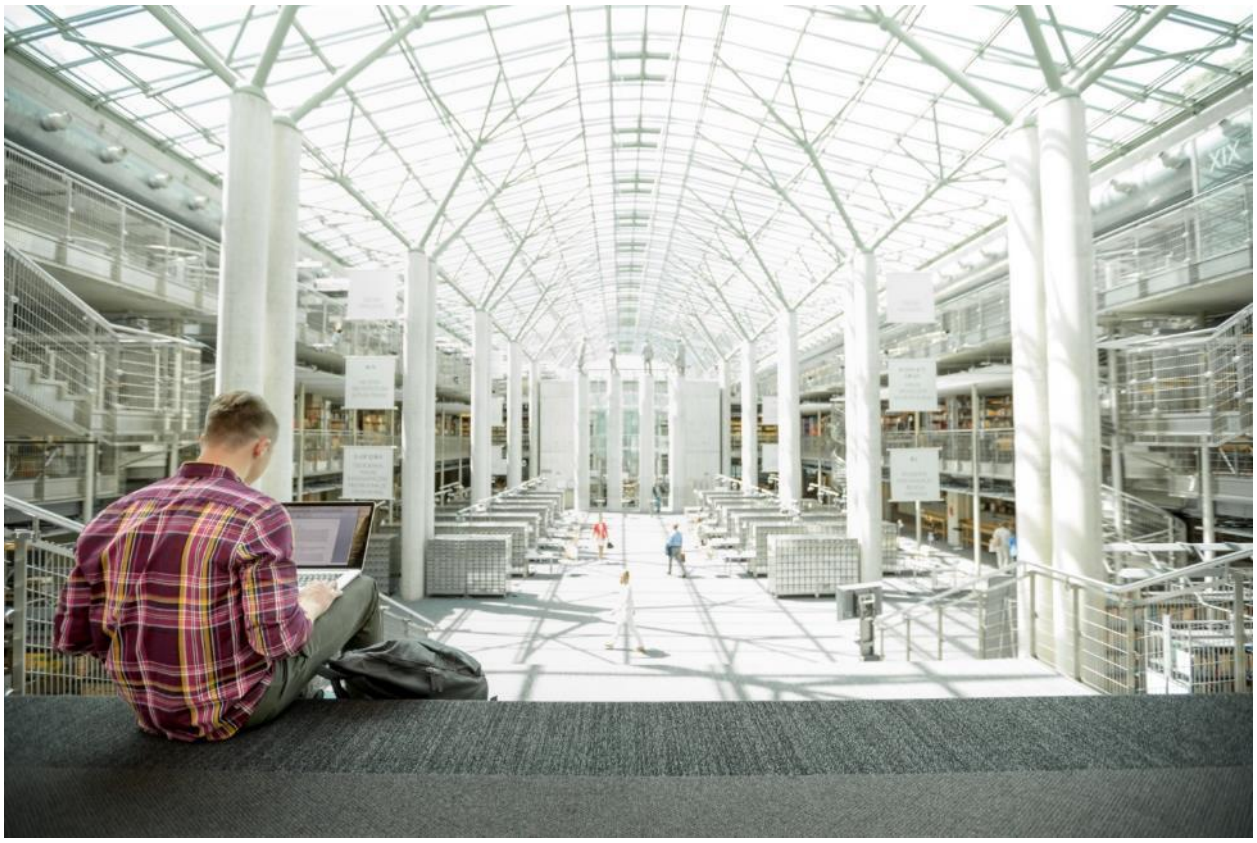


FlexPod Datacenter with Microsoft SQL Server 2017 on Linux Virtual Machine Running on VMware and Hyper-V Design Guide

Last Updated: December 5, 2019

CISCO VALIDATED DESIGN

About the Cisco Validated Design Program

The Cisco Validated Design (CVD) program consists of systems and solutions designed, tested, and documented to facilitate faster, more reliable, and more predictable customer deployments. For more information, go to:

http://www.cisco.com/go/designzone.

Table of Contents

Executive Summary

Cisco Validated Designs consist of systems and solutions that are designed, tested, and documented to facilitate and improve customer deployments. These designs incorporate a wide range of technologies and products into a portfolio of solutions that have been developed to address the business needs of our customers.

This document describes the Cisco and NetApp FlexPod solution, which is a validated approach for deploying Cisco and NetApp technologies as shared cloud infrastructure. This validated design provides a framework for deploying SQL Server 2017 on Linux Virtual Machine running on VMware vSphere and Windows Server 2016 Hyper-V based clusters. The detailed deployment steps are explained in the FlexPod Datacenter with Microsoft SQL Server 2017 on Linux VM Running on VMware and Hyper-V Deployment Guide.

FlexPod is a leading integrated infrastructure supporting broad range of enterprise workloads and use cases. This solution enables customers to quickly and reliably deploy SQL Server 2017 databases on Linux virtual machines running on VMware vSphere and Windows Hyper-V based private cloud on integrated infrastructure.

The recommended solution architecture is built on Cisco UCS using the unified software release to support the Cisco UCS hardware platforms including Cisco UCS B Series Blade Servers, Cisco UCS 6300 Fabric Interconnects, Cisco Nexus 9000 Series Switches, and NetApp All Flash Series Storage Arrays.

Program Summary

Cisco and NetApp have carefully validated and verified the FlexPod solution architecture and its many use cases while creating a portfolio of detailed documentation, information, and references to assist customers in transforming their data centers to this shared infrastructure model. This portfolio includes, but is not limited to the following items:

- Best practice architectural design
- Workload sizing and scaling guidance
- Implementation and deployment instructions
- Technical specifications (rules for what is a FlexPod configuration)
- Frequently asked questions and answers (FAQs)
- Cisco Validated Designs (CVDs) and NetApp Validated Architectures (NVAs) covering a variety of use cases

Cisco and NetApp have also built a robust and experienced support team focused on FlexPod solutions, from customer account and technical sales representatives to professional services and technical support engineers. The support alliance between NetApp and Cisco gives customers and channel services partners direct access to technical experts who collaborate with cross vendors and have access to shared lab resources to resolve potential issues.

FlexPod supports tight integration with virtualized and cloud infrastructures, making it the logical choice for long-term investment. FlexPod also provides a uniform approach to IT architecture, offering a well-characterized and documented shared pool of resources for application workloads. FlexPod delivers operational efficiency and consistency with the versatility to meet a variety of SLAs and IT initiatives, including:

- Application rollouts or application migrations
- Business continuity and disaster recovery
- Desktop virtualization
- Cloud delivery models (public, private, hybrid) and service models (IaaS, PaaS, SaaS)
- Asset consolidation and virtualization

Solution Overview

The current IT industry is witnessing vast transformations in the datacenter solutions. In the recent years, there is a considerable interest towards pre-validated and engineered datacenter solutions. Introduction of virtualization technology in the key areas has impacted the design principles and architectures of these solutions in a big way. It has opened the doors for many applications running on bare metal systems to migrate to these new virtualized integrated solutions.

With Microsoft SQL Server 2017, Microsoft has made a big announcement to support SQL Server deployments on Linux Operating systems. SQL Server 2017 on Linux platforms brings in support for most of the major features that are currently supported by SQL in a Windows platform. Microsoft claims that database performance of SQL on Linux should be similar as that of SQL in Windows. SQL on Linux is as secured as SQL in Windows platform. High availability features such as Always On Availability Groups and Failover Cluster Instance are well supported and integrated in Linux operating systems.

This enablement relieves Windows centric deployments of SQL Server and offers flexibility to customers to choose and deploy SQL Server databases on variety of widely used Linux operating systems.

Highlights of this Solution

The following software and hardware products distinguish this reference architecture from others:

- SQL Server 2017 deployment on RHEL 7.4 Virtual Machines in FlexPod Datacenter
- SQL Always On Availability Group configuration for high availability of databases
- Cisco UCS B200 M5 Blade Servers
- NetApp All Flash A300 storage with Data ONTAP 9.5 and NetApp SnapCenter 4.1.1 for virtual machine backup and recovery
- 40G end-to-end iSCSI networking and storage connectivity
- VMware vSphere 6.7 and Windows Server 2016 Hyper-V

iSCSI-Based Solution

FlexPod Datacenter is a flexible architecture that can suit any customer requirement. It allows you to choose a SAN protocol based on workload requirements or hardware availability. This solution in this document highlights a 40GbE iSCSI end-to-end deployment utilizing the Cisco Nexus 93180YC and Cisco UCS 6300 Series Fabric Interconnects. This design enables the workload virtual machines to have a 40 GbE end-to-end connectivity for both management and SAN traffic.

Use of iSCSI protocol end-to-end makes solution simple to manage, and hence lowers TCO. However, this solution can also be extended to other protocols such as FC/FCoE (subject to version compatibility) by adding necessary Storage Area Network switches and appropriate connections and configuration updates.

This design is suitable for customers who use IP-based architecture in their datacenter, enabling them to upgrade to end-to-end 40GbE. The components used in this design are:

- NetApp AFF A300 storage controllers
  - High Availability pair in switchless cluster configuration
  - 40GbE adapter used in the expansion slot of each storage controller
  - ONTAP 9.5
- Cisco Nexus 9332PQ switches
  - Pair of switches in vPC configuration
- Cisco UCS 6332-16UP Fabric Interconnect
- 40Gb Unified Fabric
- Cisco UCS 5108 Chassis
  - Cisco UCS 2304 IOM
  - Cisco UCS B200 M5 servers with VIC 1340

Figure 1 FlexPod Datacenter with Cisco UCS 6332-16UP Fabric Interconnects for End-to-End 40GbE iSCSI

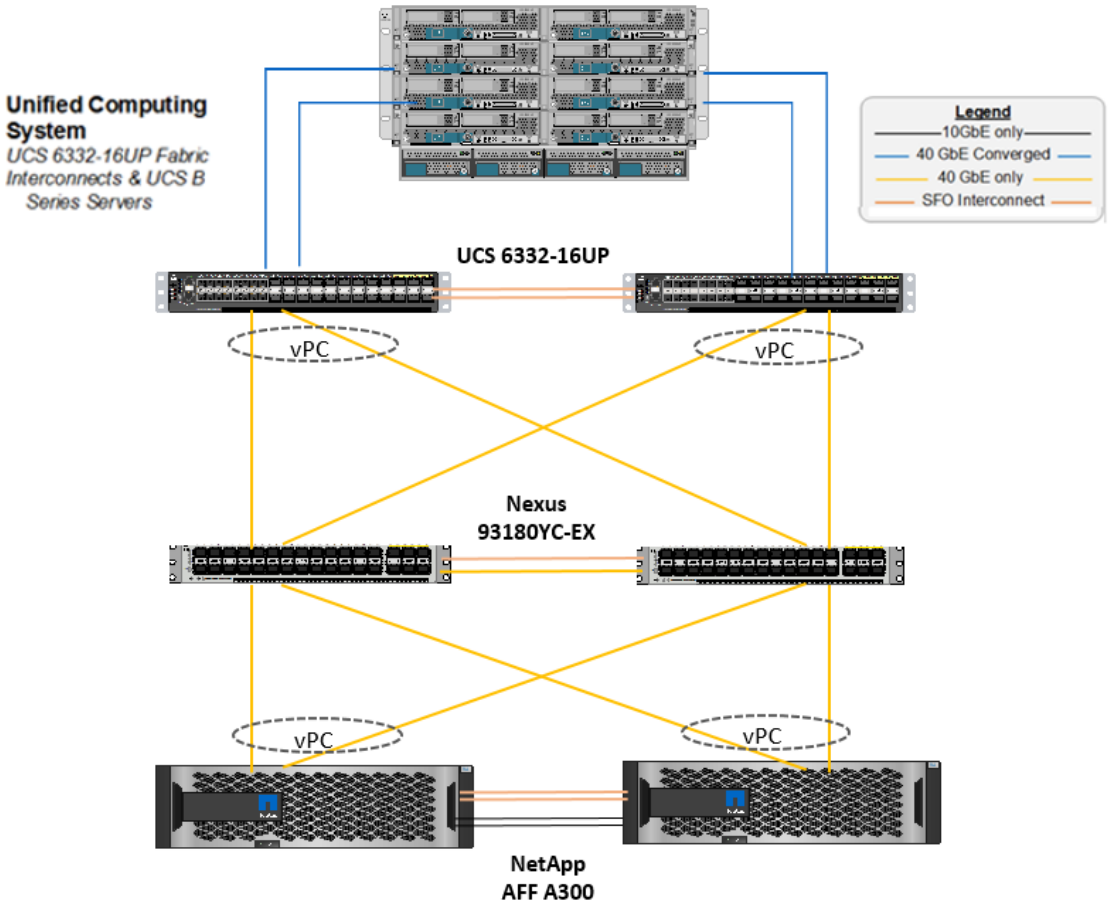


Figure 1 illustrates the FlexPod Datacenter topology that supports the end-to-end 40GbE iSCSI design. The Cisco UCS 6300 Fabric Interconnect FlexPod Datacenter model enables a high-performance, low-latency, and lossless fabric supporting applications with these elevated requirements. The 40GbE compute and network fabric increases the overall capacity of the system while maintaining the uniform and resilient design of the FlexPod solution.

Audience

The audience for this document includes, but is not limited to; sales engineers, field consultants, professional services, database administrators, IT managers, partner engineers, and customers who want to take advantage of an infrastructure built to deliver IT efficiency and enable IT innovation.

Technology Overview: Cisco Nexus 9000

Cisco Nexus series switches provide an Ethernet switching fabric for communications between the Cisco UCS , NetApp Storage controllers, and the rest of the customer's network. There are many factors to take into account when choosing the main data switch in this type of architecture to support both the scale and the protocols required for the resulting applications. All Nexus switch models including the Nexus 5000 and Nexus 7000 are supported in this design and may provide additional features such as FCoE or OTV. However, be aware that there may be slight differences in setup and configuration based on the switch used. The validation for the 40GbE iSCSI deployment leverages the Cisco Nexus 9000 series switches, which deliver high performance 40GbE ports, density, low latency, and exceptional power efficiency in a broad range of compact form factors.

Many of the most recent single-site FlexPod designs also use this switch due to the advanced feature set and the ability to support Application Centric Infrastructure (ACI) mode. When leveraging ACI fabric mode, the Nexus 9000 series switches are deployed in a spine-leaf architecture. Although the reference architecture covered in this design does not leverage ACI, it lays the foundation for customer migration to ACI in the future, and fully supports ACI today if required.

For more information, go to: Cisco Nexus 9000 Series Switches

This FlexPod design deploys a single pair of Nexus 9000 top-of-rack switches within each placement, using the traditional standalone mode running NX-OS.

The traditional deployment model delivers numerous benefits for this design:

- High performance and scalability with L2 and L3 support per port (Up to 60Tbps of non-blocking performance with less than 5 microsecond latency)
- Layer 2 multipathing with all paths forwarding through the Virtual port-channel (vPC) technology
- VXLAN support at line rate
- Advanced reboot capabilities include hot and cold patching
- Hot-swappable power-supply units (PSUs) and fans with N+1 redundancy

Cisco Nexus 9000 provides Ethernet switching fabric for communications between the Cisco UCS domain, the NetApp storage system and the enterprise network. In the FlexPod design, Cisco UCS Fabric Interconnects and NetApp storage systems are connected to the Cisco Nexus 9000 switches using virtual Port Channels.

Virtual Port Channel

A virtual Port Channel allows links that are physically connected to two different Cisco Nexus 9000 Series devices to appear as a single Port Channel. In a switching environment, vPC provides the following benefits:

- Allows a single device to use a Port Channel across two upstream devices
- Eliminates Spanning Tree Protocol blocked ports and use all available uplink bandwidth
- Provides a loop-free topology
- Provides fast convergence if either one of the physical links or a device fails

- Helps ensure high availability of the overall FlexPod system

Figure 2 Cisco 9000 Nexus Connections

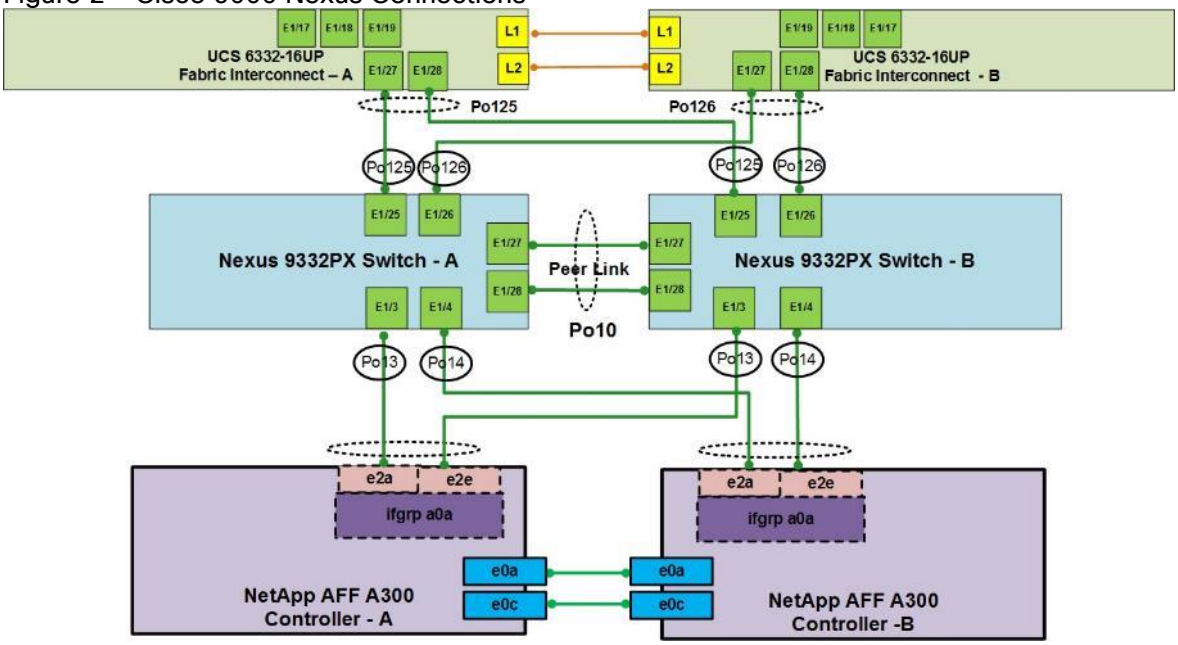


Figure 2 illustrates the connections between Cisco Nexus 9000, Cisco UCS Fabric Interconnects and NetApp AFF A300. vPC requires a " peer link" which is shown as port channel 10 in this diagram. In addition to the vPC peer-link, the vPC peer keepalive link is a required component of a vPC configuration. The peer keepalive link allows each vPC enabled switch to monitor the health of its peer. This link accelerates convergence and reduces the occurrence of split-brain scenarios. In this validated solution, the vPC peer keepalive link uses the out-of-band management network. This link is not shown in Figure 2.

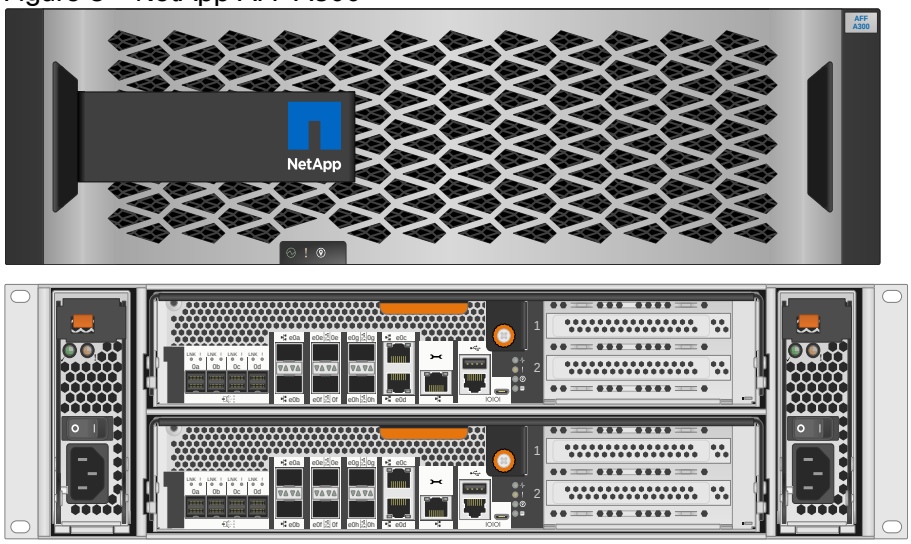
Technology Overview: NetApp A-Series All Flash FAS

NetApp AFF A300 Storage

With the new A-Series All Flash FAS (AFF) controller lineup, NetApp provides industry leading performance while continuing to provide a full suite of enterprise-grade data management and data protection features. The A-Series lineup offers double the IOPS, while decreasing the latency.

This solution utilizes the NetApp AFF A300. This controller provides high performance benefits of 40GbE and all flash SSDs, while taking up only 5U of rack space. Configured with 24 x 3.8TB SSD, the A300 provides ample performance and over 60TB effective capacity. This makes it an ideal controller for a shared workload converged infrastructure. For situations where more performance is needed, the A700s would be an ideal fit.

Figure 3 NetApp AFF A300



NetApp ONTAP 9.5

ONTAP 9.5 is the data management software that is used with the NetApp all-flash storage platforms in the solution design. ONTAP software offers unified storage for applications that read and write data over block or file-access protocols in storage configurations that range from high-speed flash to lower-priced spinning media or cloud-based object storage.

ONTAP implementations can run on NetApp engineered FAS or AFF appliances, on commodity hardware (ONTAP Select), and in private, public, or hybrid clouds (NetApp Private Storage and Cloud Volumes ONTAP). Specialized implementations offer best-in-class converged infrastructure as featured here as part of the FlexPod Datacenter solution and access to third-party storage arrays (NetApp FlexArray virtualization).

Together these implementations form the basic framework of the NetApp Data Fabric, with a common software-defined approach to data management and fast, efficient replication across platforms. FlexPod and ONTAP architectures can serve as the foundation for both hybrid cloud and private cloud designs.

The following few sections provide an overview of how ONTAP 9.5 is an industry-leading data management software architected on the principles of software defined storage.

ONTAP 9.1 provides many key features that optimize SSD performance and endurance, including the following:

- Coalesced writes to free blocks to maximize flash performance and longevity
- Flash-specific read-path optimizations that enable consistent low latency
- Advanced drive partitioning to increase storage efficiency, increasing usable capacity by almost 20%
- Support for multi-stream writes to increase write performance to SSDs

NetApp Storage Virtual Machine (SVM)

A NetApp ONTAP cluster serves data through at least one and possibly multiple storage virtual machines (SVMs; formerly called Vservers). An SVM is a logical abstraction that represents the set of physical resources of the cluster. Data volumes and network logical interfaces (LIFs) are created and assigned to an SVM and might reside on any node in the cluster to which the SVM has been given access. An SVM might own resources on multiple nodes concurrently, and those resources can be moved non-disruptively from one node in the storage cluster to another. For example, a NetApp FlexVol flexible volume can be non-disruptively moved to a new node and aggregate, or a data LIF can be transparently reassigned to a different physical network port. The SVM abstracts the cluster hardware, and thus it is not tied to any specific physical hardware.

An SVM can support multiple data protocols concurrently. Volumes within the SVM can be joined together to form a single NAS namespace, which makes all of an SVM's data available through a single share or mount point to NFS and CIFS clients. SVMs also support block-based protocols, and LUNs can be created and exported by using iSCSI, FC, or FCoE. Any or all of these data protocols can be configured for use within a given SVM. Storage administrators and management roles can also be associated with SVM, which enables higher security and access control, particularly in environments with more than one SVM, when the storage is configured to provide services to different groups or set of workloads.

Because it is a secure entity, an SVM is only aware of the resources that are assigned to it and has no knowledge of other SVMs and their respective resources. Each SVM operates as a separate and distinct entity with its own security domain. Tenants can manage the resources allocated to them through a delegated SVM administration account. Each SVM can connect to unique authentication zones such as Active Directory, LDAP, or NIS. A NetApp cluster can contain multiple SVMs. In this design, we have aligned our SVM design to the intended administrative functions, opting for three SVMs: one for infrastructure, one for SQL workload in VMware environment and one for SQL workload in Hyper-V environment. This allows administrators of the application to access only the dedicated SVMs and associated storage, increasing manageability and reducing risk. Larger orginizations with dedicated teams might opt to further segregate the administrative domains.

Storage Efficiencies

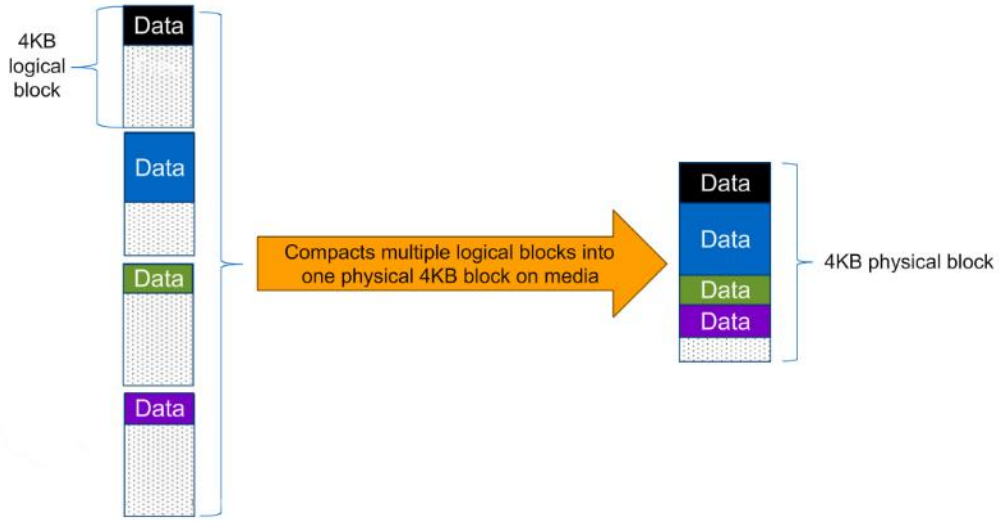
Storage efficiency has always been a primary architectural design point of ONTAP data management software. A wide array of features allows you to store more data using less space. In addition to deduplication and compression, you can store your data more efficiently by using features such as unified storage, multitenancy, thin provisioning, and utilize NetApp Snapshot technology.

Note that there is no performance penalty when using thin provisioning with ONTAP systems; data is written to available space so that write performance and read performance are maximized. Despite this fact, some products such as Microsoft failover clustering or other low-latency applications might require guaranteed or fixed provisioning, and it is wise to follow these requirements to augment support problems.

Starting with ONTAP 9, NetApp guarantees that the use of NetApp storage efficiency technologies on AFF systems reduce the total logical capacity used to store customer data by 75 percent, a data reduction ratio of 4:1. This space reduction is enabled by a combination of several different technologies, including deduplication, compression, and compaction.

Compaction, which was introduced in ONTAP 9, is the latest patented storage efficiency technology released by NetApp. In the ONTAP WAFL file system, all I/O takes up 4KB of space, even if it does not actually require 4KB of data. Compaction combines multiple blocks that are not using their full 4KB of space together into one block. This one block can be more efficiently stored on the disk to save space. This process is illustrated in Figure 4.

Figure 4 Storage Efficiency



NetApp Volume Encryption

Data security continues to be an important consideration for customers purchasing storage systems. NetApp supported self-encrypting drives in storage clusters prior to ONTAP 9. However, in ONTAP 9, the encryption capabilities of ONTAP are extended by adding an Onboard Key Manager (OKM). The OKM generates and stores keys for each of the drives in ONTAP, allowing ONTAP to provide all functionality required for encryption out of the box. Through this functionality, sensitive data stored on disk is secure and can only be accessed by ONTAP.

Beginning with ONTAP 9.1, NetApp has extended the encryption capabilities further with NetApp Volume Encryption (NVE), a software-based mechanism for encrypting data. It allows a user to encrypt data at the per-volume level instead of requiring encryption of all data in the cluster, thereby providing more flexibility and granularity to the ONTAP administrators. This encryption extends to Snapshot copies and NetApp FlexClone volumes that are created in the cluster. One benefit of NVE is that it runs after the implementation of the storage efficiency features, and, therefore, it does not interfere with the ability of ONTAP to create space savings.

For more information about encryption in ONTAP, see the NetApp Power Encryption Guide in the NetApp ONTAP 9 Documentation Center.

Quality of Service

Quality of Service (QoS) allows for managing performance on an individual LUN, volume, or file. QoS can be used to limit an unknown or "bully" virtual machine or to make sure an important virtual machine gets sufficient performance resources.

FlexClone

NetApp FlexClone technology enables instantaneous cloning of a dataset without consuming any additional storage until cloned data differs from the original.

Snapshot Copies

ONTAP offers instant Snapshot copies of a Database, virtual machine or datastore with zero performance impact to create or use the Snapshot copy. Snapshot copies are valuable by themselves to make a restoration point of a virtual machine prior to patching or for simple data protection. Note that these are different from VMware (consistency) snapshots, which are generally not recommended due to performance and other impacts. The easiest way to make an ONTAP Snapshot copy is to use the SnapCenter Plug-In for VMware vSphere to back up VMs and datastores.

SnapMirror (Data Replication)

NetApp SnapMirror is a replication technology for data replication across different sites, or within the same datacenter, or on-the-premises datacenter to cloud, or cloud to on-the-premises datacenter.

Virtual Volumes and Storage Policy-Based Management

NetApp was an early design partner with VMware in the development of vSphere Virtual Volumes (VVols), providing architectural input and early support for VVols and VMware vSphere APIs for Storage Awareness (VASA). Not only did this approach bring virtual machine granular storage management to VMFS, it also supported automation of storage provisioning through Storage Policy-Based Management. This approach allows storage architects to design storage pools with different capabilities that can be easily consumed by virtual machine administrators. ONTAP leads the storage industry in VVol scale, supporting hundreds of thousands of VVols in a single cluster, whereas enterprise array and smaller flash array vendors support as few as several thousand VVols per array. NetApp is also driving the evolution of VM granular management with upcoming capabilities in support of VASA 3.0.

NetApp SnapCenter

SnapCenter is a NetApp next-generation data protection software for tier 1 enterprise applications. SnapCenter, with its single-pane-of-glass management interface, automates and simplifies the manual, complex, and time-consuming processes associated with the backup, recovery, and cloning of multiple databases and other application workloads.

SnapCenter leverages technologies, including NetApp Snapshot copies, SnapMirror replication technology, SnapRestore data recovery software, and FlexClone thin cloning technology, that allow it to integrate seamlessly with technologies offered by Oracle, Microsoft, SAP, VMware, and MongoDB across FC, iSCSI, and NAS protocols. This integration allows IT organizations to scale their storage infrastructure, meet increasingly stringent SLA commitments, and improve the productivity of administrators across the enterprise.

SnapCenter is used in this solution for backup and restore of VMware virtual machines. SnapCenter has the capability to backup and restore SQL databases running on Windows OS. SnapCenter currently does not support SQL Server running on Linux OS, but that functionality is coming in near future. SnapCenter does not support the

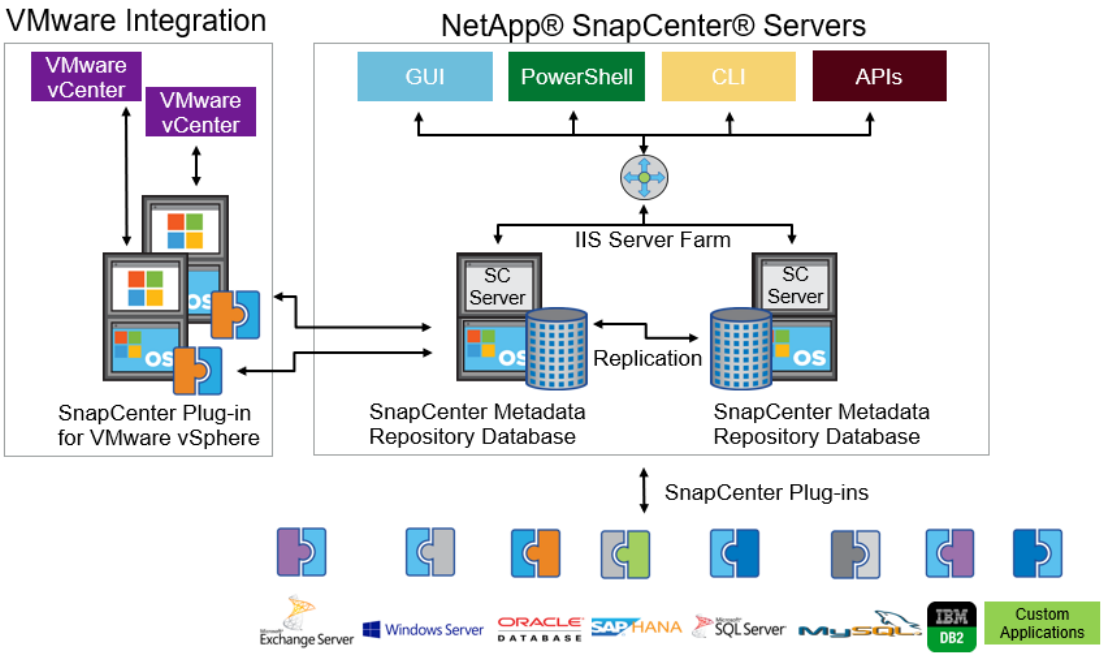
backup and restore of virtual machines running on Hyper-V. For Hyper-V virtual machine backup, NetApp SnapManager for Hyper-V can be used.

SnapCenter Architecture

SnapCenter is a centrally managed web-based application that runs on a Windows platform and remotely manages multiple servers that must be protected.

Figure 5 illustrates the high-level architecture of the NetApp SnapCenter Server.

Figure 5 SnapCenter Architecture



The SnapCenter Server has an HTML5-based GUI as well as PowerShell cmdlets and APIs.

The SnapCenter Server is high-availability capable out of the box, meaning that if one SnapCenter host is ever unavailable for any reason, then the second SnapCenter Server can seamlessly take over and no operations are affected.

The SnapCenter Server can push out plug-ins to remote hosts. These plug-ins are used to interact with an application, a database, or a file system. In most cases, the plug-ins must be present on the remote host so that application- or database-level commands can be issued from the same host where the application or database is running.

To manage the plug-ins and the interaction between the SnapCenter Server and the plug-in host, SnapCenter uses SM Service, which is a NetApp SnapManager web service running on top of Windows Server Internet Information Services (IIS) on the SnapCenter Server. SM Service takes all client requests such as backup, restore, clone, and so on.

The SnapCenter Server communicates those requests to SMCore, which is a service that runs co-located within the SnapCenter Server and remote servers and plays a significant role in coordinating with the SnapCenter plug-

ins package for Windows. The package includes the SnapCenter plug-in for Microsoft Windows Server and SnapCenter plug-in for Microsoft SQL Server to discover the host file system, gather database metadata, quiesce and thaw, and manage the SQL Server database during backup, restore, clone, and verification.

SnapCenter Virtualization (SCV) is another plug-in that manages virtual servers running on VMware and that helps in discovering the host file system, databases on virtual machine disks (VMDK), and raw device mapping (RDM).

SnapCenter Plug-In for VMware vSphere

SnapCenter Plug-In for VMware vSphere is a host-side component of the NetApp storage solution offering Backup, Restore and Cloning of databases, virtual machines and datastores.

Support for Virtualized Databases and File Systems

The Plug-in for VMware vSphere provides native backup, recovery, and cloning of virtualized applications (virtualized SQL and Oracle databases and Windows file systems) when using the SnapCenter GUI.

SnapCenter natively leverages the Plug-in for VMware vSphere for SQL Server on Windows Server, Oracle on Linux, and Windows file system data protection operations on virtual machine disks (VMDKs), raw device mappings (RDMs), and NFS datastores.

SnapCenter does not have a plug-in for SQL Server on Linux, yet.

SnapCenter Features

SnapCenter enables you to create application-consistent Snapshot copies and to complete data protection operations, including Snapshot copy-based backup, clone, restore, and backup verification operations. SnapCenter provides a centralized management environment, while using role-based access control (RBAC) to delegate data protection and management capabilities to individual application users across the SnapCenter Server and Windows hosts.

SnapCenter includes the following key features:

- A unified and scalable platform across applications and database environments and virtual and nonvirtual storage, powered by the SnapCenter Server
- Consistency of features and procedures across plug-ins and environments, supported by the SnapCenter user interface
- RBAC for security and centralized role delegation
- Application-consistent Snapshot copy management, restore, clone, and backup verification support from both primary and secondary destinations (NetApp SnapMirror and SnapVault)
- Remote package installation from the SnapCenter GUI
- Nondisruptive, remote upgrades
- A dedicated SnapCenter repository for faster data retrieval
- Load balancing implemented by using Microsoft Windows network load balancing (NLB) and application request routing (ARR), with support for horizontal scaling

- Centralized scheduling and policy management to support backup and clone operations
- Centralized reporting, monitoring, and dashboard views
- SnapCenter 4.1 support for data protection for VMware virtual machines, SQL Server Databases, Oracle Databases, MySQL, SAP HANA, MongoDB, and Microsoft Exchange
- SnapCenter Plug-in for VMware in vCenter integration into the vSphere Web Client. All virtual machine backup and restore tasks are preformed through the web client GUI

Using the SnapCenter Plug-in for VMware in vCenter you can:

- Create policies, resource groups and backup schedules for virtual machines
- Backup virtual machines, VMDKs, and datastores
- Restore virtual machines, VMDKs, and files and folders (on Windows guest OS)
- Attach and detach VMDK
- Monitor and report data protection operations on virtual machines and datastores
- Support RBAC security and centralized role delegation
- Support guest file or folder (single or multiple) support for Windows guest OS
- Restore an efficient storage base from primary and secondary Snapshot copies through Single File SnapRestore
- Generate dashboard and reports that provide visibility into protected versus unprotected virtual machines and status of backup, restore, and mount jobs
- Attach or detach virtual disks from secondary Snapshot copies
- Attach virtual disks to an alternate virtual machine

NetApp SnapManager for Hyper-V (SMHV)

SnapManager for Hyper-V provides you with a solution for data protection and recovery for Microsoft Hyper-V virtual machines residing on storage systems running ONTAP.

You can perform application-consistent and crash-consistent dataset backups according to dataset protection policies set by your backup administrator. You can also restore virtual machines from these backups. Reporting features enable you to monitor the status of the backups and get detailed information about your backup and restore jobs.

SnapManager for Hyper-V enables you to backup and restore multiple virtual machines across multiple hosts. You can create datasets and apply policies to them to automate backup tasks such as scheduling, retention, and replication. Following is the list of other tasks you can perform with SnapManager for Hyper-V:

- Group virtual machines into datasets that have the same protection requirements and apply policies to those datasets

- Backup and restore dedicated and clustered virtual machines residing on storage systems running ONTAP software
- Backup and restore virtual machines hosted on Cluster Shared Volumes (CSVs)
- Automate dataset backups using scheduling policies
- Perform on-demand backups of datasets
- Retain dataset backups for as long as you need them, using retention policies
- Update the SnapMirror destination location after a backup successfully finishes
- Specify custom scripts to run before or after a backup
- Restore virtual machines from backups
- Monitor the status of all scheduled and running jobs
- Manage hosts remotely from a management console
- Provide consolidated reports for dataset backup, restore, and configuration operations
- Perform a combination of crash-consistent and application-consistent backups
- Perform disaster recovery operations using PowerShell cmdlets

Refer to the NetApp Support site for more information about SnapManager for Hyper-V.

NetApp Virtual Storage Console

NetApp Virtual Storage Console (VSC) is shipped along with VASA Provider and SRA in a virtual appliance.

Virtual Storage Console

VSC is a vCenter plug-in that ensures ESXi host settings are optimal for using NetApp Storage over NAS or SAN. It uses best practices for quickly provisioning VMFS or NFS datastores. It includes both a VSC server appliance and user interface extensions for vCenter. VSC provides two elementary reports, one detailing Datastores and the other detailing virtual machines. It simplifies storage management and efficiency features, enhances availability, and reduces storage costs and operational overhead, whether using SAN or NAS.

NFS Plug-In for VMware VAAI

The NetApp NFS Plug-In for VMware is a plug-in for ESXi hosts that allows them to use VAAI features with NFS datastores on ONTAP. It supports copy offload for clone operations, space reservation for thick virtual disk files, and Snapshot copies for linked clones. Offloading copy operations to storage is not necessarily faster to complete, but offloads host resources such as CPU cycles, buffers, and queues. You may use VSC to install the plug-in on ESXi hosts.

VASA Provider for ONTAP

The VASA Provider for ONTAP supports the VMware vStorage APIs for Storage Awareness VASA framework. VASA Provider is combined with the VSC and SRA as a single virtual appliance for ease of deployment. VASA

Provider connects vCenter Server with ONTAP to aid in provisioning and monitoring virtual machine storage. It enables VMware Virtual Volumes (VVols) support, management of storage capability profiles and individual virtual machine VVol performance, and alarms for monitoring capacity and compliance with the profiles. VASA dashboard shows Datastore and Virtual Machine top 5 metrics which for IOPS, Latency, Uptime or Space Used.

Storage Replication Adapter

The SRA is used together with VMware Site Recovery Manager (SRM) to manage data replication between production and disaster recovery sites and test the DR replica non-disruptively. It helps automate the tasks of discovery, recovery, and reprotection. It includes both an SRA server appliance and an SRA adapter for the SRM server.

Technology Overview: Cisco Unified Computing System

Cisco UCS 6300 Fabric Interconnects

The Cisco UCS Fabric interconnects provide a single point for connectivity and management for the entire system. Typically deployed as an active-active pair, the system's fabric interconnects integrate all components into a single, highly available management domain controlled by Cisco UCS Manager (UCSM). The fabric interconnects manage all I/O efficiently and securely at a single point, resulting in deterministic I/O latency regardless of a server or virtual machine's topological location in the system.

The Fabric Interconnect provides both network connectivity and management capabilities for Cisco UCS. IOM modules in the blade chassis support power supply, along with fan and blade management. They also support port channeling and, thus, better use of bandwidth. The IOMs support virtualization-aware networking in conjunction with the Fabric Interconnects and Cisco Virtual Interface Cards (VIC).

FI 6300 Series and IOM 2304 provide a few key advantages over the existing products. FI 6300 Series and IOM 2304 support 40GbE / FCoE port connectivity that enables an end-to-end 40GbE / FCoE solution. Unified ports support 4/8/16G FC ports for higher density connectivity to SAN ports.

Cisco UCS Differentiators

Cisco Unified Computing System is revolutionizing the way servers are managed in the data center. The following are the unique differentiators of Cisco UCS and Cisco UCS Manager:

- Embedded Management—In Cisco UCS, the servers are managed by the embedded firmware in the Fabric Interconnects, eliminating need for any external physical or virtual devices to manage the servers.
- Unified Fabric—In Cisco UCS, from blade server chassis or rack servers to FI, there is a single Ethernet cable used for LAN, SAN and management traffic. This converged I/O results in reduced cables, SFPs and adapters which in turn reduce capital and operational expenses of the overall solution.
- Auto Discovery—By simply inserting the blade server in the chassis or connecting rack server to the fabric interconnect, discovery and inventory of compute resource occurs automatically without any management intervention. The combination of unified fabric and auto-discovery enables the wire-once architecture of Cisco UCS, where its compute capability can be extended easily while keeping the existing external connectivity to LAN, SAN and management networks.
- Policy Based Resource Classification—When a compute resource is discovered by Cisco UCS Manager, it can be automatically classified to a given resource pool based on policies defined. This capability is useful in multi-tenant cloud computing. This CVD showcases the policy-based resource classification of Cisco UCS Manager.
- Combined Rack and Blade Server Management—Cisco UCS Manager can manage B-Series blade servers and C-Series rack server under the same Cisco UCS domain. This feature, along with stateless computing makes compute resources truly hardware form factor agnostic.
- Model-based Management Architecture—Cisco UCS Manager Architecture and management database is model based and data driven. An open XML API is provided to operate on the management model. This enables easy and scalable integration of Cisco UCS Manager with other management systems.

- Policies, Pools, Templates–The management approach in Cisco UCS Manager is based on defining policies, pools and templates, instead of cluttered configuration, which enables a simple, loosely coupled, data driven approach in managing compute, network and storage resources.
- Loose Referential Integrity–In Cisco UCS Manager, a service profile, port profile or policies can refer to other policies or logical resources with loose referential integrity. A referred policy cannot exist at the time of authoring the referring policy or a referred policy can be deleted even though other policies are referring to it. This provides different subject matter experts to work independently from each-other. This provides great flexibility where different experts from different domains, such as network, storage, security, server and virtualization work together to accomplish a complex task.
- Policy Resolution–In Cisco UCS Manager, a tree structure of organizational unit hierarchy can be created that mimics the real-life tenants and/or organization relationships. Various policies, pools and templates can be defined at different levels of organization hierarchy. A policy referring to another policy by name is resolved in the organization hierarchy with closest policy match. If no policy with specific name is found in the hierarchy of the root organization, then special policy named "default" is searched. This policy resolution practice enables automation friendly management APIs and provides great flexibility to owners of different organizations.
- Service Profiles and Stateless Computing–a service profile is a logical representation of a server, carrying its various identities and policies. This logical server can be assigned to any physical compute resource as far as it meets the resource requirements. Stateless computing enables procurement of a server within minutes, which used to take days in legacy server management systems.
- Built-in Multi-Tenancy Support–The combination of policies, pools and templates, loose referential integrity, policy resolution in organization hierarchy and a service profiles based approach to compute resources makes Cisco UCS Manager inherently friendly to multi-tenant environment typically observed in private and public clouds.
- Extended Memory–the enterprise-class Cisco UCS B200 M5 blade server extends the capabilities of Cisco's Unified Computing System portfolio in a half-width blade form factor. The Cisco UCS B200 M5 harnesses the power of the latest Intel Xeon scalable processors product family CPUs with up to 3 TB of RAM- allowing huge virtual machine-to-physical server ratio required in many deployments or allowing large memory operations required by certain architectures like Big-Data.
- Virtualization Aware Network–VM-FEX technology makes the access network layer aware about host virtualization. This prevents domain pollution of compute and network domains with virtualization when virtual network is managed by port-profiles defined by the network administrators' team. VM-FEX also off-loads hypervisor CPU by performing switching in the hardware, thus allowing hypervisor CPU to do more virtualization related tasks. VM-FEX technology is well integrated with VMware vCenter, Linux KVM and Hyper-V SR-IOV to simplify cloud management.
- Simplified QoS–Even though Fiber Channel and Ethernet are converged in Cisco UCS fabric, built-in support for QoS and lossless Ethernet makes it seamless. Network Quality of Service (QoS) is simplified in Cisco UCS Manager by representing all system classes in one GUI panel.

Cisco Nexus 9000 Design and Best Practices

The Cisco Nexus 9000 best practices used in the validation of this FlexPod architecture are summarized below:

Cisco Nexus 9000 Features Enabled

- Link Aggregation Control Protocol (LACP part of 802.3ad)
- Cisco Virtual Port Channeling for link and device resiliency
- Cisco Discovery Protocol (CDP) for infrastructure visibility and troubleshooting

vPC Considerations

- Define a unique domain ID
- Set the priority of the intended vPC primary switch lower than the secondary (default priority is 32768)
- Establish peer keepalive connectivity. It is recommended to use the out-of-band management network (mgmt0) or a dedicated switched virtual interface (SVI)
- Enable vPC auto-recovery feature
- Enable peer-gateway. Peer-gateway allows a vPC switch to act as the active gateway for packets that are addressed to the router MAC address of the vPC peer allowing vPC peers to forward traffic
- Enable IP ARP synchronization to optimize convergence across the vPC peer link.
- A minimum of two 10 Gigabit Ethernet connections are required for vPC
- All port channels should be configured in LACP active mode

Spanning Tree Considerations

- The spanning tree priority was not modified. Peer-switch (part of vPC configuration) is enabled which allows both switches to act as root for the VLANs
- Loopguard is disabled by default
- BPDU guard and filtering are enabled by default
- Bridge assurance is only enabled on the vPC Peer Link.
- Ports facing the NetApp storage controllers and UCS are defined as " edge" trunk ports.

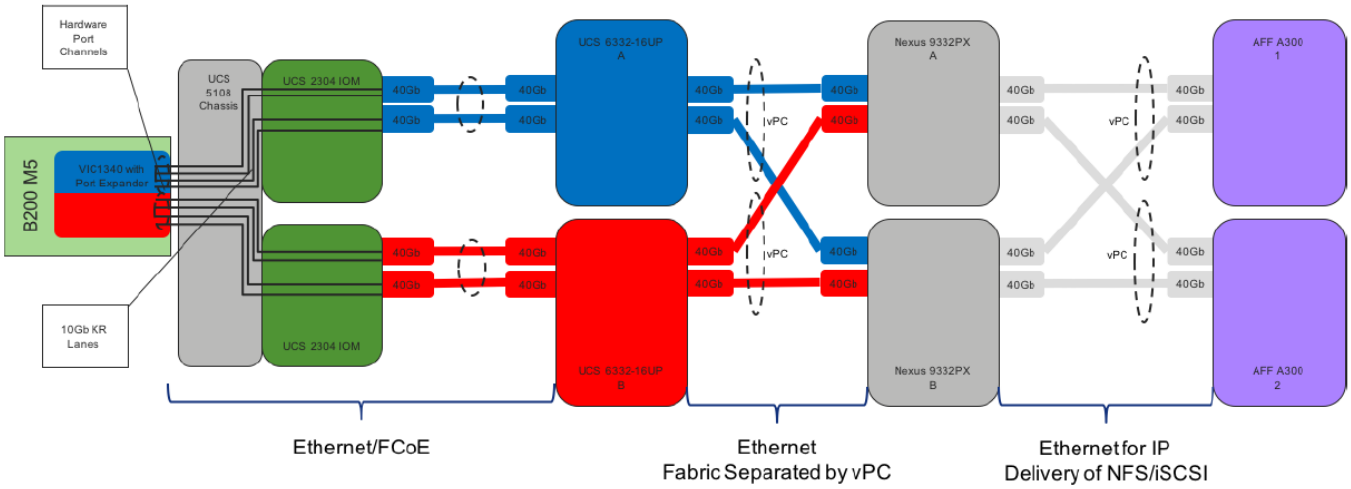
For configuration details, refer to the Cisco Nexus 9000 Series Switches Configuration guides: http://www.cisco.com/c/en/us/support/switches/nexus-9000-series-switches/products-installation-and-configuration-guides-list.html

Bringing Together 40Gb End-to-End

The Cisco Nexus 9000 is the key component bringing together the 40Gb capabilities of the other pieces of this design. vPCs extend to both the AFF A300 Controllers and the Cisco UCS 6332-16UP Fabric Interconnects. Passage of this traffic is shown in Figure 6, from left to right, is as follows:

- From the Cisco UCS B200 M5 server, equipped with a VIC 1340 adapter and a Port Expander card, allowing for 40Gb on each side of the fabric (A/B) into the server.
- Pathing through 10Gb KR lanes of the Cisco UCS 5108 Chassis backplane into the Cisco UCS 2304 IOM (Fabric Extender).
- Connecting from each IOM to the Fabric Interconnect with pairs of 40Gb uplinks automatically configured as port channels during chassis association.
- From the Cisco UCS 6332-16UP Fabric Interconnects into the Cisco Nexus 9332PX with a bundle of 40Gb ports presenting each side of the fabric from the Nexus pair as a common switch using a vPC.
- Ending at the AFF A 300 Controllers with 40Gb bundled vPCs from the Nexus switches now carrying both sides of the fabric.

Figure 6 vPC, AFF A300 Controllers, and Cisco UCS 6332-16UP Fabric Interconnect Traffic



NetApp Storage Design and Best Practices

Clustered Data ONTAP 9.5

Licenses on ONTAP

The following licenses are required for VSC, on storage systems that run ONTAP 9.5:

- Protocol licenses (NFS and iSCSI)
- NetApp FlexClone (for provisioning and cloning only)
- NetApp SnapRestore (for backup and recovery)
- The NetApp SnapManager Suite

Configuration Worksheet

Before running the setup script, complete the cluster setup worksheet from the ONTAP 9.5 Software Setup Guide in the ONTAP 9 Documentation Center. You must have access to the NetApp Support site to open the cluster setup worksheet.

Customize the cluster detail values with the information applicable to your deployment.

Table 1 ONTAP Software Installation Prerequisites

Cluster Detail	Cluster Detail Value
Cluster node 01 IP address	<node01-mgmt-ip>
Cluster node 01 netmask	<node01-mgmt-mask>
Cluster node 01 gateway	<node01-mgmt-gateway>
Cluster node 02 IP address	<node02-mgmt-ip>
Cluster node 02 netmask	<node02-mgmt-mask>
Cluster node 02 gateway	<node02-mgmt-gateway>
Data ONTAP 9.5 URL	<url-boot-software>

Pursuant to best practices, NetApp recommends the following command on the LOADER prompt of the NetApp controllers to assist with LUN stability during copy operations. To access the LOADER prompt, connect to the controller through serial console port or Service Processor connection and press Ctrl-C to halt the boot process when prompted: setenv bootarg.tmgr.disable_pit_hp 1

For more information about Windows Offloaded Data Transfers see: https://technet.microsoft.com/en-us/library/hh831628(v=ws.11).aspx.

High Availability

NetApp Storage Cluster design provides High Availability at every level; Cluster Nodes, backend storage connectivity, RAID-DP that can sustain two Disk Failures, physical connectivity to two physical networks from each node, and providing multiple data paths to storage LUNs and Volumes.

Secure Multitenancy

NetApp SVM provides a virtual storage array construct to separate security domain, policies, and virtual networking. It's recommended to create separate SVM for each tenant organization hosting data on storage cluster.

NetApp Storage Best Practices

Here are NetApp Storage best practices to consider:

- Always enable NetApp AutoSupport, which sends support summary information to NetApp through HTTPS.
- Make sure that a logical interface (LIF) is created for each SVM on each node in the ONTAP cluster for maximum availability and mobility. ALUA is used to parse paths and identify active optimized (direct) paths versus active non-optimized paths. ALUA is used for both FC/FCoE and iSCSI.
- Block protocols (iSCSI, FC, and FCoE) access LUNs by using LUN IDs and serial numbers, along with unique names (FC/FCoE use worldwide names [WWNNs and WWPNs], and iSCSI uses iSCSI qualified names [IQNs]). The path to LUNs inside the storage is meaningless to the block protocols and is not presented anywhere in the protocol. Therefore, a volume that contains only LUNs does not need to be internally mounted at all, and a junction path is not needed for volumes that contain LUNs used in datastores.
- If Challenge-Handshake Authentication Protocol (CHAP) is used in ESXi for target authentication, it must also be configured in ONTAP using the CLI (`vserver iscsi security create`) or with OnCommand System Manager (edit Initiator Security under Storage>SVMs>SVM Settings>Protocols>iSCSI).

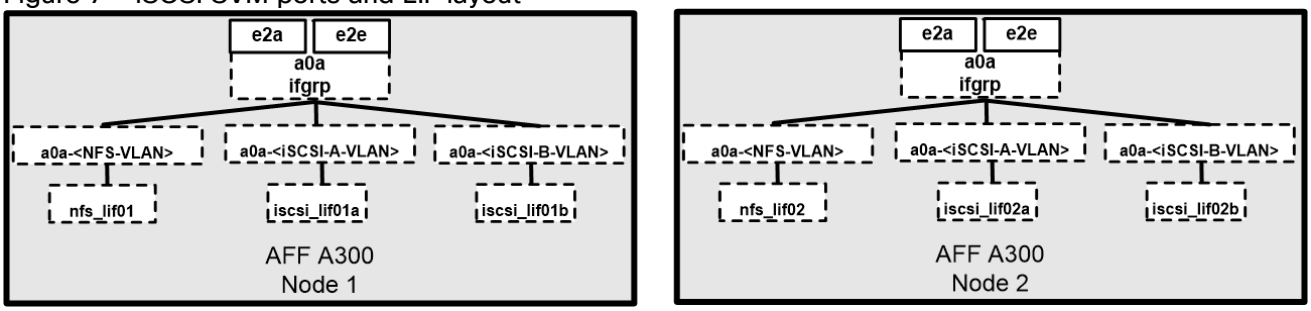
SAN Boot

NetApp recommends implementing SAN boot for Cisco UCS servers in the FlexPod Datacenter solution. Doing so enables the operating system to be safely secured by the NetApp AFF storage system, providing better performance. The design outlined in this solution uses iSCSI SAN boot.

In iSCSI SAN boot, each Cisco UCS server is assigned two iSCSI vNICs (one for each SAN fabric), which provide redundant connectivity all the way to the storage. The storage ports, in this example, e2a and e2e, which are connected to the Cisco Nexus switches, are grouped together to form one logical port called an interface group (igroup) (in this example, a0a). The iSCSI VLANs are created on the igroup, and the iSCSI logical interfaces (LIFs) are created on iSCSI port groups (in this example, a0a-<iSCSI-A-VLAN>). The iSCSI boot LUN is exposed to the

servers through the iSCSI LIF using igroups. This enables only the authorized server to have access to the boot LUN. Refer to Figure 7 for the port and LIF layout.

Figure 7 iSCSI SVM ports and LIF layout



Unlike NAS network interfaces, the SAN network interfaces are not configured to fail over during a failure. Instead, if a network interface becomes unavailable, the host chooses a new optimized path to an available network interface. ALUA is a standard supported by NetApp, provides information about SCSI targets, which allows a host to identify the best path to the storage.

Storage Efficiency and Thin Provisioning

NetApp has led the industry with storage efficiency innovation such as the first deduplication for primary workloads, and inline data compaction, which enhances compression and stores small files and I/Os efficiently. ONTAP supports both inline and background deduplication, as well as inline and background compression.

To realize the benefits of deduplication in a block environment, the LUNs must be thin provisioned. Although the LUN is still seen by the VM administrator as taking the provisioned capacity, the deduplication savings are returned to the volume to be used for other needs. NetApp recommends deploying these LUNs in FlexVol volumes that are also thin provisioned with a capacity that is two times the size of the LUN. When the LUN is deployed in this manner, the FlexVol volume acts merely as a quota. The storage consumed by the LUN is reported in FlexVol and its containing aggregate.

For maximum deduplication savings, consider scheduling background deduplication. However, these processes use system resources when running, so ideally should be scheduled during less active times (such as weekends) or run more frequently to reduce the amount of changed data to be processed. Automatic background deduplication on AFF systems has much less impact on foreground activities. Background compression (for hard disk–based systems) also consumes resources, so should only be considered for secondary workloads with limited performance requirements.

Quality of Service

Systems running ONTAP software may use the ONTAP storage QoS feature to limit throughput in MBps and/or I/Os per second (IOPS) for different storage objects such as files, LUNs, volumes, or entire SVMs. Adaptive QoS is used to set an IOPS floor (QoS min) and ceiling (QoS max), which dynamically adjust based on the datastore capacity and used space.

Throughput limits are useful in controlling unknown or test workloads before deployment to make sure they don't affect other workloads. They may also be used to constrain a bully workload after being identified. Minimum levels of service based on IOPS are also supported to provide consistent performance for SAN objects in ONTAP.

With an NFS datastore, a QoS policy may be applied to the entire FlexVol volume or individual VMDK files within it. With VMFS datastores (Clustered Shared volumes in Hyper-V) using ONTAP LUNs, the QoS policies may be applied to the FlexVol volume that contains the LUNs or individual LUNs, but not individual VMDK files because ONTAP has no awareness of the VMFS file system. When using VVols with VSC 7.1 or later, maximum QoS may be set on individual VMs using the storage capability profile. To assign a QoS policy to a LUN, including VMFS or Clustered Shared volume (CSV), the ONTAP SVM (displayed as Vserver), LUN path, and serial number can be obtained from the Storage Systems menu on the VSC home page. Select the storage system (SVM), then Related Objects > SAN. Use this approach when specifying QoS using one of the ONTAP tools.

The QoS maximum throughput limit on an object can be set in MBps and/or IOPS. If both are used, the first limit reached is enforced by ONTAP. A workload can contain multiple objects, and a QoS policy can be applied to one or more workloads. When a policy is applied to multiple workloads, the workloads share the total limit of the policy. Nested objects are not supported (for example, a file within a volume cannot each have their own policy). QoS minimums can only be set in IOPS.

Best Practices for SQL Server with NetApp Storage

Best practice considerations are aligned with the suggestions for Microsoft SQL Server deployment, optimal backup and restore of databases using NetApp Snapshots technology. To optimize both technologies, it is vital to understand the SQL Server I/O pattern and characteristics. A well-designed storage layout for a SQL Server database supports the performance of SQL Server and the management of the SQL Server infrastructure. A good storage layout also allows the initial deployment to be successful and the environment to grow smoothly over time as the business grows.

Segregation of user database layout into different volumes, influences database performance and also backup and restore performances. Separate volumes for data and log files significantly improve the restore time as compared to a single volume hosting multiple user data files. Similarly, user databases with I/O-intensive applications might experience increased backup time.

SQL Database and log files are stored in Storage LUNs mapped to Linux virtual machines over iSCSI protocol. Design of underlying NetApp Storage logical constructs is important for performance, management, high availability, and disaster recovery of the Databases.

SQL File Groups

Each database has one primary data file, which, by default, has the .mdf extension. In addition, each database can have secondary database files. Those files, by default, have .ndf extensions.

All database files are grouped into file groups. A file group is the logical unit, which simplifies database administration. They allow the separation between logical object placement and physical database files. When you create the database objects tables, you specify in what file group they should be placed without worrying about the underlying data file configuration.

The ability to put multiple data files inside the file group allows us to spread the load across different storage devices, which helps to improve I/O performance of the system. The transaction log, in contrast, does not benefit from the multiple files because SQL Server writes to the transaction log in a sequential manner.

Specify the initial file size and auto growth parameters at the time when you create the database or add new files to an existing database. SQL Server uses a proportional fill algorithm when choosing into what data file it should write data. It writes an amount of data proportionally to the free space available in the files. The more free space in the file, the more writes it handles.

NetApp recommends that all files in the single file group have the same initial size and auto growth parameters, with grow size defined in megabytes rather than percentages. This helps the proportional fill algorithm evenly balance write activities across data files.

Storage LUN

The following are NetApp recommendations on LUN design for optimal performance, management and backup:

- Create separate LUN to store TempDB and SystemDB files. Create TempDB and SystemDB LUNs with Space reservation disabled.
- Create separate LUN to storage Database data files.
- Create a separate LUN for each instance to store Microsoft SQL server log backups. The LUNs can be part of the same volume.
- Create separate LUNs to store full text-related files and file-streaming-related files.
- Create thin provisioned (Space Reservation Disabled) LUNs for Database files and Log files.

Follow the guidelines in section Storage Volume for the placement of the LUNs for Storage Volume.

Storage Volume

The following are the NetApp recommendations on volume design for optimal performance, management and backup.

- Databases with I/O intensive queries throughout the day should be isolated in different volumes and eventually have separate jobs to back them up.
- Large databases and databases that have minimal RTO should be placed in separate volumes for faster recovery.
- Small to medium-size databases that are less critical or that have fewer I/O requirements should be consolidated into a single volume. Backing up a large number of databases residing in the same volume results in fewer Snapshot copies to be maintained. NetApp also recommends consolidating Microsoft SQL Server instances to use the same volumes to control the number of backup Snapshot copies taken.
- System databases store database server metadata, configurations, and job details; they are not updated frequently. System databases and tempdb should be placed in separate drives or LUNs. Do not place system databases in same volume as user databases. User databases have different backup policies and the frequency of user database backups is not same as for system databases.
- In the case of Microsoft SQL Server AG setup, the data and log files for replicas should be placed in an identical folder structure on all nodes.
- Use flexible volumes to store Microsoft SQL Server database files, and don't share volumes between hosts.
- Configure a volume auto size policy, when appropriate, to help prevent out-of-space conditions.

- When the SQL Server database I/O profile consists mostly of large sequential reads, such as with decision support system workloads, enable read reallocation on the volume. Read reallocation optimizes the blocks for better performance.
- Set the Snapshot copy reserve value in the volume to zero for ease of monitoring from an operational perspective.
- Disable storage Snapshot copy schedules and retention policies. Instead, use the SnapCenter for SQL Server plug-in to coordinate Snapshot copies of the Microsoft SQL Server data volumes.
- Place user data files (`.mdf`) on separate volumes because they are random read/write workloads. It is common to create transaction log backups more frequently than database backups. For this reason, place transaction log files (`.ldf`) on a separate volume or VMDK from the data files so that independent backup schedules can be created for each. This separation also isolates the sequential write I/O of the log files from the random read/write I/O of data files and significantly improves Microsoft SQL Server performance.
- Microsoft SQL Server uses the system database `tempdb` as a temporary workspace, especially for I/O intensive database consistency checker (DBCC) `CHECKDB` operations. In large environments where volume count is a challenge, you can consolidate `tempdb` into the same volume as other system databases. This procedure requires careful planning. Data protection for `tempdb` is not a high priority because this database is re-created every time the SQL Server is restarted.

Aggregate Layout

Aggregates are the primary storage containers for NetApp storage configurations and contain one or more RAID groups consisting of both data disks and parity disks.

NetApp has performed various I/O workload characterization tests using shared and dedicated aggregates with data files and transaction log files separated. The tests show that one large aggregate with more RAID groups and drive (HDD or Solid State) optimizes and improves storage performance and is easier for administrators to manage for two reasons:

- One large aggregate makes the I/O abilities of all drives available to all files.
- One large aggregate enables the most efficient use of disk space.

For high availability, place the SQL Server AlwaysOn availability group secondary synchronous replica on a separate storage virtual machine (SVM) in the aggregate. For disaster recovery purposes, place the asynchronous replica on an aggregate that is part of a separate storage cluster in the DR site, with content replicated using NetApp SnapMirror® technology.

NetApp recommends having at least 10% free space available in an aggregate for optimal storage performance.

The storage aggregate layout for the A300 with two shelves of 24 960GB SSDs, is as follows:

- Keep 2 spare drives
- Use Advanced drive partitioning to create three partition on each drive, root, data and data.
- 20 data partitions + 2 parity partitions for each aggregate
- RAID DP total aggregate size = 7.2TB

- In two aggregates, total usable space of 14.4TB

The layout changes with the IO characteristics and size of the databases.

Backup SQL Databases

Currently, NetApp SnapCenter does not support backups of Databases under SQL Server instance running on Linux. The proper combination of scripting to quiesce database and then taking NetApp Snapshot, is required.

Best Practices for VMware with NetApp Storage

vSphere Storage Considerations

In this solution architecture, iSCSI VMkernel ports are used to mount the NetApp volumes as iSCSI VMFS datastores. These datastores are used to store virtual machine configuration files and virtual machine disks. SQL Server Databases are stored in File System on the LUNs accessible thru Linux guest OS iSCSI initiator connecting the virtual machine directly to the storage system. This reduces the storage processing CPU load on the ESXi hosts and enables more granular management of those specific datasets for backup and recovery, DR replication, and secondary processing tasks.

Table 2 Storage Configuration in VMware ESXi 6.7 Environment

Data Objects	Storage Protocol	Storage Container
ESXi Hypervisor Boot volumes	iSCSI	OS Boot from Storage LUN mapped to iSCSI initiator in ESXi
Virtual machines config and Virtual Disks	iSCSI	Virtual machine's Config and VMDK files stored on storage LUN mapped to all ESXi hosts and configured as shared VMFS Datastore
SQL DATA files	iSCSI	Storage LUNs directly mapped to iSCSI initiator in RHEL virtual machines for storing Database data files.
SQL LOG files	iSCSI	Storage LUNs directly mapped to iSCSI initiator in RHEL virtual machines for storing Database log files.

- For iSCSI networks, use multiple VMkernel network interfaces on different network subnets that use NIC teaming when multiple virtual switches are used. Or use multiple physical NICs connected to multiple physical switches to provide high availability and increased throughput. In ONTAP, configure either a single-mode interface group for failover with two or more links that are connected to two or more switches or use LACP or other link-aggregation technology with multimode interface groups to provide high availability and the benefits of link aggregation.

- vSphere includes built-in support for multiple paths to storage devices, referred to as native multipathing (NMP). NMP includes the ability to detect the type of storage for supported storage systems and automatically configure the NMP stack to support the capabilities of the storage system in use. Both NMP and NetApp ONTAP support Asymmetric Logical Unit Access (ALUA) to negotiate optimized and non-optimized paths. In ONTAP, an ALUA-optimized path follows a direct data path, using a target port on the node that hosts the LUN being accessed. ALUA is turned on by default in both vSphere and ONTAP. The

NMP recognizes the ONTAP cluster as ALUA, and it uses the ALUA storage array type plug-in (VMW_SATP_ALUA) and selects the round robin path selection plug-in (VMW_PSP_RR).

Space Reclamation

Space may be reclaimed for other use when virtual machines are deleted within a datastore. For LUN-based VMFS datastores, ESXi can issue VAAI UNMAP primitives to the storage (again, when using thin provisioning) to reclaim space. NetApp Storage supports UNMAP commands.

vSphere 6.5 onwards, when using VMFS 6, space should be automatically reclaimed asynchronously, but may also be run manually, using the `esxcli storage vmfs unmap` command, if needed.

Virtual Machine and Datastore Cloning

Cloning a storage object allows you to quickly create copies for further use, such as provisioning additional virtual machines, backup/recovery operations, and so on. In vSphere, you may clone a virtual machine, virtual disk, VVol, or datastore. After being cloned, the object may be further customized, often through an automated process.

Cloning may be offloaded to systems running ONTAP software through several mechanisms, typically at the virtual machine, VVol, or datastore level. These include:

- VVols using the NetApp vSphere APIs for Storage Awareness (VASA) Provider. ONTAP clones are used to support VVol Snapshot copies managed by vCenter that are space-efficient with minimal I/O impact to create and delete them. Virtual machines may also be cloned using vCenter, and these are also offloaded to ONTAP, whether within a single datastore/volume or between datastores/volumes.
- vSphere cloning and migration using vSphere APIs – Array Integration (VAAI). Virtual machine cloning operations can be offloaded to ONTAP in both SAN and NAS environments (NetApp supplies an ESXi plug-in to enable VAAI for NFS). Storage vMotion operations are also offloaded to ONTAP for SAN, but this offload is not supported by VMware for NAS. ONTAP uses the most efficient approach based on source, destination, and installed product licenses. This capability is also used by VMware Horizon View.
- Storage Replication Adapter (used with VMware Site Recovery Manager). For this solution, clones are used to test recovery of the DR replica non-disruptively.
- Backup and recovery using NetApp tools such as SnapCenter. Virtual machine clones are used to verify backup operations, as well as to mount a virtual machine backup so that individual files may be copied.
- ONTAP supports industry standard Copy offload functionality, so offloaded cloning can be invoked by VMware, NetApp, and third-party tools. Clones that are offloaded to ONTAP have several advantages. They are space-efficient in most cases, needing storage only for changes to the object; there is no additional performance impact to read and write them, and in some cases, performance is improved by sharing blocks in high-speed caches; and they offload CPU cycles and network I/O from the ESXi server.

Recommended ESXi Host and Other ONTAP Settings

NetApp has developed a set of ESXi host multipathing and HBA timeout settings for proper behavior with ONTAP based on NetApp testing. These are easily set using the Virtual Storage Console. (From the VSC Summary dashboard, click Edit Settings in the Host Systems portlet or right-click the host in vCenter, then VSC > Set Recommended Values.) Table 3 lists the recommended host settings.

Table 3 ONTAP Settings for ESXi Hosts

Host Setting	NetApp Recommended Value
Path selection policy	Set to `RR` (round robin) for all iSCSI paths. Setting this value to `RR` helps provide load balancing across all active/optimized paths.
Disk.QFullSampleSize	Set to `32` for all configurations. Setting this value helps prevent I/O errors.
Disk.QFullThreshold	Set to `8` for all configurations. Setting this value helps prevent I/O errors.

Provisioning by Virtual Storage Console

Use VSC to create and manage LUNs and igroups for VMware ESXi. VSC automatically determines WWPNs of servers and creates appropriate igroups. VSC also configures LUNs according to best practices and maps them to the correct igroups.

The Virtual Storage Console also specifies certain default settings when creating ONTAP FlexVol volumes:

- Snapshot reserve (-percent-snapshot-space) 0
- Fractional reserve (-fractional-reserve) 0
- Access time update (-atime-update) False
- Minimum readahead (-min-readahead) False
- Scheduled Snapshot copies None
- Storage efficiency Enabled

Best Practices for Hyper-V with NetApp Storage

Hyper-v Storage Considerations

In this solution architecture, iSCSI initiator in Hyper-V is used to mount the NetApp LUNs as Clustered Shared Volume. These Clustered Shared Volumes are used to store virtual machine configuration files and virtual machine disks. SQL Server Databases are stored in File System on the LUNs accessible thru Linux guest OS iSCSI initiator connecting the virtual machine directly to the storage system. This reduces the storage processing CPU load on the Hyper-V hosts and enables more granular management of those specific datasets for backup and recovery, DR replication, and secondary processing tasks.

Table 4 Storage Configuration in Windows Server 2016 Hyper-V Environment

Data Objects	Storage Protocol	Storage Container
Hyper-V Boot volumes	iSCSI	OS Boot from Storage LUN mapped to iSCSI initiator in Hyper-V
Virtual machines config and Virtual Disks	iSCSI	Virtual machine's Config and VHD files stored on storage LUN mapped to all Hyper-V hosts and configured as Clustered Shared Volume

Data Objects	Storage Protocol	Storage Container
SQL DATA files	iSCSI	Storage LUNs directly mapped to iSCSI initiator in RHEL virtual machines for storing Database data files.
SQL LOG files	iSCSI	Storage LUNs directly mapped to iSCSI initiator in RHEL virtual machines for storing Database log files.

NetApp Host Utilities Kit

Host Utilities are a set of software programs and documentation that enable you to connect host computers to virtual disks (LUNs) on NetApp storage systems. Installation of the Host Utilities Kit sets timeout and other operating system-specific values to their recommended defaults and includes utilities for examining LUNs provided by NetApp storage. See the NetApp Interoperability Matrix Tool for complete details for a given tested and supported NetApp configuration.

For the FlexPod solution with Hyper-V, the latest version of the Windows Host Utilities kit for Hyper-V is installed.

iSCSI Initiators

Use two iSCSI initiators in the Hyper-V Host support path to LUNs through each iSCSI Fabric in Cisco UCS. Create sessions from each initiator to the target portal on SVM.

Host Multipathing

Host Utilities can be used to support multiple paths, and you can use a combination of protocols between the host and storage controllers. Configuring multiple paths can provide a highly available connection between the Windows host and the storage system.

Multipath I/O (MPIO) software is required any time a Windows host has more than one path to the storage system. The MPIO software presents a single disk to the operating system for all paths, and a device-specific module (DSM) manages path failover. Without MPIO software, the operating system might see each path as a separate disk, which can lead to data corruption.

There is a native DSM provided with Windows Server 2016. It offers active/active and active/passive load balance policies for both the FC and iSCSI protocols. ALUA is enabled on the storage system by default.

Use two iSCSI initiator portals on each Windows Host to create iSCSI sessions to each iSCSI LIF in SVM.

Provisioning by SnapCenter

Use SnapCenter and SnapCenter Plug-in for Windows to create and manage LUNs and igroups for Hyper-V. SnapCenter automatically determines WWPNs of servers and creates appropriate igroups on SVM. SnapCenter also configures LUNs according to best practices and maps them to the correct igroups.

NetApp SMI-S Agent

The NetApp ONTAP SMI-S Agent allows administrators to manage and monitor NetApp FAS storage systems through open-standard protocols and classes as defined by Distributed Management Task Force (DMTF) and Storage Networking Industry Association (SNIA).

The ONTAP SMI-S Agent is a command-based interface that detects and manages platforms that run ONTAP. The SMI-S Agent uses web-based enterprise management protocols, which allow you to manage, monitor, and report on storage elements. SMI-S integration is designed to perform the following tasks:

- End-to-end discovery of logical and physical objects and the associations between them
- The addition of capacity to hosts and clusters
- The rapid provisioning of virtual machines by using the SAN and the SMB 3.0 protocol

The SMI-S Agent interface can also be used to accomplish simple tasks. Administrators can use Microsoft SCVMM to create and deploy new storage to individual hosts or clusters.

NetApp SnapManager for Hyper-V

NetApp SnapManager for Hyper-V provides a solution for data protection and recovery for Microsoft Hyper-V virtual machines running on the NetApp Data ONTAP operating system. You can perform application-consistent and crash-consistent dataset backups according to protection policies set by your backup administrator. You can also restore virtual machines from these backups. Reporting features enable you to monitor the status of and get detailed information about your backup and restore jobs.

Key benefits include the following:

- Simplified, automated backup, restores, and replication of virtual machines
- Increased operational efficiency with a built-in backup scheduler and policy-based retention
- Simplified user experience for users familiar with Microsoft Management Console interface
- Support across Fibre Channel, iSCSI, and SMB 3.0 protocols
- Best Practices for Hyper-V with NetApp Storage

Best Practices for Red Hat Linux with NetApp Storage

Install NetApp Linux Unified Host Utilities

NetApp Linux Unified Host Utilities software package include sanlun utility and documentation. The sanlun utility helps manage LUNs and HBAs. Host Utilities also configures proper timeouts for LUN access.

iSCSI Initiators, Sessions, and LUN Mapping

Configure at least two network adapters for iSCSI LUN access, one through Fabric Interconnect A, and another through Fabric Interconnect B.

Install iSCSI software initiator, discover iSCSI targets from NetApp Storage virtual machine, and then create one session from each initiator to each target portal in an SVM.

Create initiator group on NetApp Storage SVM for the initiator IQN, and map LUN to the initiator group.

Enable and Configure Multipathing in RHEL Virtual Machine

Enable the multipathd service and create entries in /etc/multipath.conf file as prescribed in deployment guide. Use ALUA and "round robin" path selection policy as in sample multipath.conf file installed by NetApp Host Utilities.

For Red Hat Enterprise Linux Operating System, Device Mapper Multipath (DM Multipath) modules are needing to be installed for configuring multipathing. DM Multipath allows you to configure multiple I/O paths between server nodes and storage arrays into a single device. Multipathing aggregates the I/O paths, creating a new device that consists of the aggregated paths. For more information about installing and configuring multipath device mappers in RHEL, see: https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/7/html/dm_multipath/mpio_setup

Figure 8 shows the sample multipath.conf file of RHEL virtual machine running SQL Server in it. As shown, three different LUNs (for sql data, log and backups) created and accessed by the RHEL virtual machine.

Figure 8 Sample Multipath.conf File in RHEL

```
##
blacklist {
        wwid 26353900f02796769
  devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
  devnode "^hd[a-z]"
}
multipaths {
  multipath {
    wwid      3600a098038303862535d4d69772f7853
    alias     sqllog
    path_grouping_policy  multibus
    path_selector   "round-robin 0"
    failback    manual
    rr_weight   priorities
    no_path_retry   5
          prio      alua
        }
    multipath {
    wwid      3600a098038303862535d4d69772f784f
    alias     sqldata
    path_grouping_policy  multibus
    path_selector   "round-robin 0"
    failback    manual
    rr_weight   priorities
    no_path_retry   5
          prio      alua
       }
        multipath {
    wwid      3600a0980383038625a244d6b79323762
    alias     sqlbackup
    path_grouping_policy  multibus
    path_selector   "round-robin 0"
    failback    manual
    rr_weight   priorities
    no_path_retry   5
     prio       alua
       }
}
```

NetApp SnapCenter

The SnapCenter platform is based on a multitier architecture that includes a centralized management server (SnapCenter Server) and a SnapCenter host agent.

SnapCenter Server Requirements

Table 5 lists the minimum requirements for installing SnapCenter Server and plug-in on a Windows server. For the latest version compatibility and other plug-in information, refer to the NetApp Interoperability Matrix Tool.

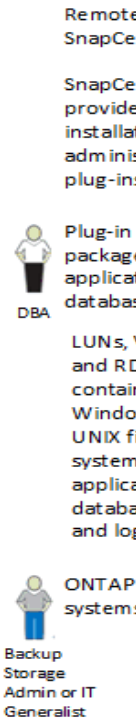
For detailed information, refer to the SnapCenter 4.1.1 Installation and Setup Guide.

Table 5 SnapCenter Server Requirements

Component	Requirements
Minimum CPU count	4 cores/vCPUs
Memory	Minimum: 8GB Recommended: 32GB
Storage space	Minimum space for installation: 10GB Minimum space for repository: 20GB
Supported operating systems	Windows Server 2012 Windows Server 2012 R2 Windows Server 2016
Software packages	.NET 4.5.2 or later Windows Management Framework 4.0 or later PowerShell 4.0 or later Java 1.8 (64-bit)
Active Directory domain membership	Windows Host must be joined to AD domain.
Database for SnapCenter repository	MySQL Server 5.7.22 (installed as part of the SnapCenter installation)
Port	Requirement
443	vCenter Server to SnapCenter Server API Access over HTTPS
8144	SnapCenter GUI to SnapCenter Plug-in for VMware
8145	SnapCenter Server Core API
8146	For SnapCenter server REST API
3306	MySQL
443	vCenter Server to SnapCenter Server API Access over HTTPS

Table 6
ort
Re
quir
em
ent
s

Figure 9
etA
pp
Sna
psh
ot
De
plo
ym
ent



Host and Privilege Requirements for the SnapCenter Plug-in for VMware vSphere

Review the following requirements before you install the SnapCenter Plug-in for VMware vSphere:

- You must have SnapCenter admin privileges to install and manage the SnapCenter GUI.

- You must install the Plug-in for VMware vSphere on a Windows host (virtual host or physical host). The Plug-in for VMware vSphere must be installed on a Windows host regardless of whether you use the plug-in to protect data on Windows systems or Linux systems.
- When installing a plug-in on a Windows host, if you specify a Run As account that is not built-in or if the Run As user belongs to a local workgroup user, you must disable UAC on the host.
- Do not install the Plug-in for VMware vSphere on the vCenter Server appliance.
- You must not install the Plug-in for VMware vSphere on the vCenter Server appliance, which is a Linux host. You can only install the Plug-in for VMware vSphere on Windows hosts.
- You must not install other plug-ins on the host on which the Plug-in for VMware vSphere is installed.
- You must install and register a separate, unique instance of the Plug-in for VMware vSphere for each vCenter Server.
  - Each vCenter Server, whether or not it is in Linked mode, must be paired with a separate instance of the Plug-in for VMware vSphere.
  - Each instance of the Plug-in for VMware vSphere must be installed on a separate Windows host. One instance can be installed on the SnapCenter Server host.
  - vCenters in Linked mode must all be paired with the same SnapCenter Server.

For example, if you want to perform backups from six different instances of the vCenter Server, then you must install the Plug-in for VMware vSphere on six hosts (one host can be the SnapCenter Server host) and each vCenter Server must paired with a unique instance of the Plugin for VMware vSphere.

Cisco UCS Design Choices and Best Practices

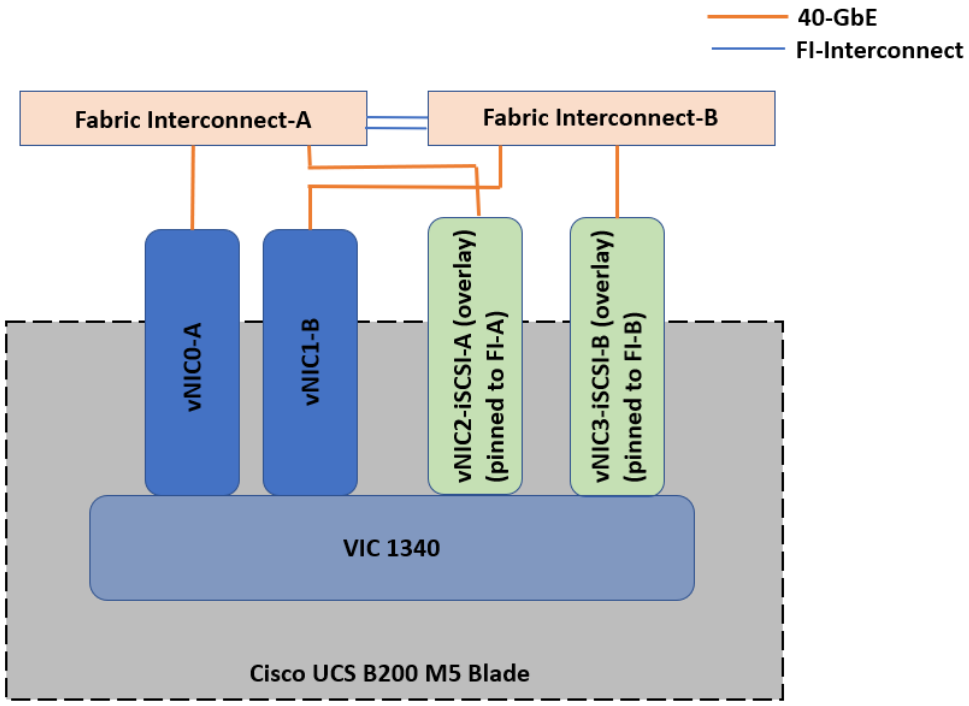
Cisco UCS Virtual Network Interface Cards

This FlexPod architecture uses two Cisco UCS Virtual Network Interface Cards (vNICs). One vNIC pinned to Fabric - A and the other pinned to Fabric - B to recover in case of a path failure. These two vNICs provisioned to the servers are teamed in the Hypervisors for failover and load balancing using the teaming technologies such as vSphere Distributed Switches (vDS) in VMware and Switch Embedded Teaming (SET) in Windows Server 2016. Note that, Cisco UCS fabric failover is not enabled on the vNICs.

In addition to these two vNICs, this FlexPod architecture requires two more additional overlay vNICs for iSCSI boot which the hypervisors use to boot from the NetApp storage. Each iSCSI vNIC is pinned to a separate fabric path in Cisco UCS to recover from path failures. These iSCSI vNICs are not teamed in neither for VMware nor for Hyper-V environments. In case of VMware environment, these iSCSI vNICs connected to Standard Switches (not migrated to vDS) and in case of Hyper-V environment, they are seen as iSCSI interfaces by the parent Operating System (Windows Server 2016). On both the environments, multipathing feature will takes care of failover and load balancing of iSCSI traffic.

Cisco UCS fabric failover is not enabled on these vNICs.

Figure 10 Cisco UCS vNICs for Hypervisors

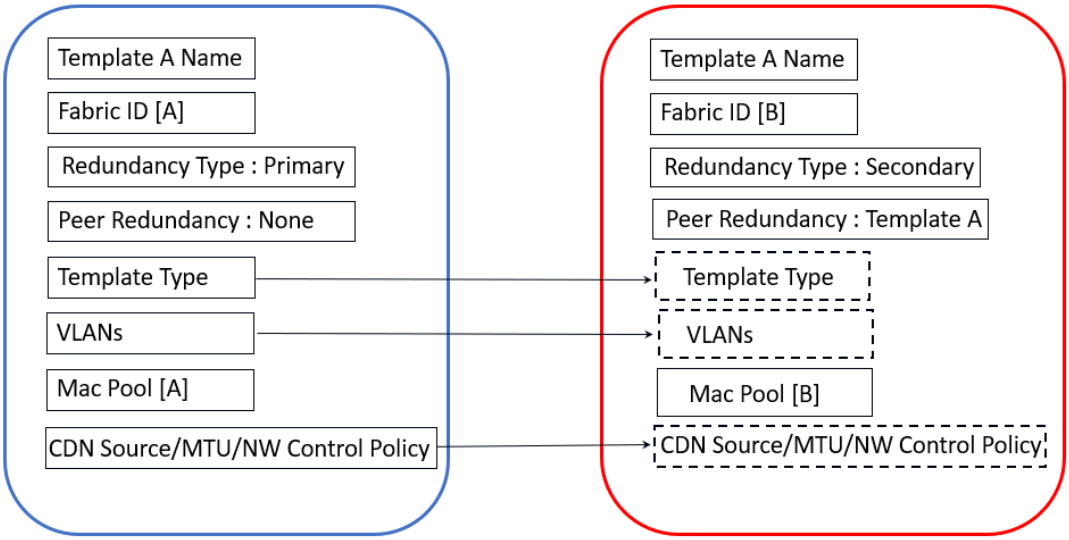


Cisco UCS vNIC Template Redundancy

An additional Cisco UCS vNIC feature included in this design is vNIC Template Redundancy. Optional configuration of vNIC Template Redundancy reduces configuration steps for the Secondary Template within a pair of configured

vNIC Templates and allows for better consistency between the two as future changes to VLANs and specific settings made upon the Primary Template are automatically propagated to the Secondary Template as shown in Figure 11.

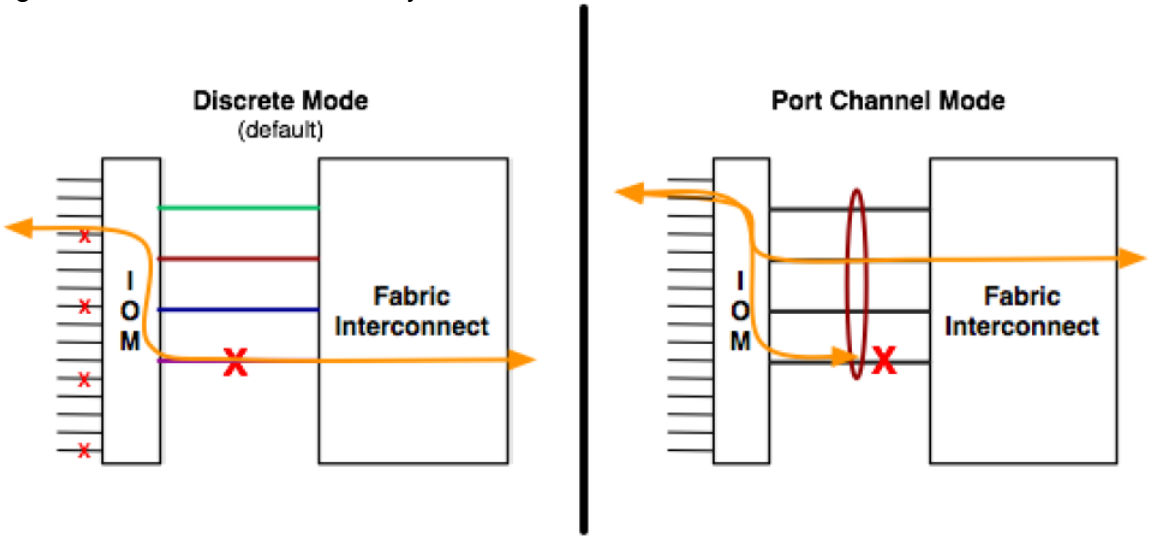
Figure 11 Cisco UCS vNIC Template Redundancy Between Two vNIC Templates



Cisco Unified Computing System Chassis/FEX Discovery Policy

Cisco UCS can be configured to discover a chassis using Discrete Mode (Link Grouping Preference of None) or the Port-Channel mode (Figure 12). In Discrete Mode each FEX KR connection and therefore server connection is tied or pinned to a network fabric connection homed to a port on the Fabric Interconnect. In the presence of a failure on the external " link" , all the KR connections are disabled within the FEX I/O module. In Port-Channel mode, the failure of a network fabric link allows for redistribution of flows across the remaining port channel members. Port-Channel mode therefore is less disruptive to the fabric and hence recommended in the FlexPod designs.

Figure 12 Chassis Discover Policy – Discrete Mode vs. Port Channel Mode



Cisco Unified Computing System – QoS and Jumbo Frames

FlexPod accommodates a myriad of traffic types (vMotion, NFS, FCoE, control traffic, and so on.) and is capable of absorbing traffic spikes and protect against traffic loss. Cisco UCS and Nexus QoS system classes and policies deliver this functionality. In this validation effort the FlexPod was configured to support jumbo frames with an MTU size of 9000. Enabling jumbo frames allows the FlexPod environment to optimize throughput between devices while simultaneously reducing the consumption of CPU resources. In this FlexPod system, the default Best Effort QoS is used for validating and testing the SQL virtual machines and Always On availability failover capabilities.

When setting up Jumbo frames, it is important to make sure MTU settings are applied uniformly across the stack to prevent packet drops and negative performance.

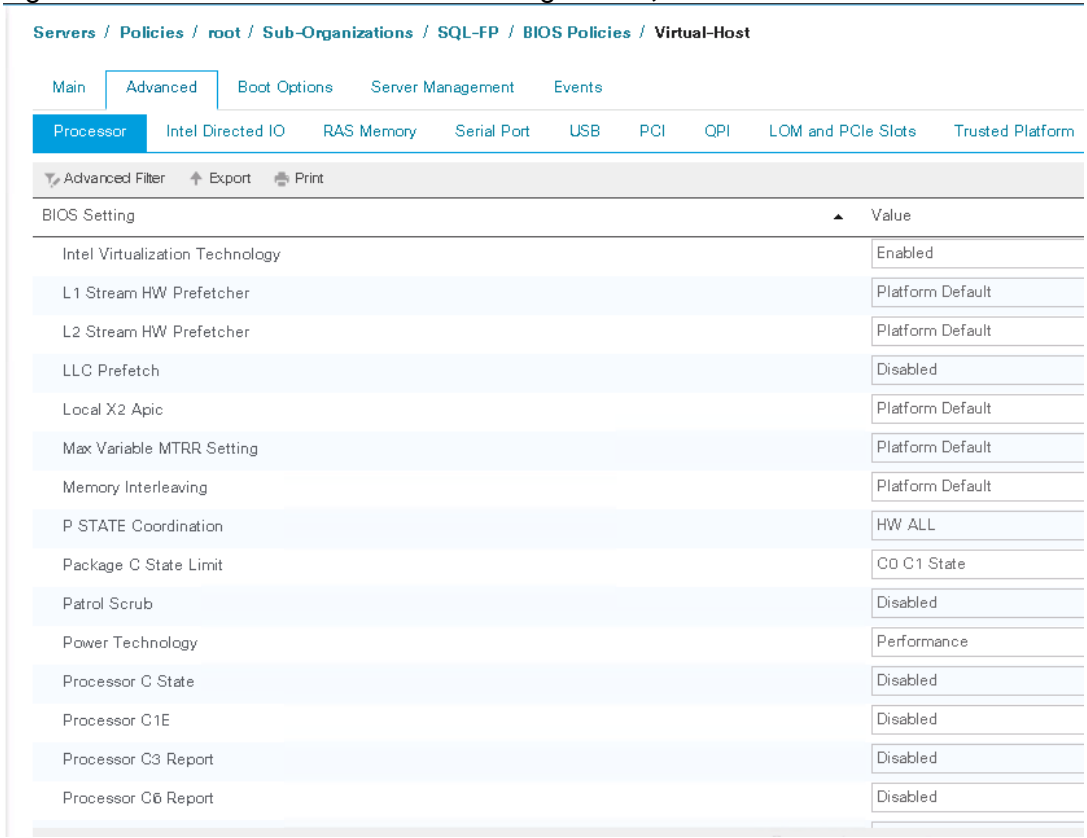
Cisco Unified Computing System – Adapter Policy

The Transmit Queues, Receive Queues defined in the default Adapter policy may eventually get exhausted as more SQL Server databases are consolidated on the FlexPod System. It is recommended to use higher queues on the vNICs that are used for Guest iSCSI storage traffic for better storage throughput. Create a new adapter policy with higher transmit and receive queues and apply it on the vNICs that are used for iSCSI guest storage traffic. These queue values are to be arrived by testing for your application performance requirements.

Cisco Unified Computing System – BIOS Policy

The underlying server's BIOS settings plays vital role in achieving optimal performance from the workloads. Tuning the BIOS settings for a system will vary from workload to workload. Therefore, make sure to use appropriate bios settings to obtain optimal performance from workloads hosted in the virtual machines. Figure 13 shows the critical bios settings used for latency intensive and critical OLTP database workloads on Cisco B200 M5 (Intel Xeon Scalable CPU – Skylake family) blade server.

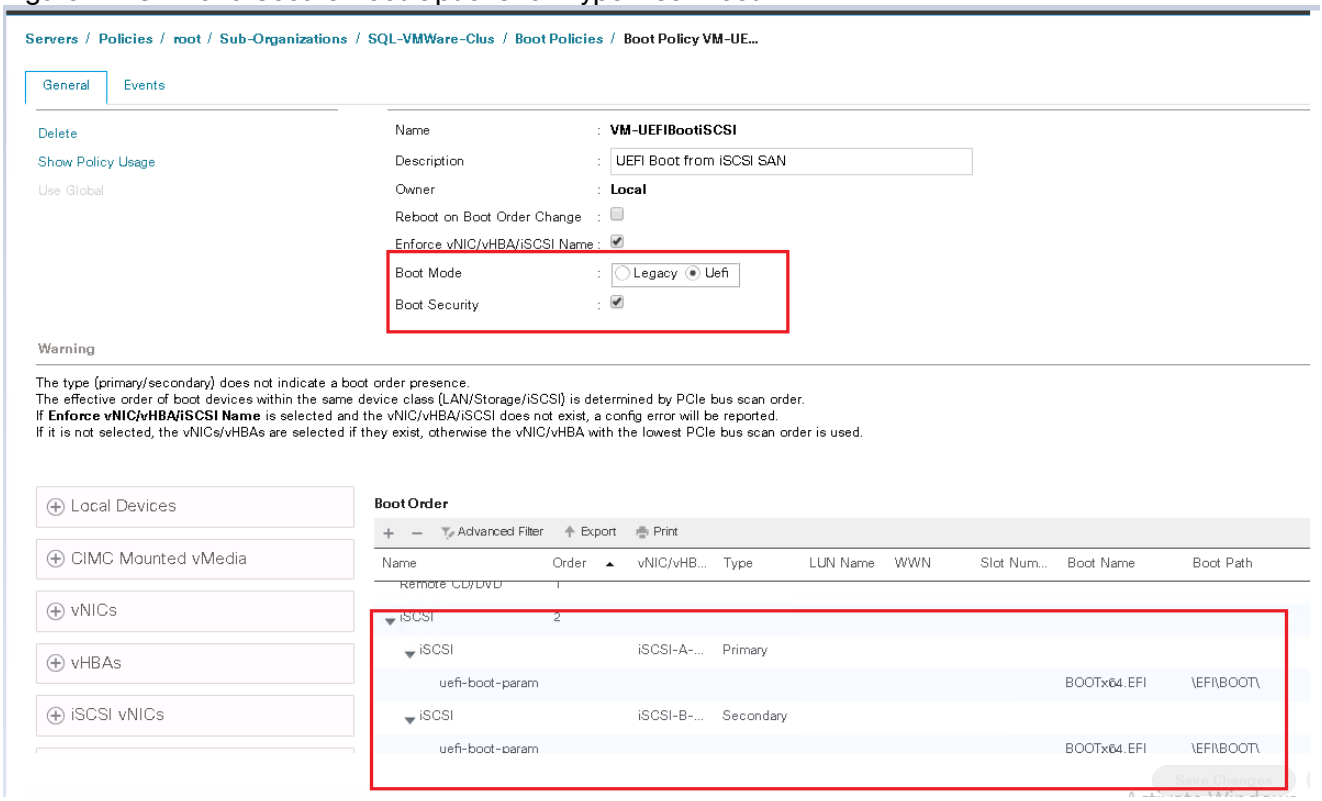
Figure 13 Cisco UCS B200 M5 BIOS Settings for SQL Server Workloads



Cisco Unified Computing System – UEFI Boot Policy

Unified Extensible Firmware Interface (UEFI) is a modern specification for a software program that connects a server's firmware to its operating system/Hypervisor. VMware vSphere 6.7 and Windows Server 2016 supports the UEFI boot mode and secure boot options. This FlexPod solution uses UEFI boot and secure boot options for the hypervisors as well as the Linux virtual machines. Figure 14 shows the UCSM boot policy that uses UEFI boot and secure boot options for booting VMware ESXi 6.7 hypervisor.

Figure 14 UEFI and Secure Boot Options for Hypervisor Boot



For Windows Deployment, set the boot path as “/EFI/BOOT/”.

Cisco UCS Physical Connectivity

In this FlexPod architecture, each Cisco UCS Fabric Interconnect is configured with a port-channels to Cisco Nexus 9000 series switches. This port channel will carry all types of traffic originated from the workloads deployed in the FlexPod system. The validated design utilized two uplinks from each FI to the Nexus switches to create the port-channels, for an aggregate bandwidth of 160GbE (4 x 40GbE) with the 6332-16UP. The number of links can be easily increased based on customer data throughput requirements.

VMware vSphere 6.7 Design and Best Practices

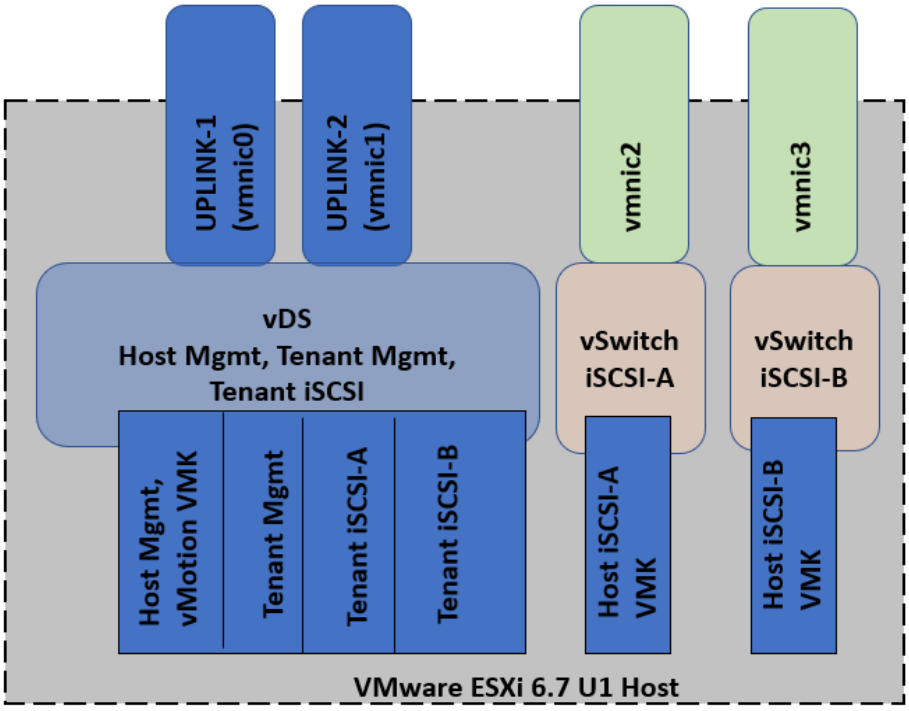
VMware vSphere is a virtualization platform for holistically managing large collections of infrastructure (resources-CPUs, storage and networking) as a seamless, versatile, and dynamic operating environment. Unlike traditional operating systems that manage an individual machine, VMware vSphere aggregates the infrastructure of an entire data center to create a single powerhouse with resources that can be allocated quickly and dynamically to any application in need. For more information on VMware vSphere and its components, refer to: http://www.vmware.com/products/vsphere.html

For additional information about vSphere 6.7, refer to: https://blogs.vmware.com/vsphere/files/2018/09/Whats-New-vSphere-6.7.pdf

Logical Network Diagram

This FlexPod solution uses a VMware Virtual Distributed Switch(vDS) for primary virtual switching. The vNIC0-A and vNIC1-B will be configured as UPLINK-1 and UPLINK-2 for the vDS. The required Distributed Port Groups and VMkernel (vmk) network interfaces for different traffics such as Host/Guest management, vMotion, iSCSI (for Guests only) will be created as required. The solution also uses two additional overlay vNICs for booting the VMware host using iSCSI protocol and these two vNICs will not be teamed; they will be part of two different Standard Switches which are pinned to each Fabric Interconnects. These interfaces will be tagged with two different VLANs as "native" for the storage connectivity. For example, vmnic2 will be tagged with VLAN ID 3012 as native VLAN and connects to NetApp storage Controllers through Fabric Interconnect A. Similarly, vmnic3 will be tagged with VLAN ID 3022 as native VLAN and connects to NetApp storage Controllers through Fabric Interconnect B.

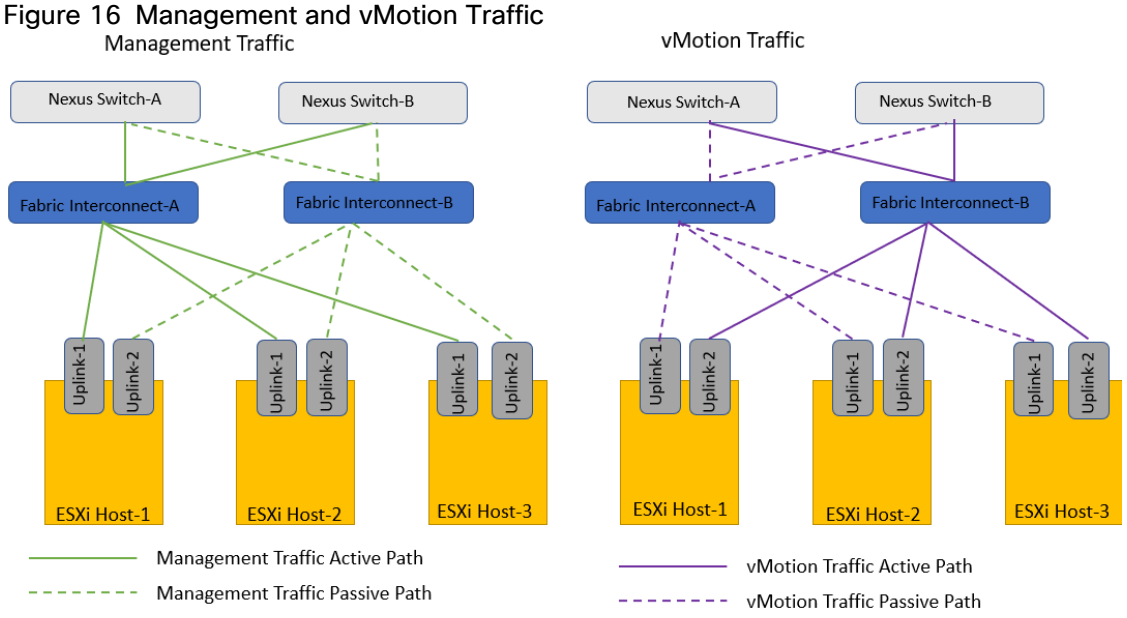
Figure 15 VMware vDS and Standard Switches in ESXi Host



Cisco UCS Infrastructure Traffic with vSphere Hosts in this Design

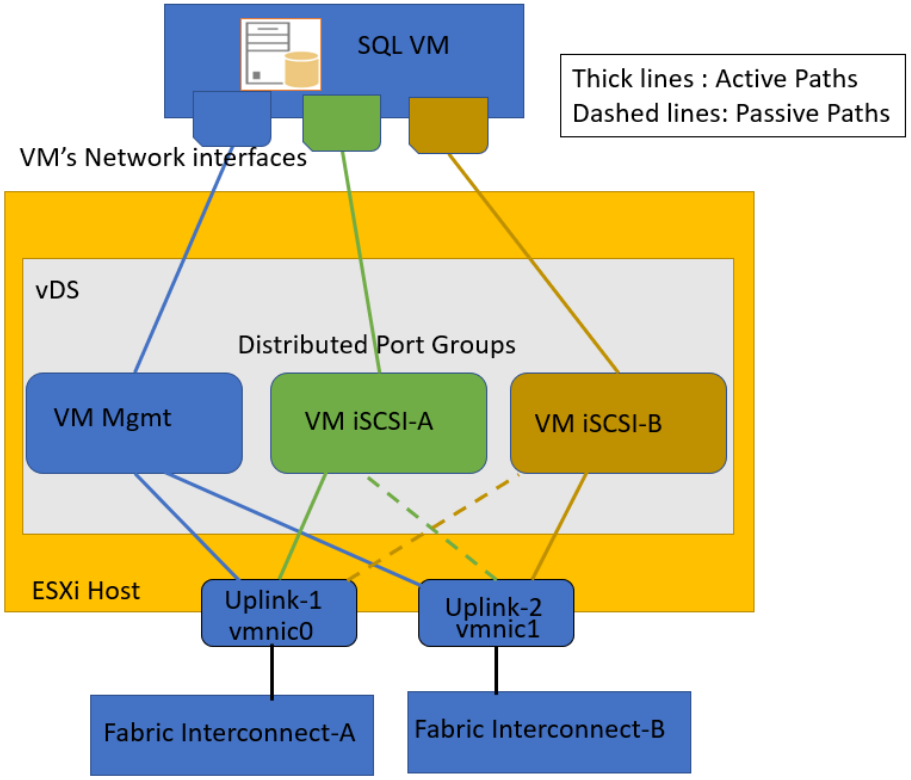
This FlexPod design incorporates active, standby, and unused uplinks to create certain traffic patterns for certain port groups in the vDS that provide better efficiency through containment to a certain fabric, or required by functionality. The Host management and vMotion VMkernels are setup with active/standby teaming within their port groups. For vMotion traffic, the port group is configured with vmnic1 (uplink-2, B side of the fabric) as active, and vmnic0 (uplink-1, A side of the fabric) as standby. For host management traffic, the port group has been setup in the alternate layout with vmnic0 (uplink-1) set as active and vmnic1 (uplink-2) set as standby. This allows traffic from vMotion and management to be contained within one side of the fabric interconnect and to prevent it from hair-pinning up through the upstream Nexus switch if the differing hosts were allowed to randomly choose between the two fabrics.

Figure 16 Management and vMotion Traffic



Similarly, for Tenant iSCSI-A traffic, the port group it is configured with vmnic0 (uplink-1, A side of the fabric) as active, and vmnic1 (uplink-2, B side of the fabric) as standby. For Tenant iSCSI-B traffic, the port group is configured with vmnic1 (uplink-2, B side of the fabric) as active, and vmnic1 (uplink-1, A side of the fabric) as standby. Figure 17 illustrates the how SQL virtual machine's network interfaces cards are configured. As shown, the virtual machine management traffic will be active on both the uplinks. For storage access, two different port groups have been configured. Each port group will be active on one uplink and passive on other uplink. For those SQL virtual machines that are critical and needs higher storage bandwidth, two iSCSI NICs can be configured (one from each iSCSI port group). For less critical SQL virtual machines single iSCSI NIC using any one of the iSCSI port groups can be configured.

Figure 17 SQL Virtual Machine's Management and iSCSI Storage Traffic



For Hypervisor boot, both iSCSI networks are set within dedicated vSwitches that are only connecting to the appropriate fabric side for the iSCSI traffic as shown in Figure 18. Using a vSwitch allows for easier configuration when using iSCSI SAN boot than using a vDS.

Figure 18 iSCSI Traffic for Host Boot

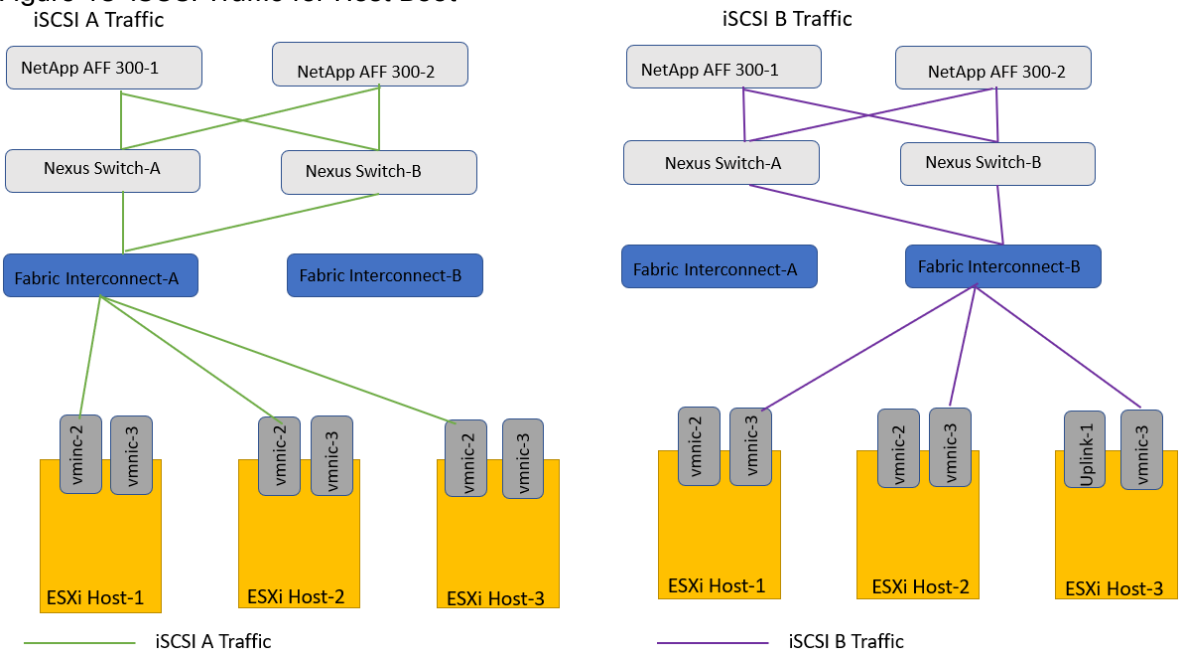
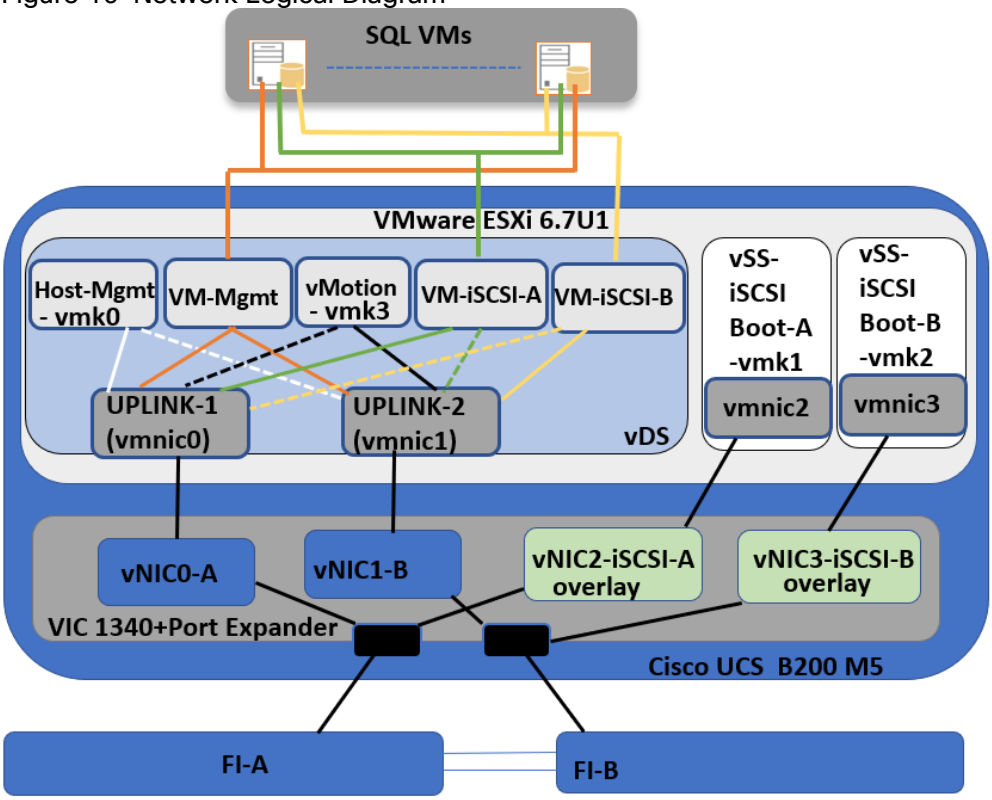


Figure 19 illustrates the complete logical networking with in the VMware vSphere ESXi host.

Figure 19 Network Logical Diagram

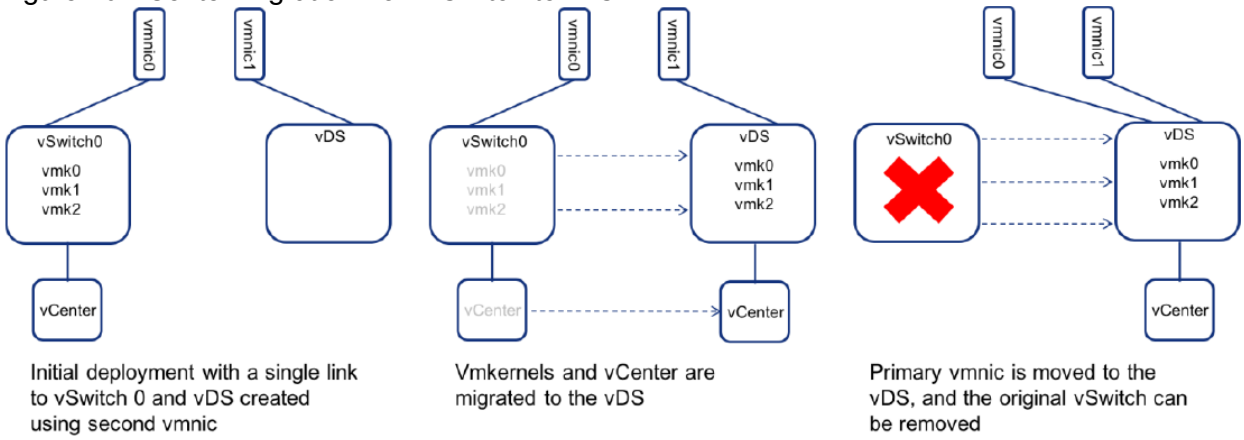


Migration of vSphere vSwitch to vDS

The VMware vDS in this release of VMware vSphere has shown reliability not present in previous vDS releases, leading this FlexPod design to bring host management VMkernels and the vCenter network interface into the vDS that is hosted from the vCenter.

There is still some care required in the step by step deployment to avoid isolating the vCenter when migrating the vCenter over from a vSwitch to a vDS hosted by the vCenter as shown in Figure 20.

Figure 20 vCenter Migration from vSwitch to vDS

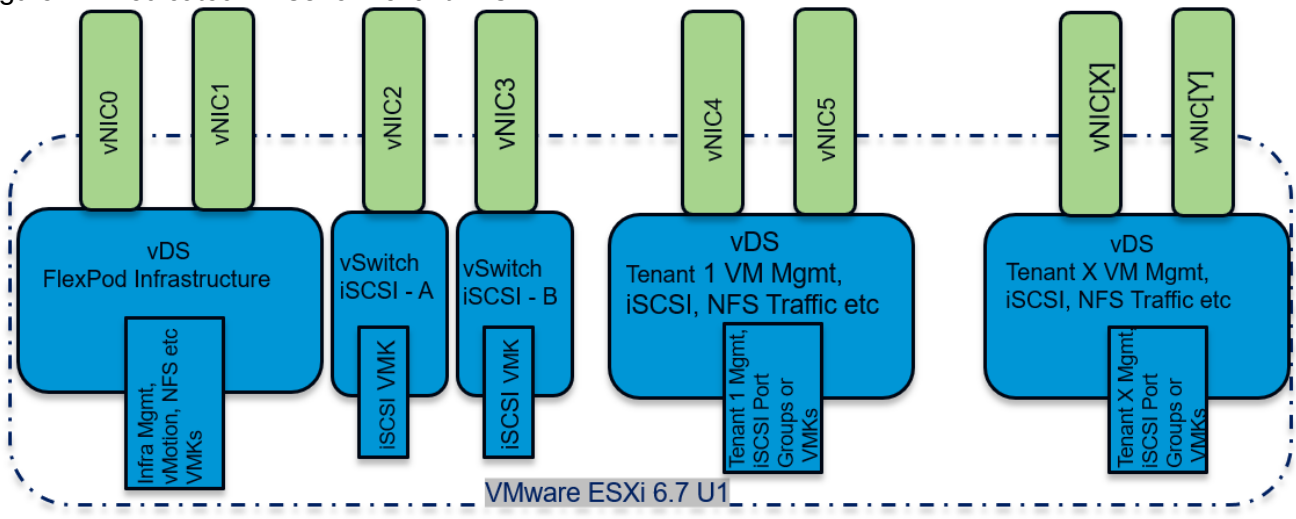


vCenter maintains connectivity during this migration as detailed in deployment guide [FlexPod Datacenter with Microsoft SQL Server 2017 on Linux VM Running on VMware and Hyper-V]. This is a requirement in FlexPod setup hosting their own vCenter, and not required when the vCenter is on a separate management cluster outside of FlexPod.

Dedicated vNICs for Tenant vDS

Tenant networks do not have a requirement to co-exist on the same vDS that we use for managing hosts and other virtual machines. They can optionally be pulled into a separate vDS. Tenant vDS can be deployed with dedicated vNIC uplinks as shown in Figure 21, allowing for RBAC of the visibility and/or adjustment of the vDS to the respective tenant manager in vCenter. Based on the purpose and workload traffic they carry, appropriate QoS and MTU settings needs to be configured on these additional vNICs.

Figure 21 Dedicated vNICs for Tenant vDS



VMware ESXi Host Power Settings

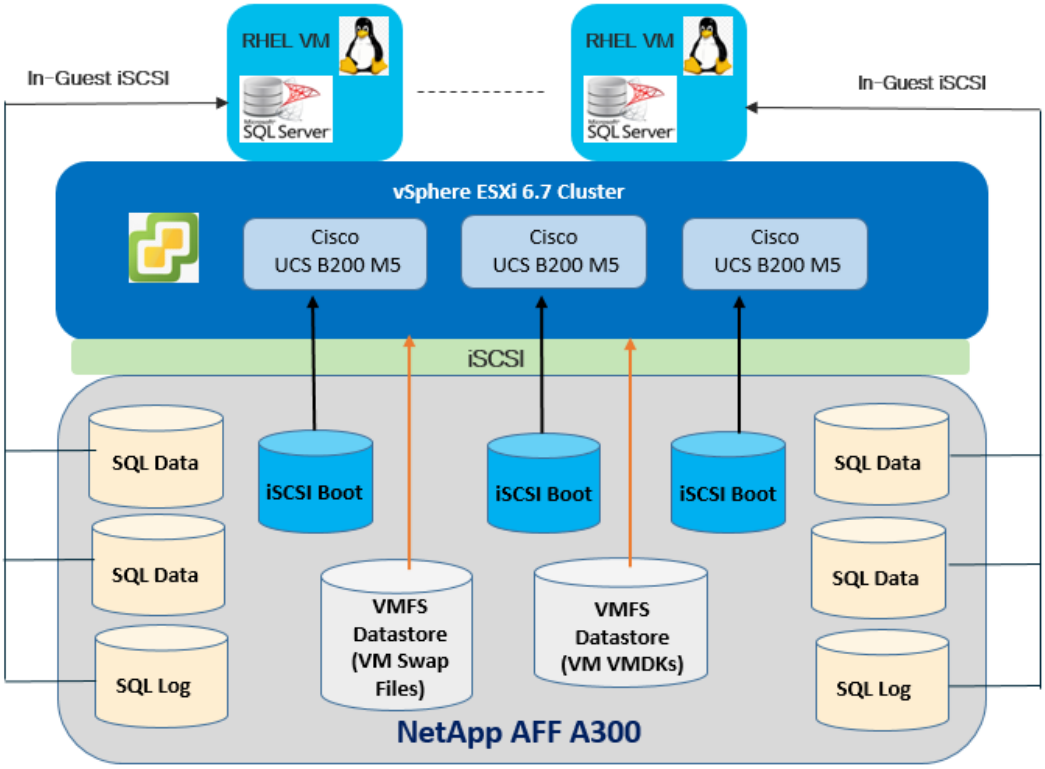
ESXi has been heavily tuned for driving high I/O throughput efficiently by utilizing fewer CPU cycles and conserving power; therefore the Power setting on the ESXi host is set to “Balanced.” However, for critical database

deployments, it is recommended to set the power setting to "High Performance." Selecting "High Performance" causes the physical cores to run at higher frequencies and thereby it will have positive impact on the database performance.

Logical Layout – Storage

Figure 22 illustrates how the NetApp storage LUNs are mapped and accessed by VMware ESXi hosts and SQL server virtual machines hosted on VMware cluster.

Figure 22 Virtual Machine's Storage Access on VMware Cluster



The VMware ESXi hypervisor is installed on the dedicated iSCSI Boot LUNs. These boot LUNs are accessed by the ESXi servers using their iSCSI VMkernel interfaces. Two additional LUNs are created and formatted with VMFS filesystem for storing virtual machine's VMDK files. The first datastore is used for storing virtual machines VMDKs where the Guest OS and SQL binaries of the virtual machines will be installed. The other datastore is used for storing virtual machine's swap files. These two vmfs datastores will also help in proper functioning of datastore heartbeat detection as vSphere high availability requires minimum of two shared datastores shared between all the ESXi hosts in a cluster. The virtual machines deployed on the vSphere clusters are provisioned with two iSCSI interfaces to directly connect to the NetApp storage array. SQL Virtual Machine directly access the NetApp storage LUNs using the in-guest iSCSI initiator which bypass the ESXi kernel TPC/IP stack which improves IO performance and provides insights in performance monitoring.

Windows Server 2016 Hyper-V Design and Best Practices

Logical Layout – Network

Switch Embedded Teaming Overview

Switch Embedded Teaming (SET) is a Microsoft's new feature introduced in Windows Server 2016. It is an alternative NIC teaming solution that can be used in environments that include Hyper-V and Software Defined Networking (SDN) stack in Windows Server 2016. It integrates some NIC teaming functionality into the Hyper-V switch. It only operates in Switch Independent Mode and all the members in the team are active and none in standby mode.

The physical NICs requirements for SET are as follows:

- Any Ethernet NIC that have passed the Windows Hardware Qualification and Logo (WHQL) test can be used to group in SET.
- All NICs in SET team must be identical (that is, same manufacture, same model, same firmware and driver).
- Supports between one and eight physical NICs in a SET team. The NICs can be on same or different physical switches.

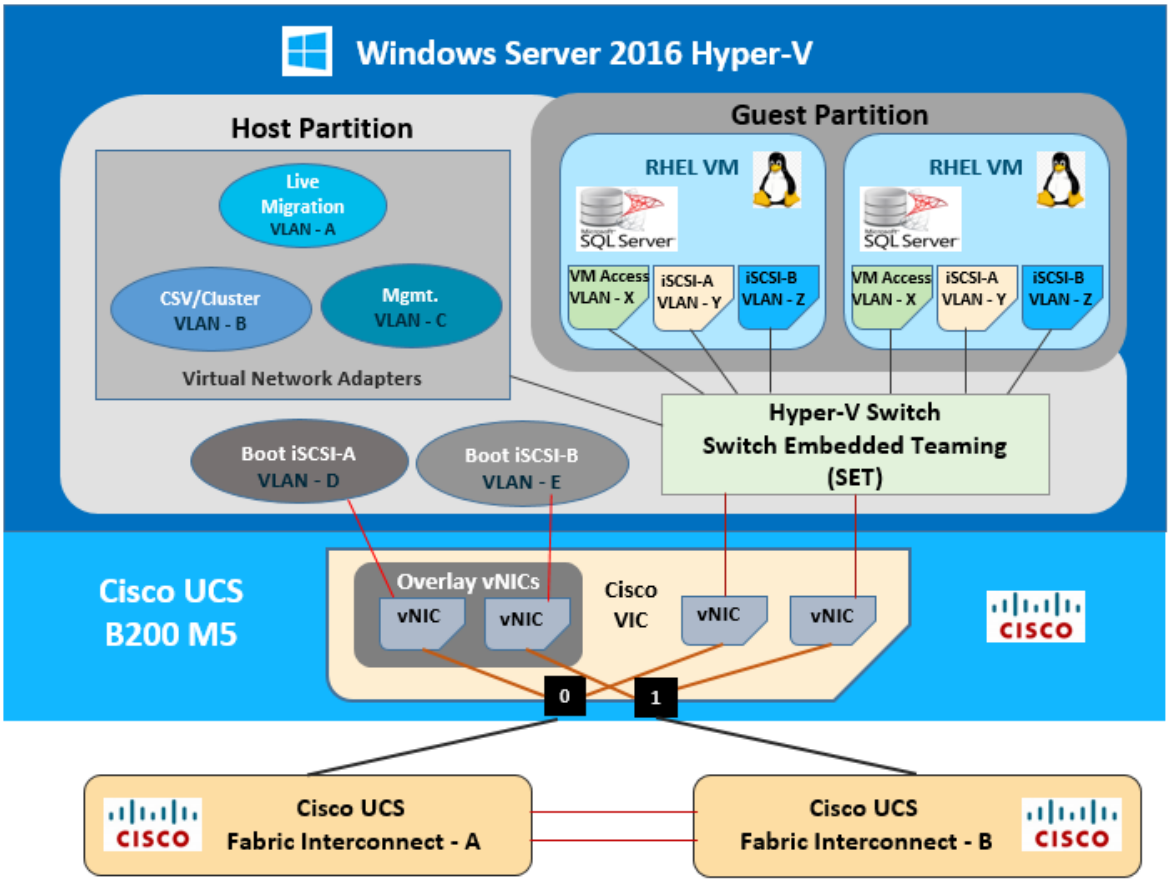
For more information about Switch Embedded Teaming (SET), refer to: https://technet.microsoft.com/en-us/library/mt403349.aspx

Figure 23 illustrates the SET enabled virtual switch deployed on all the Hyper-V hosts. SET can be managed using SCVMM or PowerShell. A total of four vNICs are created in the UCSM and exposed to the OS/hypervisor. Two NICs on the Hyper-V host are presented to the SET enabled virtual switch, on which multiple virtual network adapters (for live migration, management, cluster traffic, etc.) are created for the host partition as per the requirement. Virtual machines also connect to this virtual switch. This configuration is common where a SET enabled virtual switch is deployed on all the Hyper-V hosts using PowerShell. Apart from the two NICs used for SET enabled virtual switch, Hyper-V hosts uses two more additional iSCSI vNICs (overlay vNICs) to connect to the iSCSI storage array to boot from SAN.

If there are multiple paths configured to the boot LUN, only one path should be enabled during the installation of Windows OS. After the completion of Windows installation; enable MPIO, install NetApp Host utilities and then enable other remaining paths.

For more information, see the iSCSI boot section in the Cisco UCS Manager Server Management Guide, Release 4.0.

Figure 23 Logical Switch with SET Enabled for FlexPod Architecture



Hyper-V Failover Cluster Network

Deploying the Hyper-V cluster requires planning for several types of network traffic. Table 7 lists the different traffic types used for this solution.

Table 7 VLAN ID and Description

Network Traffic Type	VLAN ID	Description
Boot iSCSI	VLAN -A and E	For iSCSI SAN boot of Hyper-V hosts using overlay vNICs
Management	VLAN A	Used to manage the Hyper-V management operating system and virtual machines.
Cluster/CSV	VLAN B	Used for inter-node cluster communication such as the cluster heartbeat and Cluster Shared Volumes (CSV) redirection.
Live migration	VLAN C	Used for virtual machine live migration.
Virtual Machine Access	VLAN X	Used for virtual machine connectivity. Typically requires external network connectivity to service client requests

Network Traffic Type	VLAN ID	Description
Storage (In-Guest iSCSI)	VLAN – Y/Z	Used for iSCSI traffic for accessing SQL server data and log LUNs

For consistent performance and functionality, and to improve network security, it is recommended that you isolate the different types of network traffic like management, cluster, live migration, storage networks

To isolate inter-node cluster traffic, you can configure a network to either allow or not allow cluster network communication. For a network that allows cluster network communication, you can also configure whether to allow clients to connect through the network. (This includes client and management operating system access

Table 8 lists the recommended settings (with Windows PowerShell cmdlet) for each cluster network traffic using the cluster role property.

Virtual machine management traffic is not listed because these networks are isolated from the management operating system by using VLANs that are not exposed to the host. Therefore, virtual machine networks should not appear in Failover Cluster Manager as cluster networks.

Table 8 Windows Failover Cluster Network Types

Cluster Network Type	Configure Cluster Network Role Property (PowerShell Cmdlet)	Role value Network setting
Management	(Get-ClusterNetwork -Name "Mgmt_Network").Role = 3	3 - Allow cluster network communication and client connectivity
Cluster/CSV	(Get-ClusterNetwork -Name "Cluster_Network").Role = 1	1 - Allow cluster network communication only
Live Migration	(Get-ClusterNetwork -Name "LM_Network").Role = 1	1 - Allow cluster network communication only
Storage	(Get-ClusterNetwork -Name "iSCSI-A_Network").Role = 0 (Get-ClusterNetwork -Name "iSCSI-B_Network").Role = 0	0 - Do not allow cluster network communication

By default, live migration traffic uses the cluster network topology to discover available networks and to establish priority. However, you can manually configure live migration preferences to isolate live migration traffic to only the networks that you define using failover cluster manager or PowerShell.

For network recommendations for a Hyper-V cluster, refer to:

https://docs.microsoft.com/en-us/previous-versions/windows/it-pro/windows-server-2012-R2-and-2012/dn550728(v=ws.11)

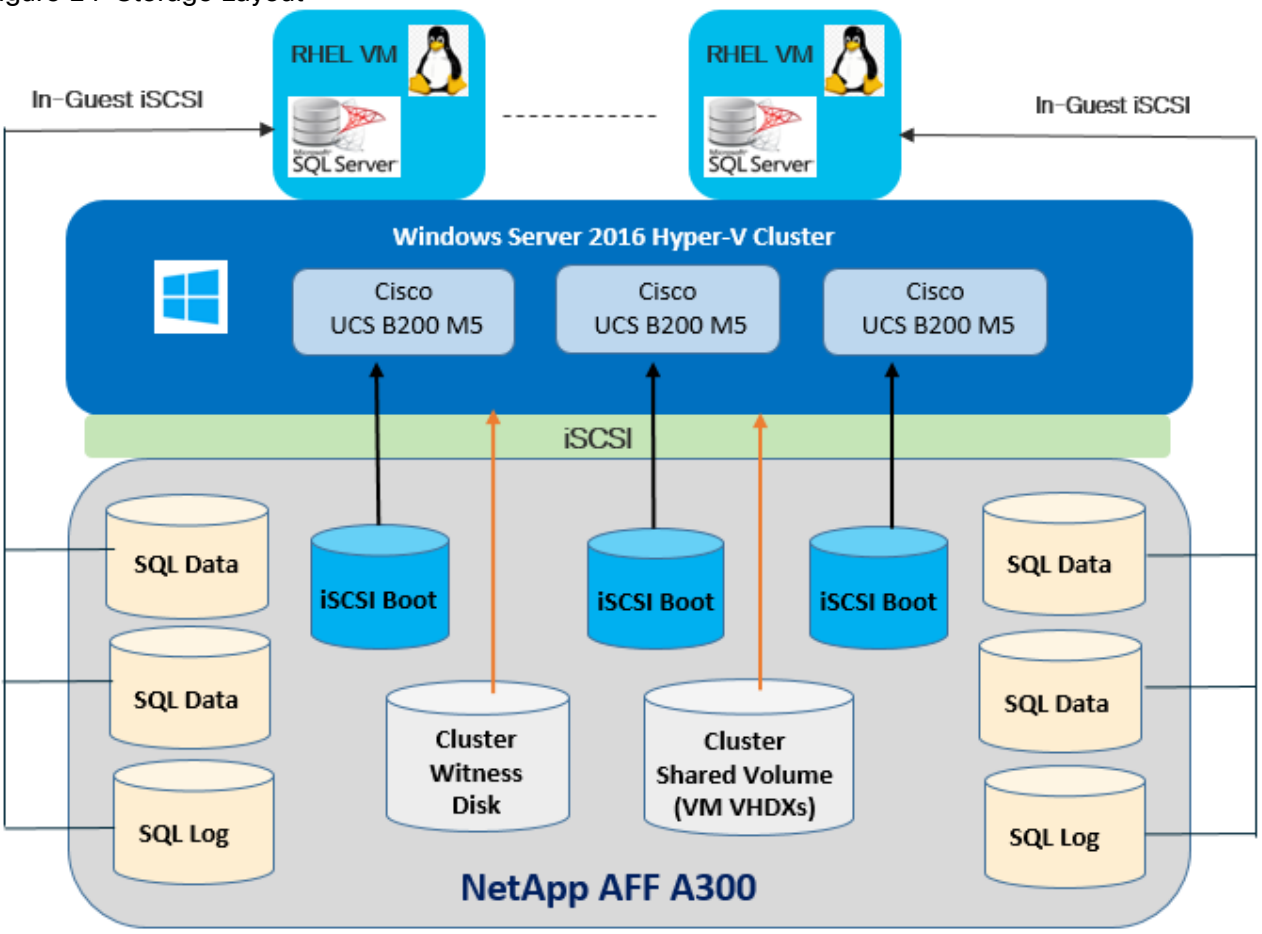
RHEL virtual machines are configured with three Virtual Machine NICs – one Virtual Machine NIC used for management traffic and two Virtual Machine NICs are used for iSCSI storage traffic. Different traffic types are segregated using individual NICs with VLAN tagging as shown in Figure 24.

Logical Layout – Storage

Figure 24 illustrates the storage layout for this solution. Cisco UCS B200 M5 servers boot from SAN using iSCSI protocol. Hyper-V hosts with Windows failover clustering provide highly available, fault tolerant environment for virtual machines running as clustered roles. A typical Hyper-V hosts consists of two or more servers sharing clustered storage and network resources. In this solution, the Hyper-V failover cluster is presented with two shared disks – witness and cluster shared volume (CSV) using the Windows iSCSI software initiator. Here, the CSV is used as a repository to store virtual machine's virtual hard disks (VHDXs) containing OS and other data files.

RHEL virtual machines running SQL server 2017 OLTP database workload boot from the VHDX residing in CSV. SQL server data and log LUNs on the storage array are directly accessed using the in-guest iSCSI feature using multiple paths.

Figure 24 Storage Layout



CSV Cache is a feature which allows you to allocate system memory (RAM) as a write-through cache. This feature is enabled by default and generally recommended to be left on as it will improve the performance of virtual machines in a Hyper-V failover cluster using CSV. A cache size of 512 MB is a good starting point/minimal value.

Since system memory is a contented resource on a Hyper-V cluster, it is recommended to keep the CSV cache size moderate, such as 512 MB, 1GB or 2 GB.

SQL Server Configuration Best Practices

For optimal SQL Server database performance, it is recommended to follow the virtual machine configuration best practices. Some of the configuration best practices are explained below.

Virtual Machine Configuration Options

vCPU: Cores Per Socket

It is recommended to align the vNUMA in virtual machine with physical NUMA of the Hypervisor host. Exposing a virtual NUMA topology into a virtual machine allows the guest operating system and any NUMA-aware applications running within it to take advantage of the NUMA performance optimizations, just as they would when running on a physical computer.

In case of VMware ESXi cluster, ensure to set one core per socket. Which enables vNUMA to select and present the best virtual NUMA topology to the guest Operating system. This will benefit the virtual machine performance by avoiding remote data reference calls. In Windows Server 2016, Hyper-V presents a virtual NUMA topology to virtual machines. By default, this virtual NUMA topology is optimized to match the NUMA topology of the underlying host computer.

Memory Reservation

SQL Server database transactions are usually CPU and memory intensive. In heavy OLTP database systems, it is recommended to reserve the required memory (instead of dynamic growth) for the SQL Virtual Machines. This makes sure that the assigned memory to the SQL Virtual Machine is committed and will eliminate the possibility of ballooning and disk swapping issues. In case of vSphere cluster, ensure to enable memory reservation. In case of Hyper-V deployments, ensure to use static memory which commits and reserves the memory for SQL Virtual Machine.

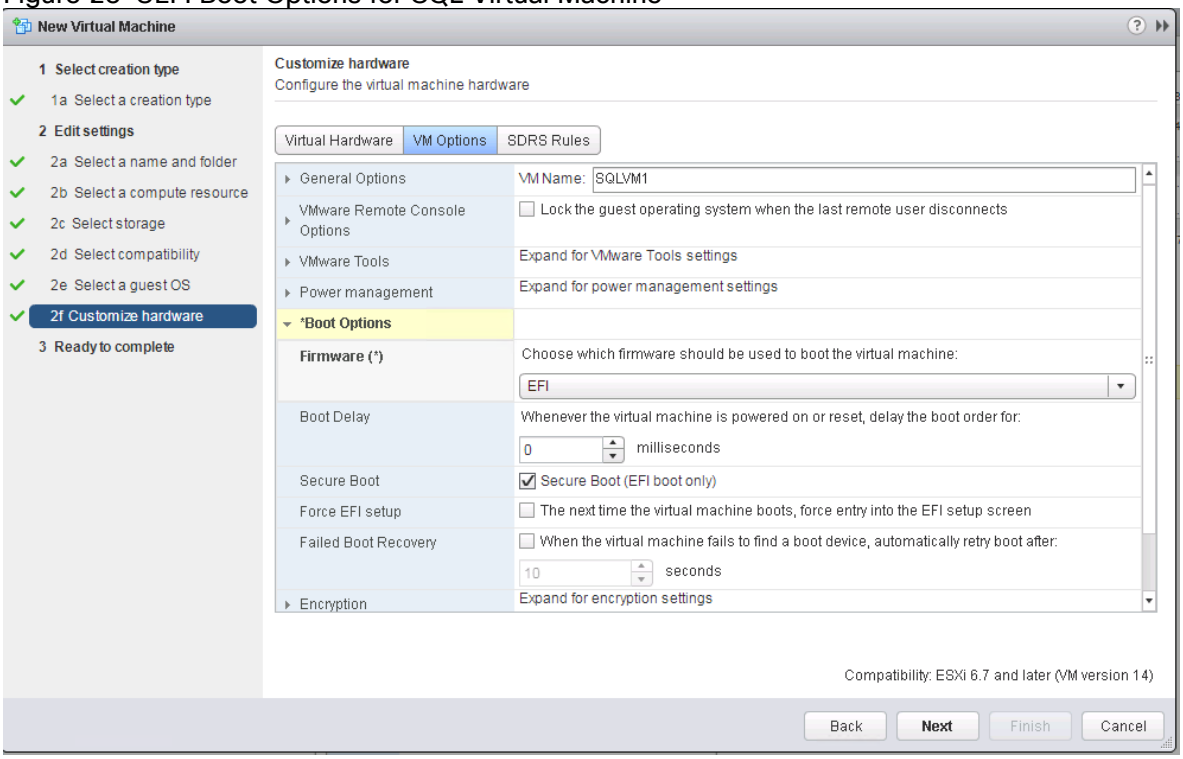
Network Adapter Type

In vSphere deployments, it is highly recommended to configure virtual machine network adaptors with "VMXNET3". VMXNET 3 is the latest generation of Para-Virtualized NICs designed for performance. It offers several advanced features including multi-queue support, Receive Side Scaling, IPv4/IPv6 offloads, and MSI/MSI-X interrupt delivery.

UEFI Boot Option for Virtual Machines

VMware vSphere cluster and Hyper-V support legacy BIOS and UEFI boot modes. If the Guest OS supports the UEFI boot mode, the virtual machines can be configured to use the UEFI boot mode along with secure boot option. If enough support is not available from the guest OS, the legacy BIOS boot option can be used. In this FlexPod system, SQL Virtual Machines are configured to use UEFI boot mode along with secure boot option as these options are supported by the guest OS RHEL 7.4. The following screen shot shows how to enable UEFI boot mode and secure boot mode options on the virtual machine's properties settings in VMware cluster.

Figure 25 UEFI Boot Options for SQL Virtual Machine



For more information about virtual machine configuration best practices for SQL Server databases workloads in a VMware environment, refer to: https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/solutions/sql-server-on-vmware-best-practices-guide.pdf

For the best practices to configure virtual machines for SQL Server workloads on Hyper-V environments, refer to: https://docs.microsoft.com/en-us/windows-server/administration/performance-tuning/role/hyper-v-server/processor-performance

Microsoft SQL Server 2017 Deployment Options

This section explains different ways to deploy and configure SQL Server for achieving high availability on RHEL virtual machines hosted on either VMware or Hyper-V cluster. SQL Server database files are stored in NetApp storage volumes. These volumes are directly accessed by the RHEL OS using in-guest software iSCSI initiator.

In this FlexPod solution, only Standalone SQL instance and Always On Availability Groups are validated and tested.

Standalone or Single SQL Instance Deployment

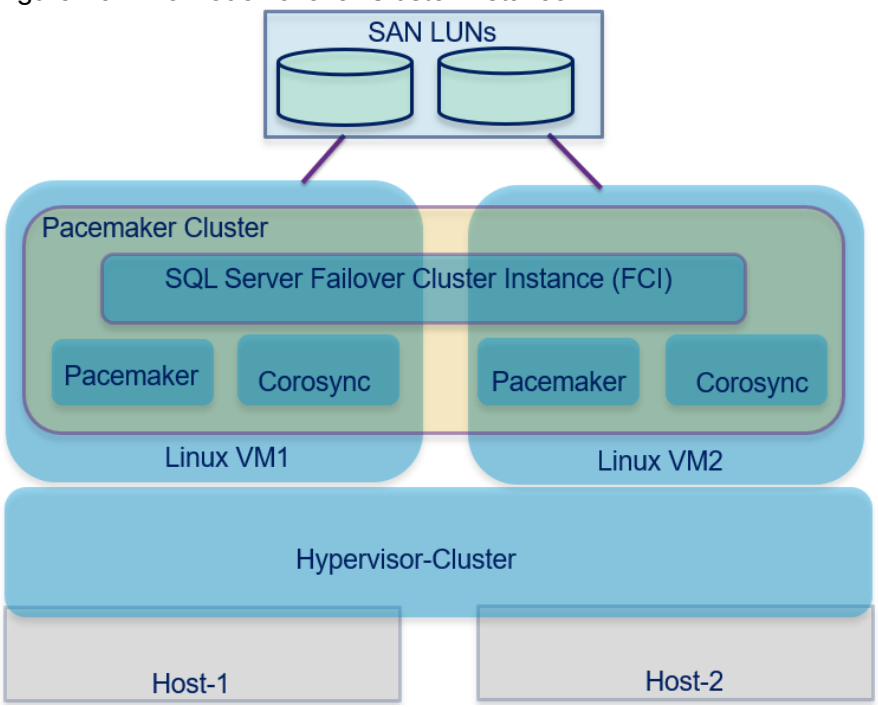
Microsoft SQL Server can be installed in a standalone virtual machine running RHEL Operating System. The SQL binaries will be installed on the RHEL virtual machine and the database files will be placed on the SAN LUNs which are directly accessed by the virtual machine using software iSCSI initiator. Once SQL engine is installed and configured, it can be connected and managed using either T-SQL or SQL Server Management studio installed on any client machine. The FlexPod solution leverages underlying hypervisor's High Availability features, such VMware

High Availability and Windows Server Failover Cluster (WSFC), to provide high availability to the stand-alone SQL Server instances. Since SQL virtual machines are directly connected to the storage LUNs using in-guest iSCSI initiator, the virtual machines can be seamlessly migrated from one hypervisor node to other node with in the cluster in case of planned or unplanned down times. Migration of a SQLVM from one node to other node involves restart of complete virtual machine during which database services will not be available.

Highly Available Failover Clustered SQL Instances on Linux

SQL Server 2017 supports Failover Cluster Instance (FCI) feature on RHEL for higher SQL instance level high availability. A Failover Cluster Instance on Linux environment leverages underlying RHEL clustering technology called pacemaker. Pacemaker is a robust and powerful open source cluster resource manager which is shipping with Red Hat Enterprise Linux 7 as high availability add on. For more information about configuring a basic two-node SQL Server Failover Clustered Instance, refer to: https://docs.microsoft.com/en-us/sql/linux/sql-server-linux-shared-disk-cluster-configure?view=sql-server-2017 The Failover Cluster Instance running on a two (more) node cluster will need the database LUNs to be shared between the nodes. The FCI will be primarily running on one node while the other node will be waiting to take over the database services in case of unplanned or unplanned downtime of primary node. The failover is automatic and no manual intervention is required. The failover times of FCI is much faster. This reduces database downtime during the failover. Figure 26 illustrates the two-node failover cluster instance on Linux virtual machines.

Figure 26 Two Node Failover Cluster Instance



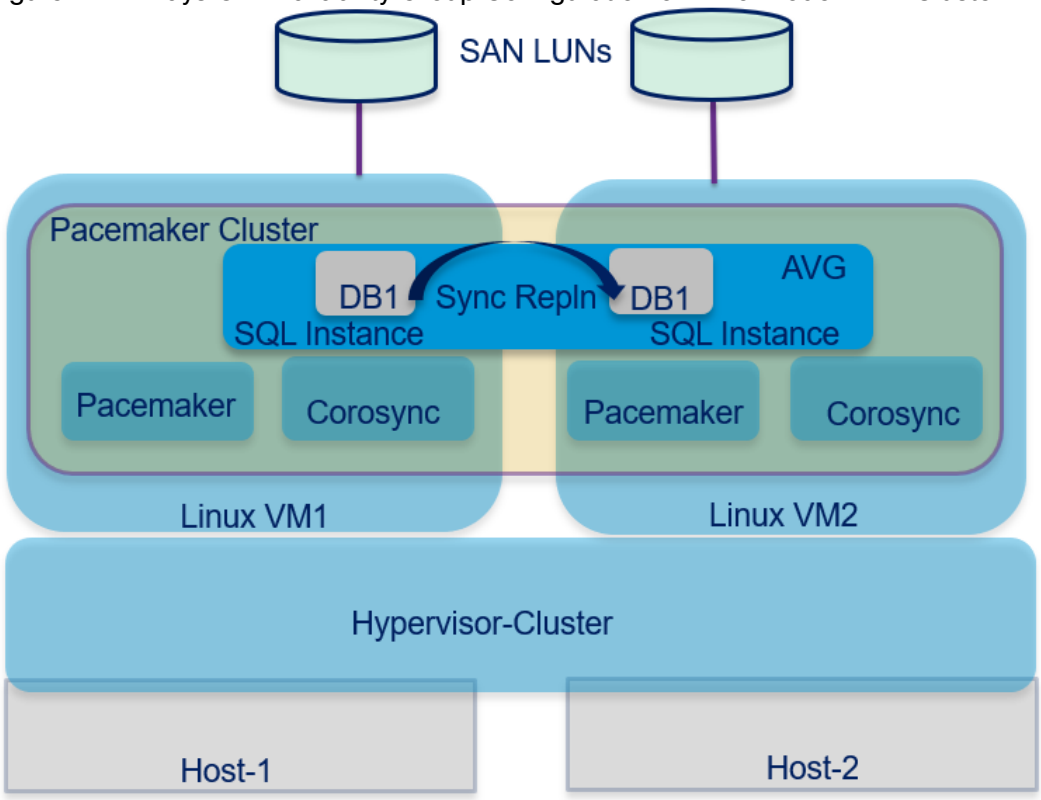
Always On Availability Group Deployment

Microsoft SQL Server 2017 also supports the Always On Availability Group feature on Linux for achieving High availability of databases that are configured for high availability. This feature can also be used as Disaster Recovery (DR) technique for achieving data recovery in case of complete unavailability of primary datacenter. This feature does not require the database LUNs to be shared between nodes. However, the standalone SQL instances running inside the clustered RHEL virtual machines will form the Availability Group and leverages the underlying

pacemaker cluster platform for managing the clustered Availability Groups. The Synchronous replication option is used for high availability of database where each user transaction will be committed on all the synchronous replicas. Asynchronous replication is used for DR purpose. For more information about the SQL Server 2017 Always On Availability Group, refer to: https://docs.microsoft.com/en-us/sql/linux/sql-server-linux-ha-basics?view=sql-server-2017

Figure 27 illustrates the two node SQL Server Always On Availability Groups configured with synchronous replication on Linux virtual machines.

Figure 27 Always On Availability Group Configuration on Two-Node RHEL Cluster



Always On Availability Group for Read-Scale on Linux

Read-Scale availability groups are designed to for specific use case where customers do not want the availability group for high-availability, instead, they want to use it to create multiple copies of the databases that span across multiple servers allowing for the spreading of a large read-only workload. Read-Scale availability groups do not require the pacemaker cluster component and does not provide high-availability or disaster recovery to the databases, it only acts as a mechanism (availability groups) to facilitate the synchronization of the databases across multiple servers. For more information about the Read-Scale Availability group, refer to: https://blogs.msdn.microsoft.com/sql_pfe_blog/2017/11/17/sql-server-2017-read-scale-availability-groups/

Validation

A high-level summary of the FlexPod Datacenter Design validation is provided in this section. The solution was validated for basic data forwarding by deploying virtual machines running the SQL Server databases. The system was validated for resiliency by failing various aspects of the system under load. The solution was also validated for virtual machine backup and restore in ESXi environment. Examples of the types of tests executed include:

- Failure and recovery of iSCSI booted Hypervisor hosts in a cluster
- Rebooting of iSCSI booted hosts
- Service Profile migration between blades
- Failure of partial and complete IOM links
- Failure and recovery of iSCSI paths to AFF nodes, Nexus switches and fabric interconnects
- Several failure tests for testing SQL Server Always On Availability Automatic failover functionality
- Backup and restore of VMs running on ESXi using SnapCenter.

Validated Hardware and Software

Table 9 lists the hardware and software versions used during solution validation. It is important to note that Cisco, NetApp, VMware, Microsoft and RedHat have interoperability matrixes that should be referenced to determine support for any specific implementation of FlexPod. Click the following links for more information:

NetApp Interoperability Matrix Tool: http://support.netapp.com/matrix/

Cisco UCS Hardware and Software Interoperability Tool: http://www.cisco.com/web/techdoc/ucs/interoperability/matrix/matrix.html

VMware Compatibility Guide: http://www.vmware.com/resources/compatibility/search.php

Microsoft Interop Matrix: https://www.windowsservercatalog.com/

RedHat Ecosystem: https://access.redhat.com/ecosystem/

Table 9 Validated Software Versions

Layer	Device	Image	Components
Compute	Cisco UCS third-generation 6332-16UP	4.0(1c)	Includes Cisco 5108 blade chassis with UCS 2304 IO Modules Cisco UCS B200 M5 blades with Cisco UCS VIC 1340 adapter
Network Switches	Includes Cisco Nexus 93180YC	NX-OS: 9.2.2	
Storage Controllers	NetApp AFF A300 storage controllers	Data ONTAP 9.5	

Layer	Device	Image	Components
Software	Cisco UCS Manager	4.0(1c)	
	Cisco UCS Manager Plugin for VMware vSphere Web Client	2.0.4	
	VMware vSphere ESXi	6.7 U1	
	VMware vCenter	6.7 U1	
	Microsoft Windows Hyper-V	Windows Server 2016	
	NetApp Virtual Storage Console (VSC)	7.2.1	
	NetApp SnapCenter	4.1.1	
	NetApp Host Utilities Kit for RHEL 7.4 & Windows 2016	7.1	
	Red Hat Enterprise Linux 7.4	7.4	
	Microsoft SQL Server	2017	

Summary

FlexPod is the optimal shared infrastructure foundation to deploy a variety of IT workloads. Cisco and NetApp have created a platform that is both flexible and scalable for multiple use cases and applications. FlexPod can efficiently and effectively support business-critical applications like databases running simultaneously from the same shared infrastructure. The flexibility and scalability of FlexPod also enable customers to start out with a right-sized infrastructure that can ultimately grow with and adapt to their evolving business requirements.

About the Authors

Gopu Narasimha Reddy, Technical Marketing Engineer, Compute Systems Product Group, Cisco Systems, Inc.

Gopu Narasimha Reddy is a Technical Marketing Engineer working with Cisco UCS Datacenter Solutions group. His current focus is to develop, test and validate solutions on Cisco UCS platform for Microsoft SQL Server databases on Microsoft Windows and VMware platforms. He is also involved in publishing TPC-H database benchmarks on Cisco UCS servers. His areas of interest include building and validating reference architectures, development of sizing tools in addition to assisting customers in SQL deployments.

Sanjeev Naldurgkar, Technical Marketing Engineer, Compute Systems Product Group, Cisco Systems, Inc.

Sanjeev has been with Cisco for six years focusing on delivering customer-driven solutions on Microsoft Hyper-V and VMware vSphere. He has over 16 years of experience in the IT Infrastructure, Server virtualization, and Cloud Computing. He holds a Bachelor's Degree in Electronics and Communications Engineering, and leading industry certifications from Microsoft and VMware.

Atul Bhalodia, Sr. Technical Marketing Engineer, NetApp Cloud Infrastructure Engineering, NetApp

Atul Bhalodia is a Sr. Technical Marketing Engineer in the NetApp and Cloud Infrastructure Engineering team. He focuses on the Architecture, Deployment, Validation, and documentation of cloud infrastructure solutions that include NetApp products. Prior to his current role, he was a Software Development Lead for NetApp SnapDrive and SnapManager Products. Atul has worked in the IT industry for more than 20 years and he holds Master's degree in Computer Engineering from California State University, San Jose, CA.

Acknowledgements

For their support and contribution to the design, validation, and creation of this Cisco Validated Design, the authors would like to thank:

- John George, Cisco Systems, Inc.

References

Products and Solutions

Cisco Unified Computing System: http://www.cisco.com/en/US/products/ps10265/index.html

Cisco UCS 6300 Series Fabric Interconnects: https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6300-series-fabric-interconnects/index.html

Cisco UCS 5100 Series Blade Server Chassis: http://www.cisco.com/en/US/products/ps10279/index.html

Cisco UCS B- Series Blade Servers: http://www.cisco.com/en/US/partner/products/ps10280/index.html

servers/index.html Cisco UCS Adapters: http://www.cisco.com/en/US/products/ps10277/prod_module_series_home.html

Cisco UCS Manager: http://www.cisco.com/en/US/products/ps10281/index.html

Cisco Nexus 9000 Series Switches: http://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/index.html

VMware vCenter Server: http://www.vmware.com/products/vcenter-server/overview.html

VMware vSphere: https://www.vmware.com/products/vsphere

Microsoft Windows Server 2016: https://docs.microsoft.com/en-us/windows-server/windows-server-2016

Microsoft SQL Server 2017: https://www.microsoft.com/en-cy/sql-server/sql-server-2017

NetApp ONTAP 9: https://www.netapp.com/us/products/data-management-software/ontap.aspx

NetApp AFF A300:http://www.netapp.com/us /products/storage-systems/all-flash-array/aff-a-series.aspx

NetApp OnCommand: http://www.netapp.com/us/products/management-software/

NetApp VSC: http://www.netapp.com/us/products/management-software/vsc/

NetApp SnapManager: http://www.netapp.com/us/products/management-software/snapmanager/