White Paper

FlexPod for Windows Server 2012 Hyper-V Design Guide

Contents

FlexPod for Windows Server 2012 Hyper-V

Goal of This Document

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.

The purpose of this document is to describe the Cisco and NetApp FlexPod solution, which is a validated approach for deploying Cisco and NetApp technologies as a shared cloud infrastructure.

Audience

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

Changes in FlexPod

The following design elements distinguish this version of FlexPod from previous models:

- End-to-end Fibre Channel over Ethernet (FCoE) delivering a unified Ethernet fabric.
- Single-wire Cisco Unified Computing System™ Manager management for Cisco UCS® C-Series M3 Rack Servers and the Cisco UCS Virtual Interface Card (VIC) 1225. These features effectively double the server density per I/O module while reducing cabling costs.
- Cisco Data Center Virtual Machine Fabric Extender (VM-FEX) for Windows Server 2012 Hyper-V offloads the task switching VM network traffic from the hypervisor. All switching is performed by the external fabric interconnect, which can switch not only between physical ports, but also between virtual interfaces (VIFs) that correspond to the virtual network interface cards (vNICs) on the VMs.

Technology Overview

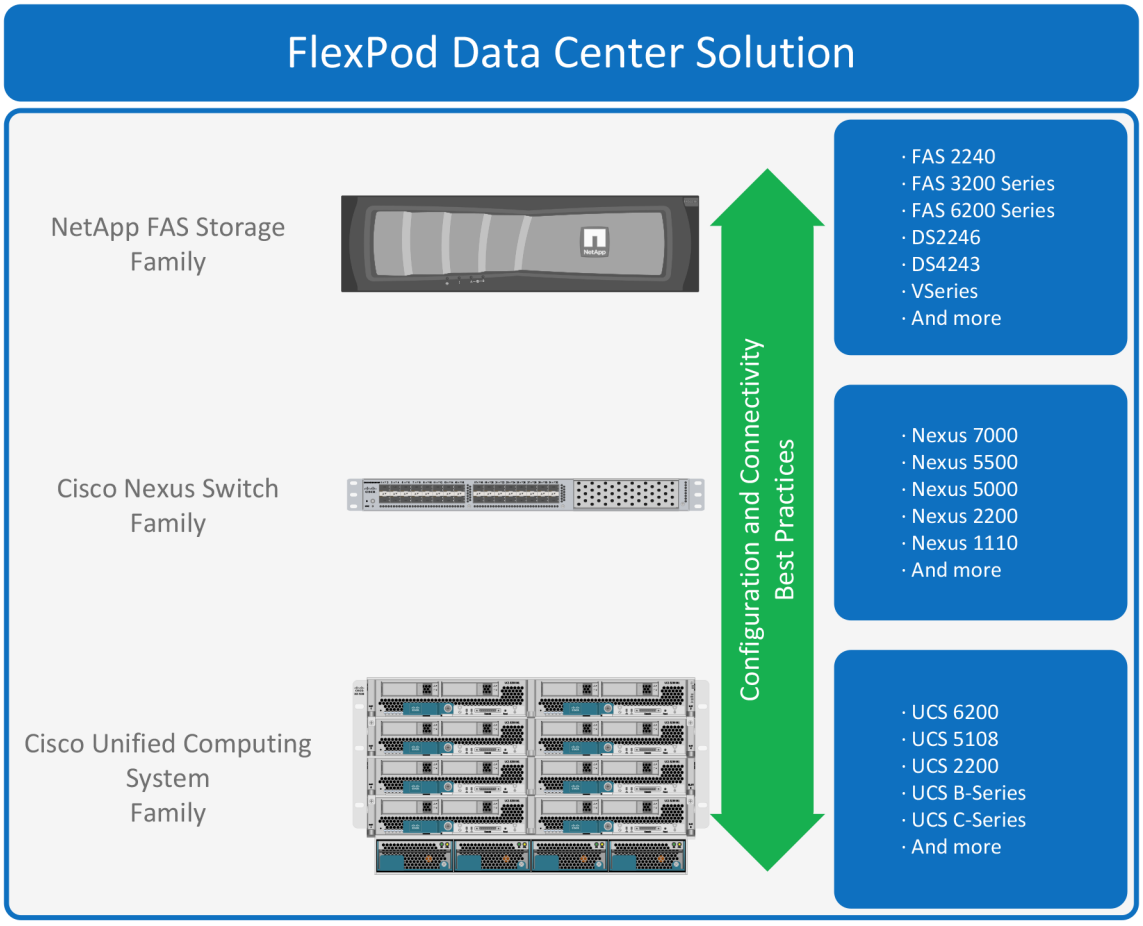
Industry trends indicate a vast data center transformation toward shared infrastructure and cloud computing. Enterprise customers are moving away from isolated centers of IT operation toward more cost-effective virtualized environments.

The objective of the move toward virtualization, and eventually to cloud computing, is to increase agility and reduce costs.

Especially because companies must address resistance to change in both their organizational and technical IT models, achieving this transformation can seem daunting and complex. To accelerate the process and simplify the evolution to a shared cloud infrastructure, Cisco and NetApp have developed a solution called Windows Server 2012 Hyper-V on FlexPod®.

FlexPod is a predesigned, best practice data center architecture that is built on the Cisco® Unified Computing System™ (Cisco UCS™), the Cisco Nexus® family of switches, and NetApp® fabric-attached storage (FAS) or V-Series systems (see Figure 1). FlexPod is a suitable platform for running a variety of virtualization hypervisors as well as bare metal operating systems and enterprise workloads. FlexPod delivers not only a baseline configuration, but also the flexibility to be sized and optimized to accommodate many different use cases and requirements.

Figure 1. FlexPod Component Families.



This document describes FlexPod for Windows Server 2012 Hyper-V model from Cisco and NetApp and discusses design choices and deployment best practices using this shared infrastructure platform.

Customer Challenges

As customers transition toward shared infrastructure or cloud computing, they face a number of questions, such as the following:

- How do I start the transition?
- What will be my return on investment?
- How do I build a future-proof infrastructure?
- How do I transition from my current infrastructure cost-effectively?
- Will my applications run properly in a shared infrastructure?
- How do I manage the infrastructure?

The FlexPod architecture is designed to help you answer these questions by providing proven guidance and measurable value. By introducing standardization, FlexPod helps customers mitigate the risk and uncertainty involved in planning, designing, and implementing a new data center infrastructure. The result is a more predictive and adaptable architecture capable of meeting and exceeding customers' IT demands.

FlexPod Program Benefits

Cisco and NetApp have thoroughly tested 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, and what is not, a FlexPod configuration)
- Frequently asked questions (FAQs)
- Cisco® Validated Designs (CVDs) and NetApp Validated Architectures (NVAs) focused on a variety of use cases

Cisco and NetApp have also built an experienced support team focused on FlexPod solutions, from customer account and technical sales representatives to professional services and technical support engineers. The support alliance provided by NetApp and Cisco provides customers and channel services partners with direct access to technical experts who collaborate with multiple 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. The following IT initiatives are addressed by the FlexPod solution:

Integrated Systems

FlexPod is a pre-validated infrastructure that brings together computing, storage and network to simplify, accelerate and minimize the risk associated with data center builds and application roll-outs. These integrated systems provide a standardized approach in the data center supports staff expertise, application onboarding, and automation, as well as operational efficiencies that are important for compliance and certification.

Fabric Infrastructure Resilience

FlexPod is a highly available and scalable infrastructure that IT can evolve over time to support multiple physical and virtual application workloads. FlexPod contains no single point of failure at any level, from the server through the network to the storage. The fabric is fully redundant and scalable providing seamless traffic failover should any individual component fail at the physical or virtual layer.

Fabric Convergence

The Cisco Unified Fabric is a data center network that supports both traditional LAN traffic and all types of storage traffic, including the lossless requirements for block-level storage transport over Fibre Channel. The Cisco Unified Fabric creates high-performance, low-latency, and highly available networks serving a diverse set of data center needs.

FlexPod uses Cisco Unified Fabric to offer a wire-once environment that accelerates application deployment. Cisco Unified Fabric also offers the efficiencies associated with infrastructure consolidation, including:

- Cost savings from the reduction in switches (LAN/SAN switch ports), associated cabling, rack space , all of which reduce capital expenditures (CapEx)
- Cost savings on power and cooling, which reduce operating expenses (OpEx)
- Migration to the faster 10 Gigabit Ethernet network, and in the future, to 40 Gigabit Ethernet and 100 Gigabit Ethernet
- Evolution to a converged network with little disruption to operations FlexPod with Cisco Unified Fabric helps you preserve investments in existing infrastructure, management tools, and staff training and expertise
- Simplified cabling, provisioning, and network maintenance to improve productivity and operational models

Network Virtualization

FlexPod delivers the capability to securely separate and connect virtual machines into the network. Using technologies such as VLANs, QoS and VM-FEX, this solution allows network policies and services to be uniformly applied within the integrated compute stack. This capability enables the full utilization of FlexPod while maintaining consistent application and security policy enforcement across the stack even with workload mobility.

FlexPod 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 versatility to meet a variety of SLAs and IT initiatives, including:

- Application roll outs or application migrations
- Business Continuity/Disaster Recovery
- Desktop Virtualization
- Cloud delivery models (public, private, hybrid) and service models (IaaS, PaaS, SaaS)
- Asset Consolidation and Virtualization

FlexPod

System Overview

FlexPod is a best practice data center architecture that is built with three components:

- Cisco Unified Computing System™ (Cisco UCS®)
- Cisco Nexus® switches
- NetApp fabric-attached storage (FAS) systems

These components are connected and configured according to the best practices of both Cisco and NetApp to provide the 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 need multiple consistent deployments (rolling out additional FlexPod stacks). FlexPod delivers not only a baseline configuration but also the flexibility to be sized and optimized to accommodate many different use cases.

Typically, the more scalable and flexible a solution is, the more difficult it becomes to maintain a single unified architecture capable of offering the same features and functionality across each implementation. This is one of the key benefits of FlexPod. Each of the component families shown in Figure 1 (Cisco UCS, Cisco Nexus, and NetApp FAS) 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 Principles

FlexPod addresses four primary design principles scalability, elasticity, availability and manageability. These architecture goals as follows:

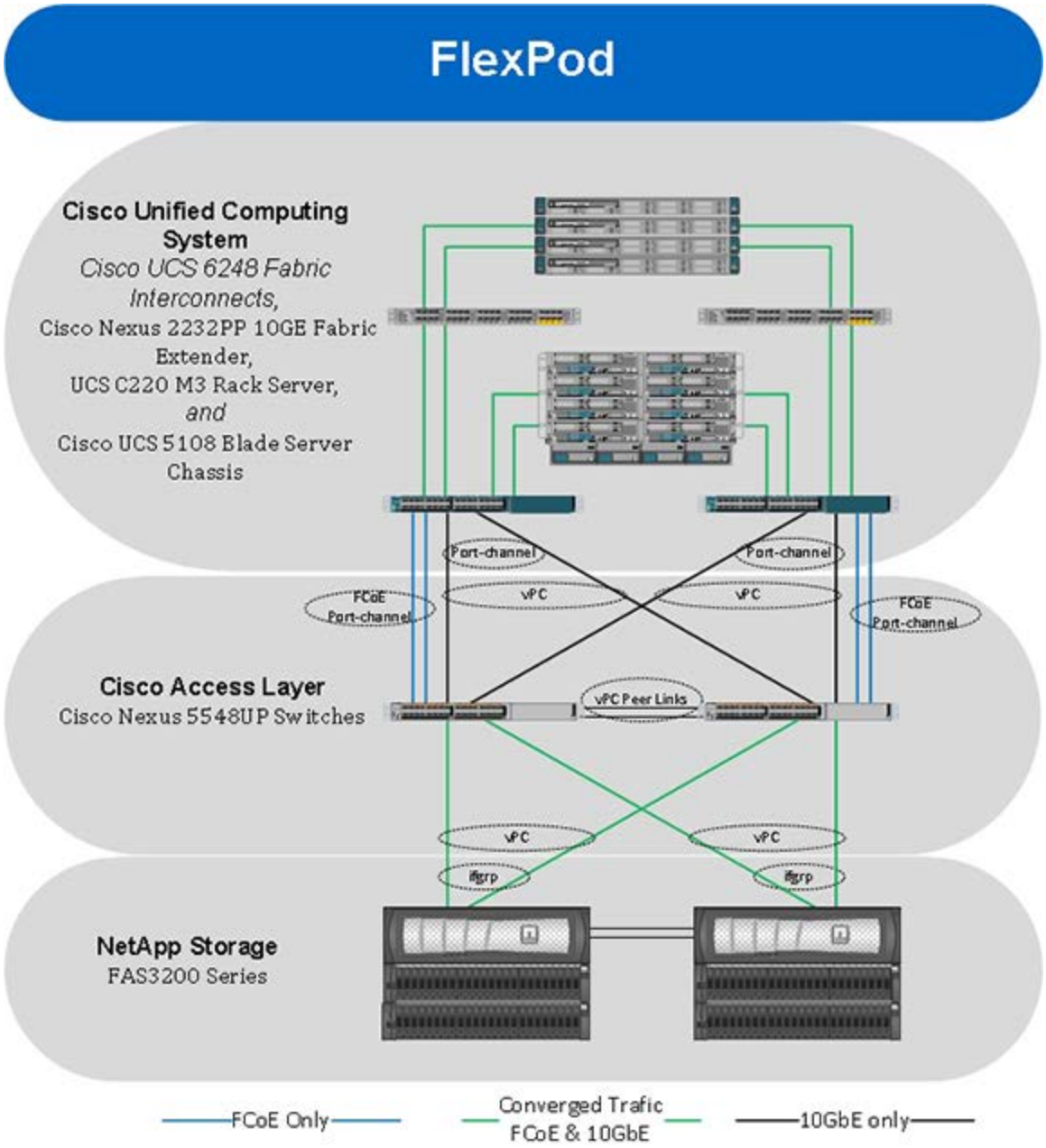
- **Application availability**—ensure accessible and ready to use services
- **Scalability**—addressing increasing demands with appropriate resources
- **Flexibility**—provide new services or recovered resources without requiring infrastructure modification
- **Manageability**—facilitate efficient infrastructure operations through open standards and APIs

Note: Performance and Security are key design criteria that were not directly addressed in this project but have been addressed in other collateral, benchmarking and solution testing efforts. Functionality and basic security elements were validated.

FlexPod—Discrete Uplink Design

Figure 2 represents the FlexPod Distinct Uplink Design with Data ONTAP operating in 7-Mode. Data On-Tap operating in 7-Mode is NetApp's traditional functional model. As depicted, the FAS devices are configured in an HA pair delivering five nines availability. Scalability is achieved through the addition of storage capacity (disk/shelves) as well as through additional controllers whether they be FAS 2200, 3200, or 6200 series. The controllers are only deployed in HA pairs meaning more HA pairs can be added for scalability but each pair is managed separately.

Figure 2. FlexPod—Discrete Uplink Design with 7-Mode Data ONTAP



The FlexPod Discrete Uplink design is an end-to-end Ethernet transport solution supporting multiple LAN protocols and most notably FCoE. The solution provides a unified 10 Gigabit enabled fabric defined by dedicated FCoE uplinks and dedicated Ethernet uplinks between the Cisco UCS Fabric Interconnects and the Cisco Nexus switches, as well as, converged connectivity between the NetApp storage devices and the same multi-purpose Cisco Nexus platforms.

The Discrete Uplink design does not employ a dedicated SAN switching environment and requires no dedicated Fibre Channel connectivity. The Cisco Nexus 5500 Series Switches are configured in N Port ID Virtualization (NPIV) mode providing storage services for the FCoE based-traffic traversing its fabric.

As illustrated in Figure 2, link aggregation technologies plays an important role providing improved aggregate bandwidth and link resiliency across the solution stack. The NetApp storage controllers, Cisco Unified Computing System and Nexus 5500 platforms all support active port channeling using 802.3ad standard Link Aggregation Control Protocol (LACP). Port channeling is a link aggregation technique offering link fault tolerance and traffic distribution (load balancing) for improved aggregate bandwidth across member ports. In addition, the Cisco Nexus 5000 series features virtual PortChannel (vPC) capabilities. vPC allows links that are physically connected to two different Cisco Nexus 5500 Series devices to appear as a single “logical” port channel to a third device, essentially offering device fault tolerance. vPC addresses aggregate bandwidth, link and device resiliency. The Cisco UCS Fabric Interconnects and NetApp FAS controllers benefit from the Nexus vPC abstraction gaining link and device resiliency, as well as, full utilization of a non-blocking Ethernet fabric.

Note: The Spanning Tree protocol does not actively block redundant physical links in a properly configured vPC-enabled environment so all ports are should forward on vPC member ports.

From a storage traffic (FCoE) perspective both standard LACP and Cisco’s vPC link aggregation technologies play an important role in the FlexPod Discrete Uplink design. showcases the use of dedicated FCoE uplinks between the UCS Fabric Interconnects and Nexus 5500 unified switches. The Cisco UCS Fabric Interconnects operate in N-Port Virtualization (NPV) mode meaning the servers FC traffic is either manually or automatically pinned to a specific FCoE uplink, in this case either of the two FCoE port channels. The use of distinct FCoE port channels with distinct VSANs allows an organization to maintain the traditional SAN A/B separation best practices. vPC links between the Cisco Nexus 5500 and NetApp storage controllers’ Unified Target Adapters (UTA) are converged, supporting both FCoE and traditional Ethernet traffic at 10 Gigabit providing a robust “last mile” connection between initiator and target.

Organizations with the following characteristics or needs may wish to use the 7-Mode design:

- Existing Data ONTAP 7G and Data ONTAP 8.x 7-Mode customers who are looking to upgrade
- Midsize enterprises: customers who are primarily interested in the FAS2000 series
- Customers who absolutely require SnapVault ®, synchronous SnapMirror ®, MetroCluster™, SnapLock ® software, IPv6, or Data ONTAP Edge

Note: It is always advisable to seek council from experts, please consider reaching out to your NetApp account team or partner for further guidance.

The Logical Build section provides more details regarding the design and of the physical components and of the virtual environment consisting of Windows Server 2012 with Hyper-V, Cisco Unified Computing System and NetApp storage controllers.

Integrated System Components

The following components are required to deploy the Discrete Uplink design:

- Cisco Unified Compute System
- Cisco Nexus 5500 Series Switch
- NetApp FAS and Data ONTAP
- Windows Server 2012 with Hyper-V Role

Cisco Unified Compute System

The Cisco Unified Computing System is a next-generation for blade and rack server computing. The Cisco UCS is a innovative data center platform that unites compute, network, storage access, and virtualization into a cohesive system designed to reduce total cost of ownership (TCO) and increase business agility. The system integrates a low-latency, lossless 10 Gigabit Ethernet unified network fabric with enterprise-class, x86-architecture servers. The system is an integrated, scalable, multi-chassis platform in which all resources participate in a unified management domain. Managed as a single system whether it has one server or 160 servers with thousands of virtual machines, the Cisco UCS decouples scale from complexity. The Cisco UCS accelerates the delivery of new services simply, reliably, and securely through end-to-end provisioning and migration support for both virtualized and non-virtualized systems.

The Cisco Unified Computing System consists of the following components:

- Cisco UCS 6200 Series Fabric Interconnects is a family of line-rate, low-latency, lossless, 10-Gbps Ethernet and Fibre Channel over Ethernet interconnect switches providing the management and communication backbone for the Cisco Unified Computing System. Cisco UCS supports VM-FEX technology.
- Cisco UCS 5100 Series Blade Server Chassis supports up to eight blade servers and up to two fabric extenders in a six rack unit (RU) enclosure.
- Cisco UCS B-Series Blade Servers increase performance, efficiency, versatility and productivity with these Intel based blade servers.
- Cisco UCS Adapters wire-once architecture offers a range of options to converge the fabric, optimize virtualization and simplify management. Cisco adapters support VM-FEX technology.
- Cisco UCS C-Series Rack Mount Server deliver unified computing in an industry-standard form factor to reduce total cost of ownership and increase agility.
- Cisco UCS Manager provides unified, embedded management of all software and hardware components in the Cisco UCS.

For more information, see: http://www.cisco.com/en/US/products/ps10265/index.html

Cisco Nexus 5500 Series Switch

The Cisco Nexus 5000 Series is designed for data center environments with cut-through technology that enables consistent low-latency Ethernet solutions, with front-to-back or back-to-front cooling, and with data ports in the rear, bringing switching into close proximity with servers and making cable runs short and simple. The switch series is highly serviceable, with redundant, hot-pluggable power supplies and fan modules. It uses data center-class Cisco® NX-OS Software for high reliability and ease of management.

The Cisco Nexus 5500 platform extends the industry-leading versatility of the Cisco Nexus 5000 Series purpose-built 10 Gigabit Ethernet data center-class switches and provides innovative advances toward higher density, lower latency, and multilayer services. The Cisco Nexus 5500 platform is well suited for enterprise-class data center server access-layer deployments across a diverse set of physical, virtual, storage-access, and high-performance computing (HPC) data center environments.

The switch used in this FlexPod architecture, the Cisco Nexus 5548UP. The following specifications describe the Nexus 5548UP switch:

- A 1-rack-unit, 1/10 Gigabit Ethernet switch
- 32 fixed Unified Ports on base chassis and one expansion slot totaling 48 ports
- The slot can support any of the three modules: Unified Ports, 1/2/4/8 native Fibre Channel, and Ethernet or FCoE
- Throughput of up to 960 Gbps

For more information, see: http://www.cisco.com/en/US/products/ps9670/index.html

Cisco Nexus 2232PP 10GE Fabric Extender

The Cisco Nexus 2232PP 10 Gigabit provides 32 10 Gigabit Ethernet and Fibre Channel Over Ethernet (FCoE) Small Form-Factor Pluggable Plus (SFP+) server ports and eight 10 Gb Ethernet and FCoE SFP+ uplink ports in a compact 1 rack unit (1RU) form factor.

The built-in standalone software, Cisco Integrated Management Controller (CIMC), manages cisco UCS C-Series Rack-Mount Servers. When a C-Series Rack-Mount Server is integrated with Cisco UCS Manager, via the Nexus 2232 platform, the CIMC does not manage the server anymore. Instead it is managed with the Cisco UCS Manager software. The server is managed using the Cisco UCS Manager GUI or Cisco UCS Manager CLI. The Nexus 2232 provides data and control traffic support for the integrated C-Series server.

Cisco VM-FEX

Cisco VM-FEX technology collapses virtual switching infrastructure and physical switching infrastructure into a single, easy-to-manage environment. Benefits include:

- **Simplified operations:** Eliminates the need for a separate, virtual networking infrastructure
- **Improved network security:** Contains VLAN proliferation
- **Optimized network utilization:** Reduces broadcast domains
- **Enhanced application performance:** Offloads virtual machine switching from host CPU to parent switch application-specific integrated circuits (ASICs)

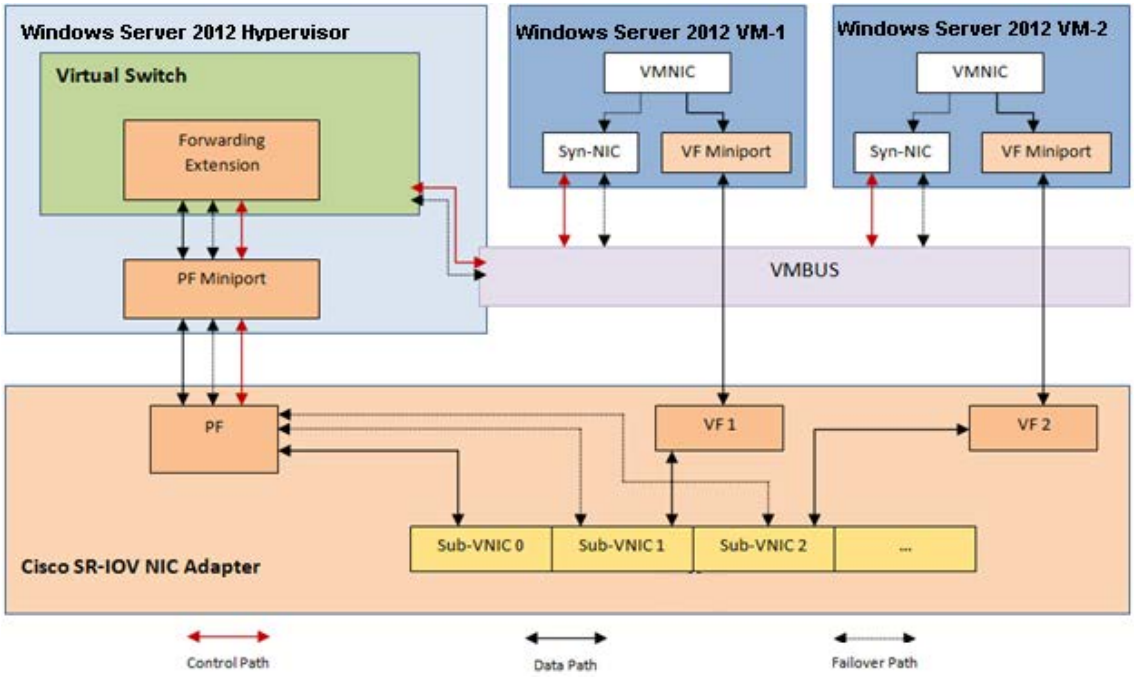
VM-FEX is supported on Windows Server 2012 Hyper-V hypervisors and fully supports workload mobility through Quick Migration and Live Migration.

VM-FEX eliminates the virtual switch within the hypervisor by providing individual Virtual Machines (VMs) virtual ports on the physical network switch. VM I/O is sent directly to the upstream physical network switch that takes full responsibility for VM switching and policy enforcement. This leads to consistent treatment for all network traffic, virtual or physical. VM-FEX collapses virtual and physical switching layers into one and reduces the number of network management points by an order of magnitude.

The SR-IOV specs do however describe how a hardware device can expose multiple “light-weight” hardware surfaces for use by virtual machines. These are called Virtual Functions or VFs for short. VFs are associated with a Physical Function (PF). The PF is what the parent partition uses in Hyper-V and is equivalent to the regular BDF (Bus/Device/Function) addressed Personal Computer Interconnect (PCI) device you may have heard of before. The PF is responsible for arbitration relating to policy decisions (such as link speed or MAC addresses in use by VMs in the case of networking) and for I/O from the parent partition itself. Although a VF could be used by the parent partition, in Windows Server 2012, VFs are used only by virtual machines. A single PCI Express device can expose multiple PFs, each with their own set of VF resources.

While software based devices work extremely efficiently, they still have an unavoidable overhead to the I/O path. Consequently, software based devices introduce latency, increase overall path length and consume computing cycles. With SR-IOV capability, part of the network adapter hardware is exposed inside the virtual machine and provides a direct I/O path to the network hardware. For this reason a vendor specific driver needs to be loaded into the virtual machine in the in order to utilize the VF network adapter (Figure 3).

Figure 3. VM-FEX from the Hyper-V Node Point Perspective



As illustrated in Figure 3, the I/O data path from the VF does not go across VMBus or through the Windows hypervisor. It is a direct hardware path from the VF in the VM to the NIC.

Also note the control path for the VF is through VMBus (back to the PF driver in the parent partition).

NetApp FAS and Data ONTAP

NetApp solutions are user friendly, easy to manage, and quick to deploy, offering increased availability while consuming fewer IT resources. This means that they dramatically lower the lifetime total cost of ownership. Where other manage complexity, NetApp eliminates it. A NetApp solution consists of hardware in the form of controllers and disk storage and NetApp’s Data ONTAP operating system, the #1 storage OS.

NetApp offers a Unified Storage Architecture. The term “unified” refers to a family of storage systems that simultaneously support storage area network (SAN), network-attached storage (NAS), and iSCSI across many operating environments such as VMware, Windows, and UNIX. This single architecture provides access to data by using industry-standard protocols, including NFS, CIFS, iSCSI, FCP, SCSI, FTP, and HTTP. Connectivity options include standard Ethernet (10/100/1000 or 10GbE) and Fibre Channel (1, 2, 4, or 8 Gb/sec). In addition, all systems can be configured with high-performance solid state drives (SSDs) or serial ATA (SAS) disks for primary storage applications, low-cost SATA disks for secondary applications (backup, archive, and so on), or a mix of the different disk types.

A storage system running Data ONTAP has a main unit, also known as the chassis or Controller, which is the hardware device that receives and sends data. This unit detects and gathers information about the hardware configuration, the storage system components, the operational status, hardware failures, and other error conditions.

A storage system uses storage on disk shelves. The disk shelves are the containers or device carriers that hold disks and associated hardware such as power supplies, connectivity interfaces, and cabling.

If storage requirements change over time, NetApp storage offers the flexibility to change quickly as needed without expensive and disruptive “forklift” upgrades. For example, a LUN can be changed from FC access to iSCSI access without moving or copying the data. Only a simple dismount of the FC LUN and a mount of the same LUN using iSCSI would be required. In addition, a single copy of data can be shared between Windows and UNIX systems while allowing each environment to access the data through native protocols and applications. If a system was originally purchased with all SATA disks for backup applications, high-performance SAS disks could be added to support primary storage applications such as Oracle, Microsoft Exchange, or ClearCase.

NetApp storage solutions provide redundancy and fault tolerance through clustered storage controllers, hot swappable redundant components (such as cooling fans, power supplies, disk drives, and shelves), and multiple network interfaces. This highly available and flexible architecture enables customers to manage all data under one common infrastructure while achieving mission requirements. NetApp Unified Storage Architecture allows data storage with higher availability and performance, easier dynamic expansion, and more unrivalled ease of management than any other solution.

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

- **Thin provisioning.** Volumes are created using “virtual” sizing. They appear to be provisioned at their full capacity, but are actually created much smaller and use additional space only when it is actually needed. Extra unused storage is shared across all volumes, and the volumes can grow and shrink on demand.
- **Snapshot copies.** Automatically scheduled point-in-time copies that write only changed blocks, with no performance penalty. The Snapshot copies consume minimal storage space, since only changes to the active file system are written. Individual files and directories can easily be recovered from any Snapshot copy, and the entire volume can be restored back to any Snapshot state in seconds.
- **FlexClone volumes.** Near-zero space, instant “virtual” copies of datasets. The clones are writable, but only changes to the original are stored, so they provide rapid, space-efficient creation of additional data copies ideally suited for test/dev environments.

- **Deduplication.** Removes redundant data blocks in primary and secondary storage with flexible policies to determine when the deduplication process is run.
- **Compression.** Compresses data blocks. Compression can be run whether or not deduplication is enabled and can provide additional space savings, whether run alone or together with deduplication.
- **SnapMirror.** Volumes can be asynchronously replicated either within the cluster or to another cluster.

For more information see: http://www.netapp.com/us/products/platform-os/data-ontap-8/index.aspx

Windows Server 2012—Hyper-V

Windows Server® 2012 Hyper-V provides significant increases to scalability and expands support for host processors and memory. New features include support for up to 64 processors and 1 terabyte (TB) of memory for Hyper V virtual machines, in many cases supporting fourto16 times the density of processors, memory, cluster nodes and running virtual machines. In addition, Windows Server® 2012 Hyper-V supports innovative server features including the ability to project a virtual non-uniform memory access (NUMA) topology onto a virtual machine to provide optimal performance and workload scalability in large virtual machine configurations. Windows Server® 2012 Hyper-V also provides improvements to Dynamic Memory including Minimum Memory and Hyper-V Smart Paging. Minimum Memory allows Hyper V to reclaim the unused memory from virtual machines to allow for higher virtual machine consolidation numbers. Smart Paging is used to bridge the memory gap between minimum and startup memory by allowing virtual machines to start reliably when the minimum memory setting has indirectly led to an insufficient amount of available physical memory during restart. In addition to these memory enhancements, Windows Server® 2012 Hyper-V allows for runtime configuration of memory settings, including increasing the maximum memory and decreasing the minimum memory of running virtual machines. These updated features help ensure that your virtualization infrastructure can support the configuration of large, high-performance virtual machines to maintain demanding workloads.

Windows Server® 2012 Hyper-V includes an update to the virtual hard disk format called VHDX. VHDX provides a higher capacity (up to 64 terabytes of storage), helps provide additional protection from corruption from power failures and prevents performance degradation on large-sector physical disks by optimizing structure alignment.

Windows Server® 2012 Hyper-V also includes virtual Fibre Channel support allowing virtual machines to have unmediated access to SAN logical unit numbers (LUNs). Virtual Fibre Channel enables scenarios, such as running the Windows Failover Cluster Management feature inside the guest operating system of a virtual machine connected to shared Fibre Channel storage. Virtual Fibre Channel supports multipath input/output (MPIO), N_Port ID Virtualization (NPIV) for one-to-many mappings and up to four virtual Fibre Channel adapters per virtual machine.

Windows Server® 2012 introduces several networking enhancements, including support for Single Root Input/Output Virtualization (SR-IOV), third-party extensions to the Hyper-V extensible switch, QoS minimum bandwidth, network virtualization, and Datacenter Bridging (DCB).

The virtualization layer is one of the primary enablers in environments with greater IT maturity. The decoupling of hardware, operating systems, data, applications, and user state opens a wide range of options for easier management and distribution of workloads across the physical infrastructure. The ability of the virtualization layer to migrate running virtual machines from one server to another without downtime and many other features that are provided by hypervisor-based virtualization technologies enable a rich set of solution capabilities. These capabilities can be utilized by the automation, management, and orchestration layers to maintain desired states, proactively address decaying hardware, or other issues that would otherwise cause faults or service disruptions.

Like the hardware layer, the automation, management, and orchestration layers must be able to manage the virtualization layer. Virtualization provides an abstraction of software from hardware that moves the majority of management and automation to software instead of requiring people to perform manual operations on physical hardware.

As described above, Windows Server® 2012 Hyper-V introduces a number of improvements in both virtualization features and scale with this release. A comparison chart of scale improvements and feature enhancements are provided in the table below as a reference.

Table 1 provides a feature comparison list of Hyper-V capabilities

Table 1. Comparison List of Hyper-V Capabilities

	Windows Server® 2008	Windows Server® 2008 R2	Windows Server® 2012
Scale			
HW Logical Processor (LP) Support	16 LPs	64 LPs	320 LPs
Physical Memory Support	1 TB	1 TB	4 TB
Cluster Scale	16 Nodes up to 1000 VMs	16 Nodes up to 1000 VMs	64 Nodes up to 4000 VMs
Virtual Machine Processor Support	Up to 4 Virtual Processors (VPs)	Up to 4 VPs	Up to 64 VPs
VM Memory	Up to 64 GB	Up to 64 GB	Up to 1 TB
Live Migration	Yes, one at a time	Yes, one at a time	Yes, with no limits. As many as hardware will allow.
Live Storage Migration	No. Quick Storage Migration via VMM.	No. Quick Storage Migration via VMM.	Yes, with no limits. As many as hardware will allow.
Servers in a Cluster	16	16	64
VP:LP Ratio	8:1	8:1 for Server	No limits. As many as hardware will allow.
		12:1 for Client (Virtual Desktop Infrastructure (VDI))	
Storage			
Live Storage Migration	No. Quick Storage Migration via VMM.	No. Quick Storage Migration via VMM.	Yes, with no limits. As many as hardware will allow.
VMs on File Storage	No	No	Yes, SMB 3
Guest Fiber Channel	No	No	Yes
Virtual Disk Format	Virtual Hard Disk (VHD) up to 2 TB	VHD up to 2 TB	VHD up to 2 TB
			VHDX up to 64 TB
VM Guest Clustering	Yes, via Internet Small Computer System Interface (iSCSI)	Yes, via iSCSI	Yes, via iSCSI, Fibre Channel and SMB
Native 4k Disk Support	No	No	Yes
Live VHD Merge	No, offline.	No, offline.	Yes
Live New Parent	No	No	Yes
Secure Offloaded Data Transfer (ODX)	No	No	Yes
Networking			
NIC Teaming	Yes, via partners	Yes, via partners	Windows Network Interface Controller (NIC) Teaming in box.
VLAN Tagging	Yes	Yes	Yes
MAC Spoofing Protection	No	Yes, with R2 SP1	Yes
Address Resolution Protocol (ARP) Spoofing Protection	No	Yes, with R2 SP1	Yes

	Windows Server® 2008	Windows Server® 2008 R2	Windows Server® 2012
SR-IOV Networking	No	No	Yes
Network QoS	No	No	Yes
Network Metering	No	No	Yes
Network Monitor Modes	No	No	Yes
IPsec Task Offload	No	No	Yes
VM Trunk Mode	No	No	Yes
Manageability			
Hyper-V PowerShell	No	No	Yes
Network PowerShell	No	No	Yes
Storage PowerShell	No	No	Yes
SCONFIG	No	Yes	Yes
Enable/Disable Shell	No	No	Yes, MinShell
	(Server Core at Operating System (OS) Setup)	(Server Core at OS Setup)	
VMConnect Support for Microsoft® RemoteFX™	N/A	No	Yes

The Hyper-V host cluster requires different types of network access, as described in the table 2.

Table 2. Host Cluster Networks

Network access type	Purpose of the network access type	Network traffic requirements	Recommended network access
Virtual machine access	Workloads running on virtual machines usually require external network connectivity to service client requests.	Varies	Public access that can be teamed for link aggregation or to fail over the cluster.
Cluster and Cluster Shared Volumes	Preferred network used by the cluster for communications to maintain cluster health. Also, used by CSV to send data between owner and non-owner nodes. If storage access is interrupted, this network is used to access CSV or to maintain and back up CSV.	Usually low bandwidth and low latency. Occasionally, high bandwidth	Private access
Live Migration	Transfer virtual machine memory and state.	High bandwidth and low latency during migrations	Private access
Storage	Access storage through iSCSI.	High bandwidth and low latency	Usually, dedicated and private access.
Management	Managing the Hyper-V management operating system. This network is used by Hyper-V Manager.	Low bandwidth	Public access that can be teamed to fail over the cluster.

Highly available host servers are one critical component of a dynamic, virtual infrastructure. A Hyper-V host failover cluster is a group of independent servers that work together to increase the availability of applications and services. The clustered servers (nodes) are connected physically. If one of the cluster nodes fails, another node begins to provide service. In the case of a planned Live Migration, users experience no perceptible service interruption.

Domain and Element Management

This section of the document provides general descriptions of the domain and element mangers used during the validation effort. The following managers were used:

- Cisco UCS Manager
- Cisco UCS Power Tool
- Cisco VM-FEX Port Profile Configuration Utility
- NetApp OnCommand System Manager
- NetApp SnapDrive for Windows
- NetApp SnapManager for Hyper-V
- Windows Server 2012 Hyper-V Manager

Cisco UCS Manager

Cisco UCS Manager provides unified, centralized, embedded management of all Cisco Unified Computing System software and hardware components across multiple chassis and thousands of virtual machines. Administrators use this software to manage the entire Cisco UCS as a single logical entity through an intuitive GUI, a command-line interface (CLI), or an XML API.

The Cisco UCS Manager resides on a pair of Cisco UCS 6200 Series Fabric Interconnects using a clustered, active-standby configuration for high availability. The software gives administrators a single interface for performing server provisioning, device discovery, inventory, configuration, diagnostics, monitoring, fault detection, auditing, and statistics collection. Cisco UCS Manager service profiles and templates support versatile role- and policy-based management, and system configuration information can be exported to configuration management databases (CMDBs) to facilitate processes based on IT Infrastructure Library (ITIL) concepts.

Service profiles let server, network, and storage administrators treat Cisco UCS servers as raw computing capacity to be allocated and reallocated as needed. The profiles define server I/O properties and are stored in the Cisco UCS 6200 Series Fabric Interconnects. Using service profiles, administrators can provision infrastructure resources in minutes instead of days, creating a more dynamic environment and more efficient use of server capacity.

Each service profile consists of a server software definition and the server's LAN and SAN connectivity requirements. When a service profile is deployed to a server, Cisco UCS Manager automatically configures the server, adapters, fabric extenders, and fabric interconnects to match the configuration specified in the profile. The automatic configuration of servers, network interface cards (NICs), host bus adapters (HBAs), and LAN and SAN switches lowers the risk of human error, improves consistency, and decreases server deployment times.

Service profiles benefit both virtualized and non-virtualized environments. The profiles increase the mobility of non-virtualized servers, such as when moving workloads from server to server or taking a server offline for service or an upgrade. Profiles can also be used in conjunction with virtualization clusters to bring new resources online easily, complementing existing virtual machine mobility.

For more Cisco UCS Manager information, visit: http://www.cisco.com/en/US/products/ps10281/index.html

Cisco UCS PowerTool

Cisco UCS PowerTool is a PowerShell module that helps automate all aspects of Cisco UCS Manager including server, network, storage and hypervisor management. PowerTool enables easy integration with existing IT management processes and tools.

Cisco UCS PowerTool is a flexible and powerful command line toolkit that includes more than 1500 PowerShell cmdlets providing customers with an efficient, cost effective and easy to use interface to integrate and automate UCS management with Microsoft products and many 3rd-party products, Cisco UCS PowerTool lets you take advantage of the flexible and powerful scripting environment offered by Microsoft PowerShell.

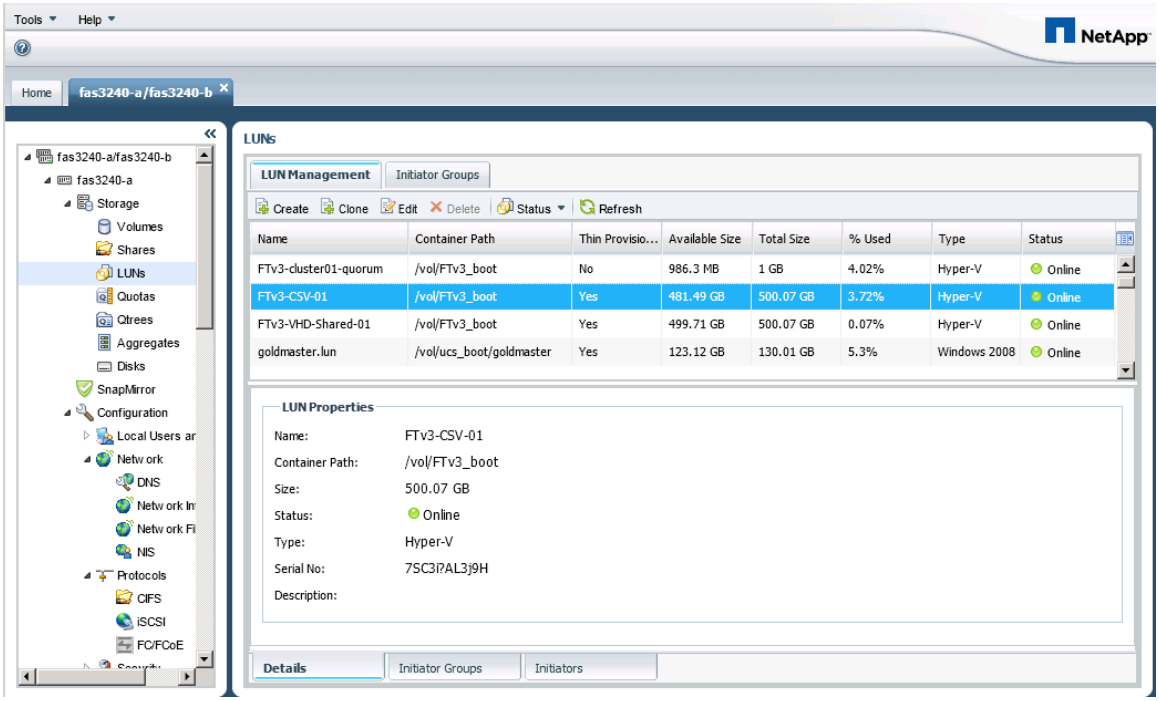
Cisco VM-FEX Port Profile Configuration Utility

The Cisco VM-FEX Port Profile Configuration Utility maps a UCS port profile to the virtual switch port that connects a virtual machine NIC to a virtual function. This utility is available in an MMC Snap-In and a set of PowerShell comdlets.

NetApp OnCommand System Manager

NetApp OnCommand System Manager makes it possible for administrators to manage individual or clusters of NetApp storage systems through an easy-to-use browser-based interface. System Manager comes with wizards and workflows, simplifying common storage tasks such as creating volumes, LUNs, qtrees, shares, and exports, which saves time and prevents errors. System Manager works across all NetApp storage: FAS2000, FAS3000, and FAS6000 series as well as V-Series systems. Figure 4 shows a sample screen in NetApp OnCommand System Manager.

Figure 4. NetApp OnCommand System Manager Example



NetApp SnapDrive for Windows

NetApp SnapDrive for Windows is an enterprise-class storage and data management application that simplifies storage management and increases availability of application data. The key functionality includes storage provisioning, file system–consistent data Snapshot copies, rapid application recovery, and the ability to manage data easily. SDW complements the native file system and volume manager and integrates seamlessly with the clustering technology supported by the host OS.

NetApp SnapManger for Hyper-V

NetApp SnapManager for Hyper-V automates and simplifies backup and restore operations for VMs running in Microsoft Windows Server 2012 Hyper-V environments hosted on Data ONTAP storage systems. SMHV enables application-consistent dataset backups according to protection policies set by the storage administrator. VM backups can also be restored from those application-consistent backups.

SnapManager for Hyper-V makes it possible to back up and restore multiple VMs across multiple hosts. Policies can be applied to the datasets to automate backup tasks such as scheduling, retention, and replication.

A Closer Look at FlexPod Discrete Uplink Design

Physical Build:Hardware and Software Revisions

The following table describes the hardware and software versions used during validation. It is important to note that Cisco, NetApp and Microsoft have interoperability matrixes that should be referenced to determine support for any one specific implementation of FlexPod. Please refer to the following links:

- NetApp Interoperability Matrix Tool
- Cisco UCS Hardware and Software Interoperability Tool
- Windows Server 2012 Catalog

Table 3. Validated Software and Firmware Versions

Layer	Device	Image	Comments
Compute	Cisco UCS Fabric Interconnects 6200 Series	2.1(1b)	Includes UCSM
	Cisco UCS 5108 Chassis	2.1(1b)	Includes the UCS-IOM 2208XP
	Cisco Nexus 2232 Fabric Extender	2.1(1b)	
	Cisco UCS B200 M3 and UCS C220 M3	2.1(1b)	B200 M3 using Cisco UCS VIC 1240 C220 M3 using Cisco VIC 1225
	Cisco E-NIC	2.2.0.13	Ethernet NIC Driver
	Cisco F-NIC	2.2.0.17	HBA Driver
	Cisco VMFex Switch	2.2.0.11	VM-FEX Forwarding Extensions for Hyper-V Virtual Switch
Network	Cisco Nexus 5000 NX-OS	5.2(1)N1(2a)	
Storage	NetApp FAS Model 3240-AE	ONTAP 8.1.2	
Software	Windows Server 2012	6.02.9200	
	Cisco UCS PowerTool	0.9.11.0	
	Cisco UCS VM-FEX Port Profile Manager	2.3.0.2	
	NetApp OnCommand	2.0.2	
	NetApp SnapDrive for Windows	6.5	
	NetApp SnapManger for Hyper-V	1.2	

Logical Build

Figure 2 above illustrates the Discrete Uplink design structure. The design is physically redundant across the stack addressing Layer 1 high availability requirements, but there are additional Cisco and NetApp technologies and features that make for an even more effective solution. This section of the document discusses the logical configuration validated for FlexPod. The topics covered include:

- FlexPod—Discrete Uplink Design with Data ONTAP Operating in 7-Mode

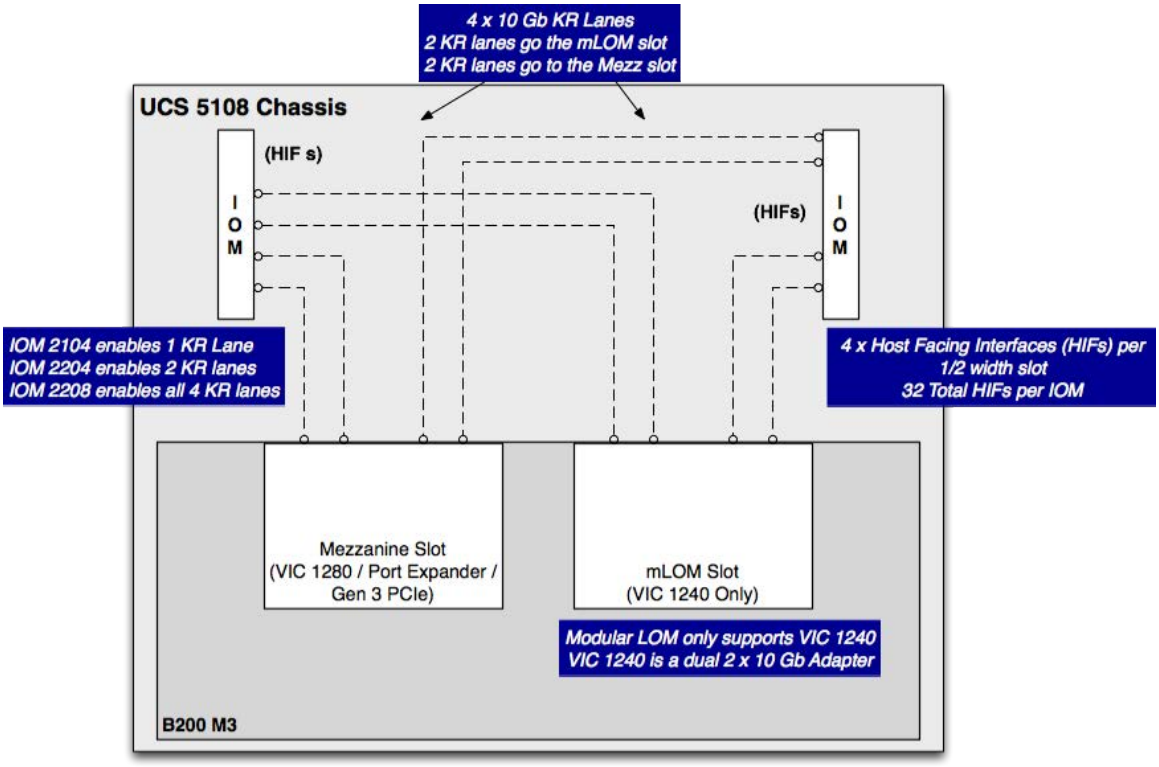
FlexPod allows organizations to adjust the individual components of the system to meet their particular scale or performance requirements. FlexPod continues this practice. One key design decision in the Cisco UCS domain is the selection of I/O components. There are numerous combinations of I/O adapter, IOM and Fabric Interconnect available so it is important to understand the impact of these selections on the overall flexibility, scalability and resiliency of the fabric.

Figure 5 illustrates the available backplane connections in the UCS 5100 series chassis. As the illustration shows, each of the two fabric extenders (I/O module) has four 10GBASE KR (802.3ap) standardized Ethernet backplane paths available for connection to the half-width blade slot. This means that each half-width slot has the potential to support up to 80Gb of aggregate traffic. What is realized depends on several factors namely:

- Fabric Extender model (2204 or 2208)
- Modular Lan on Motherboard (mLOM) card
- Mezzanine Slot card

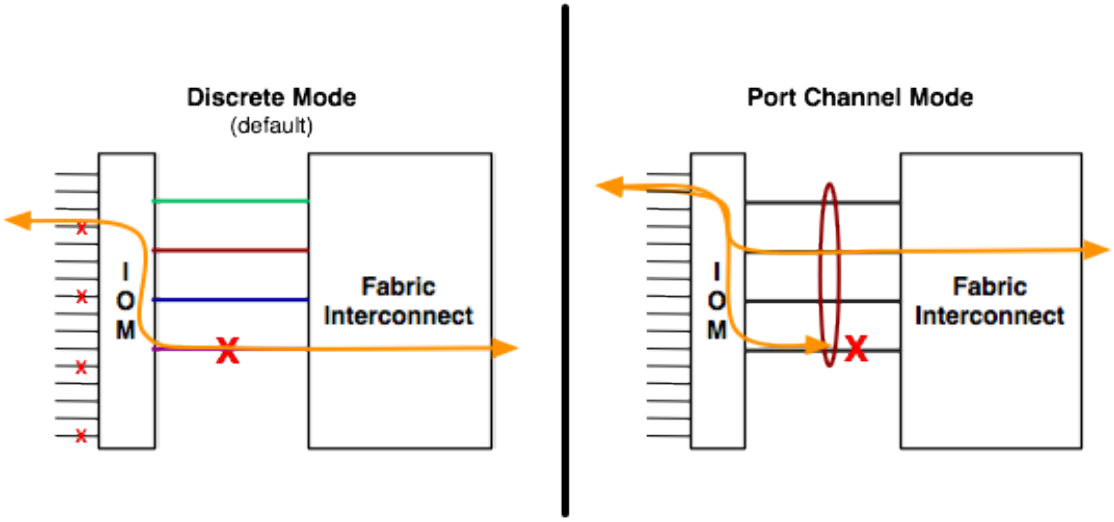
The Cisco UCS 2208XP Series Fabric Extenders have eight 10 Gigabit Ethernet, FCoE-capable, Enhanced Small Form-Factor Pluggable (SFP+) ports that connect the blade chassis to the fabric interconnect. The Cisco UCS 2204 has four external ports with identical characteristics to connect to the fabric interconnect. Each Cisco UCS 2208XP has thirty-two 10 Gigabit Ethernet ports connected through the midplane KR lanes to each half-width slot in the chassis, while the 2204XP has 16. This means the 2204XP enables two KR lanes per half-width blade slot while the 2208XP enables all four. The number of KR lanes indicates the potential I/O available to the chassis and therefore blades.

Figure 5. Cisco UCS B-Series M3 Server Chassis Backplane Connections



Port aggregation is supported by the second-generation Cisco UCS 6200 Series Fabric Interconnects, 2200 series Fabric Extenders and 1200 Series Virtual Interface Cards (VIC) support port aggregation. This capability allows for workload rebalancing between these devices providing link fault tolerance in addition to increased aggregate bandwidth within the fabric. It should be noted that in the presence of second generation VICs and Fabric Extenders fabric port channels will automatically be created in the fabric. Fabric port channels between the fabric extenders and fabric interconnects are controlled via the Chassis/FEX discovery policy. Figure 6 illustrates the two modes of operation for this policy. 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 the case of a fabric port channel discovery policy, the failure of a network fabric link allows for redistribution of flows across the remaining port channel members. This is less disruptive to the fabric.

Figure 6. Example of Discrete Mode Versus Port Channel Mode



Note: First generation Cisco UCS hardware is compatible with the second-generation gear but it will only operate in discrete mode.

Figure 7 represents one of the Cisco UCS B200-M3 backplane connections validated for the FlexPod. The B200M3 uses a VIC 1240 in the mLOM slot with an empty mezzanine slot. The FEX 2204XP enables 2 KR lanes to the half-width blade while the global discovery policy dictates the formation of a fabric port channel.

Figure 7. Validated UCS Backplane Configurations—VIC 1240 Only

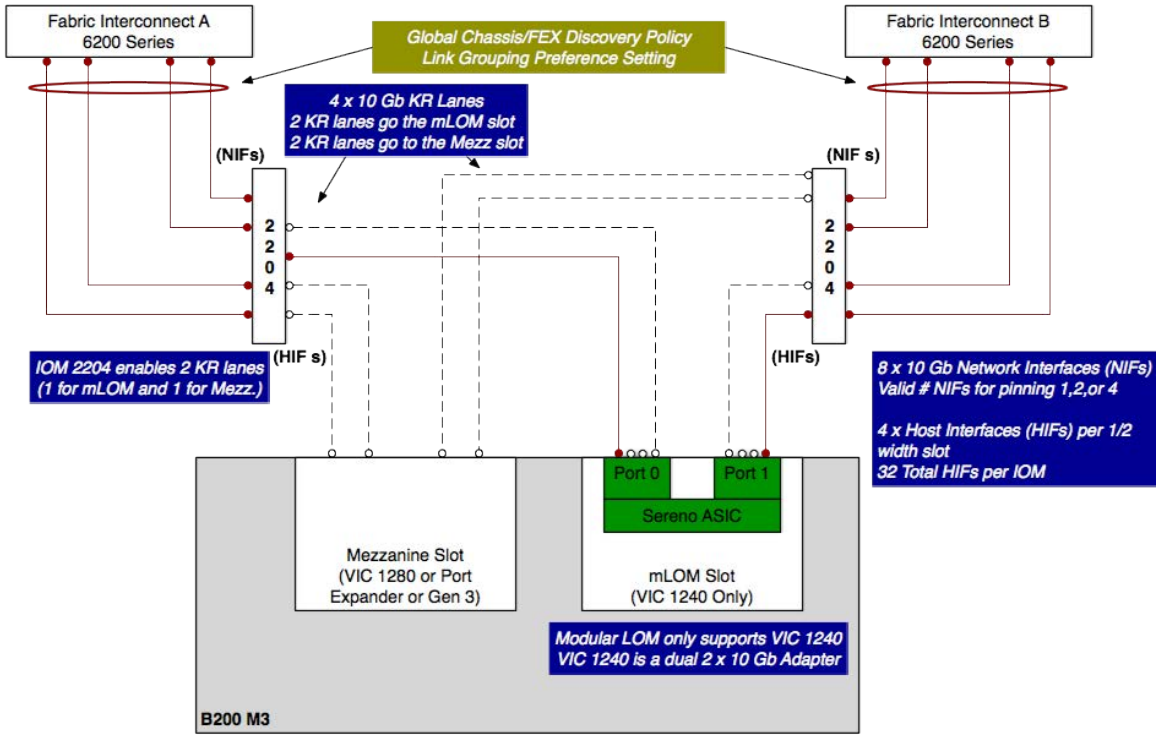
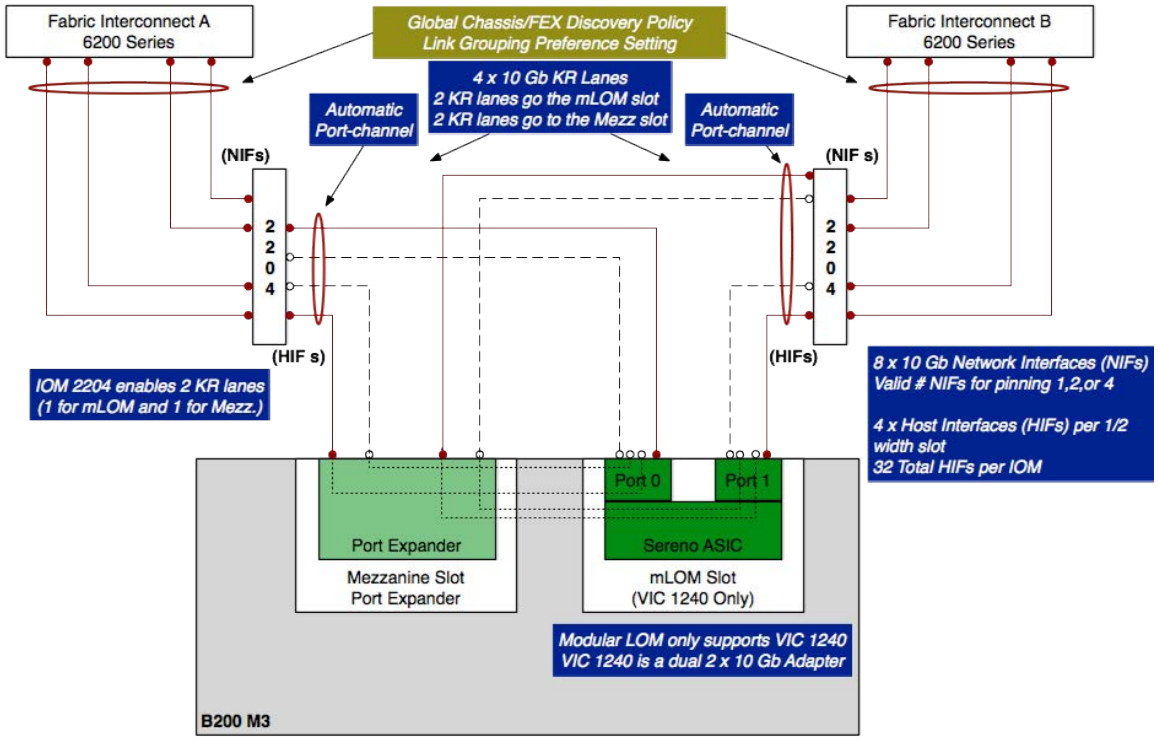


Figure 8 illustrates another Cisco UCS B200-M3 instance in the test bed. In this instance the mezzanine slot is populated with the port expander option. This passive device provides connectivity for the unused ports on the UCS VIC 1240, essentially enabling the 40-Gb potential of the mLOM card. Beyond the raw capacity improvements is the creation of two more automatic port channels between the Fabric Extender and the server. This provides link resiliency at the adapter level and double the bandwidth available to the system. (dual 2x10Gb).

Figure 8. Validated UCS Backplane Configuration—VIC1240 with Port Extender



Note: See the appendix for additional combinations of UCS second-generation hardware and the connectivity options they afford.

The FlexPod defines two FCoE port channels (Po1 & Po2) and two LAN port channels (Po13 & Po14). The FCoE port channels only carry only Fibre Channel traffic that is associated to a VSAN/VLAN set, with the set in turn supported only on one side of the fabric A or B. As in this example, the vHBA “FABRIC-A” is defined in the service profile. The vHBA uses a virtual circuit, VC 737, to traverse the Cisco UCS unified fabric to port channel Po1 where FCoE traffic egresses the UCS domain and enters the Cisco Nexus 5500 platform. Fabric A supports a distinct VSAN, which is not present on Fabric B, thus maintaining fabric isolation.

It has been said that design is the art of compromise; however with the FlexPod architecture there is very little sacrifice. Availability and performance are present the question becomes what combination meets the application and business requirements of the organization. Table 4 describe the availability and performance aspects of the second-generation UCS I/O gear.

Table 4. B-Series M3 FEX 2204XP and 2208XP Options

Reliability Technique	Fabric Failover & Adapter Redundancy & Port Channel				VIC 1240 & VIC 1280
	Fabric Failover & Adapter Redundancy		VIC 1240 & VIC 1280		
	Fabric Failover & Port Channel		VIC 1240 with Port Expander		VIC 1240 with Port Expander
	Fabric Failover	VIC 1240	VIC 1240		
		20Gb	40Gb	60Gb	80Gb
		Aggregate Bandwidth (Performance)			

*Dark Gray shading indicates the FEX 2208XP is in use. All other values are based on the FEX 2204XP model.

Note: Table 4 assumes presence of Cisco UCS 6200 series Fabric Interconnects.

Note: Third-Party Gen-3 PCIe adapters were not validated.

A balanced fabric is critical within any data center environment. Given the myriad of traffic types (Live Migration, CSV, FCoE, Public, control traffic, etc..) the FlexPod must be able to provide for specific traffic requirements while simultaneously be able to absorb traffic spikes and protect against traffic loss. To address these requirements the Cisco UCS QoS system classes and Cisco Nexus QoS policies should be configured. In this validation effort the FlexPod was configured to support jumbo frames with an MTU size of 9000. This class was assigned to the Best-Effort class. In regards to Jumbo frames it is important to make sure MTU settings are applied uniformly across the stack to prevent fragmentation and the negative performance implications inconsistent MTUs may introduce.

Cisco Unified Computing System—C-Series Server Design

Cisco UCS Manager 2.1 provides two connectivity modes for Cisco UCS C-Series Rack-Mount Server management. The following are the two connectivity modes:

- **Dual-wire management (Shared LOM):** This management mode is supported in the Cisco UCS Manager releases earlier than 2.1. Shared LAN on Motherboard (LOM) ports on the rack server are used exclusively for carrying management traffic. A separate cable connected to one of the ports on the PCIe card carries the data traffic. Using two separate cables for managing data traffic and management traffic is also referred to as dual-wire management.
- **Single-wire management (Sideband):** Cisco UCS Manager release version 2.1 introduces an additional rack server management mode using Network Controller Sideband Interface (NC-SI). Cisco UCS VIC1225 Virtual Interface Card (VIC) uses the NC-SI, that can carry both data traffic and management traffic on the same cable. This new feature is referred to as single-wire management and will allow for denser server to FEX deployments.

Note: The FlexPod Distinct Uplink design is capable of supporting both single and dual wire management. In the lab both implementations were used but at the time of this CVD the NetApp IMT only supports the dual-wire option. Please verify single-wire support with your trusted advisors or via the NetApp IMT tool.

Figure 9 illustrates the connectivity of the Cisco UCS C-Series server into the UCS domain. From a functional perspective the 1 RU Nexus FEX 2232PP replaces the UCS 2204 or 2208 IOM that are located with the UCS 5108 blade chassis. Each 10 Gigabit Ethernet VIC port connects to Fabric A or B via the FEX. The FEX and Fabric Interconnects form port channels automatically based on the chassis discovery policy providing a link resiliency to the C-Series server. This is identical to the behavior of the IOM to fabric interconnect connectivity. From a logical

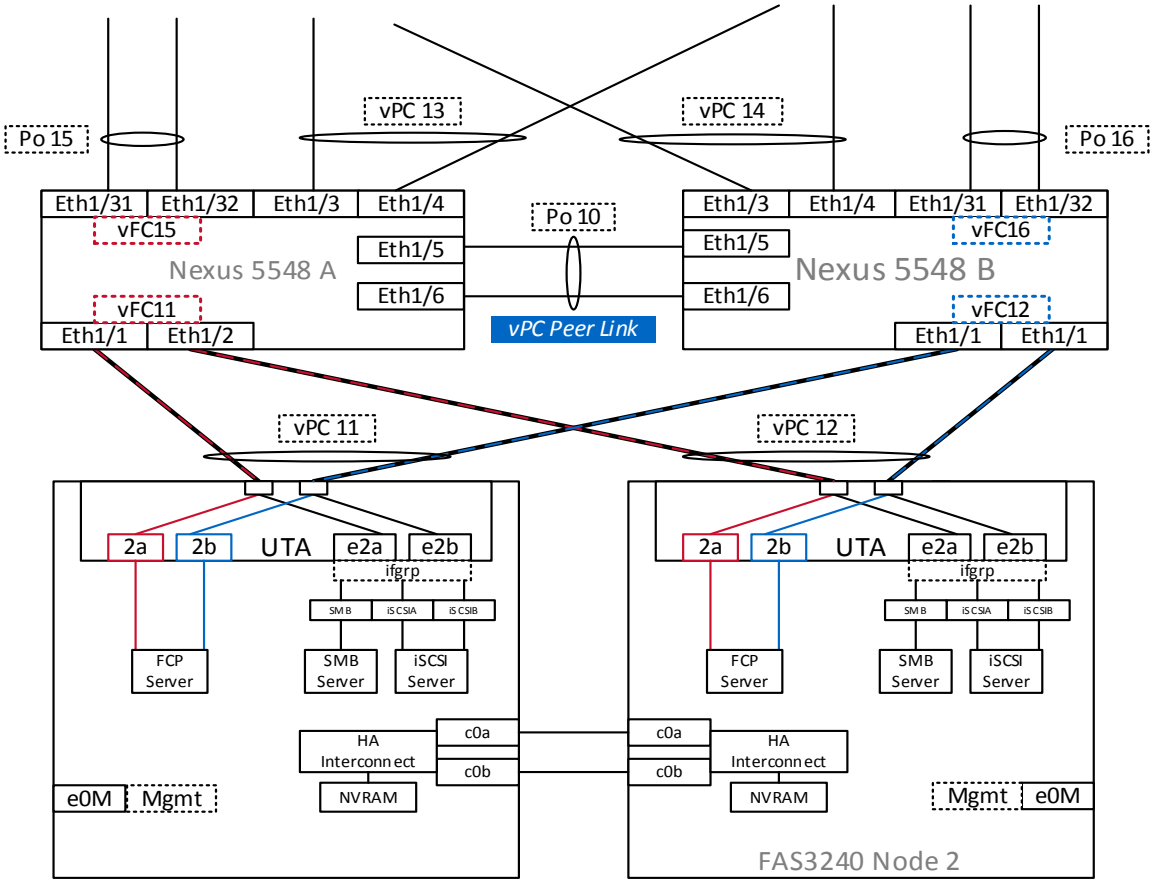
perspective the virtual circuits formed within the UCS domain are consistent between B-Series and C-Series deployment models and the virtual constructs formed at the Hyper-V.

Cisco Nexus 5500 Series Switch

As Figure 9 shows, the Cisco Nexus 5500 Series Switch provides Ethernet and in particular FCoE connectivity for the Cisco UCS domain as well as for the NetApp storage controllers. From an Ethernet perspective, the Nexus 5500 uses virtual PortChannel (vPC) allowing links that are physically connected to two different Cisco Nexus 5000 Series devices to appear as a single PortChannel to a third device in this case the UCS Fabric Interconnects and NetApp controllers. vPC provides the following benefits:

- Allows a single device to use a port channel across two upstream devices
- Eliminates Spanning Tree Protocol blocked ports
- Provides a loop-free topology
- Uses all available uplink bandwidth
- Provides fast convergence if either the link or a device fails
- Provides link-level resiliency
- Helps ensure high availability

Figure 9. Discrete Uplink Design: Nexus 5500 and NetApp Storage Focus



vPC requires a “peer link” which is documented as port channel 10 in Figure 10. It is important to note that the VLAN associated with the FCoE traffic does not traverse this peer link. Remember that the FCoE VLAN is associated or mapped to a VSAN typically using the same numeric ID. It is crucial that the fabrics do not mix, maintaining SAN A/B isolation best practices. In addition, the vPC links facing the UCS fabric interconnects, vPC13 and vPC14, do not carry any FCoE traffic. Do not define any FCoE VLANs on these links. However, the vPCs connected to the NetApp UTA’s are converged supporting both FCoE and all other VLANs associated with LAN protocols.

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 the Figure 10.)

Each Cisco Nexus 5500 Series Switch defines a port channel dedicated to FCoE and connected to the UCS Fabric Interconnects, in this instance Po15 and Po16. Each discrete port channel supports a single VLAN associated with Fabric A or Fabric B. A virtual Fiber Channel interface (vfc) is then bound to the logical port channel interface. This same construct is applied to the vPC’s facing the NetApp storage controllers, in this example vfc11 and vfc12. This assures universal accessibility of the fabric to each NetApp storage node in case of failures. To maintain SAN A/B isolation vfc 11 and 12 are associated to a different VLAN/VSAN pairing, meaning the vPCs facing the NetApp storage systems support all LAN and FCoE traffic but have unique FCoE VLANs defined on each Nexus switch.

Note: It is considered a best practice to name your vfc for the port channel it is residing on, for example vfc15 is on port channel 15.

The Nexus 5500 in the FlexPod design provides Fibre Channel services to the UCS and NetApp FAS platforms. Internally the Nexus 5500 platforms are performing FC zoning to enforce access policy between UCS-based initiators and FAS-based targets.

FlexPod is a converged infrastructure platform. This convergence is possible due to the support of Ethernet enhancements across the integrated compute stack in regards to bandwidth allocation and flow control based on traffic classification. As such it is important to implement these QoS techniques to ensure quality of service in the FlexPod.

- Priority Flow Control (PFC) 802.1Qbb—Lossless Ethernet using a PAUSE on a per Class of Service (CoS)
- Enhanced Transmission Selection (ETS) 802.1Qaz—Traffic Protection through bandwidth management
- Data Center Bridging Capability Exchange (DCBX)—Negotiates Ethernet functionality between devices (PFC, ETS and CoS values)

The Nexus 5500 supports these capabilities through QoS policy. QoS is enabled by default and managed using Cisco MQC (Modular QoS CLI) providing class based traffic control. The Nexus system will instantiate basic QoS classes for Ethernet traffic and a system FCoE class (class-fcoe) when the FCoE feature is enabled. It is important to align the QoS setting (CoS, MTU) within the Nexus 5500 and the UCS Fabric Interconnects. Realize that DCBX signaling can impact the NetApp controller be sure to allocate the proper bandwidth based on the sites application needs to the appropriate CoS classes and keep MTU settings consistent in the environment to avoid fragmentation issues and improve performance.

The following summarizes the best practices used in the validation of the FlexPod architecture:

- Nexus 5500 features enabled
  - Fibre Channel over Ethernet (FCoE) which uses the Priority Flow Control (802.1Qbb), Enhanced Transmission Selection (802.1Qaz) and Data Center Bridging eXchange (802.1Qaz) to provide a lossless fabric
  - N-Port ID Virtualization (NPIV) allows the network fabric port (N-Port) to be virtualized and support multiple fibre channel initiators on a single physical port
  - Link Aggregation Control Protocol (LACP part of 802.3ad)
  - Cisco Virtual Port Channeling (vPC) for link and device resiliency
  - Link Layer Discovery Protocol (LLDP) allows the Nexus 5000 to share and discover DCBX features and capabilities between neighboring FCoE capable devices
  - Enable 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 IP arp synchronization to optimize convergence across the vPC peer link

    Note: Cisco Fabric Services over Ethernet (CFSoE) is responsible for synchronization of configuration, Spanning Tree , MAC and VLAN information which removes the requirement for explicit configuration. The service is enabled by default.
  - 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
  - Ensure the path cost method is set to long. This setting accounts for 10Gbe Ethernet links in the environment
  - The spanning tree priority was not modified. The assumption being this is an access layer deployment
  - 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 5000 series switches configuration guides at http://www.cisco.com/en/US/products/ps9670/products_installation_and_configuration_guides_list.html.

Hyper-V

The Hyper-V role enables you to create and manage a virtualized computing environment by using virtualization technology that is built in to Windows Server 2012. Installing the Hyper-V role installs the required components and optionally installs management tools. The required components include Windows hypervisor, Hyper-V Virtual Machine Management Service, the virtualization Windows Management Interface (WMI) provider, and other virtualization components such as the virtual machine bus (VMbus), virtualization service provider (VSP) and virtual infrastructure driver (VID).

The management tools for the Hyper-V role consist of:

- GUI-based management tools: Hyper-V Manager, a Microsoft Management Console (MMC) snap-in, and Virtual Machine Connection, which provides access to the video output of a virtual machine so you can interact with the virtual machine.
- Hyper-V-specific cmdlets for Windows PowerShell. Windows Server 2012 includes a Hyper-V module, which provides command-line access to all the functionality available in the GUI, as well functionality not available through the GUI.

Windows Server 2012 introduced many new features and enhancements for Hyper-V. The following are some of the more notable enhancements that are used in this design.

- Host Scale–Up—greatly expands support for host processors and memory. New features include support for up to 64 virtual processors and 1 TB of memory for Hyper-V guests, a new VHDX virtual hard disk format with larger disk capacity of up to 64 TB, and additional resiliency. These features help ensure that your virtualization infrastructure can support the configuration of large, high-performance virtual machines to support workloads that might need to scale up significantly.
- Single Root Input/Output Virtualization (SR-IOV)–capable network devices and lets an SR-IOV virtual function of a physical network adapter be assigned directly to a virtual machine. This increases network throughput and reduces network latency while also reducing the host CPU overhead that is required for processing network traffic. The following figure shows the architecture of SR-IOV support in Hyper-V.

Cisco Virtual Machine Fabric Extender (VM-FEX)

Cisco Virtual Machine Fabric Extender (VM-FEX) is a technology that addresses both management and performance concerns in the data center by unifying physical and virtual switch management. The use of Cisco's VM-FEX collapses both virtual and physical networking into a single infrastructure, reducing the number of network management points and enabling consistent provisioning, configuration and management policy within the enterprise. This integration point between the physical and virtual domains of the data center allows administrators to efficiently manage both their virtual and physical network resources. The decision to use VM-FEX is typically driven by application requirements such as performance and the operational preferences of the IT organization.

The Cisco UCS Virtual Interface Card (VIC) offers each VM a virtual Ethernet interface or vNIC. This vNIC provides direct access to the Fabric Interconnects and Nexus 5500 series switches where forwarding decision can be made for each VM using a VM-FEX interface. Cisco VM-FEX technology for Hyper-V provides SR-IOV (Single Root I/O Virtualization) networking devices. SR-IOV works in conjunction with system chipset support for virtualization technologies. This provides remapping of interrupts and DMA and allows SR-IOV capable devices to be assigned directly to a virtual machine. Hyper-V in Windows Server 2012 enables support for SR-IOV–capable network devices and allows an SR-IOV virtual function of a physical network adapter to be assigned directly to a virtual

machine. This increases network throughput and reduces network latency, while also reducing the host CPU overhead required for processing network traffic.

For more information on the configuration limits associated with VM-FEX go to http://www.cisco.com/en/US/partner/docs/unified_computing/ucs/sw/configuration_limits/2.1/b_UCS_Configuration_Limits_2_1.html

FlexPod—Discrete Uplink Design with Data ONTAP Operating in 7-Mode

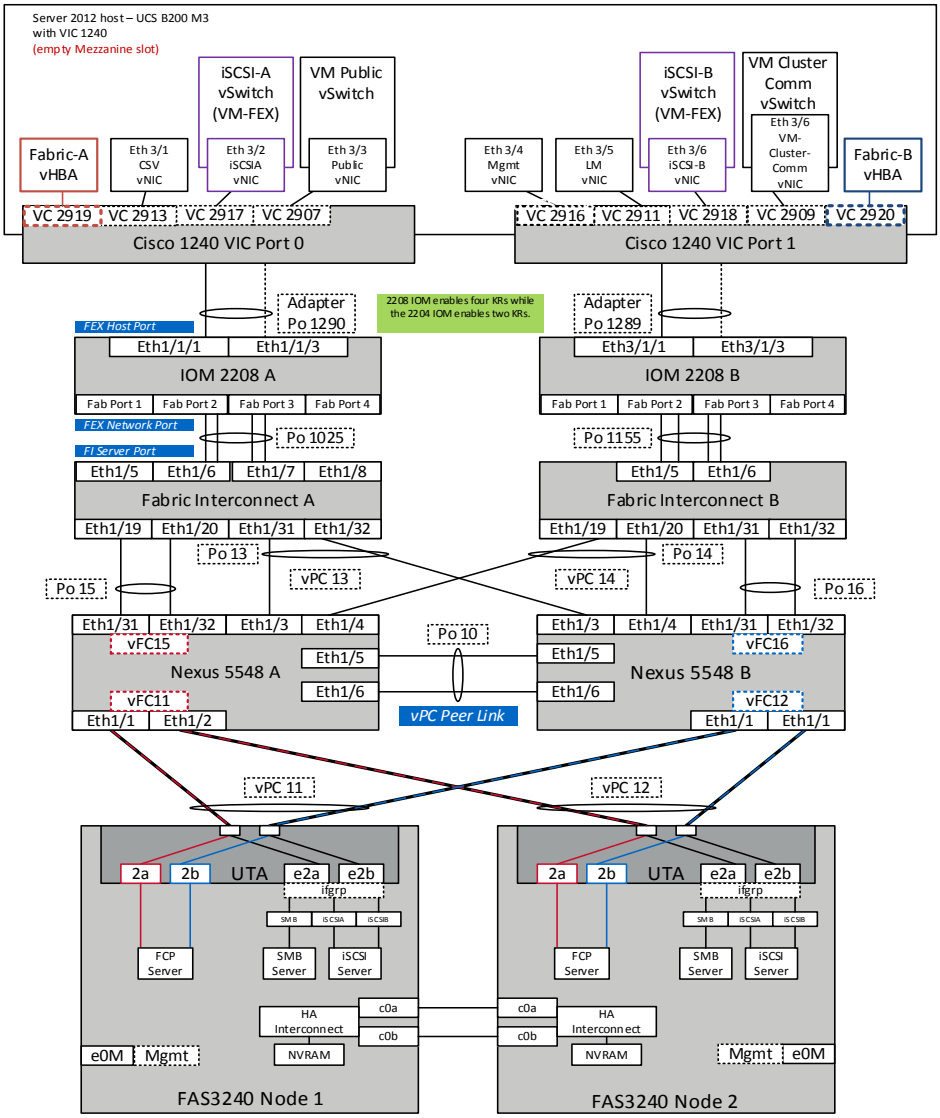
Figure 10 shows the FlexPod with Data ONTAP operating in 7-mode. 7-mode consists of only two storage controllers with shared media. The NetApp FAS controllers use redundant 10Gb converged adapters configured in a two-port interface group (ifgrp). Each port of the ifgrp is connected to one of the upstream switches, allowing multiple active paths by utilizing the Nexus vPC feature. IFGRP is a mechanism that allows the aggregation of a network interface into one logical unit. Combining links aids in network availability and bandwidth. NetApp provides three types of IFGRPs for network port aggregation and redundancy:

- Single Mode
- Static MultiMode
- Dynamic MultiMode

It is recommended to use Dynamic MultiMode IFGRPs due to the increased reliability and error reporting and is also compatible with Cisco Virtual Port Channels. A Dynamic MultiMode IFGRP uses Link Aggregation Control Protocol (LACP) to group multiple interfaces together to act as a single logical link. This provides intelligent communication between the storage controller and the Cisco Nexus switches and enables load balancing across physical interfaces as well as failover capabilities.

From a Fibre Channel perspective the SAN A(red in Figure 10) and SAN B (blue in Figure 10) fabric isolation is maintained across the architecture with dedicated FCoE channels and virtual interfaces. The 7-Mode design allocates Fibre Channel interfaces with SAN A and SAN B access for each controller in the HA pair.

Figure 10. Discrete Uplink Design with Data ONTAP Operating in 7-Mode.



Conclusion

FlexPod is the optimal shared infrastructure foundation on which to deploy a variety of IT workloads. Cisco and NetApp have created a platform that is both flexible and scalable for multiple use cases and applications. One common use case is to deploy Windows Server 2012 with Hyper-V as the virtualization solution, as described in this document. From virtual desktop infrastructure to Microsoft Exchange Server , Microsoft SharePoint Server, Microsoft SQL Server, and 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.

Appendix: Validated Bill of Materials

The following product information is provided for reference and will require modification depending on specific customer environments. Considerations include optics, cabling preferences, application workload and performance expectations.

Note: FlexPod architecture can support numerous Cisco and NetApp configurations that are not covered in this UCS BOM. Please work with your Cisco, NetApp or partner account team if you need assistance.

Table 5. NetApp Platform Bill of Materials

Product	Product Description	Qty

Table 6. Cisco Unified Computing Bill of Materials

1.1 Custom Name	SKU	Description	Qty
Blade Server UCSB-B200-M3	UCSB-B200-M3	UCS B200 M3 Blade Server w/o CPU, memory, HDD, mLOM/mezz	4
	UCS-CPU-E5-2690	2.90 GHz E5-2690/135W 8C/20MB Cache/DDR3 1600MHz	8
	UCS-MR-1X162RY-A	16GB DDR3-1600-MHz RDIMM/PC3-12800/dual rank/1.35v	64
	UCS-VIC-M82-8P	Cisco UCS VIC 1280 dual 40Gb capable Virtual Interface Card	4
	UCSB-MLOM-40G-01	Cisco UCS VIC 1240 modular LOM for M3 blade servers	4
	N20-BBLKD	UCS 2.5 inch HDD blanking panel	8
	UCSB-HS-01-EP	CPU Heat Sink for UCS B200 M3 and B420 M3	8
Chassis N20-C6508	N20-C6508	UCS 5108 Blade Svr AC Chassis/0 PSU/8 fans/0 fabric extender	2
	CAB-C19-CBN	Cabinet Jumper Power Cord, 0 VAC 16A, C20-C19 Connectors	8
	UCS-IOM-2204XP	UCS 2204XP I/O Module (4 External, 16 Internal 10Gb Ports)	4
	N01-UAC1	Single phase AC power module for UCS 5108	2
	N20-CAK	Access. kit for 5108 Blade Chassis including Railkit, KVM dongle	2
	N20-FAN5	Fan module for UCS 5108	16
	N20-PAC5-00W	00W AC power supply unit for UCS 5108	8
	N20-FW010	UCS 5108 Blade Server Chassis FW package	2
Fabric Interconnect UCS-FI-6248UP	UCS-FI-6248UP	UCS 6248UP 1RU Fabric Int/No PSU/32 UP/ 12p LIC	2
	UCS-ACC-6248UP	UCS 6248UP Chassis Accessory Kit	2
	UCS-PSU-6248UP-AC	UCS 6248UP Power Supply/100-240VAC	4
	N10-MGT010	UCS Manager v2.1	2
	CAB-9K12A-NA	Power Cord, 1VAC 13A NEMA 5-15 Plug, North America	2
	UCS-LIC-10GE	UCS 6200 Series ONLY Fabric Int 1PORT 1/10GE/FC-port license	20
	UCS-FAN-6248UP	UCS 6248UP Fan Module	4
	UCS-FI-DL2	UCS 6248 Layer 2 Daughter Card	2

1.2 Name	Catalog Num	Description	Qty
Nexus 5548UP	N5K-C5548UP-OSM.P	Nexus 5548UP Storage Solutions Bundle, Full Stor Serv Lic, OSM	2
	N55-48PO-SSK9.P	Nexus 5500 Storage License, 48 Ports, OSM	2
	N55-DL2.P	Nexus 5548 Layer 2 Daughter Card	2
	N55-M-BLNK.P	Nexus 5500 Module Blank Cover	2
	N55-PAC-750W.P	Nexus 5500 PS, 750W, Front to Back Airflow(Port-Side Outlet)	4
	N5548P-FAN.P	Nexus 5548P and 5548UP Fan Module, Front to Back Airflow	4
	CAB-C13-C14-2M.P	Power Cord Jumper, C13-C14 Connectors, 2 Meter Length	4
	N5548-ACC-KIT.P	Nexus 5548 Chassis Accessory Kit	2
	N5KUK9-521N1.1.P	Nexus 5000 Base OS Software Rel 5.2(1)N1(1)	2

References

Cisco Unified Computing System

Cisco UCS 6200 Series Fabric Interconnects

Cisco UCS 5100 Series Blade Server Chassis

Cisco UCS B-Series Blade Servers

Cisco UCS Adapters

Cisco UCS Manager

Cisco Nexus 5000 Series Switches

Authors

John George, Reference Architect, Infrastructure and Cloud Engineering, NetApp

John George is a Reference Architect in the NetApp Infrastructure and Cloud Engineering team and is focused on developing, validating, and supporting cloud infrastructure solutions that include NetApp products. Before his current role, he supported and administered Nortel's worldwide training network and VPN infrastructure. John holds a Master's degree in computer engineering from Clemson University.

Mike Mankovsky, Cisco Systems

Mike Mankovsky is a Cisco Unified Computing System architect, focusing on Microsoft solutions with extensive experience in Hyper-V, storage systems, and Microsoft Exchange Server. He has expert product knowledge in Microsoft Windows storage technologies and data protection technologies.

Chris O'Brien, Technical Marketing Engineer, Server Access Virtualization Business Unit, Cisco Systems

Chris O'Brien is currently focused on developing infrastructure best practices and solutions that are designed, tested, and documented to facilitate and improve customer deployments. Previously, O'Brien was an application developer and has worked in the IT industry for more than 15 years.

Chris Reno, Reference Architect, Infrastructure and Cloud Engineering, NetApp

Chris Reno is a reference architect in the NetApp Infrastructure and Cloud Enablement group and is focused on creating, validating, supporting, and evangelizing solutions based on NetApp products. Before being employed in his current role, he worked with NetApp product engineers designing and developing innovative ways to perform Q&A for NetApp products, including enablement of a large grid infrastructure using physical and virtualized compute resources. In these roles, Chris gained expertise in stateless computing, netboot architectures, and virtualization.

Glenn Sizemore, NetApp

Glenn Sizemore is a Private Cloud Reference Architect in the Microsoft Solutions Group at NetApp, where he specializes in Cloud and Automation. Since joining NetApp, Glenn has delivered a variety of Microsoft based solutions ranging from general best practice guidance to co-authoring the NetApp Hyper-V Cloud Fast Track with Cisco reference architecture.

Lindsey Street, Systems Architect, Infrastructure and Cloud Engineering, NetApp

Lindsey Street is a systems architect in the NetApp Infrastructure and Cloud Engineering team. She focuses on the architecture, implementation, compatibility, and security of innovative vendor technologies to develop competitive and high-performance end-to-end cloud solutions for customers. Lindsey started her career in 2006 at Nortel as an interoperability test engineer, testing customer equipment interoperability for certification. Lindsey has her Bachelors of Science degree in Computer Networking and her Master's of Science in Information Security from East Carolina University.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C07-727095-00 03/13