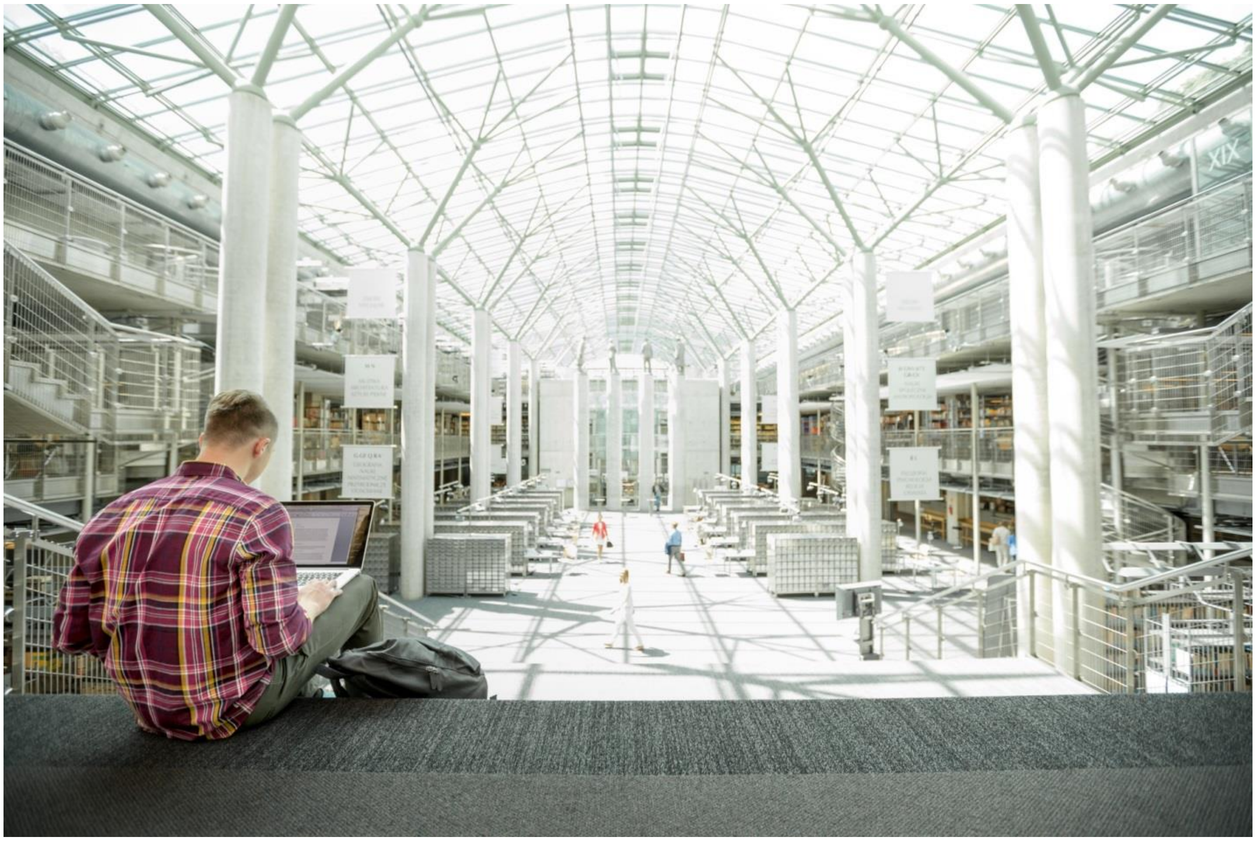


FlexPod Datacenter with Microsoft Hyper-V Windows Server 2016 Design Guide

Last Updated: August 30, 2017

About the Cisco Validated Design (CVD) Program

The CVD program consists of systems and solutions designed, tested, and documented to facilitate faster, more reliable, and more predictable customer deployments. For more information visit

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

ALL DESIGNS, SPECIFICATIONS, STATEMENTS, INFORMATION, AND RECOMMENDATIONS (COLLECTIVELY, "DESIGNS") IN THIS MANUAL ARE PRESENTED "AS IS," WITH ALL FAULTS. CISCO AND ITS SUPPLIERS DISCLAIM ALL WARRANTIES, INCLUDING, WITHOUT LIMITATION, THE WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE. IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THE DESIGNS, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

THE DESIGNS ARE SUBJECT TO CHANGE WITHOUT NOTICE. USERS ARE SOLELY RESPONSIBLE FOR THEIR APPLICATION OF THE DESIGNS. THE DESIGNS DO NOT CONSTITUTE THE TECHNICAL OR OTHER PROFESSIONAL ADVICE OF CISCO, ITS SUPPLIERS OR PARTNERS. USERS SHOULD CONSULT THEIR OWN TECHNICAL ADVISORS BEFORE IMPLEMENTING THE DESIGNS. RESULTS MAY VARY DEPENDING ON FACTORS NOT TESTED BY CISCO.

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 Windows Server 2016 Hyper-V, the popular virtualization platform in enterprise class data centers, on FlexPod.

FlexPod is a leading integrated infrastructure supporting broad range of enterprise workloads and use cases. This solution enables customers to quickly and reliably deploy 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 and C-Series rack servers, Cisco UCS 6300 or 6200 Fabric Interconnects, Cisco Nexus 9000 Series switches, Cisco MDS Fibre channel switches, and NetApp All Flash series storage arrays. In addition to that, it includes Windows Server 2016 Hyper-V, which provides a number of new features for optimizing storage utilization and facilitating private cloud.

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

Introduction

Industry trends indicate a vast data center transformation toward shared infrastructure and cloud computing. Business agility requires application agility, so IT teams need to provision applications quickly and resources need to be able to scale up (or down) in minutes.

FlexPod Datacenter is best practice data center architecture, designed and validated by Cisco and NetApp to meet the needs of enterprise customers and service providers. FlexPod Datacenter is built on NetApp All Flash FAS, Cisco Unified Computing System (UCS), and the Cisco Nexus family of switches. These components combine to enable management synergies across all of a business's IT infrastructure. FlexPod Datacenter has been proven to be the optimal platform for virtualization and workload consolidation, enabling enterprises to standardize their IT infrastructure.

Changes in FlexPod

FlexPod Datacenter with Windows Server 2016 Hyper-V introduces new hardware and software into the portfolio, enabling end-to-end 40GbE along with native 16Gb FC via the Cisco MDS Fibre Channel switch. New pieces which enable the move to 40Gb include:

- NetApp AFF A300
- Cisco Nexus 9332PQ
- Windows Server 2016 Hyper-V
- NetApp ONTAP® 9.1
- Cisco UCS 3.1(3a)

Audience

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

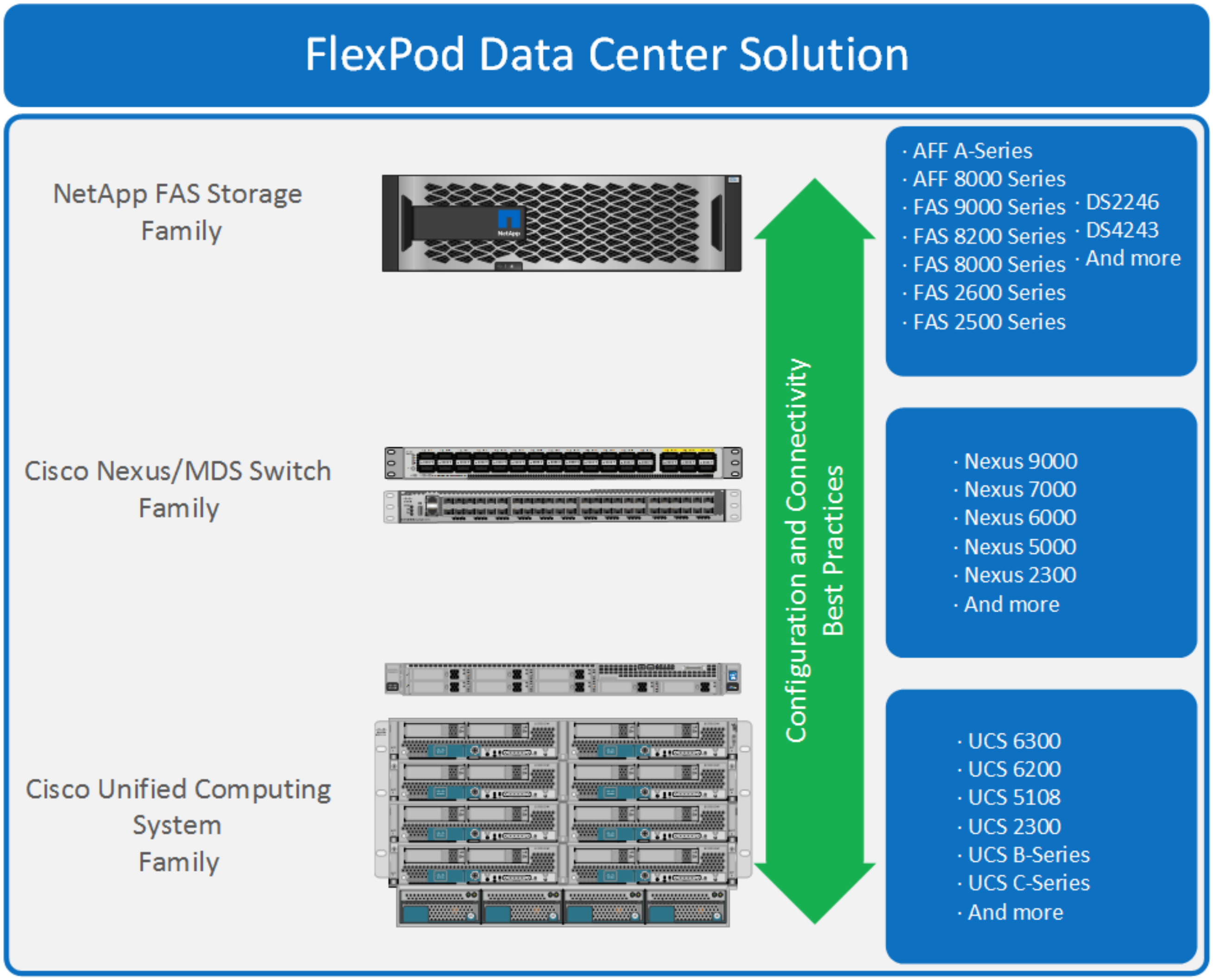
Technology Overview

FlexPod System Overview

FlexPod is a best practice datacenter architecture that includes the following components:

- Cisco Unified Computing System (Cisco UCS)
- Cisco Nexus switches
- Cisco MDS switches
- NetApp All Flash FAS (AFF) systems

Figure 1 FlexPod Component Families



These components are connected and configured as per the recommended guidelines of both Cisco and NetApp to provide an ideal platform for running a variety of enterprise workloads with confidence. FlexPod can scale up for greater performance and capacity (adding compute, network, or storage resources individually as needed), or it can scale out for environments that require multiple consistent deployments (such as rolling out of additional FlexPod stacks). The reference architecture covered in this document leverages Cisco Nexus 9000 series for network switching and Cisco MDS 9000 series for SAN switching.

One of the key benefits of FlexPod is its ability to maintain consistency during scale. Each of the component families shown (Cisco UCS, Cisco Nexus, and NetApp AFF) offers platform and resource options to scale the infrastructure up or down, while supporting the same features and functionality that are required under the configuration and connectivity best practices of FlexPod.

Design Choices

SAN Architecture

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 highlights the below listed FlexPod architectures designed based on the combination of hardware availability and the usage of SAN protocols to deploy Windows Server 2016 Hyper-V clusters, utilizing the Cisco UCS 6332-16UP Fabric Interconnect. This solution also highlights the deployment of Windows Server 2016 Hyper-V clusters using Cisco UCS 6248UP Fabric Interconnect.

The FlexPod architectures for Windows Server 2016 Hyper-V are:

- Design 1 - 16Gb FC with 40GbE IP Architecture
- Design 2 - 8Gb FC with 10GbE IP Architecture
- Design 3 – 40GbE End-to-End IP Architecture
- Design 4 - 10GbE IP Architecture
- Design 5 – 10Gb FCoE and 40GbE IP Architecture
- Design 6 – 10Gb FCoE and 10GbE IP Architecture

All the above designs discussed in this design guide provide Hyper-V host access to both block-level and file-level storage simultaneously.

Access to block-level storage (FC/FCoE/iSCSI) is limited to LUNs for SAN boot and cluster witness disk; and file-level storage is used to access the SMB/CIFS shares for generic virtual machine (VM) workloads.

The use of Microsoft System Center 2016 - Virtual Machine Manager and Operations Manager in this solution simplify the deployment, management and monitoring of Hyper-V hosts and clusters. And with Cisco UCS management integrations with Microsoft, it further simplifies the tool set to monitor and manage bare-metal, virtualized and cloud environments

Design 1 – 16Gb FC with 40GbE IP Architecture

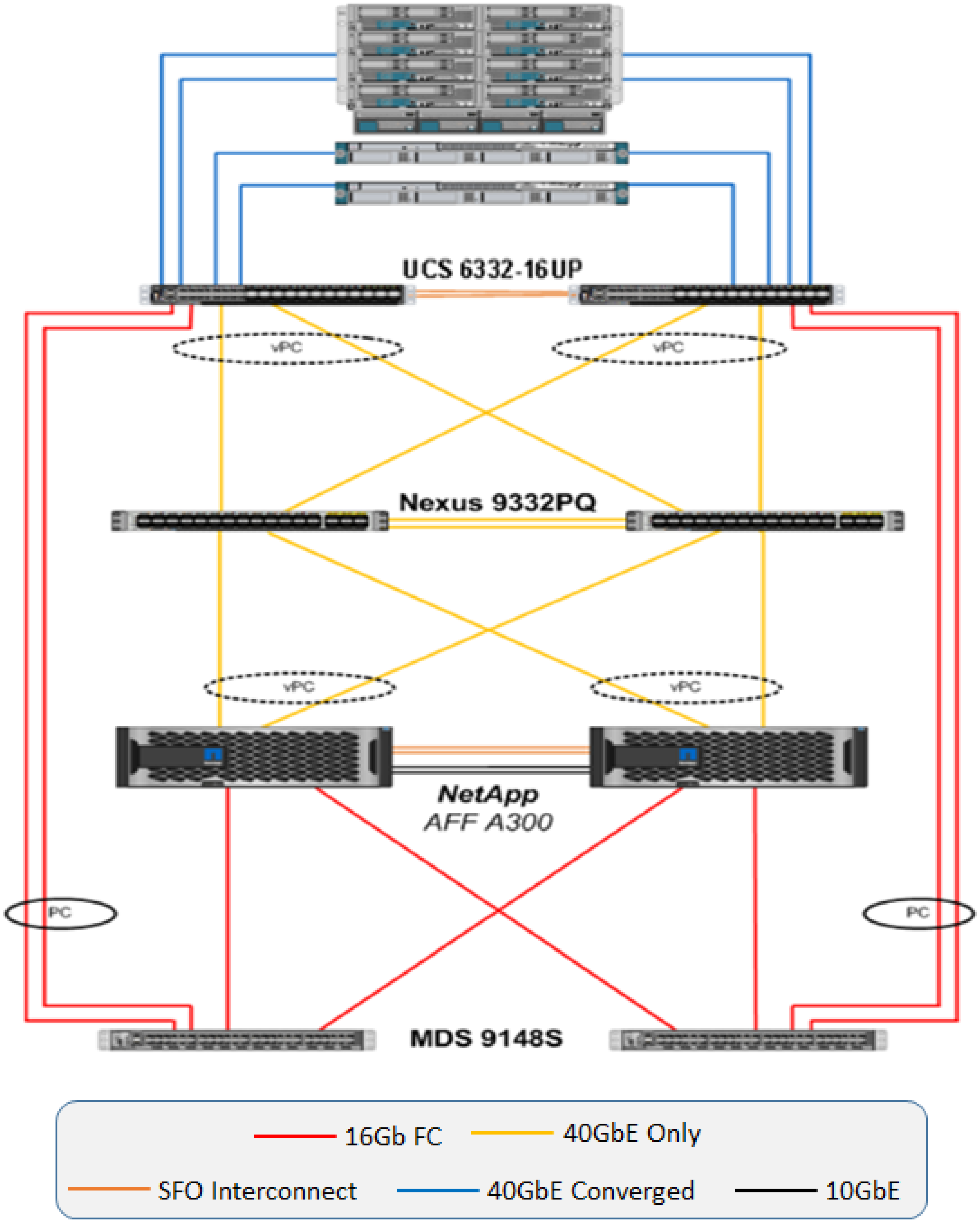
The components used in this design are:

- NetApp AFF A300 storage controllers
  - High Availability (HA) pair in switchless cluster configuration
  - Onboard 16G Unified Target Adapter 2 ports of each storage controller
  - ONTAP 9.1
- Cisco Nexus 9332PQ switches

  - Pair of switches in vPC configuration
- Cisco MDS 9148s
- Cisco UCS 6332-16UP Fabric Interconnect
- 40Gb Unified Fabric
- Cisco UCS 5108 Chassis
  - Cisco UCS 2304 IOM
  - Cisco UCS B200 M4 servers with VIC 1340
  - Cisco UCS C220 M4 servers with VIC 1385

Figure 2 illustrates the FlexPod Datacenter topology that supports 16Gb FC and 40GbE IP design. The MDS provides 4/8/16G FC switching capability. The FC storage connects to the MDS at 16Gbps and the IP storage connects to the Cisco Nexus at 40GbE. The 16Gb FC connectivity provides access to block-based LUNs and the 40GbE IP connectivity provides access to file-based SMB/CIFS shares on the storage. 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 16Gb FC and 40GbE compute and network fabric increases the overall capacity of the system.

Figure 2 FlexPod Datacenter with Cisco UCS 6332-16UP Fabric Interconnects for 16Gb FC with 40GbE IP Architecture



Design 2 – 8Gb FC with 10GbE IP Architecture

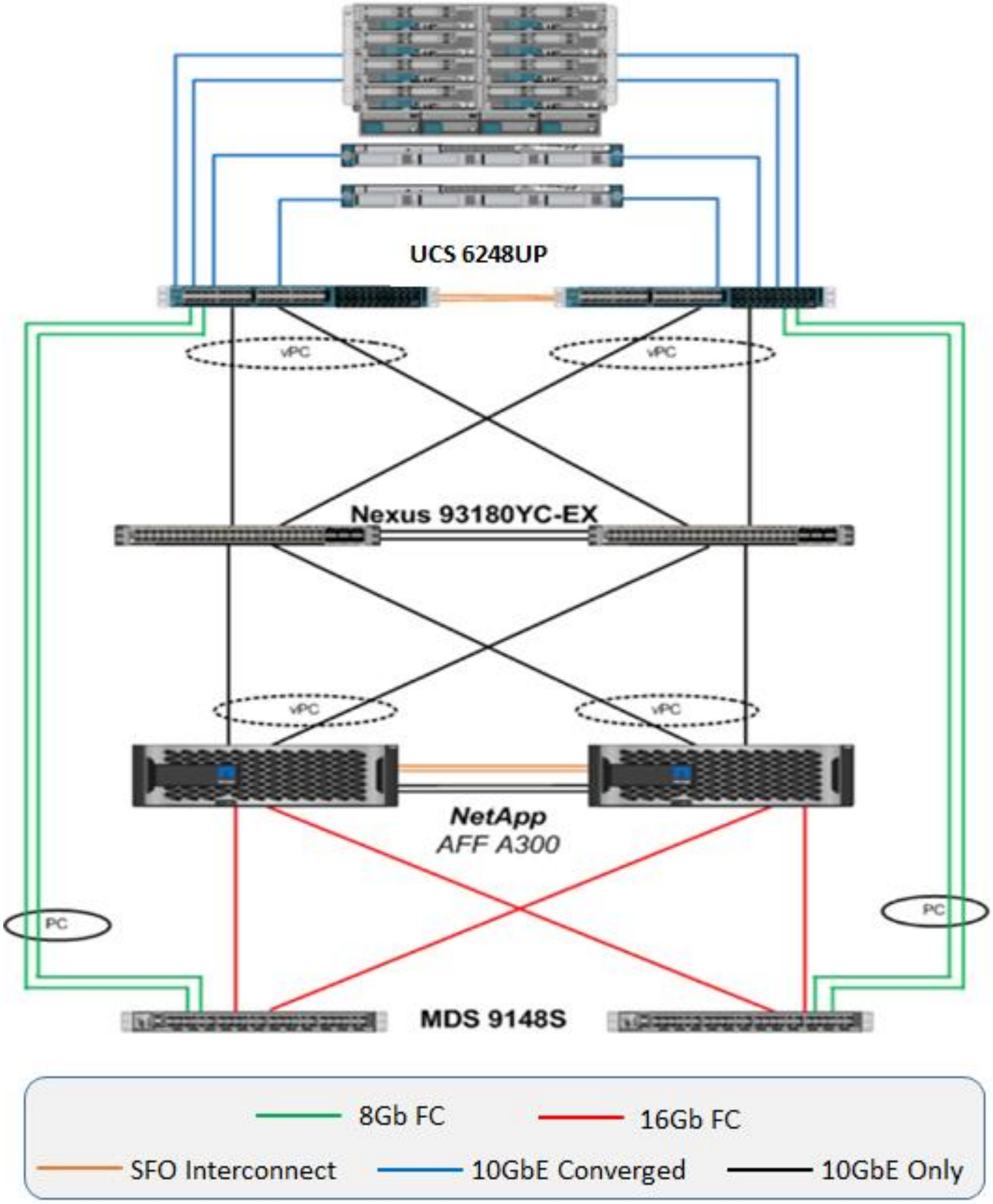
The components used in this design are:

- NetApp AFF A300 storage controllers

  - HA pair in switchless cluster configuration
  - Onboard 16G Unified Target Adapter 2 ports of each storage controller
  - ONTAP 9.1
- Cisco Nexus 93180YC-EX switches
  - Pair of switches in vPC configuration
- Cisco MDS 9148s
  - 16Gb connectivity between NetApp AFF A300 and MDS
  - 8Gb connectivity between MDS and Cisco UCS 6248UP Fabric Interconnects
- Cisco UCS 6248UP Fabric Interconnect
- 10Gb Unified Fabric
- Cisco UCS 5108 Chassis
  - Cisco UCS 2204/2208 IOM
  - Cisco UCS B200 M4 servers with VIC 1340
  - Cisco UCS C220 M4 servers with VIC 1227

Figure 3 illustrates the FlexPod Datacenter topology that supports 8Gb FC and 10GbE IP design. The MDS provides 4/8/16Gb FC switching capability. The 8Gb FC connectivity provides access to block-based LUNs and the 10GbE IP connectivity provides access to file-based SMB/CIFS shares on the storage. The FC storage connects to the MDS at 16Gbps and FC compute connects to the MDS at 8Gbps. The IP storage connects to the Cisco Nexus at 10GbE. Additional links can be added to the SAN port channels between the MDS switches and Cisco UCS fabric interconnects to equalize the bandwidth throughout the solution.

Figure 3 FlexPod Datacenter with Cisco UCS 6248UP Fabric Interconnects for 8Gb FC and 10GbE IP Architecture



Design 3 – 40GbE End-to-End IP Architecture

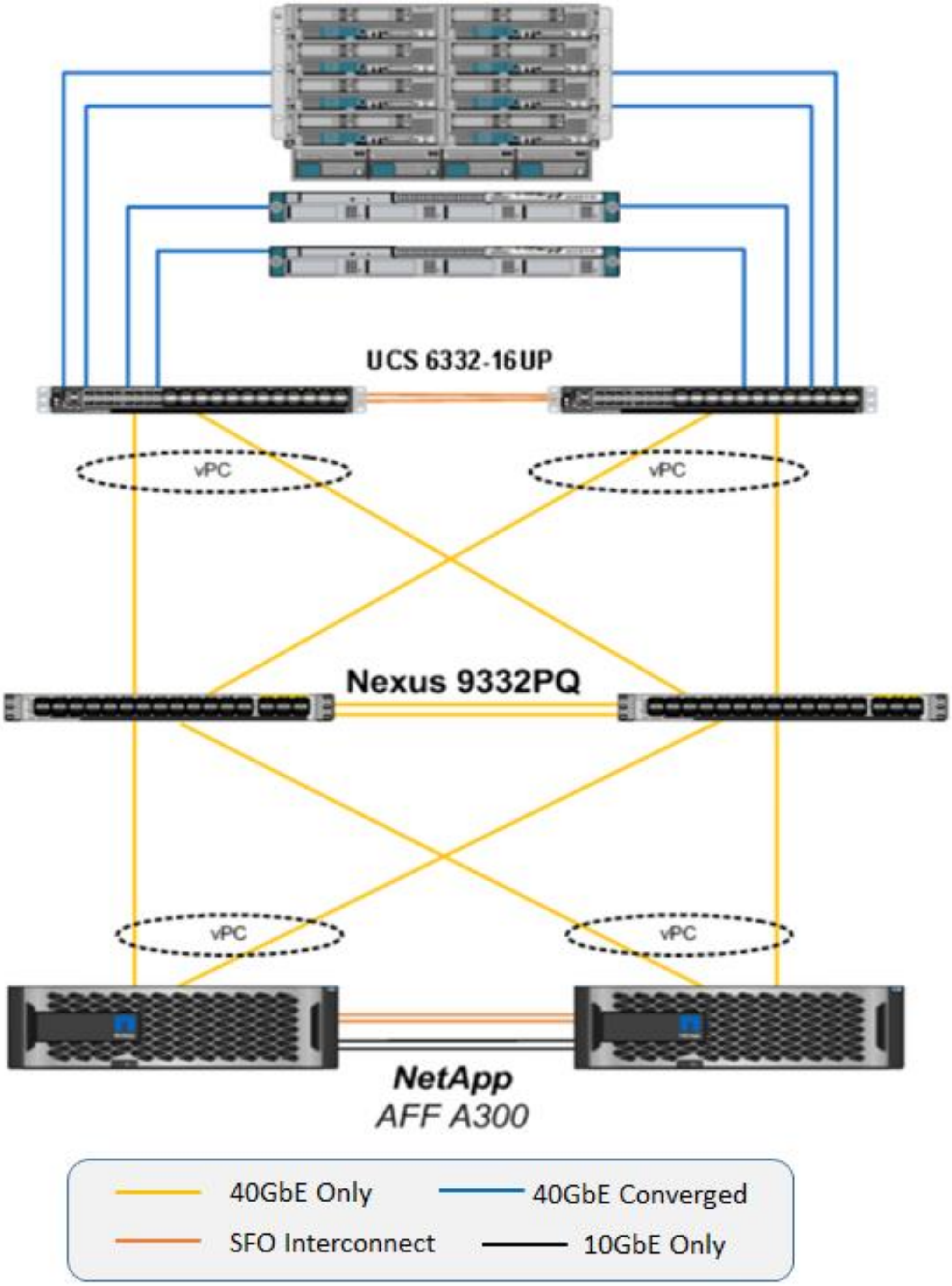
The 40GbE IP architecture is covered in the Appendix section of FlexPod Datacenter with Microsoft Hyper V Windows Server 2016 Deployment Guide: http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpod_mspc_ws2016.pdf

The components used in this design are:

- NetApp AFF A300 storage controllers
  - HA pair in switchless cluster configuration
  - 40GbE adapter used in the expansion slot of each storage controller
  - ONTAP 9.1
- 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 M4 servers with VIC 1340
  - Cisco UCS C220 M4 servers with VIC 1385

Figure 4 illustrates the FlexPod Datacenter topology that supports the end-to-end 40GbE IP design. 40GbE IP connectivity provides access to both block-based iSCSI LUNs and file-based SMB/CIFS shares on the storage. 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 solutions.

Figure 4 FlexPod Datacenter with Cisco UCS 6332-16UP Fabric Interconnects for 40GbE end-to-end IP Architecture



Design 4 – 10GbE IP Architecture

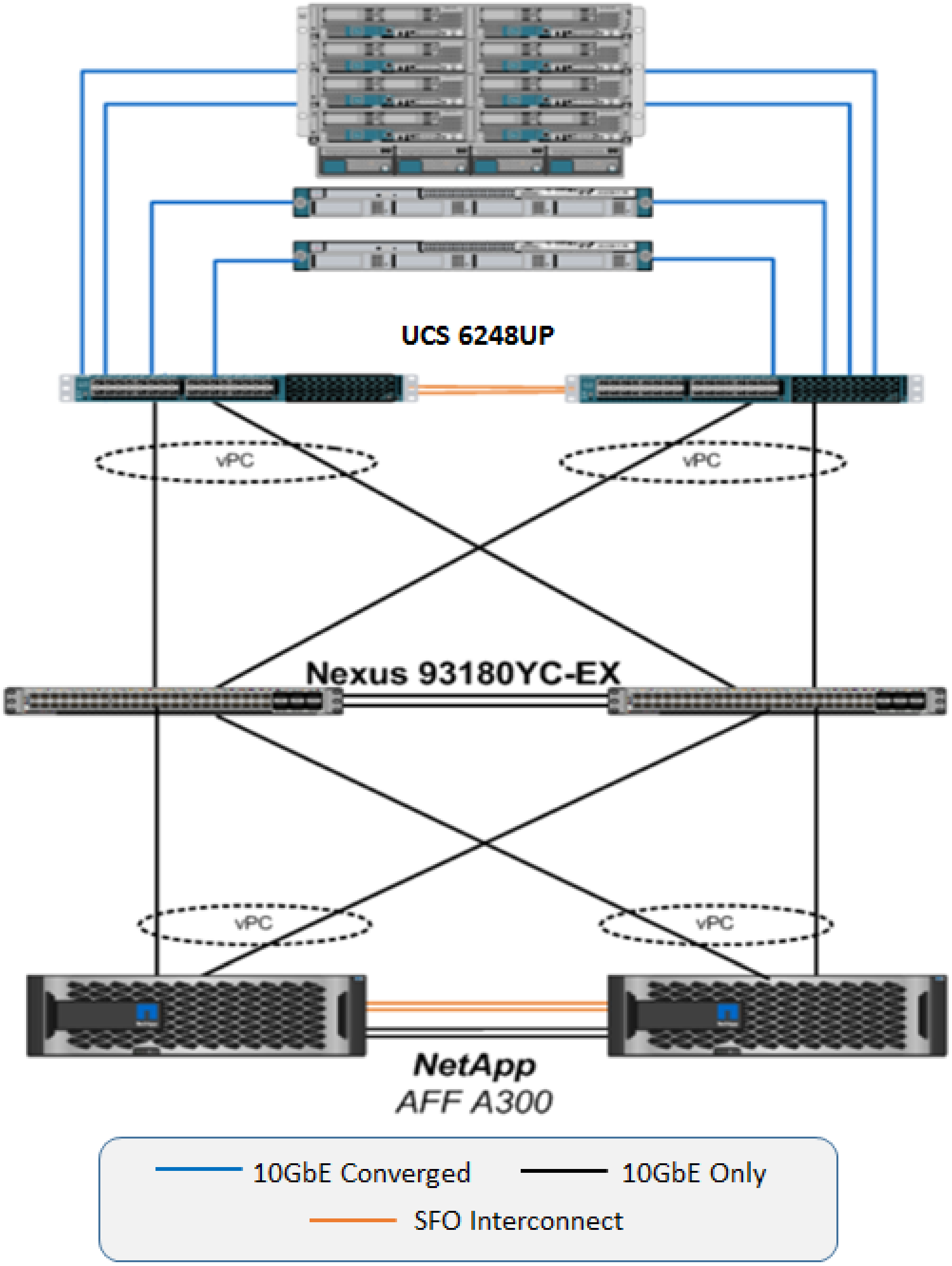
The 10GbE IP architecture is covered in the Appendix section of FlexPod Datacenter with Microsoft Hyper V Windows Server 2016 Deployment Guide: http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpod_mspc_ws2016.pdf

The components used in this design are:

- NetApp AFF A300 storage controllers
  - HA pair in switchless cluster configuration
  - Onboard 10GbE Unified Target Adapter 2 ports of each storage controller
  - ONTAP 9.1
- Cisco Nexus 93180YC-EX switches
  - Pair of switches in vPC configuration
- Cisco UCS 6248UP Fabric Interconnect
- 10Gb Unified Fabric
- Cisco UCS 5108 Chassis
  - Cisco UCS 2204/2208 IOM
  - Cisco UCS B200 M4 servers with VIC 1340
  - Cisco UCS C220 M4 servers with VIC 1227

Figure 5 illustrates the FlexPod Datacenter topology that supports 10GbE IP design. The 10GbE IP connectivity provides access to block-based iSCSI LUNs and also provides access to file-based SMB/CIFS shares on the storage.

Figure 5 FlexPod Datacenter with Cisco UCS 6248UP Fabric Interconnects for 10GbE IP Architecture



Design 5 – 10Gb FCoE and 40GbE IP Architecture

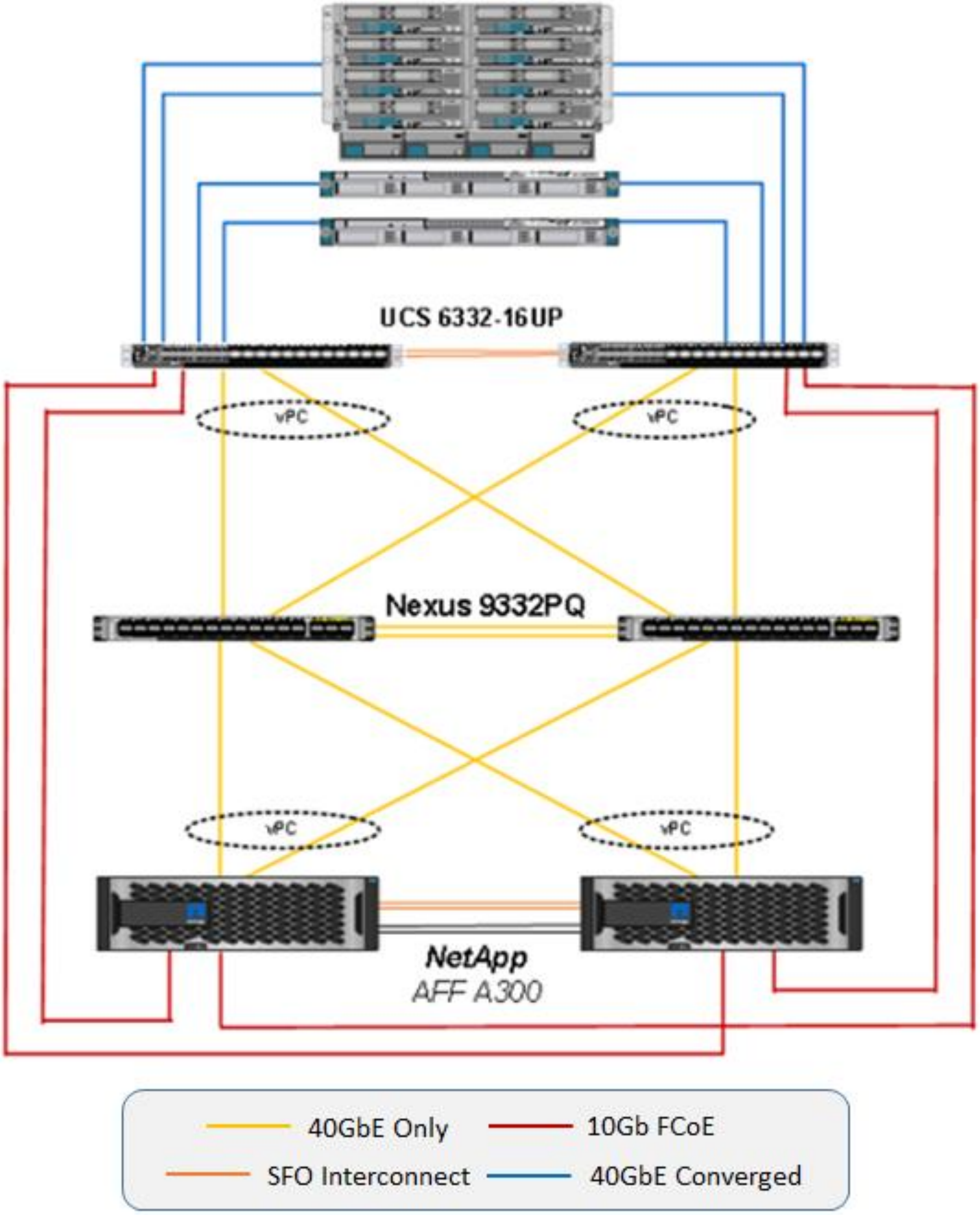
The 10Gb FCoE and 40GbE IP architecture is covered in the Appendix section of FlexPod Datacenter with Microsoft Hyper V Windows Server 2016 Deployment Guide: http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpod_mspc_ws2016.pdf

The components used in this design are:

- NetApp AFF A300 storage controllers
  - HA pair in switchless cluster configuration
  - Onboard 10G Unified Target Adapter 2 ports of each storage controller
  - ONTAP 9.1
- Cisco Nexus 9332PQ switches
  - Pair of switches in vPC configuration
- Cisco UCS 6332-16UP Fabric Interconnect
- Cisco UCS 5108 Chassis
  - Cisco UCS 2304 IOM
  - Cisco UCS B200 M4 servers with VIC 1340
  - Cisco UCS C220 M4 servers with VIC 1385

Figure 6 illustrates the FlexPod Datacenter topology that supports the 10Gb FCoE and 40GbE IP design. The 10Gb FCoE connectivity provides access to block-based FCoE LUNs and the 40GbE IP connectivity provides access to file-based SMB/CIFS shares on the storage. FCoE ports from storage are directly connected to the Cisco UCS Fabric Interconnects. Fiber channel zoning is done in the fabric interconnect, which is in fiber channel switching mode. This design would also support FC connectivity between storage and the fabric interconnects.

Figure 6 FlexPod Datacenter with 10Gb FCoE and 40GbE IP Architecture using Direct Connect Storage



Design 6 – 10Gb FCoE and 10GbE IP Architecture

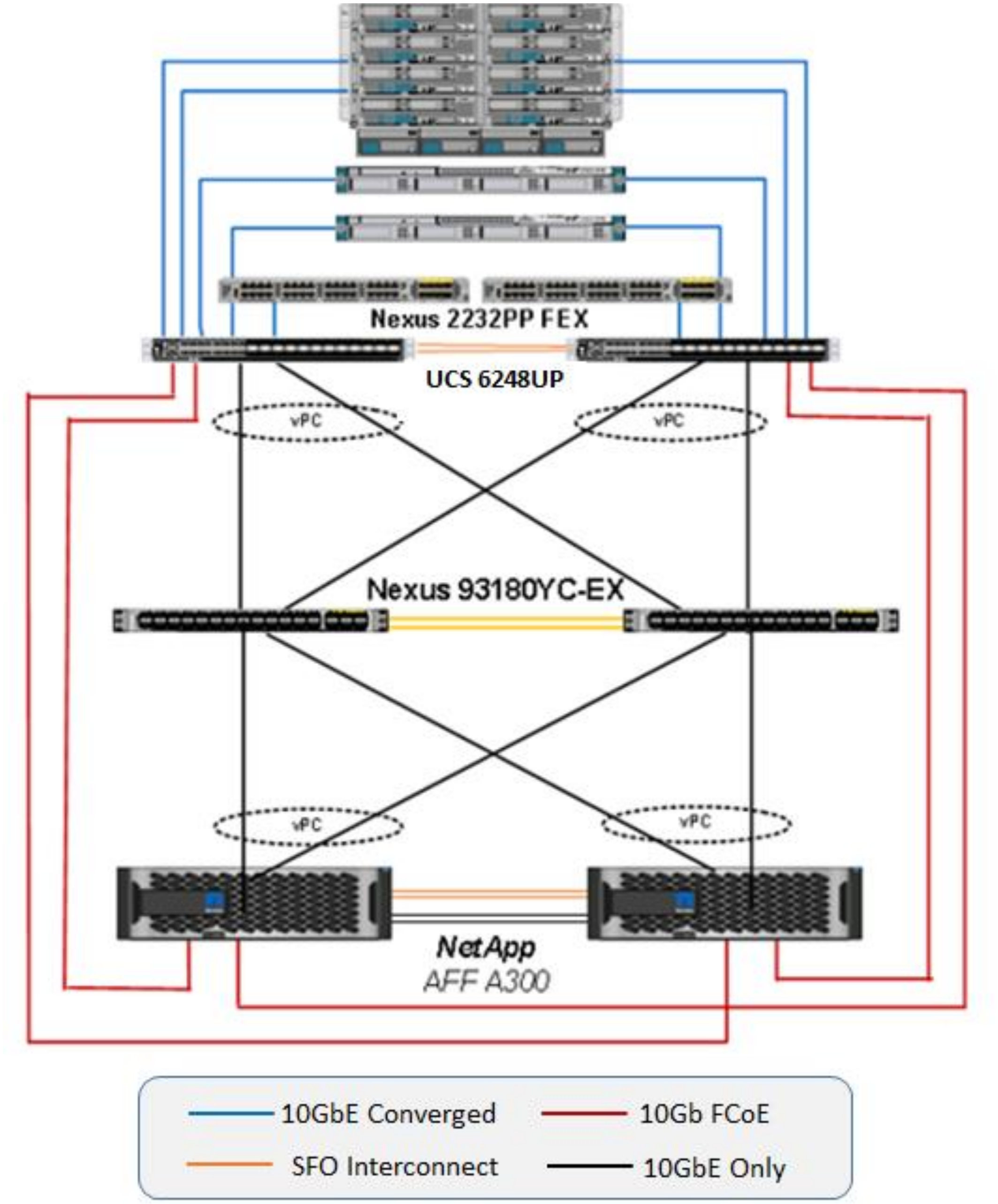
The 10Gb FCoE and 10GbE IP architecture is covered in the Appendix section of FlexPod Datacenter with Microsoft Hyper V Windows Server 2016 Deployment Guide: http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpod_mspc_ws2016.pdf

The components used in this design are:

- NetApp AFF A300 storage controllers
  - HA pair in switchless cluster configuration
  - Onboard 10G Unified Target Adapter 2 ports of each storage controller
  - ONTAP 9.1
- Cisco Nexus 93180YC-EX switches
  - Pair of switches in vPC configuration
- Cisco UCS 6248UP Fabric Interconnect
- Cisco UCS 5108 Chassis
  - Cisco UCS 2204/2208 IOM
  - Cisco UCS B200 M4 servers with VIC 1340
  - Cisco UCS C220 M4 servers with VIC 1227

Figure 7 illustrates the FlexPod Datacenter topology that supports the 10Gb FCoE and 10GbE IP design. The 10Gb FCoE connectivity provides access to block-based FCoE LUNs and the 10GbE IP connectivity provides access to file-based SMB/CIFS shares on the storage. FCoE ports from storage are directly connected to the Cisco UCS Fabric Interconnects. Fiber channel zoning is done in the fabric interconnect, which is in fiber channel switching mode. This design would also support FC connectivity between storage and the fabric interconnects.

Figure 7 FlexPod Datacenter with 10Gb FCoE and 10GbE IP Architecture using Direct Connect Storage



Logical Layout

This solution uses Microsoft System Center 2016 VMM to deploy and manage the Hyper-V hosts and cluster. This section explains the VMM networking and storage fabric setup for this deployment guide.

Network Layout

The figure below illustrates the building blocks that make up the VMM network fabric used for this solution. These building blocks help to consistently configure network adapters with identical capabilities for all Hyper-V hosts by using port profiles and logical switches. Port profiles and logical switches act as containers for the properties or capabilities that you want your network adapters to have. Instead of configuring individual properties or capabilities for each network adapter, you can specify the capabilities in port profiles and logical switches, which you can then apply to the appropriate adapters. This is more like creating a template with the different building blocks and then applying it to Hyper-V host network adapters, thus simplifying the configuration process.

Logical Networks

The first building block is to create logical network. Logical networks are modelled based on the physical network. Six networks (for management, cluster, live migration, SMB and VM traffic) required to validate the Hyper-V cluster in the lab are modelled as logical networks. All the six networks are on the same physical network but separated by VLANs that is controlled by setting the VLAN ID on the virtual network adapter. The physical ports on the switch are configured in trunk mode to carry all the various VLAN traffic

- MS-IB-MGMT: This logical network will be used for management traffic and has its own IP subnet and VLAN.
- MS-Cluster: This network will be used for Microsoft Hyper-V cluster communication and will have its own IP subnet and VLAN.
- MS-LVMN: This network will be used for Live Migration traffic and will have its own IP subnet and VLAN
- MS-SMB-1 and MS-SMB-2: This network will be used for SMB file share access/traffic and has its own IP subnet and VLAN.
- MS-Tenant-VM (optional): This network will be used for all the VM traffic.

VM Networks

VM networks help to abstract the virtual machines from the underlying logical network. VM Network is configured as per the VLAN-based configuration on top of the logical network. For this configuration, one VM Network for each network site (and VLAN) is created to match the underlying physical and logical networks. A VMs virtual network adapter connects only to a VM network.

Port Profiles and Logical Switches

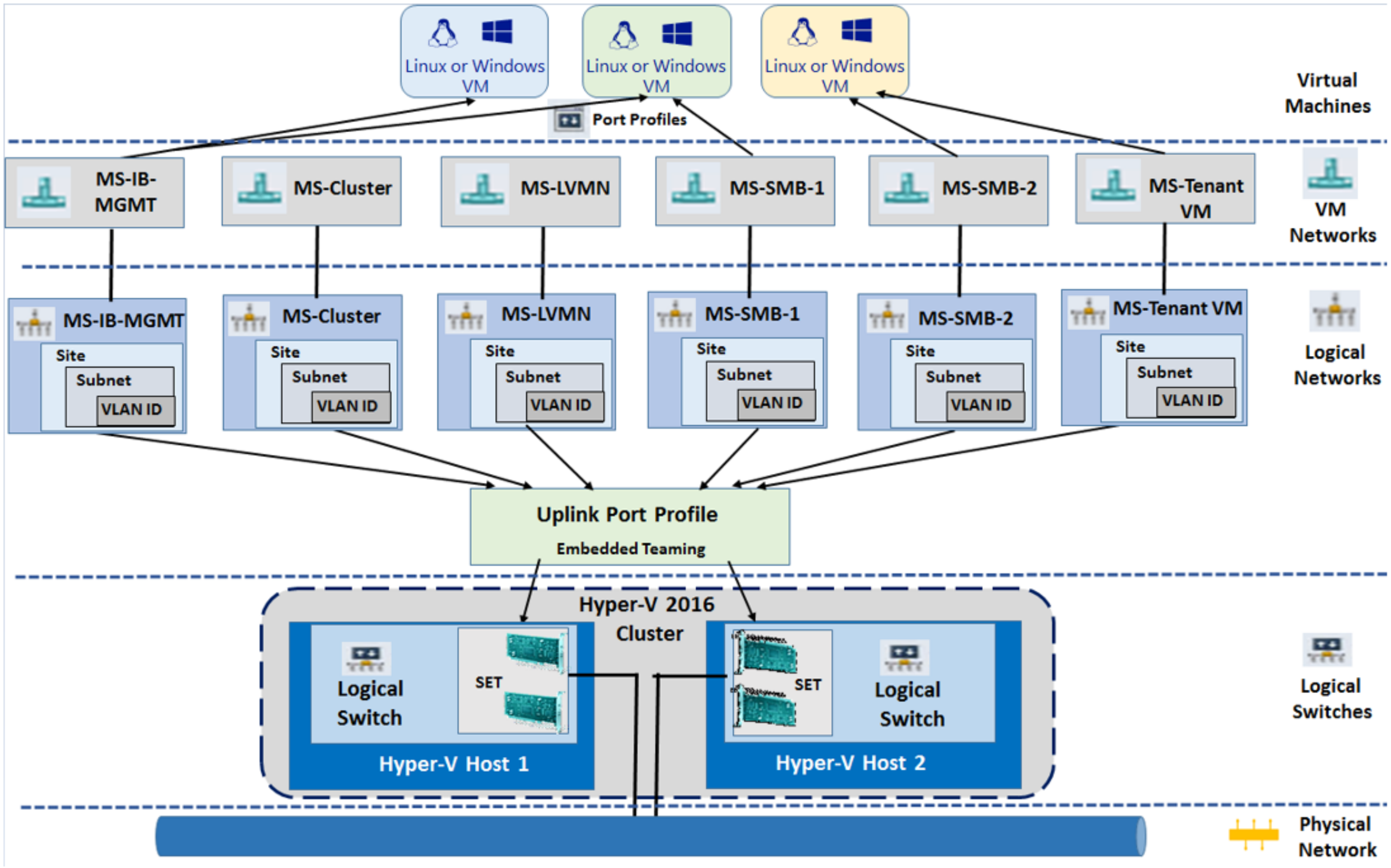
The virtual port profiles are configured with built-in profiles in VMM for common network adapter use cases. For defining uplink port profiles, host defaults for load balancing and switch independent for teaming mode is used.

The logical switch brings together port profiles for uplinks (for physical adapters), port profiles and port classification (for virtual network adapters), so that it can be applied to multiple network adapters across all Hyper-V hosts. The logical switch created uses the new embedded team as the uplink mode to deploy the switch with SET-based teaming. This logical switch is deployed to the Hyper-V hosts.

The below URL provides more information on planning and managing networking fabric in VMM:

https://docs.microsoft.com/en-us/system-center/vmm/plan-network https://docs.microsoft.com/en-us/system-center/vmm/manage-networks

Figure 8 VMM Networking fabric setup



Switch Embedded Teaming Overview

Embedded team selected as the uplink mode while creating the logical switch in VMM is Microsoft's new feature called "Switch Embedded Teaming (SET)" introduced in Windows Server 2016.

SET 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.

Physical NICs requirement for SET:

- 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.

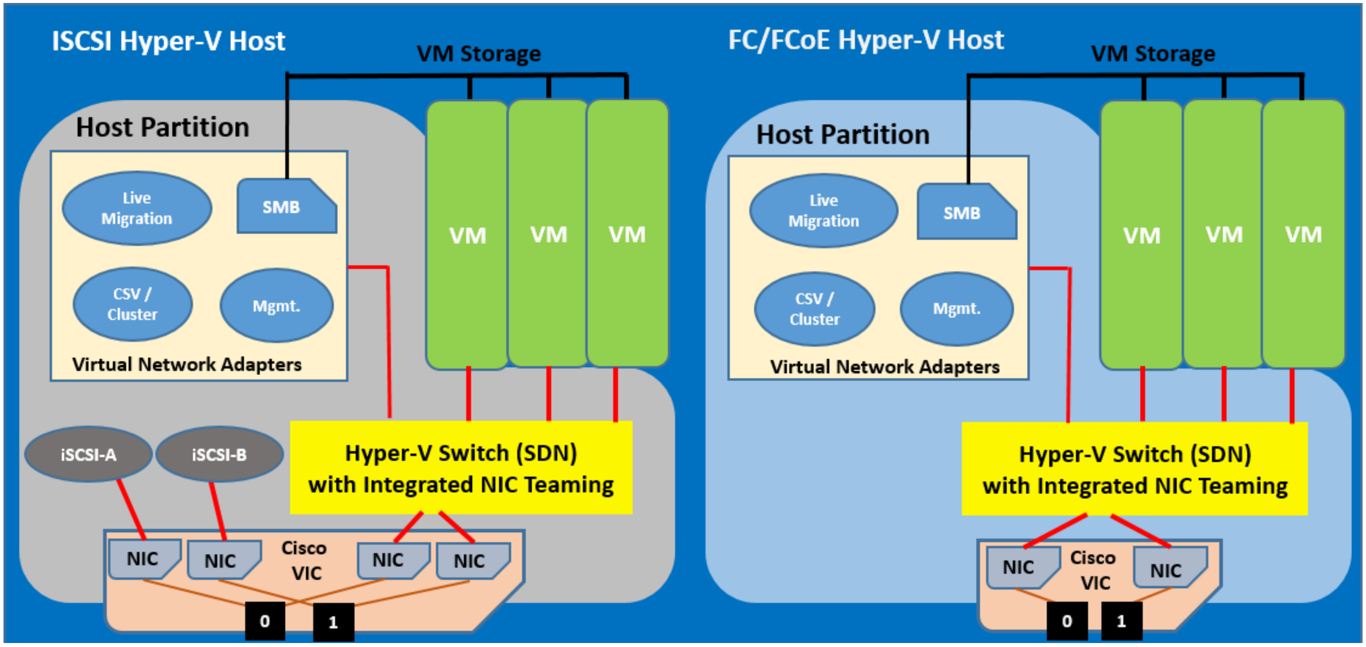
The below URL provides more information on Switch Embedded Teaming (SET):

https://technet.microsoft.com/en-us/library/mt403349.aspx

The figure below illustrates logically the SET based logical switch deployed on all the Hyper-V hosts using System Center 2016 VMM.

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 logical switch is deployed on all the Hyper-V hosts using SCVMM. Apart from the two NICs used for SET, Hyper-V hosts that boot from SAN using iSCSI protocol use two more additional NICs to connect to the iSCSI storage array. Hyper-V hosts that boot from SAN using FC/FCoE protocol use virtual host bus adapters (vHBA) to connect to the FC/FCoE storage array.

Figure 9 Logical Switch with SET enabled for FlexPod Architecture



NetApp A-Series All Flash FAS

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 via 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 the workaround, please see the NetApp public report. Note that a NetApp login is required to view the report: http://nt-ap.com/2w6myr4

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

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 reducing latency. The A-Series lineup includes the A200, A300, A700, and A700s. These controllers and their specifications are listed in Table 1. For more information about the A-Series AFF controllers, see the A-Series product page and the NetApp Hardware Universe:

Scalable and fully featured, NetApp ONTAP data management software is ideal for your converged infrastructure. It allows you to consolidate all workloads onto a single system. All NetApp storage controllers are shipped with ONTAP installed and are ready to begin managing your data.

Table 1 NetApp A-Series controller specifications

	AFF A200	AFF A300	AFF A700	AFF A700s
NAS scale-out	2-8 nodes	2-24 nodes	2-24 nodes	2-24 nodes
SAN scale-out	2-8 nodes	2-12 nodes	2-12 nodes	2-12 nodes
Per HA Pair Specifications (Active-Active Dual Controller)				
Maximum SSDs	144	384	480	216
Maximum raw capacity	2.2PB	5.9PB	7.3PB	3.3PB
Effective capacity	8.8PB	23.8PB	29.7PB	13PB
Chassis form factor	2U chassis with 2 HA controllers and 24 SSD slots	3U chassis with two HA controllers	8U chassis with 2 HA controllers	4U chassis with 2 HA controllers and 24 SSD slots
ONTAP 9 Base Bundle	✓	✓	✓	✓
ONTAP 9 Premium Bundle (FlexClone®, SnapMirror®, SnapVault®, SnapCenter®, and more)	✓	✓	✓	✓

This solution uses the NetApp AFF A300, as seen in 10 and 11. This controller provides the high-performance benefits of 40GbE and all flash SSDs, offering better performance than comparable options, while taking up less space in the rack. Combined with a shelf of 3.8TB disks, this solution provides ample horsepower and over 90TB of capacity, all while taking up only 5U of valuable rack space. This makes it an ideal controller for a shared workload converged infrastructure. As an infrastructure's capacity or performance needs grow, the NetApp AFF A300 can increase capacity with additional storage shelves and performance by adding additional controllers to the cluster. A cluster can scale up to 24 nodes.

The 40GbE cards are installed in expansion slot 2. The ports are e2a and e2e.

Figure 10   NetApp A300 Storage Controller front view

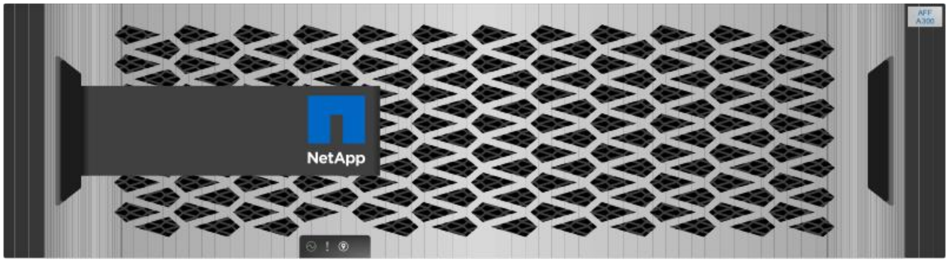
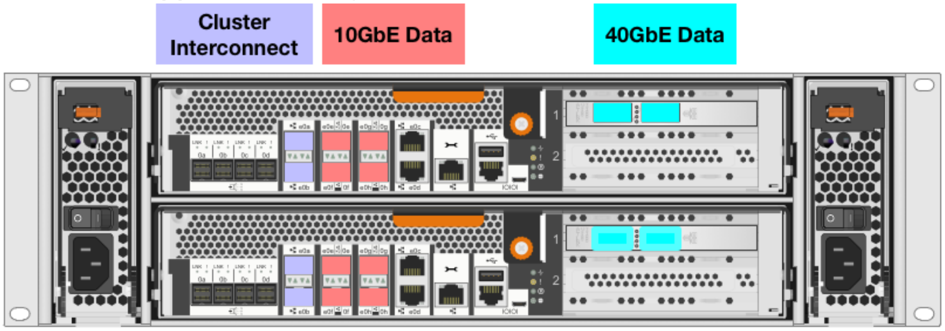


Figure 11   NetApp A300 Storage Controller Rear View



Cluster Storage and High Availability Pairs

FlexPod Datacenter architectures usually deploy NetApp FAS or AFF storage controllers running ONTAP data management software. Each controller, its storage, its network connectivity, and the instance of ONTAP running on that controller is referred to as a node.

Nodes are paired for high availability (HA). Together these pairs (up to 12 nodes for SAN and 24 nodes for NAS) comprise the cluster. Nodes communicate with each other over a private, dedicated cluster interconnects.

Nodes in an HA pair must use the same storage array model. However, a cluster can use any supported combination of controllers. You can scale out for capacity by adding nodes with like storage array models, or you can scale up for performance by adding nodes with higher-end storage arrays.

An internal HA interconnect allows each node to continually check whether its partner is functioning and to mirror log data for the other node’s non-volatile memory. When a write request is made to a node, it is logged in NVRAM on both nodes before a response is sent back to the client or host. On failover, the surviving partner commits the failed node's uncommitted write requests to disk, protecting data consistency.

Depending on the controller model, node storage consists of flash disks, capacity drives, or both. Network ports on the controller provide access to data. Physical storage and network connectivity resources are virtualized and only visible to cluster administrators. They are not visible to NAS clients or SAN hosts.

Connections to the other controller’s storage media allow each node to access the other's storage in the event of a takeover. Network path failover mechanisms make sure that clients and hosts continue to communicate with the surviving node.

You can scale up in all the traditional ways as well by upgrading disks or controllers as needed. ONTAP's virtualized storage infrastructure makes it easy to move data nondisruptively, meaning that you can scale vertically or horizontally without downtime.

NetApp All-Flash FAS A-Series Design

NetApp A-Series all-flash controllers were designed from the ground up with flash storage in mind. They provide industry leading density, scalability, and network connectivity, allowing customers to do more with their flash storage. The design of FlexPod Datacenter with Windows Server 2016 leverages the NetApp AFF A300 with 1.6TB SSDs. This controller provides a rich set of data management features as well as industry leading performance in a 2U form factor. 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 multistream writes to increase write performance to SSDs

The FlexPod Datacenter converged infrastructure supports a variety of NetApp FAS controllers, including the AFF A-Series, AFF8000, FAS9000, FAS8000, FAS2600 and FAS2500 platforms. For a full list of supported controllers, see the NetApp Interoperability Matrix Tool (IMT) and the FlexPod Technical Specifications.

Beginning with ONTAP 9.1, the X1134A adapter is supported on AFF A300 platforms. The X1134A is a 2-port 32Gb FC target-only adapter. NetApp ONTAP 9.1 also supports adding 40Gb Ethernet adapters as PCI expansion cards on AFF A300 systems. 40G Ethernet adapters are also a part of the FlexPod Datacenter with Microsoft Windows 2016 design.

For more information about the NetApp AFF A-series product family, see the A-Series product page.

NetApp ONTAP 9.1

In addition to the NetApp AFF A300, this design leverages ONTAP 9.1 data management software. ONTAP offers unified storage for applications that read and write data over block-access or file-access protocols in storage configurations that range from high-speed flash to lower-priced spinning media and cloud-based object storage.

ONTAP implementations run on NetApp-engineered Fabric-Attached Storage (FAS) or AFF appliances; on commodity hardware (ONTAP Select); and in private, public, or hybrid clouds (NetApp Private Storage and ONTAP Cloud). Specialized implementations offer best-in-class converged infrastructure (FlexPod Datacenter) and access to third-party storage arrays (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 can serve as the foundation for hybrid cloud and virtualization designs.

NetApp Storage Virtual Machines

A 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 can reside on any node in the cluster to which the SVM has been given access. An SVM may own resources on multiple nodes concurrently, and those resources can be moved nondisruptively from one node to another. For example, a flexible volume can be nondisruptively 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 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.

Because it is a secure entity, an SVM is only aware of the resources that are assigned to it, and it 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. If you have multiple SVMs, you can delegate an SVM to a specific application. This allows administrators of the application to access only the dedicated SVMs and associated storage, increasing manageability and reducing risk.

Server Message Block in ONTAP 9.1

Server Message Block (SMB) provides low cost, ease of deployment, ease of administration, and rich integration with Windows server and Active Directory. These features make it the default file-sharing protocol for Microsoft Hyper-V ecosystems. Features such as scale-out, transparent failover, persistent handles, and witness enable continuously available (CA) shares. ONTAP has supported SMB 3.0 by default since clustered Data ONTAP 8.2.1. This capability allows NetApp customers to use SMB 3.0 features introduced in Windows Server 2012 and allows Microsoft Hyper-V to use ONTAP volumes to host VM virtual disks and configuration settings.

Beginning with ONTAP 9, SMB 3.1.1 and enhanced features for SMB 2.0 and later are supported. The following enhancements are supported in SMB 2.0 and later:

- Workgroup authentication: You can now configure a CIFS server in a workgroup with CIFS clients that authenticate to the server by using locally defined users and groups.
- Large maximum transmission unit (MTU): Increased efficiency and throughput are now enabled by packet sizes up to 1MB. Previously, the maximum size was 64KB. Note: Large MTU values must be enabled through the CIFS server. Small packet sizes might result in performance degradation.
- Support for CIFS server in workgroup mode: Beginning with ONTAP 9, you can configure a CIFS server in a workgroup. This allows you to configure the CIFS server when the Microsoft Active Directory domain infrastructure is not available. A CIFS server in workgroup mode supports only Windows NT LAN Manager authentication and does not support Kerberos authentication. Certain CIFS

features are not supported by a CIFS server in workgroup mode. For more information about CIFS management, see the ONTAP 9 CIFS Reference.

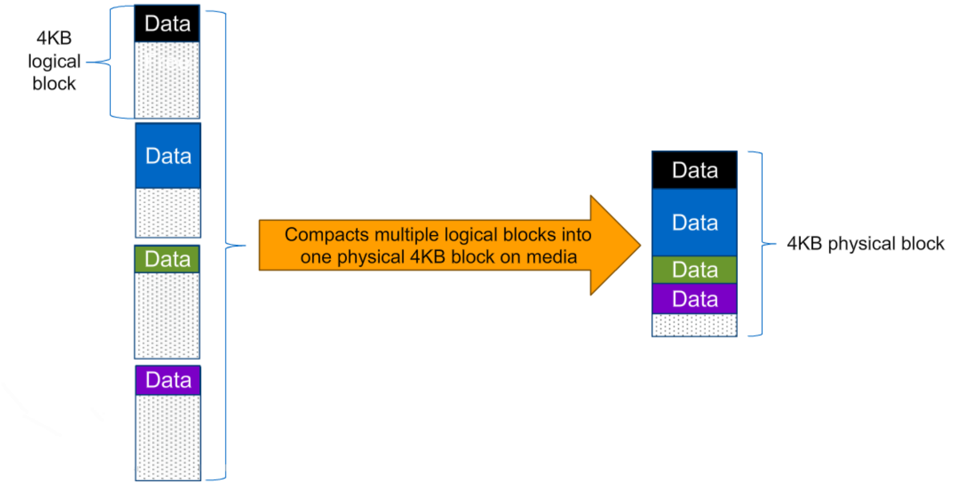
- Scale-Out. ONTAP is scale-out by design: A file share can be accessed over multiple nodes to provide better utilization of network bandwidth.
- Transparent failover: In case either partner of an HA pair experiences a failover (or upgrade event), the partner node takes over CA shares. Transparent failover allows SMB 3 clients to rapidly establish connections with CA shares on the partner node.
- Witness: ONTAP enables witness by correctly configuring SVM and CA with certain requirements. On a CA share, witness provides rapid and nondisruptive operations by notifying SMB 3 clients that a session has been lost and redirecting them to a different data LIF.
- Persistent handles: Persistent handles are enabled by default in ONTAP. Even though a new session is established on a new node, the lock state is preserved and operations are nondisruptive to the client.
- AES-128-CCM Encryption: ONTAP allows SMB encryption to be enabled at an SVM level or at the share level.

Storage Efficiency

Storage efficiency has always been a primary architectural design point of ONTAP. A wide array of features allows businesses to store more data using less space. In addition to deduplication and compression, businesses can store their data more efficiently by using features such as unified storage, multi-tenancy, and NetApp Snapshot® technology. Storage efficiency features available with ONTAP 9 include the following:

- Thin provisioning: For a thin-provisioned volume or LUN, storage is not reserved in advance. Rather, storage is allocated dynamically as it is needed. Free space is released back to the storage system when data in the volume or LUN is deleted.
- Deduplication: Deduplication reduces the amount of physical storage required for a volume (or all the volumes in an AFF aggregate) by discarding duplicate blocks and replacing them with references to a single shared block. Reads of deduplicated data typically incur no performance charge. Writes incur a negligible charge except on overloaded nodes.
- Compression: Compression reduces the amount of physical storage required for a volume by combining data blocks in compression groups, each of which is stored as a single block. Reads of compressed data are faster than in traditional compression methods because ONTAP decompresses only the compression groups that contain the requested data, not an entire file or LUN.
- Compaction: Compaction (introduced in ONTAP 9) is the latest patented storage efficiency technology released by NetApp. In the ONTAP WAFL (Write Anywhere File Layout) 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-disk to save space. See Figure 12 for an illustration of compaction.
- FlexClone volumes, files, and LUNs: FlexClone technology references Snapshot metadata to create writable, point-in-time copies of a volume. Copies share data blocks with their parents, consuming no storage except what is required for metadata until changes are written to the copy. FlexClone files and FlexClone LUNs use identical technology, except that a backing Snapshot copy is not required.

Figure 12   Compaction in NetApp ONTAP 9


4KB logical block
Data
Data
Data
Data
Compacts multiple logical blocks into one physical 4KB block on media
Data
Data
Data
Data
4KB physical block

NetApp Volume Encryption

Data security continues to be an important consideration for customers purchasing storage systems. NetApp has supported self-encrypting drives in storage clusters prior to ONTAP 9. In ONTAP 9 however, the encryption capabilities of ONTAP are extended by adding an Onboard Key Manager (OKM). The OKM generates and stores the 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 disks is secure and can only be accessed by ONTAP.

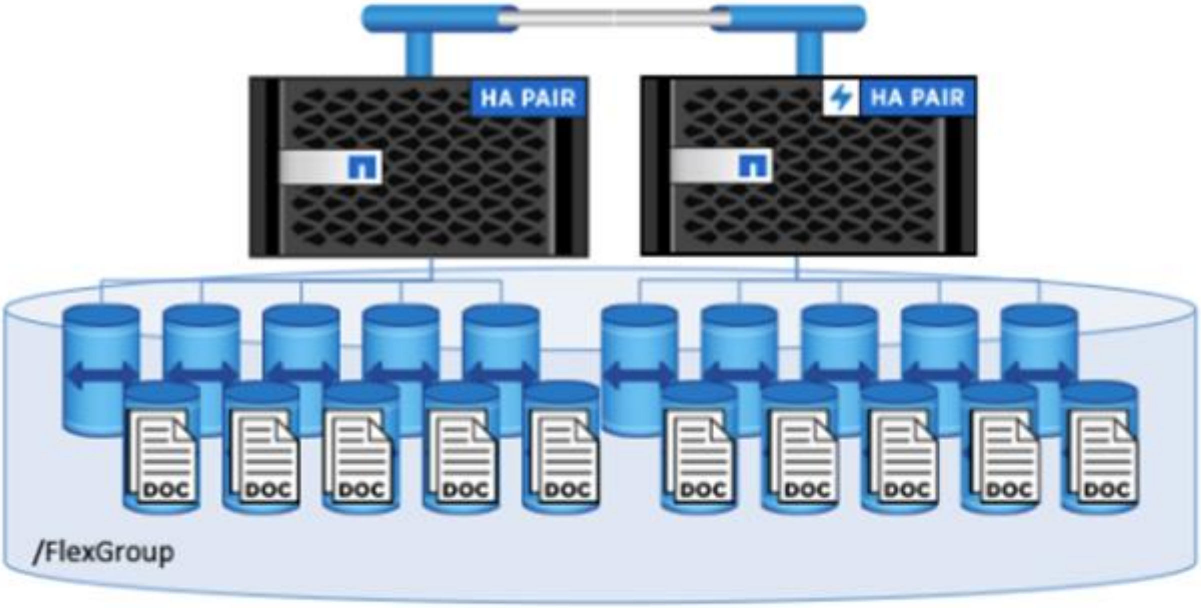
In ONTAP 9.1, NetApp extends encryption capabilities further with NetApp Volume Encryption (NVE). NVE is a software-based mechanism for encrypting data. It allows a user to encrypt data at the volume level instead of requiring encryption of all data in the cluster, thereby providing more flexibility and granularity to ONTAP administrators. After it is enabled, this encryption extends to Snapshot copies and FlexClone volumes created in the cluster. To preserve all storage efficiency benefits provided by ONTAP, NetApp Volume Encryption executes after all storage efficiency features. This makes sure that customers have secure encrypted data while enjoying the benefits of reduced capacity utilization in the cluster.

For more information about encryption in ONTAP 9.1, see the NetApp Encryption Power Guide.

NetApp FlexGroup

Beginning in ONTAP 9.1, NetApp introduced the FlexGroup, a scale-out NAS container that provides high performance, automatic load distribution, and scalability. A FlexGroup volume contains several FlexVol volumes that automatically and transparently share traffic in the cluster.

Figure 13 NetApp FlexGroup Volume



Files in a FlexGroup volume are allocated to individual member volumes and are not striped across volumes or nodes. When a client adds files and subdirectories to a FlexGroup volume, ONTAP automatically determines the best FlexVol member to use for storing each new file and subdirectory. The FlexGroup volume attempts to organize the files, both for fastest accesses and for better throughput performance.

Advantages of the FlexGroup include the following:

- Massive capacity: FlexGroup volumes can scale up to multiple petabytes with high file counts (hundreds of billions of files). The only limiting factors are the physical limits of the hardware and the total volume limits of ONTAP. For example, a 10-node cluster can have a 20PB FlexGroup volume that can handle 400 billion files.
- Predictable low latency for high-metadata workloads: A FlexGroup volume utilizes all the cluster resources. For example, it can use multiple aggregates, nodes, CPU cores, and other hardware assets. This capability enables multiple volume affinities for a single storage container for metadata-intensive workloads.
- Ease of management: A FlexGroup volume can provision storage across every node and aggregate in a cluster (without any junction path or capacity management overhead) through the FlexGroup tab in NetApp OnCommand® System Manager.

Backup and Replication

Traditionally, ONTAP replication technologies served the need for disaster recovery (DR) and data archiving. With the advent of cloud services, ONTAP replication has been adapted to data transfer between endpoints in the NetApp Data Fabric. The foundation for all these uses is ONTAP Snapshot technology:

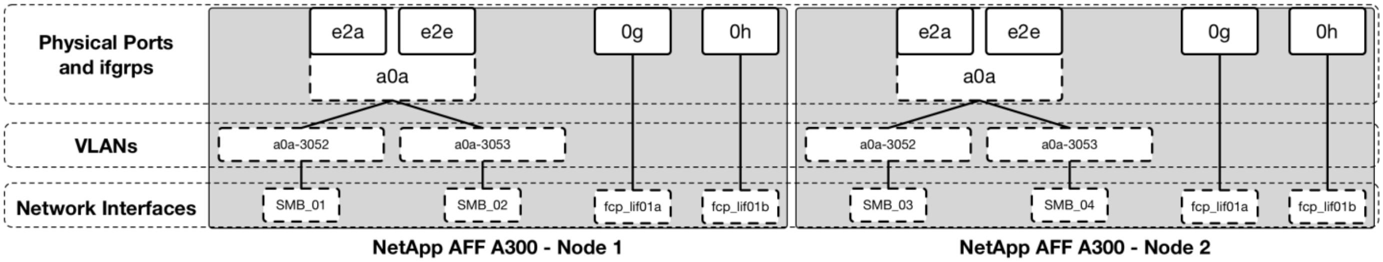
- NetApp Snapshot copies: A Snapshot copy is a read-only, point-in-time image of a volume. The image consumes minimal storage space and incurs negligible performance overhead because it records only the changes to the active file system since the last Snapshot copy. A volume can contain up to 255 Snapshot copies.
- NetApp SnapMirror disaster recovery and data transfer: SnapMirror is a disaster recovery technology designed for failover from primary storage to secondary storage at a geographically remote site. As its name implies, SnapMirror creates a replica, or mirror, of your working data in secondary storage from which you can continue to serve data in the event of a catastrophe at the primary site.
- NetApp SnapVault archiving: SnapVault is an archiving technology designed for disk-to-disk Snapshot copy replication for standards compliance and other governance-related purposes. In contrast to a SnapMirror relationship, in which the destination volume contains only the Snapshot copies currently in the source volume, a SnapVault destination volume typically retains point-in-time Snapshot copies created over a much longer period.
- NetApp MetroCluster® continuous availability: MetroCluster configurations protect data by implementing two physically separate, mirrored clusters. Each cluster synchronously replicates the data and SVM configuration of the other. In the event of a disaster at one site, an administrator can activate the mirrored SVM and begin serving data from the surviving site.

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 All Flash FAS storage system and provides better performance. In this design, SAN boot was validated with Fibre Channel (FC), iSCSI, and FCoE. Please see the SAN Architecture section for cabling differences between the different validations.

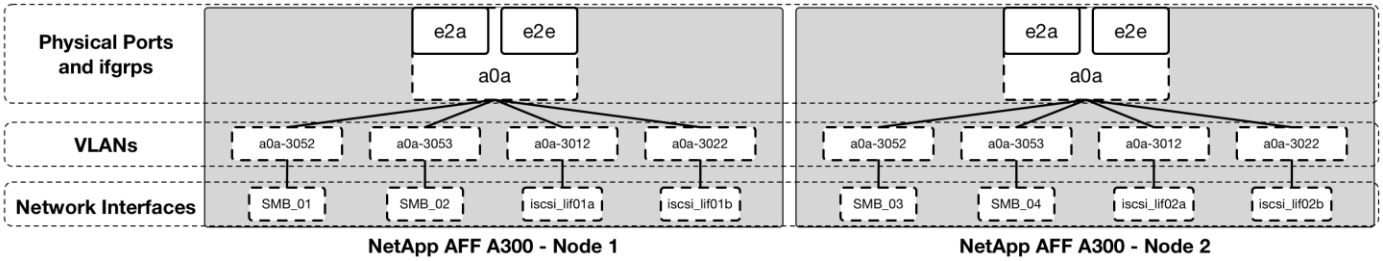
In the FC validation, 4 FC network interfaces were created, providing each Windows server 2 active optimized paths and 2 active unoptimized paths to its boot LUN. The Cisco MDS switch was used for FC traffic. SMB file share access was provided by creating an interface group on the two 40GbE ports, creating VLANs on the interface groups, and creating the network interfaces on the VLANs. This network interface and port layout can be seen in Figure 14. The networking layout from the storage point of view is the same for both FC and FCoE configurations. The differences in the designs being the direct connection to the Fabric Interconnect instead of a connection to the Cisco MDS switch.

Figure 14   Fibre Channel SAN Boot



In the iSCSI validation, the interface groups are used for all traffic. VLAN interfaces for each iSCSI VLAN are created. Network interfaces are created on each VLAN interface. Cisco UCS service profiles are created with two iSCSI vNICs - one for each iSCSI VLAN. SMB file share access was provided by creating an interface group on the two 40GbE ports, creating VLANs on the interface groups, and creating the network interfaces on the VLANs. This network interface and port layout can be seen in Figure 15.

Figure 15 iSCSI SAN Boot



In all SAN boot configurations mentioned above, Multipath I/O (MPIO) in combination with a Device-Specific Module (DSM) is configured in Windows to establish multiple connections to boot LUN on the NetApp controllers. MPIO software is required in any configuration which provides more than a single path to the LUN. Because there are multiple paths to the LUN, SAN network interfaces are not configured to fail over to an available port. Instead, the network interface will be disabled and the host chooses a new optimized path to a different network interface. Asymmetric Logical Unit Access (ALUA) is an industry standard protocol for identifying optimized paths between a storage system and a host. ALUA enables the Windows initiator to query the target about path attributes, such as primary and secondary paths. ALUA can be enabled on an ONTAP igroup and must be enabled for all configurations described in this guide.

SMB vs iSCSI for Virtual Machine Disks

Microsoft Hyper-V allows you to store VM files, including configuration, virtual hard disk files, and snapshots, on SMB file shares or a SAN LUN. In this design, both options were validated to provide adequate performance and resiliency. Although there are reasons you might want to use a SAN LUN for VM files and it may be required for certain applications, the preferred method in this design was using the SMB 3.0 file share. The capabilities in ONTAP 9.1 support SMB 3.0 features like scale-out, transparent failover, and persistent handles and writes, enabling the file share to serve as a CA share.

Advantages of using SMB file shares for VM files include the following:

- Ease of provisioning and management: CA shares can be easily provisioned using Virtual Machine Manager (through the NetApp Storage Management Initiative Specification [SMI-S] provider) with appropriate authorization and authentication rules.
- High Availability: SMB shares also enable simpler management of storage resources in a failover cluster without having to create a Clustered Shared Volume (CSV) to have simultaneous read-write access to the same LUN (disk).
- Swift VM migration: Since VM disk and configuration is stored on the same storage accessible by multiple Hyper-V hosts, only the VM information must be moved among Hyper-V hosts.
- Reduced expenditures: Multiprotocol support for storage access in ONTAP and rich integration with Microsoft SCVVMM result in reduced capital and operating expenses.

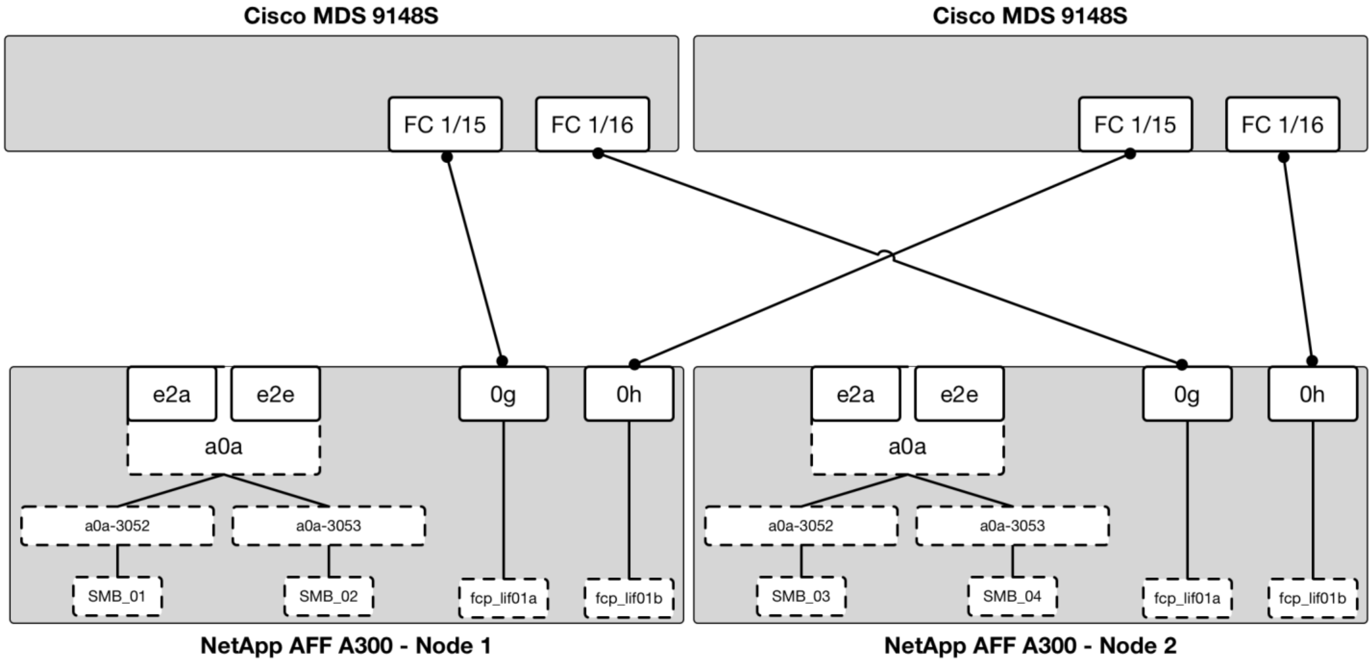
FC and iSCSI: Network and Storage Connectivity

In the FC design, the storage controller is connected to a Cisco MDS SAN switching infrastructure for FC boot and a Cisco Nexus 9332PQ for Ethernet traffic. Smaller configurations can use the directly attached FC/FCoE storage method of the Cisco UCS fabric interconnect. However, the Cisco MDS provides increased scalability for larger configurations with multiple Cisco UCS domains and the ability to connect to an enterprise SAN such as the NetApp AFF A-Series.

An ONTAP storage controller uses N_Port ID virtualization (NPIV) so that each network interface can log into the FC fabric using a separate worldwide port name (WWPN). This allows a host to communicate with an FC target network interface, regardless of where that network interface is placed. Use the show npiv status command on your MDS switch to make sure that NPIV is enabled.

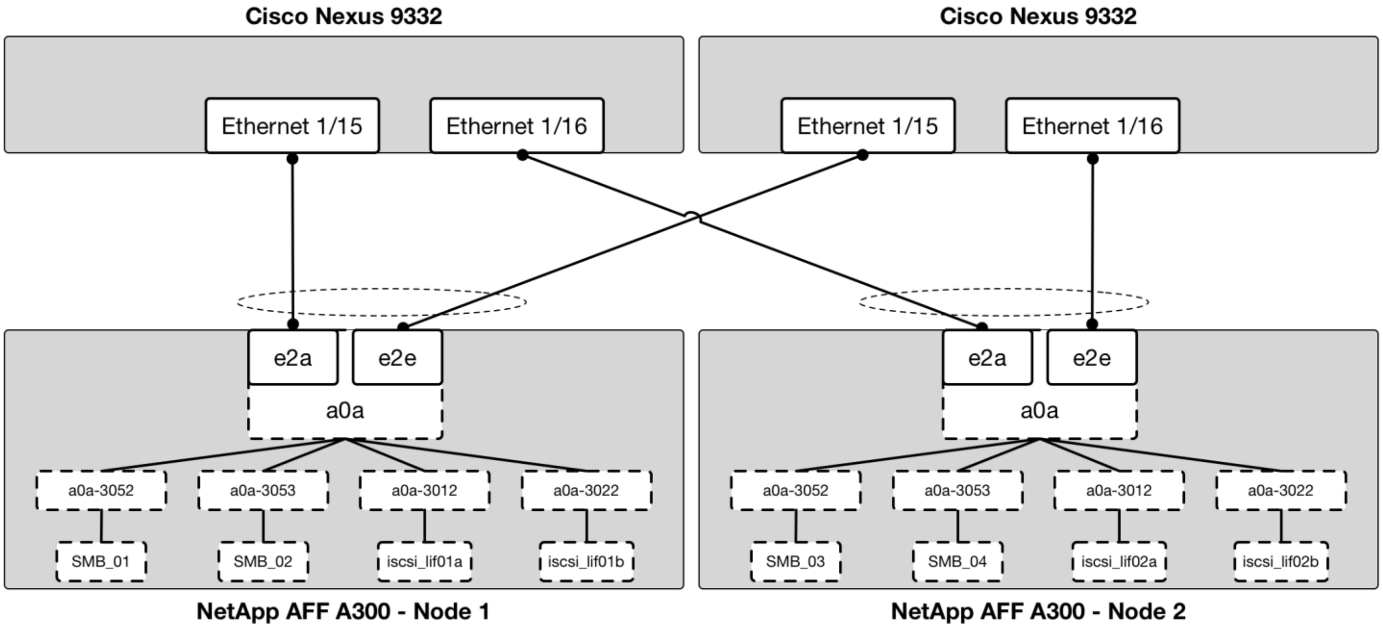
Each controller is equipped with onboard CNA ports that can operate in 10GbE or 16G FC. For the FC design, the onboard ports are modified to FC target mode and the ports are connected to two SAN fabrics. This method provides increased redundancy to make sure that the paths from the host to its storage LUNs are always available. Figure 16Error! Reference source not found. shows a connection diagram of port and interface assignments for the AFF storage to the Cisco MDS SAN fabrics. This FlexPod design uses the following port and interface assignments. In this design, iSCSI uses 40Gb and FC uses 16Gb bandwidth.

Figure 16 FC networking



In the iSCSI design, the storage controller's 40GbE ports are directly connected to the Cisco Nexus 9332PQ switches. Each controller is equipped with 40GbE cards, which has two physical ports each, in the expansion slot 2. Each storage controller is connected to two SAN fabrics. This method provides increased redundancy to make sure that the paths from the host to its storage LUNs are always available. Figure 17 shows a connections diagram of port and interface assignments for the AFF storage to the Cisco Nexus 9332PQ SAN fabrics. This FlexPod design uses the port and interface assignments as shown in the figure below. In this design iSCSI traffic utilize the 40Gb bandwidth.

Figure 17 iSCSI networking



Cisco Nexus

Cisco Nexus series switches provides an Ethernet switching fabric for communications between the Cisco UCS, NetApp storage controllers, and the rest of a 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 Cisco Nexus 9000 series switches are deployed in leaf-spine 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, refer to http://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/index.html

This FlexPod design deploys a single pair of Nexus 9000 top-of-rack switches at each site in traditional deployment mode running NX-OS, essentially creating individual FlexPods at both locations.

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 PortChannels (vPC)

Virtual Port Channel (vPC)

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

- Allows a single device to use a PortChannel 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 18 Cisco Nexus 9000 Connections

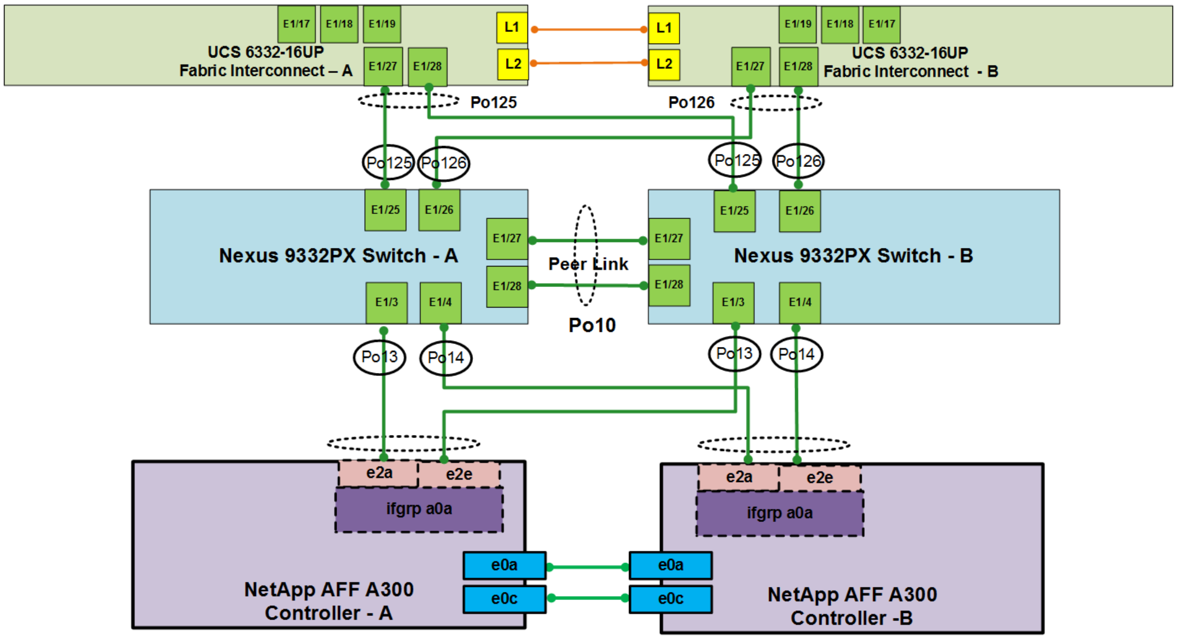


Figure 18 shows the connections between Cisco Nexus 9000 switches, Cisco UCS Fabric Interconnects and NetApp AFF A300 storage controllers. vPC requires a “peer link” which is shown as port channel 10 in this diagram. In addition to the vPC peer-link, the vPC peer keepalive link is a required component 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.

Cisco Nexus 9000 Best Practices

Nexus 9000 related best practices used in the validation of the FlexPod architecture are summarized below:

Cisco Nexus 9000 features enabled

- Link Aggregation Control Protocol (LACP part of 802.3ad)
- Cisco Virtual Port Channeling (vPC) 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 controller and UCS are defined as “edge” trunk ports

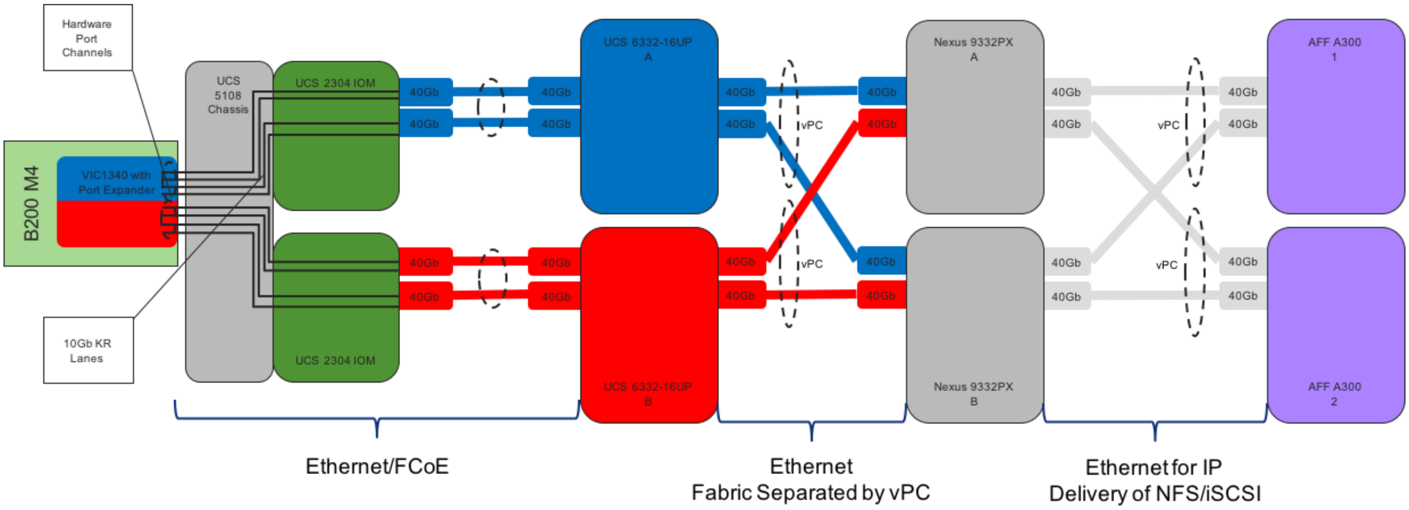
For configuration details, refer to the Cisco Nexus 9000 Series Switches Configuration guides: http://www.cisco.com/c/en/us/support/switches/nexus-9000-series-switches/products-installation-and-configuration-guides-list.html

Bringing Together 40Gb End to End

The 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 UCS 6332-16UP Fabric Interconnects. Passage of this traffic shown in Figure 19 going from right to left is as follows:

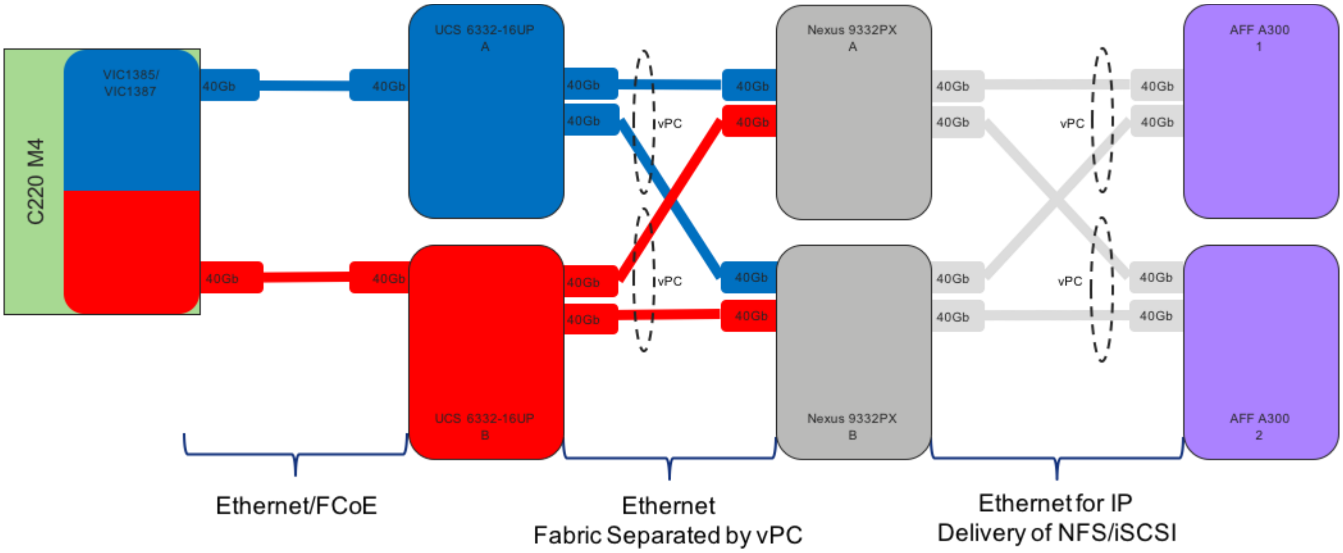
- vPC representation of the pair of Nexus switches to each AFF A300 Controller.
- Continuing onto the UCS 6332-16UP Fabric Interconnects with equivalent vPCs to each fabric side.
- Connecting from each Fabric Interconnect to the UCS 2304 IOM (Fabric Extender) with pairs of 40Gb uplinks automatically configured as port channels during chassis association.
- Pathing through 10Gb KR lanes from the IOM into the UCS 5108 Chassis backplane.
- Coming into the VIC 1340 adapter of the UCS B200 M4 server that is increased to 40Gb for each side with the addition of a Port Expander card.

Figure 19 vPC, AFF A300 Controllers, and Cisco UCS 6332-16UP Fabric Interconnect Traffic



The equivalent view for a UCS C-Series server is shown in Figure 20 which uses the same primary connectivity provided by the Nexus switches, but will not go through UCS 2304 IOM and UCS 5108 Chassis:

Figure 20   Cisco UCS C-Series Server



Cisco MDS

The Cisco® MDS 9148S 16G Multilayer Fabric Switch is the next generation of the highly reliable, flexible, and low-cost Cisco MDS 9100 Series switches. It combines high performance with exceptional flexibility and cost effectiveness. This powerful, compact one rack-unit (1RU) switch scales from 12 to 48 line-rate, 16 Gbps Fibre Channel ports. The Cisco MDS 9148S delivers advanced storage networking features and functions with ease of management and compatibility with the entire Cisco MDS 9000 Family portfolio for reliable end-to-end connectivity.

For more information on the MDS 9148S please see the product data sheet at: http://www.cisco.com/c/en/us/products/collateral/storage-networking/mds-9148s-16g-multilayer-fabric-switch/datasheet-c78-731523.html

MDS Insertion into FlexPod

The MDS 9148S is inserted into the FlexPod design to provide Fibre Channel switching between the NetApp AFF A300 controllers, and UCS managed B-Series and C-Series servers connected to the UCS 6332-16UP and UCS 6248UP Fabric Interconnects. Adding the MDS to the FlexPod infrastructure allows for:

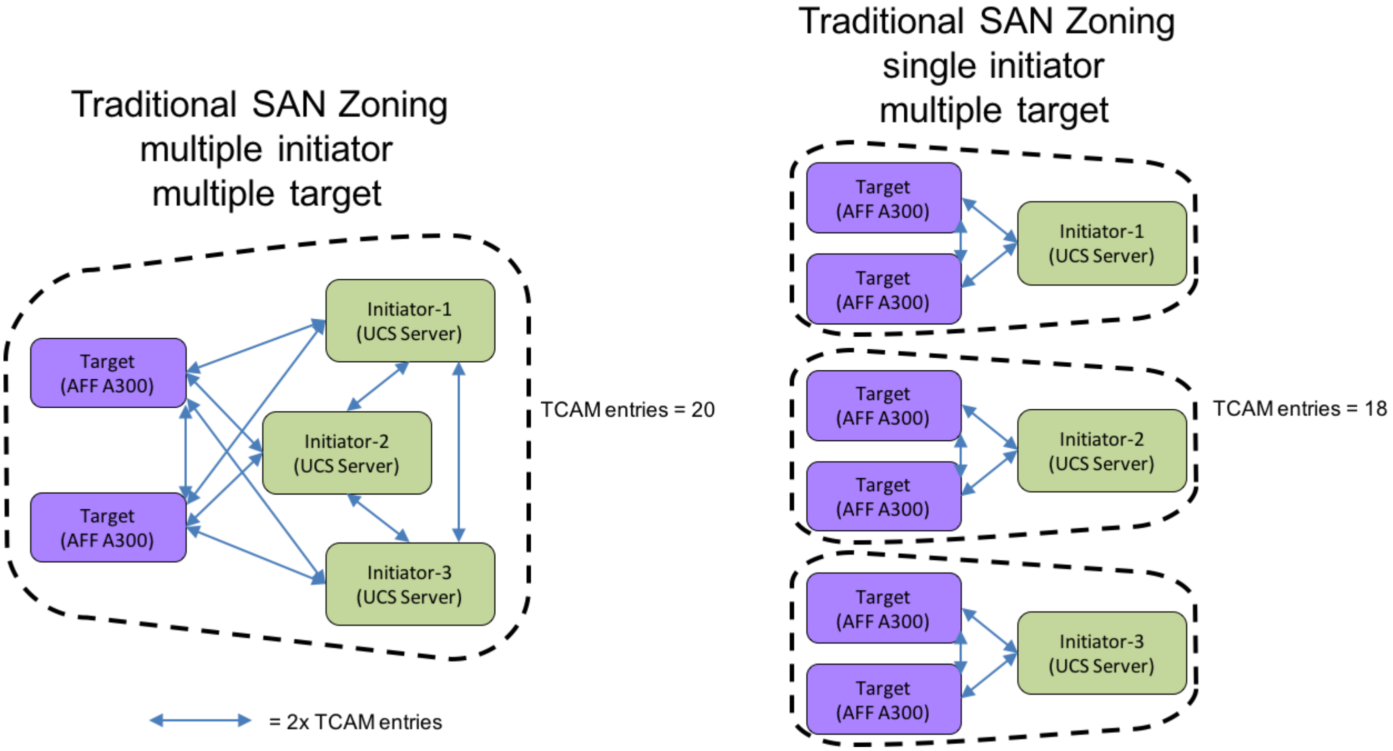
- Increased scaling of both the NetApp storage and the Cisco UCS compute resources
- Large range of existing models supported from the 9100, 9200, 9300, and 9700 product lines
- A dedicated network for storage traffic
- Increased tenant separation
- Deployments utilizing existing qualified SAN switches that might be within the customer environment

FC and FCoE direct attached (the latter covered within the appendix of the FlexPod FC Deployment Guide associated with this Design Guide) are configurable for FlexPods, using the UCS Fabric Interconnects for the SAN switching. This model will not require an MDS, but will have reduced scaling capacity, and have limited options for extending SAN resources outside of the FlexPod to elsewhere in the customer data center.

Smart Zoning with MDS

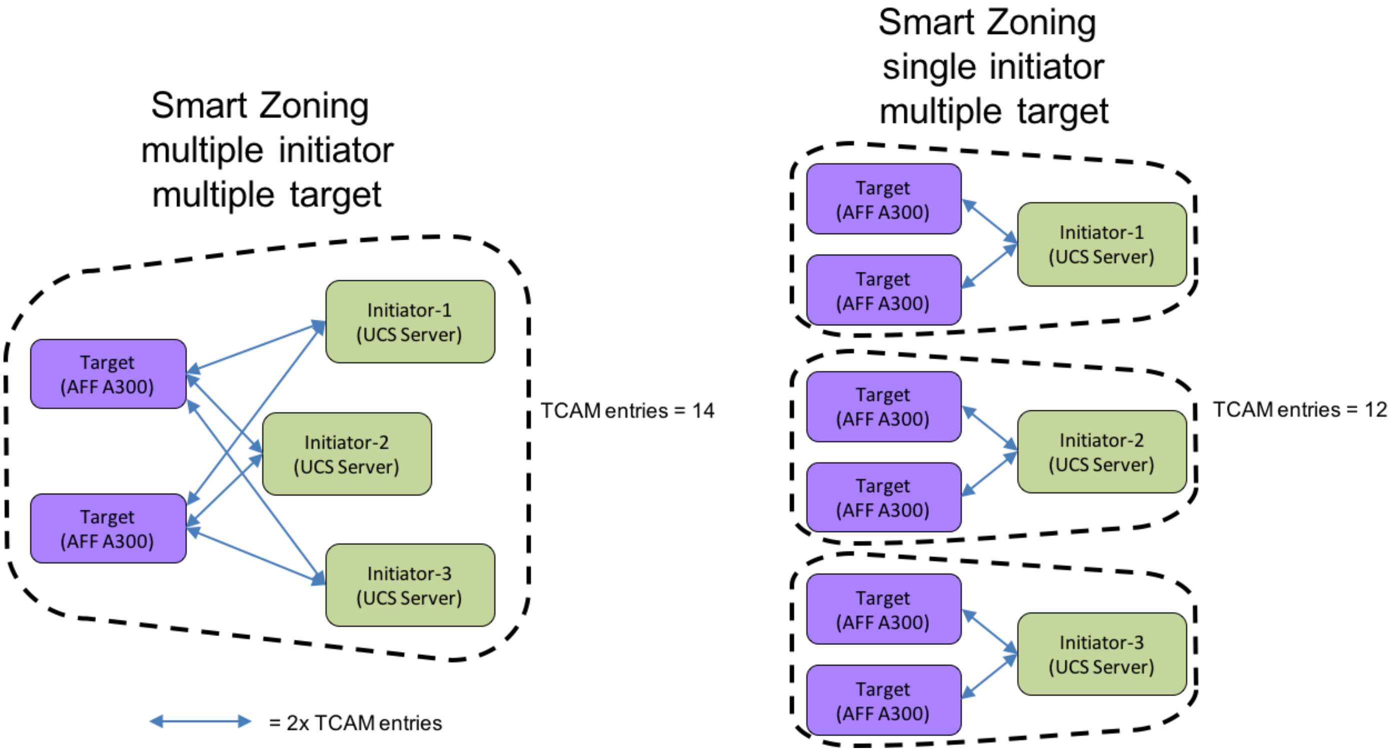
Configuration of the Cisco MDS within this FlexPod design takes advantage of Smart Zoning to increase zoning efficiency within the MDS. Smart Zoning allows for reduced TCAM (ternary content addressable memory) entries, which are fabric ACL entries of the MDS allowing traffic between targets and initiators. When calculating TCAMs used, two TCAM entries will be created for each connection of devices within the zone. Without Smart Zoning enabled for a zone, targets will have a pair of TCAMs established between each other, and all initiators will additionally have a pair of TCAMs established to other initiators in the zone as shown in Figure 21 below:

Figure 21 Traditional Zoning with MDS

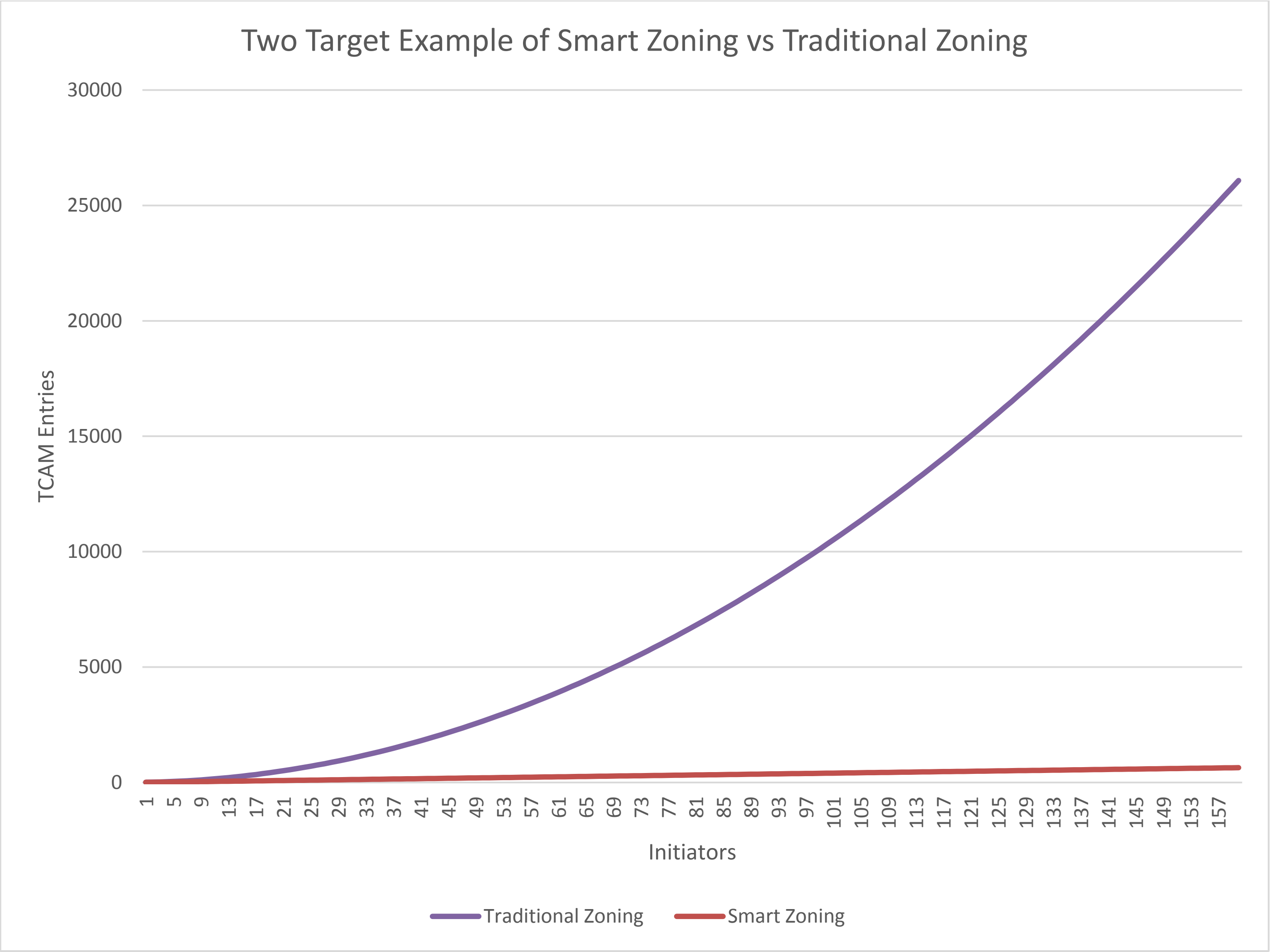


Using Smart Zoning, Targets and Initiators are identified, reducing TCAMs needed to only occur Target to Initiator within the zone as shown in Figure 22 below:

Figure 22 Using Smart Zoning



Within the traditional zoning model for a multiple initiator, multiple target zone shown in the first diagram set of figure 21, the TCAM entries will grow rapidly, representing a relationship of TCAMs = (T+I)*(T+I)-1) where T = targets and I = initiators. For Smart Zoning configuration, this same multiple initiator, multiple target zone shown in the second diagram set of figure 22 will instead have TCAMs = 2*T*I where T = targets and I = initiators. This exponential difference can be seen for the traditional zoning on the graph below showing the example maximum 160 initiators of servers configurable to a single UCS zone connecting to 2 targets represented by a pair of NetApp controllers.



For more information on Smart Zoning, see: http://www.cisco.com/c/en/us/td/docs/switches/datacenter/mds9000/sw/7_3/configuration/fabric/fabric/zone.html

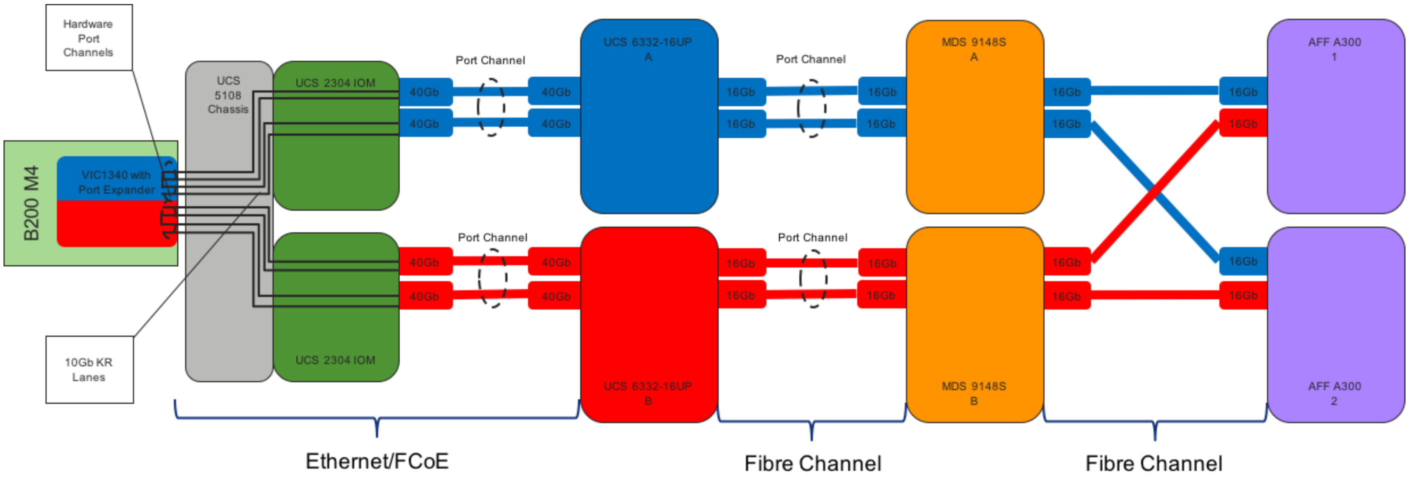
End to End design with MDS

The End to End storage network design with the MDS will look similar to the 40Gb End to End design seen with the Nexus 9300 switches. For the UCS 6332-16UP based implementation connecting to a UCS B-Series server, the view of the design shown in Figure 23 can be seen starting from the right as:

- Each NetApp AFF A300 controller will uplink into each MDS switch with a 16Gb Fibre Channel adapter
- The MDS will extend a port channel of two 16Gb uplinks from each fabric side of the fabric to its corresponding UCS 6332-16UP Fabric Interconnect
- Leaving the Fabric Interconnect through the 40Gb uplinks connecting to the UCS 2304 IOM (Fabric Extender) of the UCS 5108 Chassis, will have the FC traffic encapsulated as FCoE and sent along the common fabric with the network traffic going to the servers.

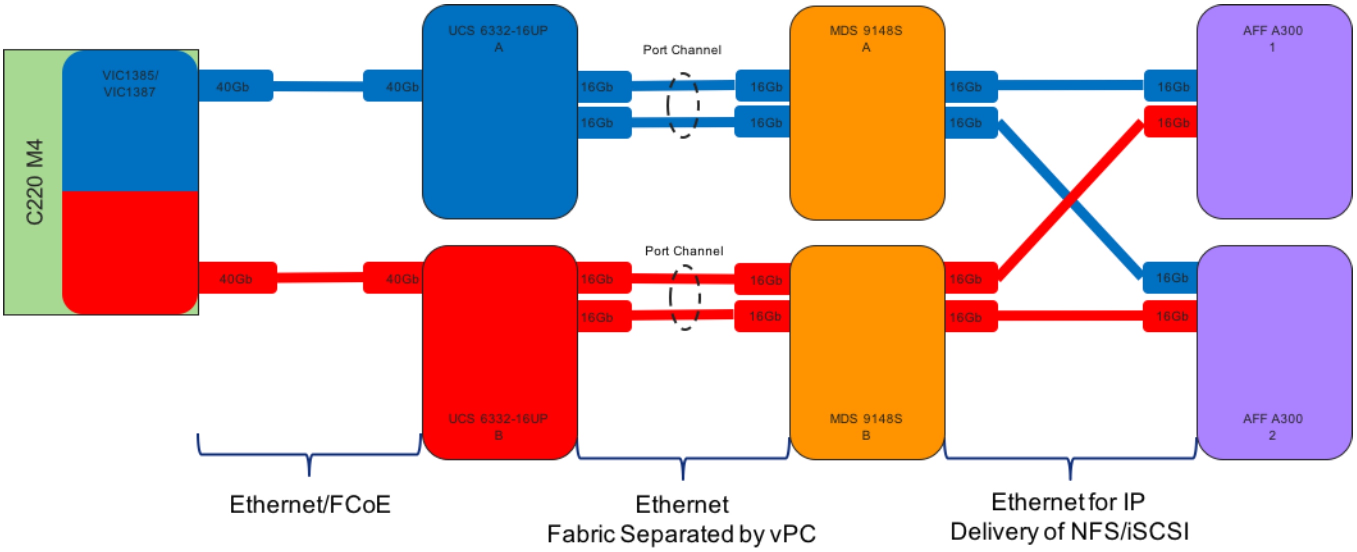
- Entering the UCS 5108 from the UCS 2304 IOM, the traffic will be sent through the 10Gb KR lanes of the chassis into the VIC 1340 Adapter with Port Expander.

Figure 23 End-to-End Design with MDS



The equivalent view for a UCS C-Series server is shown in Figure 24 below, which uses the same primary connectivity provided by the MDS Switches, but will not go through UCS 2304 IOM and UCS 5108 Chassis:

Figure 24 End-to-End Design with Cisco UCS C-Series Server



Cisco Unified Computing System

Cisco UCS 6300 and UCS 6200 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. 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 the UCS system. 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.

Table 2 The key differences between FI 6200 series and FI 6300 series

	FI 6200 Series		FI 6300 Series	
Features	6248	6296	6332	6332-16UP
Max 10G ports	48	96	96* + 2**	72* + 16
Max 40G ports	-	-	32	24
Max unified ports	48	96	-	16
Max FC ports	48 x 2/4/8G FC	96 x 2/4/8G FC	-	16 x 4/8/16G FC

** Using 40G to 4x10G breakout cables*

*** Requires QSA module*

Cisco UCS Differentiators

Cisco's Unified Compute System is revolutionizing the way servers are managed in data-center. Following are the unique differentiators of Cisco Unified Computing System and Cisco UCS Manager.

- Embedded Management – In Cisco UCS 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 – reducing capital and operational expenses of 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 UCS, where compute capability of UCS can be extended easily while keeping the existing external connectivity to LAN, SAN and management networks.

- Policy Based Resource Classification – Once a compute resource is discovered by 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 UCS Manager.

- Combined Rack and Blade Server Management – Cisco UCS Manager can manage Cisco UCS 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 UCS Manager inherently friendly to multi-tenant environment typically observed in private and public clouds.

- Extended Memory – The enterprise-class Cisco UCS B200 M4 blade server extends the capabilities of Cisco's Unified Computing System portfolio in a half-width blade form factor. The Cisco UCS B200 M4 harnesses the power of the latest Intel® Xeon® E5-2600 v4 Series processor family CPUs with up to 1536 GB of RAM (using 64 GB DIMMs) - allowing huge VM to physical server ratio required in many deployments, or allowing large memory operations required by certain architectures like big data.

- Virtualization Aware Network – Cisco 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 Fibre 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.

UCS Design Options within this FlexPod

Cisco UCS vNICs

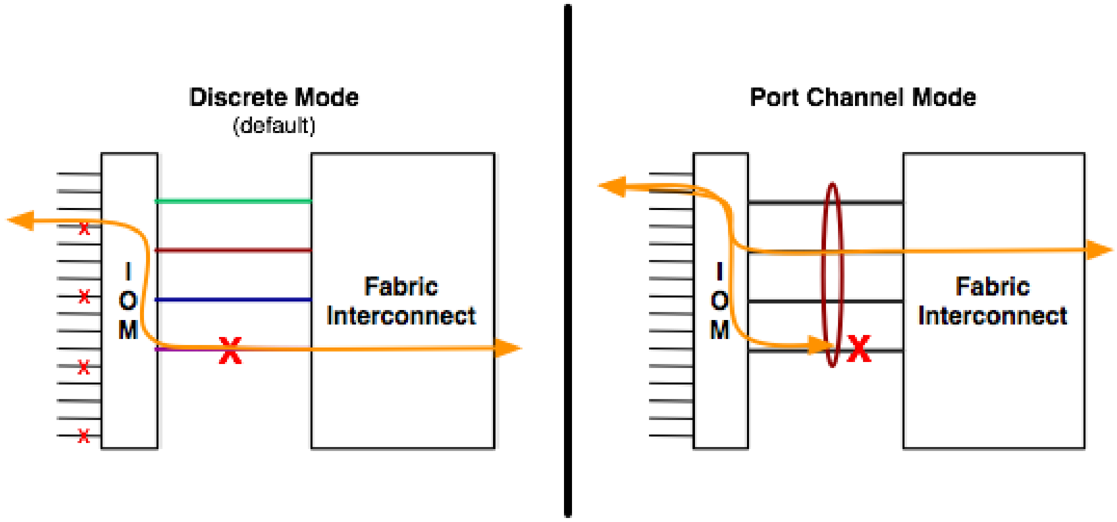
The FC and FCoE FlexPod architectures use just two Cisco UCS vNICs. One vNIC pinned to Fabric-A and the other pinned to Fabric-B to recover in case of a path failure. These two vNICs provisioned to the servers are teamed in the OS for failover and load balancing using the new "Switch Embedded Teaming" feature introduced in Windows Server 2016 Hyper-V. Note that, Cisco UCS fabric failover is not enabled on the vNICs.

In addition to the two vNICs described above, the iSCSI FlexPod architecture requires two more additional vNICs for iSCSI traffic. Each iSCSI vNIC is pinned to a separate fabric path in Cisco UCS to recover from path failure. As per Microsoft's best practice, the iSCSI vNICs are not teamed. Multipathing feature in the host takes care of failover and load balancing. Note that, Cisco UCS fabric failover is not enabled on the vNICs.

Cisco Unified Computing System Chassis/FEX Discovery Policy

A UCS system can be configured to discover a chassis using Discrete Mode or the Port-Channel mode (Figure 25). 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 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 25 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, etc.) 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.

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 UCS Physical Connectivity

Cisco UCS Fabric Interconnects are configured with two port-channels, one from each FI, to both Cisco Nexus 9000s and 16Gbps to the corresponding MDS 9148s switch. The Fibre Channel connections carry the Fibre Channel boot and data LUNs from the A and B fabrics to each of the MDS switches which then connect to the storage controllers. The port-channels carry the remaining data and storage traffic originated on the Cisco Unified Computing 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.

Cisco Unified Computing System – C-Series Server Design

Cisco UCS Manager 3.1 provides two connectivity modes for Cisco UCS C-Series Rack-Mount Server management. Cisco UCS Manager release version 2.2 introduces an additional rack server management mode using Network Controller Sideband Interface (NC-SI). Cisco UCS VIC 1385 Virtual Interface Card (VIC) uses the NC-SI, which can carry both data traffic and management traffic on the same cable. Single-wire management allows for denser server to FEX deployments.

For configuration details refer to the Cisco UCS configuration guides at: http://www.cisco.com/c/en/us/support/servers-unified-computing/ucs-manager/products-installation-and-configuration-guides-list.html

Windows Server 2016

Windows Server 2016 is the latest version of Microsoft server operating system. It is a cloud-ready server operating system designed to run both traditional and cloud-native workloads whether on-premises, in a hybrid environment, or in any public environment equally well using the new capabilities such as Windows Server Containers and the lightweight Nano Server.

Windows Server 2016 comes in three editions - Datacenter, Standard and Essentials. Datacenter edition comes with unlimited virtualization along with new features including Shielded Virtual Machines, Software-Defined Storage and Software-Defined Networking. Standard edition comes with limited virtualization and the Essentials edition is for small-to-medium sized customers with upto 50 users.

Hyper-V is Microsoft's virtualization product. The Hyper-V server role in Windows Server enables you to create a virtualized server computing environment where you can create and manage virtual machines. It allows you to install and run multiple operating systems independently and in isolation on one physical computer.

Below table compares the maximum configuration for the various components that apply to Windows Server 2012 R2 and Windows Server 2016:

Table 3 Windows Server 2012 R2 and Windows Server 2016 Maximums Comparison

System	Components	Windows Server 2012 R2	Windows Server 2016
Physical Server (Host)	Memory per host	4 TB	24 TB
	Logical processors per host	320	512
	Virtual processors per host	2048	2048
	Virtual machines per host	1024	1024
Virtual Machine	Memory support per VM	1 TB	12 TB
	Virtual processors per VM	64	240
	Virtual disk capacity	64 TB	64 TB
	Virtual SCSI disks	256	256
Cluster	Nodes per cluster	64	64
	VMs per cluster	8000	8000

This solution is limited to the scope of Windows Server 2016 Hyper-V.

New features for Windows Server 2016 Hyper-V include:

- Hyper-V is a supported role on Nano Server
- Windows containers
- Shielded Virtual Machines (see Security section of this document)
- Virtualization Based Security
- Virtual Machine Resiliency
- Production checkpoints
- Cluster OS Rolling Upgrade for Hyper-V clusters
- Storage Quality of Service (QoS)
- PowerShell Direct
- Compatible with Connected Standby
- Discrete device assignment
- Hot add and remove for network adapters
- Hot add and remove for fixed memory
- Hyper-V Manager improvements
- Integration services delivered through Windows Update
- Linux Secure Boot
- Nested virtualization
- Networking features
- Updated virtual machine file formats
- Allow running down-level virtual machines

Management and Best Practices with Windows 2016 Hyper-V

The Microsoft ecosystem of products and third party solutions offers a plethora of options for management, orchestration, monitoring, and back-ups. This section describes some of those choices and best-practice considerations.

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.

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 Microsoft Windows Server 2008, Windows Server 2008 R2, Windows Server 2012, Windows Server 2012 R2, and Windows Server 2016. It offers active/active and active/passive load balance policies for both the FC and iSCSI protocols. ALUA must be enabled on the storage system igroup for FC.

NetApp SnapDrive for Windows

With NetApp SnapDrive for Windows, you can automate storage provisioning tasks and manage data in Microsoft Windows environments. You can run SnapDrive on Windows hosts in either a physical or a virtual environment.

Several components are integrated into the SnapDrive software and are automatically installed. These components allow you to manage LUNs, Windows volumes, or SMB shares. You can use these components together to create SnapDrive workflows, including provisioning, Snapshot copy management, backup, restore, and mounting operations.

The following SnapDrive components are integrated in the software and are automatically installed during installation:

- SnapDrive "snap-in"
- SnapDrive command-line interface
- PowerShell cmdlets
- Underlying SnapDrive service
- Data ONTAP Volume Shadow Copy Service (VSS) Hardware Provider on Windows Server hosts

Microsoft System Center

Microsoft System Center 2016 Virtual Machine Manager

Virtual machine manager, a part of the Microsoft System Center suite, is a virtualization and cloud management platform. It allows an administrator to configure and manage the servers, network, and storage resources. With SCVMM, you can manage your applications and services across multiple hypervisors and across hybrid cloud infrastructure to deliver flexible and cost-effective IT services. It includes the following capabilities:

- Datacenter: Configure and manage your datacenter components as a single fabric in VMM. Datacenter components include virtualization servers, networking components, and storage resources. VMM provisions and manages the resources needed to create and deploy virtual machines and services to private clouds.

- Virtualization hosts: VMM can add, provision, and manage Hyper-V and VMware virtualization hosts and clusters.

- Networking: Add networking resources to the VMM fabric, including network sites defined by IP subnets, virtual LANs (VLANs), logical switches, static IP address and MAC pools. VMM provides network virtualization, including support for creating and manage virtual networks and network gateways. Network virtualization allows multiple tenants to have isolated networks and their own IP address ranges for increased privacy and security. Using gateways, VMs on virtual networks can connect to physical networks in the same site or in different locations.

- Storage: Microsoft SCVMM can discover, classify provision, allocate, and assign local and remote storage. VMM supports block storage (Fibre Channel, iSCSI, and Serial Attached SCSI (SAS) storage area networks (SANs)).

- Library resources: The VMM fabric retains a library of file-based and non file-based resources that are used to create and deploy VMs and services on virtualization hosts. File-based resources include virtual hard disks, ISO images, and scripts. Non file-based resources include templates and profiles that are used to standardize the creation of VMs. Library resources are accessed through library shares.

Microsoft System Center 2016 Operation Manager

Operations Manager, a component of Microsoft System Center suite provides infrastructure monitoring that is flexible and cost-effective. It enables you to monitor services, devices, and operations for many computers in a single console. Operators can gain rapid insight into the state of the IT environment and the IT services running across different systems and workloads by using numerous views that show state, health, and performance information, as well as alerts generated for availability, performance, configuration, and security situations.https://docs.microsoft.com/en-us/system-center/scom/welcome

Cisco UCS Management and Microsoft System Center Integration

The Cisco Unified Computing System™ (Cisco UCS™) with its underlying programmable technology and management tools integrates with Microsoft systems management offerings to help optimize operations. You can simplify the tool set needed to monitor and manage bare-metal, virtualized, and cloud environments.

With Cisco UCS and Microsoft integration, you can boost operation efficiency and improve visibility and control.

Cisco and Microsoft have jointly developed and enhanced management packs for Microsoft System Center Operations Manager (SCOM) and PowerShell, so you can automate data center setup, provisioning, and management. This programmatic control is central to the provisioning of hybrid cloud infrastructure to enable self-service provisioning.

Cisco UCS Management Suite for Microsoft SCOM

A management pack is a definition file with predefined monitoring settings that enables you to monitor a specific service or application in an operations manager. The Cisco UCS Management Pack Suite for Microsoft System Center Operations Manager is a set of management packs that provide visibility into the health, performance, and availability of Cisco UCS servers managed by Cisco UCS Manager, Cisco UCS Central Software, and Cisco® Integrated Management Controller (IMC). Cisco UCS Manager provides embedded management for all software and hardware components in Cisco UCS. It supports Cisco UCS B-Series Blade Servers, C-Series Rack Servers, and S-Series Storage Servers; Cisco UCS Mini, and Cisco HyperFlex™ hyperconverged infrastructure. The IMC supports C-Series servers in standalone environments.

In addition to monitoring the health of Cisco UCS, you can do the following:

View Cisco UCS physical and virtual resources, service profiles, operating systems, and virtual machines (VMs).

Correlate faults and events between Cisco UCS infrastructure and both bare-metal and virtualized OSs that you manage with SCOM.

Visually correlate blades, service profiles, and host OSs by using SCOM status views to quickly determine the way an event will affect the overall system.

Cisco UCS PowerTool Suite

The Cisco UCS PowerTool Suite consists of a set of Microsoft Windows PowerShell modules for Cisco UCS Manager, Cisco IMC, and Cisco UCS Central Software that helps you configure and manage Cisco UCS domains and solutions. This command-line toolkit is based on PowerShell and is used to automate and integrate Cisco UCS management with Microsoft servers and applications. The UCS PowerTool module for UCS Manager includes over 1,970 cmdlets. In addition, by using Cisco UCS PowerTool together with similar management tools based on PowerShell from third-party vendors, you can manage and operate all components programmatically.

Cisco UCS Manager Add-In for Microsoft System Center Virtual Machine Manager

Cisco UCS Manager Add-in for SCVMM provides an extension to the Virtual Machine Manager user interface. The extended interface enables you to manage the UCS servers (blade servers and rack-mount servers). You can perform the following tasks using the extension:

- Viewing and acknowledging pending activities
- Mapping of the hypervisor host with the blades or rack servers on which it is installed
- Adding or removing a UCS domain

- Viewing the server details
- Assigning proxy settings to the domain.
- Viewing the firmware details of a server
- Viewing the fault information
- Viewing service profiles and service profile templates details
- Viewing the registered UCS domains details
- Launching the host (server) KVM console
- Managing the desired power state of the servers
- Switching on and off of the locator LEDs
- Associating service profile to a server
- Associating service profile to a server pool
- Changing the service profile association with a host firmware package
- Disassociating a service profile from a server
- Creating a service profile from a service profile template
- Uploading and upgrading the firmware on the servers

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 two organizations:

- Distributed Management Task Force (DMTF)
- 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 VMs 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. Compliance with SMI-S standards is defined by the Conformance Test Program and set by SNIA.

NetApp OnCommand Plug-in for Microsoft

The NetApp OnCommand Plug-In for Microsoft integrates with Microsoft System Center 2016 management of NetApp storage. The NetApp OnCommand Plug-In for Microsoft allows administrators to perform the following tasks:

- Monitor server and storage availability and capacity for Microsoft Windows Server Hyper-V VMs.
- Isolate problems using System Center Operations Manager (SCOM) alerts and health explorer views.
- Enable high availability and load balancing of management servers with OpsMgr Management Server Resource Pool.
- Leverage System Center Virtual Machine Manager and System Center Orchestrator for workflow automation.
- Report on ONTAP MetroCluster storage

OnCommand Plug-in for Microsoft enables administrators to discover and monitor NetApp SAN and SMB shares. You can use the included Virtualization Pack to discover and monitor Windows Server Hyper-V VMs and alert VM administrators of potential storage issues.

NetApp SnapManager for Hyper-V

NetApp SnapManager® for Hyper-V provides a solution for data protection and recovery for Microsoft Hyper-V VMs 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 VMs 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 VMs
- 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

SCVMM vs CLI

VMM cmdlets: Commadlets not only enable automation of repeatable actions but also standardize the workflows. All the tasks in SCVMM can also be accomplished using cmdlets. The SCVMM cmdlets can be invoked directly using VMM command shell or importing the VMM module to Windows powershell manually. Each cmdlet noun is preceeded with an “SC”. For example, Install-SCVMHostCluster creates a failover-cluster from Hyper-V hosts. Additionally Hyper-V specific cmdlets, enable creating and managing virtual machine resources. For example, Add-VMNetworkAdapter enables adding virtual network adapter to a virtual machine.

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 IOMeter tool. The system was validated for resiliency by failing various aspects of the system under load. Examples of the types of tests executed include:

- Failure and recovery of SAN booted Windows Hyper-V hosts in a cluster
- Rebooting of SAN booted hosts
- Service Profiles migration between blades
- Failure of partial and complete IOM links
- Failure and recovery of FC paths to AFF nodes, MDS switches, and fabric interconnects
- SSD removal to trigger an aggregate rebuild
- Storage link failure between one of the AFF nodes and the Cisco MDS
- Storage controller failure and takeover by surviving controller.
- Load was generated using the IOMeter tool and different IO profiles were used to reflect the different profiles that are seen in customer networks

Validated Hardware and Software

Table 4 Error! Reference source not found.lists the hardware and software versions used during solution validation. It is important to note that Cisco, NetApp, and Microsoft 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
- Microsoft Interop Matrix

Table 4 Validated Software Versions

Layer	Device	Image	Comments
Compute	• Cisco UCS Fabric Interconnects 6200 and 6300 Series. • UCS B-200 M4, UCS C-220 M4	• 3.1(3a) - Infrastructure Bundle • 3.1(2f) – Server Bundle	Includes the Cisco UCS-IOM 2304 Cisco UCS Manager, Cisco UCS VIC 1340 and Cisco UCS VIC 1385

Network	Cisco Nexus 9000 NX-OS	7.0(3)I4(5)	
Storage	NetApp AFF A300	ONTAP 9.1	
	Cisco MDS 9148S	7.3(0)D1(1)	
Software	Cisco UCS Manager	3.1(3a)	
	Microsoft System Center Virtual Machine Manager	2016 (version: 4.0.2051.0)	
	Microsoft Hyper-V	2016	
	Microsoft System Center Operation Manager	2016 (version: 7.2.11878.0)	

Summary

FlexPod Datacenter with Windows Server 2016 Hyper-V is the optimal shared infrastructure foundation to deploy a variety of IT workloads that is future proofed with 16 Gb/s FC or 40Gb/s iSCSI, with either delivering 40Gb Ethernet connectivity. Cisco and NetApp have created a platform that is both flexible and scalable for multiple use cases and applications. From virtual desktop infrastructure to SAP®, FlexPod can efficiently and effectively support business-critical applications 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.

References

Products and Solutions

Cisco Unified Computing System:

http://www.cisco.com/en/US/products/ps10265/index.html

Cisco UCS 6300 Series Fabric Interconnects:

http://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

Cisco UCS C-Series Rack Mount Servers:

http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-c-series-rack-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 UCS Management with Microsoft Systems Center:

https://communities.cisco.com/docs/DOC-38133

Cisco Nexus 9000 Series Switches:

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

Cisco MDS 9000 Multilayer Fabric Switches:

http://www.cisco.com/c/en/us/products/storage-networking/mds-9000-series-multilayer-switches/index.html

Microsoft Windows Server 2016:

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

Microsoft System Center 2016:

https://docs.microsoft.com/en-us/system-center/

NetApp ONTAP 9:

http://www.netapp.com/us/products/platform-os/ontap/index.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/

Interoperability Matrixes

Cisco UCS Hardware Compatibility Matrix:

https://ucshcltool.cloudapps.cisco.com/public/

Microsoft Server Catalog:

https://www.windowsservercatalog.com/

NetApp Interoperability Matrix Tool:

http://support.netapp.com/matrix/

About the Authors

Sanjeev Naldurgkar, Technical Marketing Engineer, Data Center Solutions Engineering, Cisco Systems Inc.

Sanjeev Naldurgkar is a Technical Marketing Engineer with Cisco UCS Datacenter Solutions Group. He has been with Cisco for 5 years focusing with delivery of customer-driven solutions on Microsoft Hyper-V and VMware vSphere. Sanjeev has over 16 years of experience in the IT Infrastructure, Server virtualization and Cloud Computing. He holds a Bachelor Degree in Electronics and Communications Engineering, and leading industry certifications from Microsoft and VMware.

Aaron Kirk, Technical Marketing Engineer, Converged Infrastructure Engineering, NetApp Inc.

Aaron Kirk is a Technical Marketing Engineer in the NetApp Converged Infrastructure Engineering team. He focuses on producing validated reference architectures that promote the benefits of end-to-end data center solutions and cloud environments. Aaron has been at NetApp since 2010, previously working as an engineer on the MetroCluster product in the Data Protection Group. Aaron holds a Bachelor of Science in Computer Science and a Masters of Business Administration from North Carolina State University.

Acknowledgements

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

- Suhas Rao
- Rajendra Yogendra
- John George
- Chris Reno
- Karthick Radhakrishnan