEV-H

- Create good passwords. For guidelines, refer to https://access.redhat.com/kb/docs/DOC-9128.
- Allow access only to virtualization administrators. Regular users should not be able to access anything on the RHEV-H host, and they should not be given logins to RHEV-H.
- RHEV-H is already stripped of unnecessary services, packages, and inherently nonsecure protocols.
- Segregate the management interface on a separate VLAN, preferably one that is not routable.

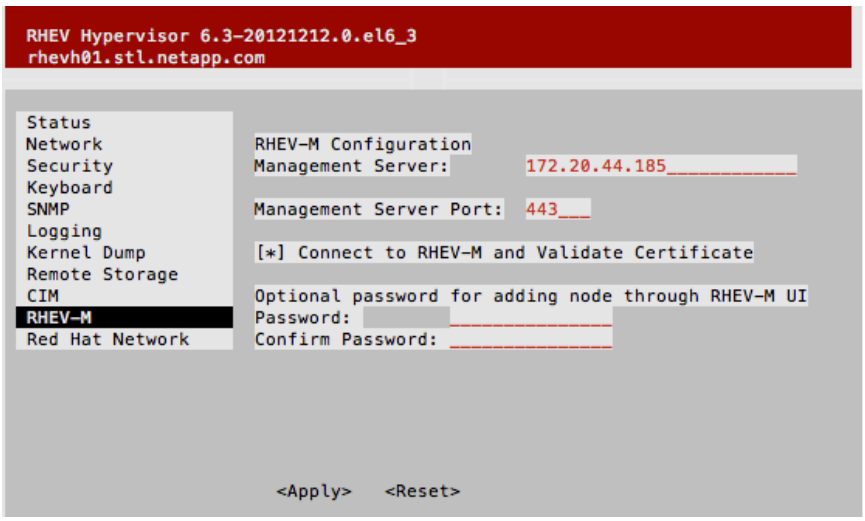
Configure RHEV-H

After the RHEV-H node has been installed and rebooted, it needs to be configured.

1. Log in to the RHEV-H node using the login admin and the password that were created during the installation process.
2. Click the Network tab and configure the host name, DNS servers, NTP servers, and network interface. Click Apply when done.



3. Click the Security tab to configure a password for SSH access to the RHEV-H node. Click Apply when done.
4. Click the RHEV-M tab to point the RHEV-H node to the RHEV-M server. Click Apply when done.



5. Click the Status tab and then select Log Off.

5.3 RHEL 6 KVM Deployment and Best Practices

Use Case for RHEL 6 KVM

Red Hat KVM can also be managed by RHEV-M, and it provides its own benefits and use cases. Although RHEL 6 can't be stripped down as much as RHEV-H, it can be stripped down to some extent. The single biggest case for using Red Hat KVM in RHEV is the ability to install a backup agent such as the NetApp Snap Creator agent, or even a monitoring agent.

For any backup requirement that includes quiescing the virtual machine before triggering a Snapshot copy, this is a benefit.

Note: Virtual machines hosted on RHEV-M can be quiesced by using the RHEV API. However, if an additional monitoring agent is needed on the hypervisor, this would be possible only on the thick hypervisor.

Deploying RHEL 6 KVM and Security Best Practices

Refer to TR-4104: Best Practices for RHEL 6, KVM, and Clustered Data ONTAP for the remaining guidelines for deployment and security regarding Red Hat KVM.

6 Red Hat Enterprise Virtualization Datastores

6.1 Raw Disk Image Files and Qcow2 Disk Image Files

RHEV supports the use of two different disk image formats, raw and Qcow2. A raw disk image is faster and is the default for preallocated virtual disks. Qcow2 has some advanced features such as virtual machine-based Snapshot copies, but at the cost of performance. The best practice when deploying virtual machines on RHEV is to preallocate disk files and allow the NetApp controller to thin provision,

deduplicate data, and create Snapshot copies. It can do so much more quickly and efficiently without involving the hypervisor CPU cycles.

6.2 LUN-Based Datastores

Red Hat Enterprise Virtualization uses Logical Volume Manager (LVM2) for managing LUN-based datastores. LVM2 is typically used as a layer of abstraction between a block-based storage device (local disk, direct attached, SAN, and so on) and the file system. However, it is used in RHEV as a means of managing LUNs and individual virtual machines without using a file system. Essentially, a LUN is configured as an LVM volume group and virtual machines are created as individual logical volumes within the volume group. This precludes the need for a clustered file system or for adding extra steps to any automation scripts for starting and stopping virtual machines.

6.3 NFS Datastores

Red Hat Enterprise Virtualization allows customers to leverage enterprise-class NFS arrays to provide datastores with concurrent access by all of the nodes that are using the shared storage. The NetApp NFS provides high performance, the lowest per-port storage costs, and advanced data management capabilities.

NFS Datastores on NetApp

Deploying RHEV with NetApp advanced NFS results in a high-performance, easy-to-manage implementation that provides VM-to-datastore ratios that cannot be accomplished with other storage protocols, such as FC. This architecture can result in a 10x increase in datastore density, with a correlated reduction in the number of datastores. With NFS, the virtual infrastructure receives operational savings because there are fewer storage pools to provision, manage, back up, replicate, and so forth.

Through NFS, customers receive an integration of Red Hat KVM virtualization technologies with WAFL®, the NetApp advanced data management and storage virtualization engine. This integration provides transparent access to VM-level storage virtualization offerings such as production-use data deduplication, array-based thin provisioning, automated policy-based datastore resizing, and direct access to array-based Snapshot copies.

Virtual machines that are backed by NFS-based datastores are created as simple files that act as block devices.

6.4 Datastore Comparison Tables

Differentiating what is available with each type of datastore and storage protocol requires considering many points. Table 4 compares the features available with each storage option.

Table 4) Datastore supported features.

Capability/Feature	FC/FCoE	iSCSI	NFS
Format	LVM	LVM	NetApp WAFL
Optimal queue depth per LUN or file system	Preconfigured on RHEV-H	Preconfigured on RHEV-H	N/A
Available link speeds	4GbE and 8GbE FC and 10GbE	1GbE and 10GbE	1GbE and 10GbE

Table 5 compares the storage-related functionality of Red Hat KVM features across different protocols. (These are the default settings for all VMs on NetApp NFS.)

Table 5) Red Hat-supported storage-related functionality.

Capacity/Feature	FC/FCoE	iSCSI	NFS
Live migration	Yes	Yes	Yes
Data deduplication	Yes	Yes	Yes
Resize datastore	Grow only	Grow only	Grow, autogrow, shrink
Thin provision datastores	Yes	Yes	Yes
NetApp Snapshot	Yes	Yes	Yes
Restore datastores and VMs from NetApp SnapMirror and SnapRestore®	Yes	Yes	Yes
Boot from SAN	Yes	Yes with HBAs	No

6.5 Create a RHEV Storage Domain

When NetApp storage is presented to RHEV by using SAN or NAS, it has to be initialized by RHEV-M. The following procedures walk you through the required steps to configure storage for virtual machines and guest operating system ISO images.

The following procedures assume that the NetApp storage has already been provisioned and that proper access has been granted for RHEV-M.

6.6 Create NFS Data Storage

1. The NFS export must be owned by the user and by group 36 prior to mounting and initializing through the RHEV-M portal. This is accomplished by manually mounting the NetApp export on one of the hypervisor nodes long enough to run the following command.

```
# chown -R 36:36 /mount/point/netapp_export
```

2. It may also be necessary to change permissions on the NFS export while it is manually mounted by entering the following command.

```
# chmod -R 755 /mount/point/netapp_export
```

3. Unmount the export and continue to the next step. (Failure to do this results in RHEV-M failing to mount and initialize the NFS storage domain.)
4. Log in to the RHEV-M portal and click the Storage tab.
5. Select New Domain.
6. Enter a name for the storage domain and then select the data center that will use the storage domain from the drop-down menu.
7. Select Data/NFS from the Function/Type drop-down menu.
8. Select the hypervisor that to attach to the storage for the initialization.
9. Enter the export path from the NetApp controller.
10. Click OK.

The initialized NFS data domain is ready to be attached to the RHEV data center.

6.7 Create iSCSI Data Storage

1. Log in to the RHEV-M portal and click the Storage tab.
2. Click the New Domain button to view the New Storage dialog box.
3. Enter a name for the storage domain and then select the data center that will use the storage domain from the drop-down menu.
4. Select Data/iSCSI from the Function/Type drop-down menu and then select the host in the data center that will attach to the storage domain first. Enter the IP address or host name of the NetApp controller, and keep the default port of 3260. If CHAP authentication is used, select the box and enter the user name and password.
5. Click Discover. All available iSCSI LUNs are listed. Select the LUN or LUNs to be used and click OK. RHEV-M initialize these LUNs for use as a virtual machine storage domain.

The initialized iSCSI data domain is ready to be attached to the RHEV data center.

6.8 Create FCP and FCoE Data Storage

1. Log in to the RHEV-M portal and click the Storage tab.
2. Click the New Domain button to open the New Storage dialog box.
3. Enter a name for the storage domain and then select the data center that will use the storage domain from the drop-down menu.
4. Select Data/FCP from the Function/Type drop-down menu and then select the host in the data center that will attach to the storage domain first. This triggers the host to scan for FC LUNs.
5. The list of available FC or FCoE LUNs appears. Select the LUN or LUNs to use and click OK. RHEV-M initializes the LUNs for use as a virtual machine storage domain.

The initialized FC data domain is ready to be attached to the RHEV data center.

6.9 Create an ISO Domain

1. The NFS export must be owned by the user and by group 36 prior to mounting and initializing through the RHEV-M portal. This is accomplished by manually mounting the NetApp export on one of the hypervisor nodes long enough to run the following command.

```
# chown -R 36:36 /mount/point/netapp_export
```

2. It may also be necessary to change permissions on the NFS export while it is manually mounted by entering the following command.

```
# chmod -R 755 /mount/point/netapp_export
```

3. Unmount the export and continue to the next step. (Failure to do this results in RHEV-M failing to mount and initialize the NFS storage domain.)
4. Log in to the RHEV-M portal and click the Storage tab.
5. Click New Domain.
6. Enter a name for the storage domain and then select the data center that will use the storage domain from the drop-down menu.
7. Select ISO/NFS from the Function/Type drop-down menu.
8. Select the hypervisor that will attach to the storage for the initialization.
9. Enter the export path from the NetApp controller.
10. Click OK.

The initialized ISO domain is ready to be attached to the RHEV data center.

6.10 Attach a Storage Domain to an RHEV Data Center

1. Log in to the RHEV-M portal and click the Data Centers tab.
2. Select the RHEV data center to which the storage domain will be attached and then click the Storage subtab.
3. Click the Attach Domain button to view the Attach Storage Domain dialog box. Select the storage domain and click OK. Select the newly attached storage domain and click the Activate button.

The storage domain is now ready to store virtual machines.

6.11 Attach an ISO Domain to an RHEV Data Center

1. Log in to the RHEV-M portal and click the Data Centers tab.
2. Select the RHEV data center to which the ISO domain will be attached and then click the Storage subtab.
3. Click the Attach ISO button to view the Attach Storage Domain dialog box. Select the storage domain and click OK. Select the newly attached storage domain and click the Activate button.
4. The storage domain is now ready to store ISO images.

6.12 Populate an ISO Domain

1. Log in as `root` to the RHEL 6 host that RHEV-M runs on.
2. Copy any ISO images that need to be uploaded to the ISO domain to `/tmp`.
3. Upload the ISO image by entering the following command.

```
# rhevm-iso-uploader --iso-domain=/tmp/name_of_iso_domain upload name_of_image.iso
```

4. It may be necessary to start the `rpcbind` service in order to get the upload tool to work properly, by entering the following command.

```
# service rpcbind start
```

The ISO image can now be used by RHEV-M to create guest operating systems.

7 RHEV Guest Configuration

7.1 File System Alignment Overview

In any virtual environment, there are a number of layers of abstraction between the physical disks and the VM's virtual disk. Each layer in turn is organized into blocks to make the most efficient use of storage. The focus is not on the size of the block, but rather on the starting offset.

To avoid latency caused by extra reads and writes, the starting offset of a file system on a virtual machine should line up with the start of the block at the next layer. This alignment should continue all the way down to data blocks at the aggregate layer on the NetApp controller.

This is not unique to NetApp; it is true for any storage vendor. It is a simple by-product of legacy partitioning schemes. For the full explanation of disk alignment in virtual environments, see TR-3747: Best Practices for File System Alignment in Virtual Environments.

Without proper alignment, significant latency occurs because the storage controller has to perform additional reads and writes for the misaligned blocks. For example, most modern operating systems such as RHEL and Windows 2000 and 2003 use a starting offset of sector 63. Pushing the offset to sector 64 or sector 128 causes the blocks to align properly with the succeeding layers.

Microsoft Windows Server® 2008, Windows 7, and RHEL 6 align properly by default and require no additional consideration. However, earlier versions of Microsoft Windows (Server 2003, XP, and so on) and RHEL (3, 4, and 5) require additional steps at deployment for proper file system alignment.

Thick or Thin Provisioning of KVM Guests

Red Hat KVM allows both thick and thin provisioning of guest virtual disks. Red Hat recommends thick provisioning production guests and thin provisioning desktops and dev/test guests to balance performance and storage capacity.

However, when coupled with NetApp, the underlying storage can be thin provisioned and deduplicated. This allows all Red Hat KVM guests to be thick provisioned but to still maintain an efficient storage environment. The best practice is to thick provision the KVM guests and to thin provision and deduplicate the underlying storage.

7.2 Create an RHEV Guest Template

Instead of creating a new virtual machine guest each time, it is a best practice to create it first with Kickstart, and then clone the subsequent instantiations.

The concept is almost identical to creating a template for Red Hat KVM hosts. The base image is created by using Kickstart, and the image is then made generic. When new guests are needed, the RHEV native cloning tools can be used to provision new virtual machines. For Microsoft guests, sysprep is used to make the guest sufficiently generic.

7.3 Kickstart

It is a best practice to use Kickstart to build an RHEV guest for the first time. Kickstart provides a semiautomated method to deploy Red Hat Enterprise Linux in a highly customizable way. Once an RHEV guest is created and made generic, it is a best practice to repeat the deployment by using NetApp FlexClone.

7.4 Guest Timing Issues

All RHEV (Windows and RHEL) guests must be configured to use NTP to avoid issues that arise from time skew.

7.5 Security Considerations

Similar to Red Hat host security, the focus is on firewall, mandatory access control, unnecessary services, and insecure services.

RHEL guests should have iptables running and SELinux enabled. As a best practice, configure the necessary ports or controls rather than disabling that layer of security. Limit the packages installed to those that are necessary. Don’t use RSH, Telnet, and FTP; rather, use of SSH, SCP, and SFTP.

Most Windows guests have some form of firewall. It is a best practice to open the necessary port, rather than disabling that layer of security. Additionally, the services running should be reviewed and unnecessary services should be shut down.

7.6 Tuning RHEL Guests for Virtualization

RHEL 6.3 offers the ability to tune the operating system to suit a particular role. To tune the operating system for use as a virtual guest, enter the following commands.

```
# yum -y install tuned*
# chkconfig tuned-adm on
```

```
# service tuned start
# tuned-adm profile virtual-guest
```

7.7 RHEV Virtualized Guest Support

RHEV presently supports the following virtualized guest operating systems:

- Red Hat Enterprise Linux 3 (32-bit and 64-bit)
- Red Hat Enterprise Linux 4 (32-bit and 64-bit)
- Red Hat Enterprise Linux 5 (32-bit and 64-bit)
- Red Hat Enterprise Linux 6 (32-bit and 64-bit)
- Windows XP Service Pack 3 and later (32-bit only)
- Windows 7 (32-bit and 64-bit)
- Windows Server 2003 Service Pack 2 and later (32-bit and 64-bit)
- Windows Server 2008 (32-bit and 64-bit)
- Windows Server 2008 R2 (64-bit only)
- Windows Server 2012 (when using RHEL/RHEV-H 6.4 or later)

RHEV supports the following virtual hardware configurations:

- Virtual CPUs, up to 160 per guest
- Virtual RAM, up to 2TB per 64-bit guest
- Virtual RAM, up to 4GB per 32-bit guest
- Virtual storage devices, up to eight per guest
- Virtual network devices, up to eight per guest
- Virtual PCI devices, up to eight per guest

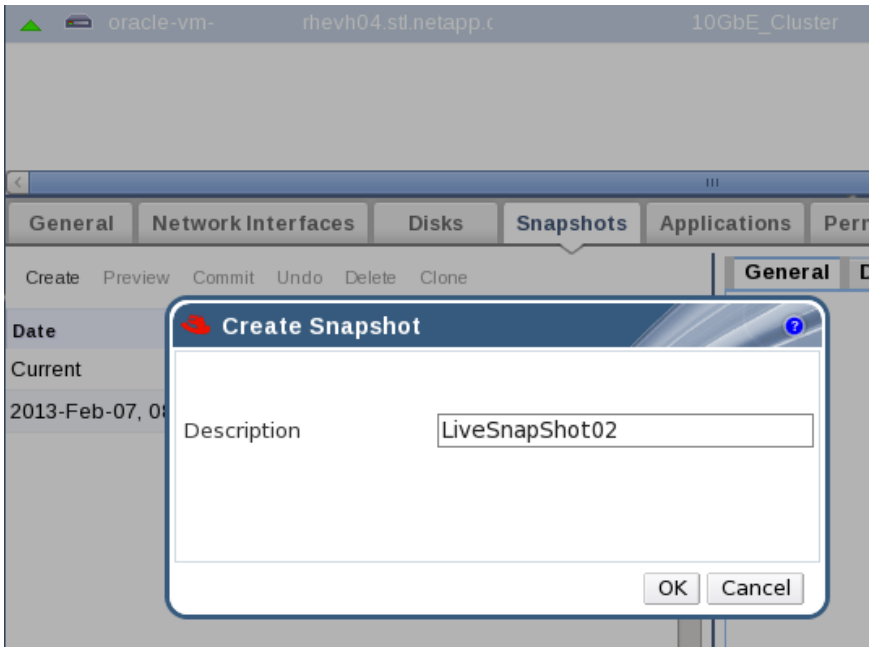
7.8 Adding a Logical Network to an RHEV Hypervisor or RHEL 6 Server

1. Log in to the RHEV-M portal and click the Hosts tab.
2. Click the Details subtab and then click the Network Interfaces tab.
3. Select the physical interface to add the Logical Network to, and then click Add/Edit.
4. From the drop-down menu, select the Logical Network name.
5. Select Static IP and enter the IP address and Netmask for the logical network interface.
6. Select Save Network Configuration and click OK.

7.9 Performing a Live Snapshot of an RHEV Guest

In the context of Red Hat Enterprise Virtualization, a Snapshot is a point-in-time copy of a single virtual machine. The new Live Snapshot feature in RHEV 3.1 allows a Snapshot copy to be taken without stopping or pausing the virtual machine. These Snapshot copies can be used to restore a virtual machine to an earlier known good version.

1. Log in to the RHEV-M portal.
2. Click the Virtual Machines tab.
3. Select the virtual machine that requires a Snapshot copy.
4. Under the virtual machine menu, click the Snapshots tab.
5. Click the Create button, enter a description, and click OK.



Use Cases for RHEV Snapshot and NetApp Snapshot Copy

As mentioned earlier, an RHEV-based snapshot is of a single virtual machine. In the context of NetApp, a Snapshot copy is a point-in-time copy of an entire volume; for virtualization, this means an entire group of virtual machines.

An RHEV snapshot may be better suited for testing a new configuration or patch set on a single virtual machine. This option is appealing for a dev/test environment where such changes do not require strict change management, because it gives the virtualization administrator the ability to quickly roll back the virtual machine without affecting the other virtual machines in the volume.

The other primary use case is to create a new virtual machine from an RHEV snapshot. This is useful when updates to an existing virtual machine constitute the new standard image. This saves time when compared to creating a new virtual machine and then adding the new updates and/or configuration.

These are good use cases only in the context of snapshot copies for one virtual machine at a time. The other factors to consider are backup and archive. A snapshot copy of any kind cannot be considered a backup until it is stored on a separate storage array from the source. Otherwise, any calamity involving the source storage array affects not only the virtual machine, but also its copy.

A NetApp Snapshot copy is preferred when a near-instantaneous copy of an entire group of virtual machines is required. A Snapshot copy makes it trivial to restore the entire volume but still allows the restoration of a single virtual machine. NetApp Snapshot copies are also equally effective on the three primary Snapshot consistency levels; they are crash consistent, application consistent, and fully consistent.

Note: Review section 12.2, “Snap Creator,” which explains the three main levels of Snapshot consistency and how to achieve them.

NetApp Snapshot copies are also well suited for backup. When used in conjunction with NetApp SnapMirror, NetApp volumes and their Snapshot copies are asynchronously mirrored to other NetApp storage arrays. These arrays may be in a different area of the date center, or in different locations altogether. These mirrored volumes can be easily restored to their source volumes for full recovery.

8 NetApp Storage Best Practices for RHEV

8.1 64-Bit Aggregates in Clustered Data ONTAP

Overview

NetApp Data ONTAP 8.x supports a new type of aggregate, called a *64-bit aggregate,* which supports much larger sizes than the 32-bit limit of 16TB. FlexVol volumes created in 64-bit aggregates support much larger volumes, ranging in size from 30TB to 100TB, based on the storage system. For more information, refer to the following documents:

- System configuration guides for maximum supported 8.x volume and aggregates sizes for specific hardware platforms.
- TR-3786: A Thorough Introduction to 64-Bit Aggregates.

Benefits

Using 64-bit aggregates offers the following benefits:

- **Better performance from more disks in the aggregate.** When the disks are the bottleneck in improving performance, using 64-bit aggregates with a higher spindle count can boost performance.
- **Ability to add disks to an aggregate in fully populated RAID groups.** Using 64-bit aggregates provides maximum storage efficiency, while also offering all the data protection benefits of NetApp RAID-DP.
- **Ability to add more disk shelves into an aggregate.** Because of the larger size of 64-bit aggregates, storage administrators can manage a system with fewer aggregates, thus reducing the overhead of managing multiple aggregates.
- **All the advantages and features of Data ONTAP.** Space efficiency, thin provisioning, deduplication, FlexVol volumes, and many other features of Data ONTAP can be used on 64-bit aggregates with the same ease and simplicity as on 32-bit aggregates.
- **Bigger aggregate and volume sizes.** Use 64-bit aggregates to build a flexible storage solution with improved scalability and ease of management as your storage needs grow.
- **Easier data management.** Use 64-bit aggregates for applications that store and access huge amounts of data. The larger-size thresholds of FlexVol volumes in 64-bit aggregates provide the ability to store all the data in a single FlexVol volume, making the data easier to manage.

A 64-bit aggregate containing the root volume is created during the Data ONTAP setup process. To create additional 64-bit aggregates, determine the aggregate name, the node on which to create it, and how many disks it will contain.

Execute the following command to create a new aggregate.

```
aggr create -aggregate new_aggr -nodes node_name -B 64 -diskcount num_of_disks
```

Note: Leave at least one disk (select the largest disk) in the configuration as a spare. A best practice is to have at least one spare for each disk type and size.

8.2 Vservers

NetApp clustered Data ONTAP supports multiple types of virtual storage servers (Vservers). This section describes their uses.

Types of Vservers in Data ONTAP 8.1

Data ONTAP 8.1 Vservers include cluster admin Vservers, cluster Vservers, and node Vservers.

Cluster Admin Vserver

The cluster admin Vserver is used solely for administrative access to the cluster deployment. Storage clients do not access a cluster admin Vserver. Cluster admin Vserver features and limitations include:

- One Vserver per cluster
- Clusterwide scope and authority
- Does not own volumes or LUNs
- Cannot export or share file systems or provide access to LUNs

Cluster Vserver

A cluster Vserver provides SAN and/or NAS access to storage clients in a secure multi-tenant fashion. To storage clients, a cluster Vserver looks like a separate entity, but it is actually a virtualized storage controller. All storage access is provided at the cluster Vserver level. Cluster Vserver features and limitations include:

- One or more Vservers per cluster
- SAN or NAS data services require at least one Vserver
- Scope limited to the objects it owns
- Owns a set of clustered Data ONTAP volumes and LUNs and uses them to provide SAN or NAS data access
- Directly accessible by means of its own optional Vserver management LIF

Create a Vserver

1. Run the Vserver setup wizard.

```
Welcome to the Vserver Setup Wizard, which will lead you through
the steps to create a virtual storage server that serves data to clients.

You can enter the following commands at any time:
"help" or "?" if you want to have a question clarified,
"back" if you want to change your answers to previous questions, and
"exit" if you want to quit the Vserver Setup Wizard. Any changes
you made before typing "exit" will be applied.

You can restart the Vserver Setup Wizard by typing "vserver setup". To accept a default
or omit a question, do not enter a value.
Step 1. Create a Vserver.
You can type "back", "exit", or "help" at any question.
```

2. Enter the Vserver name.

```
Enter the Vserver name:
```

3. Select the Vserver data protocols to configure. If you aren't sure, select all of them.

```
Choose the Vserver data protocols to be configured {nfs, cifs, fcp, iscsi}:
```

4. Select the Vserver client services to configure. If you aren't sure, select all of them.

```
Choose the Vserver client services to configure {ldap, nis, dns}:
```

5. Enter the root volume aggregate of the Vserver.

```
Enter the Vserver's root volume aggregate {aggr01, aggr02} [aggr02]:
```

6. Enter the Vserver language setting. English is the default language [C].

```
Enter the Vserver language setting, or "help" to see all languages [C]:
```

7. Answer No to Do you want to create a data volume?

```
Vserver creation might take some time to finish....

Vserver test01 with language set to C created. The permitted protocols are
nfs, cifs, fcp, iscsi.

Step 2: Create a data volume
You can type "back", "exit", or "help" at any question.

Do you want to create a data volume? {yes, no} [yes]:
```

8. Enter No for all the remaining questions regarding service configuration. Individual services can be configured later.

Note: As an alternative to the individual configuration steps, create a Vserver by using a one-line command, as shown in the following example.

```
vserver create vserver_name -rootvolume name_of_rootvol -aggregate name_of_aggr -ns-switch file -
language en_US -rootvolume-security-style unix -snapshot-policy none
```

Node Vserver

A node Vserver exists only to provide direct access to a physical node. Node Vserver features include.

- Managing the physical nodes in the cluster, if needed
- One Vserver per controller
- Identifying by `<node_name>` and D-blade UUID
- Owning one or more 7-Mode volumes (including the root volume) and zero or more qtrees on its controller

8.3 FlexVol for Clustered Data ONTAP

FlexVol Limits and Maximums

Refer to NetApp product documentation associated with the specific version of Data ONTAP for a detailed explanation of product limits and scaling considerations that are consistent with various storage efficiency and other optional settings.

FlexVol Default Snap Reserve

Be sure to take into account the storage space that snap reserve can consume (20% of the volume by default when a FlexVol volume is created).

Thin Provisioning

Thin provisioning is covered in section 8.5, “Thin Provisioning NAS with Clustered Data ONTAP.”

Create a FlexVol Volume in Clustered Data ONTAP

The following information is required to create a flexible volume: the name and size of the volume and the aggregate on which it will exist. For example, to create a volume called `vol_name` on `aggregate aggr_name` with a size of 10GB, run the following command.

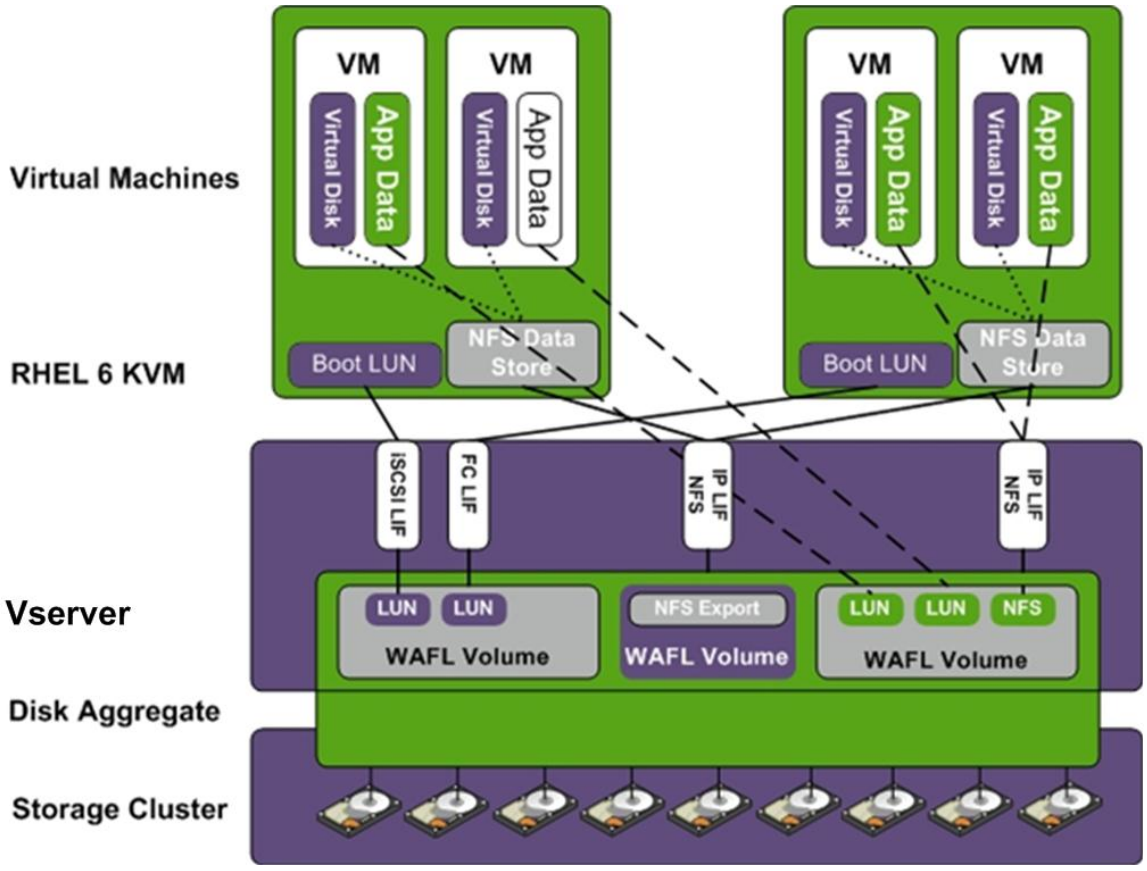
```
volume create -vserver vserver_name -volume vol_name -aggregate aggr_name -size 10g -state online
-type RW -policydefault -snapshot-policy none
```

8.4 LUN with Clustered Data ONTAP

Overview

Figure 10 shows how LUNs relate to volumes and Vservers and how they can be accessed through iSCSI or FC LIFs. This diagram also shows how a LUN serves as a VM or an application data container.

Figure 10) LUNs mapping to hosts.



Clustered Data ONTAP LUNs and Root Volume

LUNs cannot be created on the root aggregate of the node. NetApp does not recommend creating LUNs on the root volume of the Vserver.

Clustered Data ONTAP LUNs and Snapshot Copies

NetApp recommends grouping LUNs in a single volume with similar Snapshot requirements to facilitate the management of Snapshot copies.

Clustered Data ONTAP LUNs and Space Guarantees

NetApp recommends thin provisioning all LUNs, which means setting no space guarantees. Enable space guarantees only it is critical that writes to a LUN always complete.

Clustered Data ONTAP LUNs and Fractional Reserve

Fractional reserve controls the amount of space a volume reserves for overwrites to space-reserved LUNs when the volume has filled. NetApp recommends setting the fractional reserve to 0%.

8.5 Thin Provisioning NAS with Clustered Data ONTAP

Setting Up NAS Thin Provisioning

When thin provisioning is enabled, primary data and space for the associated Snapshot copy are allocated on demand. This variant achieves the best ratio of storage efficiency for provisioning applications from scratch. NetApp recommends that customers choose the thin-provisioning method to increase storage efficiency. Thin provisioning follows a 100% allocate-on-demand concept.

The thin-provisioning method has the following characteristics:

- Volumes are created without a space guarantee.
- The size of the volume follows the formula X + Δ, where:
  - X is the size of the primary data (sum of all user files and directories in the volume).
  - Δ is the amount of space needed to hold Snapshot copy data.
  - Sizing the volume defines a container with a virtual size for the consumers. NAS users are familiar with fixed-sized file shares and with the following considerations:
- Space used for Snapshot copies can grow unexpectedly. Administrators can use the autosize function to make space available when reaching a certain volume threshold or when the space reserved for user data becomes low.
- Space reserved for Snapshot copies is used to hide the capacity taken up by Snapshot copies from the consumers (NAS clients).
- For volumes with deduplication enabled, volume autogrow is a mandatory option.
- NetApp does not normally recommend using the autodelete option in NAS environments. Reserving a certain amount of space for Snapshot copies for file versioning or file restores is part of the SLAs defined for file services.

Table 6 lists thin-provisioning volume options.

Table 6) Thin-provisioning volume options.

Volume Option	Recommended Value	Notes
space-guarantee	none	This is the key setting for thin provisioning.
fractional-reserve	0	The default is 0%.
autosize	on	Set autosize to on. No artificial limited volume must be monitored. The autosize function allows the user data to grow beyond the guaranteed space limit.
space-mgmt-try-first	volume_grow	Increasing the size of the volume does not destroy any data or information. Therefore the volume size can be increased or decreased as needed. For some configurations, automatic volume growth might not be desired.

Table 7 lists thin-provisioning volume Snapshot options.

Table 7) Thin-provisioning volume Snapshot options.

Volume Snapshot Option	Recommended Value	Notes
percent-snapshot-space	0	The value depends on the number of Snapshot copies and the change rate within the volume. Displaying only the committed usable space using the SLA is the preferred way to provision NAS storage. However, in some situations, the Snapshot copy reserve area might be omitted.
autodelete	false	Deleting Snapshot copies is not recommended in most NAS environments, so that data is not lost when SnapManager® products are being used. However, if running out of space is deemed more critical than losing Snapshot copies, then change the value to true to enable this option.

Thin Provisioning NFS with Clustered Data ONTAP

Run the following commands to modify the volume options and the volume Snapshot options. These commands set the default FlexVol volume to be thin provisioned.

```
vol modify -vserver vserver_name -volume vol_name -space-guarantee none -fractional-reserve  0 -
autosize true -max-autosize (2*SIZE)g -space-mgmt-try-first volume_grow -percent-snapshot-space 0
vol snap autodelete modify -vserver vserver_name -volume flexvol_name -enabled false
```

Note: Setting the autosize option to two times the original size of the volume allows the volume to double its original size by using more space from the aggregate. It is important to understand the consequences of this setting and monitor free space in all aggregates.

Volume Size Considerations

Because physical allocation of data in a thin-provisioned volume is done on demand, theoretically, the volume size can be set to a very high value, to easily store all application data and Snapshot copies. All other applications can benefit from the shared pool of unallocated storage, because the unallocated space in the volume is not exclusively reserved for the volume itself.

NetApp recommends that customers size the volume to the expected size of its containing objects and use the autogrow option to let it grow on demand. The advantage of this method is that the commitment rate acts as a metric for data consolidation.

Note: The commitment rate reflects the amount of logical data consolidation. This metric is suitable for deciding when to leave space for organic growth.

In addition, the volume size limits should be taken into account when using deduplication because the maximum sizes depend on the storage controllers.

Requirement for Storage Monitoring

A high degree of data consolidation can be achieved by using thin provisioning. Because this effect depends on the usage characteristics of the corresponding applications, monitoring the aggregate is critical.

8.6 Deduplication with Clustered Data ONTAP

NetApp deduplication provides block-level deduplication in the entire flexible volume. Essentially, deduplication removes duplicate blocks, storing only unique blocks in the flexible volume, and creates a small amount of additional metadata in the process.

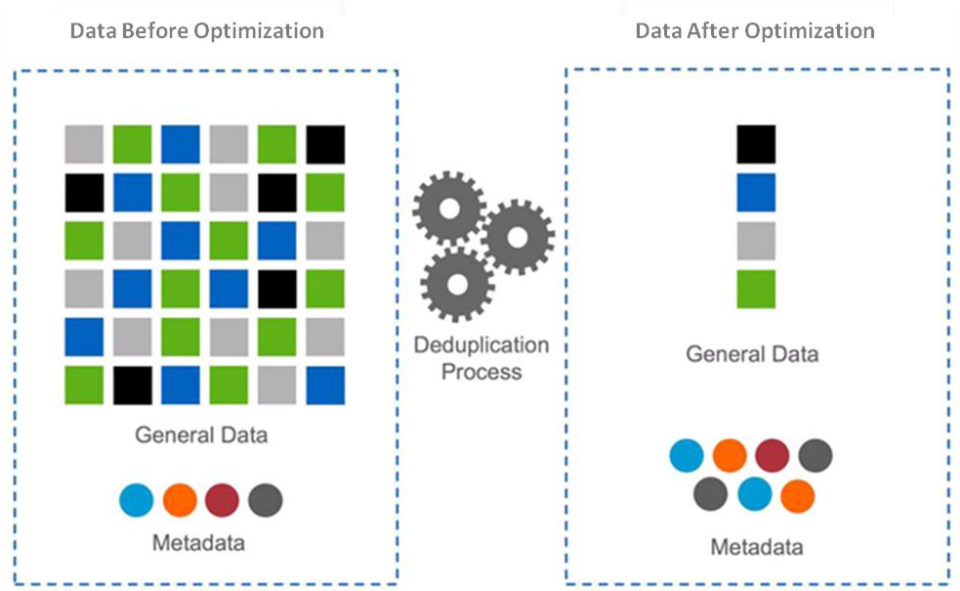
Overview

NetApp deduplication has the following characteristics:

- It works with a high degree of granularity at the 4KB block level.
- It operates on the active file system of the flexible volume. Any block referenced by a Snapshot copy is not made available until that Snapshot copy expires.
- It is a background process that can be configured to run automatically; be scheduled; or be run manually through the CLI, OnCommand System Manager, or OnCommand Provisioning Manager.
- It is application transparent and therefore can be used for the deduplication of data originating from any application that uses the NetApp system.

Figure 11 shows the NetApp deduplication process at the highest level.

Figure 11) NetApp deduplication process at the highest level.



Deduplication and Compression

Data ONTAP 8.1 and later offer deep integration of compression and deduplication. Deduplication must first be enabled on a volume before compression can be enabled; compression cannot be enabled without deduplication.

Deduplication Limitations

A maximum of eight deduplication processes can run concurrently on eight different volumes in the same NetApp storage system. If an attempt is made to run an additional deduplication process beyond the maximum number, the additional operations are placed in a pending queue, and they are automatically started when free processes become available.

Although deduplication can provide substantial storage savings in many environments, it is associated with a small amount of storage overhead. In Data ONTAP 8.1, the deduplication metadata for a volume continues to be located inside the aggregate; however, a copy of this metadata is stored in the volume. The guideline for the amount of additional space that should be left in the volume and aggregate for the deduplication metadata overhead is as follows:

- **Volume deduplication overhead.** For each volume with deduplication enabled, up to 4% of the logical amount of data written to that volume is required to store volume deduplication metadata.
- **Aggregate deduplication overhead.** For each aggregate that contains any volumes with deduplication enabled, up to 3% of the logical amount of data contained in all of those volumes with deduplication enabled is required to store the aggregate deduplication metadata.

Enable Deduplication in Clustered Data ONTAP

1. To enable deduplication on a volume, run the following volume efficiency command.

```
volume efficiency on –vserver vserver_name -volume vol_name
```

2. To start deduplication manually, run the following volume efficiency command.

```
volume efficiency start –vserver vserver_name -volume vol_name
```

8.7 NFSv3 with Clustered Data ONTAP

NFSv3 Overview

NFS environments manage security at a high level in the following ways:

- NFS exports data to a client system. NFS relies on the client to authenticate its own users. Given the complete local control a user might have over a client system, the NFS security model can be subverted relatively easily.
- NFS is derived from UNIX® and therefore implements a UNIX style set of file permissions that are limited to three categories of users (user, group, and other) and three types of file permissions (read, write, and execute).
- An NFS file server can apply various options (for example, restricting all access to read-only access) when exporting. An NFS client has similar options when mounting from an NFS file server.
- A NetApp system controller can export to specified NFS clients, netgroups (lists of clients), or even network subnets. Exporting by subnet is not always available on conventional NFS file servers from other vendors.

With UNIX NFS clients, file systems are mounted from the server to be logically indistinguishable from local file systems. This can happen without any user activity such as starting up the client, for example. The NFS server grants or denies the mount request based on the host name of the client as dictated by the restrictions specified. No password or other authentication is required by an NFS server.

User authentication is performed by the client, not the server. The client-authenticated user's UID must be passed across the network with all NFS requests sent to the server. Therefore, if a user can convince the NFS client of the user name identity associated with the UID `n`, then the NFS server accepts that any NFS requests issued by that client on that user's behalf are for UID `n`. One exception to this rule occurs when `n` corresponds to the super user (`root`). In this case, the NFS server assigns unauthenticated status (the user `nobody`) to the NFS session that was established, unless root access privileges have been assigned to that client.

Each time an export rule is loaded into memory, it undergoes a straightforward conversion from a string to an entry in a hash table. Here is the basic string format.

```
            Policy          Rule    Access   Client                RO
```

```
Vserver      Name            Index   Protocol Match                Rule
------------ --------------- ------  -------- -------------------- ---------
vs2_nfs4     default         1       any      0.0.0.0/0            any
vs2_nfs4     nfs4_policy1    1       any      0.0.0.0/0            any
```

Enable NFSv3 in Clustered Data ONTAP

1. Run all commands to configure NFS on the Vserver.
2. Install the NFS license.

```
license add -license-code license_key
```

8.8 Fibre Channel with Clustered Data ONTAP

FC Fabric Design Recommendations

NetApp requires the use of multipathing in a Fibre Channel (FC) environment when using clustered Data ONTAP. NetApp also requires Asymmetric Logical Unit Access (ALUA) with clustered Data ONTAP so that multipath selection and cluster operations such as cluster failover and nondisruptive volume movement work correctly. ALUA is automatically enabled on the storage.

Two general topologies are available in an FC environment:

- **Single fabric.** All ports of the initiator and target connect to a single switch or a fabric of switches.
- **Multifabric**. Some ports of the initiator and/or target connect to separate fabrics for redundancy.

All of these configurations include ALUA and are detailed in the Fibre Channel and iSCSI Configuration Guide. This guide includes diagrams and supported topologies for different NetApp platforms.

NetApp recommends using redundant components for any SAN solution to reduce or eliminate single points of failure. Therefore, use multiple HBAs, switches or fabrics, and storage clustering for FC.

Refer to the Fibre Channel and iSCSI Configuration Guide for the switch vendor maximum hop count.

Cascade, mesh, and core-edge fabrics are all industry-accepted and supported methods of connecting FC switches into a fabric.

Enable FC with Clustered Data ONTAP

1. License FCP.

```
system license add -license-code license_key
```

2. Create the FC service on each Vserver, if required.. This command also starts the FC service and sets the FC alias as the name of the Vserver.

```
fcp create -vserver node_name
```

3. Start the FC service on each Vserver, if required.

```
fcp start -vserver node_name
```

4. Verify whether the FC ports are targets or initiators.

```
node run -node node_name fcadmin config
```

5. Make an FC port into a target to allow connections into the node, if required.

   Note: Only FC ports that are configured as targets can be used to connect to initiator hosts on the SAN.

6. For example, to convert a port called `fc_target`, use the following syntax.

```
node run -node node_name fcadmin config -t target fc_target
```

Note: If an initiator port is made into a target port, a reboot is required. NetApp recommends rebooting after completing the entire configuration because other configuration steps might also require a reboot.

NetApp Host Utilities

NetApp provides a SAN host utilities kit for every supported OS. This set of data collection applications and configuration scripts includes SCSI and path timeout values and path retry counts. Also included are tools to improve the supportability of the host in a NetApp SAN environment.

8.9 iSCSI with Clustered Data ONTAP

Overview

This solution enables the iSCSI protocol on a Vserver and sets the default policy to `deny`. NetApp requires the use of multipathing in an iSCSI environment when using clustered Data ONTAP. NetApp also requires ALUA with clustered Data ONTAP so that multipath selection and cluster operations such as cluster failover and nondisruptive volume movement work correctly. ALUA is automatically enabled on the storage.

All configurations include ALUA and are detailed in the Data ONTAP 8.1 SAN Configuration Guide for Clustered Data ONTAP. This guide includes diagrams and supported topologies for different NetApp platforms.

Solution Notes

- This solution constitutes a basic enablement of iSCSI on a Vserver.
- The iSCSI alias defaults to the name of the Vserver. Having the same name for the iSCSI alias and the Vserver makes it easier to identify the correct controller when selecting a target from among the iSCSI initiator interfaces.
- All deployment steps must be run on all Vservers that share data through iSCSI.
- NetApp recommends the use of password-protected CHAP entries for each iSCSI host that must authenticate with the storage system. Password-protected session authentication prevents iSCSI LUNs from being mounted through spoofing attacks over the network. Without password-protected session authentication, a storage environment with iSCSI enabled is not compliant with security regulations such as HIPAA, PCI DSS, and Sarbanes-Oxley.

Enable iSCSI with Clustered Data ONTAP

The following steps configure the iSCSI service on a Vserver. These steps do not include any host-specific configuration tasks.

1. From the cluster shell, license the iSCSI protocol.

```
system license add -license-code license_key
```

2. Create the iSCSI service on each Vserver, if required. This command also starts the iSCSI service and sets the iSCSI alias to the name of the Vserver.

```
iscsi create -vserver node_name vserver_name
```

3. Start the iSCSI service on each Vserver, if required.

```
iscsi start -vserver node_name
```

8.10 LUN Creation with Clustered Data ONTAP

After the iSCSI and/or FCP licenses are installed, LUNs can be created in a volume on a Vserver.

1. To create a 2GB LUN for RHEL in volume `vol_name` called `lun_name` on Vserver `vserver_name`, execute the following command.

```
lun create –vserver vserver_name -volume vol_name -lun lun_name –size 2g –ostype linux
```

2. To provide access to LUN, an initiator group (igroup) must be created and then mapped to the LUN. Create the igroup by executing the following command.

```
igroup create -igroup igroup_name -protocol iscsi -ostype linux -initiator initiator01
```

3. The protocol option must be fcp or iscsi, depending on the requirement. The initiator names are either WWPN (FC or FCoE) or IQN (iSCSI initiator names).
4. Map the LUN to the igroup by executing the following command.

```
lun map -volume vol_cmode_nfs -lun testlun -igroup bob -lun-id 0
```

9 Storage Network Best Practices for RHEV

9.1 Storage Architecture Concepts

Production Ethernet Storage Networks

The goal of any storage network is to provide uninterrupted service to all nodes that are connected to it. This section focuses primarily on how to establish a highly available Ethernet storage network.

Regardless of storage protocol, a storage network must address the following three goals:

- Be redundant across switches in a multiswitch environment
- Use as many available physical paths as possible
- Be scalable across multiple physical interfaces or ports

10GbE for Ethernet Storage

NetApp Data ONTAP, Red Hat Enterprise Linux, and Red Hat Enterprise Virtualization provide support for 10GbE. An advantage of 10GbE is the ability to reduce the number of network ports in the infrastructure; especially (but not limited to) blade servers. Additionally, 10GbE can handle several VLANs simultaneously. It is a NetApp best practice to use 10GbE, especially for storage.

9.2 IFGRP LACP with Clustered Data ONTAP

Overview

Dynamic multimode interface groups are based on Link Aggregation Control Protocol (LACP), as described by the IEEE 802.3ad (dynamic) standard. They are similar to static multimode interface groups that contain two or more active interfaces, share a single MAC address, and require configuration on both ends of the connection. However, in a dynamic multimode interface group, additional negotiation parameters are passed between the devices using an LACP protocol data unit (PDU). This allows the two connected devices to dynamically add and remove links from the channel for reasons other than physical link failure. This is an important distinction, because dynamic multimode interface groups can detect and compensate for a lost link, and also for loss of data flow. This feature makes dynamic multimode interface groups compatible with HA environments.

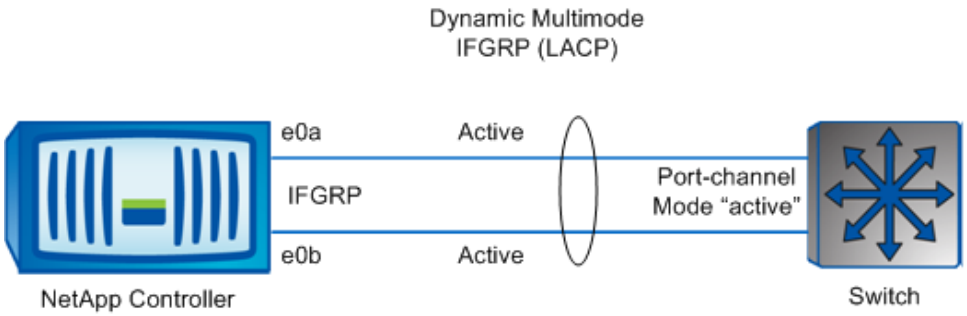
Dynamic multimode interface groups can continue operating even if all but one link has been lost. This allows higher throughput than a single-mode interface group and still provides redundancy. Multiple methods of load balancing are available for outgoing packets from the NetApp FAS device, including:

- **IP address load balancing.** This method uses a fast-hashing algorithm on the source and destination IP addresses to equalize traffic on multimode interface groups.
- **MAC address load balancing.** This method uses a fast-hashing algorithm on the source and destination MAC addresses to equalize traffic on multimode interface groups.
- **Round robin.** This method is normally used for load balancing a single connection's traffic across multiple links to increase single-connection throughput.
- **Port (Data ONTAP 7.3.2 and later).** This method uses a fast-hashing algorithm on the source and destination IP addresses, along with the transport layer port number. Port-based load balancing adds a second hash for every unique source and destination port, which can result in more efficient load balancing of traffic over LACP links. NetApp recommends this distribution function in mixed workload environments.

Dynamic multimode interface groups must be connected to a switch that supports LACP.

Figure 12 shows the dynamic multimode interface group (LACP) configuration.

Figure 12) Dynamic multimode interface group (LACP).



Advantages

Dynamic multimode interface groups are compatible with an HA environment because they can detect the loss of link status and also a loss of data flow.

The advantages are:

- LACP allows higher aggregate bandwidth based on the load-balancing algorithm chosen, because all ports in the channel are active.
- The LACP PDUs that are used enable the dynamic multimode interface group to detect a loss of link on either side and to detect a loss of data flow. This avoids queue wedge and traffic black hole problems.
- Many newer switches support multichassis LACP, allowing interface groups to be connected to multiple switches. This setup increases redundancy across the switch infrastructure.

Disadvantages

Dynamic multimode interface groups might not be compatible with some environments.

The disadvantages are:

- Older switches might not support the LACP standard.

- Older switches might not support multichassis LACP, which means that LACP interface groups would be limited to a single switch.

This type of interface group requires two or more Ethernet interfaces and a switch that supports LACP. Therefore, make sure that the switch is configured properly.

```
Run the following command on the command line. This example assumes that there are two network
interfaces called e0a and e0b.ifgrp create -node node_name -ifgrp name_of_vif -distr-func ip -
mode multimode_lacp
network port ifgrp add-port -node name_of_node -ifgrp name_of_vif -port e0a
network port ifgrp add-port -node name_of_node -ifgrp name_of_vif -port e0b
```

Note: All interfaces must be in the down status before being added to an interface group.

Note: The interface group name must follow the standard naming convention of `x0x`.

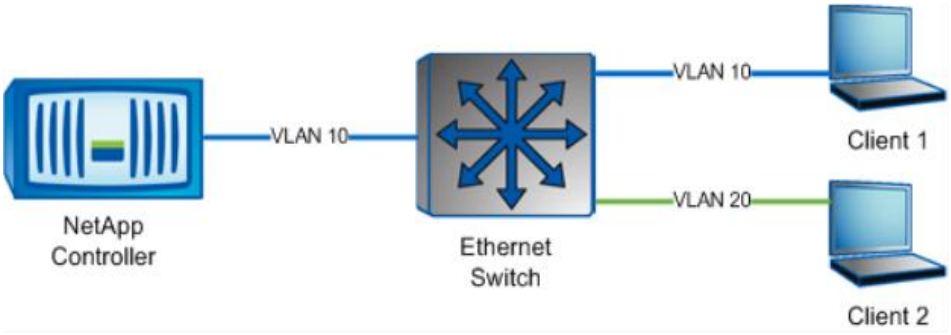
9.3 VLANs with Clustered Data ONTAP

Overview

VLANs operate at the data link layer (L2) of the OSI model; therefore, traffic is completely isolated between VLANs unless a bridge or a router (L3) is used to connect the networks. Although exceptions exist, a one-to-one relationship between an IP subnet and a VLAN helps simplify the network and facilitates management.

Figure 13 shows an example of VLAN connectivity.

Figure 13) VLAN connectivity example.



Implementing VLANs offers the following advantages in the network.

- **Higher utilization of switch infrastructure.** Allowing separate logical devices to live on the same physical switch eliminates the need for a completely separate hub or switch for each network. Higher density and more efficient switches can be used to aggregate devices.
- **Lower management costs.** Moving physical devices or cables is no longer a requirement. An administrator can logically assign devices from the management console of the switch to different networks, if required.
- **Security and stability of the network.** An L3 device is required to communicate between VLANs; therefore, L3 access lists can be applied to prevent communication between certain networks. Broadcast storms and the effects of unicast flooding become isolated to a single VLAN. A malicious user with a packet capture tool will not be able to intercept traffic that is not destined for that user's host.

NetApp recommends deploying VLANs to separate data networks from storage networks.

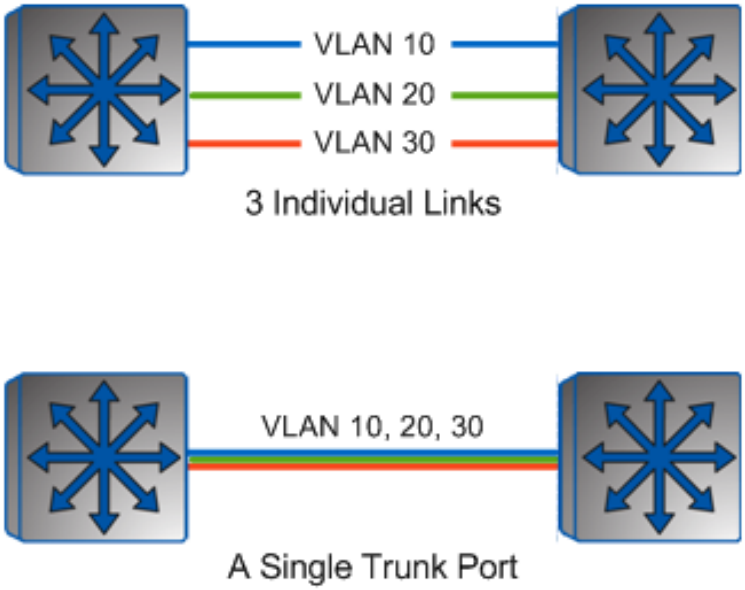
VLAN Trunking

Organizations typically require the use of multiple switches for redundancy and port density. Often, a logical VLAN might be required across multiple switches. A standard access port configuration allows only a single VLAN to be defined on a port. However, multiple ports, each assigned to a single VLAN, are required to pass VLANs across switches. This method does not scale well and is highly inefficient. The IEEE 802.1q standard offers a solution to this problem with a feature called VLAN trunking.

VLAN trunking allows a single link to carry multiple VLANs by tagging each packet with a 4-byte tag. This tag defines the VLAN to which each packet is assigned, as it travels throughout the network between VLAN-aware devices. Common VLAN-aware devices are network switches, routers, certain servers, and NetApp storage controllers. When the frame reaches an endpoint or access port, the VLAN tag is removed, and the original frame is then sent to the end device. The tag is simply a forwarding instruction to make sure that the frame is delivered to the proper broadcast domain or VLAN.

Figure 14 illustrates VLAN trunking.

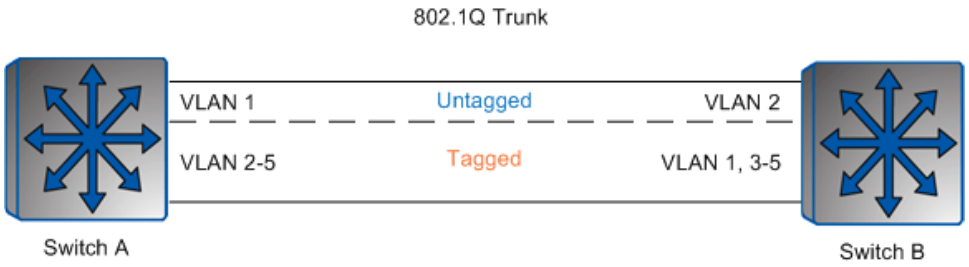
Figure 14) VLAN trunking.



Native VLAN is an important feature that is used in 802.1q trunking. Frames assigned to the VLAN, which are configured as the native VLAN, are sent untagged across the trunk link. All other VLANs that are configured in the trunk are tagged with their respective VLAN IDs. For this reason, make sure that the native VLAN configuration is the same on both ends of the connection. Improper configuration of the native VLAN can result in limited or no connectivity for the administratively defined native VLANs.

Figure Figure 15 shows an example of a faulty configuration.

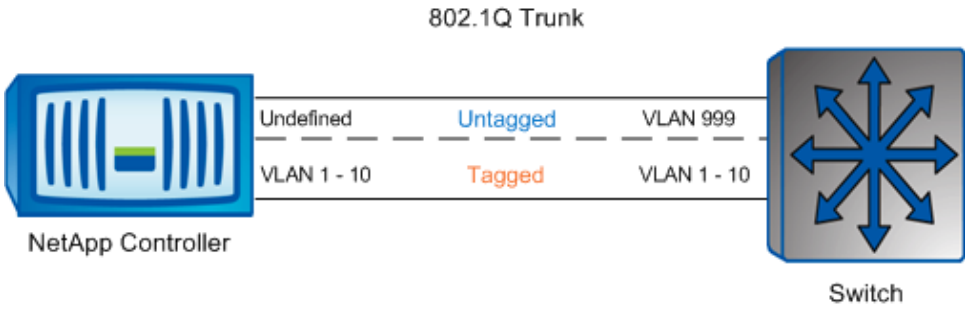
Figure 15) VLAN faulty configuration example.



As mentioned earlier, NetApp storage controllers also provide the ability to configure multiple VLANs and VLAN trunking. This allows greater flexibility and configuration options in the controller itself. For example, a NetApp controller with a 10GB interface might need to serve multiple functions, such as iSCSI boot traffic and standard NAS traffic. The iSCSI boot network might require additional security and control, but because of the relatively low traffic requirement, an administrator might not want to dedicate an entire 10GB link to this function. When VLAN trunking is implemented on the NetApp controller and switch, the single 10GB link can be shared between the two functions, while still maintaining isolation between VLANs. When configuring VLANs on a NetApp controller, make sure that the Data ONTAP operating system does not currently use the native VLAN feature. For this reason, the native VLAN on the switch port connected to a NetApp controller must be assigned to an unused VLAN to help forward data correctly.

Figure 16 shows the native VLAN NetApp configuration.

Figure 16) Native VLAN NetApp configuration.



9.4 Jumbo Frames with Clustered Data ONTAP

Overview

By default, a standard Ethernet frame has a 1,500-byte payload. The Ethernet header and the CRC checksum data add 18 bytes to the Ethernet frame for a total Ethernet frame size of 1,518 bytes, or 1,522 bytes with a VLAN tag (IEEE 802.3ac). Because the header and the CRC checksum create a fixed amount of overhead per packet, sending a larger payload is more efficient. Jumbo frames can be created by changing the MTU from the standard 1,500 bytes up to 9,000 bytes. Larger frames increase the network throughput by reducing the number of packets processed for the same amount of data.

It's important to implement jumbo frames carefully in an Ethernet storage network. An incorrect design or configuration can result in poor performance and even limited or no connectivity at all. When configuring a NetApp FAS storage controller for jumbo frames, make sure that the following elements are properly configured:

- The NetApp storage device's network ports and any associated IFGRPs or VLAN network ports
- Individual ports and port-channel interface on the external Ethernet switch, if applicable
- VLAN and ports on all switches and layer 3 routers between the FAS device and the clients

Ports with a standard MTU size and ports with a jumbo MTU size should never be mixed on the same VLAN. For example, consider a host and a storage controller that are configured on the same VLAN, where the FAS controller is configured for jumbo frames and the host is not. The host can communicate with the FAS device by using the standard 1,500-byte frames, but the reply from the FAS device will be in 9,000-byte frames. Because the two machines are located on the same VLAN, a device does not fragment the FAS device frames into the standard 1,500-byte size for consumption by the host.

To allow the NetApp storage controller to support both standard and jumbo frame requests from hosts, one option is to place a router in between the FAS device and the hosts, because routers are able to fragment jumbo frames into smaller 1,500-byte increments. Devices that can use jumbo frames are placed onto a separate VLAN that is configured to pass jumbo frames directly to the NetApp storage controller. However, hosts that can only accept standard frames are placed onto a VLAN whose traffic is passed through a router for fragmentation. This configuration allows all hosts to properly communicate with the NetApp storage controller.

Another method is to directly connect VLANs for standard frame traffic and for jumbo frame traffic to separate ports on the NetApp storage controller. This method has the advantage of allowing traffic (with the DF bit value set to true) to always reach its destination. The following scenarios make use of dedicated VLANs:

- **Local management VLAN (MTU 1,500).** Supports SNMP, Operations Manager, SSH, RLM, and so on. Storage traffic never runs across this network.
- **Storage traffic (MTU 9,000).** Isolated, nonrouted VLAN for NFS, CIFS, or iSCSI data.
- **Replication network (MTU 9,000).** Isolated, nonrouted VLAN for high-speed storage replication such as SnapMirror and SnapVault® data. Separating this traffic allows more granular monitoring and the ability to support different WAN MTU sizes, depending on the links used.
- **Intersite replication (MTU 1,500 or lower).** Useful for off-site backups where the required WAN connections have different MTU values.

Jumbo Frame Recommendations

Using jumbo frames in an Ethernet storage network can significantly increase performance. Complete these steps to improve performance:

1. Configure jumbo frames throughout the network, from the Ethernet storage controller to the host.
2. Segment traffic with jumbo frames onto a different VLAN to achieve optimal network interface performance.
3. To configure a clustered Data ONTAP network port to use jumbo frames (which usually have an MTU of 9,000 bytes), run the following command from the clustershell.

```
network port modify –node node_name -port <network_port> -mtu 9000
WARNING: Changing the network port settings will cause a serveral second interruption in carrier.
Do you want to continue? {y|n}: y
```

The network port identified by `-port` can be a physical network port, a VLAN network port, or an IFGRP.

9.5 Firewall for Clustered Data ONTAP

Each LIF type has an attached default role and firewall policy. Table 8 lists these default firewall policies.

Table 8) Default firewall policies.

Firewall Policy	DNS	HTTP	HTTPS	NDMP	NTP	SNMP	SSH	Telnet
Cluster	Allow	Allow	Allow	Allow	Allow	Allow	Allow	Allow
Management	Allow	Allow	Allow	Allow	Allow	Allow	Allow	Block
Data	Allow	Block	Block	Allow	Block	Block	Block	Block
Intercluster	Block	Block	Block	Allow	Block	Block	Block	Block

9.6 Failover Groups for NAS with Clustered Data ONTAP

A failover group is a list of network ports that are available for use if a network port failure occurs. A network interface or a LIF can subscribe to a failover group and be automatically configured with a list of failover rules for each physical port in the failover group. As ports are added or removed from the failover group, each LIF subscribed to that failover group is automatically configured to have a set of failover rules that are consistent with the ports present in the group.

Run the following commands in a failover pair to enable storage failover:

1. Enter the cluster license on both nodes.

```
node run first_node_name license add license_key
node run second_node_name license add license_key
```

2. Enable failover on one of the two nodes.

```
storage failover modify -node first_node_name -enabled true
```

Note: Enabling failover on one node enables it on both nodes.

3. Enable HA mode for two-node clusters only.

Note: Do not run this command for clusters with more than two nodes because it will cause problems with failover.

```
cluster ha modify -configured true
```

4. If prompted to enable SFO, select Yes.

9.7 FCP LIF with Clustered Data ONTAP

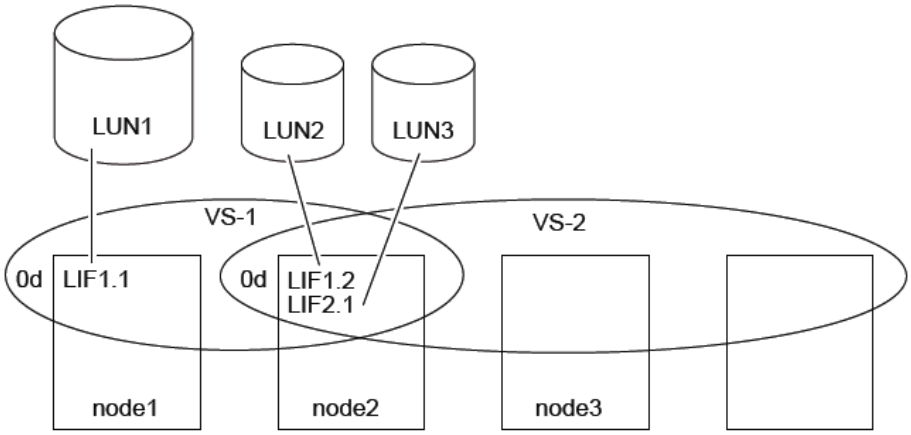
Overview

An FC LIF is a logical interface that belongs to a single Vserver and is assigned to a physical port. An FC LIF has its own unique WWPN and has the WWNN of the Vserver. The FC LIF permits access to LUNs that belong to the Vserver.

Multiple LIFs, whether from the same or different Vservers, can be assigned to a single physical port. In general, FC LIFs use a common ALUA LUN presentation model and rely on MPIO on hosts. For more information, refer to the Data ONTAP 8.1 Clustered Data ONTAP Block Access Management Guide for iSCSI and FC.

Figure 17 shows an example of FC and iSCSI LIFs in a clustered Data ONTAP environment.

Figure 17) FC and iSCSI LIFs in clustered Data ONTAP.



Limits

Table 9 lists the limits for the number of FC LIFs per port and for the FC ports per node.

Table 9) FC LIF limits.

Type	FC LIFs per Port	FC Ports per Node
FC LIF	8	32

Create an FCP LIF

Create FC LIFs named `lif_name_1` and `lif_name_2`. In this example, the cluster consists of two nodes, `node_name_1` and `node_name_2`, and two physical ports, `port_name_1` and `port_name_2`.

```
network interface create -vserver vserver_name -lif lif_name_1 -role data -data-protocol fcp -
home-node node_name_1 -home-port port_name_1
network interface create -vserver vserver_name -lif lif_name_2 -role data -data-protocol fcp -
home-node node_name_2 -home-port port_name_2
```

9.8 iSCSI LIF with Clustered Data ONTAP

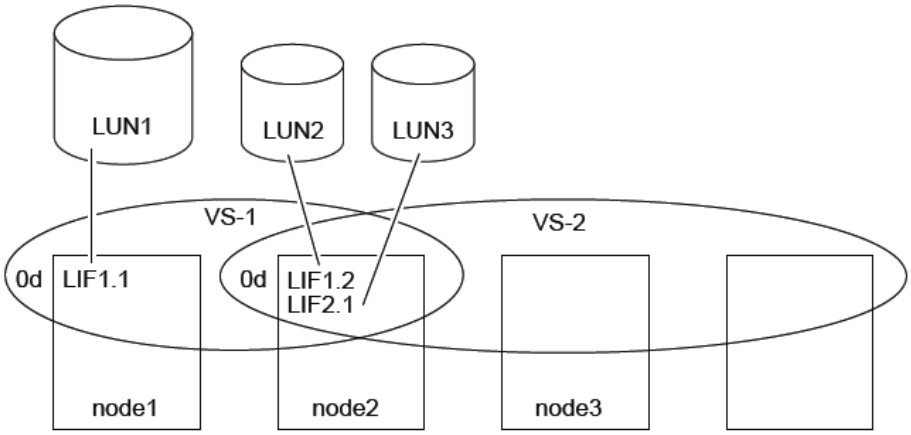
Overview

An iSCSI LIF is an IP-based logical interface that belongs to a single Vserver and is assigned to a physical port. The iSCSI LIF permits access to LUNs that belong to the Vserver.

Multiple LIFs, whether from the same or different Vservers, can be assigned to a single physical port. For more information, refer to the Data ONTAP 8.1 Clustered Data ONTAP Block Access Management Guide for iSCSI and FC.

Figure shows an example of FC and iSCSI LIFs in a clustered Data ONTAP environment.

Figure 18) FC and iSCSI LIFs in clustered Data ONTAP.



Limits

Table 10 lists the limits for the number of IP LIFs per port and for the Ethernet ports per node.

Table 10) IP LIF limits.

Type	IP LIFs per Port	Ethernet Ports per Node
iSCSI LIF	8	24

Create an iSCSI LIF

Create an iSCSI LIF named `lif_name` in the `vserver_name` Vserver, and assign the `ip_address` IP address.

Note: The home node is `node_name`, and the physical port is `port_name`.

```
network interface create -vserver vserver_name -lif lif_name -role data -data-protocol iscsi -
home-node node_name -home-port port_name -address ip_address -netmask netmask_address
```

9.9 Intercluster LIF with Clustered Data ONTAP

Overview

An intercluster LIF is a logical interface that belongs to a single Vserver and is assigned to a physical port. The intercluster LIF permits access to transfer data between Vservers on different clusters. Intercluster LIFs can be configured on data ports or intercluster ports. An intercluster LIF must be created on each node in the cluster before a cluster-peering relationship can be established.

These LIFs can fail over to data or intercluster ports on the same node, but they cannot be migrated or failed over to another node in the cluster.

Limits

Table 11 lists the limits on the minimum number of intercluster LIFs per node if cluster peering is enabled and on the maximum total number of LIFs per node.

Table 11) Intercluster LIF limits.

Type	Minimum Intercluster LIFs per Node If Cluster Peering Is Enabled	Maximum Total LIFs per Node
Intercluster LIF	1	262

Create an Intercluster LIF

To create intercluster LIFs, follow these steps:

1. Create intercluster LIFs on either data ports or intercluster ports. Optionally, change one or more ports to the intercluster role to support the intercluster LIF.

```
network port modify -node node_name -port port_name -role intercluster
```

2. Create an intercluster LIF.

```
network interface create -vserver vserver_name -lif lif_name -role intercluster -home-node
node_name -home-port port_name -address ip_address -netmask netmask_address
```

Repeat step 2 for any additional LIFs.

10 Storage Network Services and Access

10.1 DNS with Clustered Data ONTAP

Domain Name Service

NetApp recommends that DNS be configured and enabled on all Vservers in the cluster.

DNS is a service that is used to convert host names to IP addresses. Host names contain alphabetic or numeric characters and are not case sensitive. Host names can also contain hyphens, but other special characters are not permitted. A fully qualified domain name (FQDN) consists of the host name plus the domain name in the format host-name.domain.com.

The DNS service sends queries to the DNS that stores the information for specific domains and the hosts within those domains. This process enables routing by using the host name instead of the IP address; it is particularly useful when the IP address of various hosts must be changed. Applications that rely on communications to those hosts can continue to access the devices by using the host names without having to modify IP configurations at the application level.

Configure DNS

To configure DNS, run the following command for the Vserver:

```
vserver services dns create -vserver vserver_name -domains domain_name -name-servers dns_server -
state enabled
```

Note: To modify an existing entry, replace the word `create` with `modify` in the command.

10.2 NTP with Clustered Data ONTAP

NTP is a protocol that is used to synchronize the time of any network-attached device to a network time source. In this case, the time between the nodes and the chosen time server is synchronized. Time synchronization allows logs and events to be correctly correlated with other network devices. Time synchronization is also required to connect to an Active Directory server.

Configure NTP

1. Configure NTP for each node in the cluster.

```
system services ntp server create -node node_name -server ntp_server_ip_address
```

2. Enable NTP for the cluster.

```
system services ntp config modify -enabled true
```

10.3 SNMP with Clustered Data ONTAP

Overview

SNMP is used as a general management protocol for polling and asynchronous notifications (traps). SNMP is a critical part of the interface to Operations Manager and must be configured for polling and traps, if Operations Manager is used. Other upstream applications might also use SNMP for performance and fault management.

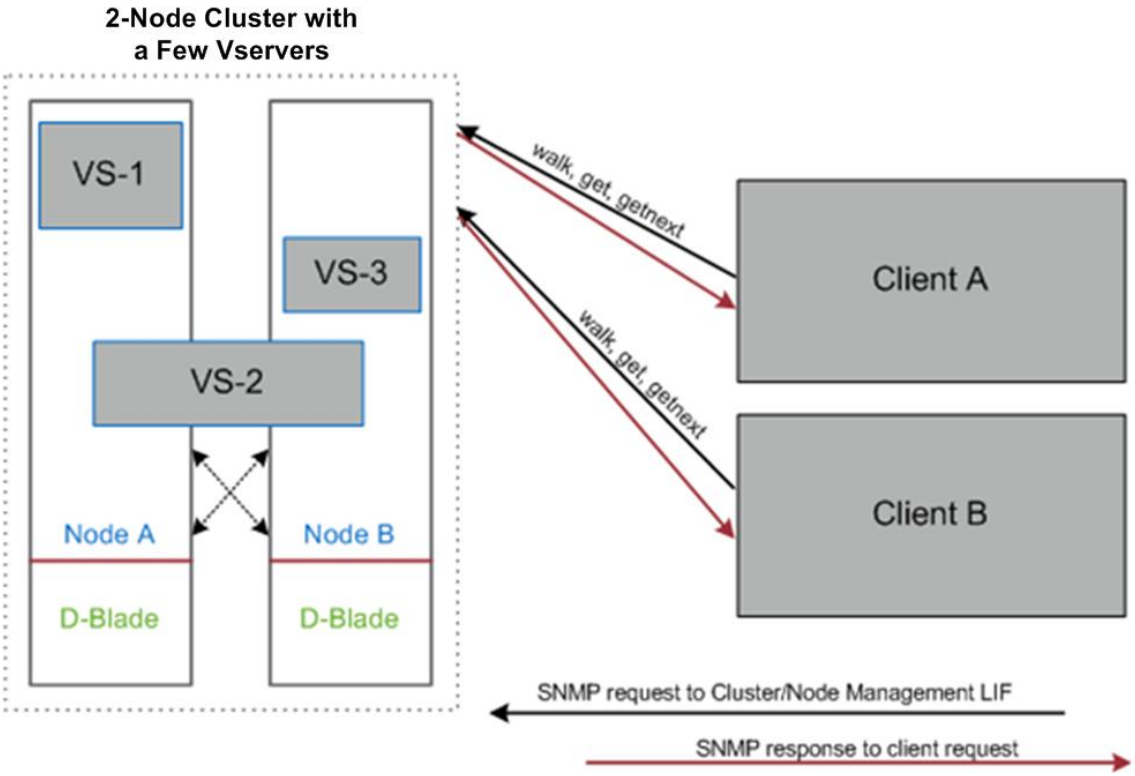
SNMP is a widely used network monitoring and control protocol. Data is passed from SNMP agents, which are hardware and/or software processes reporting activity in each network device (for example, hubs, routers, and bridges), to the workstation console being used to oversee the network. The agents return information contained in a management information block (MIB), which is a data structure that defines what is obtainable from the device and what can be controlled (for example, what can be turned off and on). SNMP originated in the UNIX community and has become widely used on all major platforms.

MIBs describe the structure of the management data of a device subsystem; they use a hierarchical namespace containing OIDs. Each OID identifies a variable that SNMP can read or set.

Note: NetApp does not support SNMP-set operations. Also, SNMP support is only available clusterwide and is not available on a per-Vserver basis. SNMP support will be available on a per-Vserver basis in releases after Data ONTAP Version 8.1.

Figure shows an SNMP diagram in clustered Data ONTAP.

Figure 19) SNMP in clustered Data ONTAP.



SNMP Polling

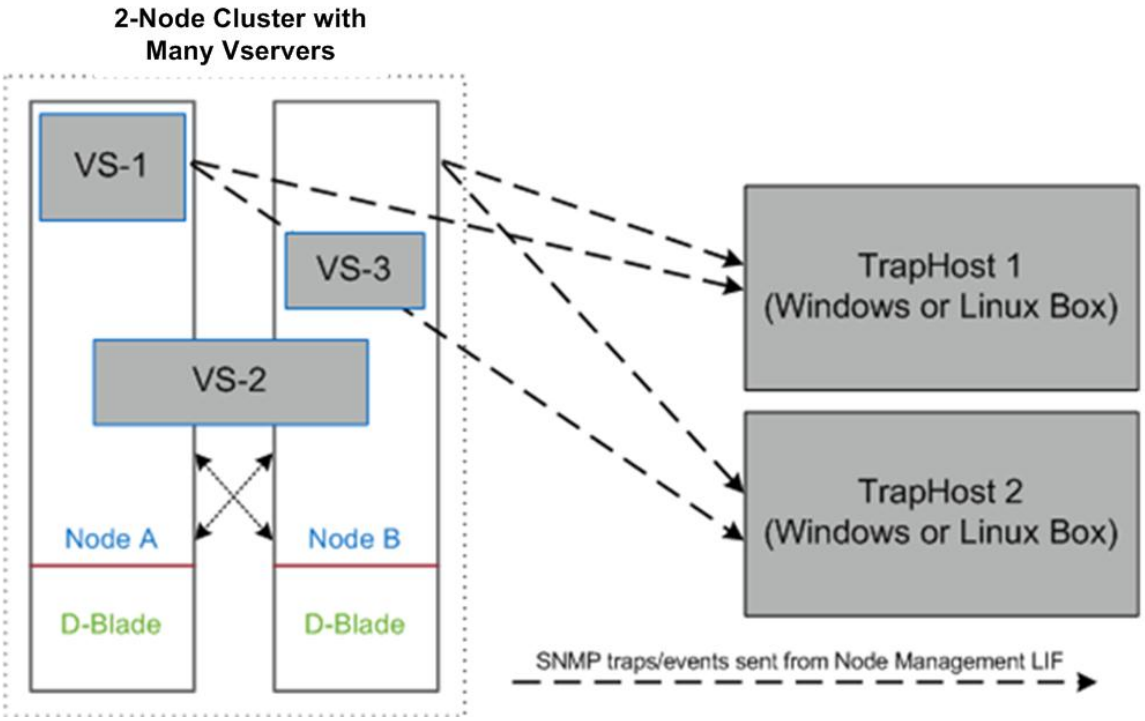
SNMP is used for polling. Data ONTAP supports two versions of SNMP, version 1 and version 3. NetApp recommends using SNMPv3 because it has a stronger authentication mechanism. Earlier versions of SNMP are based on plain text authentication methods, which can be intercepted.

SNMP Traps

SNMP traps are configured to be sent to Operations Manager and other fault management stations. Data ONTAP sends SNMP traps as SNMPv1 traps.

Figure 20 shows SNMP traps in clustered Data ONTAP.

Figure 20) SNMP traps in clustered Data ONTAP.



10.4 AutoSupport HTTPS with Clustered Data ONTAP

AutoSupport™ is used to send configuration summary information and critical event information to NetApp Support. The transport mechanism is over HTTPS. Outgoing secure web connections must be allowed from the storage controllers.

AutoSupport reports can be viewed in the My AutoSupport section of the NetApp Support site.

Configure AutoSupport HTTPS

AutoSupport sends support summary information to NetApp through HTTPS.

Execute the following command to configure AutoSupport.

```
system node autosupport modify -node * -state enable -transport https -support enable -noteto
storage_admin_email
```

10.5 User Access for Clustered Data ONTAP

Overview

Data ONTAP provides several methods for specifying how a user can access the storage system. Use the `-application` parameter of the `security login` command to specify a method. The supported access methods include:

- System console (console)
- HTTP(S) (http)
- Data ONTAP API (ONTAPI®)
- SNMP (snmp)

- SP or RLM (service-processor)
- SSH (ssh)
- Telnet (telnet)

Note: By default, Telnet is hidden and disabled.

If a firewall is enabled, the access method must be added in the firewall policy to allow requests to go through the firewall. For more information, see the system services firewall policy main pages.

Note: Vserver user accounts cannot use console, snmp, service-processor, or telnet as access methods.

Configure User Access for Clustered Data ONTAP

There are two default administrative accounts: admin and diag. The admin account serves the role of administrator and is allowed access using all applications. The best practice is to create a new account and then delete or lock the default admin account.

1. Create a login method for a new administrator from clustershell.

```
security login create –user name new_account -authmethod password -role admin -application ssh
security login create –user name new_account -authmethod password -role admin -application http
security login create –user name new_account -authmethod password -role admin -application
console
security login create –user name new_account -authmethod password -role admin -application ontapi
security login create –user name new_account -authmethod password -role admin -application
service-processor
```

2. Lock the default admin account.

```
security login lock –user name admin
```

Clustered Data ONTAP Authentication Methods for User Accounts

Data ONTAP provides several methods to specify how a user account is authenticated. Use the `-authmethod` parameter of the security login command to specify how a user account is authenticated. The supported authentication methods include:

- SNMP community strings (community)
- Windows Active Directory authentication (domain)
- LDAP or NIS authentication (nsswitch)
- User password (password)
- SSH public key authentication (publickey)
- SNMP user-based security model (usm)

For Windows Active Directory authentication, a CIFS server must be created for the Vserver. Windows domain users must be mapped to Vserver access control roles by using the `security login create` command with the `-authmethod` parameter set to `domain`. To use LDAP or NIS authentication, Vserver users must be mapped to Vserver access control roles using the `security login create` command with the `-authmethod` parameter set to `nsswitch`.

Cluster Administrator Compared to Vserver Administrator

Vserver administrators administer only their own data Vservers. Cluster administrators can administer the entire cluster and its resources. They can also set up data Vservers and delegate Vserver administration to Vserver administrators. Cluster administrators' specific capabilities depend on their access control

roles. By default, a cluster administrator with the admin account name or role name has all of the capabilities for managing the cluster and Vservers.

In contrast, Vserver administrators can only administer their own data Vservers' storage and network resources, such as volumes, protocols, LIFs, and services. Vserver administrators' specific capabilities depend on the access control roles that are assigned by cluster administrators. For more information about Vserver administrator capabilities, refer to the Data ONTAP 8.1 Clustered Data ONTAP Vserver Administrator Capabilities Overview Guide.

Default Cluster Context User Account Roles

The predefined roles for the cluster context are Admin, Readonly, and None, as shown in Table 12.

Table 12) Cluster context default roles and capabilities.

Role	Access Level to Command Directories	Capabilities
Admin	All	All
Readonly	Readonly	Readonly
None	None	None

Default Vserver Context User Account Roles

A Vserver can have its own user and administration authentication domain. A Vserver administrator's management of a Vserver can be delegated.

Table 13 describes the four predefined roles for a Vserver administrator.

Table 13) Vserver context predefined roles and capabilities.

Role	Default Capabilities
vsadmin	• Manage own user account, local password, and public key • Manage volumes, quotas, qtrees, Snapshot copies, FlexCache® files, and files • Manage LUNs • Configure protocols • Configure services • Monitor jobs • Monitor network connections and network interface • Monitor the health of a Vserver
vsadmin-volume	• Manage volumes, quotas, qtrees, Snapshot copies, FlexCache files, and files • Manage LUNs • Configure protocols • Configure services • Monitor network interface • Monitor the health of a Vserver

Role	Default Capabilities
vsadmin-protocol	• Configure protocols • Configure services • Manage LUNs • Monitor network interface • Monitor the health of a Vserver
vsadmin-readonly	• Monitor the health of a Vserver • Monitor network interface • View volumes and LUNs • View services and protocols

Local Account Attributes

Table 14 lists the modifiable account password attributes for clustered Data ONTAP.

Table 14) Account password attributes.

Attribute	Range	Default	Recommended
Length	3–64	8	Minimum 8
Alphanumeric enforcement	Enabled/disabled	Disabled	Enabled
Password history	6	6	
Minimum age	0	0	

Default Administrative Accounts

The two default administrative accounts are admin and diag.

Administrative Account

The admin administrative account serves the role of administrator and has the ability to access all applications. To enable this capability, first create another account (`security login create`) and configure the admin role for each application that the new admin account has permission to use.

Note: Do not use the off-box authentication method for the primary admin account. This is to make sure that a networking problem does not lock the new admin account out of all administrative access.

After the new admin account has been tested, either delete (`security login delete`) or lock (`security login lock`) the account.

Note: The admin console entry cannot be deleted or locked if it is the only account with permissions to that application.

Diagnostic Account

The diag diagnostic account is provided with the storage system; it can be used to perform troubleshooting tasks in the systemshell. The diag account and the systemshell are intended only for low-level diagnostic purposes and should be used only with guidance from technical support. The diag account is the only account that can be used to access the systemshell, through the advanced command `system node systemshell`. Before accessing the systemshell, set the diag account password by

using the `security login password` command. The diag account and the systemshell are not intended for general administrative purposes.

Network Security

Ports are either physical ports (NICs) or virtualized ports, such as interface groups or VLANs. A LIF communicates over the network through the port to which it is currently bound. A LIF is essentially an IP address with the following associated characteristics:

- Role
- Home node
- Home port
- Routing group
- Firewall policy
- Failover policy

Roles for Network Ports

Network ports can have roles that define their purpose and their default behavior. Port roles limit the types of LIFs that can be bound to a port. Network ports can have four roles:

- Node management
- Cluster
- Data
- Intercluster

Note: These four roles can be modified to obtain an optimal configuration.

Node Management Ports

Node management ports are used by administrators to connect to and manage a node. VLANs and interface groups can be created on node management ports. Some platforms have a dedicated management port (`e0M`). The role of these ports cannot be changed, and these ports cannot be used for data traffic. The management policy is applied by default; it allows DNS, HTTP, HTTPS, NDMP, NTP, SNMP, and SSH traffic. It blocks all Telnet traffic.

Cluster Ports

Cluster ports are used for intracluster traffic only. By default, each node has two cluster ports. Cluster ports should reside on 10GbE ports and be enabled for jumbo frames. VLANs and IFGRPs cannot be created on cluster ports. The cluster policy is applied by default; it allows DNS, HTTP, HTTPS, NDMP, NTP, SNMP, SSH, and Telnet traffic. It does not block any traffic.

Data Ports

Data ports are used for data traffic. These ports are accessed by NFS, CIFS, FC, and iSCSI clients for data requests. By default, each node has a minimum of one data port. VLANs and IFGRPs can be created on data ports. They possess the data role by default; however, their port role cannot be modified. The data policy is applied by default; it allows DNS and NDMP traffic. It blocks HTTP, HTTPS, NTP, SNMP, SSH, and Telnet traffic.

Intercluster Ports

Intercluster ports are used for cross-cluster communication. An intercluster port should be routable to another intercluster port or data port of another cluster. The intercluster policy is applied by default; it allows NDMP traffic only. It blocks DNS, HTTP, HTTPS, NTP, SNMP, SSH, and Telnet traffic.

SNMP

Enabling SNMP provides a mechanism to monitor a cluster to avoid issues before they occur, and to respond to issues when they do occur.

SNMP management includes performing the following tasks:

- Enabling SNMP
- Configuring SNMP users
- Configuring SNMP trap hosts for specific events

If SNMP is enabled in Data ONTAP, SNMP managers can query the storage system's SNMP agent for information. The SNMP agent then gathers information and forwards it to the SNMP managers. The SNMP agent also generates trap notifications whenever specific events occur. The SNMP agent on the storage system has read-only privileges that cannot be used for any set operations or for taking a corrective action in response to a trap. SNMP is enabled clusterwide in clustered Data ONTAP.

For diagnostic and other network management services, Data ONTAP provides an SNMP agent that is compatible with SNMP v1, v2c, and v3. SNMPv3 offers advanced security by using pass phrases and encryption; it is a secure protocol when compared to SNMPv1 and SNMPv2c. The SNMPv3 user must be configured to run the SNMP utilities from the SNMP manager.

Use the `security login create` command to create an SNMPv3 user.

The security level options include:

- Authentication, no privacy
- Authentication, privacy
- No authentication, no privacy

When prompted, enter the following information:

- **Engine ID** (default value is local EngineID)
- **Authentication protocol** (none, md5, sha)
- **Authentication password** (minimum eight characters)
- **Privacy protocol** (none, des)
- **Privacy protocol password** (passphrase for encryption)

10.6 HTTPS Access with Clustered Data ONTAP

The storage controller is accessed by using a secure protocol. In this case, the protocol is HTTPS (HTTP over SSL). Other nonsecure access methods such as Telnet and HTTP must be disabled.

Configure HTTPS Access in Clustered Data ONTAP

Secure access to the storage controller must be configured:

1. Increase the privilege level to access the certificate commands.

```
set -privilege advanced
```

2. Generally, a self-signed certificate is already in place. Verify this by using the following command.

```
security certificate show
```

3. If a self-signed certificate does not exist, run the following command as a one-time command to generate and install a self-signed certificate.

```
security certificate create -vserver vserver_name -common-name lab.companyname.com -size 2048 -
country US -state CA -locality Sunnyvale -organization IT -unit Software -email-addr
user@example.com
```

4. Configure and enable SSL and HTTPS access and disable Telnet access.

```
system services web modify -external true -sslv3-enabled true
system services firewall policy delete -policy mgmt -service http -action allow
system services firewall policy create -policy mgmt -service http -action deny -ip-list 0.0.0.0/0
system services firewall policy delete -policy mgmt -service telnet -action allow
system services firewall policy create -policy mgmt -service telnet -action deny -ip-list
0.0.0.0/0
```

Note: It is normal for some of these commands to return an error message stating that the entry does not exist.

11 Management Best Practices for RHEV and NetApp

The following management servers can be virtualized on a separate group of infrastructure hosts. By virtualizing the management servers, they gain the same benefits as the production virtual machines, such as mobility, availability, and centralized data management on the NetApp controllers.

11.1 RHEV-M

Red Hat Enterprise Virtualization Manager (RHEV-M) is the primary means of interacting with RHEV-H, virtual machines, logical networks, RHEV datastores, and all other aspects of the RHEV environment. Administrative access, poweruser access, and VDI access should all go through RHEV-M.

11.2 REST API

RHEV-M uses RESTful API, which allows interaction with all aspects of RHEV without having to go through the web portal. All automation tools and scripts should write to this API.

11.3 RHN and RHN Satellite

Red Hat Network (RHN) is a secure web portal as well as the source for all Red Hat related packages, updates, errata, and management tools for RHEL servers. It is a Red Hat and NetApp best practice to subscribe all Red Hat systems to RHN to keep up with security patches as well as compliance.

RHN Satellite is essentially an onsite instantiation of RHN. Instead of many systems subscribed to RHN, only the Satellite server is subscribed. All Red Hat systems are then subscribed to the Satellite server. This has many added benefits, such as reduced external network traffic as well as the ability to highly customize software channels and user management. Additionally, RHN Satellite can be used to provision new RHEL servers. RHN Satellite can be deployed on a RHEL KVM guest in the infrastructure RHEL 6 KVM hosts.

11.4 Kickstart Server

Kickstart is a means of providing semi-automated RHEL installations; it can be deployed by using a CD and answer file, HTTP, NFS, or FTP. It is a best practice to deploy Kickstart as a server on an RHEL KVM guest or as part of an RHN Satellite server in the infrastructure hosts.

11.5 NetApp OnCommand System Manager 2.0x RHEL

For an in-depth explanation of NetApp OnCommand System Manager, see OnCommand System Manager 2.0.1.

Overview

NetApp System Manager is a simple tool based on Java® technology that enables administrators to easily manage NetApp storage systems. System Manager provides recommendations based on NetApp best practices for common storage management tasks and workflows to save time and prevent configuration errors.

System Manager is the primary element-management GUI for Data ONTAP 8.1 and later. It replaces FilerView® for 7G management. It encompasses most of the element-management features supported by the current Cluster Manager for clustered Data ONTAP. Data ONTAP CLI and autogenerated web GUI are required only for advanced operations, because most of the common goals can be achieved from System Manager. System Manager is an out-of-the-box management application that can be installed on laptops, desktops, and server systems.

Administrative Capabilities

System Manager enables administrators to easily control the powerful capabilities of NetApp storage systems. These capabilities include:

- Managing disks, pooled storage, shares or exports, Snapshot copies, and network configuration
- Provisioning storage for SAN (iSCSI or FCP) and NAS (CIFS or NFS), for both physical and virtualized server environments
- Leveraging space efficiency features such as thin provisioning and deduplication
- Getting real-time views of system performance

Features

System Manager provides these features:

- **Storage management.** Volume, aggregate, SAN, NAS management, Vserver, and vFiler® unit setup
- **Diagnostics.** Cluster health dashboard
- **Configuration.** Network settings, protocol setup, user setup, security, syslog setup, and AutoSupport
- **Data protection.** SnapMirror setup

Guidelines

The following guidelines apply to System Manager:

- RHEL 5 is supported and should have Firefox 2.5 or 3.x. or later.
- The installer detects JRE, and installation does not proceed unless JRE 1.6 or later is on the system.
- The Linux platform requires Sun JRE 1.6 or later. If the system contains non-Sun based JRE, the installer does not proceed.
- The number of instances of the application is restricted to one.
- A unique session ID is used every time the application is launched to prevent multiple users from accessing the same instance of System Manager.
- System Manager uses HTTPS as the default protocol for communication with the storage systems.

11.6 Operations Manager

Product Overview

For all NetApp specific management tools described in the following sections, and for detailed explanations and deployment requirements, refer to NetApp TR-3710: Operations Manager, Provisioning Manager, and Protection Manager Best Practices Guide.

Operations Manager provides NetApp customers with a comprehensive monitoring and management dashboard for IT storage infrastructure. It also enables scalable storage management for NetApp NAS and SAN storage assets. From a central point of control, the solution offers alerts, reports, and configuration tools that help customers keep storage infrastructure in line with business requirements and policies to maximize availability and reduce the total cost of ownership. Operations Manager provides comprehensive reports of utilization and trend information to support capacity planning, space usage, and data backup space allocation.

11.7 NetApp Management Console 3.0

Overview

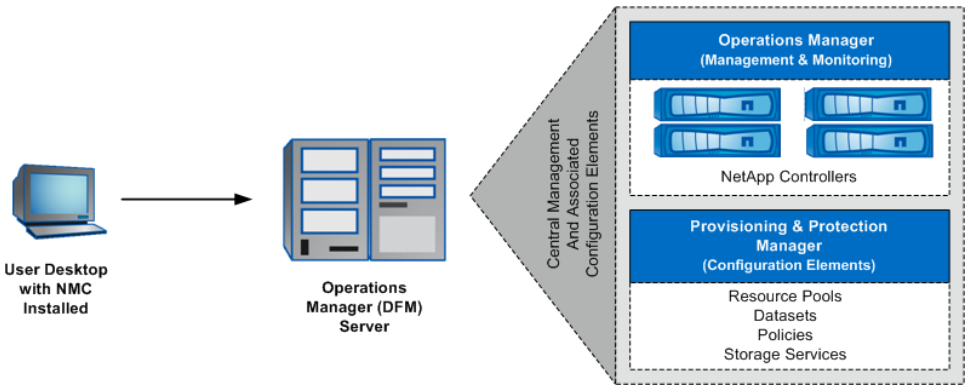
The NetApp management console (NMC) is used by administrators to carry out management tasks. It is aided by the DataFabric® Manager, but runs on Windows or Linux systems separate from the server on which the DataFabric Manager is installed.

NMC allows storage, application, and server administrators to perform management tasks such as data backup protection, space management, resource provisioning, data migration, and performance tuning, without having to switch between separate user interfaces.

The DataFabric Manager server provides infrastructure services (such as discovery, monitoring, RBAC, auditing, and logging for products in the storage and data suites) for NetApp Manageability Software client applications. The DataFabric Manager software runs on a separate server and is itself managed through Operations Manager, the web-based user interface of DataFabric Manager.

Figure 21 shows the overview of NMC.

Figure 21) NMC overview.



Applications That Run in NetApp Management Console

Performance Advisor, the licensed Protection Manager, and Provisioning Manager applications run in NMC.

11.8 Performance Advisor

This application provides a single location from which to view comprehensive information about storage system and MultiStore® vFiler unit performance and to perform short-trend analysis. The application also helps identify the causes and potential causes of reduced performance in the data infrastructure.

Performance Advisor is automatically enabled with the Operations Manager core license.

11.9 Protection Manager

This application provides a policy-based management tool to help unify and automate backup and mirroring operations. The application uses a holistic approach to data protection. It offers end-to-end workflow-based design and seamless integration of SnapVault, SnapMirror, and Open Systems SnapVault, making it possible to easily manage large-scale deployments.

The disaster recovery feature of the protection application enhances data protection services by making it possible to continue providing data access to users, even in the event of a mishap or disaster that disables or destroys the storage systems in the primary data node. If the disaster recovery license is installed, secondary storage systems can quickly be enabled to provide primary data storage access to users with little or no interruption, until the primary storage systems are reenabled or replaced.

11.10 Provisioning Manager

This application helps to simplify and automate the tasks of provisioning and managing storage. The application provides policy-based provisioning and conformance of storage in datasets. It also makes it possible to manually add volumes or qtrees to a dataset at any time, provides manual controls for space and capacity management of existing storage and newly provisioned storage, and allows the migration of datasets and vFiler units to a new storage destination.

The deduplication feature of the provisioning application enhances data provisioning services by making it possible to eliminate duplicate data blocks to reduce the amount of storage space used to store active data.

11.11 Storage Efficiency Dashboard

Overview

The NetApp Operations Manager Storage Efficiency Dashboard plug-in answers a number of questions related to storage utilization and storage efficiency savings. It also identifies ways to improve storage utilization. The dashboard is a script that can be installed on Operations Manager 3.8.1 and later. It produces a graphical dashboard report of the utilization and efficiency of NetApp storage environments.

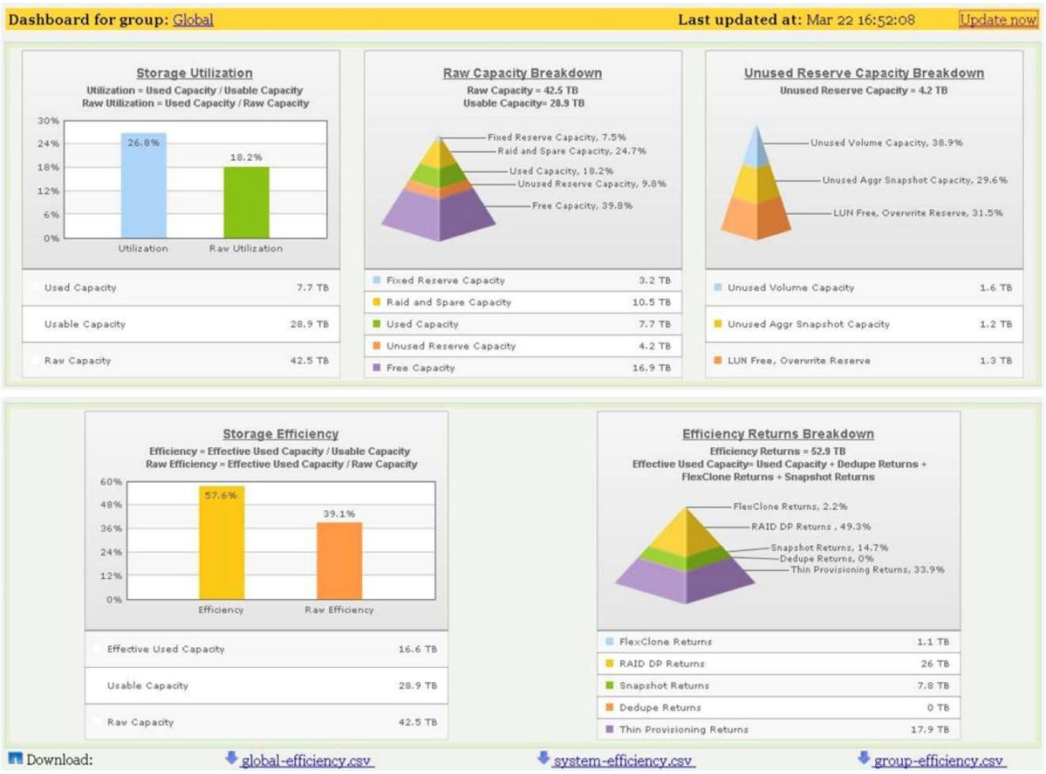
The script can be installed on an Operations Manager instance running on Windows or Linux servers and can be scheduled to generate a set of web pages that provide an efficiency dashboard for the NetApp storage systems managed by Operations Manager. It produces two primary charts, with additional charts available to provide detailed breakdowns of how the storage space is consumed. These charts represent all storage systems monitored by Operations Manager, groups of storage systems as grouped in Operations Manager, or a single storage system.

The two primary chart types are:

- **Utilization charts.** These charts provide the utilization data for all systems in a resource group or for an individual storage system.
- **Efficiency charts.** These charts show the effect of NetApp storage efficiency technologies such as deduplication, thin provisioning, and FlexClone, across resource groups or individual storage systems.

Figure 22 **Error! Reference source not found.**is a sample screenshot of the Storage Efficiency Dashboard in NetApp Operations Manager.

Figure 22) Storage efficiency dashboard.



11.12 RLM with Clustered Data ONTAP

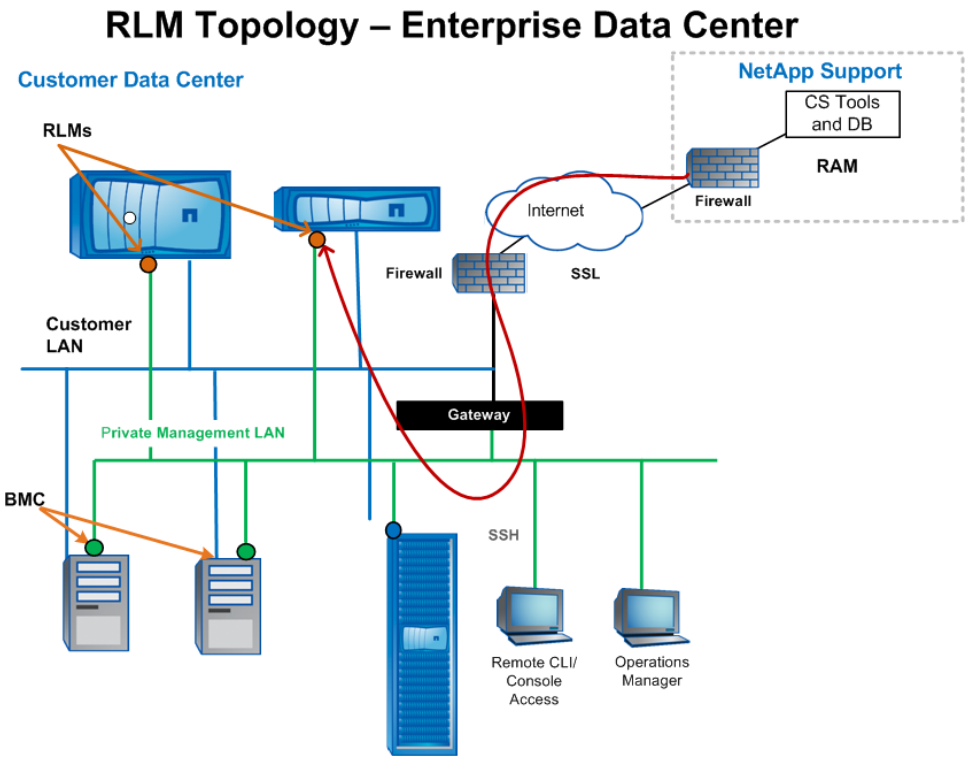
Overview

RLM is a physical card incorporated on NetApp FAS and V-Series storage systems that provides remote access and management capabilities for monitoring, troubleshooting, logging, and sending alert notifications. The V-Series and FAS3xxx and FAS6000 series storage systems have an expansion slot with a dedicated Ethernet interface for connecting the RLM card to the system controller. The card stays operational regardless of the operating state of the storage system, enabling storage administrators to remotely access their storage systems.

With the RLM card enabled, customers can remotely manage their storage systems without the need for dedicated 24/7 on-site storage administrators, as shown in Figure . RLM technology uses existing data center IP infrastructure, eliminating the need for remote support infrastructure, such as terminal concentrators or power controllers.

Figure illustrates the RLM topology.

Figure 23) RLM topology.



Easy-to-Use Remote Management Capabilities

The RLM card provides the following easy-to-use remote management capabilities on NetApp storage systems:

- The RLM remotely administers the storage system by using the Data ONTAP CLI through the RLM's system console redirection feature.
- The RLM remotely accesses the storage system and diagnoses error conditions, even if the storage system has failed, by performing the following tasks:
  - Views the storage system console messages captured in the RLM's console log
  - Views the storage system events captured in the RLM's system event log
  - Initiates a storage system core dump
  - Power-cycles the storage system (turns it on or off)
  - Resets the storage system
  - Reboots the storage system
- The RLM remotely extends AutoSupport capabilities by sending alerts and `down system` or `down filer` notifications through an AutoSupport message when the storage system goes down, regardless of whether the storage system can send AutoSupport messages. Other than generating these messages on behalf of a system that is down and attaching additional diagnostic information to AutoSupport messages, the RLM has no effect on the storage system's AutoSupport functionality.

Note: AutoSupport configuration settings and message content behavior of the RLM are inherited from Data ONTAP.

Note: The RLM does not rely on the `-transport` parameter setting of the `system node autosupport modify` command to send notifications. The RLM uses SMTP.

- If SNMP is enabled for the RLM, the RLM generates SNMP traps to configured trap hosts for all `down system` and `down filer` events.
- The RLM remains operational regardless of the operating state of the storage system. It is powered by a standby power supply, which is available as long as the storage system has input power to at least one of its power supplies.
- The RLM has a single temperature sensor to detect ambient temperature around the RLM board. Data generated by this sensor is not used for any system or RLM environmental policies. It is used only as a reference point that might help troubleshoot storage system issues. For example, it can help a remote system administrator determine whether a system was shut down due to an extreme temperature change in the system.
- The RLM has a nonvolatile memory buffer that stores up to 4,000 system events in a system event log (SEL) to help diagnose system issues. The event list from the SEL is automatically sent by the RLM to specified recipients in an AutoSupport message.

  The records contain the following data:

  - Hardware events detected by the RLM—for example, system sensor status about power supplies, voltage, or other components.
  - Errors (generated by the storage system or the RLM) detected by the RLM—for example, a communication error, a fan failure, a memory or CPU error, or a `boot image not found` message.
  - Critical software events sent to the RLM by the storage system—for example, a system panic, a communication failure, an unexpected boot environment prompt, a boot failure, or a user-triggered `down system` as a result of issuing the `system reset` or `system power cycle` command.
- The RLM monitors the storage system console regardless of whether administrators are logged in or are connected to the console. When storage system messages are sent to the console, the RLM stores them in the console log. The console log persists as long as the RLM has power from either of the storage system's power supplies. Because the RLM operates with standby power, it remains available even when the storage system is power-cycled or turned off.
- Hardware-assisted takeover is available on systems that support the RLM and have the RLM modules set up. For more information about hardware-assisted takeover, refer to the Data ONTAP Clustered Data ONTAP High-Availability Configuration Guide.
- The RLM supports the SSH protocol for CLI access from UNIX clients and PuTTY for CLI access from PC clients.

Accounts That Can Access the RLM

Consider the following when creating accounts that can access RLM:

- User accounts that are created with the service-processor application type have access to the RLM CLI on any node of the cluster that supports the RLM. RLM user accounts are managed from Data ONTAP and are authenticated by password.
- Do not create user accounts directly from the RLM CLI. RLM user accounts are created and managed from Data ONTAP, using the `security login` commands with the `-application` parameter set to `service-processor`. RLM supports only password authentication; therefore, when the RLM user account is created, the authentication method (the `-authmethod` parameter) must be set to `password`.
- Display current RLM user accounts by using the `-application service-processor` parameter of the `security login show` command.

The cluster user account `admin` includes the service-processor application type and has access to the RLM CLI by default.

The system prevents the creation of user accounts with names that are reserved for the system, such as `root` and `naroot`.

Note: Do not use a system-reserved name to access the cluster or the RLM.

11.13 Service Processor

Overview

The service processor (SP) is a physical device that enables storage administrators to access, monitor, and troubleshoot the storage system remotely. The SP is available on all NetApp systems except for the 20xx, 30xx, 31xx, and 60xx systems. The SP remains operational regardless of the operating state of the storage system. It is powered by a standby power supply, which is available as long as the system has input power to at least one of the system's power supplies. It is also connected to the system through the serial console.

SP Remote Management Capabilities

The service processor offers the following capabilities:

- Remotely administers the storage system by using the SP CLI.
- Remotely accesses the system console to run the Data ONTAP commands.

  Note: The SP can be accessed from the system console. If the system console becomes unresponsive, press `Ctrl+G` to access the SP CLI.
- Monitors environmental sensors and logs system events.
- Has a nonvolatile memory buffer that stores up to 4,000 system events in a single SEL.
- Monitors the system console regardless of whether administrators are logged in or are connected to the console.
- Extends AutoSupport capabilities by sending alerts and `down system` or `down filer` notifications through an AutoSupport message when the storage system goes down, regardless of whether the storage system can send AutoSupport messages. Other than generating these messages on behalf of a system that is down and attaching additional diagnostic information to AutoSupport messages, the SP has no effect on the storage system's AutoSupport functionality. AutoSupport configuration settings and message content behavior of the SP are inherited from Data ONTAP.

  Note: The SP does not rely on the `-transport` parameter setting of the `system node autosupport` modify command to send notifications. The SP uses SMTP.

  Note: If SNMP is enabled for the SP, the SP generates SNMP traps to configured trap hosts for all `down system` and `down filer` events.
- The SP supports the SSH protocol for SP CLI access and the ability to execute Data ONTAP commands remotely.

12 Data Protection Best Practices

12.1 Snapshot Clustered Data ONTAP

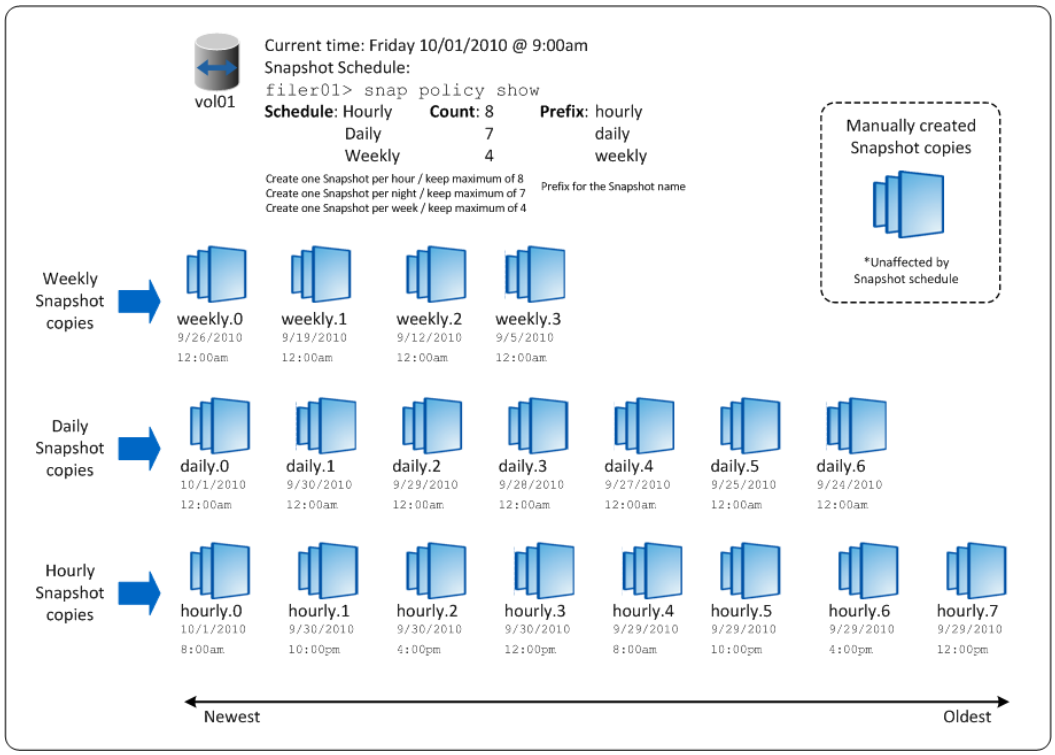
Overview

A Snapshot copy is a read-only image of a FlexVol volume or an aggregate that captures the state of the file system at a point in time.

For information about FlexVol volumes or aggregates, refer to the Data ONTAP 8.1 Clustered Data ONTAP Physical Storage Management Guide. Data ONTAP maintains a configurable Snapshot copy schedule that creates and deletes Snapshot copies automatically for each volume. Snapshot copies can also be created and deleted manually by using the `snap create` and `snap delete` commands.

Figure 24 shows an example volume (`vol01`) and the corresponding Snapshot copies that are created by the schedule.

Figure 24) Snapshot copy example that uses the snap policy show command.



Guidelines and Restrictions

Avoid scheduling Snapshot copy creation at the same time as SnapMirror updates or SnapVault activity. If these schedules conflict, Snapshot copy creation might not occur.

Stagger the Snapshot copy update schedules so that SnapMirror activity does not begin or end at exactly the same time that a Snapshot operation attempts to create a Snapshot copy.

Snapshot Copies in a SAN Environment

Administrators can use Snapshot technology to make copies in the SAN environment when the data in a Data ONTAP LUN is in a consistent state; however, Data ONTAP does not know whether an application is accessing the data inside the LUN (that is, whether or not it is in a consistent state). Therefore, before creating a Snapshot copy, administrators must quiesce the application or file system that is using the LUN. This action flushes the host file system buffers to disk and provides a consistent Snapshot copy.

One way to accomplish this is to use batch files and scripts on a host that has administrative access to the system. The SnapDrive®, SnapManager, and Snap Creator products also quiesce LUNs before creating Snapshot copies and should be used whenever possible.

Snapshot Restore Guidelines

Snapshot promote or Snapshot restore allows administrators to quickly revert a local volume or local file to the state it was in when a particular Snapshot copy was created. In most cases, reverting a file or volume is much faster than restoring files from tape or copying files from a Snapshot copy to the active file system. Snapshot restore is implemented in advanced mode.

Follow these guidelines for using Snapshot promote or Snapshot restore.

- If the amount of data to be recovered is large, Snapshot promote is the preferred recovery method, because it takes a long time to copy large amounts of data from a Snapshot copy or to restore from tape.
- If the file to be recovered needs more space than the amount of free space in the active file system, the file cannot be restored by copying it from the Snapshot copy to the active file system.
- If the volume that must be restored is a root volume, it is easier to copy the files from a Snapshot copy or to restore the files from tape than to use Snapshot promote, because rebooting is not necessary. However, if only a corrupted file on a root volume must be restored, a reboot is not necessary.
- If the entire root volume is reverted, the system reboots with the configuration files that were in effect when the Snapshot copy was created.

Create a Snapshot Copy in Clustered Data ONTAP

You can manually create a Snapshot copy from the controller CLI. In this example, a Snapshot copy named `snapshot_name` on `volume_name` is created.

```
snap create -vserver vserver_name -volume volume_name -snapshot snapshot_name
```

Restore an Entire Volume from a Snapshot Copy

To perform a Snapshot restore, run the following CLI operation.

```
vol snapshot restore -vserver vserver_name -volume volume_name -snapshot snapshot_name
```

Restore a File from a Snapshot Copy

To restore a single file from a Snapshot copy, run the following CLI operation.

```
vol snapshot restore-file -vserver vserver_name -volume volume_name -snapshot snapshot_name -path
file_path
```

12.2 Snap Creator

Snap Creator is a software framework that integrates NetApp data protection technologies with multiple applications and platforms. Snap Creator facilitates the management of backups, reduces the complexity of solutions, and provides a method to take advantage of NetApp technology with minimum programming effort on the user's part.

Overview

Snap Creator uses Snapshot, SnapRestore, and FlexClone technologies to simplify backups, restores, and cloning by communicating with NetApp storage through NetApp API commands. The framework allows triggering NetApp Snapshot copies of volumes (data stores), as well as any activities that need to occur before and after the Snapshot copy is created. It can also trigger SnapMirror activities between NetApp controllers and/or data centers to meet disaster recovery and backup requirements.

In the context of Red Hat Enterprise Virtualization, Snap Creator can be used to provide intelligent backups for virtualized applications. There is no support yet for the Snap Creator agent on RHEV-H or an RHEL 6 KVM node that is managed by RHEV-M.

12.3 Snap Creator Server

The following guidelines apply to configuring Snap Creator:

- Although the Snap Creator server is an infrastructure application and can be virtualized, do *not* place it with the other virtualized infrastructure applications. This could result in Snap Creator pausing the virtual machine that it is running on without any means of resuming itself.
- The Snap Creator agent/server configuration is a client/server type of architecture.
- The Snap Creator agent is installed on the host with the targets to be backed up in an agent/server architecture.
- The default port used by the Snap Creator agent/server architecture for communication is 9090.
- The Snap Creator server is installed on the centralized host that communicates with other hosts that have a Snap Creator agent installed and running.
- The Snap Creator agent executes any prescripts or postscripts and commands on its host.
- If plug-ins and PRE commands or POST commands are not needed, then it might not be necessary to install the agent.
- Snap Creator supports the use of both GUI and CLI. The GUI is configured when Snap Creator is installed.

Architecture Components

Figure 25 shows the components of the Snap Creator 3.x server architecture.

Figure 25) Snap Creator 3.x server architecture.

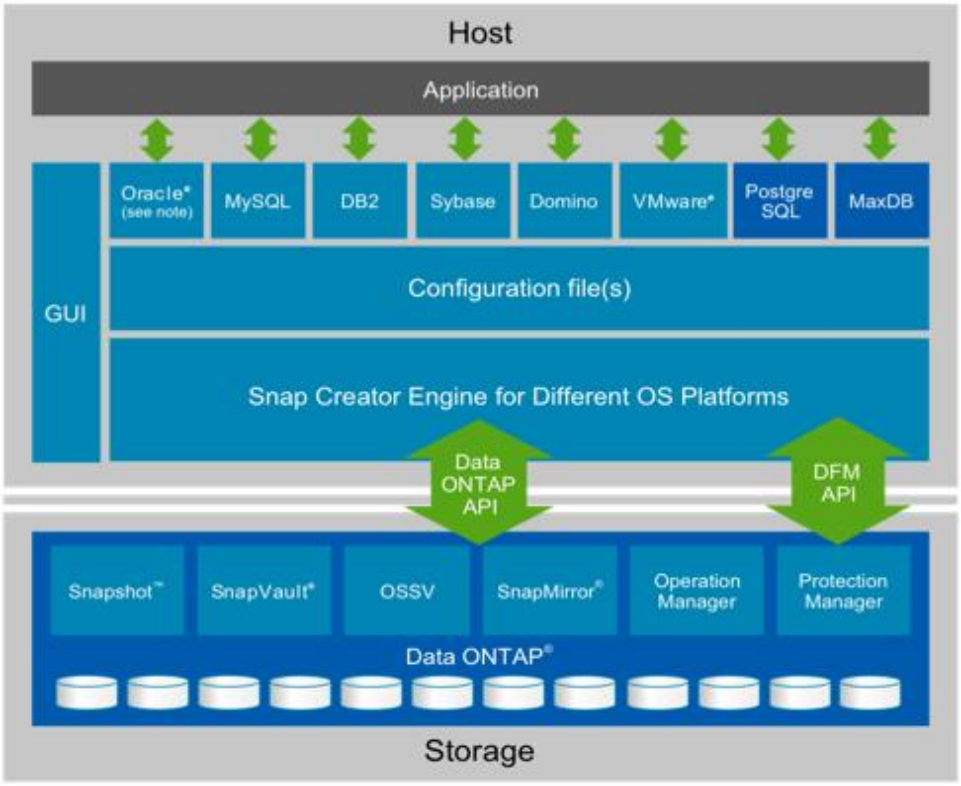


Figure shows the components of the Snap Creator 3.x agent architecture.

Figure 26) Snap Creator 3.x agent architecture.

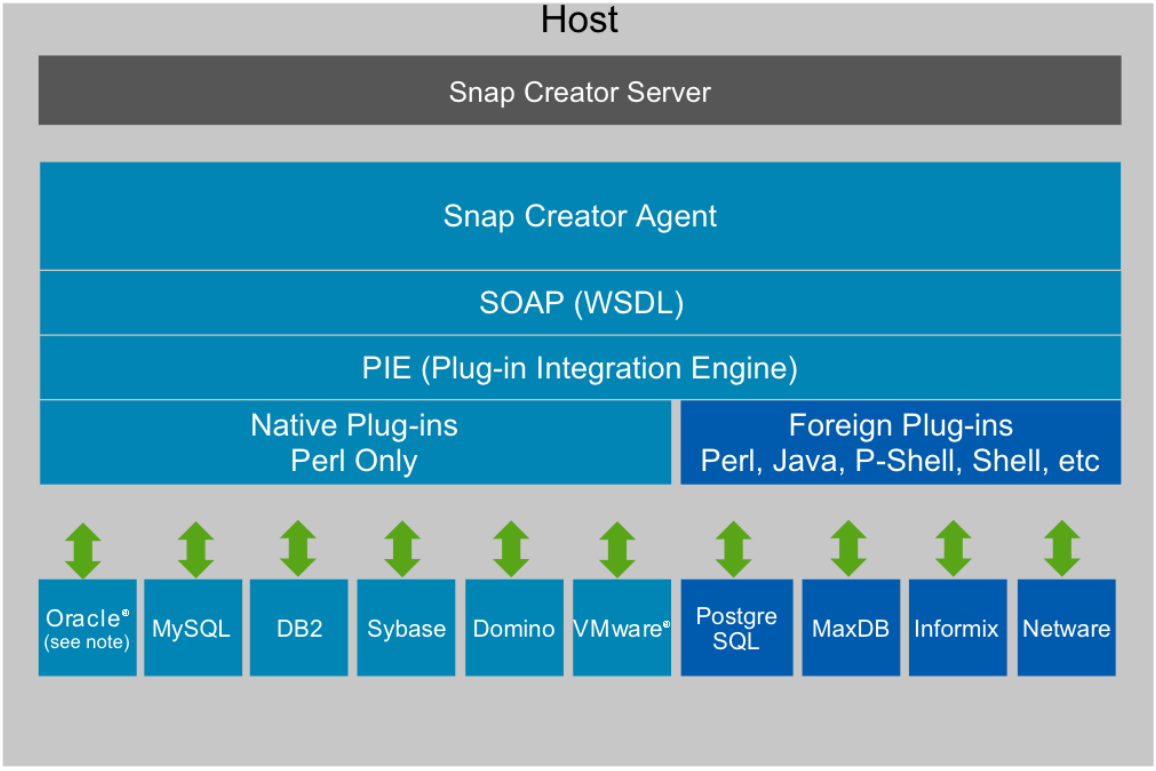
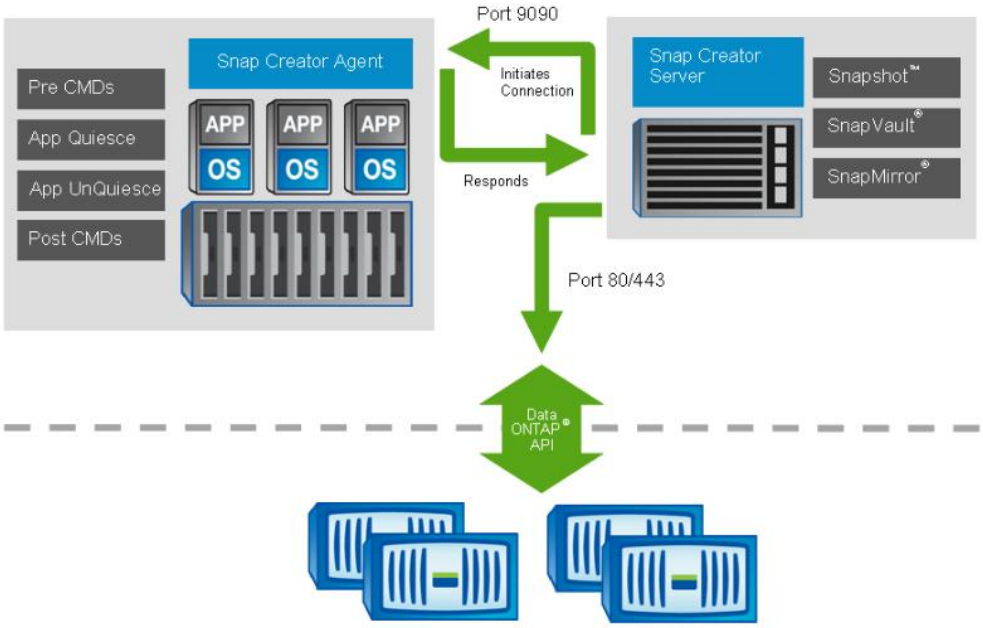


Figure 27 shows the combined agent/server architecture.

Figure 27) Snap Creator 3.x agent/server architecture.



The main components of the Snap Creator architecture are:

- **Snap Creator GUI.** The GUI allows the creation of customized configuration files and new profiles for the environment.
- **Snap Creator server.** The server is where the GUI, the configuration files, and the CLI reside. It is the main engine for the Snap Creator Framework. The Snap Creator server engine communicates with the NetApp storage controller through API commands. It also integrates with Open Systems SnapVault, Operations Manager, Protection Manager, and SnapDrive.
- **Snap Creator agent.** The agent is a lightweight daemon that runs remotely or locally and allows the Snap Creator server engine to send commands or operations to that agent based on the configuration of the workflow. The communication layer is HTTP or HTTPS, and it uses SOAP.
- **Snap Creator configuration files.** The configuration files reside in a profile. Each profile is a directory in the installation directory. The configuration file controls the workflow of Snap Creator. A profile may contain many configuration files; but only one can be active during each execution of Snap Creator. The configuration file contains information about how to name Snapshot copies, how to manage Snapshot copy retention, whether cloning is required, how many FlexClone volumes need to be retained, how to manage the data protection integration, and so forth.
- **Snap Creator application plug-ins.** Built-in plug-ins and applicable support can be found in the documentation for the applications and in the IMT. The plug-ins assist with application consistency during the creation of backups or FlexClone volumes or for the management of specific components of the application, such as Oracle® archive logs.

12.4 Install the Snap Creator Framework Server

1. Download Snap Creator Framework from the NetApp Support site.
2. Extract the Snap Creator Framework `tar` file.

```
cd /opt
tar -zxvf NetApp_Snap_Creator_Framework3.*
```

3. Run Snap Creator Framework as the root user for the initial setup.

   Note: The serial number prompted for is the serial number of the storage controller; it is optional.

```
cd /opt/scServer*
chmod 755 snapcreator
./snapcreator --profile setup
Welcome to the NetApp Snap Creator Framework!
End User License Agreement for NetApp, Inc. Software
.
.
.
Do you accept the End User License Agreement (y|n): y
Setup NetApp Snap Creator Framework Server (y|n): y
Enter serial number:
Enable GUI job monitor (Y|N): n
Please Enter GUI Administrator Username: snapcreator_admin
Please Enter password for snapcreator_admin:
Please Confirm password for snapcreator_admin:

INFO: Updated NetApp Snap Creator Framework GUI

INFO: To start GUI please do the following:

cd /opt/scServer*/gui
java -jar snapcreator.jar
or
java -jar snapcreator.jar  -db_password <db_password> -db_port <db_port> -db_user name <db_user
name> -gui_port <gui_port>

INFO: To access NetApp Snap Creator Framework GUI goto "http://unixhost:8080" or
"http://unixhost:<gui_port>"
```

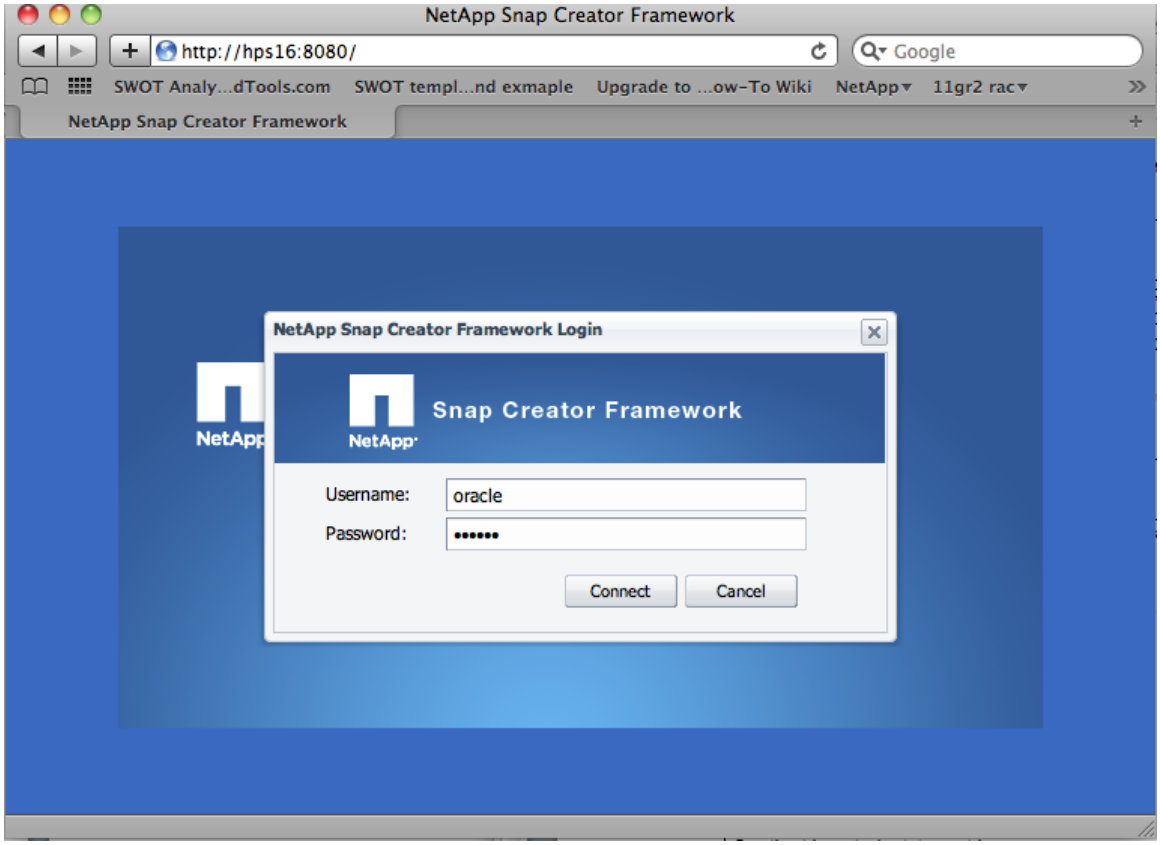
4. Start the Snap Creator Framework GUI.

   Note: Before starting the Snap Creator server, use netstat or a similar tool to verify that the network port that it uses (8080 by default) is not in use.

```
# cp /opt/scServer*/bin/scServer /etc/init.d
# chkconfig --add scServer
# chkconfig scServer on
# service scServer start
```

5. Validate the Snap Creator Framework GUI startup by navigating to the local host on port 8080.

   Note: If going through a firewall, open the network port (8080 by default).



6. If a Snap Creator user is created specifically on the NetApp cluster, then create an encrypted password. This step prevents a plain text password from being inserted into a configuration file on the host on which Snap Creator is installed.

```
./snapcreator  --cryptpasswd
Please Enter Password:
Your encrypted password is: 53616c7465645f5f614d4964d340f7f2d26eef38f443f5ea9c2f8020015a2dfa
```

12.5 Snap Creator Agent

Overview

The Snap Creator agent is a lightweight daemon that runs remotely or locally and allows the Snap Creator server to send quiesce and unquiesce operations to a given database.

The Snap Creator agent remotely handles operations on an application through the Snap Creator plug-ins. All Snap Creator configurations are centrally stored on the Snap Creator server, and all backup tasks can be scheduled from the same host. This architecture provides a “single pane of glass” for backup and restore operations.

Note: The Snap Creator agent can be installed only on an RHEV guest. It is not yet supported for RHEV-H nodes or RHEL 6 KVM nodes that are managed by RHEV-M.

The default port is 9090, but any port can be used.

Single Object Access Protocol (SOAP) over HTTP is used for communication between the agent and the server. Any SOAP client can interact with the agent through the use of the Web Services Description Language (WSDL). Currently, Apache CXF (for Java) and Windows PowerShell™ (for Windows) can be used to create an agent.

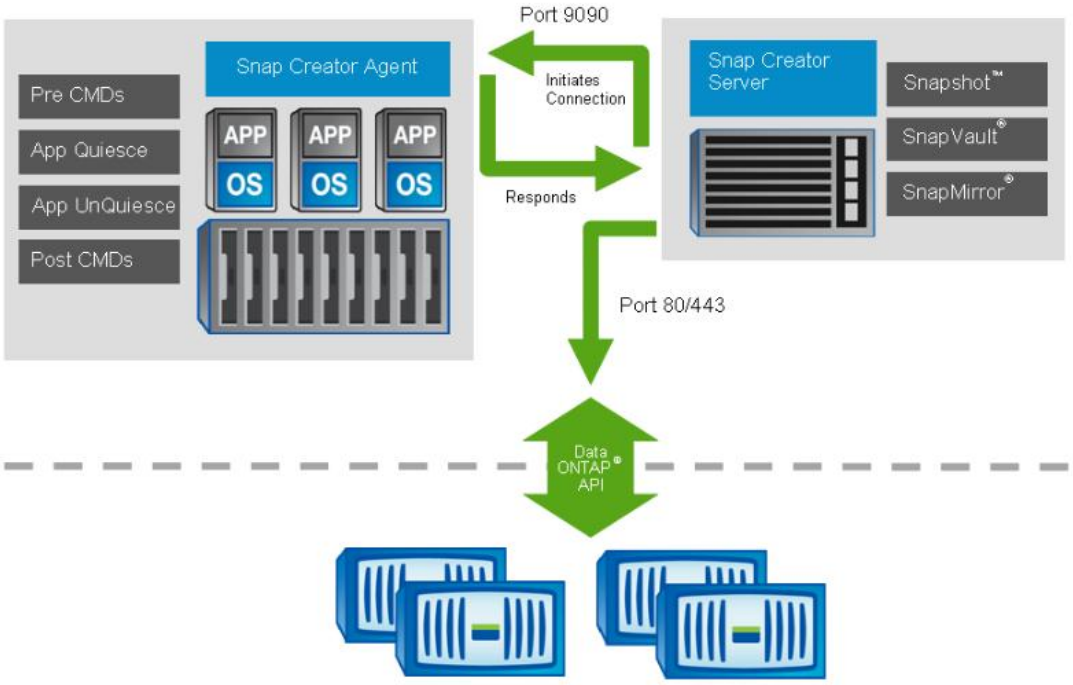
The supported application plug-ins are built into the agent. Other community plug-ins are available as source codes and can be added to the `/plug-ins` directory.

In addition to the application plug-ins, all PRE commands, POST commands, and APP commands can be executed remotely through the agent. This makes it possible to mount file systems or remotely perform additional application processing. The agent has a configuration file (`agent.conf`) in which certain commands can run. This file is located under `/path/so/ scAgent_v<#>/config/agent.conf`. By default, all commands are denied, which means that only the built-in or community plug-ins can execute commands through the agent. PRE and POST scripting commands and scripts must be added to the `agent.conf` file.

Note: Snap Creator 3.5.0 also supports Snap Creator 3.4.0 agents.

Figure shows the Snap Creator agent communication.

Figure 28) Snap Creator agent communication.



Snap Creator Agent Security

Snap Creator runs centrally and uses an agent to communicate with the database servers. The agent is a daemon that runs on any port; by default, it runs on port 9090. This port must be open between the Snap Creator server and the server that is running the agent. Communication between the agent and the Snap Creator server occurs through SOAP over HTTP. The host information and the commands should be specified in the agent configuration file.

The Snap Creator agent uses a file named `agent.conf` to secure its functionalities. The `agent.conf` file allows or restricts hosts and commands.

The `agent.conf` file is located in the same `config` subdirectory where the agent is installed: `/path/to/scAgent_v<#>/config/agent.conf`.

By default, the Snap Creator agent allows communications with any Snap Creator server, but communications can be limited to a particular server by changing the host line in the `agent.conf` file. The default host entry command in the `agent.conf` file is as follows.

```
host: scServer@*
```

The wildcard (`*`) tells Snap Creator to allow communications with any host. The wildcard can be replaced with a host name or IP address to restrict communications to a particular Snap Creator server.

12.6 Install the Snap Creator Agent

Perform these steps to install Snap Creator on Red Hat Enterprise Linux:

1. If Snap Creator has not already been downloaded, it can be found in the software download section of the NetApp Support site under Snap Creator Framework. Verify that the proper bit level for the OS is downloaded.
2. Copy the downloaded Snap Creator `tar.gz` file to `/etc`.

   If a subdirectory must be made, use the `mkdir` command with the directory name (for example, `SC`).

```
mkdir SC
```

3. Use the `cp` command to copy the Snap Creator `tar.gz` file to the newly created directory. For example, from the directory where Snap Creator is downloaded, run the command to copy the Linux Snap Creator `tar.gz file` to the newly created `SC_3.5` directory.

```
cp Snap_Creator_Community_Release_3.* /SC
```

4. If a new directory is created, the user running Snap Creator must be set as its owner. This is easiest to do before extracting the Snap Creator software, but it can be done at any point. File and folder ownership can be changed by using the `chown` command. When `chown` is used with the `–R` switch, ownership is also changed for files and folders under the directory. For example, to change ownership of all files and folders in the newly created directory for user `snapcreator_user`, run the following command.

```
chown –R snapcreator_user /SC
```

5. Use the `cd` command to change to the `/SC` directory, from which `the tar.gz` file will be extracted.

```
cd /SC
```

6. The `tar.gz` file must be unzipped before it can be extracted. To unzip the Linux Snap Creator `tar.gz` file, run the `gunzip` command.

```
gunzip Snap_Creator_Community_Release_3.*.gz
```

Use the `tar` command to extract the Linux Snap Creator `.tar` file. `Tar` is typically executed by using the `-xvf` switches, which tell `tar` to extract (`-x`) the file (`-f`) in verbose (`-v`) mode.

```
tar -xvf Snap_Creator_Community_Release_3.*.tar
```

7. Run the directory listing command (`ls`) on this directory. Two directories should be displayed: `scAgent<version#>` and `scServer<version#>`, as well as the newly extracted `.tar` file.

Deploy and Start the Snap Creator Agent

Follow these steps to start the Snap Creator agent on open systems such as AIX, Linux, and Solaris.

1. To configure the Snap Creator agent, change directories to the `/<path>/<to>/scAgent_v<#>` subdirectory and run the following command.

```
./snapcreator --profile setup
```

Note: The Snap Creator executable should already be configured before extraction, with the proper permissions to be executed. If the `profile setup` command does not work, the permissions might have to be added.

```
chmod 755 snapcreator
```

2. Executing this command starts the Snap Creator setup wizard. The first page displays the EULA. At the end of the EULA, a prompt asks for EULA acceptance. At the prompt, enter `y` and then press Enter.

```
Do you accept the End User License Agreement (y|n):
```

3. Confirm that the Snap Creator server should be set up. In this instance, it is the agent that must be configured, not the server. At the prompt, enter `n` and then press Enter.

```
Setup NetApp Snap Creator Framework 3.5.0 Server (y|n):
```

4. Confirm that the Snap Creator agent should be set up. At the prompt, enter `y` and then press Enter.

```
Setup NetApp Snap Creator Framework 3.5.0 Agent (y|n):
```

This updates the environmental variables so that the Snap Creator agent scripts will work properly. The usage information for the agent appears on the page.

```
INFO: Updated NetApp Snap Creator Framework 3.5.0 Agent
INFO: To start the NetApp Snap Creator Framework 3.5.0 Agent run
"/SC_3.5/scAgent3.5.0/bin/scAgent start"
INFO: To stop the NetApp Snap Creator Framework 3.5.0 Agent run "/SC_3.5/scAgent3.5.0/bin/scAgent
stop"
```

5. Copy the agent start/stop script, then enable it and start it.

```
# cp /opt/scAgent*/bin/scAgent /etc/init.d
# chkconfig --add scAgent
# chkconfig scAgent on
# service scAgent start
```

6. The default behavior of the Snap Creator Framework agent is to communicate with the Snap Creator server over port 9090. The port that Snap Creator uses for server/agent communication can be specified through the `SC_AGENT_PORT` environmental variable. This is a system environmental variable, so the commands to set this variable differ for each OS; check the OS documentation for information on setting environmental variables. For example, to set the `SC_AGENT_PORT` environmental variable to port number 9091 when using Red Hat Linux, the command includes 9091.

```
export SC_AGENT_PORT=9091
```

Note: NetApp recommends using netstat or a similar tool to verify that the network port (9090 by default) is not already in use.

7. Open the iptables firewall on the Snap Creator server to allow access to port 9090. If the Snap Creator server is a virtual machine, then the port needs to be opened on the virtual machine and the hypervisor.

Configure the Snap Creator Agent

The Snap Creator Agent can be installed in an RHEL 6 KVM host or a virtual machine. By default, the Snap Creator agent prevents any commands that are not part of the Snap Creator Framework or that are not part of a Snap Creator plug-in from being executed on remote agents. However, in some situations the configuration file might be set up in such a way that additional commands must be executed on a remote agent. These commands include any entries that might be added to the PRE, POST, APP, or other commands.

This is done by adding the specific commands that are required to run as part of the backup or recovery activity into the `agent.conf` file.

Because the Snap Creator agent blocks additional commands by default, the default command entry in the `agent.conf` file is as follows.

```
command:
```

Any commands or scripts that need to be given permission to run in the `agent.conf` file must be listed on a separate line. For example, to add the permission to run commands specific to affecting VM state or application state, add the following lines to the file.

```
command:sync
command:virsh *
command:/path/to/some/application
command:/path/to/some/script
```

Regular expressions can also be used to add commands to the `agent.conf` file in a more restrictive way.

Note: Although the wildcard (*) can be used to allow all commands to be executed, NetApp does not recommend this practice for security reasons.

12.7 Backup Use Cases

There are numerous backup use cases; however, the following five use cases represent the vast majority that need to be addressed in a virtual environment such as RHEL 6 KVM.

Create a Backup Profile for RHEL 6

1. Open a web browser to http://myserver.mydomain.com:8080, replacing the host and domain with appropriate values.
2. Log in to the site with the user name and password created when the server profile was created.
3. Select Management > Configurations >, then click the + button under Backup Profiles and enter a name for the profile.
4. Under Backup Profiles, right-click the newly created profile and select New Configuration to launch the configuration wizard. Click Next.
5. Enter a configuration name and select both Password Encryption and Advanced Options. Click Next.
6. Select None for the Plug-in type. (The KVM plug-in does not yet support clustered Data ONTAP.) Click “Next.”

7. Enter the IP address for the virtual machine, the agent port (9090), and a timeout of 10. Click Test agent connection. When the test succeeds, click OK. Click Next.
8. Enter the IP address of the Vserver and the login information, then select the transport protocol for the NetApp controller. If secure web access is enabled on the NetApp controller, select HTTPS; otherwise, select HTTP. Click Next.
9. Select the volume or volumes to be backed up by the profile. Click Next.
10. Enter a name for the Snapshot copies, enter a policy name, and select either Recent or Timestamp for the naming convention. Select the number of Snapshot copies to keep and how many days to keep each copy. Click Next.
11. Select Operations Manager Alert. Enter the IP and login information for the Operations Manager Server. Click Next.
12. Click Finish.

Create a Snap Creator Configuration for Red Hat

1. Open a web browser to http://myserver.mydomain.com:8080, replacing the host and domain with appropriate values.
2. Log in to the site with the user name and password created when the server profile was created.
3. Select Management > Configurations and then select the profile and configuration to be edited.
4. Scroll to the bottom of the configuration to the PRE/POST section.
5. To quiesce an application prior to the Snapshot copy, click the + button at the bottom of the Application Quiesce Command subsection. Enter the command or commands to be run. If there are multiple commands, they must be listed in proper order.
6. Make note of the commands to be run. All commands must be added to the list of allowed actions for the Snap Creator agent described in the “Configure the Snap Creator Agent,” earlier in this section.
7. To resume an application after the Snapshot copy is complete, click the + button at the bottom of the Application Un-Quiesce Command subsection. Enter the command or commands to be run. If there are multiple commands, they must be listed in proper order.
8. Make note of the commands to be run. All commands must be added to the list of allowed actions for the Snap Creator agent described in “Configure the Snap Creator Agent,” earlier in this section.
9. Scroll to the top of the configuration and click the disk icon to save the updated configuration.

Crash-Consistent Backups with Snap Creator

This use case creates a Snapshot copy of a datastore without quiescing any virtual machines or applications; that is, while the data is “in flight.” The Snapshot copy can then be mirrored to another controller for backup or archiving. This is fine for capturing the current state, but a restore would depend on file system logs on the guest operating system to replay properly.

Application-Consistent Backup with Snap Creator

This use case assumes that the application data is on a separate volume from the virtual machine datastore. Snap Creator first triggers the application to quiesce, then triggers the Snapshot copy on the application data volume, and then triggers the application to resume. The Snapshot copy can then be mirrored to another controller for backup or archive.

Fully Consistent Snapshot Backup with Snap Creator

This use case is typically an add-on to the application-consistent backup. After the application is quiesced, Snap Creator tells the guest to sync its buffers (RHEL only), then tells RHEL 6 KVM to pause

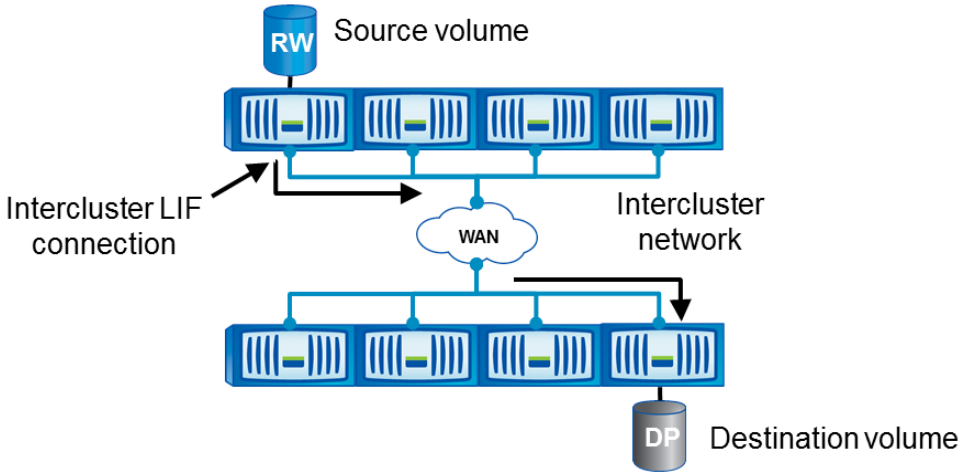
the virtual machine. It then triggers the Snapshot copies for the application volume and the virtual machine datastore, tells RHEL 6 KVM to resume the virtual machine, and finally resumes the application. The Snapshot copies can then be mirrored to another controller for backup or archiving.

12.8 Volume SnapMirror Async with Clustered Data ONTAP

Overview

SnapMirror technology is used primarily for data protection; it enables customers to copy and back up their production data from a primary or source volume over to another volume. The primary purpose of SnapMirror cross-cluster volume mirroring is to enable replication of individual volumes between independent clusters, independent of the specific network topology connecting the clusters as seen in Figure 29. This intercluster volume replication provides appropriate performance and scalability when operating across clusters and over WANs, for both data transfer throughput and overall manageability of the relationships.

Figure 29) Intercluster clustered Data ONTAP SnapMirror.



Firewall Configuration

The following ports are required for operating clustered Data ONTAP SnapMirror through a firewall:

- netapp-icmgmt: 11104
- netapp-icdata: 11105

Configure SnapMirror Async in Clustered Data ONTAP

Perform the following steps to configure and synchronize a SnapMirror relationship. Here, `cluster_name_1` is the source cluster, `vserver_name_1` is the source server, and `volume_name_1` is the source volume. Similarly, on the destination, `cluster_name_2`, `vserver_name_2`, and `volume_name_2` are the destination cluster, the destination server, and the destination volume, respectively.

1. Verify that an intercluster LIF has been created.
2. Create the destination volume to make sure that it has the correct attributes.

```
vol create -vserver vserver_name_2 -volume volume_name_2 -aggregate aggr_name -size size_of_vol -
type DP
```

3. From the cluster that contains the destination volume, run `snapmirror create`.
4. The destination volume must be of the type DP. The size of the destination volume must be the same as or larger than the size of the source volume.

```
snapmirror create -source-path cluster_name_1://vserver_name_1/volume_name_1 -destination-path
cluster_name_2://vserver_name_2/volume_name_2 -type DP
```

5. From the cluster that contains the destination volume, run `snapmirror initialize`.

```
snapmirror initialize -source-path cluster_name_1://vserver_name_1/volume_name_1 -destination-
path cluster_name_2://vserver_name_2/volume_name_2
```

12.9 Traditional Backup Methods

Although the use of NetApp Snapshot, Snap Creator, and SnapMirror technologies are best practices, they are not requirements for Red Hat Enterprise Virtualization. However, NetApp recommends using some form of data backup as a key foundation piece of enterprise data protection.

NetApp also includes two means to back up data in Data ONTAP that do not require any additional licenses. The dump and ndmpcopy tools replicate data to tape drives or to other storage respectively. This satisfies the requirement for backup utilities in Red Hat Enterprise Virtualization.

13 Conclusion

The ease of use and performance offered by Red Hat Enterprise Virtualization, combined with the highly performing, scaling, and available lustered NetApp Data ONTAP, give customers with a solid foundation for their virtualization and cloud deployments. This technical report outlines the best practices for deploying such a foundation.

Appendixes

Appendix A: HBA (FC, FCoE, and iSCSI) Configuration for SAN Boot

Booting from iSCSI (HBA)

Although RHEL 6 supports the use of a software initiator for boot, an iSCSI HBA or multiprotocol Ethernet adapter is recommended for iSCSI-based boot LUNs. Specific configuration for hardware initiators differs from vendor to vendor, but generally complies with the following workflow:

1. Boot the server to the BIOS (onboard multiprotocol Ethernet device) or HBA BIOS (iSCSI HBA).
2. Enable the device as a boot device. (It may be necessary to enable the BIOS on an iSCSI HBA.)
3. Make a note of the iSCSI initiator name or edit it to match a predetermined initiator name.
4. If necessary, edit the NetApp igroup to include the initiator name, if it has not yet been entered.
5. Scan the bus for devices.
6. Select the device from which to boot.
7. Reboot.

   RHEL 6 recognizes the device on reboot and during the installation process.

Booting from FC (FCP or FCoE HBA)

An FCP or FCoE is required for FC-based boot LUNs. Specific configuration for HBAs differs from vendor to vendor, but generally complies with the following workflow:

1. Boot the server to the HBA BIOS.
2. Enable the device as a boot device. (It may be necessary to enable the BIOS.)
3. Make a note of the WWPN.
4. If necessary, edit the NetApp igroup to include the WWPN.
5. Scan the bus for devices.
6. Select the device from which to boot.
7. Reboot.

   RHEL 6 recognizes the device on reboot and during the installation process.

Appendix B: Ports to Allow Through the Firewall

The following ports must be allowed through any relevant firewalls (host and network) when working with RHEL 6 KVM, Red Hat Enterprise Virtualization, Snap Creator, and any directory services, as described in Table 15 through Table 18.

Table 15) Ports required by RHEV-M.

Port	Protocol	Description
-	ICMP	Ping
22	TCP	SSH
80, 443	TCP	HTTP, HTTPS

Table 16) Ports required by hypervisors.

Port	Protocol	Description
22	TCP	SSH
80, 443	TCP	HTTP, HTTPS
111	TCP, UDP	Portmap
123	TCP	NTP
16514	TCP	libvirt
3260	TCP, UDP	iSCSI (optional)
53, 5353	TCP, UDP	DNS, mDNS
5432	TCP, UDP	PostgreSQL
54321	TCP	KVM interhost communication
5634-6166	TCP	Remote guest access
5900-5910	TCP	VNC consoles (optional)
5989	TCP, UDP	CIMOM
32803, 662	TCP	NFS client
49152-49216	TCP	KVM migration

Port	Protocol	Description
67, 68	TCP, UDP	DHCP
8080	TCP	Snap Creator portal
9090	TCP	Snap Creator agent
N/A	N/A	ICMP

Table 17) Ports required by Snap Creator.

Port	Protocol	Description
8080	TCP	Snap Creator portal
9090	TCP	Snap Creator agent

Table 18) Ports required by directory server.

Port	Protocol	Description
88, 464	TCP, UDP	Kerberos
389, 636	TCP	LDAP, LDAP over SSL

Appendix C: Making a Template Generic

This section describes the high-level steps to make an existing boot LUN or existing virtual machine usable as a template or golden image. Additional steps may need to be taken, depending on the application and other configurations to meet business requirements.

Removing Configuration Artifacts

This section describes static configuration artifacts and dynamic configuration artifacts:

- Static configuration artifacts are items that require explicit reconfiguration when a clone of a template is created. A server's host name is an example of an item that must be configured when a cloned RHEL 6 KVM host or guest is booted up for the first time.
- Dynamic configuration artifacts are items that automatically configure themselves. For example, if the SSH host keys are removed, they are automatically regenerated when a cloned RHEL 6 KVM host or guest is booted up for the first time.

If any other configurations are needed to support additional applications or business requirements, determine whether those configurations are static or dynamic. Also consider that many if not all of the static configuration artifacts could be converted to dynamic artifacts by using automation tools.

Perform the following steps to remove configuration artifacts.

1. Strip all static configuration artifacts.

   Note: These items need to be reconfigured when the RHEL 6 KVM host comes back up, either manually or by using a script. There may be additional items that need to be stripped out and/or accounted for, depending on business and/or application requirements.

2. Strip the host name, gateway, and IP information (`/etc/hosts`, `/etc/sysconfig/network`, `/etc/sysconfig/network-scripts/ifcfg-{eth*,br*,bond*}`).

3. Strip MAC addresses from Ethernet configuration files (`/etc/sysconfig/network-scripts/ifcfg-{eth*,br*,bond*}`).
4. Strip the system ID if registered to RHN (including RHN satellite and RHN proxy) (`/etc/sysconfig/rhn/systemid`).
5. Strip the iSCSI initiator name in `/etc/iscsi/initiatorname.conf` (can be regenerated with the `iscsi-name` command).
6. Replace the LUN WWID with either a NetApp friendly wildcard or a full wildcard in the multipath configuration file. `etc/multipath.conf. wwid 360a98000572d4273685a664462667a36` becomes `wwid 360a9*` (RHEL 6 KVM boot LUN only).
7. Clear out multipath bindings (RHEL 6) (`/etc/multipath/bindings`).
8. Rebuild `initramfs` using `dracut` (this must happen after editing the LUN WWID and multipath bindings).
9. Use the e2label tool to label the boot device, then edit `/etc/fstab` so that the boot device is identified by the label and not by a UUID (RHEL 6) or path (RHEL 5).
10. Strip out all dynamic configuration artifacts.

    Note: These items are recreated automatically when the RHEL 6 KVM host reboots. There may be additional items that need to be stripped out and/or accounted for, depending on business and/or application requirements.

11. Clear out the LVM cache (`/etc/lvm/cache/*`).
12. Remove the UDEV rule for the Ethernet device assignment (RHEL 6) (`/etc/udev/rules.d/70-persistent-net.rules`).
13. Remove the remaining persistent UDEV rules (RHEL 6) (`/etc/udev/rules.d/*-persistent-*.rules`).
14. Remove the SSH host keys (`/etc/ssh/ssh_host*`).

File System Alignment for Older RHEL Versions

In the case of RHEL-based VMs (versions 3, 4, and 5), the `%pre` section of the Kickstart file is used to create partitions that enable proper file system alignment with the underlying storage. The disk layout matches the aligned partitions in the main section of the Kickstart file.

The following example creates two partitions, the first for `/boot` and the second for everything else. Both partitions start on sectors that align correctly with the underlying storage.

1. Add the following lines at the end of the Kickstart file that is to be used to create virtual machines and/or virtual machine templates.

```
%pre
 parted /dev/sda mklabel msdos
 parted /dev/sda mkpart primary ext3 64s 208718s
 parted /dev/sda mkpart primary 208720s 100%
 parted /dev/sda set 2 lvm on
```

2. In the main section of the Kickstart file, edit the disk layout as follows.

```
zerombr yes
##clearpart --linux --drives=sda  ## comment out or remove
part /boot --fstype ext3 --onpart sda1
part pv.2 —onpart sda2
volgroup VolGroup00 --pesize=32768 pv.2
logvol swap --fstype swap --name=LogVol01 --vgname=VolGroup00 --size=1008 \--grow --maxsize=2016
logvol / --fstype ext3 --name=LogVol00 --vgname=VolGroup00 --size=1024 --grow
```

This results in both partitions on the virtual disk being properly aligned with the underlying NetApp storage.

For more information about file system alignment, see TR-3747: Best Practices for File System Alignment in Virtual Environments.

File System Alignment for Older Windows Versions

An altered Kickstart file can be used to properly align a virtual disk in preparation for a Windows guest installation (for Windows operating systems prior to Windows Server 2008).

Note: Kickstart can perform only the alignment and not the Windows installation. Although it introduces an extra step in the creation of a Windows VM template, it is still much faster than manually adjusting partitions after the VM is created.

The following example creates a single NTFS partition that takes up the entire disk but starts on a sector that enables proper file system alignment.

1. Create a typical Kickstart file (copy a known good Kickstart file) that contains valid information and sections. It doesn't matter what the content is, but it must be a valid Kickstart file.
2. Add the following lines at the end of the Kickstart file that is to be used to create Windows virtual machines and/or Windows virtual machine templates.

```
%pre
parted /dev/sda mklabel msdos
parted /dev/sda mkpart primary NTFS 128s 100%
chvt 3
echo "####################################"
echo "# Reboot with Windows Install DVD #"
echo "####################################"
sleep 5
exit
```

3. The Kickstart process parses and execute the partition creation in the `%pre` section, but exit the Kickstart process before attempting to install anything.
4. When the `Reboot` message appears, reboot the server with the Windows installation DVD. When the Windows installer runs, it recognizes the NTFS partition and permits installation to it.

References

The following references were used in this technical report.

- Operations Manager Administration Guide for Use with DataFabric Manager Server 4.0
  http://now.netapp.com/NOW/knowledge/docs/DFM_win/rel40/html/software/opsmgr/frameset.html
- TR-3446: SnapMirror Async Overview and Best Practices Guide
  http://media.netapp.com/documents/tr-3446.pdf
- TR-3747: Best Practices for File System Alignment in Virtual Environments
  http://media.netapp.com/documents/tr-3747.pdf
- TR-3800: Fibre Channel over Ethernet (FCoE) End-to-End Deployment Guide
  http://media.netapp.com/documents/TR-3800.pdf
- TR-3710: Operations Manager, Provisioning Manager, and Protection Manager Best Practices Guide
  http://www.netapp.com/us/media/tr-3710.pdf
  TR-4080: Best Practices for Scalable SAN in Clustered Data ONTAP 8.1http://www.netapp.com/us/system/pdf-reader.aspx?pdfuri=tcm:10-60386-16&m=tr-4080.pdf
- TR-4067: NFSv3/v4 in Clustered Data ONTAP 8.1 Implementation Guide

http://www.netapp.com/us/system/pdf-reader.aspx?pdfuri=tcm:10-61288-16&m=tr-4067.pdf

- TR-4104: RHEL 6, KVM, and Clustered Data ONTAP: Best Practices Guide
  http://www.netapp.com/us/system/pdf-reader.aspx?pdfuri=tcm:10-101510-16&m=tr-4104.pdf
- Data ONTAP 8.1 Clustered Data ONTAP Physical Storage Management Guide
  https://library.netapp.com/ecm/ecm_get_file/ECMM1277828
- Fibre Channel and iSCSI Configuration Guide
  https://library.netapp.com/ecm/ecm_get_file/ECMM1280845
- TR-4037: Introduction to NetApp Infinite Volume
  http://www.netapp.com/us/media/tr-4037.pdf
- TR-3786: A Thorough Introduction to 64-Bit Aggregates
  http://media.netapp.com/documents/tr-3786.pdf
- Data ONTAP 8.1 SAN Configuration Guide for Clustered Data ONTAP
  https://library.netapp.com/ecm/ecm_get_file/ECMP1140384
- Data ONTAP 8.1 Clustered Data ONTAP Block Access Management Guide for iSCSI and FC
  https://library.netapp.com/ecm/ecm_get_file/ECMM1277821
- Data ONTAP 8.1 Clustered Data ONTAP Vserver Administrator Capabilities Overview Guide
  https://library.netapp.com/ecm/ecm_get_file/ECMM1277819
- Data ONTAP 8.1 Clustered Data ONTAP High-Availability Configuration Guide
  https://library.netapp.com/ecm/ecm_get_file/ECMM1277825
- OnCommand System Manager 2.0.1
  https://library.netapp.com/ecm/ecm_get_file/ECMP1119584
- RHEV 3.1 Administration Guide
  https://access.redhat.com/knowledge/docs/en-US/Red_Hat_Enterprise_Virtualization/3.1/html/Administration_Guide/index.html
- RHEV 3.1 Hypervisor Deployment Guide
- https://access.redhat.com/knowledge/docs/en-US/Red_Hat_Enterprise_Virtualization/3.1/html/Hypervisor_Deployment_Guide/index.html
- Red Hat Enterprise Linux 6 Deployment Guide
  http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/6/html/Deployment_Guide/index.html
- Red Hat Enterprise Linux 6 Installation Guide
  http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/6/html/Installation_Guide/index.html
- Red Hat Enterprise Linux 6 Storage Administration Guide
  http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/6/html/Storage_Administration_Guide/
- Red Hat Enterprise Virtualization 3 Administration Guide
  http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Virtualization/3.0/html/Administration_Guide/index.html
- Red Hat Enterprise Virtualization 3 Installation Guide
  https://access.redhat.com/knowledge/docs/en-US/Red_Hat_Enterprise_Virtualization/3.1/html/Installation_Guide/
- SPEC's Benchmarks and Published Results
  http://www.spec.org/benchmarks.html#virtual

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 provides no representations or warranties regarding the accuracy, reliability, or serviceability of any information or recommendations provided in this publication, or with respect to any results that may be obtained by the use of the information or observance of any recommendations provided herein. The information in this document is distributed AS IS, and the use of this information or the implementation of any recommendations or techniques herein is a customer's responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. This document and the information contained herein may be used solely in connection with the NetApp products discussed in this document.

Go further, faster®

NetApp

www.netapp.com