Cisco HyperFlex System with HX Data Platform 1.8, a Hyperconverged Virtual Server Infrastructure

Design and Deployment of Cisco HyperFlex™ System for a Hyperconverged Virtual Server Infrastructure with HX Data platform 1.8 and VMWare vSphere 6.0U2

Last Updated: March 7, 2017
About Cisco Validated Designs

The CVD program consists of systems and solutions designed, tested, and documented to facilitate faster, more reliable, and more predictable customer deployments. For more information visit http://www.cisco.com/go/designzone.

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

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

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 Unity, Collaboration Without Limitation, EtherFast, EtherSwitch, Event Center, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, iQuick Study, IronPort, the IronPort logo, 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)

© 2016 Cisco Systems, Inc. All rights reserved.
Table of Contents

About Cisco Validated Designs .............................................................................................................. 2
Executive Summary ................................................................................................................................. 7
Solution Overview .................................................................................................................................. 8
  Introduction ........................................................................................................................................ 8
  Audience ........................................................................................................................................... 9
  Purpose of this Document ..................................................................................................................... 9
  What’s New? ..................................................................................................................................... 10
  Solution Summary .............................................................................................................................. 11
Technology Overview ............................................................................................................................. 13
  Cisco Unified Computing System .......................................................................................................... 13
  Cisco UCS Fabric Interconnect ............................................................................................................. 14
    Cisco UCS 6248UP Fabric Interconnect ............................................................................................... 14
  Cisco HyperFlex HX-Series Nodes ........................................................................................................ 14
    Cisco HyperFlex HX220c-M4S nodes .................................................................................................... 15
  Cisco HyperFlex HX240c-M4SX nodes .................................................................................................. 15
    Cisco VIC 1227 MLOM interface card ................................................................................................. 15
    Cisco UCS B200-M4 nodes ................................................................................................................... 16
    Cisco UCS 5108 Blade Chassis ............................................................................................................. 16
    Cisco UCS 2204XP Fabric Extender ................................................................................................... 17
  Cisco HyperFlex Converged Data Platform Software ........................................................................... 17
    Cisco HyperFlex HX Data Platform Administration Plug-in ................................................................. 18
    Cisco HyperFlex HX Data Platform Controller ................................................................................... 19
  Data Operations and Distribution ........................................................................................................ 19
Solution Design ..................................................................................................................................... 23
  Requirements ...................................................................................................................................... 23
    Physical Components ........................................................................................................................ 23
    Software Components ........................................................................................................................ 26
  Considerations .................................................................................................................................. 27
    Version Control ............................................................................................................................... 27
    VCenter Server ................................................................................................................................. 28
  Scale .................................................................................................................................................. 28
  Capacity ............................................................................................................................................. 28
  Physical Topology ............................................................................................................................... 30
Topology Overview...........................................................................................................30
Fabric Interconnects........................................................................................................32
HX-Series Rack Mount Servers........................................................................................33
Logical Topology ...............................................................................................................34
Logical Network Design .................................................................................................34
Design Elements ..............................................................................................................37
Network Design ................................................................................................................37
Cisco UCS Uplink Connectivity .......................................................................................37
VLANs and Subnets .........................................................................................................39
Jumbo Frames ..................................................................................................................40
Cisco UCS Design ...........................................................................................................40
Cisco UCS Organization ...................................................................................................41
Cisco UCS LAN Policies ...................................................................................................41
Cisco UCS Server Policies ...............................................................................................53
Cisco UCS Service Profile Templates ..............................................................................59
ESXi Host Design ............................................................................................................62
Virtual Networking Design .............................................................................................62
VMDirectPath I/O Pass-through ......................................................................................63
Storage Platform Controller VMs ....................................................................................64
Installation .......................................................................................................................69
Prerequisites ....................................................................................................................69
IP Addressing ...................................................................................................................69
DHCP vs Static IP ..........................................................................................................72
DNS .................................................................................................................................72
NTP .................................................................................................................................74
VLANs ............................................................................................................................75
Network Uplinks ..............................................................................................................76
Usernames and Passwords ..............................................................................................77
Physical Installation .......................................................................................................78
Cabling .............................................................................................................................78
Cisco UCS Installation .....................................................................................................80
Cisco UCS 6248UP Fabric Interconnect A ......................................................................80
Cisco UCS 6248UP Fabric Interconnect B ......................................................................81
Cisco UCS Manager ........................................................................................................82
Cisco UCS Configuration .................................................................................................83
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validation</td>
<td>153</td>
</tr>
<tr>
<td>Post Install Checklist</td>
<td>153</td>
</tr>
<tr>
<td>Verify Redundancy</td>
<td>153</td>
</tr>
<tr>
<td>Documentation Roadmap</td>
<td>155</td>
</tr>
<tr>
<td>Appendix</td>
<td>156</td>
</tr>
<tr>
<td>A: Cluster Capacity Calculations</td>
<td>156</td>
</tr>
<tr>
<td>B: Example Scripts</td>
<td>156</td>
</tr>
<tr>
<td>Create DHCP Reservations</td>
<td>156</td>
</tr>
<tr>
<td>Configure ESXi Post Install</td>
<td>157</td>
</tr>
<tr>
<td>About the Authors</td>
<td>160</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>160</td>
</tr>
</tbody>
</table>
Executive Summary

The previous decade has seen major shifts in the datacenter, and arguably the most significant has been the widespread adoption of virtualization of servers as the primary computing platform for most businesses. The flexibility, speed of deployment, ease of management, portability, and improved resource utilization has led many enterprises to adopt a “virtual first” stance, where all environments are deployed virtually unless circumstances make that impossible. While the benefits of virtualization are clear, the proliferation of virtual environments has shone a light on other technology stacks where they do not offer the same levels of simplicity, flexibility, and rapid deployment as virtualized compute platforms do. Networking and storage systems in particular have come under increasing scrutiny to be as agile as hypervisors and virtual servers. Cisco offers strong solutions for rapid deployment and easy management of virtualized computing platforms, including integrated networking capabilities, with the Cisco Unified Computing System (Cisco UCS) product line. Now with the introduction of Cisco HyperFlex, we bring similar dramatic enhancements to the virtualized servers and hyperconverged storage market.

Cisco HyperFlex systems have been developed using the Cisco UCS platform, which combines Cisco HX-Series x86 servers and integrated networking technologies via the Cisco UCS Fabric Interconnects, into a single management domain, along with industry leading virtualization hypervisor software from VMware, and new software defined storage technology. The combination creates a virtualization platform that also provides the network connectivity for the guest virtual machine (VM) connections, and the distributed storage to house the VMs, spread across all of the Cisco UCS x86 servers instead of using specialized components. The unique storage features of the newly developed log based filesystem enable rapid cloning of VMs, snapshots without the traditional performance penalties, data deduplication and compression, without having to purchase all-flash based storage systems. All configuration, deployment, management, and monitoring of the solution can be done with existing tools for Cisco UCS and VMware, such as CISCO UCS Manager and VMware VCenter. This powerful linking of advanced technology stacks into a single, simple, rapidly deployed solution makes Cisco HyperFlex a true second generation hyperconverged platform for the modern datacenter.
Solution Overview

Introduction

The Cisco HyperFlex System provides an all-purpose virtualized server platform, with hypervisor hosts, networking connectivity, and virtual server storage across a set of Cisco UCS C-Series x86 rack-mount servers. Legacy datacenter deployments relied on a disparate set of technologies, each performing a distinct and specialized function, such as network switches connecting endpoints and transferring Ethernet network traffic, and Fibre Channel (FC) storage arrays providing block-based storage devices via a unique storage array network (SAN). Each of these systems had unique requirements for hardware, connectivity, management tools, operational knowledge, monitoring, and ongoing support. A legacy virtual server environment was often divided up into areas commonly referred to as silos, within which only a single technology operated, along with their correlated software tools and support staff. Silos could often be divided between the x86 computing hardware, the networking connectivity of those x86 servers, SAN connectivity and storage device presentation, the hypervisors and virtual platform management, and finally the guest VM themselves along with their OS and applications. This model proves to be inflexible, difficult to navigate, and is susceptible to numerous operational inefficiencies.

A more modern datacenter model was developed called a converged architecture. Converged architectures attempt to collapse the traditional silos by combining these technologies into a more singular environment, which has been designed to operate together in predefined, tested, and validated designs. A key component of the converged architecture was the revolutionary combination of x86 rack and blade servers, along with converged Ethernet and Fibre Channel networking offered by the Cisco UCS platform. Converged architectures leverage Cisco UCS, plus new deployment tools, management software suites, automation processes, and orchestration tools to overcome the difficulties deploying traditional environments, and do so in a much more rapid fashion. These new tools place the ongoing management and operation of the system into the hands of fewer staff, with more rapid deployment of workloads based on business needs, while still remaining at the forefront of flexibility to adapt to workload needs, and offering the highest possible performance. Cisco has had incredible success in these areas with our various partners, developing leading solutions such as Cisco FlexPod, SmartStack, VersaStack, and VBlock architectures. Despite these advances, since these converged architectures relied on some legacy technology stacks, particularly in the storage subsystems, there often remained a division of responsibility amongst multiple teams of administrators. Alongside, there is also a recognition that these converged architectures can still be a somewhat complex combination of components, where a simpler system would suffice to serve the workloads being requested.

Significant changes in the storage marketplace have given rise to the software defined storage (SDS) system. Legacy FC storage arrays often continued to utilize a specialized subset of hardware, such as Fibre Channel Arbitrated Loop (FC-AL) based controllers and disk shelves along with optimized Application Specific Integrated Circuits (ASIC), read/write data caching modules and cards, plus highly customized software to operate the arrays. With the rise of Serial Attached SCSI (SAS) bus technology and its inherent benefits, storage array vendors began to transition their internal architectures to SAS, and with dramatic increases in processing power from recent x86 processor architectures, they also used fewer or no custom ASICs at all. As disk physical sizes shrank, servers began to have the same density of storage per rack unit (RU) as the arrays themselves, and with the proliferation of NAND based flash memory solid state disks (SSD), they also now had access to input/output (IO) devices whose speed rivaled that of dedicated caching devices. If servers themselves now contained storage devices and technology to rival many dedicated arrays
on the market, then the major differentiator between them was the software providing allocation, presentation and management of the storage, plus the advanced features many vendors offered. This led to the rise of software defined storage, where the x86 servers with the storage devices ran software to effectively turn one or more of them into a storage array much the same as the traditional arrays were. In a somewhat unexpected turn of events, some of the major storage array vendors themselves were pioneers in this field, recognizing the shift in the market, and attempting to profit from the software features they offered versus specialized hardware as had been done in the past.

Some early uses of SDS systems simply replaced the traditional storage array in the converged architectures as described earlier. That configuration still had a separate storage system from the virtual server hypervisor platform, and depending on the solution provider, still remained separate from the network devices. If the server that hosted the virtual servers, and also provided the SDS environment were in fact the same model of servers, could they simply do both things at once and collapse the two functions into one? This combination of resources becomes what the industry has given the moniker of a hyperconverged infrastructure. Hyperconverged infrastructures combine the computing, memory, hypervisor, and storage devices of servers into a single monolithic platform for virtual servers. There is no longer a separate storage system, as the servers running the hypervisors also provide the software defined storage resources to store the virtual servers, effectively storing the virtual machines on themselves. Now nearly all the silos are gone, and a hyperconverged infrastructure becomes something almost completely self-contained, simpler to use, faster to deploy, easier to consume, yet still flexible and with very high performance. By combining the convergence of compute and network resources provided by Cisco UCS, along with the new hyperconverged storage software, the Cisco HyperFlex system uniquely provides the compute resources, network connectivity, storage, and hypervisor platform to run an entire virtual environment, all contained in a single system.

Some key advantages of Hyperconverged infrastructures are the simplification of deployment, day to day management operations, as well as increased agility, thereby reducing the amount operational costs. Since hyperconverged storage can be easily managed by an IT generalist, this can also reduce technical debt going forward that is often accrued by implementing complex systems that need dedicated management and skillsets.

**Audience**

The intended audience for this document includes, but is not limited to, sales engineers, field consultants, professional services, IT managers, partner engineering, and customers deploying the Cisco HyperFlex System. External references are provided wherever applicable, but readers are expected to be familiar with VMware specific technologies, infrastructure concepts, networking connectivity, and security policies of the customer installation.

**Purpose of this Document**

This document describes the steps required to deploy, configure, and manage a Cisco HyperFlex system. The document is based on all known best practices using the software, hardware and firmware revisions specified in the document. As such, recommendations and best practices can be amended with later versions. This document showcases the installation and configuration of a single 8 node Cisco HyperFlex cluster, and a 4+4 hybrid Cisco HyperFlex cluster, in a typical customer datacenter environment. While readers of this document are expected to have sufficient knowledge to install and configure the products used, configuration details that are important to the deployment of this solution are provided in this CVD.
What’s New?

The Cisco HyperFlex system has some new capabilities & enhancements in version 1.8. Some highlights include:

- Cisco HyperFlex now provides support for connecting to external storage arrays. Please refer to the interoperability documentation for supported arrays.

![External Storage Arrays](image1.png)

- Cisco HyperFlex increased scales for supporting up to 8 clusters per management domain and total of 128 nodes.

- Cisco HyperFlex now support the latest generation of Cisco UCS software, Cisco UCS Manager 3.1(2b) and later.

- Cisco HyperFlex now supports more configurations for independent scaling of compute and storage allowing the use of C220 and C240 platforms to function as compute only nodes.

![Compute Node Types](image2.png)

- Cisco HyperFlex enhanced user experience with and improved installer user interface and capabilities.
**Solution Overview**

The Cisco HyperFlex system provides a fully contained virtual server platform, with compute and memory resources, integrated networking connectivity, a distributed high performance log-based filesystem for VM storage, and the hypervisor software for running the virtualized servers, all within a single Cisco UCS management domain.

**Figure 3  HyperFlex System Overview**

The following are components of a Cisco HyperFlex system:

- Cisco UCS 6248UP Fabric Interconnects
- Cisco HyperFlex HX220c- M4S rack-mount servers

Or

- Cisco HyperFlex HX240c- M4SX rack-mount servers
- Cisco HX Data Platform Software
- VMware vSphere ESXi Hypervisor
- VMware vCenter Server (end-user supplied)

Optional components for additional compute only resources are:

- Cisco UCS 5108 Chassis
- Cisco UCS 2204XP Fabric Extender
- Cisco UCS B200- M4 blade servers
- Cisco C220-M4 or C240-M4 servers

**Figure 4  Cisco HyperFlex System**

<table>
<thead>
<tr>
<th>HX220c</th>
<th>HX240c</th>
<th>B200 &amp; HX240c</th>
</tr>
</thead>
</table>

![Image of Cisco HyperFlex System configurations](image-url)
Cisco Unified Computing System

The Cisco Unified Computing System is a next-generation data center platform that unites compute, network, and storage access. The platform, optimized for virtual environments, is designed using open industry-standard technologies and aims to reduce total cost of ownership (TCO) and increase business agility. The system integrates a low-latency, lossless 10 Gigabit Ethernet unified network fabric with enterprise-class, x86-architecture servers. It is an integrated, scalable, multi-chassis platform in which all resources participate in a unified management domain.

The main components of Cisco Unified Computing System are:

- **Computing**—The system is based on an entirely new class of computing system that incorporates rack-mount and blade servers based on Intel Xeon Processors.

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

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

- **Storage access**—The system provides consolidated access to both SAN storage and Network Attached Storage (NAS) over the unified fabric. By unifying the storage access the Cisco Unified Computing System can access storage over Ethernet, Fibre Channel, Fibre Channel over Ethernet (FCoE), and iSCSI. This provides customers with choice for storage access and investment protection. In addition, the server administrators can pre-assign storage-access policies for system connectivity to storage resources, simplifying storage connectivity, and management for increased productivity.

- **Management**—The system uniquely integrates all system components which enable the entire solution to be managed as a single entity by the Cisco UCS Manager (UCSM). The Cisco UCS Manager has 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.

The Cisco Unified Computing System is designed to deliver:

- A reduced Total Cost of Ownership and increased business agility.

- Increased IT staff productivity through just-in-time provisioning and mobility support.

- A cohesive, integrated system which unifies the technology in the data center. The system is managed, serviced and tested as a whole.

- Scalability through a design for hundreds of discrete servers and thousands of virtual machines and the capability to scale I/O bandwidth to match demand.
Technology Overview

- Industry standards supported by a partner ecosystem of industry leaders.

Cisco UCS Fabric Interconnect

The Cisco UCS 6200 Series Fabric Interconnect is a core part of the Cisco Unified Computing System, providing both network connectivity and management capabilities for the system. The Cisco UCS 6200 Series offers line-rate, low-latency, lossless 10 Gigabit Ethernet, Fibre Channel over Ethernet (FCoE) and Fibre Channel functions.

The Cisco UCS 6200 Series provides the management and communication backbone for the Cisco UCS C-Series and HX-Series rack-mount servers, Cisco UCS B-Series Blade Servers and Cisco UCS 5100 Series Blade Server Chassis. All servers and chassis, and therefore all blades, attached to the Cisco UCS 6200 Series Fabric Interconnects become part of a single, highly available management domain. In addition, by supporting unified fabric, the Cisco UCS 6200 Series provides both the LAN and SAN connectivity for all blades within its domain.

From a networking perspective, the Cisco UCS 6200 Series uses a cut-through architecture, supporting deterministic, low-latency, line-rate 10 Gigabit Ethernet on all ports, 1Tb switching capacity, 160 Gbps bandwidth per chassis, independent of packet size and enabled services. The product family supports Cisco low-latency, lossless 10 Gigabit Ethernet unified network fabric capabilities, which increase the reliability, efficiency, and scalability of Ethernet networks. The Fabric Interconnect supports multiple traffic classes over a lossless Ethernet fabric from a server through an interconnect. Significant TCO savings come from an FCoE-optimized server design in which network interface cards (NICs), host bus adapters (HBAs), cables, and switches can be consolidated.

Cisco UCS 6248UP Fabric Interconnect

The Cisco UCS 6248UP 48-Port Fabric Interconnect is a one-rack-unit (1RU) 10 Gigabit Ethernet, FCoE and Fiber Channel switch offering up to 960-Gbps throughput and up to 48 ports. The switch has 32 1/10-Gbps fixed Ethernet, FCoE and FC ports and one expansion slot.

Cisco HyperFlex HX-Series Nodes

A HyperFlex cluster requires a minimum of three HX-Series nodes (with disk storage). Data is replicated across at least two of these nodes, and a third node is required for continuous operation in the event of a single-node failure. Each node that has disk storage is equipped with at least one high-performance SSD drive for data caching and rapid acknowledgment of write requests. Each node also is equipped with up to
the platform’s physical capacity of spinning disks for maximum data capacity. At first release, we offer three tested cluster configurations:

**Cisco HyperFlex HX220c- M4S nodes**

This small footprint configuration contains a minimum of three nodes with six 1.2 terabyte (TB) SAS drives that contribute to cluster storage capacity, a 120 GB SSD housekeeping drive, a 480 GB SSD caching drive, and two Cisco Flexible Flash (FlexFlash) Secure Digital (SD) cards that act as boot drives.

*Figure 6  HX220c- M4S Node*

**Cisco HyperFlex HX240c- M4SX nodes**

This capacity optimized configuration contains a minimum of three nodes, a minimum of fifteen and up to twenty-three 1.2 TB SAS drives that contribute to cluster storage, a single 120 GB SSD housekeeping drive, a single 1.6 TB SSD caching drive, and two FlexFlash SD cards that act as the boot drives.

**Cisco VIC 1227 MLOM interface card**

The Cisco UCS Virtual Interface Card (VIC) 1227 is a dual-port Enhanced Small Form-Factor Pluggable (SFP+) 10-Gbps Ethernet and Fibre Channel over Ethernet (FCoE)-capable PCI Express (PCIe) modular LAN-on-motherboard (mLOM) adapter installed in the Cisco UCS HX-Series Rack Servers (Figure 7). The mLOM slot can be used to install a Cisco VIC without consuming a PCIe slot, which provides greater I/O expandability. It incorporates next-generation converged network adapter (CNA) technology from Cisco, providing investment protection for future feature releases. The card enables a policy-based, stateless, agile server infrastructure that can present up to 256 PCIe standards-compliant interfaces to the host that can be dynamically configured as either network interface cards (NICs) or host bus adapters (HBAs). The personality of the card is determined dynamically at boot time using the service profile associated with the server. The number, type (NIC or HBA), identity (MAC address and World Wide Name [WWN]), failover policy, bandwidth, and quality-of-service (QoS) policies of the PCIe interfaces are all determined using the service profile.

*Figure 7  Cisco VIC 1227 mLOM Card*
Cisco UCS B200-M4 nodes

For workloads that require additional computing and memory resources, but not additional storage capacity, a compute-intensive hybrid cluster configuration is allowed. This configuration contains a minimum of three HX240c-M4SX Nodes with up to four Cisco UCS B200-M4 Blade Servers for additional computing capacity. The HX240c-M4SX Nodes are configured as described previously, and the Cisco UCS B200-M4 servers are equipped with boot drives. Use of the B200-M4 compute nodes also requires the Cisco UCS 5108 blade server chassis, and a pair of Cisco UCS 2204XP Fabric Extenders.

Figure 8  Cisco UCS B200 M4 Node

Cisco UCS 5108 Blade Chassis

The Cisco UCS 5100 Series Blade Server Chassis is a crucial building block of the Cisco Unified Computing System, delivering a scalable and flexible blade server chassis. The Cisco UCS 5108 Blade Server Chassis, is six rack units (6RU) high and can mount in an industry-standard 19-inch rack. A single chassis can house up to eight half-width Cisco UCS B-Series Blade Servers and can accommodate both half-width and full-width blade form factors.

Four single-phase, hot-swappable power supplies are accessible from the front of the chassis. These power supplies are 92 percent efficient and can be configured to support non-redundant, N+1 redundant, 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 for Cisco UCS Fabric Extenders. A passive mid-plane provides up to 40 Gbps of I/O bandwidth per server slot from each Fabric Extender. The chassis is capable of supporting 40 Gigabit Ethernet standards.

Figure 9  Cisco UCS 5108 Blade Chassis Front and Back Views
Cisco UCS 2204XP Fabric Extender

The Cisco UCS 2200 Series Fabric Extenders multiplex and forward all traffic from blade servers in a chassis to a parent Cisco UCS fabric interconnect over from 10-Gbps unified fabric links. All traffic, even traffic between blades on the same chassis or virtual machines on the same blade, is forwarded to the parent interconnect, where network profiles are managed efficiently and effectively by the fabric interconnect. At the core of the Cisco UCS fabric extender are application-specific integrated circuit (ASIC) processors developed by Cisco that multiplex all traffic.

The Cisco UCS 2204XP Fabric Extender has four 10 Gigabit Ethernet, FCoE-capable, SFP+ ports that connect the blade chassis to the fabric interconnect. Each Cisco UCS 2204XP has sixteen 10 Gigabit Ethernet ports connected through the midplane 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.

Figure 10  Cisco UCS 2204XP Fabric Extender

Cisco HyperFlex Converged Data Platform Software

The Cisco HyperFlex HX Data Platform is a purpose-built, high-performance, distributed file system with a wide array of enterprise-class data management services. The data platform’s innovations redefine distributed storage technology, exceeding the boundaries of first-generation hyperconverged infrastructures. The data platform has all the features that you would expect of an enterprise shared storage system, eliminating the need to configure and maintain complex Fibre Channel storage networks and devices. The platform simplifies operations and helps ensure data availability. Enterprise-class storage features include the following:

- Replication replicates data across the cluster so that data availability is not affected if single or multiple components fail (depending on the replication factor configured).
- Deduplication is always on, helping reduce storage requirements in virtualization clusters in which multiple operating system instances in client virtual machines result in large amounts of replicated data.
- Compression further reduces storage requirements, reducing costs, and the log-structured file system is designed to store variable-sized blocks, reducing internal fragmentation.
- Thin provisioning allows large volumes to be created without requiring storage to support them until the need arises, simplifying data volume growth and making storage a “pay as you grow” proposition.
- Fast, space-efficient clones rapidly replicate storage volumes so that virtual machines can be replicated simply through metadata operations, with actual data copied only for write operations.
- Snapshots help facilitate backup and remote-replication operations: needed in enterprises that require always-on data availability.
Cisco HyperFlex HX Data Platform Administration Plug-in

The Cisco HyperFlex HX Data Platform is administered through a VMware vSphere web client plug-in. Through this centralized point of control for the cluster, administrators can create volumes, monitor the data platform health, and manage resource use. Administrators can also use this data to predict when the cluster will need to be scaled. For customers that prefer a lightweight web interface there is a technical preview URL management interface available by opening a browser to the IP address of the HX cluster interface. Additionally, there is an interface to assist in running CLI commands through a web browser.

Figure 11  HyperFlex Web Client Plug-in

For the Tech Preview Web UI connect to HX controller cluster IP:

http://hx_controller_cluster_ip/ui

Figure 12  Tech Preview UI
To run CLI commands through HTTP, connect to HX controller cluster IP (Figure 13).

### Figure 13  Web CLI

![Web CLI](image)

### Cisco HyperFlex HX Data Platform Controller

A Cisco HyperFlex HX Data Platform controller resides on each node and implements the distributed file system. The controller runs in user space within a virtual machine and intercepts and handles all I/O from guest virtual machines. The platform controller VM uses the VMDirectPath I/O feature to provide PCI pass-through control of the physical server’s SAS disk controller. This method gives the controller VM full control of the physical disk resources, utilizing the SSD drives as a read/write caching layer, and the HDDs as a capacity layer for distributed storage. The controller integrates the data platform into VMware software through the use of two preinstalled VMware ESXi vSphere Installation Bundles (VIBs):

- **IO Visor**: This VIB provides a network file system (NFS) mount point so that the ESXi hypervisor can access the virtual disks that are attached to individual virtual machines. From the hypervisor’s perspective, it is simply attached to a network file system.

- **VMware API for Array Integration (VAAI)**: This storage offload API allows vSphere to request advanced file system operations such as snapshots and cloning. The controller implements these operations through manipulation of metadata rather than actual data copying, providing rapid response, and thus rapid deployment of new environments.

### Data Operations and Distribution

The Cisco HyperFlex HX Data Platform controllers handle all read and write operation requests from the guest VMs to their virtual disks (VMDK) stored in the distributed datastores in the cluster. The data platform distributes the data across multiple nodes of the cluster, and also across multiple capacity disks of each node, according to the replication level policy selected during the cluster setup. This method avoids storage hotspots on specific nodes, and on specific disks of the nodes, and thereby also avoids networking hotspots or congestion from accessing more data on some nodes versus others.
Technology Overview

Replication Factor

The policy for the number of duplicate copies of each storage block is chosen during cluster setup, and is referred to as the replication factor (RF). The default setting for the Cisco HyperFlex HX Data Platform is replication factor 3 (RF=3).

- Replication Factor 3: For every I/O write committed to the storage layer, 2 additional copies of the blocks written will be created and stored in separate locations, for a total of 3 copies of the blocks. Blocks are distributed in such a way as to ensure multiple copies of the blocks are not stored on the same disks, nor on the same nodes of the cluster. This setting can tolerate simultaneous failures 2 entire nodes without losing data and resorting to restore from backup or other recovery processes.

- Replication Factor 2: For every I/O write committed to the storage layer, 1 additional copy of the blocks written will be created and stored in separate locations, for a total of 2 copies of the blocks. Blocks are distributed in such a way as to ensure multiple copies of the blocks are not stored on the same disks, nor on the same nodes of the cluster. This setting can tolerate a failure 1 entire node without losing data and resorting to restore from backup or other recovery processes.

Data Write Operations

For each write operation, data is written to the SSD of the node designated as it’s primary, and replica copies of that write are written to the caching SSD of the remote nodes in the cluster, according to the replication factor setting. For example, at RF=3 a write will be written locally where the VM originated the write, and two additional writes will be committed in parallel on two other nodes. The write operation will not be acknowledged until all three copies are written to the caching layer SSDs. Written data is also cached in a write log area resident in memory in the controller VM, along with the write log on the caching SSDs (Figure 14). This process speeds up read requests when reads are requested of data that has recently been written.

Data Destaging, Deduplication and Compression

The Cisco HyperFlex HX Data Platform constructs multiple write caching segments on the caching SSDs of each node in the distributed cluster. As write cache segments become full, and based on policies accounting for I/O load and access patterns, those write cache segments are locked and new writes roll over to a new write cache segment. The data in the now locked cache segment is destaged to the HDD capacity layer of the node. During the destaging process, data is deduplicated and compressed before being written to the HDD capacity layer. The resulting data after deduplication and compression can now be written in a single sequential operation to the HDDs of the server, avoiding disk head seek thrashing and accomplishing the task in the minimal amount of time (Figure 10). Since the data is already deduplicated and compressed before being written, the platform avoids additional I/O overhead often seen on competing systems, which must later do a read/dedupe/compress/write cycle. Deduplication, compression and destaging take place with no delays or I/O penalties to the guest VMs making requests to read or write data.
Data Read Operations

For data read operations, data may be read from multiple locations. For data that was very recently written, the data is likely to still exist in the write log of the local platform controller memory, or the write log of the local caching SSD. If local write logs do not contain the data, the distributed filesystem metadata will be queried to see if the data is cached elsewhere, either in write logs of remote nodes, or in the dedicated read cache area of the local and remote SSDs. Finally, if the data has not been accessed in a significant amount
of time, the filesystem will retrieve the data requested from the HDD capacity layer. As requests for reads are made to the distributed filesystem and the data is retrieved from the HDD capacity layer, the caching SSDs populate their dedicated read cache area to speed up subsequent requests for the same data. This multi-tiered distributed system with several layers of caching techniques, insures that data is served at the highest possible speed, leveraging the caching SSDs of the nodes fully and equally.
Solution Design

Requirements

The following sections detail the physical hardware, software revisions, and firmware versions required to install the Cisco HyperFlex system. The components described are for a single 8 node Cisco HX cluster, or for a single 4+4 node “hybrid” Cisco HX cluster. Maximum cluster size of 16 nodes can be obtained by combining 8 converged nodes (storage nodes), and 8 compute only nodes.

Physical Components

Table 1 lists the HyperFlex System Components.

<table>
<thead>
<tr>
<th>Component</th>
<th>Hardware Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabric Interconnect</td>
<td>Two Cisco UCS 6248UP Fabric Interconnects (optional) plus 16 port UP expansion module</td>
</tr>
<tr>
<td>Servers</td>
<td>Eight Cisco HX-series HX220c-M4S servers</td>
</tr>
<tr>
<td></td>
<td>Or</td>
</tr>
<tr>
<td></td>
<td>Eight Cisco HX-series HX240c-M4SX servers</td>
</tr>
<tr>
<td></td>
<td>Or</td>
</tr>
<tr>
<td></td>
<td>Eight Cisco HX-Series HX240c-M4SX servers plus Eight Cisco UCS B200-M4 blade servers.</td>
</tr>
<tr>
<td></td>
<td>Note: you can also use C220-M4 or C240-M4 series in place of the B200-M4 for compute only nodes</td>
</tr>
<tr>
<td>Chassis</td>
<td>Cisco UCS 5108 Blade Chassis (only if using the B200-M4 servers)</td>
</tr>
<tr>
<td>Fabric Extenders</td>
<td>Cisco UCS 2204XP Fabric Extenders (required for the 5108 blade chassis and B200-M4 blades)</td>
</tr>
</tbody>
</table>

Table 2 lists some of the available processor models for the Cisco HX-Series HX220c-M4S and HX240c-M4SX servers, please refer to the links below for more information:

Compare models:

HX220c Spec Sheet:

HX240c Spec Sheet:

Table 2  HyperFlex Server CPU Options

<table>
<thead>
<tr>
<th>Model</th>
<th>Cores</th>
<th>Clock Speed</th>
<th>Cache</th>
<th>RAM Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>E5-2699 v4</td>
<td>22</td>
<td>2.2 GHz</td>
<td>55 MB</td>
<td>2400 MHz</td>
</tr>
<tr>
<td>E5-2698 v4</td>
<td>20</td>
<td>2.2 GHz</td>
<td>50 MB</td>
<td>2400 MHz</td>
</tr>
<tr>
<td>E5-2697 v4</td>
<td>18</td>
<td>2.6 GHz</td>
<td>45 MB</td>
<td>2400 MHz</td>
</tr>
<tr>
<td>E5-2690 v4</td>
<td>14</td>
<td>2.6 GHz</td>
<td>35 MB</td>
<td>2400 MHz</td>
</tr>
<tr>
<td>E5-2683 v4</td>
<td>16</td>
<td>2.1 GHz</td>
<td>40 MB</td>
<td>2400 MHz</td>
</tr>
<tr>
<td>E5-2680 v4</td>
<td>14</td>
<td>2.4 GHz</td>
<td>35 MB</td>
<td>2400 MHz</td>
</tr>
<tr>
<td>E5-2667 v4</td>
<td>8</td>
<td>3.2 GHz</td>
<td>25 MB</td>
<td>2400 MHz</td>
</tr>
<tr>
<td>E5-2660 v4</td>
<td>14</td>
<td>2.0 GHz</td>
<td>35 MB</td>
<td>2400 MHz</td>
</tr>
<tr>
<td>E5-2650 v4</td>
<td>12</td>
<td>2.2 GHz</td>
<td>30 MB</td>
<td>2400 MHz</td>
</tr>
<tr>
<td>E5-2643 v4</td>
<td>10</td>
<td>3.4 GHz</td>
<td>20 MB</td>
<td>2400 MHz</td>
</tr>
<tr>
<td>E5-2630 v4</td>
<td>10</td>
<td>2.2 GHz</td>
<td>25 MB</td>
<td>2133 MHz</td>
</tr>
<tr>
<td>E5-2620 v4</td>
<td>8</td>
<td>2.1 GHz</td>
<td>20 MB</td>
<td>2133 MHz</td>
</tr>
</tbody>
</table>

Table 3  lists the hardware component options for the HX220c-M4S server model.

Table 3  HX220c-M4S Server Options

<table>
<thead>
<tr>
<th>HX220c-M4S options</th>
<th>Hardware Required</th>
</tr>
</thead>
</table>
### HX220c-M4S options

<table>
<thead>
<tr>
<th>Hardware Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processors</td>
</tr>
<tr>
<td>Memory</td>
</tr>
<tr>
<td>Disk Controller</td>
</tr>
</tbody>
</table>
| SSD | One 120 GB 2.5 Inch Enterprise Value 6G SATA SSD  
And  
One 480 GB 2.5 Inch Enterprise Performance 6G SATA SSD (3X endurance) |
| HDD | Six 1.2 TB SAS 12Gbps 10K rpm SFF HDD |
| Network | Cisco UCS VIC1227 VIC MLOM – Dual 10 GbE port |
| Boot Devices | Two 64GB SD Cards for UCS Servers |

Table 4 lists the hardware component options for the HX240c-M4SX server model

### Table 4 HX240c-M4SX Server Options

<table>
<thead>
<tr>
<th>Hardware Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processors</td>
</tr>
<tr>
<td>Memory</td>
</tr>
<tr>
<td>Disk Controller</td>
</tr>
</tbody>
</table>
| SSD | One 120 GB 2.5 Inch Enterprise Value 6G SATA SSD (in the rear internal SAS device bay)  
And  
One 1.6 TB 2.5 Inch Enterprise Performance 6G SATA SSD (3X endurance) |
Table 5 lists the hardware component options for the B200-M4 server model.

Table 5 Cisco UCS B200-M4 Server Options

<table>
<thead>
<tr>
<th>B200-M4 options</th>
<th>Hardware Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processors</td>
<td>Chose any Intel E5-26xx v4 processor model.</td>
</tr>
<tr>
<td>Memory</td>
<td>Any supported amount of total memory using 16 GB or 64 GB DDR4 2133-2400 MHz 1.2v modules</td>
</tr>
<tr>
<td>Disk Controller</td>
<td>none</td>
</tr>
<tr>
<td>SSD</td>
<td>none</td>
</tr>
<tr>
<td>HDD</td>
<td>none</td>
</tr>
<tr>
<td>Network</td>
<td>Cisco UCS VIC1340 VIC MLOM</td>
</tr>
<tr>
<td>Boot Devices</td>
<td>Two 64GB SD Cards for UCS Servers or San Boot</td>
</tr>
</tbody>
</table>

Software Components

Table 6 lists the software components and the versions required for the Cisco HyperFlex system.
<table>
<thead>
<tr>
<th>Component</th>
<th>Software Required</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypervisor</td>
<td>VMware ESXi 5.5 update 3b, patch 1 (HX240c only)</td>
</tr>
<tr>
<td></td>
<td>Or</td>
</tr>
<tr>
<td></td>
<td>VMware ESXi 6.0 U2</td>
</tr>
<tr>
<td></td>
<td>ESXi 6.0 U2 is recommended (CISCO Custom Image for ESXi 6.0 U2 Patch 3, HX-Vmware-ESXi-60U2-4192238-Cisco-Custom-6.0.2.3.iso)</td>
</tr>
</tbody>
</table>

**Note:** Use of a published Cisco custom ESXi ISO installer file is required.

**Note:** Enterprise or Enterprise Plus licensing is required from VMware.

**Note:** VMware vSphere Standard, Essentials Plus, and ROBO editions are also supported but only for vSphere 6.0 versions.

<table>
<thead>
<tr>
<th>Management Server</th>
<th>VMware vCenter Server for Windows or vCenter Server Appliance 5.5 update 3b or later.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Cisco HyperFlex HX Data Platform</th>
<th>Cisco HyperFlex HX Data Platform Software 1.8.1c</th>
</tr>
</thead>
</table>

| Cisco UCS Firmware               | Cisco UCS Infrastructure software, B-Series and C-Series bundles, revision 3.1(2b)    |

**Considerations**

**Version Control**

Note that the software versions listed in Table 6 are the only valid and supported configuration at the time of the publishing of this validated design. Special care must be taken not to alter the revision of the hypervisor, vCenter server, Cisco HX platform software, or the Cisco UCS firmware without first consulting the appropriate release notes and compatibility matrixes to ensure that the system is not being modified into an unsupported configuration.
VCenter Server

This document does not cover the installation and configuration of VMware vCenter Server for Windows, or the vCenter Server Appliance. The vCenter Server must be installed and operational prior to the installation of the Cisco HyperFlex HX Data Platform software. The following best practice guidance applies to installations of HyperFlex 1.8.1c:

- Do not modify the default TCP port settings of the vCenter installation. Using non-standard ports can lead to failures during the installation.

- It is recommended to build the vCenter server on a physical server or in a virtual environment outside of the HyperFlex cluster. Building the vCenter server as a virtual machine inside the HyperFlex cluster environment is highly discouraged. There is a tech note for a method of deployment using USB SSD as temporary storage if not external server is available.

Scale

Cisco HyperFlex clusters currently scale up from a minimum of 3 to a maximum of 8 converged nodes per cluster, i.e. 8 nodes providing storage resources to the HX Distributed Filesystem. For the compute intensive “hybrid” design, a configuration with 3-8 Cisco HX240c-M4 or HX220c-M4 converged nodes can be combined with up to 8 compute nodes. Cisco B200-M4 blades, C220-M4, or C240-M4 servers can be used for the compute only nodes.

When the maximum size of a cluster has been reached, the environment can be “scaled out” by adding additional servers to the Cisco UCS domain, installing an additional HyperFlex cluster on them, and controlling them through the same vCenter server. A maximum of 8 HyperFlex clusters can be managed by a single vCenter server, therefore the maximum size of a single HyperFlex environment is 64 converged nodes, plus up to 64 additional compute only blades.

Capacity

Overall usable cluster capacity is based on a number of factors. The number of nodes in the cluster must be considered, plus the number and size of the capacity layer disks. Caching disk sizes are not calculated as part of the cluster capacity. The replication factor of the HyperFlex HX Data Platform also affects the cluster capacity as it defines the number of copies of each block of data written.

Disk drive manufacturers have adopted a size reporting methodology using calculation by powers of 10, also known as decimal prefix. As an example, a 120 GB disk is listed with a minimum of 120 x 10^9 bytes of usable addressable capacity, or 120 billion bytes. However, many operating systems and filesystems report their space based on standard computer binary exponentiation, or calculation by powers of 2, also called binary prefix. In this example, 2^10 or 1024 bytes make up a kilobyte, 2^10 kilobytes make up a megabyte, 2^10 megabytes make up a gigabyte, and 2^10 gigabytes make up a terabyte. As the values increase, the disparity between the two systems of measurement and notation get worse, at the terabyte level, the deviation between a decimal prefix value and a binary prefix value is nearly 10%.

The International System of Units (SI) defines values and decimal prefix by powers of 10 as follows:

<table>
<thead>
<tr>
<th>Value</th>
<th>Symbol</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) defines values and binary prefix by powers of 2 in ISO/IEC 80000-13:2008 Clause 4 as follows:

<table>
<thead>
<tr>
<th>Value</th>
<th>Symbol</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000 bytes</td>
<td>kB</td>
<td>kilobyte</td>
</tr>
<tr>
<td>1000 kB</td>
<td>MB</td>
<td>megabyte</td>
</tr>
<tr>
<td>1000 MB</td>
<td>GB</td>
<td>gigabyte</td>
</tr>
<tr>
<td>1000 GB</td>
<td>TB</td>
<td>terabyte</td>
</tr>
</tbody>
</table>

For the purpose of this document, the decimal prefix numbers are used only for raw disk capacity as listed by the respective manufacturers. For all calculations where raw or usable capacities are shown from the perspective of the HyperFlex software, filesystems or operating systems, the binary prefix numbers are used. This is done primarily to show a consistent set of values as seen by the end user from within the HyperFlex vCenter Web Plugin when viewing cluster capacity, allocation and consumption, and also within most operating systems.

Table 7 lists a set of HyperFlex HX Data Platform cluster usable capacity values, using binary prefix, for an array of cluster configurations. These values are useful for determining the appropriate size of HX cluster to initially purchase, and how much capacity can be gained by adding capacity disks. The calculations for these values are listed in Appendix A.

<table>
<thead>
<tr>
<th>HX-Series Server Model</th>
<th>Node Quantity</th>
<th>Capacity Disk Size (each)</th>
<th>Capacity Disk Quantity (per node)</th>
<th>Cluster Usable Capacity at RF=2</th>
<th>Cluster Usable Capacity at RF=3</th>
</tr>
</thead>
</table>
### Physical Topology

**Topology Overview**

The Cisco HyperFlex system is composed of a pair of Cisco UCS 6248UP Fabric Interconnects, along with up to 8 HX-Series rack mount servers per cluster. 8 compute only servers can also be added per clusters. Adding Cisco UCS 5108 Blade chassis allows use of Cisco UCS B200-M4 blade servers for additional compute resources in a hybrid cluster design. Cisco C240 and C220 servers can also be used for additional compute resources. Up to 8 separate HX clusters can be installed under a single pair of Fabric Interconnects. The Fabric Interconnects both connect to every HX-Series rack mount server, and both connect to every Cisco UCS 5108 blade chassis. Upstream network connections, also referred to as “northbound” network connections are made from the Fabric Interconnects to the customer datacenter network at the time of installation.
Figure 15  HyperFlex Standard Topology

Customer Network

UCS 6248UP
UCS 6248UP

3-8 HX220c or HX240c servers
Fabric Interconnects

Fabric Interconnects (FI) are deployed in pairs, wherein the two units operate as a management cluster, while forming two separate network fabrics, referred to as the A side and B side fabrics. Therefore, many design elements will refer to FI A or FI B, alternatively called fabric A or fabric B. Both Fabric Interconnects are active at all times, passing data on both network fabrics for a redundant and highly available configuration. Management services, including Cisco UCS Manager, are also provided by the two FIs but in a clustered manner, where one FI is the primary, and one is secondary, with a roaming clustered IP address. This primary/secondary relationship is only for the management cluster, and has no effect on data transmission.
Solution Design

Fabric Interconnects have the following ports, which must be connected for proper management of the Cisco UCS domain:

- **Mgmt:** A 10/100/1000 Mbps port for managing the Fabric Interconnect and the Cisco UCS domain through GUI and CLI tools. Also used by remote KVM, IPMI and SOL sessions to the managed servers within the domain. This is typically connected to the customer management network.

- **L1:** A cross connect port for forming the Cisco UCS management cluster. This is connected directly to the L1 port of the paired Fabric Interconnect using a standard CAT5 or CAT6 Ethernet cable with RJ45 plugs. It is not necessary to connect this to a switch or hub.

- **L2:** A cross connect port for forming the Cisco UCS management cluster. This is connected directly to the L2 port of the paired Fabric Interconnect using a standard CAT5 or CAT6 Ethernet cable with RJ45 plugs. It is not necessary to connect this to a switch or hub.

- **Console:** An RJ45 serial port for direct console access to the Fabric Interconnect. Typically used during the initial FI setup process with the included serial to RJ45 adapter cable. This can also be plugged into a terminal aggregator or remote console server device.

**HX- Series Rack Mount Servers**

The HX-Series converged servers are connected directly to the Cisco UCS Fabric Interconnects in Direct Connect mode. This option enables Cisco UCS Manager to manage the HX-Series rack-mount Servers using a single cable for both management traffic and data traffic. Both the HX220c-M4S and HX240c-M4SX servers are configured with the Cisco VIC 1227 network interface card (NIC) installed in a modular LAN on motherboard (MLOM) slot, which has dual 10 Gigabit Ethernet (GbE) ports. The standard and redundant connection practice is to connect port 1 of the VIC 1227 to a port on FI A, and port 2 of the VIC 1227 to a port on FI B (Figure 14). Failure to follow this cabling practice can lead to errors, discovery failures, and loss of redundant connectivity.

**Figure 17   HX- Series Server Connectivity**

![HX-Series Server Connectivity Diagram](image)

Cisco UCS B-Series Blade Servers

Hybrid HyperFlex clusters also incorporate 1-4 Cisco UCS B200-M4 blade servers for additional compute capacity. Like all other Cisco UCS B-series blade servers, the B200-M4 must be installed within a Cisco UCS 5108 blade chassis. The blade chassis comes populated with 1-4 power supplies, and 8 modular cooling fans. In the rear of the chassis are two bays for installation of Cisco Fabric Extenders. The Fabric Extenders (also commonly called IO Modules, or IOMs) connect the chassis to the Fabric Interconnects. Internally, the Fabric Extenders connect to the Cisco VIC 1340 card installed in each blade server across the
chassis backplane. The standard connection practice is to connect 1-4 10 GbE links from the left side IOM, or IOM 1, to FI A, and to connect the same number of 10 GbE links from the right side IOM, or IOM 2, to FI B (Figure 12). All other cabling configurations are invalid, and can lead to errors, discovery failures, and loss of redundant connectivity.

Figure 18 Cisco UCS 5108 Chassis Connectivity

Logical Topology

Logical Network Design

The Cisco HyperFlex system has communication pathways that fall into four defined zones (Figure 4):

- Management Zone: This zone comprises the connections needed to manage the physical hardware, the hypervisor hosts, and the storage platform controller virtual machines (SCVM). These interfaces and IP addresses need to be available to all staff who will administer the HX system, throughout the LAN/WAN. This zone must provide access to Domain Name System (DNS) and Network Time Protocol (NTP) services, and allow Secure Shell (SSH) communication. In this zone are multiple physical and virtual components:
  - Fabric Interconnect management ports.
- Cisco UCS external management interfaces used by the servers and blades, which answer through the FI management ports.

- ESXi host management interfaces.

- Storage Controller VM management interfaces.

- A roaming HX cluster management interface.

- VM Zone: This zone comprises the connections needed to service network IO to the guest VMs that will run inside the HyperFlex hyperconverged system. This zone typically contains multiple VLANs, that are trunked to the Cisco UCS Fabric Interconnects through the network uplinks, and tagged with 802.1Q VLAN IDs. These interfaces and IP addresses need to be available to all staff and other computer endpoints which need to communicate with the guest VMs in the HX system, throughout the LAN/WAN.

- Storage Zone: This zone comprises the connections used by the Cisco HX Data Platform software, ESXi hosts, and the storage controller VMs to service the HX Distributed Data Filesystem. These interfaces and IP addresses need to be able to communicate with each other at all times for proper operation. During normal operation, this traffic all occurs within the UCS domain, however there are hardware failure scenarios where this traffic would need to traverse the network northbound of the Cisco UCS domain. For that reason, the VLAN used for HX storage traffic must be able to traverse the network uplinks from the UCS domain, reaching FI A from FI B, and vice-versa. This zone is primarily jumbo frame traffic therefore jumbo frames must be enabled on the UCS uplinks. In this zone are multiple components:
  - A vmkernel interface used for storage traffic for each ESXi host in the HX cluster.
  - Storage Controller VM storage interfaces.
  - A roaming HX cluster storage interface.

- VMotion Zone: This zone comprises the connections used by the ESXi hosts to enable vMotion of the guest VMs from host to host. During normal operation, this traffic all occurs within the Cisco UCS domain, however there are hardware failure scenarios where this traffic would need to traverse the network northbound of the Cisco UCS domain. For that reason, the VLAN used for HX storage traffic must be able to traverse the network uplinks from the Cisco UCS domain, reaching FI A from FI B, and vice-versa.

Refer to Figure 19 for an illustration of the logical network design.
Figure 19  Logical Network Design

Solid lines represent active data paths
Dashed lines represent standby data paths
Installation of the HyperFlex system is primarily done through a deployable HyperFlex installer virtual machine, available for download at cisco.com as an OVA file. The installer VM does most of the Cisco UCS configuration work, it can be leveraged to simplify the installation of ESXi on the HyperFlex hosts, and also performs significant portions of the ESXi configuration. Finally, the installer VM is used to install the HyperFlex HX Data Platform software and create the HyperFlex cluster. Because this simplified installation method has been developed by Cisco, this CVD will not give detailed manual steps for the configuration of all the elements that are handled by the installer. Instead, the elements configured will be described and documented in this section, and the subsequent sections will guide you through the manual steps needed for installation, and how to utilize the HyperFlex Installer for the remaining configuration steps.

Network Design

Cisco UCS Uplink Connectivity

Cisco UCS network uplinks connect “northbound” from the pair of Cisco UCS Fabric Interconnects to the LAN in the customer datacenter. All Cisco UCS uplinks operate as trunks, carrying multiple 802.1Q VLAN IDs across the uplinks. The default Cisco UCS behavior is to assume that all VLAN IDs defined in the Cisco UCS configuration are eligible to be trunked across all available uplinks.

Cisco UCS Fabric Interconnects appear on the network as a collection of endpoints versus another network switch. Internally, the Fabric Interconnects do not participate in spanning-tree protocol (STP) domains, and the Fabric Interconnects cannot form a network loop, as they are not connected to each other with a layer 2 Ethernet link. All link up/down decisions through STP will be made by the upstream root bridges.

Uplinks need to be connected and active from both Fabric Interconnects. For redundancy, multiple uplinks can be used on each FI, either as 802.3ad Link Aggregation Control Protocol (LACP) port-channels, or using individual links. For the best level of performance and redundancy, uplinks can be made as LACP port-channels to multiple upstream Cisco switches using the virtual port channel (vPC) feature. Using vPC uplinks allows all uplinks to be active passing data, plus protects against any individual link failure, and the failure of an upstream switch. Other uplink configurations can be redundant, but spanning-tree protocol loop avoidance may disable links if vPC is not available.

All uplink connectivity methods must allow for traffic to pass from one Fabric Interconnect to the other, or from fabric A to fabric B. There are scenarios where cable, port or link failures would require traffic that normally does not leave the Cisco UCS domain, to now be forced over the Cisco UCS uplinks. Additionally, this traffic flow pattern can be seen briefly during maintenance procedures, such as updating firmware on the Fabric Interconnects, which requires them to be rebooted. The following sections and figures detail several uplink connectivity options.

Single Uplinks to Single Switch

This connection design is susceptible to failures at several points; single uplink failures on either Fabric Interconnect can lead to connectivity losses or functional failures, and the failure of the single uplink switch will cause a complete connectivity outage.
Figure 20  Connectivity with Single Uplink to Single Switch

Port Channels to Single Switch

This connection design is now redundant against the loss of a single link, but remains susceptible to the failure of the single switch.

Figure 21  Connectivity with Port-Channels to Single Switch

Single Uplinks or Port Channels to Multiple Switches

This connection design is redundant against the failure of an upstream switch, and redundant against a single link failure. In normal operation, STP is likely to block half of the links to avoid a loop across the two upstream switches. The side effect of this is to reduce bandwidth between the Cisco UCS domain and the LAN. If any of the active links were to fail, STP would bring the previously blocked link online to provide access to that Fabric Interconnect through the other switch. It is not recommended to connect both links from a single FI to a single switch, as that configuration is susceptible to a single switch failure breaking connectivity from fabric A to fabric B. For enhanced redundancy, the single links in the figure below could also be port-channels.
vPC to Multiple Switches

This recommended connection design relies on using Cisco switches that have the virtual port channel feature, such as Catalyst 6000 series switches running VSS, Cisco Nexus 5000 series, and Nexus 9000 series switches. Logically the two vPC enabled switches appear as one, and therefore spanning-tree protocol will not block any links. This configuration allows for all links to be active, achieving maximum bandwidth potential, and multiple redundancy at each level.

VLANs and Subnets

For the base HyperFlex system configuration, multiple VLANs need to be carried to the Cisco UCS domain from the upstream LAN, and these VLANs are also defined in the Cisco UCS configuration. The following table lists the VLANs created by the HyperFlex installer in Cisco UCS, and their functions.

<table>
<thead>
<tr>
<th>VLAN Name</th>
<th>VLAN ID</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>VLAN Name</td>
<td>VLAN ID</td>
<td>Purpose</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| hx-inband-mgmt  | Customer supplied | ESXi host management interfaces  
HX Storage Controller VM management interfaces  
HX Storage Cluster roaming management interface |
| hx-storage-data | Customer supplied | ESXi host storage vmkernel interfaces  
HX Storage Controller storage network interfaces  
HX Storage Cluster roaming storage interface |
| hx-vm-data      | Customer supplied | Guest VM network interfaces                                               |
| hx-vmotion      | Customer supplied | ESXi host vMotion vmkernel interfaces                                    |

A dedicated network or subnet for physical device management is often used in datacenters. In this scenario, the mgmt0 interfaces of the two Fabric Interconnects would be connected to that dedicated network or subnet. This is a valid configuration for HyperFlex installations with the following caveat; wherever the HyperFlex installer is deployed it must have IP connectivity to the subnet of the mgmt0 interfaces of the Fabric Interconnects, and also have IP connectivity to the subnets used by the hx-inband-mgmt VLANs listed above.

**Jumbo Frames**

All HyperFlex storage traffic traversing the hx-storage-data VLAN and subnet is configured to use jumbo frames, or to be precise all communication is configured to send IP packets with a Maximum Transmission Unit (MTU) size of 9000 bytes. Using a larger MTU value means that each IP packet sent carries a larger payload, therefore transmitting more data per packet, and consequently sending and receiving data faster. This requirement also means that the Cisco UCS uplinks must be configured to pass jumbo frames. Failure to configure the Cisco UCS uplink switches to allow jumbo frames can lead to service interruptions during some failure scenarios, particularly when cable or port failures would cause storage traffic to traverse the northbound Cisco UCS uplink switches.

**Cisco UCS Design**

This section on Cisco UCS design will describe the elements within Cisco UCS Manager that are configured by the Cisco HyperFlex installer. Many of the configuration elements are fixed in nature, meanwhile the HyperFlex installer does allow for some items to be specified at the time of creation, for example VLAN names and IDs, IP pools and more. Where the elements can be manually set during the installation, those items will be noted in <<>> brackets.
Cisco UCS Organization

During the HyperFlex Installation a Cisco UCS Sub-Organization is created named “hx-cluster”. The sub-organization is created underneath the root level of the UCS hierarchy, and is used to contain all policies, pools, templates and service profiles used by HyperFlex. This arrangement allows for organizational control using Role-Based Access Control (RBAC) and administrative locales at a later time if desired. In this way, control can be granted to administrators of only the HyperFlex specific elements of the Cisco UCS domain, separate from control of root level elements or elements in other sub-organizations.

Figure 24  Cisco UCS HyperFlex Sub-Organization

Cisco UCS LAN Policies

QoS System Classes

Specific UCS Quality of Service (QoS) system classes are defined for a Cisco HyperFlex system. These classes define Class of Service (CoS) values that can be used by the uplink switches north of the Cisco UCS domain, plus which classes are active, along with whether packet drop is allowed, the relative weight of the different classes when there is contention, the maximum transmission unit (MTU) size, and if there is multicast optimization applied. QoS system classes are defined for the entire Cisco UCS domain, the classes that are enabled can later be used in QoS policies, which are then assigned to Cisco UCS vNICs. The following table and figure details the QoS System Class settings configured for HyperFlex.

Table 9  QoS System Classes

<table>
<thead>
<tr>
<th>Priority</th>
<th>Enabled</th>
<th>CoS</th>
<th>Packet Drop</th>
<th>Weight</th>
<th>MTU</th>
<th>Multicast Optimized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platinum</td>
<td>Yes</td>
<td>5</td>
<td>No</td>
<td>4</td>
<td>9216</td>
<td>No</td>
</tr>
<tr>
<td>Gold</td>
<td>Yes</td>
<td>4</td>
<td>Yes</td>
<td>4</td>
<td>Normal</td>
<td>No</td>
</tr>
<tr>
<td>Silver</td>
<td>Yes</td>
<td>2</td>
<td>Yes</td>
<td>Best-effort</td>
<td>Normal</td>
<td>Yes</td>
</tr>
<tr>
<td>Bronze</td>
<td>Yes</td>
<td>1</td>
<td>Yes</td>
<td>Best-effort</td>
<td>9216</td>
<td>No</td>
</tr>
<tr>
<td>Best Effort</td>
<td>Yes</td>
<td>Any</td>
<td>Yes</td>
<td>Best-effort</td>
<td>Normal</td>
<td>No</td>
</tr>
<tr>
<td>Fibre Channel</td>
<td>Yes</td>
<td>3</td>
<td>No</td>
<td>5</td>
<td>FC</td>
<td>N/A</td>
</tr>
</tbody>
</table>
QoS Policies

In order to apply the settings defined in the Cisco UCS QoS System Classes, specific QoS Policies must be created, and then assigned to the vNICs, or vNIC templates used in Cisco UCS Service Profiles. The following table details the QoS Policies configured for HyperFlex, and their default assignment to the vNIC templates created.

<table>
<thead>
<tr>
<th>Policy</th>
<th>Priority</th>
<th>Burst</th>
<th>Rate</th>
<th>Host Control</th>
<th>Used by vNIC Template:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platinum</td>
<td>Platinum</td>
<td>10240</td>
<td>Line-rate</td>
<td>None</td>
<td>storage-data-a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>storage-data-b</td>
</tr>
<tr>
<td>Gold</td>
<td>Gold</td>
<td>10240</td>
<td>Line-rate</td>
<td>None</td>
<td>vm-network-a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>vm-network-b</td>
</tr>
<tr>
<td>Silver</td>
<td>Silver</td>
<td>10240</td>
<td>Line-rate</td>
<td>None</td>
<td>hv-mgmt-a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hv-mgmt-b</td>
</tr>
<tr>
<td>Bronze</td>
<td>Bronze</td>
<td>10240</td>
<td>Line-rate</td>
<td>None</td>
<td>hv-vmotion-a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hv-vmotion-b</td>
</tr>
<tr>
<td>Best Effort</td>
<td>Best Effort</td>
<td>10240</td>
<td>Line-rate</td>
<td>None</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Multicast Policy

A Cisco UCS Multicast Policy is configured by the HyperFlex installer, which is referenced by the VLANs that are created. The policy allows for future flexibility if a specific multicast policy needs to be created and applied to other VLANs, that may be used by non-HyperFlex workloads in the Cisco UCS domain. The following table and figure details the Multicast Policy configured for HyperFlex.
VLANs

VLANs are created by the HyperFlex installer to support a base HyperFlex system, with a single VLAN defined for guest VM traffic, and a VLAN for vMotion. Names and IDs for the VLANs are defined in the Cisco UCS configuration page of the HyperFlex installer web interface. The VLANs listed in Cisco UCS must already be present on the upstream network, and the Cisco UCS FIs do not participate in VLAN Trunk Protocol (VTP). The following table and figure details the VLANs configured for HyperFlex.

**Table 11  Cisco UCS VLANs**

<table>
<thead>
<tr>
<th>Name</th>
<th>ID</th>
<th>Type</th>
<th>Transport</th>
<th>Native</th>
<th>VLAN Sharing</th>
<th>Multicast Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;&lt;hx-inband-mgmt&gt;&gt;</td>
<td>&lt;user_defined&gt;</td>
<td>LAN</td>
<td>Ether</td>
<td>No</td>
<td>None</td>
<td>HyperFlex</td>
</tr>
<tr>
<td>&lt;&lt;hx-storage-data&gt;&gt;</td>
<td>&lt;user_defined&gt;</td>
<td>LAN</td>
<td>Ether</td>
<td>No</td>
<td>None</td>
<td>HyperFlex</td>
</tr>
<tr>
<td>&lt;&lt;hx-vm-data&gt;&gt;</td>
<td>&lt;user_defined&gt;</td>
<td>LAN</td>
<td>Ether</td>
<td>No</td>
<td>None</td>
<td>HyperFlex</td>
</tr>
<tr>
<td>&lt;&lt;hx-vmotion&gt;&gt;</td>
<td>&lt;user_defined&gt;</td>
<td>LAN</td>
<td>Ether</td>
<td>No</td>
<td>None</td>
<td>HyperFlex</td>
</tr>
</tbody>
</table>
Management IP Address Pool

A Cisco UCS Management IP Address Pool must be populated with a block of IP addresses. These IP addresses are assigned to the Cisco Integrated Management Controller (CIMC) interface of the rack-mount and blade servers that are managed in the UCS domain. The IP addresses are the communication endpoints for various functions, such as remote KVM, virtual media, Serial over LAN (SoL), and Intelligent Platform Management Interface (IPMI) for each rack-mount or blade server. Therefore, a minimum of one IP address per physical server in the domain must be provided. The IP addresses are considered to be an “out-of-band” address, meaning that the communication pathway uses the Fabric Interconnects’ mgmt0 ports, which answer ARP requests for the management addresses. Because of this arrangement, the IP addresses in this pool must be in the same IP subnet as the IP addresses assigned to the Fabric Interconnects’ mgmt0 ports. The default pool, named “ext-mgmt” is populated with a block of IP addresses, a subnet mask, and a default gateway by the HyperFlex installer.

Figure 28  Management IP Address Pool

Figure 29  IP Address Block
MAC Address Pools

One of the core benefits of the Cisco UCS and Virtual Interface Card (VIC) technology is the assignment of the personality of the card through Cisco UCS Service Profiles. The number of virtual NIC (vNIC) interfaces, their VLAN association, MAC addresses, QoS policies and more are all applied dynamically as part of the association process. Media Access Control (MAC) addresses use 6 bytes of data as a unique address to identify the interface on the layer 2 network. All devices are assigned a unique MAC address, which is ultimately used for all data transmission and reception. The Cisco UCS and VIC technology picks a MAC address from a pool of addresses, and assigns it to each vNIC defined in the service profile when that service profile is created.

Best practices mandate that MAC addresses used for Cisco UCS domains use 00:25:B5 as the first three bytes, which is one of the Organizationally Unique Identifiers (OUI) registered to Cisco Systems, Inc. The fourth byte (e.g. 00:25:B5:xx) is specified during the HyperFlex installation. The fifth byte is set automatically by the HyperFlex installer, to correlate to the UCS fabric and the vNIC placement order. Finally, the last byte is incremented according to the number of MAC addresses created in the pool. To avoid overlaps, you must ensure that the first four bytes of the MAC address pools are unique for each HyperFlex system installed in the same layer 2 network, and also different from other Cisco UCS domains which may exist, when you define the values in the HyperFlex installer.

The following table details the MAC Address Pools configured for HyperFlex, and their default assignment to the vNIC templates created.

Table 12  MAC Address Pools

<table>
<thead>
<tr>
<th>Name</th>
<th>Block Start</th>
<th>Size</th>
<th>Assignment Order</th>
<th>Used by vNIC Template:</th>
</tr>
</thead>
<tbody>
<tr>
<td>hv-mgmt-a</td>
<td>00:25:B5:&lt;xx&gt;:A1:01</td>
<td>100</td>
<td>Sequential</td>
<td>hv-mgmt-a</td>
</tr>
<tr>
<td>hv-mgmt-b</td>
<td>00:25:B5:&lt;xx&gt;:B2:01</td>
<td>100</td>
<td>Sequential</td>
<td>hv-mgmt-b</td>
</tr>
<tr>
<td>hv-vmotion-a</td>
<td>00:25:B5:&lt;xx&gt;:A7:01</td>
<td>100</td>
<td>Sequential</td>
<td>hv-vmotion-a</td>
</tr>
<tr>
<td>hv-vmotion-b</td>
<td>00:25:B5:&lt;xx&gt;:B8:01</td>
<td>100</td>
<td>Sequential</td>
<td>hv-vmotion-b</td>
</tr>
<tr>
<td>storage-data-a</td>
<td>00:25:B5:&lt;xx&gt;:A3:01</td>
<td>100</td>
<td>Sequential</td>
<td>storage-data-a</td>
</tr>
<tr>
<td>storage-data-b</td>
<td>00:25:B5:&lt;xx&gt;:B4:01</td>
<td>100</td>
<td>Sequential</td>
<td>storage-data-b</td>
</tr>
<tr>
<td>vm-network-a</td>
<td>00:25:B5:&lt;xx&gt;:A5:01</td>
<td>100</td>
<td>Sequential</td>
<td>vm-network-a</td>
</tr>
<tr>
<td>vm-network-b</td>
<td>00:25:B5:&lt;xx&gt;:B6:01</td>
<td>100</td>
<td>Sequential</td>
<td>vm-network-b</td>
</tr>
</tbody>
</table>
Cisco UCS Network Control Policies control various aspects of the behavior of vNICs defined in the Cisco UCS Service Profiles. Settings controlled include enablement of Cisco Discovery Protocol (CDP), MAC address registration, MAC address forging, and the action taken on the vNIC status if the UCS network uplinks are failed. Two policies are configured by the HyperFlex Installer, HyperFlex-infra is applied to the hyperFlex-vm, which is only applied to the vNIC interfaces carrying guest VM traffic. This allows for more flexibility, even though the policies are currently configured with the same settings. The following table details the Network Control Policies configured for HyperFlex, and their default assignment to the vNIC templates created:

**Table 13  Network Control Policy**

<table>
<thead>
<tr>
<th>Name</th>
<th>CDP</th>
<th>MAC Register Mode</th>
<th>Action on Uplink Fail</th>
<th>MAC Security</th>
<th>Used by vNIC Template:</th>
</tr>
</thead>
</table>

Network Control Policies

Cisco UCS Network Control Policies control various aspects of the behavior of vNICs defined in the Cisco UCS Service Profiles. Settings controlled include enablement of Cisco Discovery Protocol (CDP), MAC address registration, MAC address forging, and the action taken on the vNIC status if the UCS network uplinks are failed. Two policies are configured by the HyperFlex Installer, HyperFlex-infra is applied to the “infrastructure” vNIC interfaces of the HyperFlex system, and HyperFlex-vm, which is only applied to the vNIC interfaces carrying guest VM traffic. This allows for more flexibility, even though the policies are currently configured with the same settings. The following table details the Network Control Policies configured for HyperFlex, and their default assignment to the vNIC templates created:
<table>
<thead>
<tr>
<th>Name</th>
<th>CDP</th>
<th>MAC Register Mode</th>
<th>Action on Uplink Fail</th>
<th>MAC Security</th>
<th>Used by vNIC Template</th>
</tr>
</thead>
<tbody>
<tr>
<td>HyperFlex-infra</td>
<td>Enabled</td>
<td>Only Native VLAN</td>
<td>Link-down</td>
<td>Forged: Allow</td>
<td>hv-mgmt-a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hv-mgmt-b</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hv-vmotion-a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hv-vmotion-b</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>storage-data-a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>storage-data-b</td>
</tr>
<tr>
<td>HyperFlex-vm</td>
<td>Enabled</td>
<td>Only Native VLAN</td>
<td>Link-down</td>
<td>Forged: Allow</td>
<td>vm-network-a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>vm-network-b</td>
</tr>
</tbody>
</table>

Figure 31  Network Control Policy

vNIC Templates

Cisco UCS Manager has a feature to configure vNIC templates, which can be used to simplify and speed up configuration efforts. VNIC templates are referenced in service profiles and LAN connectivity policies, versus configuring the same vNICs individually in each service profile, or service profile template. VNIC templates contain all the configuration elements that make up a vNIC, including VLAN assignment, MAC address pool selection, fabric A or B assignment, fabric failover, MTU, QoS policy, Network Control Policy, and more.
Templates are created as either initial templates, or updating templates. Updating templates retain a link between the parent template and the child object, therefore when changes are made to the template, the changes are propagated to all remaining linked child objects. The following tables detail the settings in each of the vNIC templates created by the HyperFlex installer:

**Table 14  vNIC Template hv-mgmt-a**

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>vNIC Template Name:</td>
<td>hv-mgmt-a</td>
</tr>
<tr>
<td>Fabric ID</td>
<td>A</td>
</tr>
<tr>
<td>Fabric Failover</td>
<td>Disabled</td>
</tr>
<tr>
<td>Target</td>
<td>Adapter</td>
</tr>
<tr>
<td>Type</td>
<td>Updating Template</td>
</tr>
<tr>
<td>MTU</td>
<td>1500</td>
</tr>
<tr>
<td>MAC Pool</td>
<td>hv-mgmt-a</td>
</tr>
<tr>
<td>QoS Policy</td>
<td>silver</td>
</tr>
<tr>
<td>Network Control Policy</td>
<td>HyperFlex-infra</td>
</tr>
<tr>
<td>VLANs</td>
<td>&lt;&lt;hx-inband-mgmt&gt;&gt;</td>
</tr>
</tbody>
</table>

**Table 15  vNIC Template hv-mgmt-b**

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>vNIC Template Name:</td>
<td>hv-mgmt-b</td>
</tr>
<tr>
<td>Fabric ID</td>
<td>B</td>
</tr>
<tr>
<td>Fabric Failover</td>
<td>Disabled</td>
</tr>
<tr>
<td>Target</td>
<td>Adapter</td>
</tr>
<tr>
<td>Type</td>
<td>Updating Template</td>
</tr>
<tr>
<td>Setting</td>
<td>Value</td>
</tr>
<tr>
<td>------------------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>MTU</td>
<td>1500</td>
</tr>
<tr>
<td>MAC Pool</td>
<td>hv-mgmt-b</td>
</tr>
<tr>
<td>QoS Policy</td>
<td>silver</td>
</tr>
<tr>
<td>Network Control Policy</td>
<td>HyperFlex-infra</td>
</tr>
<tr>
<td>VLANs</td>
<td>&lt;&lt;hx-inband-mgmt&gt;&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 16  vNIC Template hv-vmotion-a</th>
</tr>
</thead>
<tbody>
<tr>
<td>vNIC Template Name: hv-vmotion-a</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabric ID</td>
<td>A</td>
</tr>
<tr>
<td>Fabric Failover</td>
<td>Disabled</td>
</tr>
<tr>
<td>Target</td>
<td>Adapter</td>
</tr>
<tr>
<td>Type</td>
<td>Updating Template</td>
</tr>
<tr>
<td>MTU</td>
<td>9000</td>
</tr>
<tr>
<td>MAC Pool</td>
<td>hv-vmotion-a</td>
</tr>
<tr>
<td>QoS Policy</td>
<td>bronze</td>
</tr>
<tr>
<td>Network Control Policy</td>
<td>HyperFlex-infra</td>
</tr>
<tr>
<td>VLANs</td>
<td>&lt;&lt;hx-vmotion&gt;&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Figure 32  vNIC Template hx-vmotion-b</th>
</tr>
</thead>
<tbody>
<tr>
<td>vNIC Template Name: hv-vmotion-b</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabric ID</td>
<td>B</td>
</tr>
</tbody>
</table>
### vNIC Template Name: hv-vmotion-b

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabric Failover</td>
<td>Disabled</td>
</tr>
<tr>
<td>Target</td>
<td>Adapter</td>
</tr>
<tr>
<td>Type</td>
<td>Updating Template</td>
</tr>
<tr>
<td>MTU</td>
<td>9000</td>
</tr>
<tr>
<td>MAC Pool</td>
<td>hv-vmotion-b</td>
</tr>
<tr>
<td>QoS Policy</td>
<td>bronze</td>
</tr>
<tr>
<td>Network Control Policy</td>
<td>HyperFlex-infra</td>
</tr>
</tbody>
</table>
| VLANs            | <<hx-vmotion>>                | Native: No

### vNIC Template storage-data-a

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>vNIC Template Name:</td>
<td>storage-data-a</td>
</tr>
<tr>
<td>Fabric ID</td>
<td>A</td>
</tr>
<tr>
<td>Fabric Failover</td>
<td>Disabled</td>
</tr>
<tr>
<td>Target</td>
<td>Adapter</td>
</tr>
<tr>
<td>Type</td>
<td>Updating Template</td>
</tr>
<tr>
<td>MTU</td>
<td>9000</td>
</tr>
<tr>
<td>MAC Pool</td>
<td>storage-data-a</td>
</tr>
<tr>
<td>QoS Policy</td>
<td>platinum</td>
</tr>
<tr>
<td>Network Control Policy</td>
<td>HyperFlex-infra</td>
</tr>
</tbody>
</table>
| VLANs            | <<hx-storage-data>>           | Native: No

**Figure 33 vNIC Template storage-data-a**
### Figure 34 vNIC Template storage-data-b

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>vNIC Template Name</td>
<td>storage-data-b</td>
</tr>
<tr>
<td>Fabric ID</td>
<td>B</td>
</tr>
<tr>
<td>Fabric Failover</td>
<td>Disabled</td>
</tr>
<tr>
<td>Target</td>
<td>Adapter</td>
</tr>
<tr>
<td>Type</td>
<td>Updating Template</td>
</tr>
<tr>
<td>MTU</td>
<td>9000</td>
</tr>
<tr>
<td>MAC Pool</td>
<td>storage-data-b</td>
</tr>
<tr>
<td>QoS Policy</td>
<td>platinum</td>
</tr>
<tr>
<td>Network Control Policy</td>
<td>HyperFlex-infra</td>
</tr>
<tr>
<td>VLANs</td>
<td>&lt;&lt;hx-storage-data&gt;&gt;</td>
</tr>
<tr>
<td></td>
<td>Native: No</td>
</tr>
</tbody>
</table>

### Figure 35 vNIC Template vm-network-a

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>vNIC Template Name</td>
<td>vm-network-a</td>
</tr>
<tr>
<td>Fabric ID</td>
<td>A</td>
</tr>
<tr>
<td>Fabric Failover</td>
<td>Disabled</td>
</tr>
<tr>
<td>Target</td>
<td>Adapter</td>
</tr>
<tr>
<td>Type</td>
<td>Updating Template</td>
</tr>
<tr>
<td>MTU</td>
<td>1500</td>
</tr>
</tbody>
</table>
**vNIC Template Name:** vm-network-a

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabric ID</td>
<td>B</td>
</tr>
<tr>
<td>Fabric Failover</td>
<td>Disabled</td>
</tr>
<tr>
<td>Target</td>
<td>Adapter</td>
</tr>
<tr>
<td>Type</td>
<td>Updating Template</td>
</tr>
<tr>
<td>MTU</td>
<td>1500</td>
</tr>
<tr>
<td>MAC Pool</td>
<td>vm-network-a</td>
</tr>
<tr>
<td>QoS Policy</td>
<td>gold</td>
</tr>
<tr>
<td>Network Control Policy</td>
<td>HyperFlex-vm</td>
</tr>
<tr>
<td>VLANs</td>
<td>&lt;&lt;hx-vm-data&gt;&gt;</td>
</tr>
</tbody>
</table>

**vNIC Template Name:** vm-network-b

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabric ID</td>
<td></td>
</tr>
<tr>
<td>Fabric Failover</td>
<td></td>
</tr>
<tr>
<td>Target</td>
<td></td>
</tr>
<tr>
<td>Type</td>
<td></td>
</tr>
<tr>
<td>MTU</td>
<td></td>
</tr>
<tr>
<td>MAC Pool</td>
<td></td>
</tr>
<tr>
<td>QoS Policy</td>
<td></td>
</tr>
<tr>
<td>Network Control Policy</td>
<td></td>
</tr>
<tr>
<td>VLANs</td>
<td></td>
</tr>
</tbody>
</table>

**LAN Connectivity Policies**

UCS Manager has a feature for LAN Connectivity Policies, which aggregates all of the vNICs or vNIC templates desired for a service profile configuration into a single policy definition. This simplifies configuration efforts by defining a collection of vNICs or vNIC templates once, and using that policy in the service profiles or service profile templates. The HyperFlex installer configures a LAN Connectivity Policy named HyperFlex, which contains all of the vNIC templates defined in the previous section, along with an Adapter Policy named HyperFlex, also configured by the HyperFlex installer. The following table details the LAN Connectivity Policy configured for HyperFlex:
### Cisco UCS Server Policies

**Adapter Policies**

Cisco UCS Adapter Policies are used to configure various settings of the Converged Network Adapter (CNA) installed in the Cisco UCS blade or rack-mount servers. Various advanced hardware features can be enabled or disabled depending on the software or operating system being used. The following figures detail the LAN Connectivity Policy configured for HyperFlex:

<table>
<thead>
<tr>
<th>Policy Name</th>
<th>Use vNIC Template</th>
<th>vNIC Name</th>
<th>vNIC Template Used:</th>
<th>Adapter Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>HyperFlex</td>
<td>Yes</td>
<td>hv-mgmt-a</td>
<td>hv-mgmt-a</td>
<td>HyperFlex</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hv-mgmt-b</td>
<td>hv-mgmt-b</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>hv-vmotion-a</td>
<td>hv-vmotion-a</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>hv-vmotion-b</td>
<td>hv-vmotion-b</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>storage-data-a</td>
<td>storage-data-a</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>storage-data-b</td>
<td>storage-data-b</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>vm-network-a</td>
<td>vm-network-a</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>vm-network-b</td>
<td>vm-network-b</td>
<td></td>
</tr>
</tbody>
</table>
Figure 37  Cisco UCS Adapter Policy Resources

Figure 38  Cisco UCS Adapter Policy Options
**BIOS Policies**

Cisco HX-Series servers have a set of pre-defined BIOS setting defaults defined in Cisco UCS Manager. These settings have been optimized for the Cisco HX-Series servers running HyperFlex. The HyperFlex installer creates a BIOS policy named “HyperFlex”, with all settings set to the defaults, except for enabling the Serial Port A for Serial over LAN (SoL) functionality. This policy allows for future flexibility in case situations arise where the settings need to be modified from the default configuration.

**Boot Policies**

Cisco UCS Boot Policies define the boot devices used by blade and rack-mount servers, and the order that they are attempted to boot from. Cisco HX-Series rack-mount servers and compute-only B200-M4 blade servers have their VMware ESXi hypervisors installed to an internal pair of mirrored Cisco FlexFlash SD cards, therefore they require a boot policy defining that the servers should boot from that location. The HyperFlex installer configures a boot policy named “HyperFlex” which specifies boot from SD card. The following figure details the HyperFlex Boot Policy configured to boot from SD card:

![Cisco UCS Boot Policy](image)

**Host Firmware Packages**

Cisco UCS Host Firmware Packages represent one of the most powerful features of the UCS platform; the ability to control the firmware revision of all the managed blades and rack-mount servers through a policy specified in the service profile. Host Firmware Packages are defined and referenced in the service profiles. Once a service profile is associated to a server, the firmware of all the components defined in the Host Firmware Package are automatically upgraded or downgraded to match the package. The HyperFlex installer creates a Host Firmware Package named “HyperFlex” which uses the simple package definition method, applying firmware revisions to all components that matches a specific UCS firmware bundle, versus defining the firmware revisions part by part. The following figure details the Host Firmware Package configured by the HyperFlex installer:
Local Disk Configuration Policies

Cisco UCS Local Disk Configuration Policies are used to define the configuration of disks installed locally within each blade or rack-mount server, most often to configure Redundant Array of Independent/Inexpensive Disks (RAID levels) when multiple disks are present for data protection. Since HX-Series converged nodes providing storage resources do not require RAID, the HyperFlex installer creates a Local Disk Configuration Policy named “HyperFlex” which allows any local disk configuration. The policy also defines settings for the embedded FlexFlash SD cards used to boot the VMware ESXi hypervisor. The following figure details the Local Disk Configuration Policy configured by the HyperFlex installer:

Maintenance Policies

Cisco UCS Maintenance Policies define the behavior of the attached blades and rack-mount servers when changes are made to the associated service profiles. The default Cisco UCS Maintenance Policy setting is “Immediate” meaning that any change to a service profile that requires a reboot of the physical server will result in an immediate reboot of that server. The Cisco best practice is to use a Maintenance Policy set to “user-ack”, which requires a secondary acknowledgement by a user with the appropriate rights within Cisco UCS, before the server is rebooted to apply the changes. The HyperFlex installer creates a Maintenance Policy named “HyperFlex” with the setting changed to “user-ack”. The following figure details the Maintenance Policy configured by the HyperFlex installer:
Power Control Policies

Cisco UCS Power Control Policies allow administrators to set priority values for power application to servers in environments where power supply may be limited, during times when the servers demand more power than is available. The HyperFlex installer creates a Power Control Policy named “HyperFlex” with all power capping disabled, and fans allowed to run at full speed when necessary. The following figure details the Power Control Policy configured by the HyperFlex installer:

Scrub Policies

Cisco UCS Scrub Policies are used to scrub, or erase data from local disks, BIOS settings and FlexFlash SD cards. If the policy settings are enabled, the information is wiped when the service profile using the policy is disassociated from the server. The HyperFlex installer creates a Scrub Policy named “HyperFlex” which has all settings disabled, therefore all data on local disks, SD cards and BIOS settings will be preserved if a service profile is disassociated. The following figure details the Scrub Policy configured by the HyperFlex installer:
Cisco UCS Serial over LAN (SoL) Policies enable console output which is sent to the serial port of the server, to be accessible through the LAN. For many Linux based operating systems, such as VMware ESXi, the local serial port can be configured as a local console, where users can watch the system boot, and communicate with the system command prompt interactively. Since many blade servers do not have physical serial ports, and often administrators are working remotely, the ability to send and receive that traffic through the LAN is very helpful. Connections to an SoL session can be initiated from Cisco UCS Manager. The HyperFlex installer creates a SoL named “HyperFlex” to enable SoL sessions. The following figure details the SoL Policy configured by the HyperFlex installer:

vMedia Policies

Cisco UCS Virtual Media (vMedia) Policies automate the connection of virtual media files to the remote KVM session of the Cisco UCS blades and rack-mount servers. Using a vMedia policy can speed up installation time by automatically attaching an installation ISO file to the server, without having to manually launch the remote KVM console and connect them one-by-one. The HyperFlex installer creates a vMedia Policy named “HyperFlex” for future use, with no media locations defined. The following figure details the vMedia Policy configured by the HyperFlex installer:
Cisco UCS Service Profile Templates

Cisco UCS Manager has a feature to configure service profile templates, which can be used to simplify and speed up configuration efforts when the same configuration needs to be applied to multiple servers. Service profile templates are used to spawn multiple service profile copies to associate with the servers, versus configuring the same service profile manually each time it is needed. Service profile templates contain all the configuration elements that make up a service profile, including vNICs, vHBAs, local disk configurations, boot policies, host firmware packages, BIOS policies and more. Templates are created as either initial templates, or updating templates. Updating templates retain a link between the parent template and the child object, therefore when changes are made to the template, the changes are propagated to all remaining linked child objects. The HyperFlex installer creates two service profile templates, named “hx-nodes” and “compute-nodes”, each with the same configuration. This simplifies future efforts, if the configuration of the compute only nodes needs to differ from the configuration of the HyperFlex converged storage nodes. The following table details the service profile templates configured by the HyperFlex installer:

<table>
<thead>
<tr>
<th>Service Profile Template Name:</th>
<th>hx-nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>and</td>
</tr>
<tr>
<td></td>
<td>compute-nodes</td>
</tr>
</tbody>
</table>
### Service Profile Template Name:

hx-nodes
and
compute-nodes

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>UUID Pool</td>
<td>Hardware Default</td>
</tr>
<tr>
<td>Associated Server Pool</td>
<td>None</td>
</tr>
<tr>
<td>Maintenance Policy</td>
<td>HyperFlex</td>
</tr>
<tr>
<td>Management IP Address Policy</td>
<td>None</td>
</tr>
<tr>
<td>Local Disk Configuration Policy</td>
<td>HyperFlex</td>
</tr>
<tr>
<td>LAN Connectivity Policy</td>
<td>HyperFlex</td>
</tr>
<tr>
<td>Boot Policy</td>
<td>HyperFlex</td>
</tr>
<tr>
<td>BIOS Policy</td>
<td>HyperFlex</td>
</tr>
<tr>
<td>Firmware Policy</td>
<td>HyperFlex</td>
</tr>
<tr>
<td>Power Control Policy</td>
<td>HyperFlex</td>
</tr>
<tr>
<td>Scrub Policy</td>
<td>HyperFlex</td>
</tr>
<tr>
<td>Serial over LAN Policy</td>
<td>HyperFlex</td>
</tr>
</tbody>
</table>

**vNIC/vHBA Placement**

In order to control the order of detection of the vNICs and vHBAs defined in service profiles, UCS allows for the definition of the placement of the vNICs and vHBAs across the cards in a blade or rack-mount server, and the order they are seen. Since HX-series servers are configured with a single Cisco UCS VIC 1227 mLOM card, the only valid placement is on card number 1. In certain hardware configurations, the physical mapping of cards and port extenders to their logical order is not linear, therefore each card is referred to as a virtual connection, or vCon. Because of this, the interface where the placement and order is defined does not refer to physical cards, but instead refers to vCons. Therefore, all the vNICs defined in the service profile templates for HX-series servers places them on vCon 1, then their order is defined.
Through the combination of the vNIC templates created (vNIC Templates), the LAN Connectivity Policy (LAN Connectivity Policies), and the vNIC placement, every VMware ESXi server will detect the network interfaces in the same order, and they will always be connected to the same VLANs through the same network fabrics. The following table outlines the vNICs, their placement, their order, the fabric they are connected to, their default VLAN, and how they are enumerated by the ESXi hypervisor:

Table 18  vNIC Placement

<table>
<thead>
<tr>
<th>vNIC</th>
<th>Placement</th>
<th>Order</th>
<th>Fabric</th>
<th>VLAN</th>
<th>ESXi interface enumeration</th>
</tr>
</thead>
<tbody>
<tr>
<td>hv-mgmt-a</td>
<td>1</td>
<td>1</td>
<td>A</td>
<td>&lt;&lt;hx-inband-mgmt&gt;&gt;</td>
<td>vmnic0</td>
</tr>
<tr>
<td>hv-mgmt-b</td>
<td>1</td>
<td>2</td>
<td>B</td>
<td>&lt;&lt;hx-inband-mgmt&gt;&gt;</td>
<td>vmnic1</td>
</tr>
<tr>
<td>storage-data-a</td>
<td>1</td>
<td>3</td>
<td>A</td>
<td>&lt;&lt;hx-storage-data&gt;&gt;</td>
<td>vmnic2</td>
</tr>
<tr>
<td>storage-data-b</td>
<td>1</td>
<td>4</td>
<td>B</td>
<td>&lt;&lt;hx-storage-data&gt;&gt;</td>
<td>vmnic3</td>
</tr>
<tr>
<td>vm-network-a</td>
<td>1</td>
<td>5</td>
<td>A</td>
<td>&lt;&lt;hx-vm-data&gt;&gt;</td>
<td>vmnic4</td>
</tr>
<tr>
<td>vm-network-b</td>
<td>1</td>
<td>6</td>
<td>B</td>
<td>&lt;&lt;hx-vm-data&gt;&gt;</td>
<td>vmnic5</td>
</tr>
<tr>
<td>hv-vmotion-a</td>
<td>1</td>
<td>7</td>
<td>A</td>
<td>&lt;&lt;hx-vmotion&gt;&gt;</td>
<td>vmnic6</td>
</tr>
<tr>
<td>hv-vmotion-b</td>
<td>1</td>
<td>8</td>
<td>B</td>
<td>&lt;&lt;hx-vmotion&gt;&gt;</td>
<td>vmnic7</td>
</tr>
</tbody>
</table>

ESXi VMDirectPath relies on a fixed PCI address for the pass-through devices. If the vNIC configuration is changed (add/remove vNICs), then the order of the devices seen in the PCI tree will change. The administrator will have to reconfigure the ESXi VMDirectPath configuration to select the 12 Gbps SAS HBA card, and reconfigure the storage controller settings of the controller VM.
ESXi Host Design

The following sections detail the design of the elements within the VMware ESXi hypervisors, system requirements, virtual networking and the configuration of ESXi for the Cisco HyperFlex HX Distributed Data Platform.

Virtual Networking Design

The Cisco HyperFlex system has a pre-defined virtual network design at the ESXi hypervisor level. Four different virtual switches are created by the HyperFlex installer, each using two uplinks, which are each serviced by a vNIC defined in the UCS service profile. The vSwitches created are:

- **vswitch-hx-inband-mgmt**: This is the default vSwitch0 which is renamed by the ESXi kickstart file as part of the automated installation. The default vmkernel port, vmk0, is configured in the standard Management Network port group. The switch has two uplinks, active on fabric A and standby on fabric B, without jumbo frames. A second port group is created for the Storage Platform Controller VMs to connect to with their individual management interfaces. The VLAN is not a Native VLAN as assigned to the vNIC template, and therefore assigned in ESXi/vSphere.

- **vswitch-hx-storage-data**: This vSwitch is created as part of the automated installation. A vmkernel port, vmk1, is configured in the Storage Hypervisor Data Network port group, which is the interface used for connectivity to the HX Datastores through NFS. The switch has two uplinks, active on fabric B and standby on fabric A, with jumbo frames required. A second port group is created for the Storage Platform Controller VMs to connect to with their individual storage interfaces. The VLAN is not a Native VLAN as assigned to the vNIC template, and therefore assigned in ESXi/vSphere.

- **vswitch-hx-vm-network**: This vSwitch is created as part of the automated installation. The switch has two uplinks, active on both fabrics A and B, and without jumbo frames. The VLAN is not a Native VLAN as assigned to the vNIC template, and therefore assigned in ESXi/vSphere.

- **vmotion**: This vSwitch is created as part of the automated installation. The switch has two uplinks, active on fabric A and standby on fabric B, with jumbo frames required. The VLAN is not a Native VLAN as assigned to the vNIC template, and therefore assigned in ESXi/vSphere.

The following table and figures help give more details into the ESXi virtual networking design as built by the HyperFlex installer:

**Table 19 Virtual Switches**

<table>
<thead>
<tr>
<th>Virtual Switch</th>
<th>Port Groups</th>
<th>Active vmnic(s)</th>
<th>Passive vmnic(s)</th>
<th>VLAN IDs</th>
<th>Jumbo</th>
</tr>
</thead>
<tbody>
<tr>
<td>vswitch-hx-inband-mgmt</td>
<td>Management Network</td>
<td>vmnic0</td>
<td>vmnic1</td>
<td>hx-inband-mgmt</td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>Storage Controller Management Network</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
VMDirectPath I/O Pass-through

VMDirectPath I/O allows a guest VM to directly access PCI and PCIe devices in an ESXi host as though they were physical devices belonging to the VM itself, also referred to as PCI pass-through. With the appropriate driver for the hardware device, the guest VM sends all I/O requests directly to the physical device, bypassing the hypervisor. In the Cisco HyperFlex system, the Storage Platform Controller VMs use this feature to gain full control of the Cisco 12Gbps SAS HBA cards in the Cisco HX-series rack-mount servers. This gives the controller VMs direct hardware level access to the physical disks installed in the servers, which they
consume to construct the Cisco HX Distributed Filesystem. Only the disks connected directly to the Cisco SAS HBA or to a SAS extender, in turn connected to the SAS HBA are controlled by the controller VMs. Other disks, connected to different controllers, such as the SD cards, remain under the control of the ESXi hypervisor. The configuration of the VMDirectPath I/O feature is done by the Cisco HyperFlex installer, and requires no manual steps.

Storage Platform Controller VMs

A key component of the Cisco HyperFlex system is the Storage Platform Controller Virtual Machine running on each of the nodes in the HyperFlex cluster. The controller VMs cooperate to form and coordinate the Cisco HX Distributed Filesystem, and service all the guest VM IO requests. The controller VMs are deployed as a vSphere ESXi agent, which is similar in concept to that of a Linux or Windows service. ESXi agents are tied to a specific host, they start and stop along with the ESXi hypervisor, and the system is not considered to be online and ready until both the hypervisor and the agents have started. Each ESXi hypervisor host has a single ESXi agent deployed, which is the controller VM for that node, and it cannot be moved or migrated to another host. The collective ESXi agents are managed through an ESXi agency in the vSphere cluster.

The storage controller VM runs custom software and services that manage and maintain the Cisco HX Distributed Filesystem. The services and processes that run within the controller VMs are not exposed as part of the ESXi agents to the agency, therefore the ESXi hypervisors nor vCenter server have any direct knowledge of the storage services provided by the controller VMs. Management and visibility into the function of the controller VMs, and the Cisco HX Distributed Filesystem is done through a plugin installed to the vCenter server or appliance managing the vSphere cluster. The plugin communicates directly with the controller VMs to display the information requested, or make the configuration changes directed, all while operating within the same web-based interface of the vSphere Web Client. The deployment of the controller VMs, agents, agency, and vCenter plugin are all done by the Cisco HyperFlex installer, and requires no manual steps.

Controller VM Locations

The physical storage location of the controller VMs differs between the Cisco HX220c-M4S and HX240c-M4SX model servers, due to differences with the physical disk location and connections on the two models of servers. The storage controller VM is operationally no different from any other typical virtual machines in an ESXi environment. The VM must have a virtual disk with the bootable root filesystem available in a location separate from the SAS HBA that the VM is controlling through VMDirectPath I/O. The configuration details of the models are as follows:

- **HX220c**: The controller VM’s root filesystem is stored on a 2.2 GB virtual disk, /dev/sda, which is placed on a 3.5 GB VMFS datastore, and that datastore is provisioned from the internal mirrored SD cards. The controller VM has full control of all the front facing hot-swappable disks through PCI pass-through control of the SAS HBA. The controller VM operating system sees the 120 GB SSD, also commonly called the “housekeeping” disk as /dev/sdb, and places HyperFlex binaries, logs, and zookeeper partitions on this disk. The remaining disks seen by the controller VM OS are used by the HX Distributed filesystem for caching and capacity layers.

- **HX240c**: The HX240c-M4SX server has a built-in SATA controller provided by the Intel Wellsburg Platform Controller Hub (PCH) chip, and the 120 GB housekeeping disk is connected to it, placed in an internal drive carrier. Since this model does not connect the 120 GB housekeeping disk to the SAS HBA, the ESXi hypervisor remains in control of this disk, and a VMFS datastore is provisioned there,
using the entire disk. On this VMFS datastore, a 2.2 GB virtual disk is created and used by the controller VM as /dev/sda for the root filesystem, and an 87 GB virtual disk is created and used by the controller VM as /dev/sdb, placing the HyperFlex binaries, logs, and zookeeper partitions on this disk. The front-facing hot swappable disks, seen by the controller VM OS through PCI pass-through control of the SAS HBA, are used by the HX Distributed filesystem for caching and capacity layers.

The following figures detail the Storage Platform Controller VM placement on the ESXi hypervisor hosts:

**Figure 49  HX220c Controller VM Placement**

---

Cisco UCS B200-M4 compute-only blades also place a lightweight storage controller VM on a 3.5 GB VMFS datastore, provisioned from the SD cards.
HyperFlex Datastores

The new HyperFlex cluster has no default datastores configured for virtual machine storage, therefore the datastores must be created using the vCenter Web Client plugin. A minimum of two datastores is recommended to satisfy vSphere High Availability datastore heartbeat requirements, although one of the two datastores can be very small. It is important to recognize that all HyperFlex datastores are thinly provisioned, meaning that their configured size can far exceed the actual space available in the HyperFlex cluster. Alerts will be raised by the HyperFlex system in the vCenter plugin when actual space consumption results in low amounts of free space, and alerts will be sent through auto support email alerts. Overall space consumption in the HyperFlex clustered filesystem is optimized by the default deduplication and compression features.
CPU Resource Reservations

Since the storage controller VMs provide critical functionality of the Cisco HX Distributed Data Platform, the HyperFlex installer will configure CPU resource reservations for the controller VMs. This reservation guarantees that the controller VMs will have CPU resources at a minimum level, in situations where the physical CPU resources of the ESXi hypervisor host are being heavily consumed by the guest VMs. The following table details the CPU resource reservation of the storage controller VMs:

**Table 20  Controller VM CPU Reservations**

<table>
<thead>
<tr>
<th>Number of vCPU</th>
<th>Shares</th>
<th>Reservation</th>
<th>Limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Low</td>
<td>10800 MHz</td>
<td>unlimited</td>
</tr>
</tbody>
</table>

Memory Resource Reservations

Since the storage controller VMs provide critical functionality of the Cisco HX Distributed Data Platform, the HyperFlex installer will configure memory resource reservations for the controller VMs. This reservation guarantees that the controller VMs will have memory resources at a minimum level, in situations where the physical memory resources of the ESXi hypervisor host are being heavily consumed by the guest VMs. The following table details the memory resource reservation of the storage controller VMs:

**Table 21  Controller VM Memory Reservations**

<table>
<thead>
<tr>
<th>Server Model</th>
<th>Amount of Guest Memory</th>
<th>Reserve All Guest Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>HX220c-M4S</td>
<td>48 GB</td>
<td>Yes</td>
</tr>
<tr>
<td>HX240c-M4SX</td>
<td>72 GB</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Cisco UCS B200 M4 compute-only blades have a lightweight storage controller VM, it is configured with only 1 vCPU and 512 MB of memory reservation.
Installation

Cisco HyperFlex systems are ordered with a factory pre-installation process having been done prior to the hardware delivery. This factory integration work will deliver the HyperFlex servers with the proper firmware revisions pre-set, a copy of the VMware ESXi hypervisor software pre-installed, and some components of the Cisco HyperFlex software already installed. Once on site, the final steps to be performed by Cisco Advanced Services or our Cisco Partner companies' technical staff are reduced and simplified due to the previous factory work. For the purpose of this document, the entire setup process is described as though no factory pre-installation work was done, yet still leveraging the tools and processes developed by Cisco to simplify the process and dramatically reduce the deployment time.

Installation of the Cisco HyperFlex system is primarily done through a deployable HyperFlex installer virtual machine, available for download at cisco.com as an OVA file. The installer VM performs most of the Cisco UCS configuration work, and it is used to simplify the installation of ESXi on the HyperFlex hosts. The HyperFlex installer VM also installs the HyperFlex HX Data Platform software and creates the HyperFlex cluster, while concurrently performing many of the ESXi configuration tasks automatically. Because this simplified installation method has been developed by Cisco, this CVD will not give detailed manual steps for the configuration of all the elements that are handled by the installer. The following sections will guide you through the prerequisites and manual steps needed prior to using the HyperFlex installer, how to utilize the HyperFlex Installer, and finally how to perform the remaining post-installation tasks.

Prerequisites

Prior to beginning the installation activities, it is important to gather the following information:

IP Addressing

IP addresses for the Cisco HyperFlex system need to be allocated from the appropriate subnets and VLANs to be used. IP addresses that are used by the system fall into the following groups:

- **Cisco UCS Manager**: These addresses are used and assigned by Cisco UCS Manager. Three IP addresses are used by Cisco UCS Manager, one address is assigned to each Cisco UCS Fabric Interconnect, and the third IP address is a roaming address for managing the active FI of the Cisco UCS cluster. In addition, at least one IP address per UCS blade or HX-series rack-mount server is required for the default ext-mgmt IP address pool, which are assigned to the CIMC interface of the physical servers. Since these management addresses are assigned from a pool, they need to be provided in a contiguous block of addresses. These addresses must all be in the same subnet.

- **HyperFlex and ESXi Management**: These addresses are used to manage the ESXi hypervisor hosts, and the HyperFlex Storage Platform Controller VMs. Two IP addresses per node in the HyperFlex cluster are required from the same subnet, and a single additional IP address is needed as the roaming HyperFlex cluster management interface. These addresses can be assigned from the same subnet at the Cisco UCS Manager addresses, or they may be separate.

- **HyperFlex Storage**: These addresses are used by the HyperFlex Storage Platform Controller VMs, and as vmkernel interfaces on the ESXi hypervisor hosts, for sending and receiving data to/from the HX Distributed Data Platform Filesystem. Two IP addresses per node in the HyperFlex cluster are required from the same subnet, and a single additional IP address is needed as the roaming HyperFlex cluster
storage interface. It is recommended to provision a subnet that is not used in the network for other purposes, and it is also possible to use non-routable IP address ranges for these interfaces. Finally, if the UCS domain is going to contain multiple HyperFlex clusters, it is possible to use a different subnet and VLAN ID for the HyperFlex storage traffic for each cluster. This is a safer method, guaranteeing that storage traffic from multiple cluster cannot intermix.

- **VMotion**: These IP addresses are used by the ESXi hypervisor hosts as vmkernel interfaces to enable vMotion capabilities. One or more IP addresses per node in the HyperFlex cluster are required from the same subnet. Multiple addresses and vmkernel interfaces can be used if you wish to enable multi-nic vMotion.

The following tables will assist with gathering the required IP addresses for the installation of an 8 node standard HyperFlex cluster, or a 4+4 hybrid cluster, by listing the addresses required, and an example configuration that will be used during the setup steps in this CVD:

### Table 22  HyperFlex Cluster IP Addressing

<table>
<thead>
<tr>
<th>Network Group</th>
<th>UCS Management</th>
<th>HyperFlex and ESXi Management</th>
<th>HyperFlex Storage</th>
<th>VMotion</th>
</tr>
</thead>
<tbody>
<tr>
<td>VLAN ID:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subnet:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subnet Mask:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gateway:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Device</th>
<th>UCS Manager Addresses</th>
<th>ESXi Management Interface</th>
<th>Storage Controller Management Interface</th>
<th>ESXi Hypervisor Storage vmkernel Interface</th>
<th>Storage Controller Storage Interface</th>
<th>VMotion vmkernel Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabric Interconnect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fabric Interconnect B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS Manager</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Cluster</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 23  HyperFlex Hybrid Cluster IP Addressing

<table>
<thead>
<tr>
<th>Network Group</th>
<th>UCS Management</th>
<th>HyperFlex and ESXi Management</th>
<th>HyperFlex Storage</th>
<th>VMotion</th>
</tr>
</thead>
<tbody>
<tr>
<td>VLAN ID:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subnet:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subnet Mask:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gateway:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Device</th>
<th>UCS Manager Addresses</th>
<th>ESXi Management Interface</th>
<th>Storage Controller Management Interface</th>
<th>ESXi Hypervisor Storage vmkernel Interface</th>
<th>Storage Controller Storage Interface</th>
<th>VMotion vmkernel Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabric Interconnect A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fabric Interconnect B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS Manager</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Cluster</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Node #8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network Group</td>
<td>UCS Management</td>
<td>HyperFlex and ESXi Management</td>
<td>HyperFlex Storage</td>
<td>VMotion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>------------------------------</td>
<td>------------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VLAN ID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subnet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subnet Mask</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gateway</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Device</th>
<th>UCS Management Address</th>
<th>ESXi Management Interface</th>
<th>Storage Controller Management Interface</th>
<th>ESXi Hypervisor Storage vmkernel Interface</th>
<th>Storage Controller Storage Interface</th>
<th>VMotion vmkernel Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabric Interconnect A</td>
<td>10.29.151.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fabric Interconnect B</td>
<td>10.29.151.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS Manager</td>
<td>10.29.151.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Cluster</td>
<td>10.29.151.189</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 24 HyperFlex Cluster Example IP Addressing**

<table>
<thead>
<tr>
<th>Network Group</th>
<th>UCS Management</th>
<th>HyperFlex and ESXi Management</th>
<th>HyperFlex Storage</th>
<th>VMotion</th>
</tr>
</thead>
<tbody>
<tr>
<td>VLAN ID</td>
<td>3175</td>
<td>3175</td>
<td>3172</td>
<td>3173</td>
</tr>
<tr>
<td>Subnet</td>
<td>10.29.151.0</td>
<td>10.29.151.0</td>
<td>172.17.72.0</td>
<td>172.17.73.0</td>
</tr>
<tr>
<td>Subnet Mask</td>
<td>255.255.255.0</td>
<td>255.255.255.0</td>
<td>255.255.255.0</td>
<td>255.255.255.0</td>
</tr>
<tr>
<td>Gateway</td>
<td>10.29.151.1</td>
<td>10.29.151.1</td>
<td>172.17.72.1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Device</th>
<th>UCS Management Address</th>
<th>ESXi Management Interface</th>
<th>Storage Controller Management Interface</th>
<th>ESXi Hypervisor Storage vmkernel Interface</th>
<th>Storage Controller Storage Interface</th>
<th>VMotion vmkernel Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabric Interconnect A</td>
<td>10.29.151.22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fabric Interconnect B</td>
<td>10.29.151.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS Manager</td>
<td>10.29.151.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HyperFlex Cluster</td>
<td>10.29.151.189</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The Cisco UCS Management, and HyperFlex and ESXi Management IP addresses can come from the same subnet, or be separate, as long as the HyperFlex installer can reach them both.

**DHCP vs Static IP**

By default, the 1.8 HX installation will now assign static IP address to the ESXi servers. Using Dynamic Host Configuration Protocol (DHCP) for automatic IP address assignment in not recommended.

**DNS**

DNS servers must be available to query in the HyperFlex and ESXi Management group. DNS records need to be created prior to beginning the installation. At a minimum, it is highly recommended to create A records for the ESXi hypervisor hosts' management interfaces. Additional A records can be created for the Storage Controller Management interfaces, ESXi Hypervisor Storage interfaces, and the Storage Controller Storage interfaces if desired.

The following tables will assist with gathering the required DNS information for the installation of an 8 node standard HyperFlex cluster, or a 4+4 hybrid cluster, by listing the information required, and an example configuration that will be used during the setup steps in this CVD:

**Table 25 DNS Server Information**

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNS Server #1</td>
<td></td>
</tr>
<tr>
<td>DNS Server #2</td>
<td></td>
</tr>
<tr>
<td>Item</td>
<td>Value</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>DNS Domain</td>
<td></td>
</tr>
<tr>
<td>VCenter Server Name</td>
<td></td>
</tr>
<tr>
<td>SMTP Server Name</td>
<td></td>
</tr>
<tr>
<td>UCS Domain Name</td>
<td></td>
</tr>
<tr>
<td>HX Server #1 Name</td>
<td></td>
</tr>
<tr>
<td>HX Server #2 Name</td>
<td></td>
</tr>
<tr>
<td>HX Server #3 Name</td>
<td></td>
</tr>
<tr>
<td>HX Server #4 Name</td>
<td></td>
</tr>
<tr>
<td>HX Server #5 Name</td>
<td></td>
</tr>
<tr>
<td>HX Server #6 Name</td>
<td></td>
</tr>
<tr>
<td>HX Server #7 Name</td>
<td></td>
</tr>
<tr>
<td>HX Server #8 Name</td>
<td></td>
</tr>
</tbody>
</table>

**Table 26  DNS Server Example Information**

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNS Server #1</td>
<td>10.29.130.112</td>
</tr>
<tr>
<td>DNS Server #2</td>
<td></td>
</tr>
<tr>
<td>DNS Domain</td>
<td>ppt.lab.cisco.com</td>
</tr>
<tr>
<td>VCenter Server Name</td>
<td>vc.ppt.lab.cisco.com</td>
</tr>
<tr>
<td>SMTP Server Name</td>
<td>10.29.130.115</td>
</tr>
<tr>
<td>UCS Domain Name</td>
<td>10.29.151.24</td>
</tr>
</tbody>
</table>
### Table 27  NTP Server Information

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HX Server #1</td>
<td>c220-one.ppt.lab.cisco.com</td>
</tr>
<tr>
<td>HX Server #2</td>
<td>c220-two.ppt.lab.cisco.com</td>
</tr>
<tr>
<td>HX Server #3</td>
<td>c220-three.ppt.lab.cisco.com</td>
</tr>
<tr>
<td>HX Server #4</td>
<td>c220-four.ppt.lab.cisco.com</td>
</tr>
<tr>
<td>HX Server #5</td>
<td>c220-five.ppt.lab.cisco.com</td>
</tr>
<tr>
<td>HX Server #6</td>
<td>c220-six.ppt.lab.cisco.com</td>
</tr>
<tr>
<td>HX Server #7</td>
<td>c220-seven.ppt.lab.cisco.com</td>
</tr>
<tr>
<td>HX Server #8</td>
<td>c220-eight.ppt.lab.cisco.com</td>
</tr>
</tbody>
</table>

### NTP

Consistent time is required across the components of the HyperFlex system, provided by reliable NTP servers, accessible for synchronization in the Cisco UCS Management network group, and the HyperFlex and ESXi Management group. NTP is used by Cisco UCS Manager, the ESXi hypervisor hosts, and the HyperFlex Storage Platform Controller VMs.

The following tables will assist with gathering the required NTP information for the installation of an 8 node standard HyperFlex cluster, or a 4+4 hybrid cluster, by listing the information required, and an example configuration that will be used during the setup steps in this CVD:

### Table 28  NTP Server Example Information

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTP Server #1</td>
<td></td>
</tr>
<tr>
<td>NTP Server #2</td>
<td></td>
</tr>
<tr>
<td>Timezone</td>
<td></td>
</tr>
</tbody>
</table>
Installation

NTP

NTP Server #1 10.29.130.112
NTP Server #2
Timezone (UTC-8:00) Pacific Time

VLANs

Prior to the installation, the required VLAN IDs need to be documented, and created in the upstream network if necessary. At a minimum there are 4 VLANs that need to be trunked to the UCS Fabric Interconnects that comprise the HyperFlex system; a VLAN for the HyperFlex and ESXi Management group, a VLAN for the HyperFlex Storage group, a VLAN for the VMotion group, and at least one VLAN for the guest VM traffic. The VLAN IDs must be supplied during the HyperFlex UCS configuration step, and the VLAN names can optionally be customized.

The following tables will assist with gathering the required VLAN information for the installation of an 8 node standard HyperFlex cluster, or a 4+4 hybrid cluster, by listing the information required, and an example configuration that will be used during the setup steps in this CVD:

**Figure 52 VLAN Information**

<table>
<thead>
<tr>
<th>Name</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;&lt;hx-inband-mgmt&gt;&gt;</td>
<td></td>
</tr>
<tr>
<td>&lt;&lt;hx-storage-data&gt;&gt;</td>
<td></td>
</tr>
<tr>
<td>&lt;&lt;hx-vm-data&gt;&gt;</td>
<td></td>
</tr>
<tr>
<td>&lt;&lt;hx-vmotion&gt;&gt;</td>
<td></td>
</tr>
</tbody>
</table>

**Table 29 VLAN Example Information**

<table>
<thead>
<tr>
<th>Name</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>hx-inband-mgmt</td>
<td>3175</td>
</tr>
<tr>
<td>hx-storage-data</td>
<td>3172</td>
</tr>
<tr>
<td>&lt;&lt;hx-vm-data&gt;&gt;</td>
<td>3174</td>
</tr>
<tr>
<td>hx-vmotion</td>
<td>3173</td>
</tr>
</tbody>
</table>
Network Uplinks

The Cisco UCS uplink connectivity design needs to be finalized prior to beginning the installation. One of the early manual tasks to be completed is to configure the UCS network uplinks and verify their operation, prior to beginning the HyperFlex installation steps. Refer to the network uplink design possibilities in the Network Design section.

The following tables will assist with gathering the required network uplink information for the installation of an 8 node standard HyperFlex cluster, or a 4+4 hybrid cluster, by listing the information required, and an example configuration that will be used during the setup steps in this CVD:

Table 30  Network Uplink Configuration

<table>
<thead>
<tr>
<th>Fabric Interconnect Port</th>
<th>Port Channel</th>
<th>Port Channel Type</th>
<th>Port Channel ID</th>
<th>Port Channel Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>☐ Yes ☐ No</td>
<td>☐ LACP</td>
<td>☐ vPC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ Yes ☐ No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ Yes ☐ No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ Yes ☐ No</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 31  Network Uplink Example Configuration

<table>
<thead>
<tr>
<th>Fabric Interconnect Port</th>
<th>Port Channel</th>
<th>Port Channel Type</th>
<th>Port Channel ID</th>
<th>Port Channel Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1/25</td>
<td>☒ Yes ☐ No</td>
<td>☐ LACP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1/26</td>
<td>☒ Yes ☐ No</td>
<td>☒ vPC</td>
<td>vpc-13-nexus</td>
</tr>
<tr>
<td></td>
<td>☐ Yes ☐ No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ Yes ☐ No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fabric Interconnect Port</td>
<td>Port Channel</td>
<td>Port Channel Type</td>
<td>Port Channel ID</td>
<td>Port Channel Name</td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------------</td>
<td>-------------------</td>
<td>-----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>B</td>
<td>1/25</td>
<td>Yes</td>
<td>LACP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1/26</td>
<td>Yes</td>
<td>vPC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Yes</td>
<td></td>
<td>vpc-14-nexus</td>
</tr>
</tbody>
</table>

**Usernames and Passwords**

Several usernames and passwords need to be defined or known as part of the HyperFlex installation process. The following tables will assist with gathering the required username and password information for the installation of an 8 node standard HyperFlex cluster, or a 4+4 hybrid cluster, by listing the information required, and an example configuration that will be used during the setup steps in this CVD:

**Table 32  Usernames and Passwords**

<table>
<thead>
<tr>
<th>Account</th>
<th>Username</th>
<th>Password</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCS Administrator</td>
<td>admin</td>
<td>&lt;&lt;ucs_admin_pw&gt;&gt;</td>
</tr>
<tr>
<td>ESXi Administrator</td>
<td>root</td>
<td>&lt;&lt;esxi_root_pw&gt;&gt;</td>
</tr>
<tr>
<td>HyperFlex Administrator</td>
<td>root</td>
<td>&lt;&lt;hx_admin_pw&gt;&gt;</td>
</tr>
<tr>
<td>VCenter Administrator</td>
<td>&lt;&lt;vcenter_administrator&gt;&gt;</td>
<td>&lt;&lt;vcenter_admin_pw&gt;&gt;</td>
</tr>
</tbody>
</table>

**Table 33  Example Usernames and Passwords**

<table>
<thead>
<tr>
<th>Account</th>
<th>Username</th>
<th>Password</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCS Administrator</td>
<td>admin</td>
<td>1q2w3e4r</td>
</tr>
<tr>
<td>ESXi Administrator</td>
<td>root</td>
<td>Cisco123</td>
</tr>
<tr>
<td>HyperFlex Administrator</td>
<td>root</td>
<td>1q2w3e4r</td>
</tr>
<tr>
<td>VCenter Administrator</td>
<td><a href="mailto:administrator@vsphere.local">administrator@vsphere.local</a></td>
<td>!QAZ2wsx</td>
</tr>
</tbody>
</table>
Physical Installation

Install the Fabric Interconnects, the HX-Series rack-mount servers, the Cisco UCS 5108 chassis, the Cisco UCS 2204XP Fabric Extenders, and the Cisco UCS B200-M4 blades according to their corresponding hardware installation guides:

Cisco UCS 6248UP Fabric Interconnect:

HX220c-M4S Server:

HX240c-M4SX Server:

Cisco UCS 5108 Chassis, Servers and Fabric Extenders:

Cabling

The physical layout of the HyperFlex system was previously described in section Physical Topology. The Fabric Interconnects, HX-series rack-mount servers, Cisco UCS chassis and blades need to be cabled properly before beginning the installation activities.

The following table provides an example cabling map for installation of a Cisco HyperFlex system, with eight HX220c-M4SX servers, and one Cisco UCS 5108 chassis.

<table>
<thead>
<tr>
<th>Device</th>
<th>Port</th>
<th>Connected To</th>
<th>Port</th>
<th>Type</th>
<th>Length</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCS6248-A</td>
<td>L1</td>
<td>UCS6248-B</td>
<td>L1</td>
<td>CAT5</td>
<td>1FT</td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>L2</td>
<td>UCS6248-B</td>
<td>L2</td>
<td>CAT5</td>
<td>1FT</td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>mgmt0</td>
<td>Customer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/1</td>
<td>HX Server #1 Name</td>
<td>mLOM port 1</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 1</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/2</td>
<td>HX Server #2 Name</td>
<td>mLOM port 1</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 2</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/3</td>
<td>HX Server #3 Name</td>
<td>mLOM port 1</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 3</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/4</td>
<td>HX Server #4 Name</td>
<td>mLOM port 1</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 4</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/5</td>
<td>HX Server #5 Name</td>
<td>mLOM port 1</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 5</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/6</td>
<td>HX Server #6 Name</td>
<td>mLOM port 1</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 6</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/7</td>
<td>HX Server #7 Name</td>
<td>mLOM port 1</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 7</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/8</td>
<td>HX Server #8 Name</td>
<td>mLOM port 1</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 8</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/9</td>
<td>2204XP #1</td>
<td>IOM1 port 1</td>
<td>Twinax</td>
<td>3M</td>
<td>Chassis 1</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/10</td>
<td>2204XP #1</td>
<td>IOM1 port 2</td>
<td>Twinax</td>
<td>3M</td>
<td>Chassis 1</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/11</td>
<td>2204XP #1</td>
<td>IOM1 port 3</td>
<td>Twinax</td>
<td>3M</td>
<td>Chassis 1</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/12</td>
<td>2204XP #1</td>
<td>IOM1 port 4</td>
<td>Twinax</td>
<td>3M</td>
<td>Chassis 1</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Device</td>
<td>Port</td>
<td>Connected To</td>
<td>Port</td>
<td>Type</td>
<td>Length</td>
<td>Note</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>------------------</td>
<td>------</td>
<td>------</td>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/25</td>
<td>LAN</td>
<td></td>
<td></td>
<td>uplink</td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/26</td>
<td>LAN</td>
<td></td>
<td></td>
<td>uplink</td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/28</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-A</td>
<td>1/32</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>L1</td>
<td>UCS6248-A</td>
<td>L1</td>
<td>CAT5</td>
<td>1FT</td>
<td></td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>L2</td>
<td>UCS6248-A</td>
<td>L2</td>
<td>CAT5</td>
<td>1FT</td>
<td></td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>mgmt0</td>
<td>Customer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/1</td>
<td>HX Server #1 Name</td>
<td>mLOM port 2</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 1</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/2</td>
<td>HX Server #2 Name</td>
<td>mLOM port 2</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 2</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/3</td>
<td>HX Server #3 Name</td>
<td>mLOM port 2</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 3</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/4</td>
<td>HX Server #4 Name</td>
<td>mLOM port 2</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 4</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/5</td>
<td>HX Server #5 Name</td>
<td>mLOM port 2</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 5</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/6</td>
<td>HX Server #6 Name</td>
<td>mLOM port 2</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 6</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/7</td>
<td>HX Server #7 Name</td>
<td>mLOM port 2</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 7</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/8</td>
<td>HX Server #8 Name</td>
<td>mLOM port 2</td>
<td>Twinax</td>
<td>3M</td>
<td>Server 8</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/9</td>
<td>2204XP #1</td>
<td>IOM2 port 1</td>
<td>Twinax</td>
<td>3M</td>
<td>Chassis 1</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/10</td>
<td>2204XP #1</td>
<td>IOM2 port 2</td>
<td>Twinax</td>
<td>3M</td>
<td>Chassis 1</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/11</td>
<td>2204XP #1</td>
<td>IOM2 port 3</td>
<td>Twinax</td>
<td>3M</td>
<td>Chassis 1</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/12</td>
<td>2204XP #1</td>
<td>IOM2 port 4</td>
<td>Twinax</td>
<td>3M</td>
<td>Chassis 1</td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UCS6248-B</td>
<td>1/20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Cisco UCS Installation

This section describes the steps to initialize and configure the UCS Fabric Interconnects, to prepare them for the HyperFlex installation.

Cisco UCS 6248UP Fabric Interconnect A

To configure Fabric Interconnect A, complete the following steps:

1. Ensure the Fabric Interconnect cabling is properly connected, including the L1 and L2 cluster links, and power the Fabric Interconnects on by inserting the power cords.

2. Connect to the console port on the first Fabric Interconnect, which will be designated as the A fabric device. Use the supplied Cisco console cable (CAB-CONSOLE-RJ45=), and connect it to a built-in DB9 serial port, or use a USB to DB9 serial port adapter.

3. Start your terminal emulator software.

4. Create a connection to the COM port of the computer’s DB9 port, or the USB to serial adapter. Set the terminal emulation to VT100, and the settings to 9600 baud, 8 data bits, no parity, 1 stop bit.

5. Open the connection just created. You may have to press ENTER to see the first prompt.

6. Configure the first Fabric Interconnect, using the following example as a guideline:

   ---- Basic System Configuration Dialog ----

   This setup utility will guide you through the basic configuration of the system. Only minimal configuration including IP connectivity to the Fabric interconnect and its clustering mode is performed through these steps.

   Type Ctrl-C at any time to abort configuration and reboot system.
   To back track or make modifications to already entered values, complete input till end of section and answer no when prompted to apply configuration.
Enter the configuration method. (console/gui) ? console

Enter the setup mode; setup newly or restore from backup. (setup/restore) ? setup

You have chosen to setup a new Fabric interconnect. Continue? (y/n): y

Enforce strong password? (y/n) [y]: y

Enter the password for "admin":
Confirm the password for "admin":

Is this Fabric interconnect part of a cluster(select 'no' for standalone)? (yes/no) [n]: yes

Enter the switch fabric (A/B) []: A

Enter the system name: SJC02-151-AAK29

Physical Switch Mgmt0 IP address : 10.29.151.22

Physical Switch Mgmt0 IPv4 netmask : 255.255.255.0

IPv4 address of the default gateway : 10.29.151.1

Cluster IPv4 address : 10.29.151.24

Configure the DNS Server IP address? (yes/no) [n]: yes

DNS IP address : 10.29.130.112

Configure the default domain name? (yes/no) [n]: yes

Default domain name : ppt.lab.cisco.com

Join centralized management environment (UCS Central)? (yes/no) [n]: no

Following configurations will be applied:

Switch Fabric=A
System Name=SJC02-151-AAK29
Enforced Strong Password=no
Physical Switch Mgmt0 IP Address=10.29.151.22
Physical Switch Mgmt0 IP Netmask=255.255.255.0
Default Gateway=10.29.115.1
Ipv6 value=0
DNS Server=10.29.130.112
Domain Name=ppt.lab.cisco.com

Cluster Enabled=yes
Cluster IP Address=10.29.151.24
NOTE: Cluster IP will be configured only after both Fabric Interconnects are initialized

Apply and save the configuration (select 'no' if you want to re-enter)? (yes/no): yes
Applying configuration. Please wait.

Configuration file - Ok

Cisco UCS 6248UP Fabric Interconnect B

To configure Fabric Interconnect B, complete the following steps:
1. Connect to the console port on the first Fabric Interconnect, which will be designated as the B fabric device. Use the supplied Cisco console cable (CAB-CONSOLE-RJ45=), and connect it to a built-in DB9 serial port, or use a USB to DB9 serial port adapter.

2. Start your terminal emulator software.

3. Create a connection to the COM port of the computer's DB9 port, or the USB to serial adapter. Set the terminal emulation to VT100, and the settings to 9600 baud, 8 data bits, no parity, 1 stop bit.

4. Open the connection just created. You may have to press ENTER to see the first prompt.

5. Configure the second Fabric Interconnect, using the following example as a guideline:

```
---- Basic System Configuration Dialog ----

This setup utility will guide you through the basic configuration of the system. Only minimal configuration including IP connectivity to the Fabric interconnect and its clustering mode is performed through these steps.

Type Ctrl-C at any time to abort configuration and reboot system.
To back track or make modifications to already entered values, complete input till end of section and answer no when prompted to apply configuration.

Enter the configuration method. (console/gui) ? console

Installer has detected the presence of a peer Fabric interconnect. This Fabric interconnect will be added to the cluster. Continue (y/n) ? y

Enter the admin password of the peer Fabric interconnect: Connecting to peer Fabric interconnect... done
Retrieving config from peer Fabric interconnect... done
Peer Fabric interconnect Mgmt0 IPv4 Address: 10.29.151.22
Peer Fabric interconnect Mgmt0 IPv4 Netmask: 255.255.255.0
Cluster IPv4 address : 10.29.151.24

Peer FI is IPv4 Cluster enabled. Please Provide Local Fabric Interconnect Mgmt0 IPv4 Address

Physical Switch Mgmt0 IP address : 10.29.151.23

Apply and save the configuration (select 'no' if you want to re-enter)? (yes/no): yes
Applying configuration. Please wait.

Configuration file – Ok
```

**Cisco UCS Manager**

Log in to the Cisco UCS Manager environment by completing the following steps:

1. Open a web browser and navigate to the UCS Manager Cluster IP address, for example [https://10.29.151.24](https://10.29.151.24)

2. Click the “Launch UCS Manager” link to download and run the UCS Manager java web applet.
3. If any security prompts appear, click Continue to download the UCS Manager JNLP file.

4. If prompted to run the downloaded java application, click Run.

5. If prompted to accept any SSL security certificates, click Accept.

6. At the login prompt, enter “admin” as the username, and enter the administrative password that was entered during the initial console configuration.

7. Click No when prompted to enable Cisco Smart Call Home, this feature can be enabled at a later time.

Cisco UCS Configuration

Configure the following ports, settings and policies in the UCS Manager interface prior to beginning the HyperFlex installation.

Cisco UCS Firmware

Your Firmware version should be correct as shipped from the factory as documented in the Software Components section. This document is based on Cisco UCS infrastructure, Cisco UCS B-series bundle, and Cisco UCS C-Series bundle software versions 3.1(2b). If the firmware version of the Fabric Interconnects is older than this version, the version must be upgraded to match the requirements prior to completing any further steps. To upgrade the Cisco UCS Manager version, the Fabric Interconnect firmware, and the server bundles, refer to the instructions at:


NTP

To synchronize the Cisco UCS environment time to the NTP server, complete the following steps:
1. In Cisco UCS Manager, click the Admin tab in the navigation pane.

2. In the navigation pane, select All > Timezone Management.

3. In the Properties pane, select the appropriate time zone in the Timezone menu.

4. Click Add NTP Server.

5. Enter the NTP server IP address and click OK.

6. Click OK.

7. Click Save Changes, and then click OK.

Uplink Ports

The Ethernet ports of a Cisco UCS 6248UP Fabric Interconnect are all capable of performing several functions, such as network uplinks or server ports, and more. By default, all ports are unconfigured, and their function must be defined by the administrator. To define the specified ports to be used as network uplinks to the upstream network, complete the following steps:

1. In Cisco UCS Manager, click the Equipment tab in the navigation pane.

2. Select Fabric Interconnects > Fabric Interconnect A > Fixed Module or Expansion Module as appropriate > Ethernet Ports

3. Select the ports that are to be uplink ports, right click them, and click Configure as Uplink Port.

4. Click Yes to confirm the configuration, and click OK.
5. Select Fabric Interconnects > Fabric Interconnect B > Fixed Module or Expansion Module as appropriate > Ethernet Ports

6. Select the ports that are to be uplink ports, right click them, and click Configure as Uplink Port.

7. Click Yes to confirm the configuration, and click OK.

8. Verify all the necessary ports are now configured as uplink ports.

### Uplink Port Channels

If the Cisco UCS uplinks from one Fabric Interconnect are to be combined into a port channel or vPC, you must separately configure the port channels using the previously configured uplink ports. To configure the necessary port channels in the Cisco UCS environment, complete the following steps:

1. In Cisco UCS Manager, click the LAN tab in the navigation pane.

2. Under LAN > LAN Cloud, click to expand the Fabric A tree.

3. Right-click Port Channels underneath Fabric A and select Create Port Channel.

4. Enter the port channel ID number as the unique ID of the port channel.

5. Enter the name of the port channel.

6. Click Next.

7. Click on each port from Fabric Interconnect A that will participate in the port channel, and click the >> button to add them to the port channel.

8. Click Finish.

9. Click OK.

10. Under LAN > LAN Cloud, click to expand the Fabric B tree.

11. Right-click Port Channels underneath Fabric B and select Create Port Channel.

12. Enter the port channel ID number as the unique ID of the port channel.

13. Enter the name of the port channel.

14. Click Next.
15. Click on each port from Fabric Interconnect B that will participate in the port channel, and click the >> button to add them to the port channel.

16. Click Finish.

17. Click OK.

18. Verify the necessary port channels have been created. It can take a few minutes for the newly formed port channels to converge and come online.

Chassis Discovery Policy

If the Cisco HyperFlex system will use B200- M4 blades as compute-only nodes in a hybrid design, additional settings must be configured for connecting the UCS 5108 blade chassis. The Chassis Discovery policy defines the number of links between the Fabric Interconnect and the UCS 2204XP Fabric Extenders which must be connected and active, before the chassis will be discovered. This also effectively defines how many of those connected links will be used for communication. The Link Grouping Preference setting specifies if the links will operate independently, or if UCS Manager will automatically combine them into port-channels. Cisco best practices recommends using link grouping, and 4 links per side per UCS 5108 chassis. To configure the necessary policy and setting, complete the following steps:

1. In Cisco UCS Manager, click the Equipment tab in the navigation pane, and click on Equipment in the top of the navigation tree on the left.

2. In the properties pane, click the Policies tab.

3. Under the Global Policies sub-tab, set the Chassis/FEX Discovery Policy to match the number of up-link ports that are cabled per side, between the chassis and the Fabric Interconnects.

4. Set the Link Grouping Preference option to Port Channel.

5. Click Save Changes.

6. Click OK.
Server Ports

The Ethernet ports of a Cisco UCS Fabric Interconnect connected to the rack-mount servers, or to the blade chassis must be defined as server ports. Once a server port is activated, the connected server or chassis will begin the discovery process shortly afterwards. Rack-mount servers and blade chassis are automatically numbered in the order which they are first discovered. For this reason, it is important to configure the server ports sequentially in the order you wish the physical servers and/or chassis to appear within UCS manager. For example, if you installed your servers in a cabinet or rack with server #1 on the bottom, counting up as you go higher in the cabinet or rack, then you need to enable the server ports to the bottom-most server first, and enable them one-by-one as you move upward. You must wait until the server appears in the Equipment tab of UCS Manager before configuring the ports for the next server. The same numbering procedure applies to blade server chassis.

To define the specified ports to be used as server ports, complete the following steps:

1. In Cisco UCS Manager, click the Equipment tab in the navigation pane.
2. Select Fabric Interconnects > Fabric Interconnect A > Fixed Module or Expansion Module 2 as appropriate > Ethernet Ports.
3. Select the first port that is to be a server port, right click it, and click Configure as Server Port.
4. Click Yes to confirm the configuration, and click OK.
5. Select Fabric Interconnects > Fabric Interconnect B > Fixed Module or Expansion Module 2 as appropriate > Ethernet Ports.
6. Select the matching port as chosen for Fabric Interconnect A that is to be a server port, right click it, and click Configure as Server Port.
7. Click Yes to confirm the configuration, and click OK.
8. Wait for a brief period, until the rack-mount server appears in the Equipment tab underneath Equipment > Rack Mounts > Servers, or the chassis appears underneath Equipment > Chassis.
9. Repeat Steps 1-8 for each server port, until all rack-mount servers and chassis appear in the order desired in the Equipment tab.
Server Discovery

As previously described, once the server ports of the Fabric Interconnects are configured and active, the servers connected to those ports will begin a discovery process. During discovery the servers’ internal hardware inventories are collected, along with their current firmware revisions. Before continuing with the HyperFlex installation processes, which will create the service profiles and associate them with the servers, wait for all of the servers to finish their discovery process and to show as unassociated servers that are powered off, with no errors. To view the servers’ discovery status, complete the following steps:

1. In Cisco UCS Manager, click the Equipment tab in the navigation pane, and click on Equipment in the top of the navigation tree on the left.

2. In the properties pane, click the Servers tab.

3. Click the Blade Servers or Rack-Mount Servers sub-tab as appropriate, and view the servers’ status in the Overall Status column.
HyperFlex Installer Deployment

The Cisco HyperFlex software is distributed as a deployable virtual machine, contained in an Open Virtual Appliance (OVA) file format. The HyperFlex OVA file is available for download at cisco.com:

https://software.cisco.com/portal/pub/download/portal/select.html?&mdfid=286305544&flowid=79522&softwareid=286305994

This document is based on the Cisco HyperFlex 1.8.1c release filename: Cisco-HX-Data-Platform-Installer-v1.8.1c-19725.ova

The HyperFlex installer OVA file can be deployed as a virtual machine in an existing VMware vSphere environment, VMware Workstation, VMware Fusion, or other virtualization environment which supports the import of OVA format files. For the purpose of this document, the process describes uses an existing ESXi server to run the HyperFlex installer OVA, and deploying it through the VMware vSphere (thick) client.

Installer Connectivity

The Cisco HyperFlex Installer VM must be deployed in a location that has connectivity to the following network locations and services:

- Connectivity to the vCenter Server or vCenter Appliance which will manage the HyperFlex cluster(s) to be installed.
- Connectivity to the management interfaces of the Fabric Interconnects that contain the HyperFlex cluster(s) to be installed.
- Connectivity to the management interface of the ESXi hypervisor hosts which will host the HyperFlex cluster(s) to be installed.
- Connectivity to the DNS server(s) which will resolve host names used by the HyperFlex cluster(s) to be installed.
- Connectivity to the NTP server(s) which will synchronize time for the HyperFlex cluster(s) to be installed.
Connectivity from the staff operating the installer to the webpage hosted by the installer, and to log in to the installer through SSH.

If the network where the HyperFlex installer VM is deployed has DHCP services available to assign the proper IP address, subnet mask, default gateway, and DNS servers, the HyperFlex installer can be deployed using DHCP. If a static address must be defined, use the following table to document the settings to be used for the HyperFlex installer VM:

**Table 35  HyperFlex Installer Settings**

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hostname</td>
<td></td>
</tr>
<tr>
<td>IP Address</td>
<td></td>
</tr>
<tr>
<td>Subnet Mask</td>
<td></td>
</tr>
<tr>
<td>Default Gateway</td>
<td></td>
</tr>
<tr>
<td>DNS Server #1</td>
<td></td>
</tr>
</tbody>
</table>

**Deploy Installer OVA**

To deploy the HyperFlex installer OVA, complete the following steps:

1. Open the vSphere (thick) client, connect and log in to a vCenter server where the installer OVA will be deployed.

2. Click File > Deploy OVF Template.

3. Click Browse and locate the Cisco-HX-Data-Platform-Installer-v1.8.1c-19725.ova file, click on the file and click Open.

4. Click Next twice.

5. Modify the name of the virtual machine to be created if desired, and click on a folder location to place the virtual machine, then click Next.

6. Click on a specific host or cluster to locate the virtual machine and click Next.

7. Click on the Thin Provision option and click Next.

8. Modify the network port group selection from the drop down list in the Destination Networks column, choosing the network the installer VM will communicate on, and click Next.
9. If DHCP is to be used for the installer VM, leave the fields blank and click Next. If static address settings are to be used, fill in the fields for the installer VM hostname, default gateway, DNS server, IP address, and subnet mask, then click Next.

10. Check the box to power on after deployment, and click Finish.

The installer VM will take a few minutes to deploy, once it has deployed and the virtual machine has started, proceed to the next step.

Figure 53  HyperFlex OVA Deployment Settings

![Deploy OVF Template](image)

**Properties**

Customize the software solution for this deployment.

- **Name**
  - **Hostname**
    - The hostname for this VM. Leave blank to try to reverse lookup the IP address.
    - [HK-installer-181c]
  - **Default Gateway**
    - The default gateway address for this VM. Leave blank if DHCP is desired.
    - [10.29.130.1]
  - **DNS**
    - The domain name servers for this VM (comma separated). Leave blank if DHCP is desired.
    - [10.29.130.112]
  - **Network 1 IP Address**
    - The IP address for this interface. Leave blank if DHCP is desired.
    - [10.29.130.230]
  - **Network 1 Netmask**
    - The netmask or prefix for this interface. Leave blank if DHCP is desired.
    - [255.255.255.0]
The HyperFlex installer Web Page

The HyperFlex installer is accessed through a webpage using your local computer and a web browser. If the HyperFlex installer was deployed with a static IP address, then the IP address of the website is already known. If DHCP was used, open the local console of the installer VM. In the console you will see an interface similar to the example below, showing the IP address that was leased:

```
Data Platform Installer.
******************************************************************************
You can start the installation by visiting the following URL:
http://10.29.130.230
******************************************************************************
```
Installation

Cisco-HX-Data-Platform-Installer login:

To access the HyperFlex installer webpage, complete the following steps:

1. Open a web browser on the local computer and navigate to the IP address of the installer VM. For example, open http://10.29.130.230
2. Click accept or continue to bypass any SSL certificate errors.
3. At the security prompt, enter the username: root
4. At the security prompt, enter the password: Cisco123
5. Verify the version of the installer in the lower right-hand corner of the Welcome page is the correct version, and click Continue.
6. Read and accept the very exciting End User Licensing Agreement, and click Login.

HyperFlex Installation of Factory Systems

The HX installer will guide you through the process of setting up your cluster. It will configure UCS profiles, and settings, as well as assign the IP addressed to ESXi servers that as the HX servers come from the factory with ESX Hypervisor preinstalled. To configure the UCS settings, policies, and templates for HyperFlex, complete the following steps:

1. On the HyperFlex installer webpage select “Cluster Creation.”
2. Enter the UCS manager and Vcenter DNS hostname or IP address, the admin username, and the password, the default Hypervisor credential which come from the factory as root password Cisco123 are already entered in the installer. You can select the option to see the passwords in clear text. Optionally, you can import a json file that has the configuration information.

3. Click continue.
4. Select the Unassociated HX server models that are to be created in the HX cluster and click Continue.

5. Enter the VLAN ID’s that are to be created in UCS, as well as the Mac pool prefix, (Example: 99). If you have not already configured out of band management IP’s in UCS, enter the information on this screen.

---

Important Note: When deploying a second or any additional clusters you must put them into a different sub-org, and you should also create new VLAN names for the additional clusters. Even if reusing the same VLAN ID, it is prudent to create a new VLAN name to avoid conflicts.

Example for 2nd cluster: Change the Cluster Name, Org Name and VLAN names as to not overwrite the original cluster information. You could use naming such as HyperflexCluster-2, HXorg-2, vlan-hx-inband-mgmt-2, vlan-storage-data-2, vlan-Hyperflex Cluster-2, vlan-hx-cluster-2
6. Enter the subnet mask, gateway, DNS, and IP addresses for the Hypervisors (ESXi hosts) as well as host names. The IP’s will be assigned through UCS to ESXi systems. Hit continue.

7. Assign the additional IP addresses for the Management and Data networks as well as the cluster IP addresses, then click Continue.
8. Enter the HX Cluster Name, replication factor setting (Recommended to keep default of 3), and the password that will be assigned to the Controller VMs, then scroll down.

9. Enter the Datacenter Name from vCenter, vCenter Cluster Name, DNS, NTP, and Time Zone as well as the Auto support Settings and scroll down.
10. Leave the defaults for Advanced Networking, and validate VDI is not checked. Jumbo Frames should be enabled. It is recommended to select Clean up disk partitions. Click Start.
11. Validation of the configuration will now start. If there are warnings, you can review and click “Skip Validation” if the warnings are acceptable. If there are no warnings, the validation will automatically continue on to the configuration process.

The HX installer will now proceed to complete the deployment and perform all the steps listed at the top of the screen along with status. This process can take approximately 1 hour or more. The process can also be monitored in the vCenter and Cisco UCS Manager GUI as well as the profiles and cluster are created.
You can review the summary screen after the install completes by selecting Summary on the top right of the window.
Post Installation Script to Complete and Validate the HX Configuration

To automate the post installation procedures and verify the HyperFlex Installer has properly configured UCS Manager, a script has been provided on the Hyperflex Installer OVA. These steps can also be performed manually in vCenter if preferred. The following procedure will use the script.

1. SSH to the installer OVA IP as root with password Cisco123,

   # ssh root@10.29.130.228

2. Cd into the scripts directory of /usr/share/springpath/storfs/misc/hx-scripts and perform an ls command

   root@Cisco-HX-Data-Platform-Installer:/usr/share/springpath/storfs/misc/hx-scripts# ls

3. Run the post_install.py

4. The Installer will have the information of the previous HX installation and it will be used by the script. Enter the HX Controller Password for the HX cluster (1q2w3e4r is the password used in this install), as well as the vCenter user name and password. You can also enter the vSphere license or complete this task later.
5. Enter “y” to enable HA/DRS

Enable HA/DRS on cluster? (y/n) y

6. Enter “y” to disable SSH warning

Disable SSH warning? (y/n) y

7. You can configure logging to a local datastore and create that datastore by entering “y” for configure ESXi logging and entering a datastore name and datastore size.

Configure ESXi logging onto HX datastore? (y/n) y
No datastores found
Creating datastore...
Name of datastore: log
Size (GB): 100
Storing logs on datastore log
Creating folder [log]/esxi_logs

8. Add the vmotion VMkernel Ports to each node by inputting “y”. Input the netmask, the vmotion VLAN, and IP addresses for each of the hosts as prompted.
9. The main installer will have already created a vm-network port group and assigned the default VM network VLAN. Enter “n” to skip this step and use the default group that was created.

If desired, additional VM network port groups with assigned VLANs in the vm-networks vSwitch can be created. This option will also create the corresponding VLANs in UCS and assign the VLAN to the vm-network vNIC-Template. This script can be rerun at later time as well to create additional VM networks & UCS Vlans.

Example: Using this option in the script to show how to add more VM networks:

VLANs are created in Cisco UCS.
VLANs are assigned to vNICs.

Port groups are created.
10. Input y for the Enable NTP option.

11. Enter yes to test the auto support email function and enter a valid email address.

12. The post install script will now check the networking configuration and jumbo frames. Enter the ESXi and UCSM passwords (example: Cisco123 and 1q2w3e4r)
13. The script will complete and provide a summary screen. Validate there are no errors and the cluster is healthy.
14. Optionally, it is recommended to enable a syslog global host. In this case we will use our vCenter server as the log host by entering the vCenter IP and port info. Select the ESXi server, configuration tab, advanced settings, syslog, and entering the value for syslog.global.loghost. Repeat for each HX host.

15. Create a datastore for virtual machines. This task can be complete by using the vSphere plugin, or by using the web management interface. Using the web interface, go to the HX-Controller cluster IP management URL. Example: [HTTP://10.29.151.189/ui](http://10.29.151.189/ui) Select Datastores in the left pane, and click Create Datastore. In the popup, enter the name, size and click create.
To use the vSphere web client, select vCenter Inventory Lists, and select the Cisco Hyperflex System, Cisco HX Data Platform, cluster-name, manage tab and the plus icon to create a datastore.

16. Create a test virtual machine on your HX datastore in order to take a snapshot and perform a cloning operation.
17. Take a snapshot of the new virtual machine through the vSphere Web Client prior to powering it on. This can be scheduled as well. In the vSphere in the web client, right click on the VM, and select Cisco HX Data Platform, and Snapshot Now. Input the snapshot name and click OK.
18. Create a few clones of our virtual machine. Right click the VM, and select Cisco HX Data Platform, ReadyClones.
19. Input the Number of clones and Prefix, then click OK to start the operation. The clones will be created in seconds.
Installation

Your cluster is now ready. You may run any other preproduction tests that you wish to run at this point. Optionally, you can change the HX controller password through the “stcli security password set” command.

```
root@SpringpathControllerYIB7YF11YT:~# stcli security password set
Enter new password for user root: 
```

Customized Installation Overview - Adding vHBA’s or vNic’s

With version 1.8 customers now have the flexibility to leverage other storage infrastructure by mapping other storage to HX systems. As an example, one can map fibre channel luns on an IBM VersaStack or Flexpod storage system, and then easily Storage vMotion systems to and from other storage types.

Figure 55 VersaStack Storage in HX
In order to connect to other storage such as a FlexPod or Versastack through iSCSI or SAN, it is recommended that vHBAs or additional vNICs be added prior to creating the HX cluster. If these are added post cluster creation, the PCI enumeration can change, causing passthrough device failure in ESX and the HX cluster to go offline. The basic workflow for this procedure is outlined as follows. This assumes that ESX is installed on the HX nodes prior to this step. A separate white paper documenting in depth the process of adding other storage to HX will be available shortly on the Cisco web site. There are 2 basic methods covered below. The first method is adding HBA’s prior to creating the cluster, and the second method is adding HBA’s after creating the cluster.

Overview of Process for Customized Install if Adding Additional vHBAs to the Cisco UCS Profile Prior to Cluster Creation

1. Use the customized version of the installation workflow by selecting the “I know what I am doing” link

![Select a Workflow](image)

2. Select the tasks Run UCS Manager Configuration which will create the UCS profiles, and the Run ESX Configuration.
3. When the process completes, log in to Cisco UCS Manager and Add HBA’s or VNIC’s according to existing documentation.

4. Back in the installer, in the top right of the screen, select the Start Over option.
5. In the custom workflow screen select the options to deploy HX software, and Create HX Cluster.

6. Run the post install configuration script to complete the HX setup. Then create LUNs and map other storage to the HX systems.

Adding vNICs or vHBAs to a Deployed Cluster

To add additional storage, such as a FlexPod, after you have already installed your cluster, use the following procedure. It is recommended you do not reboot multiple nodes at once after making these hardware changes. Validate the health state of each system before rebooting or performing the procedure on subsequent nodes. In this example, you will be adding vHBAs after a cluster is created through the UCS profile service template. You will reboot 1 ESXi HX node at a time in a rolling upgrade fashion so there will be no outage.

To add vNICs or vHBAs to a deployed cluster, complete the following steps:

1. Example of hardware change: Add vHBAs to the service profile templates for HX (refer to the Cisco UCS documentation for your storage device such as a FlexPod CVD for configuring the vHBAs).
2. After you have completed adding HBA’s to the templates, the servers will require a reboot. Do not reboot the HX servers at this time.

3. Using the vSphere Web Client, place 1 of the HX ESXi hosts in HX-maintenance mode.
4. After the system has entered maintenance mode, you can reboot the associated node in UCS to complete the addition of the new hardware.

5. In vSphere, Go to the configuration tab of the ESX server.

6. Go to hardware, and select advanced settings. In our example, there will be no devices configured for passthrough since we have change the PCI order.

7. Select Configure Passthrough.
8. Select the LSI card to be passthrough and click OK.

9. Reboot the ESXi host.

10. When the ESXi host has rebooted, you can validate your that passthrough device is available in the advanced setting screen. Exit maintenance mode.
11. Reconfigure the HX controller, by right clicking the system and selecting Edit Settings. Click OK.

12. In the hardware tab, the PCI device is shown as unavailable. Highlight the device and click Remove.

13. Click Add, select PCI device, and click Next.
14. The PCI device should be highlighted in the drop-down. Click Next.

15. Select Finish, then OK.

16. Right-click and power on the HX controller.

17. Check the health status of the cluster with the command to validate when the cluster is healthy before proceeding to the next node. Example: Run these command to the cluster IP for the HX Controllers “stcli cluster refresh” then “stcli cluster info | grep -i health”
18. Wait and check again validating the cluster is healthy.

19. Repeat the process for each node in the cluster as necessary.

**ESXi Hypervisor Installation for Reinstallation of HX**

There are various instances where you might want to redeploy or perform an install using the HX customized install method. The HyperFlex installer VM can be used to build a custom ESXi installation ISO using a kickstart file, which automates the installation process of ESXi to the HyperFlex nodes. Since the systems come from the factory with ESXi pre-installed, for a re-install we will need to install ESXi on the nodes. The HyperFlex system requires a Cisco custom ESXi ISO file to be used, which has Cisco hardware specific drivers pre-installed to ease the installation process, as detailed in section Software Components. The Cisco custom ESXi ISO file is available for download at cisco.com: [https://software.cisco.com/portal/pub/download/portal/select.html?&mdfid=286305544&flowid=79522&softwareid=286305994](https://software.cisco.com/portal/pub/download/portal/select.html?&mdfid=286305544&flowid=79522&softwareid=286305994)

A high-level example of a rebuild procedure is as follows:

1. **Clean up the existing environment by deleting the cluster in vCenter, removing vCenter MOB entries as well as Cisco UCS Manager service profiles and VLANs in Cisco UCS.**

2. **Use customized workflow and only run the “Run UCS Manager Configuration” option.**

3. **Perform a fresh ESXi installation using the custom ISO and procedure the vMedia method in this section.**
4. Use customized workflow and select the remaining 3 options ESX Configuration, Deploy HX Software, and Create HX Cluster making sure to select the delete existing partitions option in the wizard.

For more information about the various installation methods can be found in the Getting Started Guide:


The HX custom ISO is based on the Cisco custom ESXi 6.0 U2 ISO release filename: Vmware-ESXi-60U2-4192238-Cisco-Custom-6.0.2.3.iso

The kickstart process will automatically perform the following tasks with no user interaction required:

- Accept the End User License Agreement.
- Configure the root password to: Cisco123
- Install ESXi to the internal mirrored Cisco FlexFlash SD cards.
- Set the default management network to use vmnic0, and obtain an IP address through DHCP.
- Enable SSH access to the ESXi host.
- Enable the ESXi shell.
- Enable serial port com1 console access to facilitate Serial over LAN access to the host.
- Configure the ESXi configuration to always use the current hardware MAC address of the network interfaces, even if they change.
- Rename the default vSwitch to vswitch-hx-inband-mgmt.

**ESXi Kickstart ISO Creation**

A custom iso file could be required if you are re-installing the cluster, or if a newer ISO is available. ISO is available on the Cisco web site:

https://software.cisco.com/download/type.html?mdfid=286290156&catid=null
To create the custom kickstart ESXi installation ISO file, complete the following steps:

1. Copy the base ISO file, Vmware-ESXi-60U2-4192238-Cisco-Custom-6.0.2.3.iso to the HyperFlex installer VM using SCP, SFTP or any available method. Place the file in the /var/hyperflex/source-images folder.

2. Login to the HyperFlex installer VM through SSH, using PuTTY or a similar tool. Username is: root
Password is: Cisco123

3. Enter the following commands to integrate the kickstart file with the base ISO file, and create a new ISO file from them:

   ```bash
   # cd /opt/springpath/install-esxi
   # ./makeiso.sh /opt/springpath/install-esxi/cisco-hx-esxi-6.0.ks.cfg \
   /var/hyperflex/source-images/Vmware-ESXi-60U2-4192238-Cisco-Custom-6.0.2.3.iso \
   /var/www/localhost/images/HX-Vmware-ESXi-60U2-4192238-Cisco-Custom-6.0.2.3.iso
   ```

4. You will see output from the command as below:

   ```bash
   root@Cisco-HX-Data-Platform-Installer:/opt/springpath/install-esxi# ./makeiso.sh \
   /opt/springpath/install-esxi/cisco-hx-esxi-6.0.ks.cfg \
   /var/hyperflex/source-images/Vmware-ESXi-60U2-4192238-Cisco-Custom-6.0.2.3.iso \
   /var/www/localhost/images/HX-Vmware-ESXi-60U2-4192238-Cisco-Custom-6.0.2.3.iso
   mount: block device /var/hyperflex/source-images/Vmware-ESXi-60U2-4192238-Cisco-Custom-6.0.2.3.iso is write-protected, mounting read-only
   Warning: creating filesystem that does not conform to ISO-9660.
   I: -input-charset not specified, using utf-8 (detected in locale settings)
   Size of boot image is 4 sectors -> No emulation
   2.71% done, estimate finish Wed Nov  9 00:31:59 2016
   5.42% done, estimate finish Wed Nov  9 00:31:59 2016
   8.12% done, estimate finish Wed Nov  9 00:31:59 2016
   10.83% done, estimate finish Wed Nov  9 00:31:59 2016
   13.53% done, estimate finish Wed Nov  9 00:31:59 2016
   16.24% done, estimate finish Wed Nov  9 00:31:59 2016
   18.94% done, estimate finish Wed Nov  9 00:31:59 2016
   21.66% done, estimate finish Wed Nov  9 00:31:59 2016
   24.36% done, estimate finish Wed Nov  9 00:31:59 2016
   27.07% done, estimate finish Wed Nov  9 00:31:59 2016
   29.77% done, estimate finish Wed Nov  9 00:31:59 2016
   32.48% done, estimate finish Wed Nov  9 00:31:59 2016
   35.18% done, estimate finish Wed Nov  9 00:31:59 2016
   37.89% done, estimate finish Wed Nov  9 00:31:59 2016
   40.59% done, estimate finish Wed Nov  9 00:31:59 2016
   43.30% done, estimate finish Wed Nov  9 00:31:59 2016
   46.01% done, estimate finish Wed Nov  9 00:31:59 2016
   48.72% done, estimate finish Wed Nov  9 00:31:59 2016
   51.42% done, estimate finish Wed Nov  9 00:31:59 2016
   54.13% done, estimate finish Wed Nov  9 00:31:59 2016
   56.83% done, estimate finish Wed Nov  9 00:31:59 2016
   59.54% done, estimate finish Wed Nov  9 00:31:59 2016
   62.24% done, estimate finish Wed Nov  9 00:31:59 2016
   64.95% done, estimate finish Wed Nov  9 00:31:59 2016
   67.65% done, estimate finish Wed Nov  9 00:31:59 2016
   70.37% done, estimate finish Wed Nov  9 00:31:59 2016
   73.07% done, estimate finish Wed Nov  9 00:31:59 2016
   75.78% done, estimate finish Wed Nov  9 00:31:59 2016
   ```
Installation

78.48% done, estimate finish Wed Nov 9 00:31:59 2016
81.19% done, estimate finish Wed Nov 9 00:31:59 2016
83.89% done, estimate finish Wed Nov 9 00:31:59 2016
86.60% done, estimate finish Wed Nov 9 00:32:00 2016
89.30% done, estimate finish Wed Nov 9 00:32:00 2016
92.01% done, estimate finish Wed Nov 9 00:32:00 2016
94.72% done, estimate finish Wed Nov 9 00:32:00 2016
97.43% done, estimate finish Wed Nov 9 00:32:00 2016

Total translation table size: 2048
Total rockridge attributes bytes: 11105
Total directory bytes: 6144
Path table size(bytes): 50
Max brk space used: 22000
184764 extents written (360 MB)
Successfully created ESXi ISO Image /var/www/localhost/images/HX-Vmware-ESXi-60U2-4192238-Cisco-Custom-6.0.2.3.iso
root@Cisco-HX-Data-Platform-Installer:/opt/springpath/install-esxi#

5. Verify the newly created ISO file exists in the proper webpage location, by opening a web browser and navigating the the IP address of the HyperFlex installer VM, followed by /images. For example: http://10.29.130.230/images

6. The new file, named HX-Vmware-ESXi-60U2-4192238-Cisco-Custom-6.0.2.3.iso will be listed in the webpage file listing.

Cisco UCS vMedia and Boot Policies

Using a UCS vMedia policy, the mounting of the newly created custom kickstart ESXi installation ISO file can be automated. The existing vMedia policy, named “HyperFlex” must be modified to mount this file, and the boot policy must be modified temporarily to boot from the remotely mounted vMedia file. Once these two tasks are completed, the servers can be rebooted, and they will automatically boot from the remotely mounted vMedia file, installing and configuring ESXi on the servers as described above in ESXi Hypervisor Installation.

Warning: While vMedia boot is very efficient for installing multiple servers, using vMedia boot policies below can automatically re-install ESX on any existing server that is rebooted with this policy. Extreme caution is recommended. This is a destructive procedure to the hosts as it will overwrite the install location should other systems be rebooted accidentally. This procedure needs to be carefully monitored and the boot policy should be changed back to original settings immediately after intended servers are rebooted. It is recommended only for new installs/rebuilds. Alternatively you can manually select the boot device using the KVM console on startup instead of making the vMedia device the default boot selection.
To configure the Cisco UCS vMedia and Boot Policies, complete the following steps:

1. In Cisco UCS Manager, click the Servers tab in the navigation pane.
2. Expand Servers > Policies > root > Sub-Organizations > hx-cluster > vMedia Policies, and click vMedia Policy HyperFlex.
3. In the configuration pane, click Create vMedia Mount.
4. Enter a name for the mount, for example: ESX.
5. Select the CDD option.
6. Select HTTP as the protocol.
7. Enter the IP address of the HyperFlex installer VM, for example: 192.168.10.210
8. Select None as the Image Variable Name.
9. Enter HX-Vmware-ESXi-60U2-4192238-Cisco-Custom-6.0.2.3.iso as the Remote File.
10. Enter /images/ as the Remote Path.
11. Click Save Changes, and click OK.
12. Select Servers > Service Profile Templates > root > Sub-Organizations > hx-cluster > Service Template hx-nodes.
13. In the configuration pane, click the vMedia Policy Tab.
14. Click Modify vMedia Policy.
15. Chose the HyperFlex vMedia Policy from the drop down selection, and click OK twice.
17. In the navigation pane, expand the section titled CIMC Mounted vMedia.
18. Click the entry labeled Add CIMC Mounted CD/DVD.
19. Select the CIMC Mounted CD/DVD entry in the Boot Order list, and click the Move Up button until the CIMC Mounted CD/DVD entry is listed first.
20. Click Save Changes, and click OK.
Installation

Install ESXi

To begin the installation after modifying the vMedia policy, Boot policy and service profile template, the servers need to be rebooted. To monitor the progress of one or more servers, it is advisable to open a remote KVM console session to watch the installation.

To open the KVM console and reboot the servers, complete the following steps:

1. In Cisco UCS Manager, click the Servers tab in the navigation pane.
2. Expand Servers > Service Profiles > root > Sub- Organizations > hx- cluster > rack- unit- 1.
3. In the configuration pane, click KVM Console.
4. Click continue to any security alerts that appear. The remote KVM Console window will appear shortly and show the server’s local console output.
5. Repeat Steps 2-4 for any additional servers whose console you wish to monitor during the installation.
6. In Cisco UCS Manager, click the Equipment tab in the navigation pane.
8. In the configuration pane, click the first server to be rebooted, then shift+click the last server to be rebooted, selecting them all.
9. Right-click the mouse, and click Reset.
10. Click OK.
11. Select Power Cycle and click OK.
12. Click OK.
The servers you are monitoring in the KVM console windows will now immediately reboot, and boot from the remote vMedia mount and install ESXi. There may be error messages seen on screen, but they can be safely ignored.
Installation

- Equipment
- Servers
- LAN
- SAN
- VM
- Admin
- Storage

- Server 1: Ok
- Server 2: Ok
- Server 3: Ok
- Server 4: Ok
- Server 5: Ok
- Server 6: Ok
- Server 7: Ok
- Server 8: Ok

- Fabric Interconnects
  - Fabric Interconnect A (primary)
    - Fixed Module
    - Ethernet Ports
    - FC Ports
    - Expansion Module 2
    - Ethernet Ports

- Boot Server
- Shutdown Server
- Reset
- Start Fault Suppression
- Stop Fault Suppression

Press [Tab] to edit options
Automatic boot in 5 seconds...
Undo vMedia and Boot Policy Changes

Once all the servers have booted from the remote vMedia file and begun their installation process, the changes to the boot policy need to be quickly undone, to prevent the servers from going into a boot loop, constantly booting from the installation ISO file.

To revert the boot policy settings, complete the following steps:

1. Select Servers > Policies > root > Sub-O rganizations > hx-cluster > Boot Policy HyperFlex.
2. Select the CIMC Mounted CD/DVD entry in the Boot Order list, and click Delete.
3. Click Save Changes, and click OK.

HyperFlex Cluster Expansion

The process to expand a HyperFlex cluster can be used to grow an existing HyperFlex cluster with additional converged storage nodes, or to expand an existing cluster with additional compute-only nodes to create a hybrid cluster. With converged nodes, you are able to use the standard workflow wizard for cluster expansion.

The process for adding compute-only nodes differ slightly and so is covered below.

Wizard for Cluster Expansion of Converged Nodes is very similar to initial setup.

Figure 56  Cluster Expansion Converged Nodes

HyperFlex Compute-Only Cluster Expansion

HyperFlex clusters are built by first installing applying UCS profiles to the blades, installing ESX, then expanding the cluster with compute-only nodes. There are some specific differences in the processes which
Installation

will be outlined below. To expand an existing HyperFlex cluster, creating a hybrid cluster, complete the following steps:

Configure Cisco UCS

1. In Cisco UCS Manager, click the Servers tab in the navigation pane.

2. Expand Servers > Service Profile Templates > root > Sub-Organizations > hx-cluster.

3. Right-click Service Profile Template compute-nodes and click Create Service Profiles from Template.

4. Enter the naming prefix of the service profiles that will be created from the template.

5. Enter the starting number of the service profiles being created and the number of service profiles to be created.

6. Click OK twice.

7. Expand Servers > Service Profiles > root > Sub-Organizations > hx-cluster.

8. Click on the first service profile created for the additional B200-M4 compute-only nodes, and right click.

9. Click Change Service Profile Association.

10. In the Server Assignment dropdown list, change the selection to Select Existing Server.

12. In the list of available servers, choose the B200-M4 blade server to assign the service profile to and click OK.

13. Click Yes to accept that the server will reboot to apply the service profile.

14. Click OK. Repeat steps 7-14 for each additional compute-only B200-M4 blade that will be added to the cluster.
Install ESXi

1. Install and configure the ESXi hypervisor either manually using the HX cluster ISO or through the vMedia policy method.

2. Login to the KVM console of the server through Cisco UCS.

3. Enter F2 on the console for login prompt.

4. Login to ESX with root password of Cisco123.

5. Configure Management Network.

6. Select VLAN and Configure the VLAN ID to your hx-inband-management VLAN and click OK.

7. Select IPv4 Configuration

8. Set Static IP and input IP, Netmask and Gateway, then OK

9. Select the DNS configuration and enter the DNS info and hit ok.
10. Leave the network config section making sure to enter Y to apply and the changes.

**DNS Configuration**

This host can only obtain DNS settings automatically if it also obtains its IP configuration automatically.

( ) Obtain DNS server addresses and a hostname automatically
(o) Use the following DNS server addresses and hostname:

Primary DNS Server: 10.29.130.112
Alternate DNS Server: [ ]
Hostname: [ ]

<Up/Down> Select <Space> Mark Selected <Enter> OK <Esc> Cancel

---

**Configure Management Network: Confirm**

You have made changes to the host’s management network. Applying these changes may result in a brief network outage, disconnect remote management software and affect running virtual machines. In case IPv6 has been enabled or disabled this will restart your host.

Apply changes and restart management network?

<Y> Yes <N> No <Esc> Cancel

---

**Expand the Cluster**

1. Open the HyperFlex installer webpage.

2. Select the customize workflow.

3. Select both the Deploy HX Software and Expand HX Cluster options.
4. Enter the vCenter and Hypervisor Credentials and click Continue.

5. Validate the cluster IP is detected and correct or manually enter it, and click Continue.
6. Click Add Compute Server and add the IP address for the Hypervisor Management Traffic, and add the IP address for the Hypervisor Data traffic. Input the Controller VM password.

There are no storage controller IP addresses since this is a compute only node.
7. Select start to complete the installation. This process will automatically add the compute nodes to the cluster in vCenter.

8. Select the Summary screen.
9. After the install has completed, the Compute node is added to the cluster and now has access to the datastores, but requires some post installation steps manual completed. Click on the ESXi server and select the summary screen to identify issues.

10. Select the configuration tab and review the networking settings to note that there are port groups that need to be added.
11. The list of tasks required to complete setup are shown below. You can manually perform the tasks through the vSphere GUI, or use PowerCLI or other scripting. An example is provided using PowerCLI.

- NTP configuration
- Disable SSH warning
- Syslog Server Configuration
- Creating the "vswitch-hx-vm-network" portgroup
- Creating the "vMotion" kernel port

Example: PowerCLI script to complete tasks on the ESXi host.

```powershell
$i=“10.29.151.241"
Connect-VIServer -server $i -user root -password Cisco123
$vmhost = Get-VMHost -Name $i
#add NTP
$myntpserver="171.68.38.66"
Add-VmHostNtpServer -VMHost $vmhost -NtpServer $myntpserver
Get-VmHostService -VMHost $vmhost | Where-Object {$_._key -eq "ntpd"} | Start-VmHostService
Get-VmHostService -VMHost $vmhost | Where-Object {$_._key -eq "ntpd"} | Set-VmHostService -policy "automatic"
#set SSH warning
$vmhost | Set-VmHostAdvancedConfiguration UserVars.SuppressShellWarning 1
```
Installation

$vmhost | Get-VMHostService | Where {$_.Key -eq "TSM-SSH"} | Set-VMHostService -Policy "On"

$vmhost | Get-VMHostService | Where {$_ .Key -eq "TSM-SSH"} | Start-VMHostService

$vmhost |.ForEach {Start-VMHostService -HostService {$_ | Get-VMHostService | Where {$_ .Key -eq "TSM-SSH"}}}

#set syslog

$vmhost | Set-VMHostSysLogServer -SysLogServer "udp://10.29.151.102" -SysLogServerPort 514

#create portgroup for vm net

$vswitch4 = Get-VirtualSwitch -VMHost $vmhost -Name "vswitch-hx-vm-network"
New-VirtualPortGroup -VirtualSwitch $vswitch4 -Name "vm-network-3174" -VLANID 3174
$portgroupname="vmotion"

#create vmotion vmkernel

$vswitch3 = Get-VirtualSwitch -VMHost $vmhost -Name vmotion
New-VMHostNetworkAdapter -VMHost $vmhost -VirtualSwitch $vswitch3 -PortGroup $portgroupname -MTU 9000 -IP 172.17.73.241 -SubnetMask 255.255.255.0 -VMotionEnabled $true

$VMHost | Get-VirtualPortGroup -Name $portgroupname | Set-VirtualPortGroup -VLANID 3173
12. To validate our configuration, you will vmotion a VM to the new compute node. Right-click a VM and select migrate, then change host by selecting your new compute-only node as the target. Click next, next and finish. You can validate your VM in now running on the compute only node through the summary tab of the VM.
### Installation

#### VMware vSphere vCenter Server 6.5 UI

**General**
- **Guest OS:** Other 3.0 or later Linux (64-bit)
- **VM Version:** 10
- **CPU:** 4 vCPU
- **Memory:** 4096 MB
- **Memory Overhead:**
- **VMware Tools:** Running (Guest managed)
- **IP Addresses:** 10.29.145.15
- **DNS Name:** localhost
- **EVC Mode:** N/A
- **State:** Powered On
- **Host:** b2000hsq-gob-bb-cisco.com
- **vSphere HA Protection:** Protected

![Screenshot of VMware vSphere vCenter Server 6.5 UI](image-url)
Management

VCenter Web Client Plugin

The Cisco HyperFlex vCenter Web Client Plugin is installed by the HyperFlex installer to the specified vCenter server or vCenter appliance. The plugin is accessed as part of the vCenter Web Client interface, and is the primary tool used to monitor and configure the HyperFlex cluster.

To manage HyperFlex cluster using the plugin, complete the following steps:

1. Open the vCenter Web Client, and login.
2. In the home pane, from the home screen click on VCenter Inventory Lists.
3. In the Navigator pane, click on Cisco HX Data Platform.
4. In the Navigator pane, click the name of the HyperFlex cluster.

Summary

From the Web Client Plugin Summary screen, several elements are presented:

- Overall cluster usable capacity, used capacity, free capacity, datastore capacity provisioned, and the amount of datastore capacity provisioned beyond the actual cluster capacity.
- Deduplication and compression savings percentages calculated against the data stored in the cluster.
- The cluster health state, and the number of node failures that can occur before the cluster goes into read-only or offline mode.
- A snapshot of performance over the previous hour, showing IOPS, throughput, and latencies.

Monitor

From the Web Client Plugin Monitor tab, several elements are presented:

- Clicking the Performance button gives a larger view of the performance charts. If a full webpage screen view is desired, click the Preview Interactive Performance charts hyperlink.
- Clicking the Events button displays a HyperFlex event log, which can be used to diagnose errors and view system activity events.

Figure 57  HyperFlex Plug-in Performance Monitor
Manage

From the Web Client Plugin Manage tab, several elements are presented:

- Clicking the Cluster button gives an inventory of the HyperFlex cluster and the physical assets of the cluster hardware.
- Clicking the Datastores button allows datastores to be created, edited, deleted, mounted and unmounted, along with space summaries and performance snapshots of that datastore.
Figure 59  HyperFlex Plug-in Manage Cluster

<table>
<thead>
<tr>
<th>Model</th>
<th>HX220C-M4S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial Number</td>
<td>FCH1950V000</td>
</tr>
<tr>
<td>Converged Nodes</td>
<td>1</td>
</tr>
<tr>
<td>Hard Disks</td>
<td>6</td>
</tr>
<tr>
<td>SSDs</td>
<td>1</td>
</tr>
<tr>
<td>PSU 1 Status</td>
<td>PSU Power Supply 1: Running/Full Power-Enabled</td>
</tr>
<tr>
<td>PSU 2 Status</td>
<td>PSU Power Supply 2: Running/Full Power-Enabled</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>State</th>
<th>Status</th>
<th>IPMI ...</th>
<th>Controller IP</th>
<th>Host IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>hx220-02.hxlab.cisco.com</td>
<td>connected</td>
<td>Normal</td>
<td>0.0.0.0</td>
<td>192.168.51.22</td>
<td>192.168.10.12</td>
</tr>
</tbody>
</table>
Management

Management Best Practices

In this section, various best practices and guidelines are given for management and ongoing use of the Cisco HyperFlex system. These guidelines and recommendations apply only to the software versions upon which this document is based, listed in Software Components.

ReadyClones

For the best possible performance and functionality of the virtual machines that will be created using the HyperFlex ReadyClone feature, the following guidelines for preparation of the base VMs to be cloned should be followed:

- Base VMs must be stored in a HyperFlex datastore.
- All virtual disks of the base VM must be stored in the same HyperFlex datastore.
- Base VMs can only have HyperFlex native snapshots, no VMware redo-log based snapshots can be present.

For very high IO workloads with many clone VMs leveraging the same base image, it might be necessary to use a multiple copies of the same base image for groups of clones. Doing so prevents referencing the same blocks across all clones and could yield an increase in performance. This step is typically not required for most use cases and workload types.

Snapshots

HyperFlex native snapshots are high performance snapshots that are space-efficient, crash-consistent, and application consistent, taken by the HyperFlex Distributed Filesystem, rather than using VMware redo-log based snapshots. For the best possible performance and functionality of HyperFlex native snapshots, the following guidelines should be followed:

- Make sure that the first snapshot taken of a guest VM is a HyperFlex native snapshot, by using the “Cisco HX Data Platform” menu item in the vSphere Web Client, and choosing Snapshot Now or Schedule Snapshot. Failure to do so reverts to VMware redo-log based snapshots.

- A Sentinel snapshot becomes a base snapshot that all future snapshots are added to, and prevents the VM from reverting to VMware redo-log based snapshots. Failure to do so can cause performance degradation when taking snapshots later, while the VM is performing large amounts of storage IO.

- Additional snapshots can be taken through the “Cisco HX Data Platform” menu, or the standard vSphere client snapshot menu. As long as the initial snapshot was a HyperFlex native snapshot, each additional snapshot is also considered to be a HyperFlex native snapshot.

- Do not delete the Sentinel snapshot unless you are deleting all the snapshots entirely.

- Do not revert the VM to the Sentinel snapshot.

- If large numbers of scheduled snapshots need to be taken, distribute the time of the snapshots taken by placing the VMs into multiple folders or resource pools. For example, schedule two resource groups, each with several VMs, to take snapshots separated by 15 minute intervals in the scheduler window. Snapshots will be processed in batches of 8 at a time, until the scheduled task is completed.
Figure 61  vSphere Web Client HX Snapshot Now

Figure 62  HyperFlex Sentinel Snapshot
Storage vMotion

The Cisco HyperFlex Distributed Filesystem can create multiple datastores for storage of virtual machines. While there can be multiple datastores for logical separation, all of the files are located within a single distributed filesystem. As such, performing storage vMotions of virtual machine disk files has little value in the HyperFlex system. Furthermore, storage vMotions create additional filesystem consumption and generate additional unnecessary metadata within the filesystem, which must later be cleaned up through the filesystem's internal cleaner process.

It is recommended to not perform storage vMotions of the guest VMs between datastores within the same HyperFlex cluster. Storage vMotions between different HyperFlex clusters, or between HyperFlex and non-HyperFlex datastores are permitted.

Virtual Disk Placement

HyperFlex clusters can create multiple datastores for logical separation of virtual machine storage, yet the files are all stored in the same underlying distributed filesystem. The only difference between one datastore and another are their names and their configured sizes. Due to this, there is no compelling reason for a virtual machine’s virtual disk files to be stored on a particular datastore versus another.
All of the virtual disks that make up a single virtual machine must be placed in the same datastore. Spreading the virtual disks across multiple datastores provides no benefit, and can cause ReadyClone and snapshot errors.

**Maintenance Mode**

Within the vCenter Web Client, a specific menu entry for “HX Maintenance Mode” has been installed by the HyperFlex plugin. This option directs the storage platform controller on the node to shutdown gracefully, redistributing storage IO to the other nodes with minimal impact. Using the standard Maintenance Mode menu in the vSphere Web Client, or the vSphere (thick) Client can be used, but graceful failover of storage IO and shutdown of the controller VM is not guaranteed.

In order to minimize the performance impact of placing a HyperFlex converged storage node into maintenance mode, it is recommended to use the HX Maintenance Mode menu selection to enter or exit maintenance mode whenever possible.

**Figure 64  vSphere Web Client HX Maintenance Mode**
Validation

This section provides a list of items that should be reviewed after the HyperFlex system has been deployed and configured. The goal of this section is to verify the configuration and functionality of the solution, and ensure that the configuration supports core availability requirements.

Post Install Checklist

The following tests are critical to functionality of the solution, and should be verified before deploying for production:

- Verify the expected number of converged storage nodes and compute-only nodes are members of the HyperFlex cluster in the vSphere Web Client plugin manage cluster screen.
- Verify the expected cluster capacity is seen in the vSphere Web Client plugin summary screen. (See Appendix A)
- Create a test virtual machine that accesses the HyperFlex datastore and is able to perform read/write operations.
- Perform the virtual machine migration (vMotion) of the test virtual machine to a different host on the cluster.
- During the vMotion of the virtual machine, make sure the test virtual machine can perform a continuous ping to default gateway and to check if the network connectivity is maintained during and after the migration.

Verify Redundancy

The following redundancy checks can be performed to verify the robustness of the system. Network traffic, such as a continuous ping from VM to VM, or from vCenter to the ESXi hosts should not show significant failures (one or two ping drops might be observed at times). Also, all of the HyperFlex datastores must remain mounted and accessible from all the hosts at all times.

1. Administratively disable one of the server ports on Fabric Interconnect A which is connected to one of the HyperFlex converged storage hosts. The ESXi virtual switch uplinks for fabric A should now show as failed, and the standby uplinks on fabric B will be in use for the management and vMotion virtual switches. Upon administratively re-enabling the port, the uplinks in use should return to normal.

2. Administratively disable one of the server ports on Fabric Interconnect B which is connected to one of the HyperFlex converged storage hosts. The ESXi virtual switch uplinks for fabric B should now show as failed, and the standby uplinks on fabric A will be in use for the storage virtual switch. Upon administratively re-enabling the port, the uplinks in use should return to normal.

3. Place a representative load of guest virtual machines on the system. Put one of the ESXi hosts in maintenance mode, using the HyperFlex HX maintenance mode option. All the VMs running on that host should be migrated through vMotion to other active hosts through vSphere DRS, except for the
storage platform controller VM, which will be powered off. No guest VMs should lose any network or storage accessibility during or after the migration. This test assumes that enough RAM is available on the remaining ESXi hosts to accommodate VMs from the host put in maintenance mode. The HyperFlex cluster will show in an unhealthy state.

4. Reboot the host that is in maintenance mode, and exit it from maintenance mode after the reboot. The storage platform controller will automatically start when the host exits maintenance mode. The HyperFlex cluster will show as healthy after a brief time to restart the services on that node. VSphere DRS should rebalance the VM distribution across the cluster over time.

Many vCenter alerts automatically clear when the fault has been resolved. Once the cluster health is verified, some alerts may need to be manually cleared.

5. Reboot one of the two Cisco UCS Fabric Interconnects while traffic is being sent and received on the storage datastores and the network. The reboot should not affect the proper operation of storage access and network traffic generated by the VMs. Numerous faults and errors will be noted in UCS Manager, but all will be cleared after the FI comes back online.
Documentation Roadmap

For comprehensive documentation, please also refer to the following location on the Cisco UCS HX-Series Documentation Roadmap. This location includes Pre-Install Checklist, Getting Started Guide, Admin Guide, Tech Notes, Release Notes, Interoperability, CLI reference and other pertinent information:

Documentation roadmap (Cisco Web login required):


Hyperconverged Infrastructure:

http://hyperflex.io
A: Cluster Capacity Calculations

A HyperFlex HX Data Platform cluster capacity is calculated as follows:

\[((<\text{capacity disk size in GB}> \times 10^9) / 1024^3) \times <\text{number of capacity disks per node}> \times <\text{number of HyperFlex nodes}> \times 0.92) / \text{replication factor}\]

Divide the result by 1024 to get a value in TiB

The replication factor value is 3 if the HX cluster is set to RF=3, and the value is 2 if the HX cluster is set to RF=2.

The 0.92 multiplier accounts for an 8% reservation set aside on each disk by the HX Data Platform software for various internal filesystem functions.

Calculation example:

- < capacity disk size in GB> = 1200 for 1.2 TB disks
- <number of capacity disks per node> = 15 for an HX240c-M4SX model server
- <number of HyperFlex nodes> = 8
- replication factor = 3

Result: \[((1200*10^9)/1024^3)*15*8*0.92)/3 = 41127.2049\]

\[41127.2049 / 1024 = 40.16 \text{ TiB}\]

B: Example Scripts

Example PowerShell scripts are provided to assist with the configuration of the HyperFlex system. The scripts must be modified for their proper use according to the settings required by the system being installed. Additional PowerShell integrations may be required to be installed on the system where the scripts will be run, such as vSphere PowerCLI or Cisco UCS PowerTool Suite.

Create DHCP Reservations

DHCP is not recommended since VLANs are not native by default. If using DHCP, you must set the management net VLAN on the ESXi host.

# Create_DHCP_Reservations.ps1

# Description: Imports a UCS MAC address pool CSV export file and creates DHCP reservations
# Usage: Modify the $csv variable to point to the UCS CSV export file. Modify the $ip variable to set the starting IP address of the last octet. Modify the $address variable to
Appendix

# contain the first three octets of the IP addresses. Modify the -Scopeld value to the name
# of the DHCP scope where the reservations are being added. Run the script from the Windows
# DHCP server with the scope you are adding to.
#
$csv="mgmt-a.csv"
$ip=11
$macs = Import-Csv $csv | Where-Object {$_.Assigned -eq "yes"} | Select ID

Foreach ($mac in $macs) {
    $mac = ($mac.ID -split '/')[3]
    $mac = $mac -replace ":"
    $address="192.168.10."+$ip
    Write-Host "Creating reservation for MAC address:" $mac
    Write-Host "Using IP address:" $address
    Get-DhcpServerv4Scope -Scopeld 192.168.10.0 | Add-DhcpServerv4Reservation -ClientId $mac -IPAddress $address
    $ip++
}

Configure ESXi Post Install

# Configure_ESX_post_install.ps1
# Description: Configures ESXi options and settings after HyperFlex installation.
# Usage: Modify the variables to specify the ESXi root password, the servers to be
# configured, the guest VLAN ID, and the IP addresses used for the vMotion vmkernel
# interfaces.
#
Set-PowerCLIConfiguration -InvalidCertificateAction Ignore -Confirm:$false | Out-Null
$rootpw = "Cisco123"
$ip=11

Foreach ($server in $servers) {

    # connect to the ESXi host server
    Connect-VIServer -server $server -user root -password $rootpw
    $vmhost = Get-VMHost -Name $server

    # retrieve the virtual switch configurations
    $vswitch2 = Get-VirtualSwitch -VMHost $vmhost -Name vswitch-hx-vm-network
    $vswitch3 = Get-VirtualSwitch -VMHost $vmhost -Name vmotion

    # create a port group for the guest VMs
    New-VirtualPortGroup -VirtualSwitch $vswitch2 -Name "VM Network" -VLANID 100

    # create the vmotion port group and vmkernel interface
    $vmip="192.168.200."+$ip
    New-VMHostNetworkAdapter -VMHost $vmhost -VirtualSwitch $vswitch3 -PortGroup "vmotion" -Mtu 9000 -VMotionEnabled $true -IP $vmip -SubnetMask 255.255.255.0 -Confirm:$false
    $ip=$ip+1

    # configure syslog traffic to send to vCenter
    Set-VMHostSysLogServer -SysLogServer '192.168.10.101:514' -VMHost $vmhost

    Disconnect-VIServer -server $server -Confirm:$false
}

}
About the Authors

Brian Everitt, Technical Marketing Engineer, Cisco UCS Datacenter Engineering Solutions Group, Cisco Systems, Inc.

Brian is an IT industry veteran with over 18 years of experience deploying server, network, and storage infrastructures for companies around the world. During his tenure at Cisco, he has previously been a lead Advanced Services Solutions Architect for Microsoft solutions, virtualization, and SAP Hana on Cisco UCS. Currently his focus is on Cisco's portfolio of Software Defined Storage (SDS) and Hyperconverged Infrastructure solutions.

Acknowledgements

For their support and contribution to the design, validation, and creation of this Cisco Validated Design, we would like to acknowledge the significant contribution and expertise that resulted in developing this document:

- Jeffrey Fultz, Technical Marketing Engineer with the Cisco UCS Datacenter Engineering Solutions Group, Cisco Systems, Inc.