**FlexPod SolidFire with VMware 6.5 Design Guide**

Last Updated: August 15, 2017

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_2.png)

About the Cisco Validated Design (CVD) Program

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

[http://www.cisco.com/go/designzone](https://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.

CCDE, CCENT, Cisco Eos, Cisco Lumin, Cisco Nexus, Cisco StadiumVision, Cisco TelePresence, Cisco WebEx, the Cisco logo, DCE, and Welcome to the Human Network are trademarks; Changing the Way We Work, Live, Play, and Learn and Cisco Store are service marks; and Access Registrar, Aironet, AsyncOS, Bringing the Meeting To You, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, CCVP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unified Computing System (Cisco UCS), Cisco UCS B-Series Blade Servers, Cisco UCS C-Series Rack Servers, Cisco UCS S-Series Storage Servers, Cisco UCS Manager, Cisco UCS Management Software, Cisco Unified Fabric, Cisco Application Centric Infrastructure, Cisco Nexus 9000 Series, Cisco Nexus 7000 Series. Cisco Prime Data Center Network Manager, Cisco NX-OS Software, Cisco MDS Series, Cisco Unity, Collaboration Without Limitation, EtherFast, EtherSwitch, Event Center, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, iQuick Study,  LightStream, Linksys, MediaTone, MeetingPlace, MeetingPlace Chime Sound, MGX, Networkers, Networking Academy, Network Registrar, PCNow, PIX, PowerPanels, ProConnect, ScriptShare, SenderBase, SMARTnet, Spectrum Expert, StackWise, The Fastest Way to Increase Your Internet Quotient, TransPath, WebEx, and the WebEx logo are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0809R)

© 2017 Cisco Systems, Inc. All rights reserved.

Table of Contents

[Executive Summary](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567261)

[Solution Overview](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567262)

[Introduction](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567263)

[Audience](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567264)

[What’s New?](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567265)

[Solution Summary](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567266)

[Technology Overview](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567267)

[FlexPod SF with VMware vSphere 6.5 Solution Overview](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567268)

[Solution Component Overview](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567269)

[Cisco Unified Computing System](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567270)

[Cisco UCS Differentiators](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567271)

[Cisco Nexus 9000 Series Switches](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567272)

[NetApp SolidFire SF9608 Storage](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567273)

[VMware vSphere 6.5](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567274)

[Solution Design](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567275)

[FlexPod SF Design Overview](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567276)

[FlexPod SF Design Options](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567277)

[Design 1: End-to-End 10GbE iSCSI Architecture](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567278)

[Design 2: 40GbE iSCSI Architecture](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567279)

[Design 3: End-to-End 10GbE iSCSI Design using Cisco UCS Mini](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567280)

[FlexPod SF Design](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567281)

[Compute Layer Design](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567282)

[Virtualization Layer Design](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567283)

[LAN Network Design](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567284)

[Storage Design](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567285)

[High Availability](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567286)

[Scalability](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567287)

[Other Design Considerations](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567288)

[Management Connectivity](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567289)

[QoS and Jumbo Frames](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567290)

[Cisco Nexus 9000 Series – vPC Best Practices](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567291)

[Solution Validation](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567292)

[Validated Hardware and Software](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567293)

[Hardware](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567294)

[Software](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567295)

[Summary](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567296)

[References](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567297)

[Products and Solutions](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567298)

[Interoperability Matrixes](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567299)

[About the Authors](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567300)

[Acknowledgements](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.html#_Toc490567301)

**Executive Summary**

Cisco Validated Designs (CVDs) include platforms and solutions that are designed, tested, and documented to improve customer deployments. These designs include a wide range of technologies and products into a portfolio of solutions that address the business needs of customers.

This document describes the Cisco and NetApp® FlexPod® converged infrastructure with the NetApp SolidFire® solution. NetApp SolidFire is a block-based all flash storage solution designed for highly virtualized and automated environments. The FlexPod SolidFire (FlexPod SF) with VMware vSphere 6.5 solution enables rapid development and deployment of applications, services, and innovation.

The FlexPod SF with VMware vSphere 6.5 Solution is a predesigned, best-practice data center architecture built on the Cisco Unified Computing System (Cisco UCS), the Cisco Nexus 9000 Cloud Scale switches, and the NetApp SolidFire storage. The design details and best practices of this new architecture are covered in the following sections.

**Solution Overview**

**Introduction**

Industry trends indicate a vast data center transformation toward shared infrastructure and cloud computing. Business agility requires application agility, so IT teams need to provision applications in hours instead of weeks. Resources need to scale up (or down) in minutes, not hours.

Cisco and NetApp have developed a NetApp SolidFire (SF) solution to simplify application deployments and accelerate productivity.  FlexPod SF minimizes complexity making room for better, faster IT service, increased productivity and innovation across the business.

**Audience**

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

**What’s New?**

FlexPod SF infrastructure solution is a predesigned, best-practice data center architecture built on the Cisco UCS B-Series compute infrastructure, the Cisco Nexus 9000 Cloud Scale switches, and the NetApp SolidFire all flash storage.

**Solution Summary**

The FlexPod SF with VMware vSphere 6.5 infrastructure solution provides the same reliable converged infrastructure solution that Cisco and NetApp have been developing for years, with the additional benefit of the SolidFire all flash storage platform. The SolidFire storage platform allows you to build environments faster, develop environments sooner, and run multiple projects on one shared infrastructure platform.

**Technology Overview**

**FlexPod SF with VMware vSphere 6.5 Solution Overview**

FlexPod SF is a converged infrastructure platform for Enterprise datacenter and cloud deployments based on the following components:

         Cisco Unified Computing System (Cisco UCS)

         Cisco Nexus Cloud Scale Switches

         NetApp SolidFire All Flash Storage (SF9608)

         VMware vSphere 6.5

The FlexPod SF with VMware vSphere 6.5 architecture covered in this document uses the following models of the above components for validation of the solution design:

         Cisco Unified Computing System

         Cisco UCS 5108 Blade Server chassis with 2200 Series Fabric Extenders (FEX) and VIC Adapters: [Cisco UCS 5108 Blade Server Chassis](https://www.cisco.com/en/US/products/ps10279/index.html) supports up to eight blade servers and two fabric extenders in a six-rack unit (RU) enclosure. [Cisco UCS VIC Adapters](https://www.cisco.com/en/US/products/ps10277/prod_module_series_home.html) wire-once architecture offers a range of options to converge the fabric, optimize virtualization and simplify management.

         Cisco UCS B-Series Blade Servers provide performance, efficiency, versatility and productivity with a range of Intel based blade servers.

         [Cisco UCS 6200 Series Fabric Interconnects](https://www.cisco.com/en/US/products/ps11544/index.html) (FI) provide a family of line-rate, low-latency, lossless, 10Gbps Ethernet and Fibre Channel over Ethernet interconnect switches providing the management and communication backbone for the Cisco Unified Computing System. For a 40GbE design, Cisco UCS 6300 Series Fabric Interconnects can also be used. For additional information, see section Design 2: 40GbE iSCSI Architecture.

         [Cisco UCS Manager](https://www.cisco.com/en/US/products/ps10281/index.html) provides unified management of all software and hardware components in the Cisco Unified Computing System and manages servers, networking and access to storage from a single interface.

         Cisco Nexus 93180YC-EX Series Cloud Scale Switches deliver a number of innovative and advanced capabilities for next generation data centers. Cisco Nexus 93180YC-EX Switch used in this design comes with 48 1/10/25G-Gbps Small Form Pluggable Plus (SFP+) ports and 6 40/100-Gbps Quad SFP+ (QSFP+) uplink ports. All ports are line rate, delivering 3.6 Tbps of throughput in a 1-rack-unit (1RU) form factor. Cisco Nexus 93180 YC-EX can be deployed in NX-OS standalone or Application Centric Infrastructure (ACI) mode across enterprise, service provider, and Web 2.0 data centers. Cisco Nexus 9000-EX series switches are designed with innovative Cisco Cloud Scale technology for supporting highly scalable cloud architectures.

         NetApp SolidFire SF9608 All Flash Storage: The NetApp SolidFire SF9608 storage appliance is an iSCSI-based, all-flash storage solution based on the Cisco UCS C220 M4S Server platform.  Leveraging industry-leading QoS and scale-out capabilities, NetApp SolidFire can deliver guaranteed performance to hundreds of applications on a single, highly-available storage array. SolidFire storage nodes leverage high-bandwidth 10 GbE connectivity to provide up to 20Gbps throughput per storage node.  The “shared-nothing”, scale-out architecture of NetApp SolidFire facilitates scaling storage capacity and performance when adding additional nodes to the storage array, all with no downtime

The FlexPod SF architecture provides design flexibility and supports the following components as an alternative to models listed above:

         [Cisco UCS Mini](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-mini/index.html) is a smaller, more compact version of Cisco UCS, optimized for branch and remote offices and smaller IT environments. Cisco UCS Mini delivers servers, storage, and 10 Gigabit networking in an easy-to-deploy, compact form factor. The Fabric Interconnects are integrated into the chassis, reducing power and cooling needs, and overall footprint. With a server capacity of sixteen half-width blade servers, Cisco UCS Mini is ideal for customers that need a small deployment but with the management simplicity of Cisco UCS Manager.

         [Cisco UCS 6300 Series Fabric Interconnects](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6300-series-fabric-interconnects/index.html) is the third-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 high-capacity data centers. Fabric Interconnects are the backbone of the Cisco Unified Computing System, delivering a unified fabric with centralized management.

The FlexPod SF with VMware vSphere 6.5 architecture can also be enhanced to address specific operational and deployment needs through the use of the following components:

         [Cisco UCS Central](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-central-software/index.html) provides a scalable management platform for managing multiple, globally distributed Cisco UCS domains with consistency. Cisco UCS Central integrates with Cisco UCS Manager in each domain to provide global and uniform configuration and maintenance of service profiles and configurations.

         [Cisco UCS Director](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-director/index.html) replaces time-consuming, manual provisioning and de-provisioning of data center resources with automated workflows.  It supports multiple hardware vendors and multiple hypervisors to deliver heterogeneous infrastructure management.

         [Cisco Cloud Center](https://www.cisco.com/c/en/us/products/cloud-systems-management/cloudcenter/index.html), formerly Cliqr, can securely extend FlexPod SF resources between private and public cloud.  Application blueprints rapidly scale the user’s needs, while implementing them in a manner that is secure, and provides the controls for proper governance.

These components are integrated using technology and product best practices recommended by Cisco and NetApp to deliver a converged infrastructure platform capable of supporting a variety of applications, from traditional Enterprise applications to web scale applications. This FlexPod SF with VMware vSphere 6.5 solution is designed to be highly available, with redundancy in compute, networking and storage layers of the stack. There is no single point of failure from a device or traffic path perspective. For higher performance and capacity, FlexPod SF can either scale-up by upgrading or adding compute, network and storage resources individually or scale-out by deploying additional instances of the entire stack. FlexPod solutions can do this while maintaining consistency in design and component best practices.

**Solution Component Overview**

This section provides a technical overview of the individual compute, network, storage and management components of the NetApp SolidFire solution.

**Cisco Unified Computing System**

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 or 40 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 Cisco Unified Computing System consists of the following subsystems:

         **Compute** - The compute piece of the system incorporates servers based on latest Intel’s x86 processors. Servers are available in blade form factor, managed by the same Cisco UCS Manager software.

         **Network** - The integrated network fabric in the system provides low-latency, lossless, 10Gbps or 40Gbps Ethernet fabric. LANs, SANs, and high-performance computing networks which are separate networks today are a consolidated within the fabric. 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.

         **Management:** The system uniquely integrates compute, network and storage access subsystems, enabling it 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.

Cisco’s Unified Compute System has revolutionized the way servers are managed in data-center.  The next section takes a detailed look at the unique differentiators in Cisco UCS and Cisco UCS Manager. An overview of the key sub-components leveraged in this FlexPod SF architecture are also provided.

**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) 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, 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.

         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, and 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 Fabric Interconnects**

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

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

The Cisco UCS Fabric Interconnect family comprises of Cisco UCS 6200 and Cisco UCS 6300 series of Fabric Interconnects.

**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 connects to the FI using a single Ethernet cable for carrying all network, storage and management traffic. Cisco UCS Fabric Interconnects provides uniform network and storage access to servers and is 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 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.

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          A pair of Cisco UCS 6248UP Fabric Interconnects were used in validation 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](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6200-series-fabric-interconnects/index.html)

Figure 1       Cisco UCS 6248UP Fabric Interconnect  
[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_3.jpg)

**Cisco UCS 6300 Series Fabric Interconnects**

Cisco UCS 6300 Series Fabric Interconnects provide a 40Gbps unified fabric with double the switching capacity of the Cisco UCS 6200 Series and support higher workload density. When combined with the newer 40Gbps Cisco UCS 2300 Series Fabric Extenders (see the following section for more information), they provide 40GbE / FCoE port connectivity to enable end-to-end 40GbE / FCoE solution. The unified ports now support 16G FC for high speed FC connectivity to SAN.

Two 6300 Fabric Interconnect models have been introduced supporting Ethernet, FCoE, and FC ports.

Cisco UCS 6332 Fabric Interconnects is a one-rack-unit (1RU) 40 Gigabit Ethernet, and FCoE switch offering up to 2.56 Tbps throughput and up to 32 ports. The switch has 32 40 Gbps fixed Ethernet, and FCoE ports. This Fabric Interconnect is targeted for IP storage deployments requiring high-performance 40 Gbps FCoE connectivity to Cisco MDS switches.

Figure 2       Cisco UCS 6332 Fabric Interconnect – Front and Rear

[https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_4.png](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_4.png)

[https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_5.png](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_5.png)

Cisco UCS Fabric Interconnects 6332-16UP is a one-rack-unit (1RU) 40 Gigabit Ethernet/FCoE switch and 1/10 Gigabit Ethernet, FCoE and Fibre Channel switch offering up to 2.24 Tbps throughput and up to 40 ports. The switch has 24 40Gbps fixed Ethernet/FCoE ports and 16 1/10Gbps Ethernet/FCoE or 4/8/16Gbps Fibre Channel ports. This Fabric Interconnect is targeted for FC storage deployments requiring high performance 16Gbps FC connectivity to Cisco MDS switches.

Figure 3       Cisco 6332-16UP Fabric Interconnect – Front and Rear

[https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_6.png](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_6.png)

[https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_5.png](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_5.png)

The capabilities of all Fabric Interconnects are summarized below.

Table 1    Cisco UCS 6200 and 6300 series Fabric Interconnects

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

   \* Using 40G to 4x10G breakout cables          \*\* Requires QSA module

**Cisco UCS Manager**

Cisco UCS Manager provides unified, embedded management for all software and hardware components in Cisco UCS. Using [Cisco Single Connect](https://www.cisco.com/c/en/us/products/servers-unified-computing/singleconnect.html) 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 or 6300 Series Fabric Interconnects using a clustered, active-standby configuration for high availability.

Cisco UCS Manager offers a 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](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-manager/index.html)

**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 supported blade servers, see: [http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-b-series-blade-servers/index.html](https://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. Two fabric extenders can be deployed to provide higher uplink bandwidth with redundancy. A passive mid-plane provides up to 80 Gbps of I/O bandwidth per server slot and can support future 40 GbE 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](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-5100-series-blade-server-chassis/index.html)

Figure 4       Cisco UCS 5108 Blade Server Chassis

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_7.png)

**Cisco UCS Fabric Extenders**

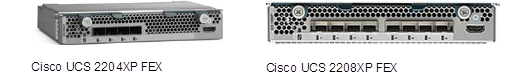
The Cisco UCS Fabric extenders multiplexes and forwards all traffic from servers in a blade server chassis to a parent Cisco UCS Fabric Interconnect over a 10Gbps or 40Gbps unified fabric links. All traffic, including traffic between servers on the same chassis, or between virtual machines on the same 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 is comprised of the Cisco UCS 2200 and 2300 series fabric extenders for the blade server chassis. The three available models of Cisco UCS Fabric Extenders for the blade server chassis are as follows:

         **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.

         **Cisco UCS 2208XP Fabric Extender** has eight 10 Gigabit Ethernet, FCoE-capable, Enhanced Small Form-Factor Pluggable (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.

Figure 5       Cisco UCS 2204XP/2208XP Fabric Extender

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_8.png)

         **Cisco UCS 2304XP Fabric Extender** has four 40 Gigabit Ethernet, FCoE-capable, Quad Small Form-Factor Pluggable (QSFP+) ports that connect the blade chassis to the fabric interconnect. Each Cisco UCS 2304XP has eight 40 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 320 Gbps of I/O to the chassis.

Figure 6       Cisco UCS 2304 Fabric Extender

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_9.png)  
For more information, see, [Cisco UCS 2200 and 2300 Series Fabric Extenders](https://www.cisco.com/c/en/us/products/servers-unified-computing/fabric-interconnects.html)

**Cisco UCS B200 M4 Servers**

The enterprise-class Cisco UCS B200 M4 Blade Server extends the capabilities of the Cisco 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](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-b200-m4-blade-server/index.html)

Figure 7       Cisco UCS B200 M4 Blade Server  
[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_10.jpg)

The Cisco UCS Virtual Interface Card (VIC) 1340 is a 40-Gbps Ethernet or 4 x 10-Gbps Ethernet, FCoE-capable modular LAN on motherboard (mLOM) designed exclusively for the Cisco UCS B200 Series Blade Servers. When used in combination with an optional port expander, the Cisco UCS VIC 1340 capabilities provides an additional 40Gbps of uplink bandwidth. 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](https://www.cisco.com/c/en/us/products/interfaces-modules/ucs-virtual-interface-card-1340/index.html)

Figure 8       Cisco VIC 1340

**Cisco Nexus 9000 Series Switches**

Cisco Nexus 9000 Series Cloud Scale switches provide an Ethernet switching fabric for building next-generation data centers. The new generation of Cisco Nexus 9000 Series Switches support the innovative Cloud Scale ASIC technology that brings the following advanced capabilities to the FlexPod SF solution.

         Cloud Scale ASICs address network congestion and application needs through enhanced queueing and traffic management features.Traditional buffer management treats both small and large flows the same leading to increased latency and impact to small flows. However, with the intelligent buffer management innovations in Cloud Scale ASICs, the performance for small flows are now similar to that of large sized flows resulting in overall application performance, including IP Storage traffic.

         Lossless Fabric which is critical to any storage traffic

         Cisco Tetration Analytics® provides pervasive application visibility and insights into the data center infrastructure. Cisco Tetration also enables advance machine learning, network and security analytics, and network forensics, allowing you to inspect every flow, every packet, at any time. Telemetry is captured from every packet at line rate, including 100G ports with zero CPU Impact. The granular visibility and insight into the data center traffic is critical for developing policies and securing the datacenter as workloads become more distributed and move between on-prem and cloud.

         Provides 7 times more endpoint density with 16nm ASIC technology for greater scale and performance at lower cost and energy savings. Optimized for spine-leaf architecture and standalone models, it can deliver 1.8 Tbps of bandwidth.

Cisco Nexus 9000 Series Cloud Scale Switches also provide the following benefits:

         In NXOS standalone mode, Nexus continues to offer a feature rich, scalable, extensible and robust platform with open API and programmability for DevOps and Automation with 3rd party tools such as Puppet, Chef and Ansible. This allows IT to be more agile and responsive in addressing business needs.

         Migration path to Application Centric Infrastructure (ACI) mode that uses an application centric policy model with simplified automation and centralized management

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          The Cisco Nexus 93180YC-EX Switches in this design are deployed in standalone mode.

The Cisco 93180YC-EX switches used in this design offers the following benefits:

         High-performance, non-blocking architecture in a 1RU form factor with 1.8 Tbps of bandwidth, latency of less than 2 microseconds

         8 fixed 1/10/25-Gbps SFP+ ports and 6 fixed 40/100-Gbps QSFP+ for uplink connectivity

         1+1 redundant hot-swappable power supplies and 3+1 hot swappable redundant fan trays

         Simplified Operations through PXE and POAP support, automation and configuration management with DevOps tools (Puppet, Chef, Ansible) and Open NX-OS, API support for HTTP(s) management through RPC, JSON or XML and Python based programmatic on-box or off-box access to switch.

         Built-in hardware sensors for Cisco Tetration Analytics enable line-rate data collection and rich telemetry. Visibility into real-time buffer utilization per port and per queue enable monitoring of micro-bursts and application traffic patterns.

         Investment Protection through 10/25 Gbps access connectivity and 40/100 Gbps uplink connectivity that facilitate migration to faster speeds. Use of Cisco 40Gbps bidirectional transceiver enables reuse of existing 10 Gigabit Ethernet multimode cabling for 40 Gigabit Ethernet.

         For more information, see: [http://www.cisco.com/c/en/us/support/switches/nexus-93180yc-ex-switch/model.html#~tab-documents](https://www.cisco.com/c/en/us/support/switches/nexus-93180yc-ex-switch/model.html#~tab-documents)

In summary, the FlexPod SF design with Cisco Nexus Cloud Scale switches can not only improve response time and performance of application traffic, including storage traffic, it can also provide innovative features and capabilities important for next generation data center.

**NetApp SolidFire SF9608 Storage**

A primary benefit of FlexPod is that it allows organizations to right-size their deployments on a granular basis as needs dictate. NetApp SolidFire adds to the FlexPod portfolio of storage solutions and delivers a scale-out, all-flash storage for the next generation datacenter. NetApp SolidFire is a programmable storage architecture to enable datacenter automation. Starting with a minimum of (4) nodes, customers can scale their storage cluster as needed based on tenant or workload requirements, all while configuring individual performance guarantees for each tenant/workload.

Figure 9 illustrates the front and rear of the SolidFire SF9608 node.

Figure 9       NetApp SolidFire SF9608

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_12.png)

[https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_13.png](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_13.png)

SolidFire SF9608 comes in a 1RU form factor with the following specifications:

         CPU: 2 x 2.6GHz CPU (E5-2640v3)

         Memory: 256 GB RAM

         8 x 960GB SSD drives (non-SED)

SF9608 comes with the following physical interfaces:

         2 x 1 GbE interfaces (Bond1G): Optionally used as the in-band management interface of the node, and/or the API endpoint

         2 x 10 GbE interfaces (Bond10G): Used for iSCSI storage and intra-cluster communication

         1 x 1 GbE interface (CIMC) – Out-of-Band management of the node

Each pair of 1GbE and 10GbE interfaces are bundled into a single logical interface (Bond1G, Bond10G) to facilitate redundancy and higher aggregate uplink bandwidth.  By default, both bonds are set into Active/Passive bond mode.  LACP bond mode is recommended and used on the storage 10G uplinks.

The bonded interfaces are used to provide storage and management connectivity as outlined below:

**Management IP Interface (MIP)**

         MIP is the “per-node” logical management interface and is used for management of the individual storage node through the WebUI, as well as the API endpoint for the node.

         By default, the MIP is bound to the Bond1G interface (2 x 1GbE ports). However, after the initial configuration, the MIP can be moved to the storage interface (Bond10G).

**Storage IP Interface (SIP)**

         SIP is the “per-node” interface used to terminate incoming iSCSI connections for iSCSI storage access. By default, the SIP is bound to the Bond10G interface (2 x 10GbE).

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          Only IPv4 is supported on SolidFire at this time.

**SolidFire Cluster**

A SolidFire storage array, also known as SolidFire cluster, is comprised of a minimum of (4) SF9608 storage nodes, scalable up to 40 nodes in a single storage cluster.

Within the cluster, each SolidFire node’s CPU, networking, and storage contributes to the available resources of the overall cluster.  As nodes are added or removed from the cluster, the cluster’s overall available resources scale linearly.

Initially all SolidFire storage nodes exist as stand-alone nodes.  When each node is configured with the proper network settings and proposed cluster name, a cluster can be created.

With the initial cluster creation, one of the nodes out of the group of nodes to comprise the cluster is elected to be the cluster master. The cluster master is responsible for overall cluster management and decision making as well as redirecting iSCSI initiators to the proper member storage node(s) for volume access.

The following cluster-wide interfaces are created on the master node:

         Cluster Management Virtual IP (MVIP) – Cluster-wide API endpoint; Cluster-wide WebUI management IP

         Cluster Storage Virtual IP (SVIP) – Cluster-wide iSCSI target interface

When all nodes are initially added to a cluster, their drives are added to the cluster-wide resource pool as storage, and dictates the overall storage capacity.  The cluster is ready for use.

Should the master node of the cluster suffer a failure or be manually removed, a new cluster master will be automatically elected from the remaining cluster nodes.  The cluster MVIP and SVIP will move to the newly selected cluster master.

**Drives**

Each SF9608 node contains 8 SSD drives that are used to store a portion of the data for the cluster. Data is distributed so that individual drive or node failures do not affect availability or performance. Two types of drives exist in a SolidFire storage node:

         Metadata drives: Store compressed information that defines and tracks each volume, clone, or NetApp Snapshot® within a cluster.

         Block drives: Store the compressed, de-duplicated data blocks for server application volumes.

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          The drive in slot one acts as both the boot and metadata drive. It is important to note that if the slot1 drive fails, the SF node will become unresponsive and require an immediate replacement of the failed drive. However, the cluster will work around the failed node and provide access to the same data through other nodes. From a client perspective, iSCSI timeouts will determine failover to other nodes to maintain access to data.

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          The drives currently available on SF9608 nodes are not self-encrypting (non-SED) drives so encryption at rest is currently not available.

**Volumes**

         Volumes (Provisioned banks of storage) in a SolidFire storage cluster, are assigned to an account and are accessed through a unique iSCSI target IQN. Volumes possess the following characteristics:

         Each volume has a unique iSCSI IQN address and a single LUN on that target.

         When a volume is provisioned, it consumes only a small amount of space on the metadata drives.

         Each volume is associated with a single tenant account, (or owner) at the time of creation. The account is used for security control, iSCSI access, and reporting purposes.

         SolidFire uses Advanced Format (4K) blocks internally. Volumes can be accessed using 4K native (4Kn) sectors for compatible operating systems, or 512 emulation (512e) for legacy operating systems. When partitions are properly aligned, there is no performance effect from using 512e mode.

         SolidFire Helix®: The metadata for an individual volume is stored on a metadata drive and replicated to a secondary metadata drive for redundancy.

**Accounts**

There are two types of accounts on a SolidFire system:

         Cluster admin accounts. These accounts are used to administer and manage a SolidFire cluster.

         Tenant accounts. These accounts are used to control access to SolidFire volumes from iSCSI initiators. When a unique account user name is created, CHAP initiator and target passwords are defined, and are used to secure volume access based on the account assigned to the volume.

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          A maximum of 2000 volumes can be assigned to a tenant account. A volume can have only one tenant account assigned to it.

As part of the base configuration, you should create an infrastructure “administrator” tenant account to host boot and other infrastructure-related volumes.

**QoS**

A SolidFire cluster allows QoS to be dynamically configured on a per-volume basis. Per-volume QOS settings can be used to control storage performance based on customer-defined SLAs.

There are three configurable parameters that define the QoS:

         Min IOPS. The minimum number of sustained inputs and outputs per second (IOPS) that are provided by the SolidFire cluster to a volume. The minIOPS configured for a volume is the guaranteed level of performance for a volume. Per-volume performance does not drop below this level.

         Max IOPS. The maximum number of sustained IOPS that are provided by the SolidFire cluster to a particular volume.

         Burst IOPS. The maximum number of IOPS allowed in a short burst scenario. If a volume has been running below the maxIOPS, burst credits are accumulated. When performance levels become very high and are pushed to maximum levels, short bursts of IOPS are allowed on the volume, beyond maxIOPS.

**Multi-tenancy**

With NetApp SolidFire, dedicated storage silos are a thing of the past.  SolidFire’s combination of advanced QoS, security, and flexible networking features facilitate mixed-workload and multi-tenant environments on a single SolidFire cluster while guaranteeing individual application performance.

Secure multi-tenancy is achieved through the following features:

         Secure Authentication: CHAP for secure volume access; LDAP for secure access to the cluster for management and reporting.

         Volume Access Groups (VAGs): Optionally, VAGs can be utilized in lieu of authentication, mapping any number iSCSI initiator specific IQNs, to one or more volumes.  In order to access a volume in a VAG, the initiator’s IQN must be in the allowed IQN list for the group of volumes.

         Tenant VLANs: At the network level, end-to-end network security between iSCSI initiators and SolidFire Storage Arrays are facilitated by using VLANs. For any VLAN created to isolate a workload or tenant, SolidFire creates a separate iSCSI target SVIP that is only accessible through the specific VLAN.

         VRF-Enabled VLANs: To further support security and scalability in the data center, SolidFire allows one to enable any tenant VLAN for “VRF-like” functionality.  This feature adds two key capabilities:

         L3 routing to a tenant SVIP – This feature allows you to situate iSCSI initiators on a separate network/VLAN from that of the SolidFire storage array.

         Overlapping/Duplicate IP subnets – Provides the ability to add a template to tenant environments, allowing each respective tenant VLAN to be assigned IP addresses from the same IP subnet.  This can be useful in service provider environments where scale and preservation of IP space is important.

**Efficiency**

The SolidFire storage system leverages key features to increase the overall storage efficiency and performance. These features are performed inline, are always on, and require no manual configuration by the user.

         Deduplication. The system only stores unique 4K blocks. Any duplicate 4K blocks are automatically associated to an already stored version of the data. Data is on block drives and is mirrored by using the SolidFire Helix feature. This system significantly reduces capacity consumption and write operations within the system.

         Compression. Compression is performed inline before data is written to NVRAM. Data is compressed and stored in 4K blocks and, once compressed, it remains compressed in the system. This significantly reduces capacity consumption, write operations, and bandwidth consumption across the cluster.

         Thin provisioning. This capability provides the right amount of storage at the time it is needed eliminating capacity consumption caused by over provisioning volumes or underutilized volumes.

         SolidFire Helix®. The metadata for an individual volume is stored on a metadata drive and replicated to a secondary metadata drive for redundancy

**VVols and VASA Provider**

The NetApp SolidFire’s implementation of VMware Virtual Volumes (VVols) version 1.0 allows SolidFire storage to seamlessly integrate with VMware vSphere.  VVols enable storage options at the virtual machine level, allowing for more granular control. VVols are a supported feature on the FlexPod SolidFire architecture, however they were not validated in this reference architecture.

VVols functionality requires vSphere APIs for Storage Awareness(VASA) provider support. The VASA provider runs natively on the SolidFire cluster.  After the VASA provider is enabled, you can then create and apply policies at the virtual machine or VMDK level.  For example, the policy can contain the minimum, maximum, and burst QoS parameters.

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          VVOL is supported in the FlexPod SF solution. Please consult Cisco HCL and NetApp IMT for additional details.

**vCenter Plug-in (VCP)**

NetApp SolidFire’s vCenter Plug-in allows you to harness the capabilities of the SolidFire storage platform from within the vCenter environment:

         Discover and manage multiple SolidFire clusters

         vSphere datastore create, extend, clone, share, and delete

         SolidFire account create, edit, and delete

         SolidFire volume create, edit, clone, delete, and access group add and remove

         SolidFire access group create, edit, delete

**SolidFire Management Node (mNode)**

An mNode is a virtual machine running a specialized version of SolidFire’s ElementOS that is used to manage SolidFire cluster upgrades, as well as to facilitate proactive system monitoring, and support through NetApp’s SolidFire Active Support program.

**SolidFire Active Support**

SolidFire Active Support model proactively monitors your SolidFire systems to ensure the highest possible level of availability and performance.

Secure assist provides support personnel secure remote access to your SolidFire Cluster when remote assistance or troubleshooting is needed.

The SolidFire Active IQ SaaS platform, a key element of Active Support, provides real-time health diagnostics, historical performance, and trending on a granular level. This holistic approach to infrastructure monitoring, combined with SolidFire’s unique abilities to upgrade, tune, and scale on-demand without disruption, redefines operational success for storage infrastructure.

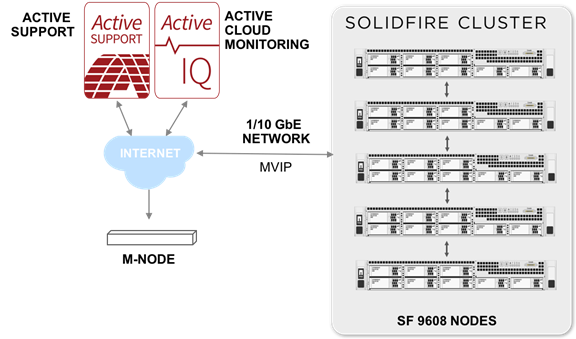
Active IQ provides real-time system health diagnostics and historical trending at the cluster, node, account and volume level, including:

         **Proactive Monitoring** and Alerting – View information on all your deployed SolidFire clusters through a single pane of glass. Real-time alerts can be customized on a variety of metrics: upgrade notification, capacity and performance thresholds, system faults, drive wear and health.

         **Granular System Data**– Telemetry data is sent to Active IQ as frequently as every 10 seconds. Active IQ aggregates a wide variety of information on all manner of system data: active volumes, deleted volumes, volume performance, system performance, system provisioned and used capacity, system efficiency, active nodes and drives, upgrade processes, system health, etc. All of this data is quickly and easily accessible in graphical format through Active IQ’s Web UI.

         **Historical Trending** – All Active IQ data is retained for five years, granularly. We do not roll up or condense older data for ease of viewing. Data points are kept at the frequency with which the system is polled, so you can quickly view system data from one hour, one week, one month, and all the way to five years back. Additionally, all that data is exportable, allowing you to run custom, extended analysis to fit your needs.

Figure 10     Active IQ

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_14.png)

**VMware vSphere 6.5**

VMware vSphere 6.5 is the validated hypervisor for the FlexPod SF design. VMware vSphere allows organizations to reduce their power and cooling footprint while confirming that the purchased compute capacity is used to its fullest. In addition, VMware vSphere allows hardware failure protection (VMware High Availability, or VMware HA) and compute resource load balancing across a cluster of vSphere hosts (VMware Distributed Resource Scheduler, or VMware DRS).

VMware vSphere 6.5 features the latest VMware innovations. The VMware vCenter Server Appliance (VCSA) that is used in this design adds a host of new features and functionality, such as VMware vSphere Update Manager integration. The VCSA also provides native vCenter High Availability for the first time. To add clustering capability to hosts and to use features such as VMware HA and VMware DRS, VMware vCenter Server is required.

VMware vSphere 6.5 also has several enhanced core features. VMware HA introduces an orchestrated restart for the first time, so virtual machines restart in the proper order in case of an HA event. In addition, the DRS algorithm has now been enhanced, and more configuration options have been introduced for more granular control of compute resources inside vSphere.

The vSphere Web Client is now the management tool of choice for VMware vSphere environments, because vSphere 6.5 marks the first release in which the C# client is no longer available. Several user enhancements have also been made to the vSphere Web Client, such as reorganization of the home screen and the inventory tree’s now being the default view upon login.

For more information about the new features of VMware vSphere 6.5, see [What’s New in VMware vSphere 6.5](https://www.vmware.com/content/dam/digitalmarketing/vmware/en/pdf/whitepaper/vsphere/vmw-white-paper-vsphr-whats-new-6-5.pdf).

**Solution Design**

**FlexPod SF Design Overview**

FlexPod SF Datacenter is a flexible architecture that facilitates evolution to the Next-Generation Datacenter. This solution utilizes an end-to-end 10GbE iSCSI deployment using Cisco Nexus 93180YC-EX Cloud Scale switches, Cisco UCS 6248 Fabric Interconnect and NetApp SolidFire SF9608 nodes for storage. This design can also support an 40GbE iSCSI architecture using the Cisco UCS 6300 series Fabric Interconnects. For more information about FlexPod Designs with Cisco UCS 6300 Fabric Interconnects, see: [FlexPod Datacenter with Cisco UCS 6300 Fabric Interconnect and VMware vSphere 6.0 U1 Design Guide](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpod_esxi60_n9k_design.html).

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          The Cisco Hardware Compatibility List (HCL) and NetApp Interoperability Matrix Tool (IMT) defines the qualified components and versions you can use to build a FlexPod SF.  At the time of release, the IMT and HCL testing of the Cisco UCS 6300 Fabric Interconnect is pending.  Please consult IMT and HCL for updated support information regarding all hardware and software elements.

From a compute perspective, the Cisco UCS 5108 blade server chassis with blade servers are used in this design but for ROBO or smaller deployments, a Cisco UCS Mini based with an integrated 6342FI can also be used. For more information about FlexPod Designs with Cisco UCS Mini, see: [FlexPod Datacenter with Cisco UCS Mini and VMware vSphere 6.0 with IP-Based Storage](https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpod_ucsmini_esxi60_ip.html).

This design uses a new generation of Cisco Nexus 9000 series data center switches in NX-OS standalone mode but provide investment protection to migrate to ACI or higher network bandwidths (1/10/25/40/50/100Gbps) while enabling innovative analytics and visibility using Tetration and automation that support in-box and off-box Python scripting and Open NX-OS that support dev-ops tools (Chef, Puppet, Ansible).

**FlexPod SF Design Options**

**Design 1: End-to-End 10GbE iSCSI Architecture**

This design is an end-to-end 10GbE IP based data center architecture with the ability to upgrade compute and network fabrics to 40GbE, maintaining investment protection by leveraging the higher network bandwidths and port densities available on the Cisco Nexus 9000-EX series switches. Cisco Nexus 9000-EX series switches also bring innovative features including software defined networking with ACI, Cloud Scale technology, Tetration Analytics, along with comprehensive programmability and tools for building next-generation automated datacenters. This CVD uses this design.

The components in this design are:

         Cisco Nexus 93180YC-EX switches in NX-OS Standalone mode

         Cisco UCS 6200 Series Fabric Interconnect (10GbE Unified Fabric)

         Cisco UCS 5108 Blade Server Chassis

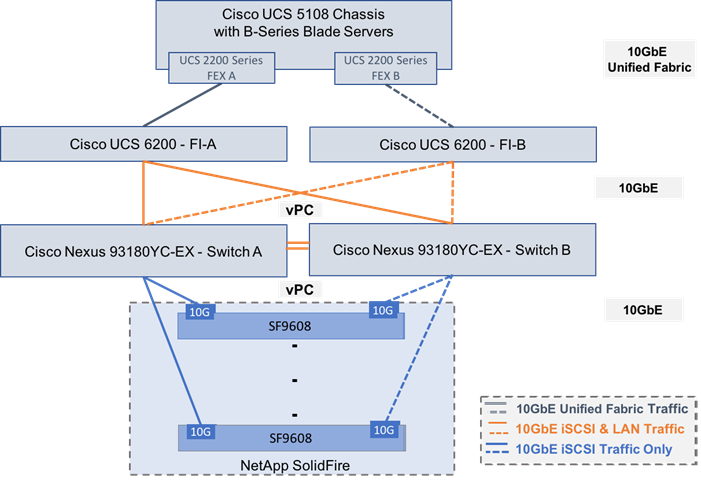
         Cisco UCS 2200 Series IOM/FEX (10GbE Uplink for Blade Server Chassis)

         Cisco UCS B200 M4 servers with VIC 1340

         NetApp SolidFire SF9608 (4-node minimum)

Figure 11 illustrates the FlexPod Datacenter topology for an end-to-end 10GbE iSCSI design. The Cisco UCS 6200 series Fabric Interconnects provide a high-performance, low-latency, and lossless fabric to support applications with these requirements. The 10GbE compute and network fabric serves as a foundation for a uniform and resilient FlexPod solution. Each network segment of the design depicted in Figure 11 represents a combination of network links in a port-channel or virtual port-channel (vPC). Additional links can be added as needed to increase the aggregate bandwidth (excluding the SF9608 nodes). NetApp SolidFire SF9608 nodes facilitate linear scaling of overall cluster network bandwidth at a rate of 20Gbps per node (2 x 10GbE).

Figure 11     End-to-End 10GbE iSCSI Architecture

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_15.png)

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          Dotted lines are used to highlight the redundancy in the design though all links are active.

**Design 2: 40GbE iSCSI Architecture**

This is a 40GbE IP based datacenter architecture with the ability to upgrade to higher network bandwidths and capabilities enabled by the use of Cisco Nexus 9000-EX series switches in the design. These include flexible port densities that support 40/100GE on uplinks and 1/10/25GE on access links. Cisco Nexus 9000-EX series switches also brings innovative features including software defined networking with ACI, Cloud Scale technology, Tetration Analytics, along with comprehensive programmability and tools for building next-generation automated datacenters.

The components in this design are:

         Cisco Nexus 93180YC-EX switches in NX-OS Standalone mode

         Cisco UCS 6300 Series Fabric Interconnect (40GbE Unified Fabric)

         Cisco UCS 5108 Blade Server Chassis

         Cisco UCS 2304 IOM/FE (40GbE Uplink for Blade Server Chassis)

         Cisco UCS B200 M4 servers with VIC 1340

         NetApp SolidFire SF9608 (4-node minimum)

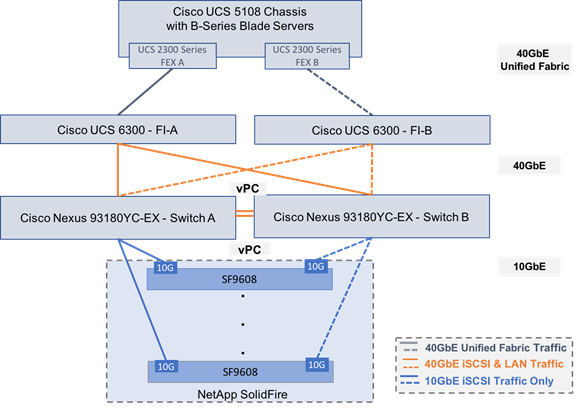
Figure 12 illustrates the FlexPod SF topology for a 40GbE IP based datacenter design. The Cisco UCS 6300 Series Fabric Interconnect enables a high-performance, low-latency, and lossless fabric for supporting applications with these requirements. In NX-OS standalone mode, the 40GE uplink ports on Cisco Nexus 93180YC-EX switches can be used to connect to access Cisco UCS 6300 Fabric Interconnects as shown in the topology below.

The 40GbE compute and network fabric increases the overall capacity of the system while maintaining the uniform and resilient design of the FlexPod solution. This design uses 40GbE links from the UCS compute domain to the Nexus switches and 10GbE from the Nexus switches to the individual SF9608 nodes. However, given that a single SolidFire cluster requires a minimum of 4xSF9608 nodes, each with 2x10GbE storage uplinks per node, the effective bandwidth from Nexus switches to the SolidFire cluster is 80Gbps. Also, given the scale-out architecture of NetApp SolidFire storage, each node added to the cluster results in an additional 20Gbps of bandwidth. NetApp SolidFire SF9608 nodes also provides granular scaling of both capacity and performance (beyond the 4-node minimum for a cluster) by being able to add nodes to the cluster, one at a time.

Link bundling is also used in the different segments of the compute and network fabric by aggregating the 40GbE connections to provide a minimum bandwidth of 80Gbps. Aggregate bandwidths of up to 320Gbps can be supported in this design by using available ports on Cisco UCS Fabric Interconnects, Cisco UCS Fabric Extenders and Cisco Nexus switches.

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          The Cisco Hardware Compatibility List (HCL) and NetApp Interoperability Matrix Tool (IMT) defines the qualified components and versions you can use to build a FlexPod SF. At the time of release, the IMT and HCL testing of the Cisco UCS 6300 Fabric Interconnect is pending. Please consult IMT and HCL for updated support information regarding all hardware and software elements.

Figure 12     40GbE iSCSI Architecture

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_16.png)

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          Dotted lines are used to highlight the redundancy in the design though all links are active.

**Design 3: End-to-End 10GbE iSCSI Design using Cisco UCS Mini**

This design uses a 10GbE IP based architecture for the datacenter with Cisco UCS Mini for compute. Cisco UCS Mini has an integrated Fabric Interconnect that can be used to connect to the data center switches.

The components in this design are:

         Cisco Nexus 93180YC-EX switches in NX-OS Standalone mode

         Cisco UCS Mini Blade Server Chassis

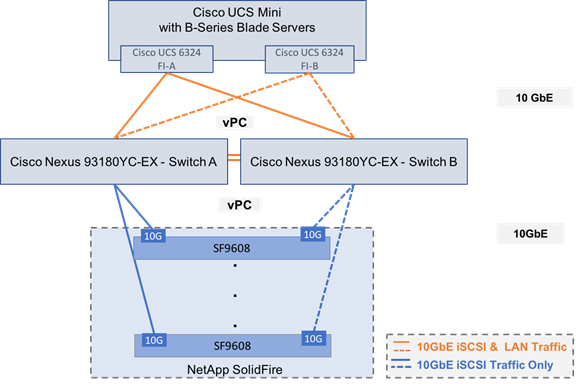
         Integrated Cisco UCS 6324 Fabric Interconnect (10GbE Unified Fabric)

         Cisco UCS B200 M4 servers with VIC 1340

         NetApp SolidFire SF9608 (4-node minimum)

Figure 13 illustrates the FlexPod SF topology for an end-to-end 10GbE iSCSI design using Cisco UCS Mini with integrated Cisco UCS 6324 Fabric Interconnects. To increase the aggregate bandwidth in the unified fabric layer and in the network layer, additional links can be bundled together. NetApp SolidFire SF9608 nodes facilitate scaling both capacity and performance linearly, one node at a time (beyond 4-node cluster).

Figure 13     End-to-End 10GbE iSCSI Architecture using Cisco UCS Mini

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_17.png)

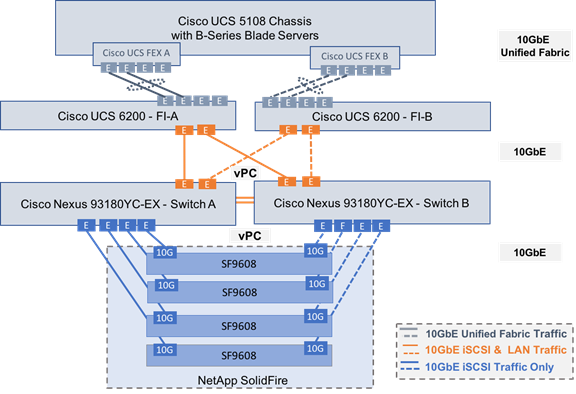
[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          Dotted lines are used to highlight the redundancy in the design though all links are active.

**FlexPod SF Design**

The FlexPod SF Solution delivers a converged infrastructure solution using scale out storage provided by NetApp SolidFire. This solution incorporates compute, network and storage layer best practices to provide a resilient, scalable data center architecture for enterprise and cloud deployments.

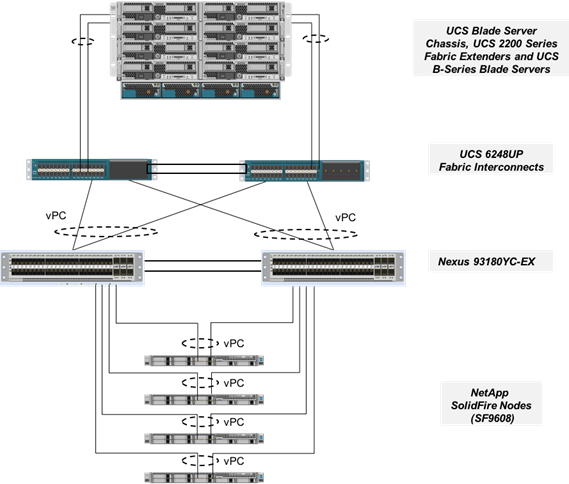
Figure 14 and Figure 15 illustrate the end-to-end FlexPod SF design in this CVD.

Figure 14     FlexPod SF Design – Logical Topology

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_18.png)

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          Dotted lines are used to highlight the redundancy in the design though all links are active.

Figure 15     FlexPod SF Design – Physical Topology

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_19.png)

**Compute Layer Design**

The compute resources supported in this design are Cisco UCS B-Series Servers. Cisco UCS B-Series are housed in a Cisco UCS 5108 Blade Server Chassis that can support up to 8 half-width or 4 full-width blades. Each Cisco UCS server is equipped with a Virtual Interface Cards (VIC) that aggregate all LAN and SAN traffic to and from the server across a single interface. Cisco VICs eliminate the need for separate physical interface cards on each server for LAN, SAN and management connectivity. Cisco VICs can be virtualized to create up to 256 virtual interfaces that can be dynamically configured as virtual network interface cards (vNICs) or virtual host bus adapters (vHBAs). These virtual interfaces will be presented and appear as standards-compliant PCIe endpoints to the OS. Cisco VIC models are available that can enable up to 80Gbps of uplink bandwidth per server.

Multiple models of Cisco VICs are available. VIC is available as a Modular LAN on Motherboard (mLOM) or Mezzanine Slot card for the blade servers. For more information on the different models of adapters available on Cisco UCS servers, see:  [http://www.cisco.com/c/en/us/products/interfaces-modules/unified-computing-system-adapters/index.html](https://www.cisco.com/c/en/us/products/interfaces-modules/unified-computing-system-adapters/index.html)

**Cisco UCS Server Connectivity to Unified Fabric**

Cisco UCS servers are typically deployed with a single VIC card for unified network and storage access. The Cisco VIC connects into a redundant unified fabric provided by a pair of Cisco UCS Fabric Interconnects. Fabric Interconnects are an integral part of the Cisco Unified Computing System, providing unified management and connectivity to all attached chassis and blade servers. Fabric Interconnects provide a lossless and deterministic Fibre Channel over Ethernet (FCoE) fabric. For the servers connected to it, the Fabric Interconnects provide LAN, SAN and management connectivity to the rest of the network.

**Blade Server Connectivity to Unified Fabric**

On a blade server chassis, Cisco UCS B-series servers connect into the fabric interconnect through Fabric Extenders (FEX) or Input Output Modules (IOM). Fabric Extenders extend the unified fabric from the FI to the chassis and serves as a consolidation point for all blade server I/O traffic. FEX is managed as an extension of the fabric interconnects, simplifying diagnostics, cabling and operations with a single point of management and policy enforcement. This approach reduces the overall infrastructure complexity and enables the Cisco UCS to scale with multiple blade servers managed as a single, highly available management domain. For information about additional benefits of FEX, see: [http://www.cisco.com/c/en/us/solutions/data-center-virtualization/fabric-extender-technology-fex-technology/index.html](https://www.cisco.com/c/en/us/solutions/data-center-virtualization/fabric-extender-technology-fex-technology/index.html)

A Cisco UCS 5100 Series Blade Server Chassis can support up to two fabric extenders that fit into the back of the chassis. Fabric Extenders operate in an active-active mode for data forwarding and active-passive mode for management functions. For the Cisco UCS blade server chassis, three FEX models are currently available. For each FEX model, the number of external 10GbE/40GbE links to fabric interconnects and internal links to blade servers within the chassis are shown in the table below. The maximum uplink I/O bandwidth for each FEX is also provided.

For additional information about the blade server FEX models, see: [http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6300-series-fabric-interconnects/datasheet-listing.html](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6300-series-fabric-interconnects/datasheet-listing.html)

Table 2    Blade Server Fabric Extender Models

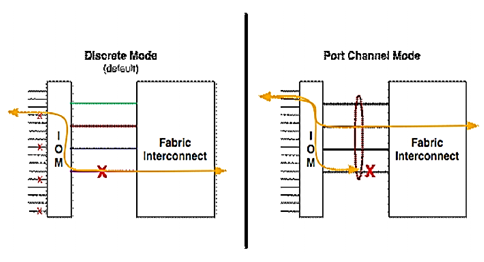
|  |  |  |  |
| --- | --- | --- | --- |
| **Blade Server Models** | **Internal Facing Links to Servers** | **External Facing Links to FI – Uplink Ports** | **Max Uplink I/O Bandwidth per FEX** |
| Cisco UCS 2204XP | 16 x 10GbE | Up to 4 x 10GbE | 40Gbps |
| Cisco UCS 2208XP | 32 x 10GbE | Up to 8 x 10GbE | 80Gbps |
| Cisco UCS 2304 | 8 x 40GbE | Up to 4 x 40GbE | 160Gbps |

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          This design supports both Cisco UCS 2200 and 2300 Series Fabric Extenders.

**Fabric Extender Connectivity to Unified Fabric**

The fabric extenders can connect to the unified fabric using multiple Ethernet (10GbE/40GbE) links – the number of links determines the uplink I/O bandwidth available through each FEX. Each FEX connect into a fabric interconnect using either 1, 2, 4 or 8 Ethernet links depending on the model of FEX used. 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 16     Fabric Extender to Fabric Interconnect Connectivity Options

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_20.png)

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          This design recommends port-channel mode connectivity between FEX and FI.

**Cisco UCS Fabric Interconnect Connectivity to Core**

Fabric Interconnects enable LAN, SAN and management connectivity for all servers in the UCS domain that connect to it. Two models of Fabric Interconnects are current available. Cisco UCS 6200 series fabric interconnects provide a unified 10GbE FCoE fabric with 10GbE uplinks for northbound connectivity to customer’s LAN/IP network. Cisco UCS 6300 series is the newest fabric interconnect model and provides a 40GbE unified FCoE fabric with 40GbE ports for uplink connectivity. The uplinks can be bundled together for higher aggregate uplink bandwidth.

This design supports Cisco UCS B-series connecting to either 6200 or 6300 series fabric interconnects.

**Validated Compute Design**

For validation, Cisco UCS B200 M4 blade servers with VIC 1240 and VIC 1340 adapters respectively, were connected to 2 x Cisco UCS 6248 Fabric Interconnects. The Cisco UCS 5108 blade server chassis, housing the blade servers, were deployed using 2 x Cisco UCS 2204 XP FEX adapters to connect to the fabric interconnects. Two 10GbE links were used for FEX to FI connectivity, one from FEX-A to FI-A and one from FEX-B to FI-B, for an aggregate access bandwidth of 20Gbps from the blade server chassis to the unified fabric.

Connectivity from each individual fabric interconnect, to all upstream or northbound (NB) networks is provided by 2 x 10G links to each of the top-of-rack Cisco Nexus switches as follows:

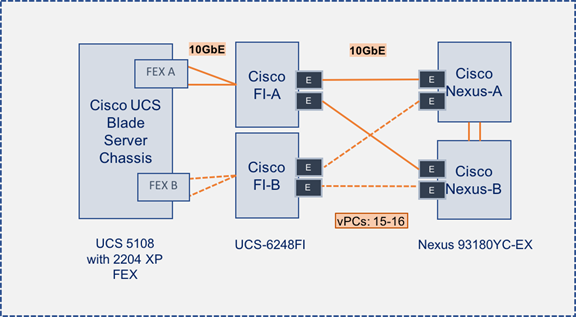
         2 x 10G uplinks from FI-A to Nexus-A and Nexus-B respectively

         2 x 10G uplinks from FI-B to Nexus-A and Nexus-B respectively

Both uplinks are configured into a single port channel, making the total aggregate bandwidth to the core switching infrastructure 20Gbps per fabric interconnect. Each port designated as a core switch connection is designated as an uplink port within Cisco UCS Manager.

Figure 17 illustrates the connectivity used in validation.

Figure 17     Connectivity - Compute Layer

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_21.png)

**Virtualization Layer Design**

The compute resources in this architecture are grouped into either a management infrastructure cluster or an application infrastructure cluster. Servers in the management infrastructure cluster would typically for management (vCenter) and other services (AD, DNS) that are necessary for deploying and managing the entire data center. Servers in the application cluster would host the virtual machines that run business applications that enterprise and cloud users will access to fulfill their business function. High availability features available in VMware vSphere are leveraged to provide virtualization layer resiliency.

**Cisco UCS Server – Virtual Networking Design**

Cisco VICs provide 256 virtual adapters that can be presented as storage (vHBAs) or networking (vNICs) adapters to operating system. The virtual adapters are presented as vPCIe devices to ESXi and are referred to as vmnics within VMware vSphere.

In this design, the following virtual adapters are used with –A vNICs connected to unified fabric A and –B vNICs to unified fabric B resulting in each ESXi node being dual homed to the external network. The virtual adapters are assigned to different virtual switches depending on the type of traffic. VMware vSwitch is used for management, vCenter HA and vMotion traffic. VMware vDS is used for application traffic with different application port groups to isolate the applications.

         Two vNICs (MGMT-A, MGMT-B) for in-band management

         Two vNICs (vMotion-A, vMotion-B) for vMotion traffic

         Two vNIC (iSCSI-A, iSCSI-B) for iSCSI storage traffic

         Two vNICs (NIC-A, NIC-B) for application VM traffic

The design also used two VMkernel NICs, each with its own port group for:

         vmk0 - ESXi management

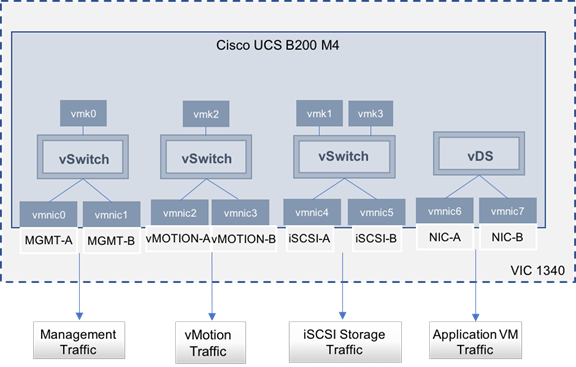
         vmk2 - vMotion

         vmk1,3 – iSCSI

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 to support VM migration between hosts. The iSCSI interfaces will be used to establish iSCSI connections to NetApp SolidFire storage.

Figure 18 illustrates the virtual networking design and connectivity within a given ESXi server.

Figure 18     Cisco UCS Server-Virtual Networking

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_22.png)

**Validated Virtualization Design**

For validation, Cisco UCS servers in the design were deployed to a VMware High Availability (HA) cluster in either management infrastructure cluster or in the application infrastructure to mitigate against host failures. VMware HA enables automatic failover of active VMs to other hosts in the cluster in the event of a host failure. Admission control is used to ensure that application VMs has the resources it needs before a failover is allowed.

The FlexPod SF design support different hypervisors and workloads but VMware vSphere 6.5 was used in validation.

**LAN Network Design**

The overall network infrastructure design focuses on providing enhanced redundancy and HA as well as adequate bandwidth to support bandwidth-intensive applications and storage traffic. The network design uses fully meshed connectivity with virtual port channels (vPC) to facilitate better utilization of available switch bandwidth, distributing traffic across all switches in the path, while still supplying adequate bandwidth to sustain operations in the event of a network failure.

The FlexPod SF design is highly resilient with redundant Unified fabric, LAN fabric and storage network that includes both component and link level redundancy. The unified fabric, LAN connections and storage network connections are all 10Gb Ethernet links. Multiple 10GbE links with Link Aggregation Control Protocol (LACP) based aggregation is leveraged to provide higher aggregate bandwidth and resiliency across the different fabrics. Use of LACP is strongly recommended over other methods when available to improve failover convergence time and protection from misconfigurations.

The FlexPod SF design uses a pair of Cisco Nexus 93180YC-EX switches that serve as the access/aggregation layer of the data center network. The Nexus 9000 series switches 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. In the FlexPod SF design, Nexus 9000 series switches are deployed in NX-OS standalone mode but provide investment protection and pathway to Cisco’s Application Centric Infrastructure and 40GbE/100GbE connectivity when needed.

Link aggregation using vPCs are used on the network links from Nexus switches to Cisco UCS Fabric Interconnects and to the NetApp SolidFire nodes. 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

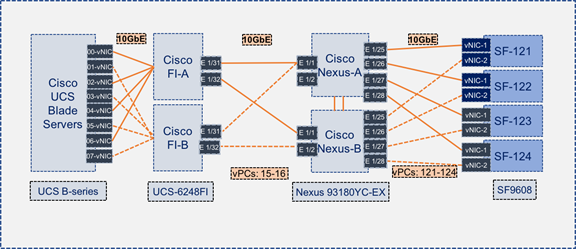
         Helps ensure high availability

Redundant 10Gbps links from each FI to each Nexus provide 20Gbps of bandwidth through each Cisco Nexus. VLAN trunking is enabled on these links as multiple application data VLANs, management, vMotion and iSCSI storage traffic traverse these links. Jumbo Frames are also enabled in the LAN network to support vMotion between multiple Cisco UCS domains and iSCSI storage traffic.

Redundant 10Gbps links are also used from each SolidFire node to each Nexus switch, providing 20Gbps of aggregate uplink bandwidth from each node.

See Design Practices section for other Cisco Nexus 9000 best practices in the design. Figure 19 illustrates the end-to-end topology used in validating this design.

Figure 19     End-to-End Network Connectivity for Validation

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_23.png)

**Storage Design**

**NetApp SolidFire**

Traditional storage area networks (SANs) were designed to use FC to guarantee high-bandwidth, low-latency connectivity between servers and storage devices. Redundant switches were implemented to provide HA for storage connectivity and to make sure there is no single point of failure for storage access. Separate network infrastructures permit performance and availability, but at the cost of additional hardware and management overhead.

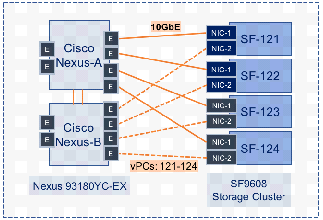
Ethernet networking has evolved and can now deliver much higher throughput with lower latency than ever before. Lossless Ethernet protocols, advanced bandwidth management and HA features like  virtual port channels (vPCs) deliver the reliability and resiliency necessary for mission-critical workloads. With these new capabilities, IT organizations are moving away from dedicated FC SAN infrastructures to Ethernet-based storage infrastructures, and protocols such as iSCSI.

SolidFire Element OS was specifically designed to leverage today’s high-bandwidth Ethernet network architectures and the iSCSI Ethernet storage protocol. SolidFire uses a 10GB iSCSI Ethernet network to connect multiple storage nodes and create a clustered storage array. HA within the overall storage solution is facilitated by a combination of physical connectivity redundancy, and SolidFire’s self-healing, Double Helix technology.

NetApp SolidFire SF9608 nodes are connected to the Cisco Nexus 93180YC-EX switches using vPCs, bonding two physical 10G ports on each SolidFire node into a single logical interface (Bond10G). These port channels carry all IP-based ingress and egress storage data traffic and cluster traffic for the NetApp SolidFire nodes.

Figure 20 illustrates the connectivity used in validation.

Figure 20     Connectivity to Storage Layer

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_24.png)

By default, separate onboard 1G ports are utilized for in-band management traffic and API automation, however in-band management can be moved to the Bond10G interface if converged networking is desired. Below are the connections or network types used in this solution.

         In-Band Management network. A network used for the administration of nodes and the API endpoint

         Out-of-Band Management network.  Network used to administer and monitor the SolidFire node hardware

         Storage network. A network used by clients to access storage via the iSCSI protocol

         Ports. Physical 10GbE Port 1 and Port 2 (identified in SolidFire Element OS as Eth0 and Eth1) are used for storage network. Physical 1GbE Port 1 and Port 2 (identified in SolidFire Element OS as Eth3 and Eth4) are used for in-band management network.

**Storage Infrastructure Connectivity**

The storage infrastructure design makes sure that there is no single point of failure in the storage path. SolidFire storage leverages the iSCSI storage protocol to deliver all storage traffic traverses the dedicated 10G switches for fast, reliable delivery of iSCSI storage traffic.

**SolidFire Storage Node Connectivity**

SolidFire nodes leverage dual-homed NIC connectivity for both storage and in-band management to provide redundancy, enhanced fault tolerance, and increased bandwidth for storage traffic:

         Storage (vPC)

         10G link to storage switch N9K-A

         10G link to storage switch N9K-B

         In-band management (Active/Passive Bond):

         1 x 1G link to management switch N9K-A

         1 x 1G link to management switch N9K-B

**iSCSI Redirect and Storage Virtual IP**

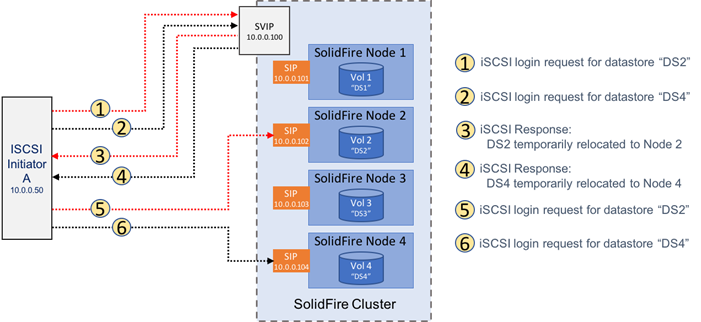
The highly-available storage virtual IP (SVIP) is the single point of access for all iSCSI traffic.  The iSCSI redirect process is a key component of SolidFire’s highly available, scale out architecture.  By leveraging iSCSI redirect, SolidFire can ensure that no single node is a point of failure or data loss.

The initiator sends the iSCSI login request to the volume via the cluster SVIP, which is logically hosted by the cluster master node.

SolidFire cluster master responds with an iSCSI redirect back to the initiator, directing it to the actual node that hosts the volume, and the initiator redirects its login request towards the designated node’s per-node storage IP (SIP), thus establishing a direct iSCSI connection to the individual node.

Figure 21 illustrates how iSCSI redirect process works with a SolidFire cluster.

Figure 21     iSCSI Redirect Process with SolidFire Cluster

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_25.png)

By leveraging iSCSI redirect, SolidFire can maintain its linear scale-out ability while maintaining high availability. As additional storage nodes are added, network bandwidth is added in addition to storage capacity, allowing client access to scale along with capacity.

If a cluster node suffers a failure, iSCSI redirect is used to redirect the initiator to the specific node that hosts the initiator’s secondary copy of its volume, while transparently recreating yet another copy in the background, quickly restoring the volume to a fully redundant state.

In a typical controller-based storage system additional drive shelves can be added, but the controllers and their client network bandwidth will eventually become a bottleneck.

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          Routing between nodes within a cluster is not supported.

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          NetApp strongly recommends that you keep all nodes on the same switch to minimize latency.

**iSCSI SAN Boot of Cisco UCS Servers**

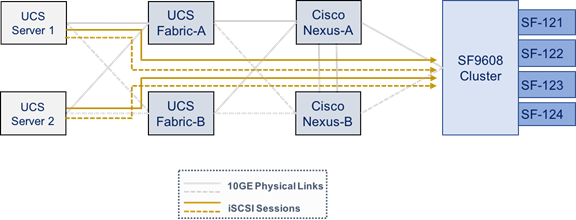
FlexPod SF design recommends SAN boot of Cisco UCS servers to fully achieve stateless computing. In this design, the servers are booted using iSCSI from boot volumes hosted on SF9608 nodes.

For iSCSI SAN boot, the service profiles used to configure and deploy the Cisco UCS B-Series servers are configured with a boot policy that direct it to boot using iSCSI. The boot policy is configured to use two iSCSI paths for booting by assigning two iSCSI vNICs, one through Fabric Interconnect A and a second through Fabric Interconnect B in order to provide redundancy across the unified fabric. The traffic from each fabric can traverse either Cisco Nexus switches through the vPC links between each FI and Cisco Nexus switches.

From the Cisco Nexus switches, the traffic can take one of two paths to reach the storage virtual IP address of the SolidFire cluster. By mapping each iSCSI vNIC to a fabric and through the vPC configuration between FI and Cisco Nexus switches and SolidFire nodes, each Cisco UCS server has two redundant paths all the way to the storage for booting. The iSCSI VLAN used to carry the boot traffic is enabled end to end across this path with an MTU of 9000 to optimize throughput and minimize CPU usage.

Figure 22 illustrates a high-level view of the connectivity for iSCSI access to storage.

Figure 22     iSCSI Connectivity - High Level

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_26.png)

On the SolidFire cluster, boot volumes are created for each Cisco UCS server and configured to allow the iSCSI Qualifier Name (IQN) of the host to access the boot volumes. Each boot volume is configured to be in a separate Volume Access Group and associated with an Infrastructure Administrator account. All boot volumes are associated with the same account but each boot volume is part of a single VAG.

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          In the current version of ElementOS, the maximum number of VAGs supported is 1000 per cluster.

The following figures illustrate the iSCSI SAN boot design.

Figure 23 highlights the design for two hosts.

Figure 23     iSCSI SAN Boot Design – For Two Cisco UCS Hosts

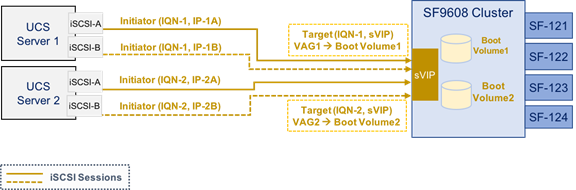
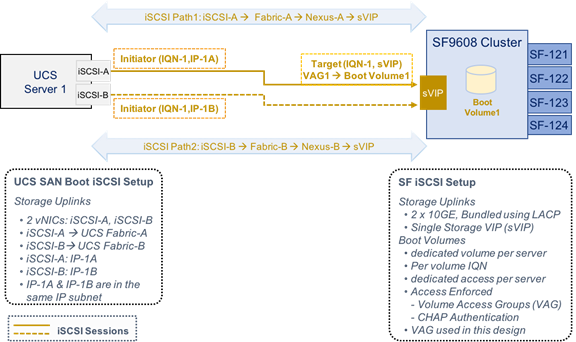
[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_27.png)

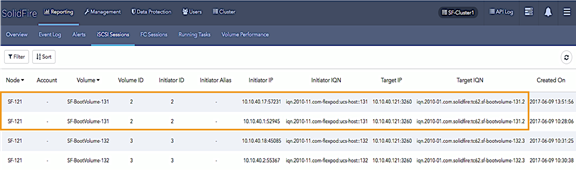
Figure 24 provides a detailed design for a single host.

Figure 24     iSCSI SAN Boot Design – Detailed

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_28.png)

The iSCSI sessions between the Cisco UCS Hosts and NetApp SolidFire are shown in Figure 25.

Figure 25     iSCSI Sessions for Boot Volumes

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_29.png)

**High Availability**

The FlexPod SF solution 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 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.

         Each fabric extender on the Cisco UCS 5108 blade server is deployed with redundant 10GbE 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 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 iSCSI multipathing to distribute load and provide redundancy when accessing volumes on the NetApp SolidFire cluster. If any component (NIC, FEX, FI, SolidFire node, 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 FlexPod SF 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.

**Storage**

         SolidFire’s Helix - The primary copy of metadata for an individual volume is replicated to a secondary metadata location on a different node for redundancy.

         iSCSI redirect is used to maintain connectivity throughout node failures, additions and upgrades.  As well as volume location changes due to automated volume rebalancing of storage resources in order to optimize performance and efficiencies.

         Redundant Network Connectivity is facilitated via redundant interfaces for both management and storage.

**Scalability**

For higher performance and capacity, FlexPod SF 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 FlexPod SF solution.

**Management**

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

**Compute**

         FlexPod SF can support multiple Cisco UCS domains with each domain supporting up to 20 blade server chassis (up to 8 blade server per chassis) per domain. Each Cisco UCS domain consists of a pair of Cisco 6200 or 6300 Series Fabric Interconnects that can provide a unified fabric and management with up to 2.56 Tbps of bandwidth. High-performance ports capable of line-rate, low-latency, lossless 1-, 10-, and 40-Gigabit Ethernet (6300 Series only) and Fibre Channel over Ethernet (FCoE), and 16-, 8-, and 4-Gbps (6300 Series) and 8-, 4-, and 2-Gbps (6200 Series) Fibre Channel.

         Each Cisco UCS domain can support multiple compute form factors (Blade Servers, Cisco UCS Mini) to meet the needs of the deployment and the applications deployed.

**Storage**

The SolidFire storage solution, based on the SF9608 is an all-flash, scale-out storage architecture for next generation data centers.

Leveraging the following key features, the solution can be scaled on demand as needed:

         **Linear Capacity and Performance scalability** - As nodes are added to a cluster, each delivers an exact amount of additional storage capacity, network bandwidth, and storage performance (IOPS) for a predictable pattern of linear scalability. Additional storage resources are readily available for use within minutes of being added to the cluster.

         **Non-disruptive scale-out / scale-in** - Add or remove nodes to a SolidFire cluster non-disruptively without compromising volume-level Quality of Service (QoS) settings. Data is automatically redistributed in the background across all nodes in the cluster, maintaining a balanced distribution of storage resources as the cluster scales.

         **Simplify capacity planning** - Initial implementations begin with a minimum of 4 node/4U cluster configurations and scale-out easily through 1U node increments, allowing performance and capacity resources to be added as needs dictate. Eliminate multi-year capacity and performance projections and scale on demand.

         **Seamless generational upgrades –**As new node platforms with more capacity and performance become available, they can simply be added to the existing cluster without disruption. Older nodes may be removed, retired, or repurposed dynamically, without disruption as well. Forklift upgrades are a thing of the past.

Table 3    SolidFire SF9608 Scalability

|  |  |  |
| --- | --- | --- |
| **SF9608** | **Single Node** | **4-node Cluster** |
| Raw Capacity | 7.7 TB | 30.8 TB |
| Effective Capacity | 15TB (5X efficiencies) to  30TB (10X efficiencies) | 60TB (5X efficiencies) to  125TB (10X efficiencies) |
| Predictable IOPS | 75,000 | 300,000 |

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          In the current version of ElementOS, the maximum number of SF9608 nodes in a single cluster is 40.

**Other Design Considerations**

The following sections outline other design considerations for the FlexPod SF solution.

**Management Connectivity**

The FlexPod SF 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, SolidFire SF9608 nodes, Cisco Nexus 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 and this is done in-band in this design. However, if out-of-band management to these components are required, the disjoint layer 2 feature on Cisco UCS Manager 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.

In-band connectivity to NetApp SolidFire nodes are also supported and would require the management vlan to be trunked on the storage links between the SF nodes and uplink Nexus switches. The type of management connectivity (in-band or out-of-band) must be specified during the creation of the SolidFire cluster.

**QoS and Jumbo Frames**

Cisco UCS, Cisco Nexus, and NetApp SolidFire storage nodes in this solution provide QoS policies and features for handling congestion and traffic spikes. The network-based QoS capabilities in these components can alleviate and provide the priority that the different traffic types require.

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          When enabling QoS on Cisco UCS or Nexus switches, it should be implemented end-to-end to avoid issues.

The FlexPod SF design also requires enabling jumbo frames support for a minimum of 9000-byte MTU across the entire storage datapath, including all LAN and Unified Fabric links. Additionally, jumbo frames should be enabled on all iSCSI and vMotion traffic paths, as detailed in this design. 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 Nexus 9000 Series – vPC Best Practices**

FlexPod SF design recommends the following vPC best practices.

**vPC Peer Keepalive Link**

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 design.

**vPC Peer Link**

         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.

**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 design.

**vPC Member Link**

         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**

         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, followed by 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 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 links provided below to avoid issues in your deployment:

         [vPC Design Guide](https://www.cisco.com/c/dam/en/us/td/docs/switches/datacenter/sw/design/vpc_design/vpc_best_practices_design_guide.pdf)

         [Cisco Nexus 9000 NX-OS Release 6.x Configuration Guide Validation](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/nexus9000/sw/7-x/interfaces/configuration/guide/b_Cisco_Nexus_9000_Series_NX-OS_Interfaces_Configuration_Guide_7x.html)

**Solution Validation**

A high-level summary of the validation done for this FlexPod SF solution is provided below. The solution was validated for data forwarding and storage access by deploying virtual machines running Vdbench tool. The system was validated for resiliency by failing various aspects of the system under load. These include the following:

         Failure and recovery of iSCSI booted ESXi hosts in a cluster

         Service Profile migration to a new server

         Failure and recovery of redundant links between IOM/FEX links and Cisco UCS FI

         Failure and recovery of redundant links between Cisco UCS FI and Cisco Nexus

         Failure and recovery of redundant links between Cisco Nexus and NetApp SolidFire nodes

         Failure and recovery of SSDs on NetApp SolidFire nodes

         Failure and recovery of nodes in the NetApp SolidFire cluster

Load was generated using Vdbench tool with IO profiles reflecting the different profiles seen in customer networks. See Table 4 for the examples of these profiles.

Table 4        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**

It is important to note that Cisco, NetApp, and VMware have interoperability matrixes that should be referenced to determine support for any specific implementation of FlexPod SF. Click the following links for more information:

         [NetApp Interoperability Matrix Tool](http://support.netapp.com/matrix/)

         [Cisco UCS Hardware and Software Interoperability Tool](https://www.cisco.com/web/techdoc/ucs/interoperability/matrix/matrix.html)

         [VMware Compatibility Guide](http://www.vmware.com/resources/compatibility/search.php)

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          FlexPod SF Solution only supports Cisco UCS B-Series for compute.

**Hardware**

Table 5 lists the hardware components used in validating the FlexPod SF with VMware vSphere 6.5 infrastructure solution.

Table 5    Validated Hardware

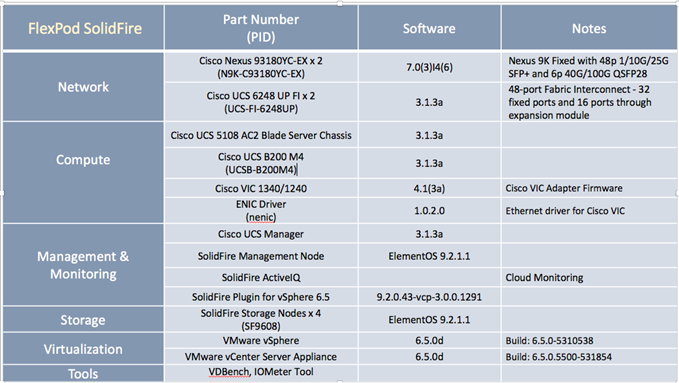
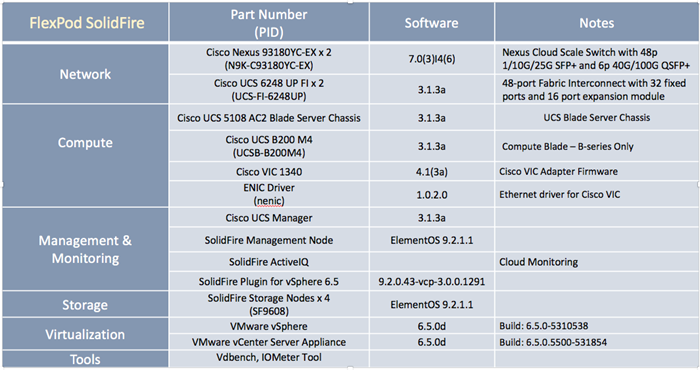
| **Hardware** | **Quantity** |
| --- | --- |
| Cisco UCS B200 M4 blades | 2 |
| Cisco UCS blade server chassis | 1 |
| Cisco VIC 1340 (1240 on B200M3) | 1 per blade |
| Cisco Nexus 93180YC-EX Ethernet switches | 2 |
| Cisco UCS 6248 Fabric Interconnect (FI) switches | 2 |
| Cisco UCS 2204XP FEX | 2 |
| SolidFire SF9608 Storage Appliance | 4 |

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          The components listed in the hardware table were validated for this reference architecture.  Customers can scale their environments based on their needs as long as they meet these minimum requirements.

**Software**

Table 6 lists the software versions used for validating the FlexPod SF infrastructure solution.

Table 6    Validated Software

[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_30.png)[](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_31.png)

[*](https://www.cisco.com/c/dam/en/us/td/docs/unified_computing/ucs/UCS_CVDs/flexpodsf_esxi65design.docx/_jcr_content/renditions/flexpodsf_esxi65design_0.png)          The components listed in Table 6 were validated for this reference architecture. Customers can use different versions of software as long as they are supported on the Cisco HCL and the NetApp IMT.

**Summary**

FlexPod SF with VMware vSphere 6.5 is the optimal shared infrastructure foundation to deploy a variety of IT workloads that is future proofed with Cisco Nexus 9000 series switches. Cisco and NetApp have created a platform that is both flexible and scalable for multiple use cases and applications. From virtual desktop infrastructure to DevOps, FlexPod SF can efficiently and effectively support business-critical applications running simultaneously from the same shared infrastructure. The flexibility and scalability of FlexPod also enable customers to start out with a right-sized infrastructure that can ultimately grow with and adapt to their evolving business requirements.

**References**

**Products and Solutions**

Cisco Unified Computing System:

[http://www.cisco.com/en/US/products/ps10265/index.html](https://www.cisco.com/en/US/products/ps10265/index.html%20)

Cisco UCS 6200 Series Fabric Interconnects:

[http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6200-series-fabric-interconnects/index.html](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6200-series-fabric-interconnects/index.html)

Cisco UCS 6300 Series Fabric Interconnects:

[http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6300-series-fabric-interconnects/index.html](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-6300-series-fabric-interconnects/index.html)

Cisco UCS 5100 Series Blade Server Chassis:

[http://www.cisco.com/en/US/products/ps10279/index.html](https://www.cisco.com/en/US/products/ps10279/index.html%20)

Cisco UCS B-Series Blade Servers:

[http://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-b-series-blade-servers/index.html](https://www.cisco.com/c/en/us/products/servers-unified-computing/ucs-b-series-blade-servers/index.html)

Cisco UCS Adapters:

[http://www.cisco.com/en/US/products/ps10277/prod\_module\_series\_home.html](https://www.cisco.com/en/US/products/ps10277/prod_module_series_home.html%20)

Cisco UCS Manager:

[http://www.cisco.com/en/US/products/ps10281/index.html](https://www.cisco.com/en/US/products/ps10281/index.html%20)

Cisco Nexus 9000 Series Switches:

[http://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/index.html](https://www.cisco.com/c/en/us/products/switches/nexus-9000-series-switches/index.html)

VMware vCenter Server:

[http://www.vmware.com/products/vcenter-server/overview.html](http://www.vmware.com/products/vcenter-server/overview.html%20)

VMware vSphere:

<https://www.vmware.com/products/vsphere>

NetApp SolidFire:

<http://www.netapp.com/us/products/storage-systems/all-flash-array/solidfire-web-scale.aspx>

NetApp SolidFire Element OS:

<http://www.netapp.com/us/products/data-management-software/element-os.aspx>

**Interoperability Matrixes**

Cisco UCS Hardware Compatibility Matrix:

[http://www.cisco.com/c/en/us/support/servers-unified-computing/unified-computing-system/products-technical-reference-list.html](https://www.cisco.com/c/en/us/support/servers-unified-computing/unified-computing-system/products-technical-reference-list.html)

VMware and Cisco Unified Computing System:

[http://www.vmware.com/resources/compatibility](http://www.vmware.com/resources/compatibility%20)

NetApp Interoperability Matrix Tool:

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