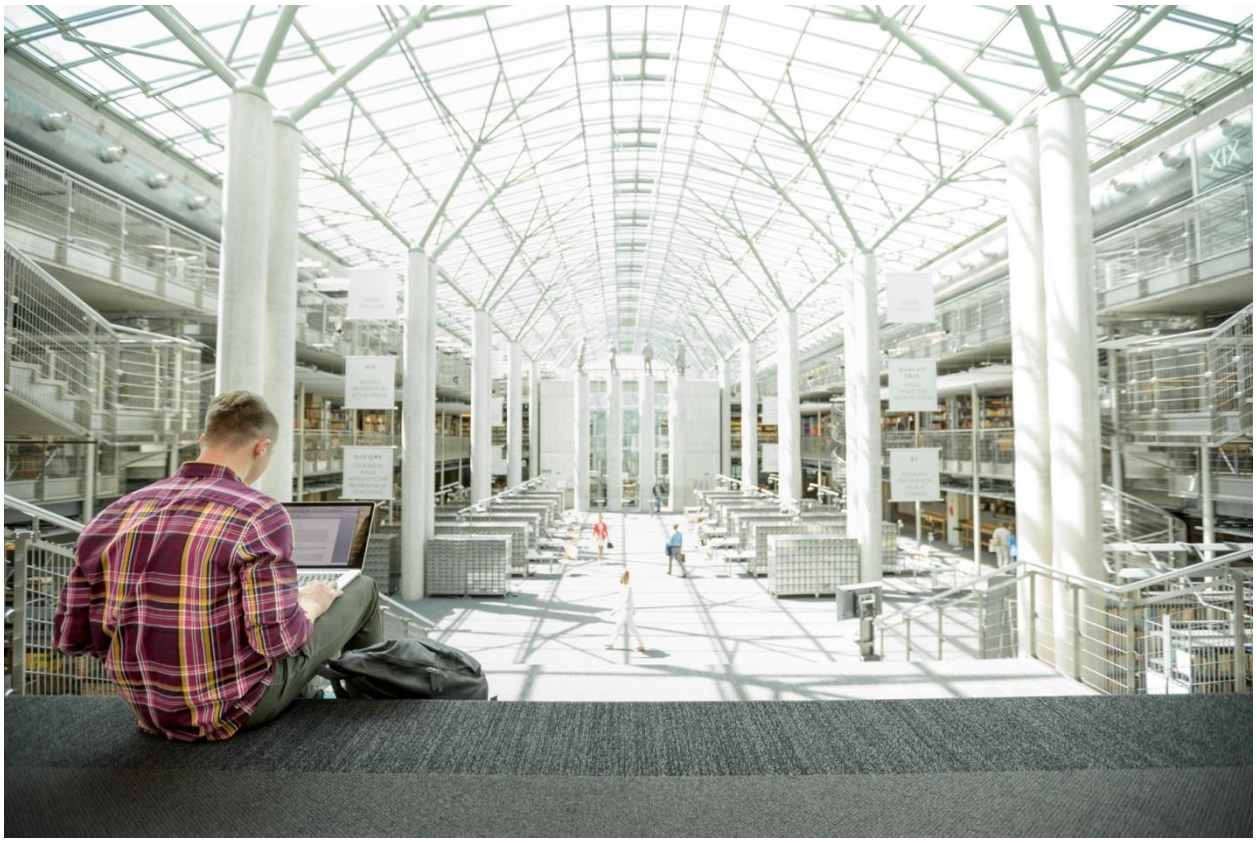


SmartStack Design Guide with Cisco UCS and Nimble CS700 Hybrid Array

Last Updated: October 11, 2016

About Cisco Validated Designs

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 (CVD) are systems and solutions that have been designed, tested and documented to facilitate and accelerate 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 a customer. CVDs deliver a validated design, documentation and support information to guide customers from design to deployment.

Cisco and Nimble have partnered to deliver a series of SmartStack™ solutions that combine Cisco Unified Computing System servers, Cisco Nexus family of switches, and Nimble Storage arrays into a single, flexible architecture. SmartStack solutions are pre-designed, integrated and validated architectures for the data center.

Customers looking to solve business problems using shared data center infrastructure face a number of challenges. A perennial infrastructure challenge is to achieve the levels of IT agility and efficiency that is necessary to meet business objectives. Addressing these challenges requires having an optimal solution with the following characteristics:

- **Availability**: Helps ensure applications and services are accessible at all times with no single point of failure
- **Flexibility**: Ability to support new services without requiring infrastructure modifications
- **Efficiency**: Facilitate efficient operation of the infrastructure through re-usable policies and API management
- **Manageability**: Ease of deployment and management to minimize operating costs
- **Scalability**: Ability to expand and grow with some degree of investment protection
- **Compatibility**: Minimal risk by ensuring optimal design and compatibility of integrated components

SmartStack enables a data center platform with the above characteristics by delivering an integrated architecture that incorporates compute, storage and network design best practices. SmartStack minimizes IT risks by testing the integrated architecture to ensure compatibility between the integrated components. SmartStack also addresses IT pain points by providing documented design guidance, deployment guidance and support that can be used in all stages (planning, designing and implementation) of a deployment.

The SmartStack solution outlined in this document delivers a converged infrastructure platform designed for Enterprise and Cloud data centers. SmartStack incorporates compute, network and storage best practices to deliver a resilient, scalable and flexible data center architecture. The design uses Cisco UCS servers for compute, VMware vSphere 6.0U1 hypervisor, Cisco Nexus and MDS switches for network and Fibre Channel-attached Nimble CS700 hybrid array for storage.

Solution Overview

Introduction

SmartStack is pre-designed, validated integrated infrastructure architecture for the data center. SmartStack solution portfolio combines Nimble® Storage arrays, Cisco® UCS servers, Cisco MDS fabric switches and Cisco Nexus switches into a single, flexible architecture. SmartStack solutions are designed and validated to minimize deployment time, project risk, and overall IT costs.

SmartStack is designed for high availability, with no single points of failure while maintaining cost-effectiveness and flexibility in design to support a variety of workloads in Enterprise and cloud data centers. SmartStack design can support different hypervisor options, bare metal and also be sized and optimized to support different use cases and requirements.

This document describes the SmartStack® integrated infrastructure solution for Enterprise and cloud deployments built using Cisco UCS, VMware vSphere 6.0U1, Cisco Nexus and MDS switches, and Fibre Channel-attached Nimble CS700 array.

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.

SmartStack Program Benefits

The SmartStack program is the result of a joint partnership between Cisco and Nimble Storage to deliver a series of infrastructure and application solutions optimized and validated on Cisco UCS, Nimble Storage and Cisco Nexus switches. Customers must use Cisco UCS, Nimble Storage and one of the approved application stacks to be a valid SmartStack solution and they must also have valid support contracts with Cisco and Nimble Storage.

Each SmartStack solution will include (though not limited to) the following line of documentation:

- CVD Design Guide with the architecture, design and best practices validated
- CVD Deployment guide with the validated hardware and software details and implementation details to deploy the solution

Cisco and Nimble Storage have a solid, joint support program focused on SmartStack solution, from customer account and technical sales representatives to professional services and technical support engineers. The support alliance provided by Cisco and Nimble Storage provides customers and channel partners with direct access to technical experts who can collaborate with cross vendors and have access to the shared lab resources to resolve the potential issues.

Technology Overview

SmartStack is a data center architecture for Enterprise or Cloud deployments and uses the following infrastructure components for compute, network and storage:

- Cisco Unified Computing System (Cisco UCS)
- Cisco Nexus and MDS Switches
- Nimble Storage arrays

The validated SmartStack design covered in this document uses the following models of the above mentioned infrastructure components.

- Cisco UCS 5108 Blade Server chassis with 2200 Series Fabric Extenders (FEX)
- Cisco UCS B-series Blade servers
- Cisco UCS C-series rack mount servers
- Cisco UCS 6200 Series Fabric Interconnects (FI)
- Cisco Nexus 9300 Series Platform switches
- Cisco MDS 9100 Series Fabric switches
- Nimble CS700 Hybrid Array

The above components are integrated using component and design best practices to deliver a converged infrastructure for Enterprise and cloud data centers.

Other optional Cisco UCS components of the SmartStack solution are:

- Cisco UCS 6300 Series Fabric Interconnects is the 3rd generation Cisco® Unified Fabric for the Cisco UCS system, delivering both Ethernet and Fibre Channel connectivity at higher speeds than previous generation models. Cisco UCS 6300 FIs provide line-rate connectivity with low-latency, lossless, 40Gbps Ethernet and 16G Fibre Channel speeds for deploying high-capacity data centers. Fabric Interconnects are the backbone of the Cisco Unified Computing System, delivering a unified fabric with centralized management.
- Cisco UCS C-Series Rack Mount Server delivers unified computing innovations and benefits to rack server to reduce total cost of ownership and increase agility with performance and density to support a wide range of workloads.
- Cisco Nexus 1000V Switches are virtual machine access switches for VMware vSphere environments that provide full switching capabilities and Layer 4 through Layer 7 services to virtual machines.
- Cisco UCS Central provides a scalable management platform for managing multiple, globally distributed Cisco UCS domains with consistency by integrating with Cisco UCS Manager to provide global configuration capabilities for pools, policies, and firmware.
- Cisco UCS Performance Manager is purpose-built data center management solution that provides a single pane-of-glass visibility of a converged heterogeneous infrastructure data center for performance monitoring and capacity planning.

Cisco UCS 6300 Series FI and Cisco UCS Central were not validated in this SmartStack design.

The next section provides a technical overview of the compute, network, storage and management components of the SmartStack solution.

Cisco Unified Compute System

The Cisco Unified Computing System™ (Cisco UCS) is a next-generation data center platform that integrates computing, networking, storage access, and virtualization resources 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 where all resources are managed through a unified management domain.

The main components of the Cisco UCS are:

Compute - The system is based on an entirely new class of computing system that incorporates rack mount servers and blade servers based on Intel processors.

Network - The system is integrated onto a low-latency, lossless, 10-Gbps unified network fabric. This network foundation consolidates LANs, SANs, and high-performance computing networks which are separate networks today. The unified fabric lowers costs by reducing the number of network adapters, switches, and cables, and by decreasing the power and cooling requirements.

Virtualization - The system unleashes the full potential of virtualization by enhancing the scalability, performance, and operational control of virtual environments. Cisco security, policy enforcement, and diagnostic features are now extended into virtualized environments to better support changing business and IT requirements.

Storage access – Cisco UCS system provides consolidated access to both SAN storage and Network Attached Storage over the unified fabric. This provides customers with storage choices and investment protection. Also, the server administrators can pre-assign storage-access policies to storage resources, for simplified storage connectivity and management leading to increased productivity. SmartStack solutions can support either iSCSI or Fibre Channel based access. This design covers only Fibre Channel connectivity.

Management: The system uniquely integrates all the system components, enabling the entire solution to be managed as a single entity through Cisco UCS Manager software. Cisco UCS Manager provides an intuitive graphical user interface (GUI), a command-line interface (CLI), and a robust application-programming interface (API) to manage all system configuration and operations. Cisco UCS Manager helps in increasing the IT staff productivity, enabling storage, network, and server administrators to collaborate on defining service profiles for applications. Service profiles are logical representations of desired physical configurations and infrastructure policies. They help automate provisioning and increase business agility, allowing data center managers to provision resources in minutes instead of days.

The Cisco Unified Computing System in the SmartStack architecture consists of the following components:

- Cisco UCS Manager provides unified management of all software and hardware components in the Cisco Unified Computing System and manages servers, networking, and storage from a single interface.
- 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 5108 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's Unified Compute System has revolutionizing the way servers are managed in data-center. The next section takes a detailed look at the unique differentiators of Cisco UCS and Cisco UCS Manager.

Cisco UCS Differentiators

- Embedded Management — Servers in the system are managed by embedded software in the Fabric Interconnects, eliminating need for any external physical or virtual devices to manage the servers.
- Unified Fabric — There is a single Ethernet cable to the FI from the server chassis (blade or rack) 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 a blade server in the chassis or connecting a rack server to the FI, discovery and inventory of compute resource occurs automatically without any intervention. Auto-discovery combined with unified fabric enables the wire-once architecture of Cisco UCS, where compute capability of Cisco UCS can be extended easily without additional connections to the external LAN, SAN and management networks.
- Policy Based Resource Classification — Once a compute resource is discovered by Cisco UCS Manager, it can be automatically classified to a resource pool based on policies defined. This capability is useful in multi-tenant cloud computing. This CVD showcases the policy based resource classification of Cisco UCS Manager.
- Combined Rack and Blade Server Management — Cisco UCS Manager can manage B-series blade servers and C-series rack servers under the same Cisco UCS domain. Along with stateless computing, this feature makes compute resources truly agnostic to the hardware form factor.
- 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 from different domains, such as network, storage, security, server and virtualization the flexibility to work independently 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 logic enables automation friendly management APIs and provides great flexibility to owners of different organizations.
- Service Profiles and Stateless Computing — A service profile is a logical representation of a server, carrying its various identities and policies. This logical server can be assigned to any physical compute resource as far as it meets the resource requirements. Stateless computing enables procurement of a server within minutes, which used to take days in legacy server management systems.
- Built-in Multi-Tenancy Support — The combination of policies, pools and templates, loose referential integrity, policy resolution in organization hierarchy and a service profiles based approach to compute resources makes Cisco UCS Manager inherently friendly to multi-tenant environment typically observed in private and public clouds.
- Extended Memory — The enterprise-class Cisco UCS B200 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 v3 and v4 Series processor family CPUs with up to 1536 GB of RAM (using 64 GB DIMMs), two solid-state drives (SSDs) or hard disk drives (HDDs), and

up to 80 Gbps throughput connectivity 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.

Cisco UCS 5108 Blade Server Chassis

The Cisco UCS 5108 Blade Server Chassis is a fundamental building block of the Cisco Unified Computing System, delivering a scalable and flexible blade server architecture. A Cisco UCS 5108 Blade Server chassis is six rack units (6RU) high and can house up to eight half-width or four full-width Cisco UCS B-series blade servers.

For a complete list of blade servers supported, see: http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-b-series-blade-servers/index.html

There are four hot-swappable power supplies that are accessible from the front of the chassis. These power supplies are 94 percent efficient and can be configured to support non-redundant, N+1, and grid-redundant configurations. The rear of the chassis contains eight hot-swappable fans, four power connectors (one per power supply), and two I/O bays that can support Cisco UCS 2000 Series Fabric Extenders. The two fabric extenders can be used for both redundancy and bandwidth aggregation. A passive mid-plane provides up to 80 Gbps of I/O bandwidth per server slot and up to 160 Gbps of I/O bandwidth for two slots. The chassis is capable of supporting future 40 Gigabit Ethernet standards.

Cisco UCS 5108 blade server chassis uses a unified fabric and fabric-extender technology to simplify and reduce cabling by eliminating the need for dedicated chassis management and blade switches. The unified fabric also reduces TCO by reducing the number of network interface cards (NICs), host bus adapters (HBAs), switches, and cables that need to be managed, cooled, and powered. This architecture enables a single Cisco UCS domain to scale up to 20 chassis with minimal complexity.

For more information, see: http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-5100-series-blade-server-chassis/index.html

Figure 1 Cisco UCS 5108 Blade Server Chassis

Front View

Back View

Cisco UCS 6200 Series Fabric Interconnects

Cisco UCS Fabric Interconnects are a family of line-rate, low-latency, lossless 1/10 Gigabit Ethernet and Fibre Channel over Ethernet (FCoE), and 4/2/1 and 8/4/2 native Fibre Channel switches. Cisco UCS Fabric Interconnects are the management and connectivity backbone of the Cisco Unified Computing System. Each chassis or rack server connects to the FI using a single Ethernet cable for carrying all network, storage and management traffic. Cisco UCS Fabric Interconnects provide uniform network and storage access to servers and are typically deployed in redundant pairs.

Cisco UCS® Manager provides unified, embedded management of all software and hardware components of the Cisco Unified Computing System™ (Cisco UCS) across multiple chassis with blade servers, rack servers and thousands of virtual machines. The Cisco UCS Management software (Cisco UCS Manager) runs as an embedded device manager in a clustered pair fabric interconnects and manages the resources connected to it. An instance of Cisco UCS Manager with all Cisco UCS components managed by it forms a Cisco UCS domain, which can include up to 160 servers.

The Cisco UCS Fabric Interconnect family is currently comprised of the Cisco 6100 Series, Cisco 6200 Series and Cisco 6300 Series of Fabric Interconnects.

Cisco UCS 6248UP Fabric Interconnects were used for this CVD.

For more information, see: http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6200-series-fabric-interconnects/index.html

Figure 2 Cisco UCS 6248UP Fabric Interconnect

Cisco UCS Fabric Extenders

The Cisco UCS Fabric extenders multiplexes and forwards all traffic from servers in a chassis to a parent Cisco UCS Fabric Interconnect over from 10-Gbps unified fabric links. All traffic, including traffic between servers on the same chassis, or between virtual machines on the same server, is forwarded to the parent fabric interconnect, where network profiles and polices are maintained and managed by the Cisco UCS Manager. The Fabric extender technology was developed by Cisco. Up to two fabric extenders can be deployed in a Cisco UCS chassis.

The Cisco UCS Fabric Extender family currently comprises of Cisco UCS 2200 and Cisco Nexus 2000 Series of Fabric Extenders. The Cisco UCS 2200 Series Fabric Extenders come in two flavors as outlined below.

- The Cisco UCS 2204XP Fabric Extender has four 10 Gigabit Ethernet, FCoE-capable, SFP+ ports that connect the blade chassis to the fabric interconnect. Each Cisco UCS 2204XP has sixteen 10 Gigabit Ethernet ports connected through the mid-plane to each half-width slot in the chassis. Typically configured in pairs for redundancy, two fabric extenders provide up to 80 Gbps of I/O to the chassis.
- The Cisco UCS 2208XP Fabric Extender has eight 10 Gigabit Ethernet, FCoE-capable, SFP+ ports that connect the blade chassis to the fabric interconnect. Each Cisco UCS 2208XP has thirty-two 10 Gigabit Ethernet ports connected through the mid-plane to each half-width slot in the chassis. Typically configured in pairs for redundancy, two fabric extenders provide up to 160 Gbps of I/O to the chassis.

Cisco UCS 2208 Fabric Extenders were used for this CVD.

For more information, see: http://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-6200-series-fabric-interconnects/data_sheet_c78-675243.html

Figure 3 Cisco UCS 2208XP Fabric Extenders

Cisco UCS Manager

Cisco Unified Computing System (UCS) Manager provides unified, embedded management for all software and hardware components in the Cisco UCS. Using Cisco Single Connect technology, it manages, controls, and administers multiple chassis for thousands of virtual machines. Administrators use the software to manage the entire Cisco Unified Computing System 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.

Cisco UCS Manager offers unified embedded management interface that integrates server, network, and storage. Cisco UCS Manger performs auto-discovery to detect inventory, manage, and provision system components that

are added or changed. It offers comprehensive set of XML API for third part integration, exposes 9000 points of integration and facilitates custom development for automation, orchestration, and to achieve new levels of system visibility and control.

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

For more information on Cisco UCS Manager, see: http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-manager/index.html

Cisco UCS B-Series M4 Servers

Cisco UCS B200 M4 blade servers with Cisco Virtual Interface Card 1340 were used for this CVD.

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 uses the power of the latest Intel® Xeon® E5-2600 v3 and v4 Series processor family CPUs with up to 1536 GB of RAM (using 64 GB DIMMs), two solid-state drives (SSDs) or hard disk drives (HDDs), and up to 80 Gbps throughput connectivity. The Cisco UCS B200 M4 Blade Server mounts in a Cisco UCS 5100 Series blade server chassis or Cisco UCS Mini blade server chassis. It has 24 total slots for registered ECC DIMMs (RDIMMs) or load-reduced DIMMs (LR DIMMs) for up to 1536 GB total memory capacity. It supports one connector for the Cisco VIC 1340 or 1240 adapters, which provides Ethernet and FCoE.

For more information, see: http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-b200-m4-blade-server/index.html

Figure 4 Cisco UCS B200 M4 Blade Server

Cisco VIC 1340

The Cisco UCS Virtual Interface Card (VIC) 1340 is a 2-port 40-Gbps Ethernet or dual 4 x 10-Gbps Ethernet, FCoE-capable modular LAN on motherboard (mLOM) designed exclusively for the M4 generation of Cisco UCS B-Series Blade Servers. When used in combination with an optional port expander, the Cisco UCS VIC 1340 capabilities is enabled for two ports of 40-Gbps Ethernet. The Cisco UCS VIC 1340 enables a policy-based, stateless, agile server infrastructure that can present over 256 PCIe standards-compliant interfaces to the host that can be dynamically configured as either network interface cards (NICs) or host bus adapters (HBAs). In addition, the Cisco UCS VIC 1340 supports Cisco® Virtual Machine Fabric Extender (VM-FEX) technology, which extends the Cisco UCS Fabric interconnect ports to virtual machines, simplifying server virtualization deployment and management.

For more information, see: http://www.cisco.com/c/en/us/products/interfaces-modules/ucs-virtual-interface-card-1340/index.html

Figure 5 Cisco VIC 1340

Cisco UCS C-Series M4 Servers

Cisco UCS C220 M4 rack server was used for this CVD.

The enterprise-class Cisco UCS C220 M4 server extends the capabilities of the Cisco Unified Computing System (UCS) portfolio in a one rack-unit (1RU) form-factor. The Cisco UCS C220 M4 uses the power of the latest Intel® Xeon® E5-2600 v3 and v4 Series processor family CPUs with up to 1536 GB of RAM (using 64 GB DIMMs), 8 Small Form-Factor (SFF) drives or 4 Large Form-Factor (LFF) drives, and up to 80 Gbps throughput connectivity. The Cisco UCS C220 M4 Rack Server can be used standalone, or as integrated part of the Unified Computing System. It has 24 DIMM for up to 1536 GB total memory capacity. It supports one connector for the Cisco VIC 1225, 1227 or 1380 adapters, which provide Ethernet and FCoE.

For more information, see: http://www.cisco.com/c/en/us/products/collateral/servers-unified-computing/ucs-c220-m4-rack-server/datasheet-c78-732386.html.

Figure 6 Cisco UCS C220 M4 Rack Server

Cisco UCS Performance Manager

Cisco UCS Performance Manager can help unify the monitoring of critical infrastructure components, network connections, applications, and business services across dynamic heterogeneous physical and virtual data centers powered by Cisco UCS. Cisco UCS Performance Manager delivers detailed monitoring from a single customizable console. The software uses APIs from Cisco UCS Manager and other Cisco® and third-party components to collect data and display comprehensive, relevant information about your Cisco UCS integrated infrastructure.

Cisco UCS Performance Manager does the following:

- Unifies performance monitoring and management of Cisco UCS integrated infrastructure
- Delivers real-time views of fabric and data center switch bandwidth use and capacity thresholds
- Discovers and creates relationship models of each system, giving your staff a single, accurate view of all components
- Provides coverage for Cisco UCS servers, Cisco networking, storage, hypervisors, and operating systems
- Allows you to easily navigate to individual components for rapid problem resolution

An Express version of Cisco UCS Performance Manager is also available for Cisco UCS-based compute platform coverage only (physical and virtual). The Express version covers Cisco UCS servers, hypervisors, and operating systems.

Cisco UCS Performance Manager provided a scalable and highly available architecture. It can be deployed in a distributed environment across geographical locations as needed. The monitored Cisco devices in the SmartStack architecture includes the following:

- Cisco UCS domain components, including Cisco UCS Mini and Cisco UCS Central software
- Cisco Nexus® 1000V Switch (optional) and 9000 Series Switches
- Cisco MDS 9000 Series Multilayer Switches
- Cisco Nexus 2000 Series Fabric Extenders

Cisco Nexus 9000 Series Platform Switches

The Cisco Nexus 9000 family of switches offers both modular (9500 switches) and fixed (9300 switches) 1/10/40/100 Gigabit Ethernet switch configurations designed to operate in one of two modes:

- Application Centric Infrastructure (ACI) mode that uses an application centric policy model with simplified automation and centralized management
- Cisco NX-OS mode for traditional architectures – the SmartStack design in this document uses this mode

Architectural Flexibility

- Delivers high performance and density, and energy-efficient traditional 3-tier or leaf-spine architectures
- Provides a foundation for Cisco ACI, automating application deployment and delivering simplicity, agility, and flexibility
-

Scalability

- Up to 60-Tbps of non-blocking performance with less than 5-microsecond latency
- Up to 2304 10-Gbps or 576 40-Gbps non-blocking layer 2 and layer 3 Ethernet ports
- Wire-speed virtual extensible LAN (VXLAN) gateway, bridging, and routing

High Availability

- Full Cisco In-Service Software Upgrade (ISSU) and patching without any interruption in operation
- Fully redundant and hot-swappable components
- A mix of third-party and Cisco ASICs provide for improved reliability and performance

Energy Efficiency

- The chassis is designed without a mid-plane to optimize airflow and reduce energy requirements
- The optimized design runs with fewer ASICs, resulting in lower energy use
- Efficient power supplies included in the switches are rated at 80 Plus Platinum

Investment Protection

- Cisco 40-Gb bidirectional transceiver allows for reuse of an existing 10 Gigabit Ethernet cabling plant for 40 Gigabit Ethernet
- Designed to support future ASIC generations
- Easy migration from NX-OS mode to ACI mode

Cisco Nexus 9372PX platform switches were used in this CVD.

For more information, refer to: http://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/index.html

Cisco Nexus 1000v Series Switches

Cisco Nexus 1000V Series Switches provide a comprehensive and extensible architectural platform for virtual machine (VM) and cloud networking. Integrated into the VMware vSphere hypervisor and fully compatible with VMware vCloud® Director, the Cisco Nexus 1000V Series provides:

- Advanced virtual machine networking using Cisco NX-OS operating system. Capabilities include PortChannels (LACP), IEEE 802.1Q VLAN trunking, Jumbo Frame support and Virtual Extensible Local Area Network (VXLAN) for cloud deployments
- Cisco vPath technology for efficient and optimized integration of Layer 4-7 virtual networking services (For example, Firewall)
- Mobile virtual machine security and network policy. Advanced security capabilities include Strom Control, BPDU Guard, Dynamic Host Configuration Protocol (DHCP) snooping, IP source guard, Dynamic Address Resolution Protocol (ARP) Inspection, and Cisco TrustSec® security group access (SGA), Security Group Tagging (SGT) and Security Group ACL (SGACL) support.
- Non-disruptive operational model for your server virtualization and networking teams
- Policy-based virtual machine connectivity
- Quality of service (QoS)
- Monitoring: NetFlow, Switch Port Analyzer (SPAN), and Encapsulated Remote SPAN (ERSPAN)
- Easy deployment using Cisco Virtual Switch Update Manager (VSUM) which allows you to install, upgrade, monitor and also migrate hosts to Cisco Nexus 1000V using the VMware vSphere web client.
- Starting with Cisco Nexus 1000V Release 4.2(1)SV2(1.1), a plug-in for the VMware vCenter Server, known as vCenter plug-in (VC plug-in) is supported on the Cisco Nexus 1000V virtual switch. It provides the server administrators a view of the virtual network and a visibility into the networking aspects of the Cisco Nexus 1000V virtual switch. The server administrator can thus monitor and manage networking resources effectively with the details provided by the vCenter plug-in. The vCenter plug-in is supported only on VMware vSphere Web Clients where you connect to VMware vCenter through a browser. The vCenter plug-in is installed as a new tab in the Cisco Nexus 1000V as part of the user interface in vSphere Web Client.

For more information, see:

- http://www.cisco.com/en/US/products/ps9902/index.html
- http://www.cisco.com/en/US/products/ps10785/index.html

Cisco MDS 9100 Series Fabric Switches

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. Cisco MDS 9148S is powered by Cisco NX-OS and 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. Cisco MDS 9148S is well suited as a:

- Top-of-rack switch in medium-sized deployments

- Edge switch in a two-tiered (core-edge) data center topology
- Standalone SAN in smaller departmental deployments

The main features and benefits of Cisco MDS 9148S are summarized in the table below.

Table 1 Cisco MDS 9148S Features and Benefits

Features	Benefits
Up to 48 autosensing Fibre Channel ports are capable of speeds of 2, 4, 8, and 16 Gbps, with 16 Gbps of dedicated bandwidth for each port. Cisco MDS 9148S scales from 12 to 48 high-performance Fibre Channel ports in a single 1RU form factor.	High Performance and Flexibility at Low Cost
Supports dual redundant hot-swappable power supplies and fan trays, PortChannels for Inter-Switch Link (ISL) resiliency, and F-port channeling for resiliency on uplinks from a Cisco MDS 9148S operating in NPV mode.	High Availability Platform for Mission-Critical Deployments
Intelligent diagnostics/Hardware based slow port detection and Cisco Switched Port Analyzer (SPAN) and Remote SPAN (RSPAN) and Cisco Fabric Analyzer to capture and analyze network traffic. Fibre Channel ping and traceroute to identify exact path and timing of flows. Cisco Call Home for added reliability.	Enhanced performance and monitoring capability. Increase reliability, faster problem resolution, and reduce service costs
In-Service Software Upgrades	Reduce downtime for planned maintenance and software upgrades
Aggregate up to 16 physical ISLs into a single logical PortChannel bundle with multipath load balancing.	High performance ISLs and optimized bandwidth utilization
Virtual output queuing on each port by eliminating head-of-line blocking	Helps ensure line-rate performance
PowerOn Auto Provisioning to automate deployment and upgrade of software images.	Reduces administrative costs
Smart zoning for creating and managing zones	Reduces consumption of hardware resources and administrative time
SAN management through a command-line interface (CLI) or Cisco Prime DCNM for SAN Essentials Edition, a centralized management tool. Cisco DCNM task-based wizards simplify management of single or multiple switches and fabrics including management of virtual resources end-to-end, from the virtual machine and switch to the physical storage.	Simplified Storage Management with built-in storage network management and SAN plug-and-play capabilities. Sophisticated Diagnostics

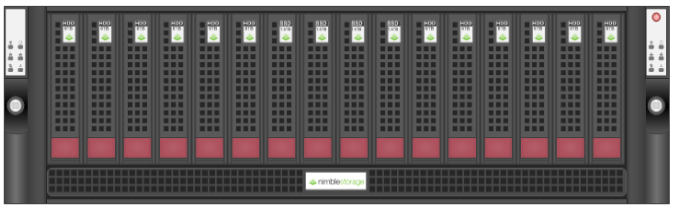
Fabric-wide per-VSAN role-based authentication, authorization, and accounting (AAA) services using RADIUS, Lightweight Directory Access Protocol (LDAP), Microsoft Active Directory (AD), and TACACS+. Also provides VSAN fabric isolation, intelligent, port-level packet inspection, Fibre Channel Security Protocol (FC-SP) host-to-switch and switch-to-switch authentication, Secure File Transfer Protocol (SFTP), Secure Shell Version 2 (SSHv2), and Simple Network Management Protocol Version 3 (SNMPv3) implementing Advanced Encryption Standard (AES). Other security features include control-plane security, hardware-enforced zoning, broadcast zones, and management access.	Comprehensive Network Security Framework
Virtual SAN (VSAN) technology for hardware-enforced, isolated environments within a physical fabric. Access control lists (ACLs) for hardware-based, intelligent frame processing. Advanced traffic management features, such as fabric-wide quality of service (QoS) and Inter-VSAN Routing (IVR) for resource sharing across vSANs. Zone-based QoS simplifies configuration and administration by using the familiar zoning concept.	Intelligent Network Services and Advanced Traffic Management for better and predictable network service without compromising scalability, reliability, availability, and network security
Common software across all platforms by using Cisco NX-OS and Cisco Prime DCNM across the fabric.	Reduce total cost of ownership (TCO) through consistent provisioning, management, and diagnostic capabilities across the fabric

Nimble Storage – CS700 Hybrid Array

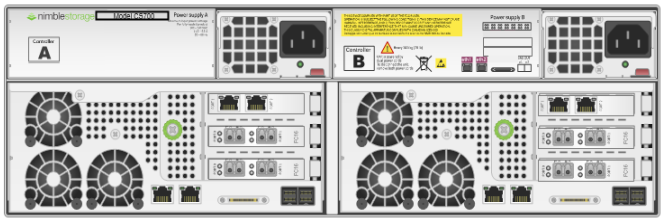
The Nimble Storage CS700 is designed for consolidating multiple large-scale critical applications with aggressive performance demands. This array handles IO-intensive workloads like larger scale VDI, Oracle or SQL server databases, while provides a compelling performance and IOPS value.

For more information, refer to: http://www.nimblestorage.com/technology-products/adaptive-flash-array-specifications/

Figure 7 Nimble CS700 Array



Front View



Back View

The Nimble Storage Predictive Flash Platform

The Nimble Storage Predictive Flash platform enables enterprise IT organizations to implement a single architectural approach to dynamically cater to the needs of varying workloads, driven by the power of predictive analytics. Predictive Flash is the only storage platform that optimizes across performance, capacity, data protection, and reliability within a dramatically smaller footprint.

Predictive Flash is built upon Nimble Cache Accelerated Sequential Layout (CASL™) architecture, NimbleOS and InfoSight™, the company's cloud-connected predictive analytics and management system. CASL scales performance and capacity seamlessly and independently. InfoSight leverages the power of deep data analytics to provide customers with precise guidance on the optimal approach to scaling flash, CPU, and capacity around changing application needs, while ensuring peak storage health.

NimbleOS Architecture

The Nimble Storage operating system, NimbleOS is based on its patented CASL™ architecture. CASL leverages the unique properties of flash and disk to deliver high performance and capacity – all within a dramatically small footprint. CASL and InfoSight™ form the foundation of the Predictive Flash platform, which allows for the dynamic and intelligent deployment of storage resources to meet the growing demands of business-critical applications.

Nimble Storage InfoSight

Using systems modeling, predictive algorithms, and statistical analysis, InfoSight™ solves storage administrators' most difficult problems. InfoSight also ensures storage resources are dynamically and intelligently deployed to satisfy the changing needs of business-critical applications, a key facet of Nimble Storage's Predictive Flash platform. At the heart of InfoSight is a powerful engine comprised of deep data analytics applied to telemetry data gathered from Nimble arrays deployed across the globe. More than 30 million sensor values are collected per day per Nimble Storage array. The InfoSight Engine transforms the millions of gathered data points into actionable information that allows customers to realize significant operational efficiency through:

- Maintaining optimal storage performance
- Projecting storage capacity needs
- Proactively monitoring storage health and getting granular alerts
- Proactively diagnoses and automatically resolves complex issues, freeing up IT resources for value-creating projects
- Ensures a reliable data protection strategy with detailed alerts and monitoring
- Expertly guides storage resource planning, determining the optimal approach to scaling cache, IOPS to meet changing SLAs
- Identifies latency and performance bottlenecks through the entire virtualization stack

For more information, refer to: http://www.nimblestorage.com/technology-products/infosight/

In-Line Compression

CASL uses fast, in-line compression for variable application block sizes to decrease the footprint of inbound write data by as much as 75 percent. Once there are enough variable-sized blocks to form a full write stripe, CASL writes the data to disk. If the data being written is active data, it is also copied to SSD cache for faster reads. Written data is protected with triple-parity RAID.

Thin-Provisioning and Efficient Capacity Utilization

Capacity is consumed as data gets written. CASL efficiently reclaims free space on an ongoing basis, preserving write performance with higher levels of capacity utilization. This avoids fragmentation issues that hamper other architectures.

Accelerated Write Performance

By sequencing random write data, CASL's writes to disk are orders of magnitude faster than other storage systems' random writes. The CS700, Nimble's top-of-the-line array, delivers double the write IOPS of a single MLC flash drive with a 7,200-RPM hard disk.

Read Performance

CASL accelerates read performance by dynamically caching hot data in flash, delivering sub-millisecond read latency and high throughput across a wide variety of demanding enterprise applications.

Adaptive Flash

CASL leverages flash as a true read cache, as opposed to a bolt-on tier. This enables Nimble arrays to easily adapt to changing workloads. As the architectural foundation of Adaptive Flash, CASL allows flash to flexibly scale for higher performance, especially benefitting those applications that work best when their entire working sets reside in the flash.

Intelligent Caching

CASL leverages flash as a true read cache, as opposed to a bolt-on tier. This enables Nimble arrays to easily adapt to changing workloads. As the architectural foundation of Adaptive Flash, CASL allows flash to flexibly scale for higher performance, especially benefitting those applications that work best when their entire working sets reside in flash.

Adaptive Flash Service Levels

Flash can be allocated to individual workloads on a per-volume basis according to one of three user-assignable service levels:

- **All Flash**: The entire workload is pinned in cache for deterministic low latency. Ideal for latency-sensitive workloads or single applications with large working sets or high cache churn.
- **Auto Flash**: Default service level where workload active data is dynamically cached. Ideal for applications requiring high performance, or a balance of performance and capacity.
- **No Flash**: No active data is cached in flash. Recommended for capacity-optimized workloads without high performance demands.

Efficient, Fully Integrated Data Protection

All-inclusive snapshot-based data protection is built into the Adaptive Flash platform. Snapshots and production data reside on the same array, eliminating the inefficiencies inherent to running primary and backup storage silos. And, InfoSight ensures that customers' data protection strategies work as expected through intuitive dashboards and proactive notifications in case of potential issues.

SmartSnap: Thin, Redirect-on Write Snapshots

Nimble snapshots are point-in-time copies capturing just changed data, allowing three months of frequent snapshots to be easily stored on a single array. Data can be instantly restored, as snapshots reside on the same array as primary data.

SmartReplicate: Efficient Replication

Only compressed, changed data blocks are sent over the network for simple and WAN-efficient disaster recovery.

Zero-Copy Clones

Nimble snapshots allow fully functioning copies, or clones of volumes, to be quickly created. Instant clones deliver the same performance and functionality as the source volume, an advantage for virtualization, VDI, and test/development workloads.

Application-Consistent Snapshots

Nimble enables instant application/VM-consistent backups using VSS framework and VMware integration, using application templates with pre-tuned storage parameters.

SmartSecure: Flexible Data Encryption

NimbleOS enables encryption of individual volumes with little to no performance impact. Encrypted volumes can be replicated to another Nimble target, and data can be securely shredded.

REST API:

NimbleOS has a RESTful API that allows for powerful, secure and scriptable management of storage objects. Using this API, an administrator can use the scripting or orchestration to interact with infrastructure components in a well-defined, repeatable manner.

For more information, see:
http://info.nimblestorage.com/rs/nimblestorage/images/nimblestorage_technology_overview.pdf

Solution Architecture

SmartStack solution delivers a converged infrastructure platform that incorporates compute, network and storage best practices from Cisco and Nimble to deliver a resilient, scalable and flexible data center architecture for Enterprise and cloud deployments.

The following platforms are integrated in this SmartStack architecture:

- Cisco Unified Computing System (UCS) – blade and rack servers
- Cisco UCS 6200 Series Fabric Interconnects (FI) – unified access to storage and LAN networks
- Cisco Nexus 9300 series switches – connectivity to users, other LAN networks and Cisco UCS domains
- Cisco MDS fabric switches – SAN fabric providing Fibre Channel (FC) connectivity to storage
- Nimble CS700 array – SAN boot of hybrid storage array with SSDs and HDDs
- Cisco Nexus 1000V (optional) – access layer switch for virtual machines
- VMware vSphere 6.0 U1b - Hypervisor

This SmartStack architecture uses Cisco UCS B-series and C-series rack mount servers for compute. Cisco UCS B-series servers are housed in a Cisco UCS 5108 blade server chassis that can support up to eight half-width blades or four full-width blades. Each server supports a number of converged network adapters (CNAs) that converge LAN and SAN traffic onto a single interface rather than requiring multiple network interface cards (NICs) and host bust adapters (HBAs) per server. Cisco's Virtual Interface Cards (VICs) support 256 virtual interfaces and supports Cisco's VM-FEX technology (see link below). Two CNAs are typically deployed on a server for redundant connections to the fabric. Cisco VIC is available as a Modular LAN on Motherboard (mLOM) card and as a Mezzanine Slot card. A half-height PCI Express (PCIe) card form-factor is also available but exclusively for Cisco UCS C-series rack-mount servers available. For more information on the different models of Cisco UCS VIC adapters available, see: http://www.cisco.com/c/en/us/products/interfaces-modules/unified-computing-system-adapters/index.html

All compute resources in the data center connect into a redundant pair of Cisco UCS fabric interconnects that provide unified access to storage and other parts of the network. Each blade server chassis requires 2 Fabric Extender (FEX) modules that extend the unified fabric to the blade server chassis and connect into the FIs using 2, 4 or 8 10GbE links. Each Cisco UCS 5108 chassis can support up to two fabric extenders that provide active-active data plane forwarding with failover for higher throughput and availability. FEX is a consolidation point for all blade server I/O traffic, which simplifies operations by providing a single point of management and policy enforcement. For other benefits of FEX and additional info, see: http://www.cisco.com/c/en/us/solutions/data-center-virtualization/fabric-extender-technology-fex-technology/index.html.

Two second generation models of FEX are currently available for the Cisco UCS blade server chassis. Cisco UCS 2204XP and 2208XP connect to the unified fabric using multiple 10GbE links. The number of 10GbE links that are supported for connecting to an external fabric interconnect and to the blade servers within the chassis are shown in the table below. The maximum I/O throughput possible through each FEX is also shown.

Table 2 Blade Server Fabric Extender Models

Blade Server Models	Internal Facing Links to Blade Servers	External Facing Links to FI	Aggregate I/O Bandwidth
Cisco UCS 2204XP	16 x 10GbE	Up to 4 x 10GbE	40Gbps per FEX
Cisco UCS 2208XP	32 x 10GbE	Up to 8 x 10GbE	80Gbps per FEX

By deploying a pair of Cisco 2208XP FEX, a Cisco UCS 5108 chassis with blade servers can get up to maximum of 160Gbps of I/O throughput for the servers on that chassis. For additional details on the FEX, see following link:

http://www.cisco.com/en/US/prod/collateral/ps10265/ps10276/data_sheet_c78-675243.html

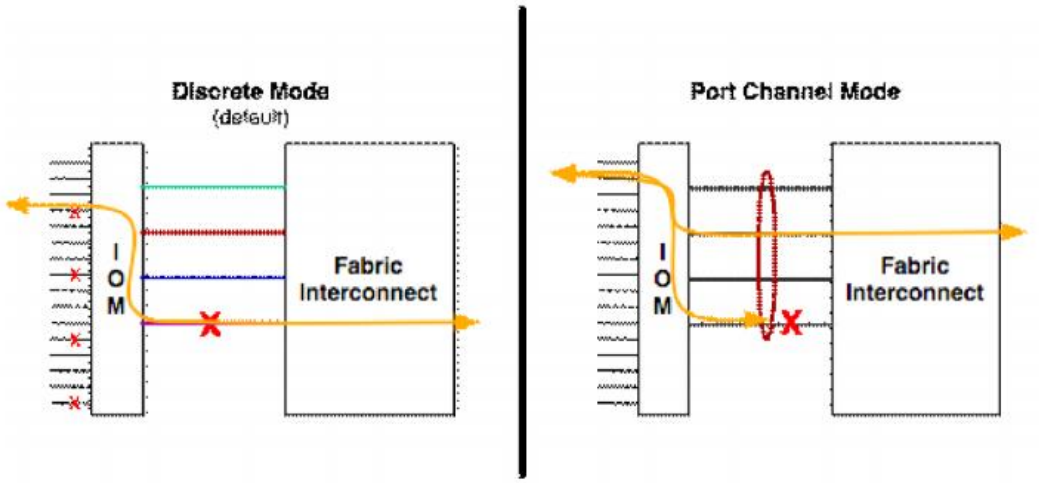
The rack mount servers can also benefit from a FEX architecture to aggregate I/O traffic from several rack mount servers but unlike the blade servers where the FEX modules fit into the back of the chassis, rack mount servers require a standalone FEX chassis. Cisco 2200 Series FEX model is functionally equivalent to the above blade server FEX models. Other than the physical cabling required to connect ports on Cisco Nexus FEX 2300 to servers and FIs, the discovery and configuration is same as that of the blade server to FI connectivity. For data centers migrating their access-switching layer to 40GbE speeds, Cisco Nexus 2300 series FEX is the newest model. For additional details on the Cisco Nexus 2000 Series FEX, see following link:

http://www.cisco.com/c/en/us/products/switches/nexus-2000-series-fabric-extenders/index.html

Rack mount servers can also be deployed by directly connecting into the FIs, without using Cisco UCS FEX chassis. However, this could mean additional operational overhead by having to manage server policies individually but nevertheless a valid option in smaller deployments.

The fabric extenders in a blade server chassis connect externally to the unified fabric using multiple 10GbE links – the number of links depends on the aggregate I/O bandwidth required. Each FEX connect into a Cisco UCS 6200 series fabric interconnect using up to 4 or 8 10GbE links depending on the model of FEX. The links can be deployed as independent links (discrete Mode) or grouped together using link aggregation (port channel mode). In discrete mode, each server is pinned to a FEX link going to a port on the fabric interconnect and if the link goes down, the server's connection also goes down through the FEX link. In port channel mode, the flows from the server will be redistributed across the remaining port channel members. This is less disruptive overall and therefore port channel mode is preferable.

Figure 8 FEX Connection in Discrete Mode and PortChannel Mode



Cisco UCS system provides the flexibility to individually select components of the system that meet the performance and scale requirements of a customer deployment. There are several options for blade and rack servers, network adapters, FEX and Fabric Interconnects that make up the Cisco UCS system.

Compute resources are grouped into an infrastructure layer and application data layer. Servers in the infrastructure layer are dedicated for hosting virtualized infrastructure services that are necessary for deploying and managing the entire data center. Servers in the application data layer are for hosting business applications and supporting services that Enterprise users will access to fulfill their business function.

The architecture can support any hypervisor but VMware vSphere will be used to validate the architectures. High availability features available in the hypervisor will be leveraged to provide resiliency at the virtualization layer of the stack.

SmartStack architecture with Nimble Storage array can support block storage using either iSCSI or Fibre Channel (FC). The focus of this CVD will be fibre channel access to Nimble Storage CS700 array. For more details on an

iSCSI based SmartStack design using Nimble CS 300 array – see following links on Cisco Design Zone for SmartStack:

- Design Guide: http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/smartstack_cs300_mini.html
- Deployment Guide: http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/smartstack_cs300_mini_deploy.html

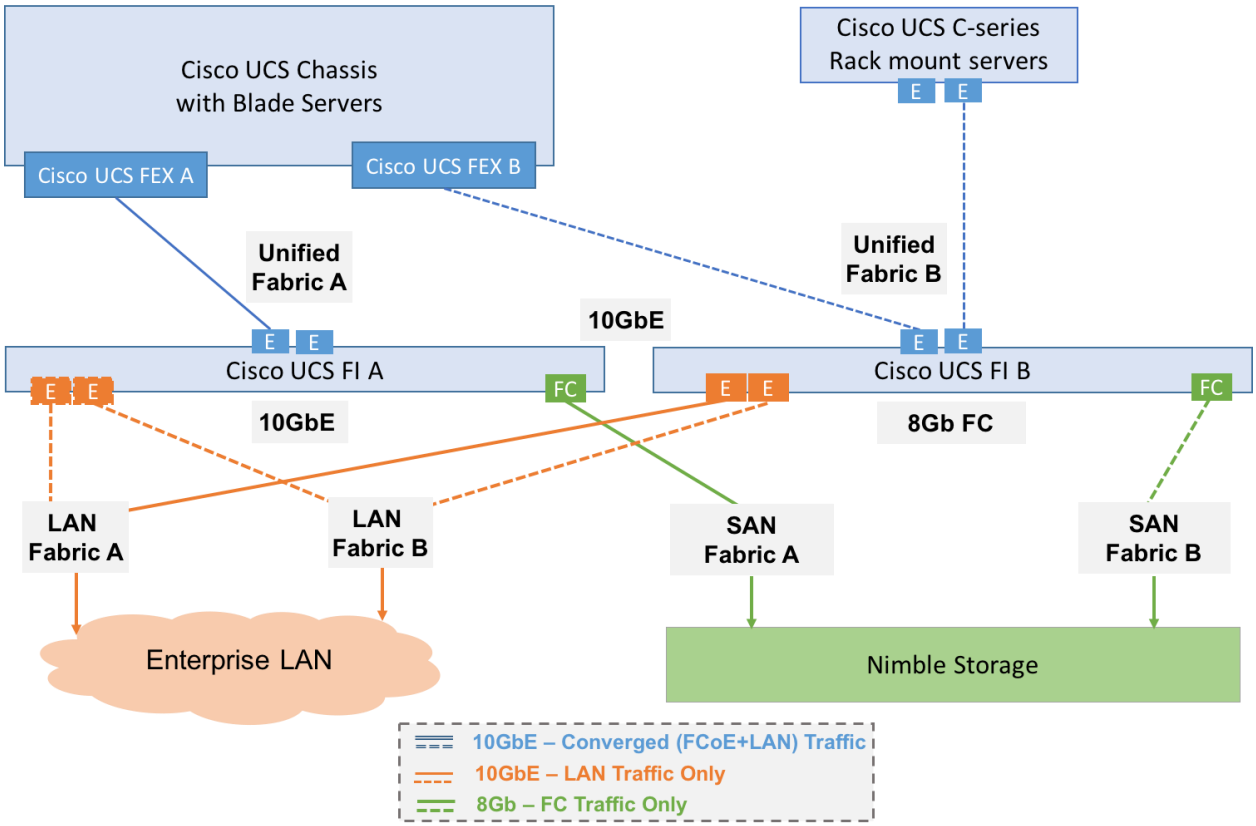
The CS-series arrays are based on Nimble Adaptive Flash CPU-driven architecture that provide enterprises with the ability to scale performance and capacity independently. The Adaptive Flash platform is based on Nimble's patented Cache Accelerated Sequential Layout (CASL™) architecture, and InfoSight™, a cloud-based management and support system. CASL maximizes the unique properties of flash and disk to deliver high performance and capacity – with a dramatically small footprint. CASL and InfoSight™ form the foundation of the Adaptive Flash platform, which allows for the dynamic and intelligent deployment of storage resources to meet the growing demands of business-critical applications.

Nimble CS700 is a hybrid array designed for large scale enterprise applications with high performance needs. The base array can support up to 12 hard disk drives (HDD) and 4 solid-state drives (SSD). The maximum raw capacity of the base array using 6TB HDD drives is 72TB, with an effective capacity of around 50-100 TB. The flash capacity of the base array using 1.6 TB SSD drives is approximately 3.2-7.6 TB. The capacity and performance can be extended using expansion shelves, with the best model providing up to 90TB of raw capacity and 66-132TB of effective capacity per shelf. The SSD expansion shelf can deliver an additional 30.72TB of flash capacity. CS700 supports up to 6 expansion shelves which can be combined to meet the performance and capacity needs of enterprise data centers. For more information on CS700, see: http://www.nimblestorage.com/technology-products/adaptive-flash-array-specifications/

If a unified fabric from compute to storage is required or preferable, iSCSI access can be used but the focus of this SmartStack solution is on a FC based access which currently requires a dedicated SAN network. The traffic does traverse a unified fabric between Cisco UCS and Fabric Interconnects but then diverge onto separate LAN and SAN networks as shown in the figure below.

This SmartStack architecture utilizes a dedicated SAN network for block storage traffic. The figure below shows a unified fabric (blue lines) between Cisco UCS servers and Fabric Interconnects which then splits into separate SAN (green lines) and LAN (orange lines) networks. If a unified fabric is required or preferable from compute to storage end-to-end, then the iSCSI block storage access can be used. Alternatively, a Cisco Nexus 5000 series or MDS switch can be used for FCoE to FC connectivity since FCoE capability is not directly available to the Nimble Storage array. The focus of this SmartStack solution is on a dedicated fibre channel based storage access design using Cisco MDS fabric switches.

Figure 9 SmartStack Architecture - High Level



The SmartStack architecture is a highly resilient design with redundant Unified, LAN and SAN fabrics that includes component and link level redundancy when accessing these fabrics. The unified fabric and LAN connections are 10Gb Ethernet links and the SAN connections are 8Gb FC links. Multiple 10GbE links and FC links are deployed in the SmartStack architecture with link aggregation using Link Aggregation Control Protocol (LACP) to provide higher aggregate bandwidth and resiliency across the different fabrics. Use of LACP is strongly recommended when available for improved failover convergence time and protection from misconfigurations. Static or manual bundling is an alternative but is less preferable and therefore not used in this architecture.

The data center LAN network in this SmartStack architecture uses a pair of Cisco Nexus 9300 switches that serve as the access/aggregation layer of the data center network. In this design, the Cisco Nexus 9300s provide reachability to users and other parts of the network. In larger deployments, an additional layer of hierarchy can be added using Cisco Nexus 9500 series switches as an aggregation/core layer in a classic data center tiered design or as a spine in a spine/leaf design. Cisco Nexus 9000 family of switches can operate in one of two modes: Application Centric Infrastructure (ACI) for newer cloud architectures or in Cisco NX-OS standalone mode for the more traditional data center architectures. In the SmartStack architecture, the Cisco 9300s are deployed in NX-OS standalone mode and provide investment protection by enabling a pathway to Cisco's Application Centric Infrastructure (ACI). Cisco ACI was designed to help Enterprises transition their data centers to a cloud architecture that is dynamic, where applications can be quickly deployed and scaled, and adapt with the needs of the business. To enable this, Cisco ACI provides a flexible, scalable, application centric, policy-based framework based on open APIs that accelerate and simplify data center deployments while providing centralized orchestration, visibility and management.

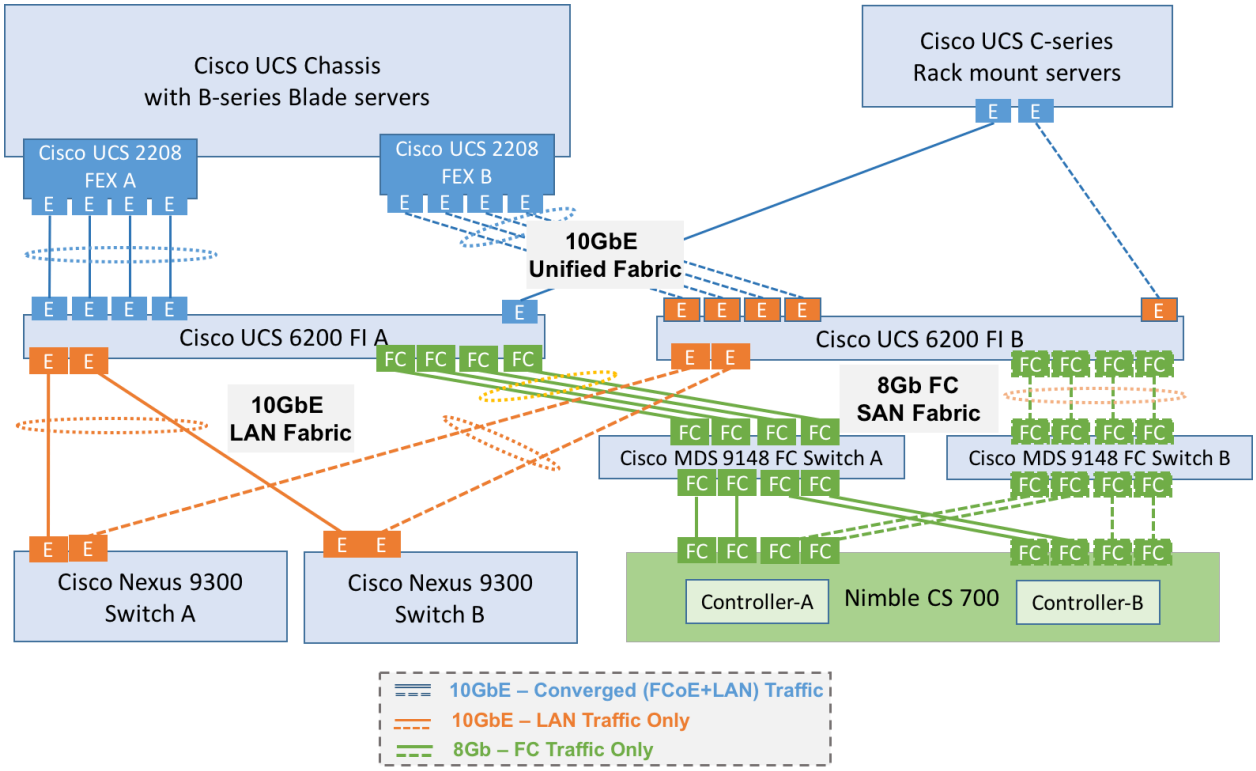
Virtual Port Channel (vPC) or link-aggregation capabilities of the Cisco Nexus 9000 family of switches are used on the network links to Cisco UCS Fabric Interconnects. A vPC allows links that are physically connected to two different Cisco Nexus 9000 Series devices to appear as a single PortChannel. vPC provides Layer 2 multipathing with load balancing by allowing multiple parallel paths between nodes that result in increased bandwidth and redundancy. A vPC-based architecture is therefore highly resilient and robust and scales the available Layer 2 bandwidth by using all available links. Other benefits of vPCs include:

- Provides a loop-free topology

- Eliminates Spanning Tree Protocol blocked ports
- Uses all available uplink bandwidth
- Provides fast convergence for both link and device failures
- Provides higher aggregate bandwidth by adding links – same as Port Channels
- Helps ensure high availability

The SAN network in this SmartStack architecture is 8Gb FC based and uses Cisco MDS 9100 switches.

Figure 10 SmartStack Architecture



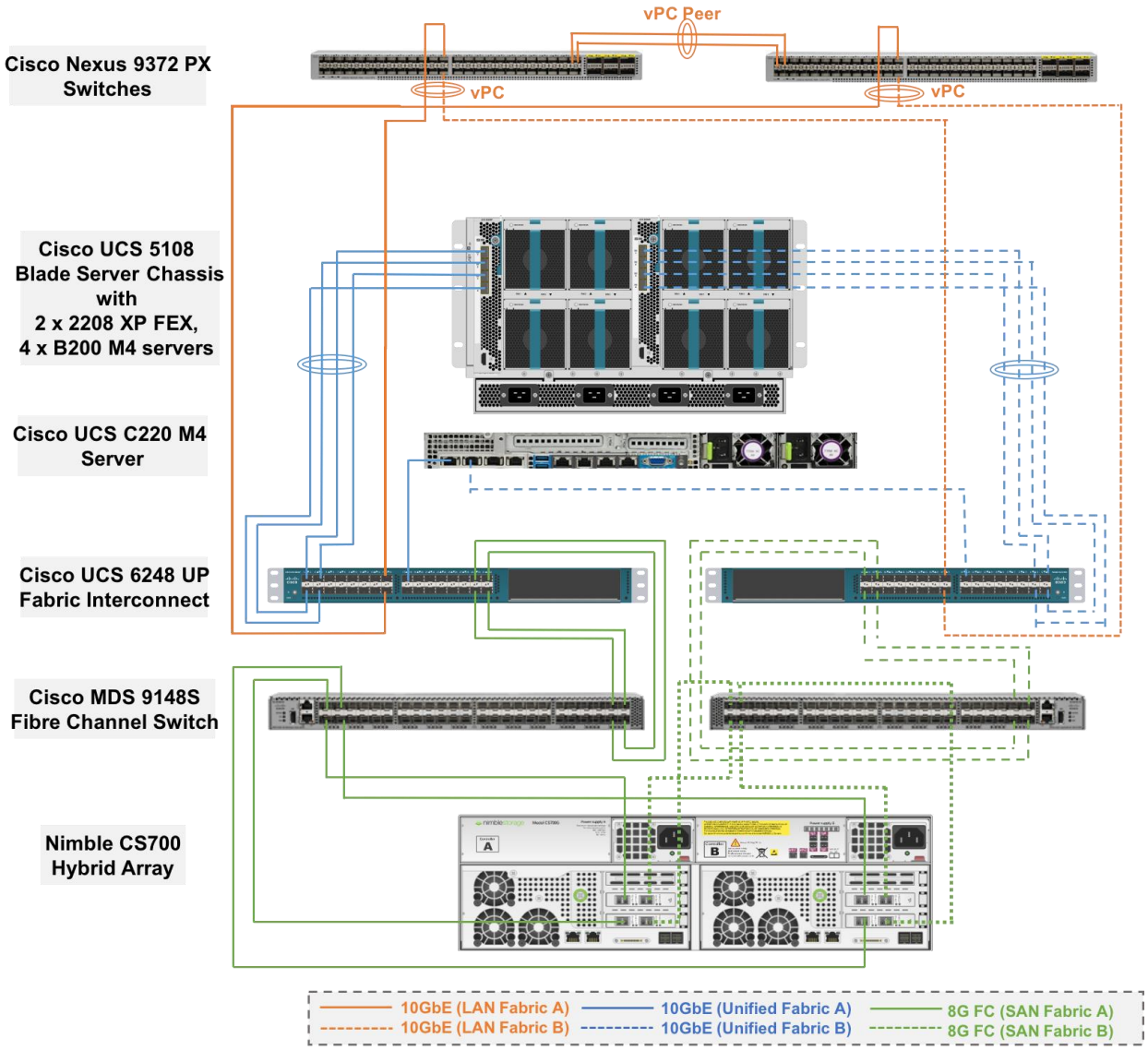
The aggregate throughput of the two fabrics depends on the number of 10GbE and 8Gb FC links deployed to meet the throughput needs of a given customer deployment. The figure above shows a 4 x 10GbE access to the unified fabric from Cisco UCS blade server chassis and 10GbE from the rack mount server. This can be scaled up to max of 8x10 GbE on the Cisco UCS blade server chassis. The LAN and SAN fabric provides a 2x10 GbE and 4x8Gb FC access respectively from each FI. This can also be scaled higher by adding additional 10GbE and 8Gb FC links.

The SmartStack platform will be managed locally using Cisco UCS Manager, VMware vCenter and Nimble Management software. The storage array will also be remotely monitored from the cloud using Nimble InfoSight™ to provide insights into I/O and capacity usage, trend analysis for capacity planning and expansion, and for proactive ticketing and notification when issues occur.

Solution Design

The end-to-end SmartStack design is shown in the figure below.

Figure 11 SmartStack Design



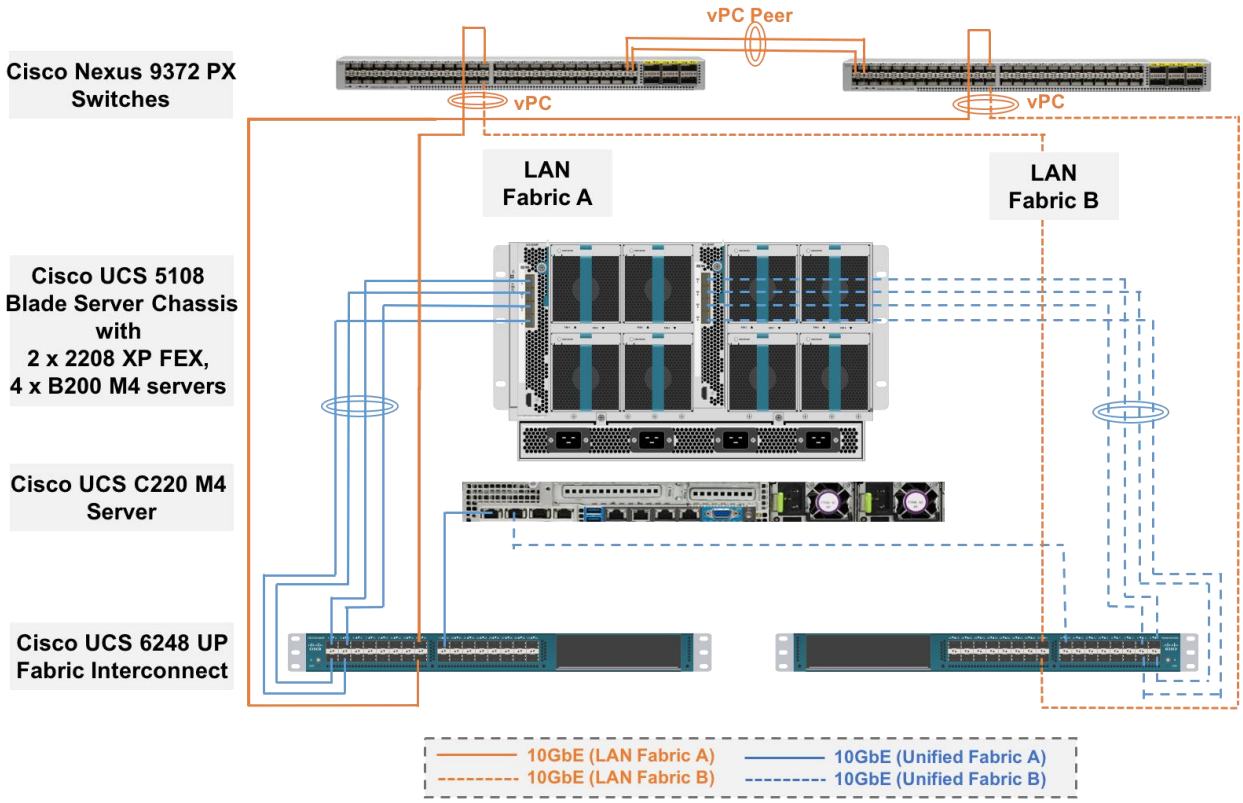
Compute

SmartStack design uses Cisco UCS with 4x Cisco B200M4 half-width blades to provide the compute resources. A Cisco C220 M4 rack mount server is also included in the design as alternative compute option. The hypervisor layer in the design is provided by VMware ESXi 6.0 U1b. Features available at the hypervisor layer (For example, VMware clustering, High availability) are leveraged where possible for a robust design. The blade server chassis is connected via FEX modules on the back of the chassis to a pair of Cisco UCS 6248 FIs. A pair of Cisco 2204 XP fabric extenders is used in this design. The FEX to FI connectivity uses 8x10GbE links, 4 from FEX-A to FI-A and 4 from FEX-B to FI-B to provide an aggregate access bandwidth of 80Gbps to the unified fabric. The FIs are connected to LAN and SAN network using 10GbE and 8Gb FC links. The FI provides 40Gbps of aggregate bandwidth to the LAN network and 64Gbps to the SAN network. Link aggregation using port channels are used on the unified fabric FEX-FI connections and virtual port channels on the Cisco Nexus-FI LAN connections.

Network

The LAN network design is shown in the figure below.

Figure 12 SmartStack LAN Fabric Design



The LAN network provides network reachability to the applications hosted on Cisco UCS servers in the data center. The infrastructure consists of a pair of Cisco Nexus 9372 PX switches deployed in NX-OS standalone mode. Redundant 10Gbps links from each Cisco Nexus switch are connected to ports on each FI and provide 20Gbps of bandwidth through each Cisco Nexus. Virtual PortChannels (vPCs) are used on the Cisco Nexus links going to each FI. VLAN Trunking is enabled on these links as multiple application data VLANs, management and vMotion traffic needs to traverse these links. Jumbo Frames are also enabled in the LAN network to support vMotion between multiple Cisco UCS domains. See Design Practices section for other Cisco Nexus 9000 best practices in the design.

The SAN network provides fibre channel connectivity to the Nimble storage array and consists of a pair of MDS switches. The MDS switches form completely separate fabrics (SAN fabric A, SAN fabric B) and use a dual vSAN (vSAN-A, vSAN-B) design to provide two redundant and completely diverse paths to the Nimble array.

Link aggregation using port channels are used to aggregate 4 x 8G FC links to provide 32G of FC bandwidth on each SAN Fabric between Cisco FI and MDS switches. Link aggregation is not used on the links to Nimble array but 2 links from each SAN fabric connects to both controllers to provide 32G of FC bandwidth to the active controller. Four links from the SAN fabric, 2 from SAN Fabric A and 2 from SAN Fabric B, also connect to the backup controller so that both controllers have 32B FC access to the SAN fabric.

Cisco MDS switches are deployed with N-Port ID Virtualization (NPIV) enabled to support the virtualized environment running on Cisco UCS blade and rack servers. NPIV is necessary to provide isolation in virtualized environments where multiple virtual machines are running on a single server but a LUN needs to be presented to only one VM and not all VMs running on the server. Without NPIV, LUNs would be presented to the host and as a result, all VMs running on that host. To support NPIV on the Cisco UCS servers, the Cisco UCS Fabric Interconnects that connect the servers to the SAN fabric, are enabled for N-Port Virtualization (NPV) mode by configuring to operate in end-host mode (as opposed to FC switching mode). NPV enables Cisco FIs to proxy

fabric login, registration and other messages from the servers to the SAN Fabric without being a part of the SAN fabric. This is important for keeping the limited number of Domain IDs that Fabric switches require to a minimum.

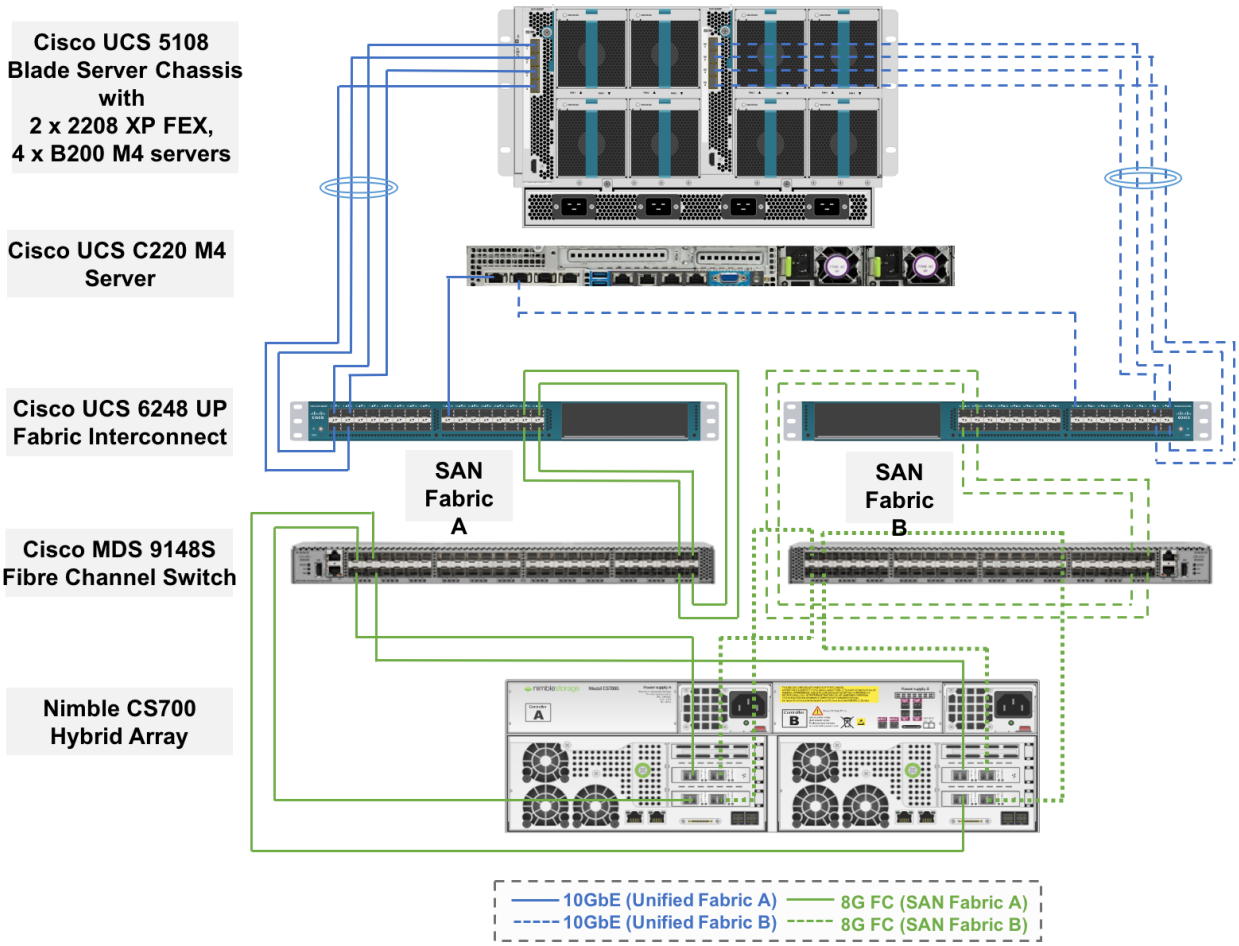
The design also uses the following best practices:

- Jumbo frames on unified fabric links between Cisco UCS and fabric interconnects
- QoS policy for traffic prioritization on the unified fabric
- Port-channels with multiple links are used in both the unified fabric and SAN network for higher aggregate bandwidth and redundancy
- Zoning is single initiator (vHBA) to multiple targets

Storage

The Nimble storage design is shown in the figure below.

Figure 13 SmartStack Storage Design



SmartStack design uses Nimble CS700 hybrid array to provide block storage. A base configuration with 54.4TB of raw capacity (12 x 4TB HDDs, 4 x 1.6TB SSDs) was deployed in the array used for validation. Nimble's CS700 supports the addition of up to 6 expansion shelves or up to 4 CS-series (any model) arrays in a scale-out cluster to increase performance and capacity.

Each Nimble Storage controller supports up to 3 FC interface cards, each with dual 16G FC ports. This SmartStack design uses 8G fabric connectivity with two FC interface cards to provide 32G of FC bandwidth per controller. For additional FC bandwidth, a third FC card can be deployed on each controller but this interface is typically used for 10GbE connections to other arrays in a scale-out cluster for data replication traffic. The links

between a pair of Cisco MDS and Fabric Interconnect switches are aggregated using 4x8G FC links to deliver 32G of bandwidth across the SAN fabric to each controller.

This SmartStack design uses FC SAN boot to boot the servers. The Service Profile used to configure and deploy Cisco UCS servers is configured to include a boot policy that points to the Nimble Storage array. The boot policy specifies a primary and secondary SAN path to the two controllers on the array where the boot volumes reside. A second boot policy is also configured but with the primary and secondary paths reversed from that of the first boot profile. The second boot policy is used to load balance SAN boot across different paths when multiple servers are booting. This is an optional aspect of the design that can be helpful in larger deployments for distributing load when multiple servers have to be simultaneously booted. Each server has a dedicated boot volume (40GB) on the Nimble storage array. Nimble Storage arrays provide an ACL at the initiator level to only allow connections from the appropriate Cisco UCS blade. During the initial SAN boot, the server attaches to all primary and secondary connections to both active and standby controller. This provides for normal boot operations even when a controller or primary path is offline. The hosts are configured with the Nimble Connection Manager and Path Selection Policy which optimize MPIO settings. This will allow for proper FC path management and failover connectivity.

The following sections provide more details on the connectivity and high availability aspects of this design.

Design Considerations

Management Connectivity

This SmartStack design uses a separate out-of-band management network to configure and manage compute, storage and network components in the solution. Management ports on each physical device (Cisco UCS FI, Nimble CS700 Array Controllers, Cisco Nexus and MDS switches) in the solution are connected to a separate, dedicated management switch.

Management access is also needed to ESXi hosts, vCenter VM and other management VMs but this is done in-band in this design. However, if out-of-band management to these components are required, the disjoint layer 2 feature available in Cisco UCS running end-host mode can be used. This would require additional uplink port(s) on the Cisco UCS FI to connect to the management switches. Additional out-of-band management VLANs and vNICs will also be needed and associated with the uplink ports. The additional vNICs are necessary since a server vNIC can only be associated with a single uplink.

QoS and Jumbo Frames

Cisco UCS, Cisco Nexus and MDS switches in the SmartStack solution provide QoS policies and features for handling congestion and traffic spikes that can occur in a SmartStack environment. SmartStack support different types of traffic (For example, vMotion, FCOE) and the QoS capabilities in these components can alleviate and provide the priority that certain types of traffic require.

SmartStack design also uses jumbo frames with an MTU of 9216 Bytes across the LAN and Unified Fabric links. Jumbo frames increase the throughput between devices by enabling larger sized frames to be sent and received on the wire while reducing the CPU resources necessary to process them. Jumbo frames were enabled during validation on the LAN network links in the Cisco Nexus switching layer and on the Unified Fabric links.

Cisco UCS C-Series Server Connectivity Options

As of Cisco UCS Manager Release 2.2, the Cisco UCS C-Series servers can be connected to the Cisco UCS Fabric Interconnects using one of the two design options and still can leverage the benefits of unified management that the Cisco UCS Manager provides.

Direct Attached Design

Customers can directly connect the Cisco UCS C-series rack servers to Cisco UCS Fabric Interconnects as of Cisco UCS Manager Release 2.2 and above releases. This design is well suited for smaller deployments with limited scale. For more details on this connectivity option, see:

http://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/c-series_integration/ucsm3-1/b_C-Series-Integration_UCSM3-1/b_C-Series-Integration_UCSM3-1_chapter_011.html

This SmartStack solution was validated with Cisco UCS C-series using the Direct Attached design.

Fabric Extender Attached Design

For larger scale deployments, Cisco UCS C-series can be connected using standalone Fabric Extenders, namely the 1RU FEX 2232PP. Functionally, the standalone FEX is identical to the Cisco UCS 2204 and 2208 IOM modules that are deployed on the Cisco UCS 5108 blade server chassis for connecting to Cisco UCS Fabric Extenders. Similarly, the Cisco VIC on each Cisco UCS C-series server connect to both Fabric Interconnects using two Cisco FEX 2232PPs. The FEX and Fabric Interconnects form port channels automatically based on the chassis discovery policy providing a link resiliency to the Cisco UCS C-series server. This is identical to the behavior of the IOM to Fabric Interconnect connectivity. Logically, the virtual circuits formed within the Cisco UCS domain are consistent between B and C series deployment models and the virtual constructs formed at the vSphere are unaware of the platform in use.

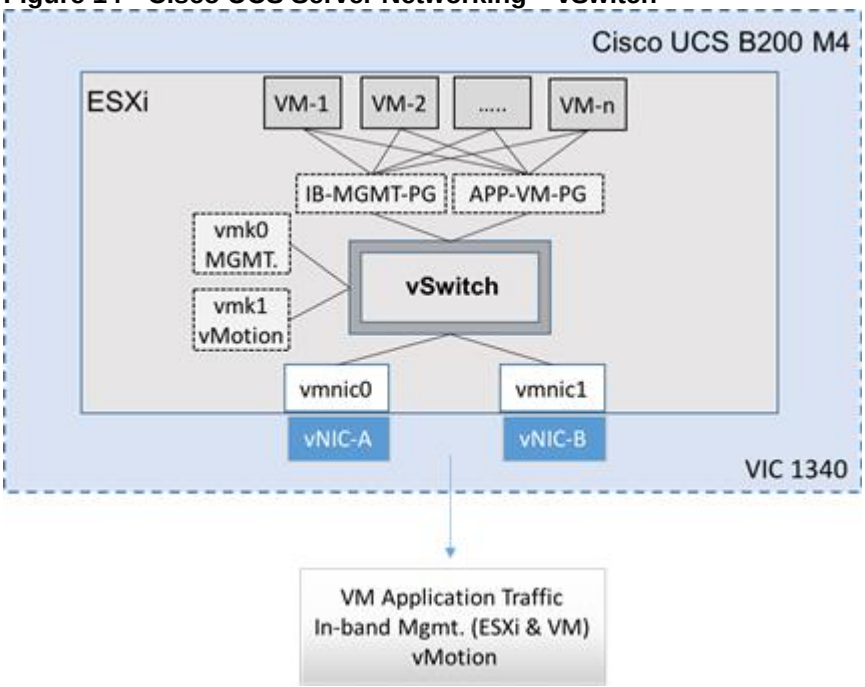
Cisco UCS Server – vSphere Configuration

Cisco UCS B200M4 blade servers with Cisco 1340 VIC and Cisco UCS C220M4 rack servers with Cisco 1227 VIC running vSphere 6.0 U1 were validated in this SmartStack design. Cisco UCS servers were assigned to a VMware High Availability (HA) cluster to mitigate against host failures. Two VMware HA clusters were used in validation – one for infrastructure management and services (For example, VMware vCenter) and one for applications that users access. The Cisco VIC on each server presents multiple vPCIe devices to ESXi. VMware vSphere identifies these virtual adapters as vmnics. In this SmartStack design, the following virtual adapters (vNICs) were used with –A connected to unified fabric A and –B to unified fabric B resulting in each ESXi node being dual homed to the external network.

Two vNICs (vNIC-A, vNIC-B) for application VM, in-band management and vMotion traffic

The connectivity within each ESXi server and the vNIC presented to ESXi are shown below.

Figure 14 Cisco UCS Server Networking – vSwitch



The SmartStack architecture uses two port groups (IB-MGMT-PG) for in-band management of the VMs and APP-VM-PG for application traffic. The design also used two VMkernel NICs (vmk), each with its own port group for host level functionality:

- vmk0 - ESXi management

- vmk1 - vMotion interface
- The ESXi management interface is for host to vCenter connectivity, direct access to ESXi shell and VMware cluster communication. The vMotion interfaces are private subnets supporting data access and VM migration across the SmartStack infrastructure.

Cisco UCS Server – Virtual Switching using Cisco Nexus 1000V (Optional)

A Cisco Nexus 1000V virtual distributed switch is used to provide connectivity between virtual machines and host connectivity to external networks. Cisco Nexus 1000v is an optional component of the SmartStack solution. Cisco Nexus 1000v is fully integrated into the VMware virtual infrastructure, including VMware vCenter and vCloud Director and extends the network edge to the hypervisor and virtual machines. Cisco Nexus 1000V is compatible with any upstream physical access layer switch that is compliant with Ethernet standards including Cisco Nexus switches and switches from other network vendors.

The Cisco Nexus 1000v comprises of the following components and operationally emulates a modular switch where:

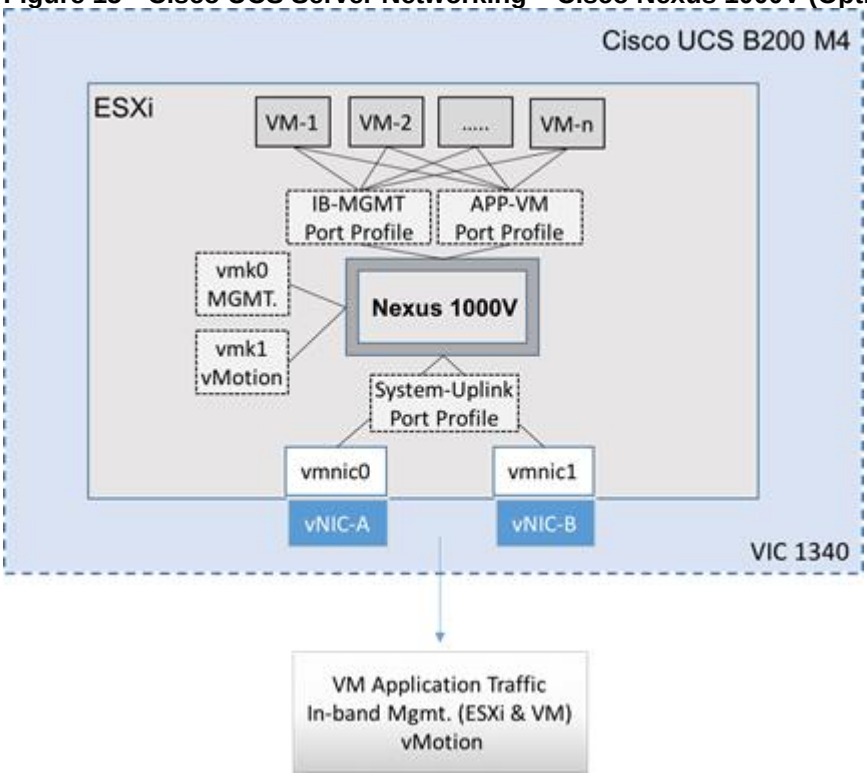
- Virtual Supervisor Module (VSM) – is the control and management plane of the virtual switch. VSM is deployed as an external virtual machine and runs NX-OS to manage multiple Virtual Ethernet Modules as one logical modular switch.
- Cisco Virtual Ethernet Module (VEM) – virtual line card or module within the virtual switch that VMs connect into. VEM is embedded in each VMware vSphere host and replaces the VMware Virtual Switch (vSwitch) functionality.
- Operating inside the VMware ESXi hypervisor, Cisco Nexus 1000V VEM uses the VMware vNetwork Distributed Switch (vDS) API, jointly developed by Cisco and VMware to provide policy-based VM connectivity, Layer 2 switching and advanced networking functions. The tight integration makes Cisco Nexus 1000V fully aware of key virtualization features such as VMware vMotion and Distributed Resource Scheduler (DRS). VEM provides switching functionality based on the configuration information it receives from the VSM. In the event of a communication loss with the VSM, VEM continues to forward traffic based on last known configuration or Nonstop Forwarding (NSF). VEM therefore provides reliability and advanced switching functionality.

Cisco Nexus 1000V VSM controls multiple VEMs as one logical modular switch with the VEM running as software on each server representing the line cards on a switch. VSM is integrated into VMware vCenter server to that the data center administrator can manage and monitor the network configuration on the Cisco Nexus 1000V switches. Configurations done through the VSM are automatically propagated to all VEMs managed by a given VSM. For high availability, VSMs can be redundantly deployed providing rapid, stateful failover as the VSMs. VSMs also provide port-profiles as a mechanism for grouping ports by category that enables the solution to scale to a high number of ports. VSM can also be accessed and managed through CLI, SNMP, XML API and CiscoWorks LAN management Solution.

The figure below shows the virtual networking within ESXi on a single Cisco UCS server. The Cisco Nexus 1000V VEM running on the ESXi node is registered to a VSM running on the infrastructure cluster and integrated into VMware vCenter. The Cisco VIC on each server presents multiple vPCIe devices to ESXi that are identified as vmnics. This SmartStack design uses two virtual adapters (vNIC-A, vNIC-B) with vNIC-A connected to unified fabric A and vNIC–B connected to unified fabric B. Host traffic (application VMs, in-band management, vMotion) are distributed across these vNICs. The ESXi vmnics are presented as Ethernet interfaces on Cisco Nexus 1000V. Cisco Nexus 1000V provides port profiles to address the dynamic nature of server virtualization from the network's perspective. Port profiles, defined on the VSM, serve as templates that define the network, security and service level policies for groups of virtual machines. Cisco Nexus 1000v aggregates the Ethernet uplinks into a single port channel named the "System-Uplink" port profile for fault tolerance and improved throughput. The port profiles can then be applied to individual virtual machine Ethernet interfaces through VMware vCenter.

Cisco Nexus 1000v provides link failure detection, disabling Cisco UCS Fabric Failover within the vNIC template is recommended.

Figure 15 Cisco UCS Server Networking – Cisco Nexus 1000V (Optional)



The SmartStack architecture uses two port profiles (IB-MGMT) for in-band management of the VMs and APP-VM for the application traffic used in validation. The design also uses two VMkernel NICs (vmk), each with its own port profile for host level functionality:

- vmk0 - ESXi management
- vmk1 - vMotion interface
- The ESXi management interface is for host to vCenter connectivity, direct access to ESXi shell and VMware cluster communication. The vMotion interfaces are private subnets supporting data access and VM migration across the SmartStack infrastructure.

The Cisco Nexus 1000v also supports Cisco's MQC to assist in uniform operation and enforcement of QoS policies across the infrastructure. The Cisco Nexus 1000v supports marking at the edge and policing traffic from VM-to-VM.

For more information about "Best Practices in Deploying Cisco Nexus 1000V Series Switches on Cisco UCS B and C Series Cisco UCS Manager Servers" refer to:
http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9902/white_paper_c11-558242.html

Cisco Nexus 9000 Series – vPC Best Practices

SmartStack incorporates the following Cisco Nexus 9000 design best practices and recommendations.

vPC Peer Keepalive Link Considerations

It is recommended to have a dedicated 1Gbps layer 3 link for vPC peer keepalive, followed by out-of-band management interface (mgmt0) and lastly, routing the peer keepalive link over an existing Layer3 infrastructure between the existing vPC peers. vPC peer keepalive link should not be routed over a vPC peer-link. The out-of-band management network is used as the vPC peer keepalive link in this SmartStack design.

vPC Peer Link Considerations

- Only vPC VLANs are allowed on the vPC peer-links. For deployments that require non-vPC VLAN traffic to be exchanged between vPC peer switches, deploy a separate Layer 2 link for this traffic.
- Only required VLANs are allowed on the vPC peer links and member ports – prune all others to minimize internal resource consumption.
- Ports from different line cards should be used to provide redundancy for vPC peer links. It was not possible to do this on the fixed module Cisco Nexus 9372PX switches used in this SmartStack design.

vPC General Considerations:

- vPC peer switches deployed using same bridge-id and spanning tree VLAN priority by configuring the **peer-switch** command on both vPC peer switches. This feature improves convergence and allows peer switches to appear as a single spanning-tree root in the Layer 2 topology.
- vPC role priority specified on **both** Cisco Nexus peer switches. vPC role priority determines which switch will be primary and which one will be secondary. The device with the lower value will become the primary. By default, this value is 32677. Cisco recommends that the default be changed on both switches. Primary vPC devices are responsible for BPDU and ARP processing. Secondary vPC devices are responsible for shutting down member ports and VLAN interfaces when peer-links fail.
- vPC convergence time of 30s (default) was used to give routing protocol enough time to converge post-reboot. The default value can be changed using **delay-restore <1-3600>** and **delay-restore interface-VLAN <1-3600>** commands. If used, this value should be changed globally on both peer switches to meet the needs of your deployment.
- vPC peer switches enabled as peer-gateways using peer-gateway command on both devices. 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.
- vPC **auto-recovery** enabled to provide a backup mechanism in the event of a vPC peer-link failure due to vPC primary peer device failure or if both switches reload but only one comes back up. This feature allows the one peer to assume the other is not functional and restore the vPC after a default delay of 240s. This needs to be enabled on both switches. The time to wait before a peer restores the vPC can be changed using the command: **auto-recovery reload-delay <240-3600>**.
- Cisco NX-OS can synchronize ARP tables between vPC peers using the vPC peer links. This is done using a reliable transport mechanism that the Cisco Fabric Services over Ethernet (CFSoE) protocol provides. For faster convergence of address tables between vPC peers, **IP ARP synchronize** command was enabled on both peer devices in this SmartStack design.

vPC Member Link Considerations

- LACP used for port channels in the vPC. LACP should be used when possible for graceful failover and protection from misconfigurations
- LACP mode active-active used on both sides of the port channels in the vPC. LACP active-active is recommended, followed by active-passive mode and manual or static bundling if the access device does not support LACP. Port-channel in mode active-active is preferred as it initiates more quickly than port-channel in mode active-passive.
- LACP graceful-convergence disabled on port-channels going to Cisco UCS FI. LACP graceful-convergence is ON by default and should be enabled when the downstream access switch is a Cisco Nexus device and disabled if it is not.
- Only required VLANs are allowed on the vPC peer links and member ports – prune all others to minimize internal resource consumption.

- Source-destination IP, L4 port and VLAN are used load-balancing hashing algorithm for port-channels. This improves fair usage of all member ports forming the port-channel. The default hashing algorithm is source-destination IP and L4 port.

vPC Spanning Tree Considerations:

- Bridge Assurance enabled on vPC peer links by specifying **spanning-tree port type network**. Bridge Assurance should be **disabled** on vPC member ports.
- Spanning port type specified as **edge** or **edge trunk** on host facing interfaces connecting to Cisco UCS FI.
- BPDU Guard feature enabled on host-facing interfaces connecting to Cisco UCS FI. This was done by globally enabling it on all edge ports using the **spanning-tree port type edge bpduguard default** command.
- BPDU Filtering feature enabled on host-facing interfaces connecting to Cisco UCS FI. This was done by globally enabling it on all edge ports using the **spanning-tree port type edge bpdufilter default** command.
- Loop Guard was disabled (default setting) in this design. If necessary, they can be enabled globally using **spanning-tree loopguard default** or at the interface level using **spanning-tree guard loop**.
- Root Guard enabled on vPC member ports connected to access devices to ensure that vPC peer switches remain the spanning tree root – using interface level command **spanning-tree guard root**

Other Considerations

Unidirectional Link Detection (UDLD) was enabled globally using **feature UDLD** and on vPC peer links and member ports to Cisco UCS FI.

- HSRP specific
  - Interface VLANs should be defined as passive interfaces to avoid routing peer information
  - Disable IP redirection on HSRP interface VLANs
  - Use default timer for HSRP/VRRP

If the SmartStack design outlined in this CVD is connected to additional aggregation/core layer Cisco Nexus switches in a two-tiered design for scalability or other expansion purposes, the following guidelines should be followed.

- In a two-tiered data center design using Cisco Nexus switches, vPCs can also be used between the Cisco Nexus switches in each tier using a double-sided vPC topology. In such a design, the vPC domain identifiers must be different as this information is exchanged through LACP protocol and using the same vPC domain identifiers will generate continuous flaps on vPC between the different Cisco Nexus network layers.
- If modular Cisco Nexus switches are used, redundancy should be provided by using ports from different line cards.
- Deploy dedicated Layer 3 link(s) for exchanging routing information between peer switches in a two-tiered design or other topologies where Layer 3 is used between the tiers. The vPC peer-link should not be used.

Last but not least, review the criteria for vPC Type-1 and Type-2 consistency checks in the link provided below to avoid issues in your deployment.

1. vPC Design Guide: http://www.cisco.com/c/dam/en/us/td/docs/switches/datacenter/sw/design/vpc_design/vpc_best_practices_design_guide.pdf

2. Cisco Nexus 9000 NX-OS Release 6.x Configuration Guide: http://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/sw/6x/interfaces/configuration/guide/b_Cisco_Nexus_9000_Series_NXOS_Interfaces_Configuration_Guide.html

High Availability

SmartStack platform was designed for maximum availability of the complete infrastructure (compute, network, storage, and virtualization) with no single points of failure.

Compute and Virtualization

Cisco UCS system provides redundancy at the component and link level and end-to-end path redundancy to storage array and LAN network.

Cisco UCS 5108 blade server platform is highly redundant with redundant power supplies, fans and fabric extender modules. C-series rack mount servers also have redundant power supplies and fans. The rack mount servers are directly connected to the upstream fabric interconnects in this SmartStack design.

Each fabric extender on the Cisco UCS 5108 blade server is deployed with 4x10GbE links to the unified fabric. The links are part of a port channel to ensure rerouting of flows to other links in the event of a link failure.

Each server is deployed using vNICs and vHBAs that provide redundant connectivity to the unified fabric. NIC failover is enabled between Cisco UCS Fabric Interconnects using Cisco UCS manager. This is done for all management and virtual machine vNICs.

VMware vCenter is used to deploy VMware HA clusters to allow VMs to failover in the event of a server failure. VMware vMotion and VMware HA are enabled to auto restart VMs after a failure. Host Monitoring is enabled to monitor heartbeats of all ESXi hosts in the cluster for faster detection. Admission Control is also enabled on the blade servers to ensure the cluster has enough resources to accommodate a single host failure.

VMware vSphere hosts use SAN multipathing to access LUNs on the Nimble array. If any component (NIC, HBA, FEX, FI, MDS, Nimble controller, cables) along a path fails, all storage traffic will reroute to an alternate path. When both paths are active, traffic is load balanced.

Network

Link aggregation using port channels and virtual port channels are used throughout the SmartStack design for higher bandwidth and availability.

Port channels are used on unified fabric links between fabric extender and fabric interconnects. Virtual port channels are used between FIs and Cisco Nexus switches. vPCs provider higher availability than port channels as it can continue to forward traffic even if one of the Cisco Nexus switches fail because vPCs distribute member links of port-channel across different Cisco Nexus switches.

Cisco Nexus 9000 series switches are used in the data center LAN fabric to provide redundancy in the event of a switch failure. Cisco MDS switches are used in the SAN fabric to provide redundancy in the event of a switch failure.

MDS and Cisco Nexus switches are highly redundant with redundant power supplies, fans and have out-of-band management access.

The two MDS switches form two separate fabrics and provide two distinct physical paths to storage for redundancy. FI-A to MDS-A to Nimble array is SAN Fabric A and FI-B to MDS-B to Nimble array is SAN Fabric B. Dual VSANs are used across these fabrics with vSAN-A on Fabric A and vSAN-B on Fabric B. The dual vSANs represent two redundant paths to the storage array with traffic load balanced across both vSANs when there is no failure.

Storage

The Nimble CS700 array has redundant storage controllers which allow for an active and standby configuration.

The CS700 has redundant power supplies with diverse cabling and data paths to each controller.

Each Nimble storage controller is redundantly connected to the SAN fabric. Each controller is connected using 4x8Gb FC links to upstream MDS switches with 2x8Gb links going to MDS-A switch and 2x8Gb links going to MDS-B switch in the SAN fabric. This will allow 32GB network bandwidth for each controller.

FC target connections are configured in a Dual fabric / dual vSAN switch fabric. This configuration is used across the SAN fabric and unified fabric for redundant connectivity to storage.

Each Service Profile has a boot profile with redundant paths to primary and secondary FC targets on the Nimble Storage array.

All VMware datastore volumes utilize Nimble PSP_Directed for proper path failover and load distribution.

Scalability

For higher performance and capacity, SmartStack solutions can scale up by adding compute, network, storage or management subsystems individually or scale out with consistent and predictable performance by deploying additional units of the complete SmartStack solution.

Management

Cisco UCS Manager residing on a clustered pair of Cisco UCS Fabric Interconnects that makes up a Cisco UCS domain can manage up to 160 servers (8 servers per chassis x 20 chassis) in a single Cisco UCS domain.

Cisco UCS Central can be deployed to provide a single pane for managing multiple Cisco UCS domains – for up to 10,000 servers. Cisco UCS Central complements Cisco UCS Manager to provide centralized inventory, faults, and logs, global policies and pools, firmware management, global visibility and flexible administrative controls. Cisco UCS Central is a manager of managers that interfaces with Cisco UCS Manager in each domain to manage multiple, globally distributed Cisco UCS domains.

Storage

Scale-to-Fit

With Nimble Storage's CS700 Predictive Flash platform, it is easy to accommodate application growth by scaling performance, capacity, or both—efficiently and non-disruptively. With Nimble Storage scale-to-fit, organizations can:

- Flexibly scale flash to accommodate a wide variety of application working sets. With the addition of an All Flash Shelf, the CS700 can support up to 38 TB in SSD.
- Scale capacity by adding additional HDDs or expansion shelves. The CS700 platform can support an effective capacity of up to 892 TB with expansion shelves.
- Scale up performance by upgrading compute for greater throughput and IOPS
- Scale capacity and performance together by clustering any combination of Nimble Storage arrays – see next section for Nimble's scale-out capabilities

Scale-Out

Nimble Storage's Predictive Flash platform features a scale-out architecture that makes it easy to scale capacity and performance beyond the physical limitations of a single array. With Nimble Storage scale-out, organizations can:

- Group up to 4 Nimble arrays (any model) for higher scalability and performance
- One management console to administer all storage hardware resources in the group as a single entity
- Dynamic load balancing across arrays in a group to eliminate performance hot spots
- Multi-array data striping, enabling any application to fully leverage the collective hardware resources of the scale-out group

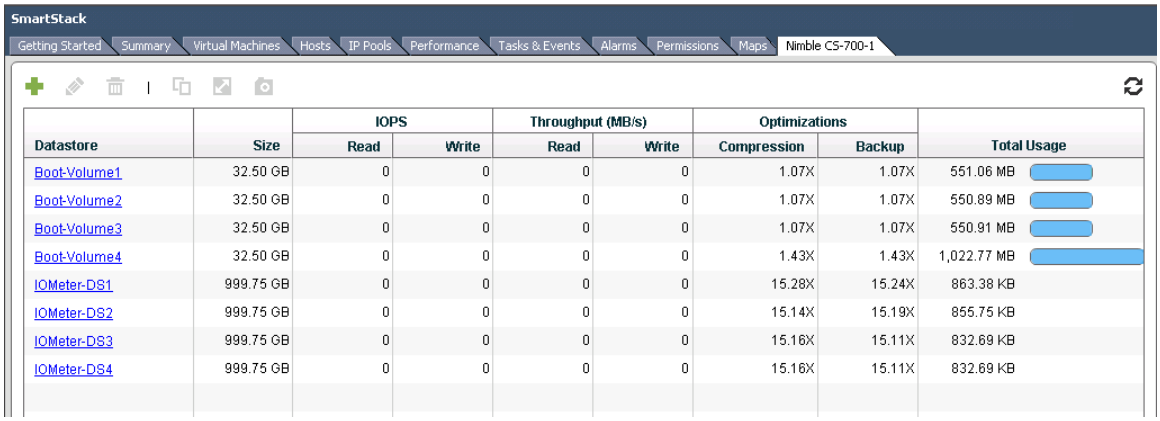
- Flexible configuration of any combination of Nimble Storage arrays in the group, maximizing storage ROI
- Seamless reconfiguration and hardware refreshes, without downtime

Storage management Nimble vCenter Plugin

Nimble Storage provides a plugin that works specifically with vCenter to manage datastores residing on Nimble arrays. You can use either the desktop vCenter plugin, or the web-based vCenter plugin. Both provide the same functionality, with slight variances in the user views, and minor enhancements.

The Nimble vCenter plugin allows the following capability directly from vCenter:

- create, clone, grow, and edit datastores
- take, clone, and delete snapshots
- add Nimble-specific capabilities to vCenter server which can be used to create vCenter roles
- edit protection schedules



Datastore	Size	IOPS		Throughput (MB/s)		Optimizations		Total Usage
		Read	Write	Read	Write	Compression	Backup	
Boot-Volume1	32.50 GB	0	0	0	0	1.07X	1.07X	551.06 MB
Boot-Volume2	32.50 GB	0	0	0	0	1.07X	1.07X	550.89 MB
Boot-Volume3	32.50 GB	0	0	0	0	1.07X	1.07X	550.91 MB
Boot-Volume4	32.50 GB	0	0	0	0	1.43X	1.43X	1,022.77 MB
IOMeter-DS1	999.75 GB	0	0	0	0	15.28X	15.24X	863.38 KB
IOMeter-DS2	999.75 GB	0	0	0	0	15.14X	15.19X	855.75 KB
IOMeter-DS3	999.75 GB	0	0	0	0	15.16X	15.11X	832.69 KB
IOMeter-DS4	999.75 GB	0	0	0	0	15.16X	15.11X	832.69 KB

Validation

A high level summary of the validation done for the SmartStack design is provided in this section. The solution was validated for basic data forwarding by deploying virtual machine running IO Meter tool. The system was validated for resiliency by failing various aspects of the system under load. Examples of the types of tests executed are as follows:

- Failure and recovery of ESXi hosts in a cluster (For example, rebooting of hosts, shutting down of hosts)
- Failure and recovery of redundant links between Cisco UCS FI and Cisco MDS
- Failure and recovery of redundant MDS switches
- Failure and recovery of redundant links between Cisco MDS and Nimble controllers
- Failure and recovery of SSD and spinning disks on Nimble CS700
- Managed failover and recovery of backup and active Nimble controllers from Nimble management interface
- Failure and recovery of backup and active Nimble controllers from Nimble management interface – simulating a hard failure
- Upgrading Nimble OS while system is actively processing I/O
- Failure and recovery of Cisco Nexus switches
- Failure and recovery of redundant links between Cisco UCS FI and Cisco Nexus switches

Load was generated using IOMeter tool and different IO profiles were used to reflect the different profiles that are seen in customer networks. See table below for the profiles used.

Table 3 Traffic Profiles

IO Profiles	IO Performance
I/O Profile 1	8k size, Random, 75% Read, 25% Write,16 Outstanding IOs
I/O Profile 2	4k size, Sequential, 100% Reads
I/O Profile 3	8k Size, Random, 50% Read, 50% Write

Validated Hardware and Software

Table below lists all the components and software versions used in validating the SmartStack design.

Table 4 Validated Infrastructure Components and Software Revisions

	Components	Software Version	Comments

Network	Cisco Nexus 9372PX (N9k-9372PX)	6.1(2)I3(5)	Cisco Platform Switch for ToR, MoR, EoR deployments; Provides connectivity to users and other networks and deployed in NX-OS Standalone mode
	Cisco Nexus 1000V	5.2(1)SV3(1.15)	(Optional) Distributed Virtual Switch
	Cisco UCS 6248UP FI	3.1.1g	Fabric Interconnect with embedded management
	Cisco MDS 9148S	6.2(13a)	16G Multilayer Fabric Switch
Compute	Cisco UCS 5108	3.1.1g	Blade Server Chassis
	Cisco UCS B200M4 servers with E5-2600 v4 processors	3.1.1g	Blade Servers
	Cisco UCS C220M4 servers with E5-2600 v4 processors	3.1.1g	Rack mount Server
	Cisco ENIC Driver	2.3.0.7	Cisco VIC Ethernet driver
	Cisco FNIC Driver	1.6.0.25	Cisco VIC FCoE driver
	Adapter Firmware	4.1(1d)	Cisco VIC Adapter Firmware
Management	Cisco UCS Manager	3.1.1g	Embedded Management
	vCenter plugin for Nimble Storage	Follows Nimble OS	
	Cisco UCS Performance Manager	2.0.0	Performance Monitoring Tool
Storage	Nimble CS700	NimbleOS 2.3.14	
	Nimble NCM for ESXi	2.3.1	Build: 2.3.1-600006
Virtualization	VMware vSphere	6.0 U1b	Custom Cisco ISO with updated driver
	VMware vCenter Server	6.0 U1	Appliance
Tools	Workload – IOMeter Tool		
Other	Microsoft Active Directory/DNS		

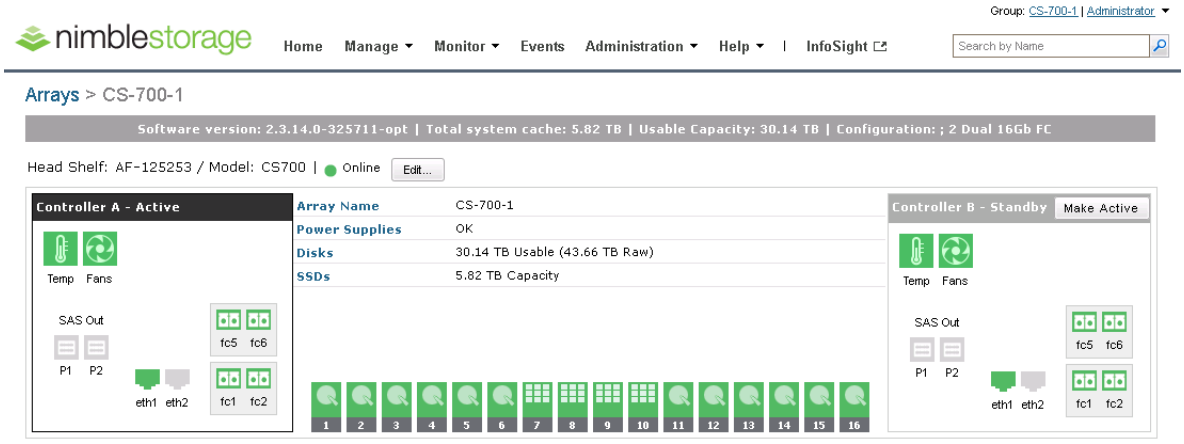
Supported SmartStack Software Versions

SmartStack was validated using the software versions outlined in the table above but other software versions are also supported provided the components and releases are listed in the interoperability matrices below that Cisco and Nimble Storage provides.

- Cisco UCS Hardware and Software Interoperability Tool: http://www.cisco.com/web/techdoc/ucs/interoperability/matrix/matrix.html
- Interoperability Matrix for Cisco Nexus and MDS 9000 Products: http://www.cisco.com/c/en/us/support/storage-networking/mds-9000-nx-os-san-os-software/products-device-support-tables-list.html
- Nimble Support Matrix: https://infosight.nimblestorage.com/InfoSight/media/cms/active/pubs_support_matrix_2_3_rev_k.pdf (requires login)
- VMware Compatibility Guide: http://www.vmware.com/resources/compatibility/search.php

The figure below shows a detailed view of the Nimble storage array configuration.

Figure 16 Nimble Storage Array Configuration



Bill of Materials (BOM)

The BOM below lists the major components validated but it is **not** intended to be a comprehensive list.

Table 5 SmartStack Bill of Materials

Line	SKU	Description	Quantity
1.0	**UCSB-B200-M4**	**Cisco UCS B200 M4 w/o CPU, mem, drive bays, HDD, mezz**	**1**
1.1	UCS-CPU-E52660E	2.00 GHz E5-2660 v4/105W 14C/35MB Cache/DDR4 2400MHz	2
1.2	UCS-MR-1X161RV-A	16GB DDR4-2400-MHz RDIMM/PC4-19200/single rank/x4/1.2v	8
1.3	UCSB-MLOM-40G-03	Cisco UCS VIC 1340 modular LOM for blade servers	1
1.4	UCSB-HS-EP-M4-F	CPU Heat Sink for Cisco UCS B200 M4/B420 M4 (Front)	1
1.5	UCSB-HS-EP-M4-R	CPU Heat Sink for Cisco UCS B200 M4/B420 M4 (Rear)	1
1.6	UCSB-LSTOR-BK	FlexStorage blanking panels w/o controller, w/o drive bays	2
1.7	C1 UCS-OPT-OUT	Cisco ONE Data Center Compute Opt Out Option	1
2.0	**UCSB-5108-AC2-UPG**	**Cisco UCS 5108 Blade Server AC2 Chassis, 0 PSU/8 fans/0 FEX**	**1**
2.1	N01-UAC1	Single phase AC power module for Cisco UCS 5108	1
2.2	N20-FAN5	Fan module for Cisco UCS 5108	8
2.3	N20-CBLKB1	Blade slot blanking panel for Cisco UCS 5108/single slot	7
2.4	N20-CAK	Accessory kit for Cisco UCS 5108 Blade Server Chassis	1
2.5	N20-FW014	Cisco UCS Blade Server Chassis FW Package 3.1	1
2.6	Cisco UCSB-5108-PKG-HW	Cisco UCS 5108 Packaging for chassis with half width blades.	1
2.7	Cisco UCSB-PSU-2500ACDV	2500W Platinum AC Hot Plug Power Supply - DV	2
2.8	CAB-C19-	Cabinet Jumper Power Cord, 250 VAC 16A, C20-C19	2

	CBN	Connectors	
2.9	Cisco UCS-IOM-2208XP	Cisco UCS 2208XP I/O Module (8 External, 32 Internal 10Gb ports)	2
3.0	**UCS-FI-6248UP-UPG**	**Cisco UCS 5108 Blade Server AC2 Chassis, 0 PSU/8 fans/0 FEX**	**1**
3.1	N10-MGT014	Cisco UCS Manager 3.1	2
3.2	Cisco UCS-ACC-6248UP	Cisco UCS 6248UP Chassis Accessory Kit	1
3.3	Cisco UCS-PSU-6248UP-AC	Cisco UCS 6248UP Power Supply/100-240VAC	2
3.4	Cisco UCS-BLKE-6200	Cisco UCS 6200 Series Expansion Module Blank	1
3.5	Cisco UCS-FAN-6248UP	Cisco UCS 6248UP Fan Module	2
3.6	Cisco UCS-FI-DL2	Cisco UCS 6248 Layer 2 Daughter Card	1
3.7	CAB-9k12A-NA	Power Cord, 125VAC 13A NEMA 5-15 Plug, North America	2
3.8	Cisco UCS-L-6200-10G-C	2nd Gen FI License to connect C-direct only	1
4.0	**UCSC-C220-M4S**	**Cisco UCS C220 M4 SFF w/o CPU, mem, HD, PCIe, PSU, rail kit**	**1**
4.1	Cisco UCS-CPU-E52620D	2.40 GHz E5-2620 v3/85W 6C/15MB Cache/DDR4 1866MHz	2
4.2	Cisco UCS-MR-1X162RU-A	16GB DDR4-2133-MHz RDIMM/PC4-17000/dual rank/x4/1.2v	8
4.3	Cisco UCSC-PCIE-CSC-02	Cisco VIC 1225 Dual Port 10Gb SFP+ CNA	1
4.4	Cisco UCSC-RAIL-NONE	NO RAIL KIT OPTION	1
4.5	Cisco UCSC-PSU1-770W	770W AC Hot-Plug Power Supply for 1U C-Series Rack Server	2

4.6	CAB-9K12A-NA	Power Cord, 125VAC 13A NEMA 5-15 Plug, North America	2
4.7	Cisco UCSC-HS-C220M4	Heat sink for Cisco UCS C220 M4 rack servers	2
4.8	Cisco UCSC-MLOM-BLK	MLOM Blanking Panel	1
4.9	N20-BBLKD	Cisco UCS 2.5 inch HDD blanking panel	8
4.10	C1UCS-OPT-OUT	Cisco ONE Data Center Compute Opt Out Option	1
5.0	**DS-C9148S-D12P8K9**	**MDS 9148S 16G FC switch, w/ 12 active ports + 8G SW SFPs**	**1**
5.1	DS-SFP-FC8G-SW	8Gbps Fibre Channel SW SFP+, LC	12
5.2	DS-9148S-KIT-CSCO	MDS 9148S Accessory Kit for Cisco	1
5.3	CAB-9K12A-NA	Power Cord, 125VAC 13A NEMA 5-15 Plug, North America	2
6.0	**N9K-C9372PX**	**Nexus 9300 with 48p 10G SFP+ and 6p 40G QSFP+**	**1**
6.1	N9KDK9-612I3.5	Nexus 9500 or 9300 Base NX-OS Software Rel 6.1(2)I3(5)	1
6.2	N3K-C3064-ACC-KIT	Nexus 3K/9K Fixed Accessory Kit	1
6.3	NXA-FAN-30CFM-F	Nexus 2K/3K/9K Single Fan, port side exhaust airflow	4
6.4	CAB-9K12A-NA	Power Cord, 125VAC 13A NEMA 5-15 Plug, North America	2
6.5	N9K-PAC-650W-B	Nexus 9300 650W AC PS, Port-side Exhaust	2
7.0	**CS700-4F-48T-6400FS**	**Nimble CS700 FC Connectivity with 12 x 4TB HDDs and 4 x 1.6TB SSDs**	**1**

Summary

SmartStack delivers an infrastructure platform for Enterprise and cloud data centers using Cisco and Fibre Channel-attached Nimble CS700 array. SmartStack is designed and validated using compute, network and storage best practices and high availability to reduce deployment time, project risk and IT costs while maintaining scalability and flexibility for addressing a multitude of IT initiatives.

About Authors

Archana Sharma, Technical Leader, Cisco UCS Solutions Engineering, Cisco Systems Inc.

Archana Sharma has 20 years of experience at Cisco focused on Data Center, Desktop Virtualization, Collaboration and related technologies. Archana has been working on Enterprise and Service Provider systems and solutions and delivering Cisco Validated designs for over 10 years. Archana holds a CCIE (#3080) in Routing and Switching and a Bachelor's degree in Electrical Engineering from North Carolina State University.

Steve Sexton, Technical Marketing Engineer, Nimble Storage Inc.

Steve Sexton has over 15 years of experience in both Network and Storage industries. For the last five years he has been focused on Cisco UCS technologies and integration efforts. He holds Bachelor's and Master's degrees in Industrial Technology from Appalachian 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:

- Chris O' Brien, Director, Cisco UCS Solutions Technical Marketing Team, Cisco Systems Inc.
- Jawwad Memon, Product Manager, Cisco UCS Product Management and Solutions, Cisco Systems Inc.
- Ashish Prakash, VP Solutions Engineering, Nimble Storage Inc.
- Arun Garg, Director, Solutions Product Management, Nimble Storage Inc.
- Matt Miller, Senior Product Marketing Manager, Nimble Storage Inc.